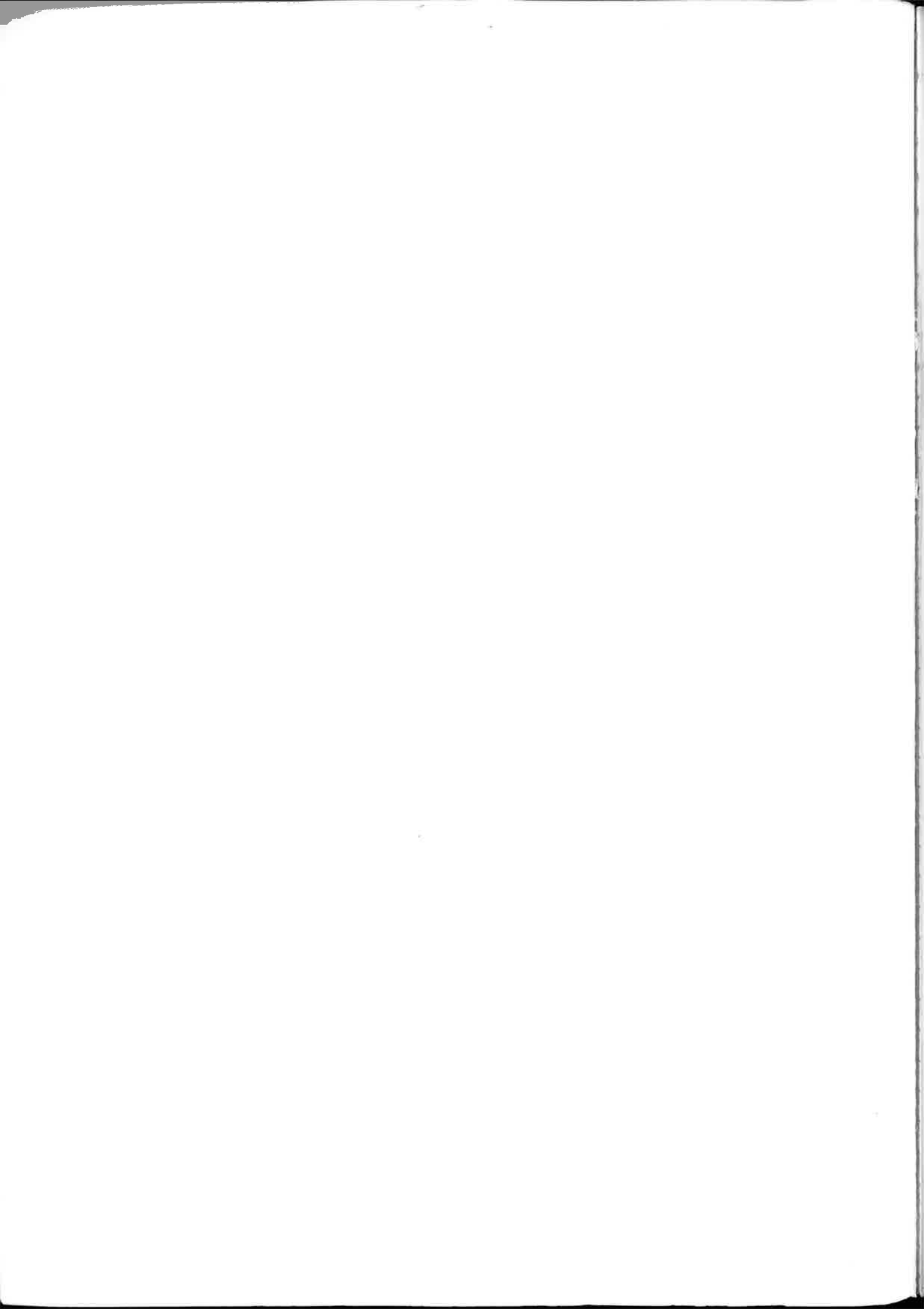


Computing at CERN in the 1990s

July 1989



Computing at CERN in the 1990s

Executive Summary
and
Summary Report of the Steering Committee

Computing for Experiments

CERN Data Networking Requirements in the
Nineties

Accelerator Computing Requirements in the Nineties

Computing for Engineering at CERN in the 1990s

MIS Computing in the 90s

Theory Computing in the 1990s

July 1989

Computing at CERN in the 1990s

Introduction

1. Introduction

2. The CERN Computing Centre

3. The CERN Computing Centre in the 1990s

4. The CERN Computing Centre in the 1990s

5. The CERN Computing Centre in the 1990s

6. The CERN Computing Centre in the 1990s

7. The CERN Computing Centre in the 1990s

Foreword

This report is the result of considerable effort on the part of many people. Computer users both on the CERN staff and from Universities and Institutes involved in the CERN programme have contributed. I should like to thank them all for the invaluable work they have done.

The main purpose of the report is to provide information and recommendations which can be used by the CERN management to build up a plan for computing for all aspects of work in the laboratory over approximately the next five years. As such it is the natural successor to the report "Computing at CERN in the LEP Era" published in 1983.

Since computing at CERN cannot be viewed in isolation, there is, inevitably, much that is relevant to users needs outside CERN. This applies in particular to networking and to the analysis of data from experiments. I hope therefore that the report will be of value to the HEP community as a whole and also that it will help in the efforts already being made to provide more coherence and better coordination in the general provision and use of computing resources for HEP in the future.

J.J. Thresher

Introduction to "Computing at CERN in the 1990s"

This report, entitled *Computing at CERN in the 1990s*, contains a review of the CERN computing needs up until the middle of the next decade. *Its purpose is to provide CERN Management with information and with recommendations which can be used for a policy and a plan for Computing at CERN in the next five years.* It is therefore the natural successor to the report entitled "Computing at CERN in the LEP Era", published in May 1983, and often referred to as the Green Book.

Although this is primarily a CERN report, a major part of it inevitably concerns the whole community of high energy physicists engaged in CERN-based experiments. The recommendations in many areas are therefore addressed not only to the CERN Management but also to the managements of all the European Institutions involved in the CERN experimental programme, and also to the High Energy Physics Computing Coordinating Committee, HEP-CCC. Indeed, the need for improved collaboration and pooling of computing resources at a European level is a major conclusion of this study.

This study was begun in 1988. Most of the work was carried out in six working groups which covered the following areas:-

- Computing for Accelerators
- Computing for Engineering
- Computing for Experiments
- Computing for Theorists
- Computing for MIS
- Networking

The working groups comprised both local and external experts in the various fields. An Editorial Board and a Steering Committee were responsible for setting in place the appropriate working groups, integrating the available material, and for defining an overall strategy. Considerable effort was spent in soliciting outside comment, organizing public meetings and polling country representatives, for example through ECFA.

The rôle of the Steering Committee¹, beyond setting in place the appropriate Working Groups in order that the user requirements were analyzed and documented, has been to review and integrate the material into an overall report for the management, including an Executive Summary.

This Full Report contains seven Reports and an Appendix:-

- The Report of the Steering Committee, containing:-
 - an Executive Summary of the major recommendations in the Summary Report,
 - a Summary Report, which reviews the working group reports, and presents the recommendations of the Steering Committee,
- the six complete Reports of the Working Groups above.
- an Appendix containing letters from Member State Representatives

Additional general background material is available as a CERN DD Report, number 89/9.

¹ The committee formed by the merger of the original Editorial Board and original Steering Committee

Composition of the Steering Committee

At the outset of work on this report early in 1988, the original Steering Committee mandated an Editorial Board to produce the set of Working Group reports. Once these reports were essentially complete, the Editorial Board merged with the Steering Committee in order to prepare the Report of the Steering Committee, and to produce this complete report *Computing at CERN in the 1990s*.

Steering Committee

R. Billinge: PS Division Leader

J.J. Bunn: Secretary

F. Dydak: EP Division Leader

J. Ferguson: Head of the MIS Unit

C.M.S. Jones: Editor of the Report from February 1989

J. Sacton: Professor, the Free University of Bruxelles, and Chairman of ECFA from 1984 to 1987.

J.J. Thresher: Chairman of the Steering Committee, and Director responsible for Computing

D.O. Williams: Chairman of the Editorial Board, and DD Division Leader from January 1989.

P. Zanella: DD Division Leader until 31st. December 1989, and member until March 1989

Editorial Board

J.V. Allaby: Chairman, Networking Working Group

R. Billinge: Chairman, Computing for Accelerators Working Group

R.K. Bock: Chairman, Computing for Experiments Working Group

J.J. Bunn: Secretary of the Editorial Board

T.E.O. Ericson: Chairman, Computing for Theorists Working Group

J. Ferguson: Chairman, Computing for Management Information Services Working Group

D. Jacobs: Chairman, Computing for Engineers Working Group

D.O. Williams: Chairman of the Editorial Board, and Editor of the Report until January 1989.

Table of Contents

Foreword	iii
Introduction to "Computing at CERN in the 1990s"	v
Composition of the Steering Committee	vi
Steering Committee	vi
Editorial Board	vi
Report 1: Executive Summary and Summary Report of the Steering Committee	1
Part 1: Executive Summary	1
Part 2: The Planning for Computing at CERN	7
Part 3: The Review of the Working Group Reports	13
Appendix A: Existing CERN Computing Facilities	45
Appendix B: The Cost of Computing at CERN over the past decade	47
Appendix C: An Estimate of Computing Capacity potentially available to CERN HEP in Europe	50
Report 2: Computing for Experiments	53
Part 1: Executive Summary	53
Part 2: Summary of Recommendations	55
Part 3: Data Acquisition and Real-Time computing	73
Part 4: Central and Distributed Computing, Software Support	93
Report 3: CERN Data Networking Requirements in the Nineties	183
Report 4: Accelerator Computing Requirements in the Nineties	199
Report 5: Computing for Engineering at CERN in the 1990s	213
Report 6: MIS Computing in the 90's	253
Report 7: Theory Computing in the 1990's	271
Appendix A: Letters from Member State Representatives	277

Computing at CERN in the 1990s

The Report of the Steering Committee including the Executive Summary

J. Allaby, R. Billinge, R. Bock, J. Bunn,
F. Dydak, T. Ericson, J. Ferguson, D. Jacobs,
C. Jones, J. Sacton, J. Thresher, D. Williams.

14th July 1989

Table of Contents

Introduction to the Report of the Steering Committee	v
Part 1: Executive Summary of the Recommendations of the Steering Committee	1
1.1 Introduction	1
1.2 The Resources for Computing in the Laboratory	1
1.3 Computing Policy at CERN and in European HEP	1
1.3.1 A fresh approach	1
1.3.2 Allocation of Production Resources	2
1.3.3 Large "Private" computers at CERN	2
1.4 The Key Elements of the Computing Plan	2
1.4.1 Long-distance networking	2
1.4.2 An evolution to more Distributed Computing	2
1.4.3 Data Storage and Handling	3
1.4.4 Processor Power	3
1.4.5 Computing at the Experiments	3
1.4.6 Software diversity	4
1.4.7 MIS	4
1.4.8 Computing for Engineers	4
1.4.9 Accelerator Control	4
1.4.10 Computing for Theorists	4
1.5 Staffing issues	5
1.6 Education/Training/Development	5
1.7 The Next Steps	6
Part 2: The Planning for Computing at CERN	7
2.1 Introduction	7
2.1.1 CERN General Computing Infrastructure	7
2.1.2 Computing Support for the Physics Programme	8
2.2 The Existing CERN Computing Facilities and Staffing	8
2.2.1 Physics production resources	8
2.2.2 Growth of Services	9
2.2.3 Staffing	10
2.3 The Cost of Computing at CERN over the past Decade.	10
Part 3: The Review of the Working Group Reports	13
3.1 Introduction	13
3.2 Computing for Experiments	13
3.2.1 Overview	13
3.2.2 The Major Challenges to be faced by the HEP Community	13
3.2.3 Recommended Resource Levels of the Working Group	14
3.2.4 The need for Overall Management of Computing for European HEP	15
3.2.4.1 The Coordination of Production Resources at a European Level	15
3.2.4.2 The Coordination of HEP Computing at a European Level	16
3.2.5 An Evolution to a Distributed Computing Environment	17
3.2.6 Data Storage and Handling	18

3.2.6.1	The Requirements	18
3.2.6.2	An Analysis of these Requirements	18
3.2.7	CPU Power	20
3.2.7.1	Conventional Mainframe Computing	20
3.2.7.2	Parallel computing facilities	20
3.2.7.3	The Cost of the Conventional Mainframe CPU Power	21
3.2.8	Support for Interactive Services and Workstations	22
3.2.8.1	Workstation Support Services	22
3.2.8.2	The Central VAX Cluster	22
3.2.9	The Challenge facing Real-Time Computing at the Experiments	22
3.2.10	A need to reduce Software Diversity	23
3.2.11	"Private" computing resources at CERN	24
3.3	CERN Data Networking Requirements in the Nineties	26
3.3.1	Overview	26
3.3.2	Requirements concerning On-going Activities	26
3.3.3	New Requirements	27
3.3.3.1	High-speed long distance links	27
3.3.3.2	Involvement in HEPnet and general purpose network activities	28
3.3.3.3	Special Solutions	28
3.3.4	Recommended Resource Levels of the Working Group	28
3.4	Computing for Accelerators	30
3.4.1	Overview	30
3.4.2	A General Model for Accelerator Computing	30
3.4.3	Increased dependence on Central Services	30
3.4.4	Standard Software Solutions	31
3.4.5	Network Support	31
3.4.6	Accelerator Design	31
3.4.7	Recommended Resource Levels of the Working Group	31
3.5	Computing for Engineers	33
3.5.1	Overview	33
3.5.2	Support for Microprocessor Users	34
3.5.3	CPU-intensive computations	34
3.5.3.1	Structural Analysis	34
3.5.3.2	Software Support for Field Computations	34
3.5.4	Analog Electronics CAE/CAD	34
3.5.5	Digital Electronics CAE/CAD	34
3.5.6	Database support for Engineering	35
3.5.7	Computer Aided Mechanical Engineering and related fields	35
3.5.8	Recommended Resource Levels of the Working Group	36
3.6	Management Information Services	38
3.6.1	Overview	38
3.6.2	The current situation	38
3.6.2.1	Administrative data processing	38
3.6.2.2	Office systems	39
3.6.3	The Working Group Report	39
3.6.4	Resources required	40
3.7	Computing for Theorists	41
3.7.1	Overview	41
3.7.2	General infrastructure	41
3.7.3	Use of Supercomputers and Parallel Systems	41
3.7.4	Recommended Resource Levels of the Working Group	42
3.8	General Remarks and Conclusions of the Review of the Working Group Reports	43
3.8.1	Staffing issues	43
3.8.2	Education/Training/Development	43

3.8.3 The Next Steps	44
Appendix A: Existing CERN Computing Facilities	45
A.1 The Computer Centre	45
A.2 Networking	46
A.3 Computers at Experiments	46
A.4 Computers at Accelerators	46
A.5 MIS	46
A.6 Macintoshes and PCs	46
A.7 Workstations	46
Appendix B: The Cost of Computing at CERN over the past decade	47
B.1 Total CERN Material Costs for Computing over the past decade	47
B.2 Total CERN Personnel Costs for Computing over the past decade	48
B.3 Total CERN Material and Personnel Costs	48
B.4 An Alternative Approach for the Total Costs of Computing	49
B.5 The Base Line Pattern of Expenditure	49
Appendix C: An Estimate of Computing Capacity potentially available to CERN HEP in Europe	50
C.1 Objectives and restrictions	50
C.2 Reported capacity in CERN units.	51

Introduction to the Report of the Steering Committee

This Report of the Steering Committee sets out not to make a précis of the content of the working group reports, which already contain summaries of their own, but rather to put the requirements into an overall picture. It indicates the key directions that emerge, and the consequences they have on resources both at CERN and at the level of European High Energy Physics, (HEP).

Part 1 of the report contains an Executive Summary of the recommendations of the Steering Committee.

Part 2 reviews the existing computing facilities, and overall trends and directions, distinguishing between the computing requirements of the CERN laboratory infrastructure and the needs of the experimental programme. Estimates are given of the total expenditure on computing within the laboratory over the past decade, and of the computing power available to HEP elsewhere in Europe.

Part 3 reviews the six working group reports and indicates the key directions and challenges. It makes recommendations to the managements concerned on resources, the evolution of activities in certain areas and next steps in the planning.

Finally, three appendices cover:-

- Existing CERN Computing Facilities
- The Cost of Computing at CERN over the past decade
- An Estimate of Computing Capacity potentially available to CERN HEP in Europe

Part 1

Executive Summary of the Recommendations of the Steering Committee

1.1 Introduction

This Executive Summary of the Recommendations of the Steering Committee necessarily abbreviates the arguments and recommendations of the Summary Report. In particular, 17 recommendations of the Summary Report have been compressed into one major recommendation 5, "Major Components of the Computing Plan", which cover the whole of section 1.3 below. For this reason, cross-references are given to the relevant sections of the Summary Report and to the Summary Report Recommendations (SRR).

1.2 The Resources for Computing in the Laboratory

The Steering Committee endorses the recommendations of the Working Group reports within their context as *requirements documents*. It recognizes the strong level of agreement and support from the user community for the requirements. However, the *total resources required for full implementation* of these recommendations, in particular those of the physics programme, are *extremely large relative to current expenditure and staffing levels*, and in some areas are a factor of two to five higher.

The Steering Committee is aware that there is a widespread shortage of money across the whole laboratory, and a commitment to reduce the overall staff complement. It understands that it is unrealistic, in the present CERN budgetary situation, to expect that all of the requested resources could be provided on the timescales mentioned. Nevertheless, it believes that it is essential to find a means to increase the resources allocated to computing in order to carry out some reasonable fraction of the recommendations.

Executive Summary Recommendation 1 : Resources

It is strongly recommended that the resources allocated to computing be increased above the present level, and that this be reviewed in the context of the overall CERN scientific programme. (SRR 1, section 3.2.3)

1.3 Computing Policy at CERN and in European HEP

1.3.1 A fresh approach

The organization required for collaboration on computing matters at the European level needs a fresh approach. The need for easy movement of programs and data between home institutes and accelerator laboratories, for improved facilities when working remotely, and to make good use of all available resources, makes improved collaboration vital.

Executive Summary Recommendation 2 : HEP-CCC

The present HEP-CCC should evolve into a fully representative body with a strong mandate to coordinate European HEP computing. In order to help accomplish this, a set of advisory groups which carry out the technical coordination in the different areas of HEP computing concerned is recommended. (SRR 3, section 3.2.4)

1.3.2 Allocation of Production Resources

We strongly support CERN's long-standing policy that, as a guideline, two-thirds of HEP computing resources related to the analysis of data from CERN experiments should be located in the home institutes. However, until now, application of this policy has been limited to pressure on the individual experiments and post-facto reporting. The demand for processing resources, especially from the LEP experiments, will be so considerable that an improved integration and planning on behalf of the experiments and the computing centres providing the resources is essential

Summary Report Recommendation 3 : European Resource Coordination

A new method is needed of organizing and planning the allocation of mainframe production resources amongst the larger European centres, for the major CERN experiments. This should seek to minimize the barriers to moving production work, for instance through improved telecommunications facilities and standardized working environments. (SRR 2, section 3.2.4)

1.3.3 Large "Private" computers at CERN

The presence of large private computers on the CERN site goes against the spirit of the one-third:two-third policy above. In practice it also places an additional load on CERN resources. The committee recognizes the difficulty of making hard rules in such cases, and of refusing to accept the installation of private resources at a time when CERN's general resources are scarce.

Executive Summary Recommendation 4 : Private Computing Facilities

The setting up or expansion of "private" computing facilities at CERN should be carefully controlled. In addition, the allocation of central facilities to an experiment should take into account the availability of "private" facilities. (SRR 15, section 3.2.11)

1.4 The Key Elements of the Computing Plan

Executive Summary Recommendation 5 : Major Components of the Computing Plan

We recommend that the plan for computing in the next five years must provide for the following key elements:-

1.4.1 Long-distance networking

1. **Communications facilities** between CERN and the collaborating institutes must be improved in order to facilitate decentralized physics data processing and analysis. The bandwidth proposed for the major paths has been expressed as one glass-fibre equivalent, currently meaning 2 Mbits/s. (SRR 17, section 3.3.3)

1.4.2 An evolution to more Distributed Computing

2. It is widely believed that physics analysis will profit from cooperative processing between powerful graphics workstations and the central computing facilities. (section 3.2.5)

3. This alone will require a **substantial upgrade of the on-site networking**, especially in achieved bandwidth. FDDI products have an essential rôle here. (SRR 16, section 3.3.2)
4. Additional, improved **centralized support for workstations** will be necessary. (SRR 9, section 3.2.8)

1.4.3 Data Storage and Handling

5. The central computer services will need to put particular emphasis on **high data volumes**. This will have major consequences on the **handling, storage and copying of data**, which will also need to be reflected in the larger processing centres outside CERN. The estimates for the overall requirements of the LEP experiments in terms of storage capacity and processing power contained in the MUSCLE report, have been largely confirmed by the working groups. From those estimates we make the following recommendation intended to cover all CERN requirements in 1991/1992.
 - Make provision for sufficient **cartridges and storage space to handle at least 50 Terabytes of data** (250,000 cartridges at current densities). (section 3.2.6)
 - Provide at CERN **an automated cartridge storage facility** capable of handling 8 Terabytes (40,000 cartridges). Similar facilities will also be necessary at the major outside institutes. (section 3.2.6)
 - Provide a new **cartridge/tape copying and distribution service at CERN**. (SRR 5, section 3.2.6)
 - Provide about **1 Terabyte of disk space at CERN** to allow fast access to the most active percentage of the data in a **central data repository**. (SRR 4, section 3.2.6)

1.4.4 Processor Power

6. The CERN physics programme will require at least 600 units¹ of CPU power by 1991-1992.
 - At least **150 units of batch processor power** should be provided on general-purpose computers in the **CERN Computer Centre**. This is to be compared with the existing 70/75 units available for batch. This power is necessary not only for conventional batch work, but in order that the centre can fulfil its increasingly demanding rôle of supporting distributed computing. (SRR 6, section 3.2.7)
 - This implies the availability of **twice this capacity, namely 300 units, in the outside institutes**. Taking into account centres equipped for production, this capacity is currently around 150 units. (Appendix C)
 - Additionally, we recommend that methods be investigated for providing substantial processing power in the form of **advanced parallel computing facilities** that offer potentially cheaper computing for relatively stable programs. Integration into normal Computer Centre operation is essential. (SRR 7, section 3.2.7)

Parallel computing systems have in addition considerable interest for on-line systems and applications for theoretical physics.

1.4.5 Computing at the Experiments

7. **Centralized support for hardware and software components** to be used in the construction of on-line systems at the experiments, as well as **design support** to aid their effective integration, will still be necessary, both for smaller experiments, and to provide components for the larger experiments. (SRR 12, section 3.2.9)

¹ The CERN unit of computing power is historically the power of an IBM 370/168 or a VAX 8600. As a rough indication, this may be taken as 3 IBM MIPS or 4 DEC VUPS.

8. The extreme environment of future hadron colliders requires major advances, and forces a tight coupling of the issues of detector design, detector digitization, data compression and triggering. We recommend setting up **coordinated pilot projects** for introducing new high-level design methodologies for both software and hardware, and for acquiring familiarity with the application of modern techniques to future experiments' real-time problems. (SRR 11, section 3.2.9)

1.4.6 Software diversity

9. The desire to reduce diversity by moving to **common software environments** is a strong theme from the working group reports, in areas ranging from operating systems to end-user packages, (SRR 14, section 3.2.10). In particular:
- **Unix** should be investigated as a common operating system, given that new initiatives (e.g. OSF, AIX) are encouraging. (SRR 13, section 3.2.10)
 - The use of the **ORACLE** DBMS should be consolidated at CERN and generalized within the community. (sections 3.2.10, 3.4.1, 3.5.6, 3.6.3)
 - A set of **CASE** tools should be selected, and a suitable license negotiated so that all participating institutes may benefit. (section 3.2.10)

1.4.7 MIS

10. The **MIS** environment at CERN should be **modernized and made coherent**. It should include a **CERN corporate data model** using a **unique DBMS (ORACLE)**. This implies the replacement of the currently used major corporate applications. A comprehensive **electronic forms handling system** should be implemented and **EDI** should be used, where possible, for corporate applications. (SRR 22 and 23, section 3.6.3)

1.4.8 Computing for Engineers

11. Experiments and accelerators depend on **front-line technology** that can only be achieved through the use of **CAE tools**. CERN is late investing in this area. (SRR 20, section 3.5.1)
12. A large percentage of CERN staff were educated before computer-aided tools became part of the syllabus. If CERN is to gain full benefit from CAE techniques, **staff at supervisory as well as technical levels will need training in modern design methodologies**. (SRR 21, section 3.5.1)

1.4.9 Accelerator Control

13. **Accelerator control systems** should be increasingly based on **commercial hardware and software**, using a layered approach with common design elements and a **three tier LAN**. (sections 3.4.1, 3.4.2)
14. **Strong central support** for the computing needs of the accelerator divisions must be consolidated. (SRR 19, section 3.4.7)

1.4.10 Computing for Theorists

15. A policy decision is required as to the fraction of CERN's vector processing power that should be devoted to theoreticians working at CERN. It is recommended that this fraction should not exceed 10% of the total available. (SRR 24, section 3.7.3)

1.5 Staffing issues

The laboratory is seriously under-staffed for the computing demands it is already asked to fulfil. The working groups make large additional demands for staff. (On face value, they ask for around 100 posts). Furthermore, in view of the overall compression of the CERN staff complement that is in progress, it will be particularly difficult to make the necessary increase in informatics staff numbers in the short term. The committee therefore feels that:-

Executive Summary Recommendation 6 : Staffing Five Year Plan

It is essential to agree on a staged plan for an acceptable increase in informatics staffing levels across the whole laboratory over, say, five years. (SRR 25, section 3.8.1)

A small, but valuable, contribution to alleviating the staffing problem is being made through students, coopérants, joint projects with industry, etc.. Some increased activity of this nature may be possible, although it is clear that the essential part of the problem cannot be solved by this means.

Executive Summary Recommendation 7 : Alternative Staffing Methods

We recommend investigation of further alternative methods for informatics staffing, such as the appointment of technical associates from companies and universities, and increased involvement of graduate students.

Joint development projects and staff exchanges with industry and academia, especially in the fields of advanced electronics and computer science, should be actively encouraged. (SRR 26, section 3.8.1)

1.6 Education/Training/Development

Despite the urgent pressures from the above situation, a massive effort is necessary in order to train and re-train existing CERN staff in informatics. If the missing computer-literate staff are to be found, a substantial part will have to come by internal movement and re-training. *This will require a training programme on a scale comparable with that found in industrial companies which have successfully tackled this problem before CERN.* New attitudes and extra resources will be required to produce the desired results.

Executive Summary Recommendation 8 : Training

CERN must set-up an organization devoted to a massive training and re-training of the existing staff in informatics technology. (SRR 27, section 3.8.2)

In an associated area of training:-

Executive Summary Recommendation 9 : Advanced Development Projects

Progress in the field of computing is so rapid that CERN must devote a fraction of its staff resources to work on **advanced development projects** that are only expected to show benefits in the medium to long-term. It is also vital that the long-term CERN staff involved with computing are kept well informed of the latest techniques. (SRR 28, section 3.8.2)

1.7 The Next Steps

Executive Summary Recommendation 10 : The Next Steps

1. The CERN management should define a policy framework and level of resources, money and staff, that it will make available for computing in the next five years.
2. The staff providing the computing services should then be allowed to optimize the services that they provide within the policy framework and the agreed level of resources. This should be done in consultation with the users, both those on the CERN staff and those from the outside institutes. (SRR 29, section 3.8.3)

Part 2

The Planning for Computing at CERN

2.1 Introduction

In making the plans for "Computing at CERN in the 1990s", it is necessary to consider **two major components**. On the one hand, there is a responsibility to provide a **healthy and broad computing infrastructure** for the 3500 or so staff, plus the users based in the laboratory. On the other hand, there is a special responsibility to provide for the **specific needs of the experimental physics programme** which involves over 4000 visiting scientists, coming from some 150 institutes in Western Europe, and also from further afield.

It is well understood that this report comes at a difficult time for several reasons. The physics programme of the next few years will be dominated by the LEP experiments whose computing needs are very demanding in resources. The laboratory computing infrastructure is still in a build-up phase. The level of agreement on the requirements in both these areas within the user community has been remarkably good. Yet these requirements come at a time when an overall policy to reduce CERN staff, leading to minimal recruitment, coincides with a period of unprecedented budgetary difficulties for the Laboratory.

Planning for CERN cannot be made in isolation from planning for the whole European High Energy Physics community. The challenge posed by LEP is such that the level of coordination and cooperation at a European level will have to increase significantly, especially if the above budgetary difficulties are to be overcome. Explicitly, the required computing resources at CERN will have to be more than matched by increased resources in the collaborating institutes, and specific steps, particularly in the area of telecommunications, will be required to facilitate the exploitation of these distributed resources.

In a planning report of this nature, it is necessary first to establish the "base-line" on which the proposed changes can be built. This chapter contains a brief review of the computing facilities already in place at CERN at the beginning of 1989, together with the staffing levels and the money spent on computing across the laboratory, over the past few years. *For more than a decade there has been an exponential growth of the installed computing power, a strong growth in the range of facilities and in numbers of users served, whilst the staff numbers and budget have remained essentially constant.*

Amongst the initiatives that have made this possible is the establishment of exceptional relationships with major manufacturers. These have led to outstanding financial conditions and to joint projects that have opened doors to advanced technology and industry collaboration. CERN has something to offer in return *provided it is prepared to keep its computing environment at the forefront of technology.* On the other hand, if CERN is unable to maintain a leading "shop window" position, it will lose the interest of these industry collaborations, the important discounts and the advanced technology that it needs.

2.1.1 CERN General Computing Infrastructure

The conscious provision for the CERN **general computing infrastructure** is relatively new. Ten years ago, such general facilities were provided only as a by-product of those installed for physics number-crunching or data acquisition. Their suitability varied with the task in hand. WYLBUR provided for simple general needs, but could not prevent isolated initiatives, leading to inefficiency typified by the existence of eleven different makes of word processor on-site. This being recognized, some actions have been taken:

- A Communications Systems group was created in the DD Division, and the Telecommunications Board established to coordinate laboratory-wide networking.

- Management Information Services (MIS), was established as a DD unit in 1987, with an MIS Board to coordinate activities across the laboratory and a high level policy committee (SCAIP).
- Computing Support for Engineers (CSE) was established as a DD unit in 1988, coordinated by a high level committee (MACCSE).

There are similarities between these three infrastructure areas. Firstly, the planning for all three has been reviewed and agreed relatively recently by the CERN management. Consequently, few major changes are revealed in this report. Secondly, working organizational structures for overseeing the activities are in place. Thirdly, the overriding policy of reducing CERN numbers hits hardest activities which are being built-up. Paradoxically, the recognition of manpower wasted in isolated efforts leads to the need for organized support, for which the staff are difficult to find. *In this situation of few new posts, it seems clear that the major part of these support posts can only come from reorganization and re-training within the laboratory itself.*

2.1.2 Computing Support for the Physics Programme

Computing support for the physics programme has changed significantly in the past five years. The range of machines, systems and facilities has had to expand dramatically, along with the physicists' requirements. The complexity and degree of interconnection has obliged the support staff to become highly specialized, and some have moved from this traditional area into the newer areas above.

Faced with this situation, many solutions have been tried; stopping services, pushing more on to the users, greater standardization, use of commercial packages etc. *Solutions at a European level, albeit needing greater organization than in the past, can no longer be ignored.*

The next five years of the physics programme will be dominated by LEP. The high statistics required will lead, both at CERN and in the home institutes, to high demands on processor power, as well as on data storage and handling, all of which will be major financial investments. Just as the manpower available to HEP at the European level will have to be pooled and organized better than in the past, so will these major computing resources, e.g. the processor time and data handling facilities available for physics production in the major European centres. Both will be scarce.

2.2 The Existing CERN Computing Facilities and Staffing

The rest of this chapter records the existing computing resources, staffing and costs as a "base-line" for the discussions of the working group reports that follow. A summary review of the computing facilities presently installed across the laboratory can be found in Appendix A. Some overall considerations are discussed here.

2.2.1 Physics production resources

A summary of the computer power currently installed in the CERN Computer Centre is given in Table 1. As a comparison, Appendix C contains an estimate of the corresponding power available to HEP for data processing in the outside European centres. Note that this estimate attempts to make the difficult distinction between resources organized for major physics production computing and the smaller machines and workstations that are more readily used for physics analysis. Taking into account the fraction of the CERN capacity that is available for physics production, say 70 to 75 units, the potential capacity in member state institutes during 1989 matches reasonably well the "one third/two thirds rule". Some further capacity is available to the LEP experiments outside the member states, and in "private computing facilities" at CERN.

Table 1: Computer Power installed for HEP in the CERN Computer Centre, Easter 1989

Computer	Number of Processors	Scalar Power per processor	Vectors	Total Scalar Power
Cray X/MP-48	4	8	Yes	32 CERN units
IBM 3090-600E	6	6.5	6 VFs	39 CERN units
Siemens 7890	2	6.5	No	13 CERN units
VAX Cluster	3	1.5	No	≈ 4.5 CERN units
CERN Total				≈ 88.5 CERN units

2.2.2 Growth of Services

Figure 1 on page 10 illustrates the processor hours delivered annually to physics production in the CERN Computer Centre in CERN units² over three decades. It does not include use of the Central VAX Cluster, nor the use of the Cray by Theory Division. It also shows the growth of available disk space in the Centre. Exponential growth of the installed mainframe processor power has been sustained over more than 30 years, with the power doubling roughly every 3.5 years since 1968. To meet demand the installed disk space has been doubling somewhat faster, roughly every 2.3 years over the past decade.

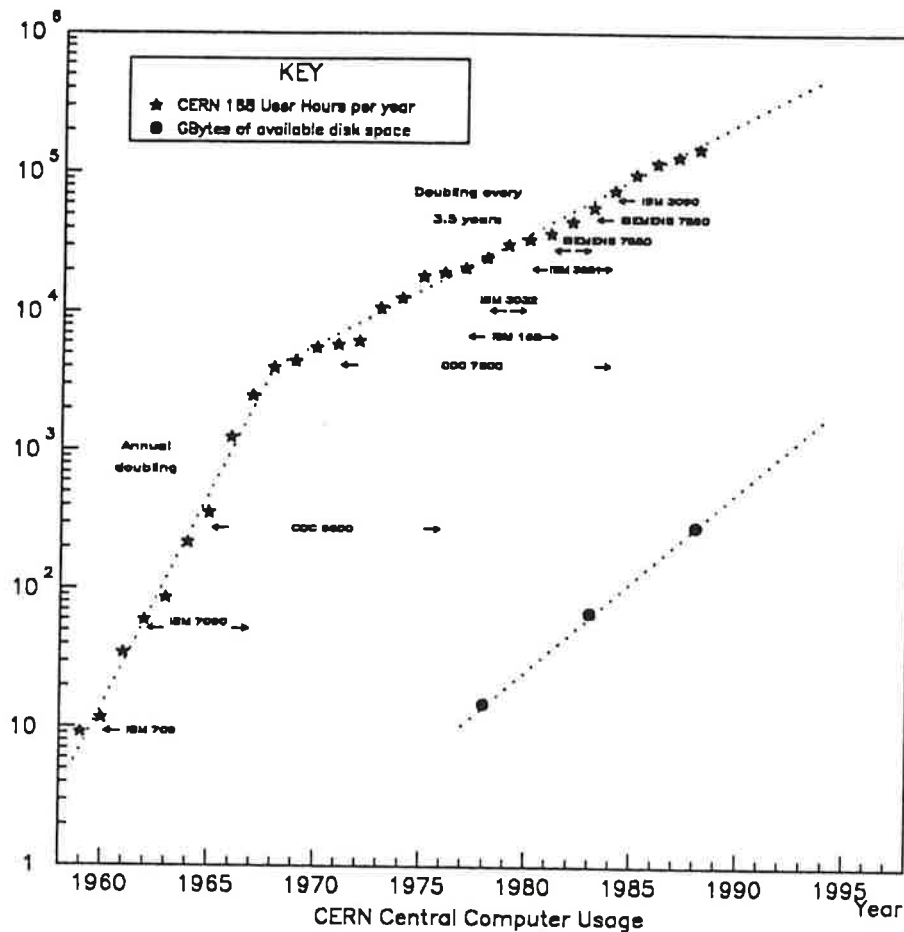
The number of users of the central computer services is high. As an example, Figure 2 shows that the current number of weekly users of the IBM service is more than twice the peak number at the height of the WYLBUR service five years ago. The VAX services have grown from essentially zero in the same period, to a level where 1500 users log in every week. The Cray has been added as an entirely new service (but the CDC services have been stopped). The range of other services provided and requiring support has also changed rather dramatically. Tape and cartridge mounts frequently exceed 1500 per day and are steadily increasing. In February 1989, the IBM 3090-600E did more work in one week than the IBM 370-168 did in the whole of 1978!

The above examples all refer to the computer power and range of services within the traditional computer centre. These are perhaps the easiest to quantify. *Yet it is above all in the past five years or so that an explosion of many other forms of computer services across the CERN site has taken place.* As an example, networking, both on and off the site, has seen dramatic growth, and the services thus made available have become a normal and fundamental requirement of every computer user whatever his or her level of skill. Standardization on one form of network has been far from possible.

A second example, namely the growth in numbers of intelligent workstations, (from Macs/PCs to high-end Apollos/VAXstations), which have been installed on CERN office desks, is a witness of the spread of computers throughout the laboratory, and their use in an ever increasing range of activities. There are certainly more than 3000 such workstations installed in offices of staff and visitors on-site, and the rate of installation is increasing. Taking as an example the pattern observed in the USA, one can expect this number to exceed one per person, and hence to grow beyond 6000.

² The CERN unit of computing power is historically the power of an IBM 370/168 or a VAX 8600. As a rough indication, this may be taken as 3 IBM MIPS or 4 DEC VUPS.

Figure 1: The Growth in Usage of the CERN Central Computers.



2.2.3 Staffing

Without attempting to quantify further all areas of computing growth at CERN, the point that needs to be recorded as a base-line in this report is that *right across the laboratory there is a general shortage of computer-literate staff, be they for direct support or for applications support, which has been estimated on several occasions to be between 80 and 100 persons.*

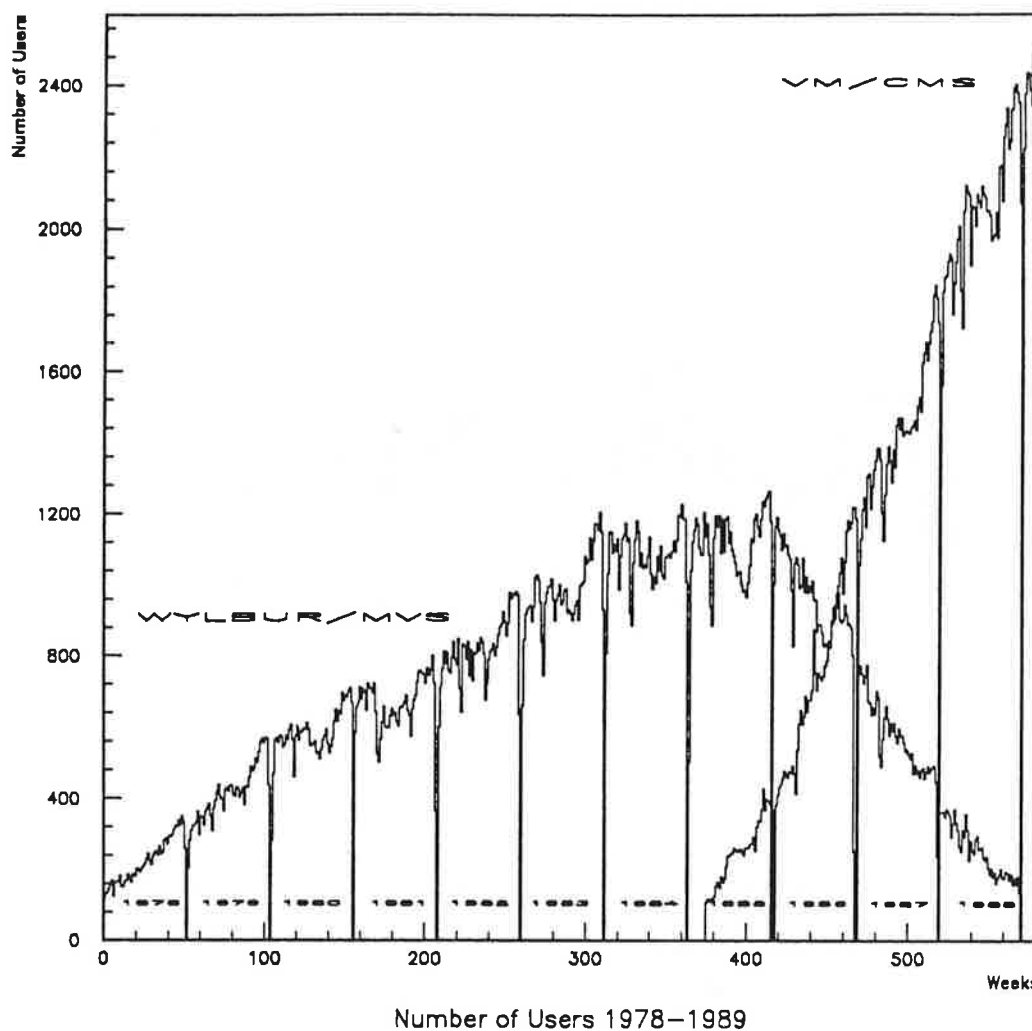
2.3 The Cost of Computing at CERN over the past Decade.

A summary review of the costs of computing at CERN over the past decade is given in Appendix B. Some general remarks are made in this section. The numbers used here and in the Appendix are taken from the CERN accounts. Where the numbers have been corrected to present day figures, the correction factors used have been the appropriate CERN official indices.

Note that the CERN accounts contain no explicit record of the "costs of computing" as such. Indeed, the term "costs of computing" needs definition in such a discussion. Three examples in common usage are, the costs of the computer centre, the full costs of the DD Division, or the costs coded in the CERN accounts under headings labelled "computers" for all divisions.

Figure 3 illustrates the yearly expenditures attributed to DD Central Computers in the accounts over twenty years, and is shown here as a companion to Figure 1. The graph also gives the breakdown into

Figure 2: Number of Users of the IBM service per week over 11 years.

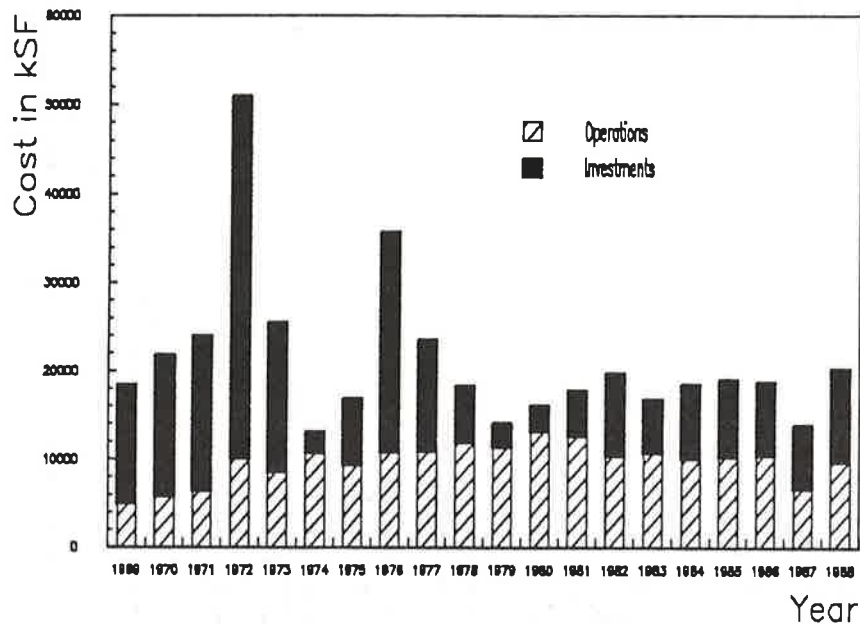


investments and operations, showing notably the acquisition of the CDC 7600 in 1972 and the new IBM system in 1976/77. These peaks apart, the annual costs have remained unchanged over the whole period, whilst the installed computer power has increased exponentially.

This rather flat expenditure profile across the years is also typical of the numbers found in Appendix B for the computing costs across the laboratory as a whole. In order to make this comparison for all divisions in a consistent manner over many years, one is obliged to take a strict definition from the accounts, (see Appendix for this definition). The recorded annual material costs for computing for the whole laboratory thus obtained, have remained (in 1988 prices) between 27 and 37 MSF with average 31 MSF over the past decade (see Table 8 on page 47 and Figure 4 on page 48). The costs of computing within DD division, again as defined in the above manner, have been a constant two thirds of the total CERN costs, averaging 21 MSF/annum. Note that this definition accounts typically for between 84 and 92% of the total DD expenditure on all items.

The personnel costs for computing are harder to determine. Whilst they are well defined and known for DD Division, their definition becomes hazier in other divisions. For example, the distinction between physicist or engineer and computer support staff can be difficult. Similarly, it is not easy to understand what fraction of the accelerator control staff should be considered as computer specialists.

Figure 3: Material costs labelled Central Computers in the CERN accounts. (Numbers corrected to 1989 prices).



Computer Centre Expenditure

Nonetheless, an estimate is made in the Appendix of these personnel costs, using a method again chosen to allow a comparison across the years. This suggests that the personnel costs for computing have changed little over the years. Whilst the accounts show that personnel costs of the DD Division have in real terms stayed remarkably constant, this is unlikely to be true for the rest of CERN where the numbers of people involved in computing activities have grown substantially.

Using the above definitions of material and personnel costs, one may sum the total costs attributed to computing across the years, (Figure 4 on page 48), producing a flat distribution with average 55 MSF over the last decade. This number certainly represents a lower bound. A second method in the Appendix takes better into account the above remark on personnel outside DD Division, and estimates the upper bound for the current total costs of computing across the whole laboratory to be 85 MSF/year.

Part 3

The Review of the Working Group Reports

3.1 Introduction

This part of the report reviews the reports of the six working groups, attempting to integrate them into a consistent picture. It gives an overview of the key directions to be followed, and the consequences on resources both at CERN and at the level of European High Energy Physics. The position and attitude of the Steering Committee in making such choices and judgements is summarized in the following points.

- *The Steering Committee endorses the recommendations of the Working Group reports within their context as requirements documents.* It recognizes the strong level of agreement and support from the user community for the requirements. However, *the total resources required for full implementation of these recommendations, in particular those of the physics programme, are extremely large relative to current expenditure and staffing levels, and in some areas are a factor of two to five higher.*
- The Steering Committee is aware that there is a widespread shortage of money across the whole laboratory, and a commitment to reduce the overall staff complement. It understands that it is unrealistic, in the present CERN budgetary situation, to expect that all of the requested resources could be provided on the timescales mentioned. Nevertheless, it believes that *it is essential to find a means to increase the resources allocated to computing in order to carry out some reasonable fraction of the recommendations.*
- In addition, improved collaboration and coordination in computing matters at a European level is essential in order to optimize the use of such resources as can be made available at CERN and in the outside institutes.

The Recommendations that appear in boxes in the review are those of the Steering Committee.

3.2 Computing for Experiments

3.2.1 Overview

"Computing for Experiments" is the largest of the working group reports and by definition the one most concerned with the physics programme. The report is the product of a number of smaller working groups whose recommendations were grouped into two main areas, separating data acquisition and all other aspects of the data processing. Part 2 of the Computing for Experiments report contains 22 recommendations, themselves a summary of a larger number of recommendations in the main body. Part 3 reviews costs and manpower whilst making further remarks on priorities. An Executive Summary of Computing for Experiments is provided in Part 1.

It is evident that detailed remarks on all aspects of this 120 page report cannot be made here, nor is it attempted to make a further précis. The major directions and challenges to be faced by the community are identified, and input to the next step of the planning is given.

3.2.2 The Major Challenges to be faced by the HEP Community

The picture emerging for the years ahead is rather consistent and well agreed upon both by physicists who use computers and computing specialists. The key elements are summarized as follows:-

- **a major evolution towards more distributed computing** across the HEP community,
- **very substantial requirements for data storage**, and associated requirements for the management and interchange of that data,

- considerable power to process and analyze that data, requiring both increasingly powerful general purpose computers and new initiatives to exploit the inherent parallelism of the event data,
- a new and rather unknown challenge in data acquisition to be faced for the next generation of experiments such as at LHC,
- a general and strong desire to reduce software diversity wherever possible,
- a need for improved coordination at a European level.

The consequences of the above directions are reviewed in the next sections.

3.2.3 Recommended Resource Levels of the Working Group

The user requirements have been clearly analysed and presented in the working group report. They have received substantial support in public user meetings and in general discussion in the community. They present the Steering Committee and the managements of the various high energy physics institutes and funding bodies with a number of serious problems, not least *a strong opinion that the present levels of funding and staffing are at a marginal or inadequate level*. The costs and manpower suggested represent significant increases over present levels, (see Tables 1 and 2 of the WG report, deliberately NOT summarized out of context here). Whilst the above tables concern mainly CERN, there are clearly described consequences for the community outside. Apart from substantial increases in processing power and disks, there are high-speed line and workstation costs to be borne by the institutes or the experiments.

It is in this report above all that the greatest proposed resource increases from the present levels are to be found, especially when the consequences on the outside institutes are included.

Any study of requirements can make estimates for the additional costing and manpower of new activities. The difficulty of such a study comes in relating those requests to the current budgets and staffing; to fold in what can be achieved by stopping activities, redeployment, imaginative negotiation, joint projects with industry, spreading the plan over more years etc. Indeed, this is exactly the task facing the laboratory.

The magnitude of this task is considerable. The working group costs their requirements at 128 MSF spread over three years, including 26 MSF estimated for high speed links to be paid by the outside institutes. At face value, the resultant CERN cost of 102 MSF is a factor three to four higher than the DD capital expenditure which has averaged 8.6 MSF/annum over the past 9 years, and 10/11 MSF/annum in the past two. This money has to be found in addition to the 5 MSF outstanding payment for the most recent IBM 3090 upgrade which is already spread until 1992.

Similar remarks apply to the staffing estimates of this working group. They suggest that for the manpower resources *"located at CERN and serving the whole of the CERN-connected community"* an increase of new staff totalling 34 in 1989, plus 20 in 1990, plus 17 in 1991 should be made. Given the probable level of new recruitment that will be available to the laboratory as a whole, it is clear that if these numbers are to be in any way approached they will have to come in major part through diversion of staff currently employed on other CERN activities, including physics or physics support. Furthermore, all other creative forms of obtaining manpower for these activities will have to be explored, such as industry involvement, European collaboration amongst HEP institutes, and closer involvement of computing expertise in the universities.

The conclusion is inescapable that the resources required to fulfil the programme foreseen by this working group represent, even after drastic cuts, a significant change from the previous decade. The principle reasons for this change are to be found in the growth in data rates and event sizes at the LEP collider, and also in the advances in computing technology, which have opened up new possibilities and allowed greater sophistication in data acquisition and detector systems.

The working group, in recognizing this increased demand, states:-

"computing expenditures provide at least equally significant potential in added physics returns when compared to spending for other sectors such as accelerators or detector construction."

Summary Report Recommendation 1 : Resources

It is strongly recommended that the resources allocated to computing at CERN be increased above the present level, and that this be reviewed in the context of the overall CERN scientific programme.

3.2.4 The need for Overall Management of Computing for European HEP

Many recommendations made in the working group's report concern the needs of the HEP community as opposed to specific needs at CERN. The CERN management has traditionally seen its rôle as that of the host laboratory, and has been cautious in addressing computing problems at the community level. The 'High-energy Physics Computing Coordination Committee' (HEP-CCC) exists as an independent, high-level representative body for community computing matters, and a number of community-wide technical groups exist, for example ECFA SG5³ and HEPVM. The working group expresses a strong view that the time has come for a more concerted effort in this area. *This growing demand is recognized, and the view of the working group endorsed. Indeed, improved coordinated planning at a European level is a major theme of "Computing at CERN in the 90s".*

This leads to the following two recommendations:-

3.2.4.1 The Coordination of Production Resources at a European Level

CERN's long-standing policy that, as a guideline, two-thirds of HEP computing resources related to CERN experiments should be located in the home institutes is strongly endorsed. It is understood that, until now, application of this policy has been limited to pressure on the individual experiments and post-facto reporting. This has led to periods where production resources available to HEP or provided specifically for HEP in outside computer centres have been under-utilized. On the other hand, note is taken of the strong view in parts of the physics community that the movement of physics production from CERN to outside institutes has traditionally introduced significant delays which are considered unacceptable at certain times, e.g. at the start-up of a new experiment when most of the key personnel involved are unavoidably at CERN.

The obstacles preventing the movement of production can be reduced. The provision of high-speed links, described in this chapter and the next, would make a significant jump in the convenience of remote working and remove many of the barriers and delays. The success of the HEPVM collaboration in an important fraction of the sites involved has already demonstrated that the time-wasting and frustrating differences between sites can be significantly reduced. There is now good evidence that attention to these convenience factors does indeed reduce the dependence on the CERN computer centre.

³ ECFA SG5 has recently evolved into the HEPnet Requirements Committee (HRC), partner to the HEPnet Technical Committee (HTC).

In the light of the very significant production resources, (e.g. CPU power, disk space, cartridge units and robots), being requested for the physics programme ahead, and the severe difficulties in funding those resources, *an improved effort must go into making the best possible use of such production resources as can be made available at the European level.* This will require a new coordination and planning on behalf of the experiments and the computing centres providing the resources.

Summary Report Recommendation 2 : European Resource Coordination

A new method is needed of organizing and planning the allocation of mainframe production resources amongst the larger European centres, for the major CERN experiments. This should seek to minimize the barriers to moving production work, for instance through improved telecommunications facilities and standardized working environments.

The initiative for setting this up should come from CERN, but it requires full support at the European level, and should involve those responsible for the management and allocation of resources in the major computer centres.

3.2.4.2 The Coordination of HEP Computing at a European Level

Summary Report Recommendation 3 : HEP-CCC

The present HEP-CCC should evolve into a fully representative body with a strong mandate to coordinate European HEP computing. In order to help accomplish this, a set of advisory groups which carry out the technical coordination in the different areas of HEP computing concerned is recommended.

In reality this is not as trivial as it perhaps sounds, as the present HEP-CCC is itself aware. The national representative physicists on the committee often have only indirect control over the resources and running of the computer centres available to HEP. Where university centres are concerned, the special needs of HEP can conflict with overriding university needs. Coordination of working environments, software and hardware choices for HEP conflict with other bodies trying to achieve coordination in a different manner at say a national level. The size of the committee would be a problem if all members states and funding bodies were to be directly represented. These problems are being addressed already by the HEP-CCC. Part of the solution can come through the establishment of further strong technical groups reporting to the HEP-CCC.

3.2.5 An Evolution to a Distributed Computing Environment

Arguably the most significant technical change ahead is the evolution towards a more distributed computing environment across the whole community, clearly described in chapter 10 of the working group report:-

"The next five years will be a challenging period in the development of computing for offline High Energy Physics, as it progresses towards a distributed computing environment, exploiting the fast-developing technologies of workstations and networks, whilst at the same time ensuring that the needs of the early years of LEP are met effectively.

The main elements in the evolution of centrally organized computing, at CERN, in national and regional centres, and in the universities, are:

- *The growth of batch capacity for LEP data analysis, involving the expansion of conventional general-purpose facilities at CERN and in major centres, the large-scale exploitation of cheap parallel computers, and the integration of these with cartridge tape and disk storage facilities on a scale new to HEP.*
- *The establishment of personal workstations as the principal tools for general-purpose interactive computing. The technological and economic conditions are ripe for this development, but the challenge will be to develop the support services, the network management and the tools to integrate workstations with each other and with central services for file management, batch processing, etc.*
- *The development of distributed co-operative processing for interactive data analysis, with different components of a single program executing simultaneously on a powerful workstation and on one or more central CPU and file servers.*
- *The introduction of high performance wide-area networking, which will enable the above developments to evolve on a European-wide basis and allow the individual physicist to work effectively from his home institute.*

It is important that this evolution take place smoothly. While we must press ahead with the introduction of new technologies for distributed computing, we must also learn how to exploit these effectively, and be careful not to abandon prematurely existing proven services."

This is clearly and widely agreed as the model for the years ahead for "physics off-line computing". It identifies a number of components and areas for work, notably:-

1. on the CERN site and in the large outside centres:-

- the central computers must become a **central data repository** on a large scale
- the increasing use of **personal workstations** as the principle tools for interactive computing will require a **centrally-organized workstation support service**,
- the **cooperative processing** between workstations and mainframe hosts will require greatly increased end-to-end bandwidth between the workstations and the mainframe, attainable using **FDDI communications products**, but only with **special attention to the overall system aspects**, including the computers at the ends of the links,
- **effort to manage these links** such that they become a transparent part of the environment for the end user,
- **essential integration of the whole system** that cannot be purchased,
- a general education process,

2. and wider than the local site:-

- **high-speed links** between CERN and the large outside centres,
- **tools for organizing the distribution of data files amongst the centres,**
- **coordination between experiments and the large centres,** in order to make as efficient use of the available resources as possible.

The recommendations as to specific areas of the above model are to be found in the following pages.

3.2.6 Data Storage and Handling

3.2.6.1 The Requirements

Data Storage and Handling are also part of the above model, but are treated separately because of the particularly important rôle they have to play. The physics to be done at LEP involves high statistics. Modern detectors produce event sizes measured in hundreds of kilobytes. The high requirements for LEP data storage were first identified in the MUSCLE⁴ report. One number to remember is that **each LEP experiment is expected to have an accumulated data volume of approximately 7000 Gigabytes by "1991"**⁵. The working group, in section 9.2.4, argues that the MUSCLE numbers are lower limits, particularly with respect to "master DSTs", and adds to them the predictions of non-LEP experiments. They consider the distribution of data to universities and regional centres and the copies that implies. They suggest 300,000 active tape cartridges⁶ in 1992 at CERN and at least that number outside CERN. This leads logically to a **proposal for automated handling of cartridges, with a capacity of 40,000 cartridges in 1991 at CERN,** with similar robots in major centres outside CERN. They propose a **centrally supported cartridge/tape copying and distribution service** such that cartridges can arrive at regional centres and universities within days of submission of an appropriately authorized electronic request. Regional centres should have similar facilities.

Following the MUSCLE report, they argue that **100 GBytes of disk space for active data per LEP experiment by end-1991,** representing 1.4% of the expected data volume, are necessary in order to facilitate physics analysis and eliminate a large fraction of the cartridge mounts. In this they follow the MUSCLE model of a central hierarchical data repository, from which subsets of data may be extracted for different physics analysis needs elsewhere on workstations or off-site. They plausibly suggest that the disk storage for all of CERN should be 8 to 10 times bigger than one LEP experiment. Adding system needs, this leads to their requirement of **1000 GBytes of disk storage by 1991.**

Tools for **managing the data in the above implied storage hierarchy** are another requirement, and should include automatic migration of data based on activity.

3.2.6.2 An Analysis of these Requirements

The cost of the **disk space** requested at CERN is realistically estimated by the working group to be 31 MSF, including the required controllers etc. Given this large sum, which represents about 30% of the total estimated costs of this working group, (excluding the long distance lines), three obvious questions are:

⁴ The Computing Needs of the LEP Experiments, CERN/DD/88/1, January 1988

⁵ The MUSCLE Report defined "1991" as being the time when the total number of Z^0 events that each LEP experiment has accumulated reaches 10 million. In practice this is not now expected to occur before the end of 1992.

⁶ at the current 200 MBytes capacity

- How would the physics suffer if only some fraction of this capacity were to be installed?
- Will the major improvements in magnetic disk capacity predicted for the coming years cause the cost/MB to drop further than expected?
- Normal academic discounts are not enough. Can a frame be constructed within which an exceptional discount be negotiated via bulk purchase or overall major "system bid"?

The following recommendation is suggested as a reasonable method to deal with this situation. It would allow some form of optimized planning which could approach the desired capacity at one third of the working group's proposed cost. It will allow some relief to a situation that will become very difficult by the end of 1989. Should the capacity which it is possible to install with this reduced budget prove largely insufficient in later years, not only will the costs per byte have fallen but the budgetary situation should have improved.

Summary Report Recommendation 4 : Disk Space

A fixed budget of the order of 2.5 MSF/annum over each of the next 4 years should be allocated at CERN for this important acquisition, such that the planning and the disk space eventually obtained can be optimized.

This is clearly a crucial matter for the detailed planning that follows.

The **automated cartridge handling device** is an important proposal which no longer requires a recommendation in this report in that the matter is already in hand. An attractive joint study agreement with a manufacturer to develop such a robot to CERN's specific needs has recently been signed, following Finance Committee approval. If successful this should fulfil the needs of the working group for a cost around 40% of their estimated 4 MSF. The funding for the robot must, nonetheless, be included in the budget planning. Note that it is possible that future announcements of higher capacity cartridges, may require money above that presently foreseen.

The exact levels of money and manpower required for a **cartridge copying and distribution service** are clearly dependent on scale and the exact requirements of the physics groups. This needs further investigation before it can be correctly costed. The working group's estimate of 0.5 MSF per year should in principle be sufficient to cover additional contract manpower, new hardware and maintenance money for a new service. The first step is to use existing equipment until it proves inadequate. One must not forget however that this is a service on which the LEP experiments are counting, and one which is completely consistent with ideas to make as best use as possible of the computing resources outside CERN.

The growing user pressure to provide for video 8 Exabyte technology as a second standard, in addition to the established 3480 cartridges (which now represent 85% of the mounts in the computer centre) is difficult to resist. However, the consequences on the costs and staffing of a centrally organized service using this technology could be considerable.

Summary Report Recommendation 5 : Cartridge Copying Service

A centrally supported cartridge/tape copying and distribution service should be provided at CERN.

3.2.7 CPU Power

3.2.7.1 Conventional Mainframe Computing

It is widely believed that *conventional mainframe computers will continue to be vital components of the HEP computing environment*. The transition from terminal to workstations will certainly allow more productive physics analysis, but it is very unlikely to reduce the steady growth on demand for mainframe power, (in fact the contrary is likely).

The model in which the central computer is the master file server and base for cooperative processing with the workstations leads to high bandwidth demands for I/O which will require significant CPU power to drive it. Present measurements indicate that an aggregate throughput from the central data store to the workstations of 1 MB/s would require more than 50% of one of the IBM central processors.

The rôle of the central mainframe, apart from traditional general-purpose computing and part of physics production with its unavoidable development and debugging, will thus be as the central data repository and data handling system, exploiting a storage hierarchy based on the cartridge robot and the above disk storage, and feeding the workstations and parallel farms described below.

Summary Report Recommendation 6 : CPU Power

The CERN physics programme will require at least 600 CERN units of CPU power by 1991-1992. It is proposed that at least 150 such units be installed in the form of conventional mainframe systems on the CERN site and that hence at least 300 units be installed and accessible to HEP in the outside centres. The balance should be provided by parallel farms, as discussed below.

By comparison the power installed in the CERN Computer Centre at the beginning of 1989 was 88.5 units, and that installed in other European centres available to HEP and equipped for production was around 150 units. (See Table 1 on page 9 and Appendix C). The recommendation is thus to continue to double the mainframe capacity in the next 3.5 years, following the line of the graph in Figure 1 on page 10.

3.2.7.2 Parallel computing facilities

There is a conviction within HEP that *the inherent parallel nature of the event-related computing problems should be exploited*. It is proposed that the balance of the CPU power requested above be provided in the form of **parallel computing facilities, which must be integrated into the normal operation of the large centres concerned**, and which can then be used for production work with relatively stable problems. On the other hand, there is no general agreement within the community as to the true cost-effectiveness of, and to how much manpower has in reality been consumed by, the various existing initiatives using parallel farms in an off-line environment up to now, (whatever their success within specific experiments or in the on-line area).

Summary Report Recommendation 7 : Parallel Farms

It is recommended that a new generation of parallel processor farms should be investigated, using as far as possible commercially-based products, by setting up a pilot project. The farms should be fully interfaced to the existing off-line environment and operational procedures.

Bearing in mind the targeted operation, the investigations will have to take into account the costs and manpower required to build and to integrate the system into normal operation, as well as reliability, available de-bugging tools, overall efficiency, etc. An initial budget of 0.5MSF for the pilot project should be reviewed according to success.

Whilst such investigations should exploit the obvious "trivial" parallelism at the event level, they should also address what other forms of parallelism could be profitable. Note that such parallel devices clearly have other areas of similar and related interest such as on-line and theoretical physics applications.

The working group suggests that around 150 units of processing power be provided in this fashion by 1991/1992, with one third of this at CERN. The working group notes that if one is unable to provide around 150 units of this type in a cost-effective and sufficiently useful manner, then this power will have to be provided in the form of conventional general purpose computers.

3.2.7.3 The Cost of the Conventional Mainframe CPU Power

At CERN the proposal is approximately to double the existing capacity by 1991-1992. The cost depends on which computers one changes, and when, and strongly on possible price negotiations. The working group has foreseen 47 MSF for this expansion over 3 years.

In order to make an estimate of what might be possible, it should be recorded that CERN has spent some 29 MSF since December 1984 in order to acquire a net increase of approximately 43 units of IBM-compatible CPU power, together with 150 GBytes of disk space, and other peripherals such as cartridge tapes and communications controllers. Assuming the mainframes themselves represented about 75% of the cost, one has spent 22 MSF on 43 computing units over 4 years. It must be emphasized that these were skillfully negotiated prices, far below the levels obtainable with normal academic discounts, and it may not be possible to match them in the future.

As a rough number, one may assume that the price of such mainframes falls by a factor 2 over 4 years, and hence one should be able to obtain a net increase of around 80 mainframe units with no peripherals for around 20 MSF spread over the 4 years, assuming a continuation of advantageous arrangements made with the manufacturers concerned. Of course, there will also be a need for peripherals other than the disks, which have been specified separately. This suggests that a sum of 5 to 6 MSF/annum each year for the four years 1990 to 1993 is the minimum that could provide the necessary capacity by the end of 1992 (which is the latest point suggested by the working group). The payments still to be made on the IBM 3090/600E until the year 1992 are in addition to the above calculation.

The above estimation is based exclusively on IBM-compatible computers. It should be noted that the Cray version would be quite different. The present Cray is on a lease/rental contract fixed at 3.6 MSF/annum. In mid-1990 CERN will have the option to change this for a Cray 2 at the same price, renegotiate for a Y-MP or a successor machine, or continue with the present X-MP. The choices will depend in part on the degree of success of use of vectorization by the physicists. Obviously, further models can be made, based on other manufacturers, or combinations of manufacturers.

Summary Report Recommendation 8 : Mainframe Funding

A capital investment budget line of 6 MSF/annum for the four years 1989/1992 for the provision of upgrades to the conventional mainframe power is recommended. This should cover the mainframes and major peripherals except disks which are specified elsewhere.

3.2.8 Support for Interactive Services and Workstations

3.2.8.1 Workstation Support Services

The evolution to a distributed computing environment will establish personal workstations as the principal tools for general-purpose interactive computing. The challenge will be to develop the support services, the network management and the tools to integrate workstations with each other and with central services for file management, batch processing, etc. Management of workstation system software, communications, and files will become a task of considerable size, probably comparable with the management of a mainframe computer centre. The working group recommends that both at CERN and in the regional centres these tasks be centrally managed for the benefit of the end users.

Summary Report Recommendation 9 : Workstation Support Services

Centrally managed services should be established to provide workstation support services for a limited range of workstations on the CERN site.

3.2.8.2 The Central VAX Cluster

The central VAX Cluster acts as a focal point for the integration of the distributed VAXes on-site, including the VMS-based personal workstations that are an important component of the above section, as well as serving a large community of visitors who are VAX-based in their home institutes. In addition, they have been of great value for interactive program development and for applications that only run, or are best performed on a VAX. Experimentalists working with VAXes used for data acquisition have profited greatly from the service.

The Cluster has provided a minor part of CERN's batch capacity and should not be upgraded in order to provide batch processing power, unless there are substantial technical changes in the future which would make this advantageous. However, the rôle that the Cluster plays as a centre of competence and focal point for the management of distributed VAXes and in particular the VAX workstations will increase in importance as described above.

Summary Report Recommendation 10 : Central VAX Cluster

It is recommended that the central VAX Cluster continue to be upgraded by a reasonable level of investment, in order to act as a VAX competence centre and management base for the VAX workstations.

Note that the possible migration from VMS to Ultrix is another parameter which could influence the development of the cluster.

3.2.9 The Challenge facing Real-Time Computing at the Experiments

The Computing for Experiments working group is of the opinion that, whilst current generations of experiments are relatively well understood in this area, *a major challenge on the horizon is that of constructing trigger and data acquisition systems for high rate hadron colliders such as the LHC*. They consider that a great deal of research, development and learning of techniques from outside HEP is necessary. This must be started as soon as possible. It leads to recommendations how to face this challenge, to set up the correct framework and to train the people concerned, summarized here as:-

Summary Report Recommendation 11 : Real-Time Projects

It is recommended to set up **coordinated pilot projects** for introducing new high-level design methodologies for both software and hardware, and for acquiring familiarity with the application of modern techniques to future experiments' real-time problems.

The working group believes that smaller experiments have needs in this area which are often as technically demanding as those of larger experiments. These needs arise both from complex triggers and from data acquisition systems containing mixtures of on-line computers, workstations, and microprocessors. Due to their size, the smaller experiments are less able to provide for themselves. They are thus interested in standard, off-the-shelf systems. The view that *if CERN wishes to continue its current broad programme, it is important that these experiments are not neglected* is fully endorsed.

Summary Report Recommendation 12 : Real-Time Support

Centralized support for hardware and software components to be used in the construction of on-line systems at the experiments, as well as **design support** to aid their effective integration, will still be necessary, both for smaller experiments, and to provide components for the larger experiments.

Very sophisticated design skills are required in these areas; the application of VLSI techniques to trigger and data acquisition problems is now of crucial importance. Additionally, the standardization of real-time kernels and associated system software should be followed closely, and taken into account in planning future data acquisition systems.

3.2.10 A need to reduce Software Diversity

The general desire to attack the diversity of software can be found in many of the working group reports. In the report for experiments the strongest recommendation in this area proposes *"to start a Unix service on the central IBM service at CERN as soon as adequate software is available from IBM."* An initial small-scale service should receive further resources according to user demand. In addition there are a number of recommendations aimed at improved organization and coordination of the HEP software community.

This is not the first time that a pilot Unix service on the IBM has been proposed. What is new is that there are now sufficient good reasons for such a pilot investigation to make sense. Apart from the penetration of Unix into the HEP world through its appearance on workstations, the Cray and various other machines, several of the traditional arguments against Unix in this rôle are decreasing in validity. One may note in this context the previously bad reputation of FORTRAN compilers available with Unix, which gave inefficient run time performance compared with those available in proprietary systems (e.g. MVS, SCOPE, VMS, VM). There has been no seriously supported version of Unix on IBM systems. The recent initiatives of IBM in the direction of the AIX operating system, (AIX will run the same FORTRAN compiler as VM/CMS and MVS), and the creation of the Open Systems Foundation⁷ by many manufacturers, show an increased interest in Unix from the computing world in general.

Nonetheless, one is obliged to take into account the special needs of the HEP community, particularly in the areas of very substantial tape handling and of batch schedulers. These are not yet guaranteed to

⁷ There are currently nine sponsoring manufacturers and more than 90 full members.

be solved by AIX on the IBM, even if they have indeed been solved on the Cray. Indeed the requirements for Unix running on such large systems have been tackled in relatively few places as yet. However, referring back to Figure 2 on page 11, it is instructive to note that the VM/CMS service at CERN was a user requirement during 1983, a Directorate decision during 1984, and an embryonic user service during 1985.

Summary Report Recommendation 13 : Pilot Unix Service

It is recommended that a pilot Unix service is started on the central IBM computers at CERN as soon as practicable.

The recommendations concerning the organization and coordination of software issues in the HEP software community include the proposal for a **Software Support Committee**. Such a body could help to reduce the diversity of software by making successful commercial and HEP-specific products more widely available. It could also evaluate the benefits of writing HEP-specific tools against the cost of manpower, were a suitable commercial package not available.

Summary Report Recommendation 14 : Software Support Committee

HEP-CCC should establish a Software Support Committee as one of its technical working groups.

The crucial rôle that data base techniques and access methods have to play in the management of the huge volumes of data expected in the next decade of computing at CERN is recognized. The working group recommends *"the consolidation of the use of ORACLE in the HEP community for this purpose."* This recommendation is strengthened by the request for HEP-wide coordination of the definition, writing and installation of data base orientated software for large physics experiments. It is felt that this coordination could also come under the umbrella of the Software Support Committee.

Additionally, the working group report highlights the success of some larger experiments in using software engineering tools to design and implement their physics codes. They recommend the support of software tools and methodologies⁸, including their selection and evaluation. *It is believed that, in view of the distributed nature of the software development effort amongst collaborating institutes, such tools will become essential for coherent and correct code.*

3.2.11 "Private" computing resources at CERN

The presence of large private computers on the CERN site goes against the spirit of the one-third:two-third policy. In practice they also place an additional load on CERN resources, and there are a number of other negative remarks that are frequently made against them. The interconnection and networking of such centres with the CERN Computer Centre causes arguments of unfair assistance to one experiment over another, whereas artificial restrictions in this area often make technical nonsense and lend weight to the argument that the money made available for the private centre would have been better invested for the community as a whole via the central pot.

To these traditional arguments against private centres, one may now add the complication of deciding what exactly is a private centre. In the case of a major mainframe this is clear, but the potential processor power of a networked set of powerful workstations, or an emulator farm, may well be

⁸ A methodology is a set of "methods", typically purchased as a collection of CASE tools.

higher. The confusion is all the greater for emulator farms used both on-line during an experimental run and off-line outside runs. *The difficulty of making hard rules in such cases is recognized.*

A further aspect of this difficult subject concerns the possible inconsistency of refusing private resources, and hence increasing the general CERN demand for resources, at a time when CERN funding for those resources is too scarce to meet the essential requirements for computing resources on site. This is definitely true in the case of funding for private centres that would otherwise not be available to CERN, or at least to the experiments concerned.

Individual cases should be resolved at the Directorate level.

Summary Report Recommendation 15 : Private Computing Facilities

The setting up or expansion of "private" computing facilities at CERN should be carefully controlled. In addition, the allocation of central facilities to an experiment should take into account the availability of "private" facilities.

3.3 CERN Data Networking Requirements in the Nineties

3.3.1 Overview

The existing mechanisms and structures for networking are well established and working, such that *the overall theme is one of continuity and expansion, and of reinforcement of the existing data communications strategy*. This does not imply that there is little to be done. The successful implementation of the networking requirements below is of the greatest importance to HEP computing and to its evolution towards cooperative distribution of function. Two major thrusts are clear:

- The networking between the outside institutes and CERN must make a substantial jump in bandwidth (to at least 2 Mbits/s) in order to provide new possibilities and to overcome barriers to distributed working.
- One of the world's largest Ethernets exists at CERN; this needs continuous upgrading in bandwidth. Presently, FDDI⁹ products offer the best way by which to proceed.

3.3.2 Requirements concerning On-going Activities

The working group report gives a general model of computing activity. It estimates the requirements up to 1993 and lists the services required. It also reviews the state of the requirements made in 1983. There is close agreement with the body of this work.

It notes that three of the requirements from 1983, namely:-

- Implementation of a general-purpose high-speed backbone network interconnecting LANs.
- Use of industrial products for the backbone, and integration with other facilities including digital telephone.
- Use of CCITT transmission standards across the site i.e. G.703 series TDM (Time Division Multiplex) equipment and optical fibres.

have not yet been implemented in their entirety due to the lack of suitable (100 MHz) industrial products for the backbone, but an intermediate backbone network has been created using slower products. G.703 equipment has been widely installed, mainly by SPS/ACC for the LEP site, and it is used both for the digital telephone and for the intermediate backbone.

Summary Report Recommendation 16 : FDDI

The Committee considers that CERN is well placed to exploit the FDDI products that will appear imminently on the market in order to provide this general-purpose high-speed backbone, and recommends active continuation of the FDDI programme.

The Committee endorses the working group's view on protocol requirements which may be summarized as, *protocols should wherever possible be manufacturer independent and standardized, but ad hoc or proprietary protocols should be used when performance or timescales demand them.*

⁹ Fibre Distributed Data Interface, ANSI X3T9.5, ISO DIS 9314

3.3.3 New Requirements

This section contains a summary of the new requirements from the working group which are felt to be most important.

3.3.3.1 High-speed long distance links

The working group recommends that

"access to CERN at speeds of at least 2 Mbits/s is required from major HEP sites in Europe, to allow decentralized analysis of the LEP data, as soon as this analysis starts in earnest. Higher speeds (equivalent to those attainable on a LAN) will be required as soon as practicable, to allow effective use of workstations over geographical distances."

The provision of these high speed links is of the highest priority, and initiatives that are being actively pursued are strongly supported. It is recognized that the implementation of 2 Mbit/s links on CERN site, namely CERNET, enabled a significant change in the number of users able to work away from the Computer Centre. The recent success of T1 (1.5 Mbits/s) long distance links within NSFNET¹⁰ in the United States is additional evidence of the quantum jump in user productivity that can be achieved through such bandwidth. Progress towards the desired goal of the HEP community to overcome geographic barriers, and to facilitate decentralized physics data processing/analysis, is dependent on the existence of such links. The required bandwidth has been expressed as "one glass-fibre equivalent".

Summary Report Recommendation 17 : Long-distance Links

Improved communication facilities between CERN and the collaborating institutes must be provided. The major paths should operate at a bandwidth of at least 2 Mbits/s.

The provision of these links goes beyond the cost of the lines themselves, itself a challenge at present PTT tariffs. Within the major centres non-negligible funding must also be found for infrastructure costs such as multiplexors and communications controllers. These costs could well amount to around 150 kSF for each end of such a high-speed line, assuming the likely model of multiple protocols on each line. CERN has a policy that the costs of such outside lines are payed by the outside institute itself. Up to now this has consisted mainly of PTT bills.

Summary Report Recommendation 18 : Payment Policy

The borderline of payment for the various components necessary to install these important lines at CERN must be urgently clarified.

The smaller national institutes will only profit if they too can obtain funding for high bandwidth into their national/regional centre.

¹⁰ The network of the American National Science Foundation, currently upgrading some links to DS3 (45 Mbits/s).

3.3.3.2 Involvement in HEPnet and general purpose network activities

The Committee endorses the associated recommendation of the working group, namely that:-

CERN should take an active role in HEPnet activities, without losing sight of the priorities of the CERN experimental programme. CERN should also, in conjunction with HEPnet, collaborate with general-purpose network activities (RARE, COSINE, EARN, EUNET) as long as this is beneficial to the HEP community.

3.3.3.3 Special Solutions

At a second level of priority, support is given for the two working group recommendations which are both aimed at stopping or discouraging solutions requiring special, (or worse, home-made), hardware or software, namely:-

- *"Dependency on CERN-designed protocols should be eliminated as rapidly as possible, in favour of industrially supported protocols.*
- *The general-purpose LAN infrastructure should take into account specific needs of the large experiments, to avoid their needing to implement their own extended LANs across the sites."*

In this context, CERNET is being actively phased out. The old TITN networks for the PS and SPS control systems should also be phased out.

3.3.4 Recommended Resource Levels of the Working Group

A summary of the resources estimated by the working group is given below. For an explanation of the financial model used and of the assumptions made the reader is referred to chapter 6 of the working group report. The financial model in Table 2 contains a certain contingency after 1990 for unforeseen capital expenditure. Money for the STK digital exchange and on-line networks (controls, DAQ) have been excluded.

Table 2: Summary of Financial Totals from the Networking Report

	1989	1990	1991	1992	1993
Capital Expenditure (MSF)	3.5	3.5	4.5	3.5	3.5
Operations Expenditure (MSF)	5.5 ¹	5.5 ¹	5.5 ¹	5.5 ¹	5.5 ¹
Totals	9.0	9.0	9.0	9.0	9.0

1. including 4.1 MSF paid by DA for telephone, telex, X.25, EARN, EUNET

Two major items are not included in the above table. The first is the cost of the high-speed external links, assumed here to be paid by the institutes which request them, following established CERN policy for conventional direct links. At normal PTT tariffs this would require an increase on the community's operations budget rising to 12 MSF in 1992. This is covered further in the Computing for Experiments report. The second item is the infrastructure in the computer centres at the end of the links, intelligent multiplexors and communications controllers.

The detailed arguments and staff movements leading to Table 3 are not reproduced here. Note that staff for on-line networks, TV, audio and safety communications are specifically excluded, and that it is assumed that all departures will be compensated, with the exception of some telephonists.

These are considered and conservative figures for the staff required to run the services.

Table 3: Summary of Staffing levels from the Networking Report

	DD ¹	PS	SPS	LEP	EP/EF	All Divs
Current Staffing Level	46	1.5	6	0	1.5	55
Future Staffing Needs	52	1.5	1	1	2	57.5

1. all DD groups, not only CS, and includes the telephone service.

3.4 Computing for Accelerators

3.4.1 Overview

The computing requirements of the accelerator divisions bring together four somewhat different areas. Firstly, there is the well-established accelerator controls area, where the accelerator divisions have long been self-sufficient providers of their own computing services, and the specialist computer staff have had little dependency on others. In the second area of accelerator design, the simulation work has been largely performed using the general Computer Centre services. The third area concerns the accelerator engineers whose requirements clearly overlap with those in the Computing for Engineers working group report. Finally, the required management information services infrastructure largely overlaps with the MIS working group report.

It is worth noting that the accelerator divisions have used database management systems in earnest for the accelerator planning and installation for many years. Indeed, one can make the general observation that those divisions have played an important rôle in pioneering areas of the computing infrastructure at CERN such as CAE, MIS and DBMS.

In overall terms, this is a well-established part of the computing programme of the laboratory, and the structures required to deal with it are mainly in place. The trend is thus towards a **common model for accelerator computing in the future**, thus reducing the diversity of solutions in the different divisions. The manpower needed is dependent on the success of this common model.

The major themes particular to the accelerator computing requirements are thus:-

- a general model for accelerator computing, involving a three-level structure of local area networks (LANs),
- a major shift in the direction of accelerator divisions becoming users of more generally supplied services,
- a general trend towards standard software solutions, in particular Unix and ORACLE.

3.4.2 A General Model for Accelerator Computing

The working group has developed a schematic view comprising three levels of LAN. The first layer is a **conventional office LAN** with workstations in every divisional office, and servers of various forms in the corridors. The second LAN is used for the **Main Control network**, and is interconnected to the office LAN and to the central computer services. The third layer, named **regional LANs**, replaces traditional equipment control highways, such as serial CAMAC. The regional LANs will be connected in a controlled fashion to the other LANs.

3.4.3 Increased dependence on Central Services

The computing infrastructure of the laboratory has been almost totally based on the needs of the experimental physicists, a situation which is necessarily changing. As this infrastructure broadens, there is a general trend for the accelerator divisions to stop the specialist services they have provided for themselves in the past, and to use the general laboratory services. Particular examples are a centralized accelerator database, central networking support, central computers and their file backup services.

This is not a trivial point. The accelerator divisions have some particular demands to make on the general services which must be taken into account. The provision of an 24-hour database service for all accelerator data, including controls, needs consideration and management, and probably additional resources. Whilst the Computer Centre has often scheduled its services around the needs of a specific experiment(s), it has never directly had to give special priority to the needs of the accelerator and its engineers during running or development periods.

3.4.4 Standard Software Solutions

The working group believes that "support for open systems such as Unix, and products that will run on a variety of support engines, should be continued and even given preference." It is clear that this allows CERN to retain a multi-vendor software policy across the site, whilst allowing common solutions in the accelerator divisions.

The communications facilities of the accelerator divisions should be unified using TCP/IP protocols.

A single central database management system for all accelerators, preferably running on the same computers, is proposed. *The accelerator divisions should capitalize on their early investment in ORACLE.*

Other standards such as EUCLID and AUTOCAD are proposed in line with the recommendations of the Computing for Engineers working group report.

3.4.5 Network Support

It is recommended that a much greater fraction of the network facilities should be provided by DD Division, including some particular demands of this working group. The substantial commitment of the accelerator divisions to IBM-compatible PCs requires good interconnection of these PCs and full network facilities, including coordination of choice of software, negotiation of site licences, distribution, and grouped ordering. This shift of support was not foreseen by the working group on networking, and should be referred to the Telecommunications Board.

3.4.6 Accelerator Design

While accelerator design and modelling will move in part to workstations, there will still be a requirement for use of the central computer services. Substantial use of computers of the power of the Cray will be required for design of complex accelerators.

3.4.7 Recommended Resource Levels of the Working Group

Table 4: Staffing Requirements (not including Fellows etc.)

Area	Present	Future	New Posts
Controls	134	134	0 ¹
Databases	15	15	0 ²
General Infrastructure	10	15	5 ³
Note 1.	The need for increased effort in end user applications will be compensated by rationalization elsewhere.		
Note 2.	Reallocation within the Accelerator Divisions.		
Note 3.	Possibly obtained by rationalization across other areas.		

The numbers in Table 4 assume that there will be strong central support from DD. This support is currently used in the fields of: networking, ORACLE system support, CSE including PRIAM for microprocessors, MIS including PC-Shop and applications development, and central computing support on VM.

It is important that the accelerator divisions can count on the continuity of this central support and indeed on its consolidation to take these needs more into account than in the past. If not, the numbers in the table would have to be increased and an inefficient situation would result. The division between central support and local user support is always hard to delineate and must be decided according to the individual circumstances of each application. Nevertheless, the benefits and savings are such as to make the effort worthwhile.

Summary Report Recommendation 19 : Central support for accelerator divisions

Strong central support for the computing needs of the accelerator divisions must be consolidated. The level of support and its interaction with local support must be decided on a case by case basis.

No additional money above current levels is requested in the working group report.

3.5 Computing for Engineers

3.5.1 Overview

Computing for Engineers is a relatively young subject at CERN. Whilst the authors of the 1983 Green Book were aware of the growing problem of providing for CERN's engineers, there was no separate working group assigned to this area, and the associated topics were but lightly considered. Five years later, the extent to which the engineering community at CERN has come to use computers as tools, and the vastly increased sophistication of its needs, has led to a specific working group on this topic and to a separate chapter in this report. It has also led to strong pressure from the engineering community for an increase in the organized computing support which they receive, and which today is at a rather low level of manpower.

In this time, an Advisory Committee on Computing Support for Engineering, (ACCSE), prepared a report for the CERN management.¹¹ Following discussion of this report, CERN has decided (June 1988) that a plan for coordinating and enhancing the facilities for engineers be implemented. The resources for that implementation, and especially the manpower requested, remain in large part still to be found, and hence the implementation is still incomplete.

The working group report is in some sense a later review of the material covered by ACCSE. The requirements and requests upon the laboratory in the two reports are thus similar and the CERN management should find no surprises. The working group divided itself into seven areas and sub-groups as follows:-

- Support for Microprocessor Users
- Software Support for Field Calculations
- Digital Electronics CAE/CAD
- Computer Aided Mechanical Engineering and Related Fields
- Structural Analysis
- Analog Electronics CAE/CAD
- Database Support for Engineering

The working group report consists largely of the seven separate sub-group reports. Nonetheless, there are two related main themes which emerge from the report as a whole, and form the basis for recommendations:-

Summary Report Recommendation 20 : CAE Tools

Experiments and accelerators depend on front-line technology that can only be achieved through the use of CAE tools. CERN is late investing in this area and is not yet investing enough in proportion with what is spent on the more traditional areas of computing.

Summary Report Recommendation 21 : CAE Training

A large percentage of CERN staff were educated before computer-aided tools became part of the syllabus. If CERN is to gain full benefit from CAE techniques, staff at supervisory as well as technical levels will need training in modern design methodologies.

¹¹ P-G. Innocenti, "Report of the Advisory Committee on Computing Support for Engineering", 30th. September 1987.

3.5.2 Support for Microprocessor Users

Microprocessor support was set in place before some of the other engineering support activities, and as a result there is a clear message that the current activity is a good model and should continue in the same spirit in the future. There are a number of technical recommendations, and a request for more resources at the level of 2.5 more people in central support and an upgrade of the central support computer costing 500 to 700 kSF.

3.5.3 CPU-intensive computations

3.5.3.1 Structural Analysis

Historically, a number of structural analysis packages have been installed on the central computers, but the present trend is to use either CASTEM¹² or ANSYS¹³ in particular. Support for the thirty or so users is on a somewhat *ad hoc* or goodwill basis. Besides recommendations of improved bridges and integration with other areas such as field calculations and CAD systems, the major request here is for user support at the level of one post (programmer/engineer) in central DD support.

3.5.3.2 Software Support for Field Computations

A story similar to that above is to be found for field computations. A number of packages have been installed over the years, but the major problems with them, (lack of pre- and post-processing, and lack of good 3-D programs), are being solved by the introduction of TOSCA¹⁴ and ANSYS. These commercial packages tend to be expensive by academic standards, and an annual budget of the order of 100 kSF is required. A larger problem comes, as above, from the request that the present 50 users need central informatics support at the level of two persons.

This area of support has been particularly neglected and merits urgent attention.

3.5.4 Analog Electronics CAE/CAD

The major message from the working group is that it is high time that the laboratory started working seriously towards microelectronics, and got itself better organized in this area. Training has a capital rôle to play in introducing this new technology. Also required are centrally supported modern CAE packages with good handling of exotic components. The establishment of relevant model libraries for these tools is very important.

3.5.5 Digital Electronics CAE/CAD

The subject is covered in a great deal of detail in the sub-group report and not repeated here. The messages too are similar. The conclusion is perhaps that this is mainly an organization question, and one in which one needs continuous investment in money and manpower.

¹² from CEA Saclay.

¹³ from Swanson Analysis Systems Inc.

¹⁴ from the Vector Field (VF) company, Oxford.

3.5.6 Database support for Engineering

The usefulness of database management systems (DBMS) has been generally accepted by engineers. The rapidly expanding user community and number of applications around the ORACLE relational DBMS system suggests that ORACLE will continue for a long time at CERN, there being no significantly better product on the market at present.

There is a strong request coming mainly from the accelerator divisions to provide a database service 24 hours a day, seven days a week. The intimate scheduling of the central computers to fit with the accelerator division needs (as opposed to the needs of the experimental physicists) does indeed have new implications on the running of the central service, and these need further discussion and study.

The major message from this area is that an **application support team** is needed to complement the existing DD-based database service support. Its tasks would include coordination of application development, database service management, application development expertise, data organization and other tasks. **The bulk of the application development effort itself can continue to be provided by the engineering groups concerned under the guidance of this team.**

3.5.7 Computer Aided Mechanical Engineering and related fields

Computer Aided Design (CAD) made a major step with the decision in 1982 to install the true 3-D solid modelling system EUCLID from Matra Datavision, and this powerful design system has been supplemented recently by a commercially widespread and simpler drawing package AUTOCAD, running on IBM-PC compatibles. This aspect of design is thus reasonably well covered technically but lacks support manpower (especially organized support for AUTOCAD).

The strongest new message coming from this area is a need for **better integration of engineering tasks**, linking drawing tools to databases and also to structural analysis.

3.5.8 Recommended Resource Levels of the Working Group

Table 5: Support Staff Requirements (not including Fellows etc.)

Figures in brackets are changes known or anticipated in 1989.

Area	Present DD	Present non-DD	Recommended
Microprocessors	4(+1)	—	6 + Fellow (DD)
Mechanical CAE	4.5	—	7.5 (DD)
Field Calculations	—	—	1 (DD)
Analog El. CAE	2(+1)	—	4 (DD)
Digital El. CAE	3(+1)	2.5(-1.5?)	6.5-9.5 (Total)
Database appl.	—	4	5 (Total)
Totals.	13.5(+3)	6.5(-1.5?)	30-33

Note 1. In the non-DD column, only those numbers are shown which affect the recommendations. Indications of the other numbers may be found in the ACCSE report.

Note 2. The range of numbers recommended for Digital El. CAE corresponds at the low end to maintaining the status quo and at the high end to satisfying the estimated future needs of the lab.

Note 3. The doubt in the non-DD effort for Digital El. CAE corresponds to the lingering uncertainty concerning P-CAD support. This should be resolved.

Clearly, the provision of the increased staffing is an on-going exercise which was recently reviewed by the laboratory management.

The increases will come partly from internal reassignment. A small number of internal posts have already been allocated and are, in part, being filled. There is a transition between the situation where engineers write their own code and one in which they become users of more sophisticated commercial packages for which centrally-organized support is necessary. It remains to be seen whether reassignment will be successful in providing all the necessary expertise in these areas.

Table 5: Central Funding Requirements (in KSF) - excluding training

Figures in brackets are 1989 requests.

Area	DD ops 88(89)	DD investments 88(89)	Other projects 88(89)	Recommended annual ops/investment
Microprocessors	-	400(400)	-	250/250
Mechanical CAE	140(0)	0(700)	1250(700)[LEP]	700/700
Field Calculations	-	0(100)	-	100
Analog El. CAE	30(90)	0(130)	-	100/150
Digital El. CAE	190(210)	0(450)	-	300-500/635-1145
Database appl.	-	-	700(1300)[LEP]	400/800
Totals.	360	400	1950	1850-2050/ 2535-3045

- Note 1. Investment figures are averaged. There must be enough flexibility to allow peaks in some years.
- Note 2. For microprocessors an additional 5-700 kSF central computer upgrade will be required in the early 90's.
- Note 3. See note below on electronic CAE.
- Note 4. For Digital El. CAE, spending in divisions certainly took place. This is difficult to quantify since it took place from budgets not earmarked for CAE.
- Note 5. For Digital El. CAE, the lower recommended values correspond to maintaining the status quo with a 4 year depreciation. The upper values correspond to "steady state" spending with a "fully" equipped lab. In each case the numbers correspond to equipment which should arguably be funded centrally (although this is not yet entirely the case).

The anomaly of having a service (Microprocessor Support) funded entirely with investment money should be noted. This situation should not be allowed to occur again in the future.

Note: The electronic CAE figures do not contain any allocation for micro-electronics development and only a preliminary estimate was made for this at the time of the report. A more recent working group headed by C. Fabjan has suggested an annual operating budget of 400 kSF for this activity (excluding direct chip development costs) and investment of 1 MSF in 1989-90 on top of an existing equipment base of 700 kSF. If 4 year depreciation is applied to all this equipment, the sustained investment rate will need to be at least 450 kSF per annum in the future.

3.6 Management Information Services

3.6.1 Overview

Management Information Services clearly fall into the category of the computing infrastructure of the laboratory, having a greater responsibility towards the correct administrative functioning of CERN itself than towards the provision of services for the physics community. Yet these two rôles cannot be entirely separated, since the physicists, engineers and those working closely with them need a computer service for their administrative and management work albeit in some cases of a specialized nature. In practice, certain areas of the MIS unit at CERN are also the only, or at least major, providers of certain important services to physicists and engineers, e.g. processing of large scientific documents, PC support and the PC shop.

Historically, MIS at CERN has been attributed a relatively low priority and central investment. This, together with a lack of a key overall strategy, has led to the existence of a wide spectrum of independent and heterogeneous solutions both at the level of the word processing equipment distributed around the site, and also at the level of the major applications running on the administrative data processing computer (ADP). The continual need to create interfaces which solve specific problems between subsets of inhomogenous solutions, and the frequent discovery of multiple uncoordinated solutions to the same CERN problem, are good examples of the inefficiency that has been created by this situation.

This being recognized, a senior management committee, SCAIP, was created in late '85 and decided to initiate an MIS Task Force in '86. This led to the creation in '87 of the MIS unit in DD Division. SCAIP continues to oversee MIS policy, and an MIS Board, consisting of representatives from all the divisions, provides technical coordination. The latter formed a natural basis for the working group producing input to this report. This group was joined by two participants from DESY.

The working group's report reflects the established programme of work of the MIS unit, and the resource levels required to carry it out. These have been discussed several times by the CERN management during the unit's short evolution, and this particular report has few surprises. The overall theme is one of continuity of the established programme and structures. This being said, it must be emphasized that not all of the problems have simple solutions and that the unit has not yet reached the staffing level required to carry out its programme of work. In addition, pressure from the CERN Member States for improved efficiency of the CERN administrative procedures and clearer presentation of the CERN accounts has increased the demands on the initial programme, and emphasized the need for an upgrade of the already saturated ADP computer.

3.6.2 The current situation

In broad terms the MIS activities at CERN divide into two areas:-

1. those concerning corporate data bases centred primarily around the ADP computer, and
2. those concerning office systems, used by MIS users in their daily work to access corporate data and for others tasks.

3.6.2.1 Administrative data processing

The ADP computer is small, an IBM 4361, installed 5 years ago and completely saturated. Response time for transactions and enquiries during the working day is unacceptable even though batch applications run outside prime shift. For historical reasons applications run under two different operating systems. There are three data base systems, none of which is Oracle, and three transaction systems. Despite serious attempts to homogenize the user interface, one is unable to hide the "context switch" that occurs when passing from one application to another. The financial data base system and

the purchasing/receiving system are different and use separate data bases. Interfacing links have been built by CERN, but cannot conceal the differences.

3.6.2.2 Office systems

The second area of office systems covers several areas. The scientists have long used the central machines for their office tasks such as mail, scientific text processing and document storage. Despite the evolution of workstation software in this area, many of these tasks will remain in part on the mainframe for some time. The two latter tasks are the responsibility of the current MIS unit.

The administrative office systems began as stand-alone word processors, independently installed by independent entities, and have formed clear barriers to document interchange and a laboratory-wide system. Scientific symbols and mathematical text processing have been needed not just by the scientists but also by the secretariats. Scientific document exchange with other laboratories and with the publishers of scientific journals is another requirement. Scientists visiting CERN want to find the text system they are used to at home. Whilst stand-alone word processors are no longer purchased, document markup standards begin to ease the situation, and Macs/PCs are provide acceptable powerful solutions, *there is still no unique solution that provides all the requirements*. The present mixed and inefficient situation has only just started to be resolved.

The investments in this area cannot be ignored. It is estimated that over 7 MSF was spent on word processors, which today still cost a recurrent 0.7 MSF in maintenance. The Mac/PC base is estimated at 20 MSF of capital investment expanding at 5 MSF per year.

Current and future office applications focus on four unavoidable operating environments into which all present effort is concentrated, which are IBM with VM/CMS, DEC VAX with VMS, PC-compatibles and MS-DOS, and Macintosh with MacOS. These are linked by the CERN Ethernet communications infrastructure. Obsolete systems will be phased out gradually.

Note that each of the above has its own particular rôle with the CERN computing picture, but that the overwhelming choice for a desktop system by those not constrained by the needs of other environments is the Macintosh. Unix systems may bring future possibilities for simplification, but it is possibly in this area above all that the manufacturers will see the need and the chance to put "added value" into their Unix products, especially in the area of the human interface.

3.6.3 The Working Group Report

The working group report analyzes the requirements in section two of its report and produces a list of required services to meet the objectives it has set, namely to:-

- *"provide a coherent set of services accessible from any office in CERN,*
- *reduce the number of different and incompatible office systems and administrative data bases and procedures,*
- *reduce paper consumption and paper shuffling,*
- *introduce modern techniques and keep up to date."*

They list 17 recommendations, separated into three main categories which may be found in detail in their report, and which lead here two summary recommendations:-

Summary Report Recommendation 22 : Upgrade ADP Computer

The ADP Computer should be upgraded, and its operation integrated with the CERN Computer Centre.

3.7 Computing for Theorists

3.7.1 Overview

The theoreticians at CERN have traditionally used computers somewhat less than the other research divisions, although this situation is now changing. A review of their present computing requirements reveals two main themes, namely:-

- improved general infrastructure, and
- usage of supercomputers or parallel systems.

3.7.2 General infrastructure.

In order to carry out their work in a satisfactory manner nowadays, the theoreticians make use of electronic mail, scientific text processing and other general tools. The CERN Theory Division is, however, somewhat humbly equipped even in the number of terminals available. *It is generally agreed that the equipping of all offices of the Theory Division with terminals, PC's, Macintoshes or workstations, together with ready access to printing facilities is of high priority.* This can be achieved to a large extent by the provision of modest financial resources.

Certain software packages have found particular acceptance within the theoretical world, notably the T_EX scientific text processing package. The request for improved support for these packages, given that CERN had chosen other standards, was a manpower issue, although this has subsequently been largely resolved.

3.7.3 Use of Supercomputers and Parallel Systems.

Recently developed techniques for solving basic non-perturbative problems in QCD by use of Monte Carlo simulation on the lattice require powerful computing facilities such as vector supercomputers or parallel processors. They require such large amounts of computer time that they could keep any number of powerful computers busy all of the time.

It is recalled that the computing resources in the Centre are there primarily to assist the experimental programme. Furthermore, in the general spirit of the one-third : two-thirds rule, this type of computation can just as well be done outside CERN.

On the other hand, one should distinguish between the development, debugging and testing phase of this research and the major simulation runs that consume the computer time. CERN theorists are involved in large international computer simulation collaborations. It is true that these collaborations have succeeded in finding, using their own initiative, the computer time they need at other installations. However, if CERN does not provide a certain level of facilities to match international standards, (at least for the development phase), its theorists will be unable to participate fully in this work and to attract others concerned.

There is therefore a question of policy to be resolved here, as to what fraction of CERN's central resources are to be devoted to this kind of theoretical work.

Summary Report Recommendation 24 : Theorists use of CERN Computers

It is recommended that such theoretical work should have access to CERN's central computing resources for development and debugging but that the fraction of resources allocated should remain modest, of the order of 10% of the total available.

3.7.4 Recommended Resource Levels of the Working Group

The financial resources required to equip the Theory Division with normal computing infrastructure are modest, of the order of 300 kSF. Nonetheless, this can probably not be handled within the normal operating budget of the division.

The requirement to find manpower at the level of one person in a user support rôle essentially for Theory Division is a request that comes on top of existing other demands on an over-stretched area. Note that this form of user support spans the boundary between MIS and the computer centre user support group.

3.8 General Remarks and Conclusions of the Review of the Working Group Reports.

A number of general points come out of the above review which are not specifically brought out in any section. Two particular examples are staffing and training.

3.8.1 Staffing issues

The laboratory is seriously under-staffed for the computing demands it is already asked to fulfil. The working groups make large *additional* demands for staff. (On face value, they ask for around 100 posts). Furthermore, in view of the overall compression of the CERN staff complement that is in progress, it will be particularly difficult to make the necessary increase in informatics staff numbers in the short term. The committee therefore feels that:-

Summary Report Recommendation 25 : Staffing Five Year Plan

It is essential to agree on a staged plan for an acceptable increase in informatics staffing levels across the whole laboratory over, say, five years.

A small, but valuable, contribution to alleviating the staffing problem is being made through co-operants, joint projects with industry, etc.. Some increased activity of this nature may be possible, although it is clear that the essential part of the problem cannot be solved by this means.

Summary Report Recommendation 26 : Alternative Staffing Methods

We recommend investigation of further alternative methods for informatics staffing, such as the appointment of technical associates from companies and universities, and increased involvement of graduate students.

Joint development projects and staff exchanges with industry and academia, especially in the fields of advanced electronics and computer science, should be actively encouraged.

3.8.2 Education/Training/Development

Despite the urgent pressures from the above situation, a massive effort is necessary in order to train and re-train existing CERN staff in informatics. If the missing computer-literate staff are to be found, a substantial part will have to come by internal movement and re-training. *This will require a training programme on a scale comparable with that found in industrial companies which have successfully tackled this problem before CERN.* Industry uses commercially available courses where appropriate, but many companies go further and establish permanent training centres. Three month courses are not unusual.

The success will vary with the area. Some areas can probably not be filled by re-training. Nonetheless, it is essential to establish a major programme to this end if CERN wishes to escape from this chronic staffing constraint. New attitudes and extra resources will be required to produce the desired results.

Summary Report Recommendation 27 : Training

CERN must set-up an organization devoted to a massive training and re-training of the existing staff in informatics technology.

In an associated area of training:-

Summary Report Recommendation 28 : Advanced Development Projects

Progress in the field of computing is so rapid that CERN must devote a fraction of its staff resources to work on **advanced development projects** that are only expected to show benefits in the medium to long-term. It is also vital that the long-term CERN staff involved with computing are kept well informed of the latest techniques.

3.8.3 The Next Steps

Summary Report Recommendation 29 : The Next Steps

1. The CERN management should define a policy framework and level of resources, money and staff, that it will make available for computing in the next five years.
2. The staff providing the computing services should then be allowed to optimize the services that they provide within the policy framework and the agreed level of resources. This should be done in consultation with the users, both those on the CERN staff and those from the outside institutes.

Appendix A

Existing CERN Computing Facilities

This Appendix is a rather brief resumé of the computing facilities in place at CERN at the beginning of 1989, intended to serve a base-line to the discussions contained in the main report. A much higher level of detail can be found in the DD Report CERN/DD/89/9.

A.1 The Computer Centre

The main batch and interactive services are provided by a six processor IBM 3090/600E with six VFs and a twin processor Siemens 7890 S (Fujitsu M382). These machines provide a total scalar processor capacity of some 52 CERN units,¹⁵ and share access to 25 tape drives, to 32 cartridge drives, and to 205 Gbytes of disk space. The main operating system is VM/XA SP2 with HEPVM¹⁶ additions. The MVS/WYLBUR operating system used for the previous ten years is being actively phased out, together with its key peripheral device, the IBM Mass Storage System (MSS). The IBM system has 5000 registered users of whom more than 2500 use the system in one week. At peak hours there are around 600 connected users of whom close to 300 are active each minute.

A Cray X-MP/48 (4 processor, 8 Mwords memory) with a solid state disk of 128 Mwords and 45 Gbytes of disk space runs the Unicos¹⁷ operating system. It provides around 32 CERN units of scalar processing power. Whilst this machine is clearly targeted at codes suitable for high levels of vector performance it is also running some scalar data processing codes from the experiments. The Cray is connected directly to the above IBM tape and cartridge drives. The Cray has around 200 registered users.

The exploitation of vector processing facilities on the Cray and IBM for the codes that concern experimental physics is the subject of much active work. The vector speed-up for such codes has been traditionally small, and new ideas are needed. The potential gains are nonetheless high.

Major interactive services are also provided under VMS on a VAXcluster which contains five powerful VAXes comprising 8800, 8700, 8650, and a number of smaller models. The key features of this popular service are compatibility with VAX systems in the experiments and the home institutes, and the suitability of the user-friendly operating system for development work. There are nearly 3000 registered users of whom 1500 use the system in one week. At peak times the two VAX machines dedicated to the general service have 200 connected users.

Besides this general service, the VAXcluster provides specialized services for LEP Project databases (based on Oracle), and mechanical CAE (based on Euclid). Another VAX, an 8530, runs an Ultrix (Unix) service for a number of specialist areas including microprocessor support.

¹⁵ the CERN unit of computing power is historically the power of an IBM 370/168 or a VAX 8600. As a rough indication, this may be taken as 3 IBM MIPS or 4 DEC VUPS.

¹⁶ HEPVM is the name of the collaboration between High Energy Physics sites which seeks to make such modifications and additions to the VM operating system that are important for HEP as part of a common effort, and which thus eases the task of moving physics code between sites. The most essential additions concern the Batch Monitor written originally at SLAC. The key member sites are CERN, IN2P3, RAL, SACLAY, and SLAC, and there are around 25 collaborating sites.

¹⁷ Cray's multiprocessor version of Unix

A.2 Networking

The INDEX terminal switching system connects around 2000 terminals to 200 hosts. No further investment in this facility is planned, but it remains a primary route for connecting terminals. Around 30 bridged Ethernets link some 1000 computers and other devices across the site. Terminal concentrators on the Ethernets provide for new connections and the eventual phase-out of INDEX.

The home-developed CERNET still has some 50 computers from several different manufacturers attached to it, but is being closed down. The on-site X.25 network is connected to the public packet switched networks, and to some 12 external leased lines, including one satellite and one under-ocean link to the USA. The Swiss EARN node is located at CERN, and is attached to a further 12 leased lines.

A.3 Computers at Experiments

Most large experiments have at least one powerful 32-bit computer, typically a large VAX, and a number of smaller machines, usually in a cluster, used for a variety of tasks including data acquisition during detector tests and calibration runs. There are at least 100 VAX and microVAX computers, and at least 20 Norsk Data computers (including six 32-bit Nord-500s) being used at experiments.

A.4 Computers at Accelerators

The PS control system consists of two Nord-570 computers and around 30 Nord-100s. The Linac and Lear are controlled by a VAX-based system, linked to the above. The SPS and LEP control systems are largely integrated, the main hardware consisting of two Nord 570s, around 80 Nord-100s, and about 100 machines of IBM-PC/AT architecture. The SPS network hardware is based on six interconnected stars, built under contract to CERN specifications by TITN about 10 years ago. The LEP network hardware is based on 16 IBM Token Rings. Workstations, notably Apollos, are being used increasingly for a variety of tasks.

A.5 MIS

CERN's ADP services run on an IBM 4361 Model 5 computer, running the DOS/VSE operating system under VM-SP. The machine is saturated and the interactive response is unacceptable. The one Gbyte of disk space used to hold the active copies of the main corporate databases, is insufficient, e.g. purchasing records older than 3 months are archived, at considerable inconvenience to users and ADP staff.

CERN has inherited a large number of diverse office computing systems. In terms of numbers of users the biggest system is the ND Notis, but there are also AES, Wang and Philips systems in use.

A.6 Macintoshes and PCs

The best estimates indicate that there are at least 2000 Macintoshes, and at least 800 IBM-compatible PCs at CERN today, and they are all spread widely throughout the organization. Many of these have been purchased by outside user groups.

A.7 Workstations

There are over 120 Apollo workstations, and at least 250 VAXstations installed currently at CERN. Of the Apollos, some 65 are on a main network, being used for software development and physics analysis by physicists and by some DD staff. A further 20 are used in LEP controls, and around seven are used for electronic CAD design. The VAXstations are also mainly used for software development and physics analysis.

Appendix B

The Cost of Computing at CERN over the past decade

This section describes the staffing levels and the money spent on computing across the laboratory over the past decade. The numbers used in this Appendix are taken from the CERN accounts. Where the numbers have been corrected to present day figures, the correction factors used have been the appropriate CERN official indices.

This being said, the CERN accounts do not contain an explicit record of the "costs of computing" as such. Indeed the term "costs of computing" needs definition in such a discussion. Three examples in common usage are, the costs of the computer centre, the full costs of the DD Division, or the costs coded in the CERN accounts under headings labelled "computers".

B.1 Total CERN Material Costs for Computing over the past decade

The following strict definition is used in order to obtain consistent numbers which may be compared across the years, and in order to handle computing expenses from all divisions, not just DD. It is assumed that a reasonable picture for material costs is given for operation money by the sum of the expenses coded under 251, 252 and 253 in the accounts (data handling: maintenance, general expenses and consumable items), and for computing capital purchases by the code 351 (data handling equipment: capital outlays). The precision of the numbers is thus dependent on the use of the correct code by the initiator of an expenditure. Computers or other data handling items purchased under other names and codes are clearly not recorded in these numbers. This definition does not cover money spent e.g. on buildings or power supplies for computers, on electrical services, on laboratory equipment, or any form of divisional infrastructure such as travel money. The numbers for DD Division are well understood and provide an indication of the strictness of the definition. In 1988, this method derives a DD total of 22.3 MSF, which can be compared with the total divisional expenditure of 25.9 MSF. This ratio varies over the years concerned between 84 and 92%.

Table 8: Material costs labelled Computers in the CERN Accounts, 1980-1988, (corrected to 1988 prices in MSF)

YEAR	1980	1981	1982	1983	1984	1985	1986	1987	1988
CERN running costs	19.7	20.0	17.1	15.9	15.1	15.8	16.2	12.0	14.9
CERN capital outlays	7.1	7.6	20.0	13.4	13.8	17.5	13.7	17.7	20.1
CERN Total	26.8	27.7	37.2	29.4	28.9	33.2	29.9	29.4	34.9
DD running costs	15.4	15.5	13.0	12.3	11.6	11.7	11.7	8.1 ¹	10.6
DD capital outlays	3.8	4.7	11.3	6.3	9.7	9.9	8.6	11.1	11.7
DD Total	19.2	20.2	24.3	18.5	21.3	21.6	20.3	19.2	22.3
Ratio DD/CERN	0.72	0.73	0.65	0.63	0.74	0.65	0.68	0.65	0.64

Note 1: 1987 was exceptional. The CDC service was stopped in order to help finance the acquisition of the Cray one year later.

From the numbers in Table 8 on page 47 one can see that the total CERN computing material costs thus defined, (adjusted to 1988 prices by using the CERN material indices), have stayed essentially constant over the past decade, as have the DD computing costs. This leads to a rather constant ratio, showing that DD represents on average 68% of the CERN total material computing costs.

B.2 Total CERN Personnel Costs for Computing over the past decade

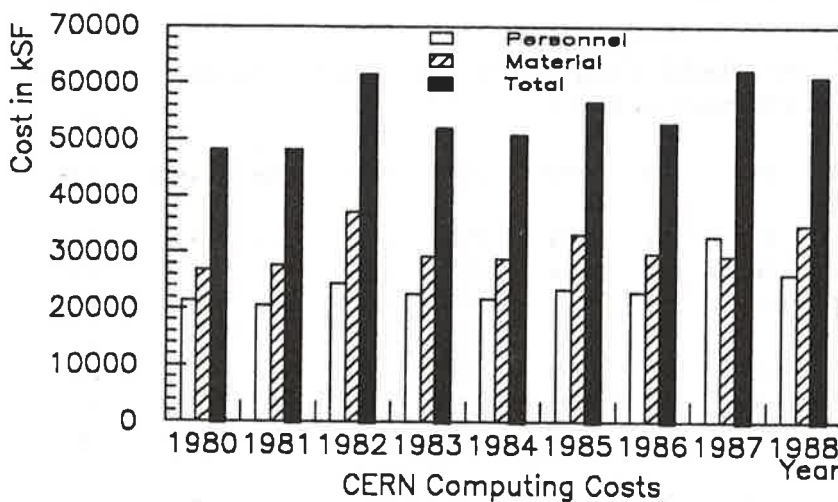
The personnel costs are well defined for the services provided by the Computer Centre, and for the DD Division as a whole. However, the true staff involvement in computing across the rest of the laboratory is not recorded, and cannot be reliably determined. For example, the distinction between physicist or engineer and computer support staff can be difficult. Similarly, it is not easy to understand what fraction of the accelerator control staff should be considered as computer specialists.

In order to estimate the total personnel costs in a manner that allows a comparison across the years, a major assumption is made that the ratio of personnel costs to material costs for the central computers is also applicable to the laboratory as a whole. This ratio varies between 0.61 and 1.09 and averages 0.75.¹⁸ The personnel costs produced in the above manner have then been corrected to real 1988 prices.

B.3 Total CERN Material and Personnel Costs

Figure 4 shows that the sum of personnel and material costs for computing at CERN thus defined have changed little over the past 10 years, and average around 55 MSF. This represents some seven percent of the total CERN budget.

Figure 4: A Comparison of CERN Computing Personnel and Material costs from 1980 until 1988. (See definitions in the text).



¹⁸ In 1987, a year of low DD material costs, this leads to an obviously incorrect personnel cost, demonstrating the limitation of the method.

B.4 An Alternative Approach for the Total Costs of Computing

The strict definition used above produces certainly a lower bound on the true costs of computing in the laboratory. By definition, the method used does not reflect the growing number of staff outside the DD Division who have been increasingly involved in computing activities. An alternative approach is offered here as an upper bound.

One assumes that all activities within DD Division relate to computing. The accounts show that the full personnel costs of the division have stayed, (with necessary corrections for changing composition and in 1988 figures), remarkably constant at 25.0 MSF/annum over 10 years. Making a rough estimate, one finds that the number of staff outside DD who are principally engaged in computing activities, (programmers in experiments, accelerator control etc.), is of the order of 200 people, and hence is similar to that working within DD. The total personnel budget involved is thus of the order of 50 MSF.

The total materials costs for all activities in DD in 1988 were 25.9 MSF. Assuming that this represents of the order of 70% of the total CERN computing costs, (see above), one may estimate the latter at around 35 MSF. This leads to a number of around 85 MSF for the current total annual CERN costs of computing, broadly defined.

B.5 The Base Line Pattern of Expenditure

In order to record the base line of expenditure for the review and recommendations in the report one may make the following breakdown. The money allocated to the DD Division for projects, essentially the whole capital investment of the DD Division, has averaged 10 to 11 MSF/annum over the past few years. This money can be divided into a pattern of expenditure the major components of which are:-

- mainframe computer purchase, 3.5 MSF/annum
- additional peripherals for above, 1.5 MSF/annum
- networking, 2 MSF/annum
- MIS/ADP, 1 MSF/annum
- computing support for engineers in DD, .2 MSF/annum
- DD Infrastructure, .5 MSF/annum
- direct support for experiments, 1.7 MSF/annum

In addition, the Cray computer is funded via the operational budget on a lease/rental contract at the level of 3.6 MSF/year.

Appendix C

An Estimate of Computing Capacity potentially available to CERN HEP in Europe

C.1 Objectives and restrictions

In the following we review the capacity potentially available for processing of data from CERN experiments in European centres outside CERN. The term **data processing** is understood here as requiring a minimal set of organized resources, including not only cpu power but also directly connected reliable tape/cartridge units, and operational facilities such that the processing power may be used a major fraction of the week. In the spirit of this definition, many physics departmental computers and workstations are believed to contribute more to the **physics analysis** than to the data processing of an experiment.

The numbers given below in CERN units¹⁹ are the result of a survey of physics users and staff of the computer centres known to us. The list is an integration of the centres declared to COCOTIME, those involved in HEPVM, and centres declared by the experiments to a MEDDLE survey as a source of capacity. There are reasons to exercise care when reading these numbers:-

1. Many centres do not have an established method of maintaining the actual use by HEP groups for analysis of data from experiments.
2. The experiments generally give their intended use. Actual use can be very different from this and often not easy to establish.
3. There are many VAXes in physics departments. In general one finds physics analysis rather than major organized production on these machines.
4. The growing potential of parallel farms makes such estimates difficult. These farms have been harder work to use than large centres and their usage has been proportional to the needs and persistence of those concerned. Since the effective usage promises to become easier with commercial engines, and full Monte Carlo grows in importance, one will see more work done on farms in the future. Nonetheless their use in Europe is considered to be a small part of the overall capacity at present.
5. It is often difficult to estimate or verify the capacity available to HEP groups at certain centres which are not under HEP control. Some subjective judgement is unavoidable.
6. Attempts to produce these numbers in the past, (e.g. COCOTIME, HEP-CCC), have not provided a complete picture. It requires more time than we have had for this report.

We have thus concentrated on major centres, set up for production, having tape/cartridge units, disks and operational facilities to run overnight. All with close connections to HEP have replied. Some with distant relations have not yet given answers. Beyond the major centres there are smaller physics department machines and physics' share of the general university machines, whose total capacity is estimated from those that are well known. (Typically one or two CERN units each).

The capacity available generally at Universities and research institutes is growing, including that which is available and useful to HEP. Thus the table below will soon be out of date and should be treated as giving lower limits.

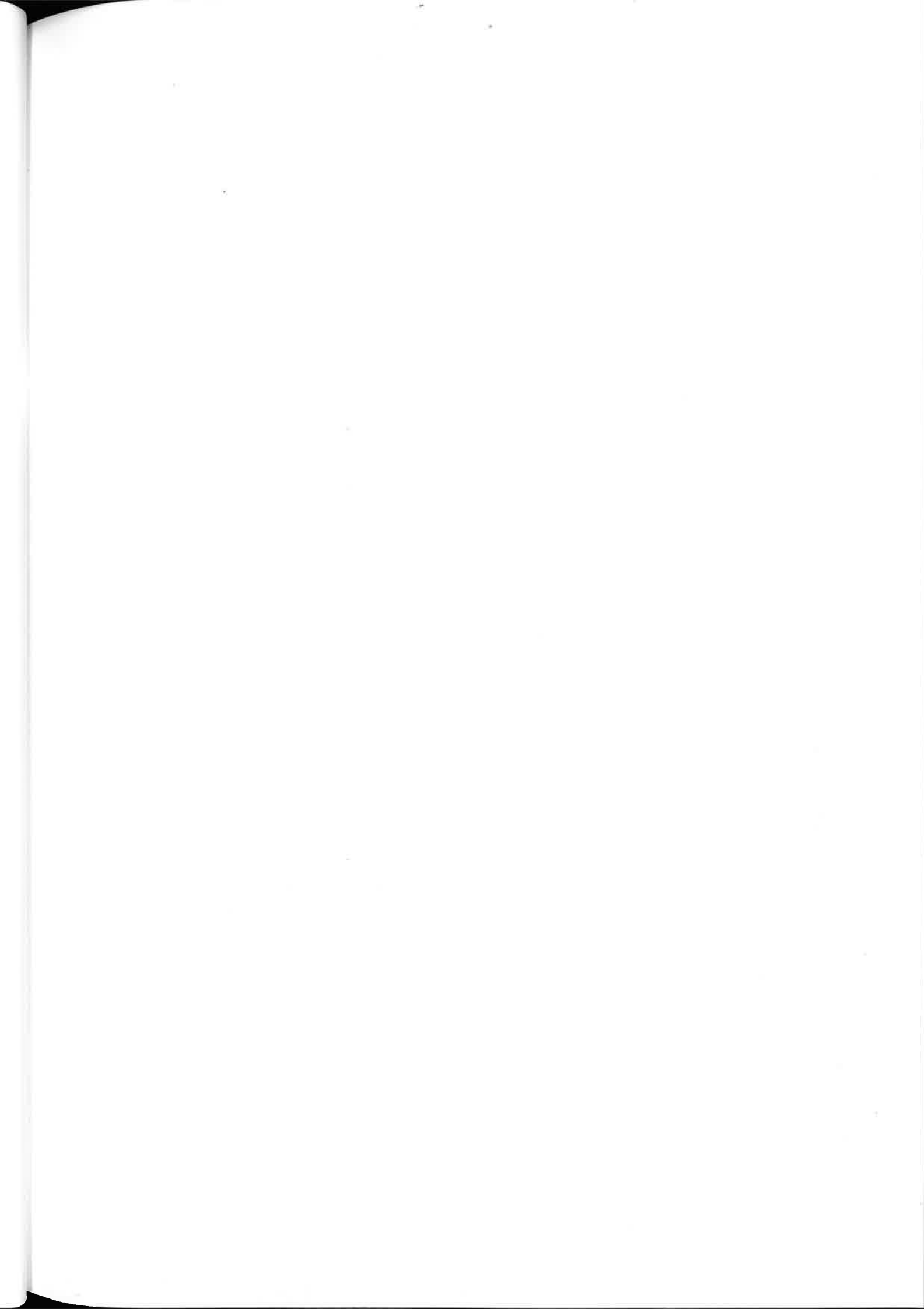
¹⁹ The CERN unit of computing power is historically the power of an IBM 370/168 or a VAX 8600. As a rough indication, this may be taken as 3 IBM MIPS or 4 DEC VUPS.

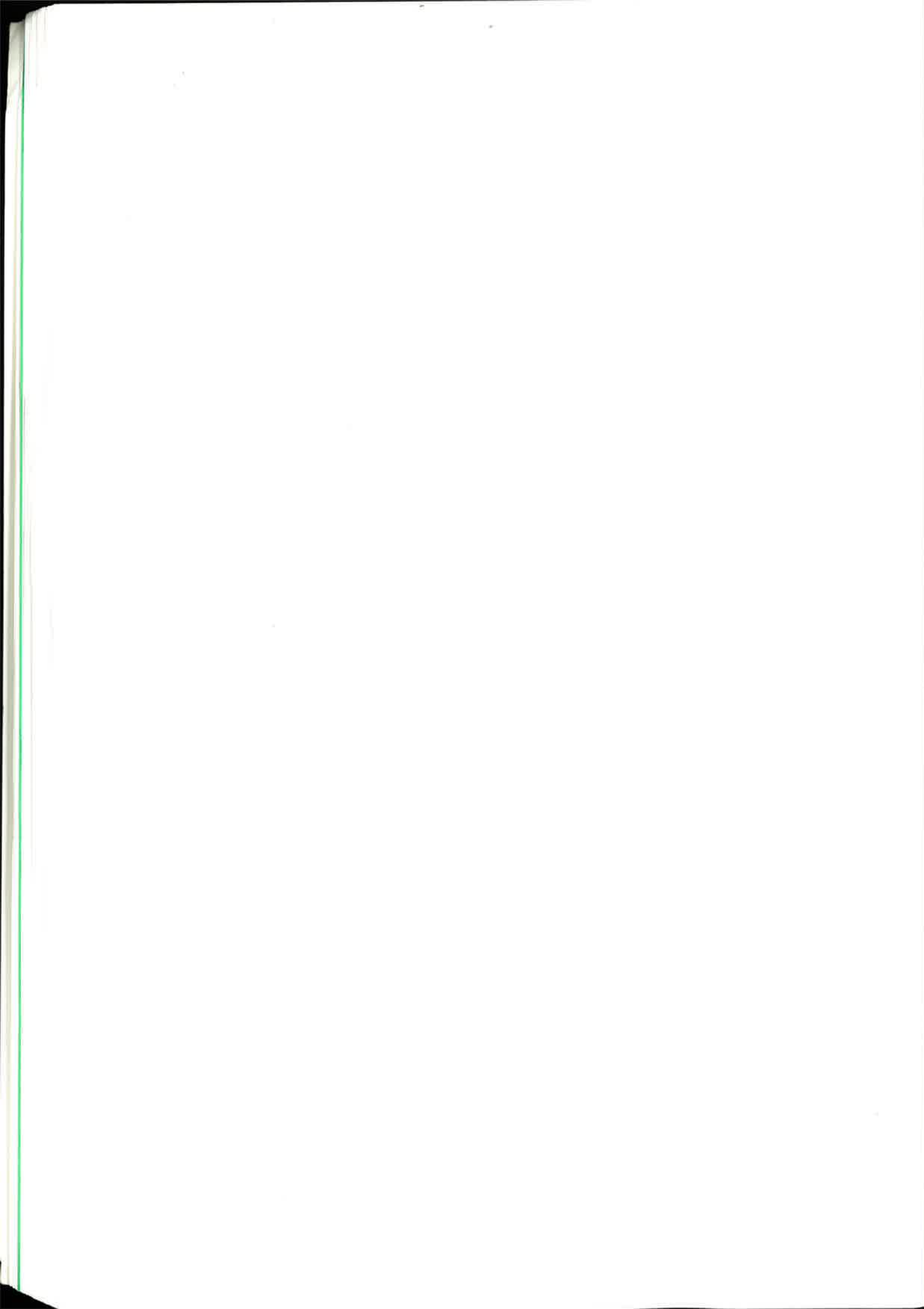
C.2 Reported capacity in CERN units.

Table 9: Reported capacity for HEP experimental work in CERN units.

The following lists the capacity that could be made available in the European data centres for processing of HEP data originating at CERN. It excludes work for theoreticians, data from DESY, CERN data processed outside Europe, nuclear physics data, and "private facilities" at CERN.

Centre	Town/ Country	Machine June '89	HEP share June '89	Machine End '89	Potential HEP End '89
CINECA	Bologna/I	IBM 3090-600E	13	Same with XA	24
CINECA	Bologna/I	Cray X/MP	≈ 0	Same	≈ 0
CILEA	Milano/I	IBM 3090-200EVF	9	200S	11
CNUCE	Pisa/I	IBMs	≈ 0	Same	≈ 0
INFN + SNS	Pisa/I	IBM 4381's	1.5	Same	1.5
Rome U.	Rome/I	IBM 3090-300SVF	≈ 0	600SVF	24
RAL	Oxford/GB	IBM 3090-600EVF	6.5	Same	max. 22
GB Univs.	Univs./GB	Many	10	Same	10
IN2P3	Lyon/F	IBM 3090-600S	25	Same	25
DPHPE	Saclay/F	IBM 3090-400E	13	Same	13 +
CIEMAT	Madrid/E	IBM 3090-150E	2	Same	2
E Univs.	Univs./E	IBMs	2	Same	2
IPP/MPI	Garching/D	IBMs	2	Same	2
MPI	Munich/D	IBM 4381	1	Same	1
R-WTH	Aachen/D	IBM 3090-600SVF	8	Same	8
IEK	Karlsruhe/D	IBM 3090-300S	2.5	Same	2.5 +
D Univs.	Univs./D	IBMs, emul.	10	Same	10
NIKHEFH	Amsterdam/H	Gould	6	+ DN1000	6 + ?
Nijmegen	Nijmegen/H	IBM 3090 ?	2?	+ DN1000	2 + ?
Geneva U.	Geneva/CH	IBM 3090 180VF	2	Same	2
ETH	ETHZ/CH	IBM 3090-180	4	Same	4 +
ETH	ETHZ/CH	Cray	≈ 0	Same	≈ 0
U. Bern	Bern/CH	IBM 3090 ?	1	Same	1
Vienna	Vienna/A	NAS + IBM	1	Same	1
Belgium	Belgium	Cray/IBM/VAX	1	Same	2
Sweden	Sweden	IBM/Alliant	2	Same	2
Portugal	Portugal	Several	2	Same	2
Norway	Norway	NDs + VAXs	≈ 0	+ Cray	max 1
N.Bohr I.	Denmark	VAX, others	≈ 0	Same	≈ 0
Greece	Greece	Several	≈ 0	Same	≈ 0
Total Share CERN Units			June 1989 ≈ 130		Potential, end '89 ≈ 180





Computing at CERN in the 1990s
Computing for Experiments
Final Report

January 26th. 1989

THE UNIVERSITY OF CHICAGO
LIBRARY

Contents

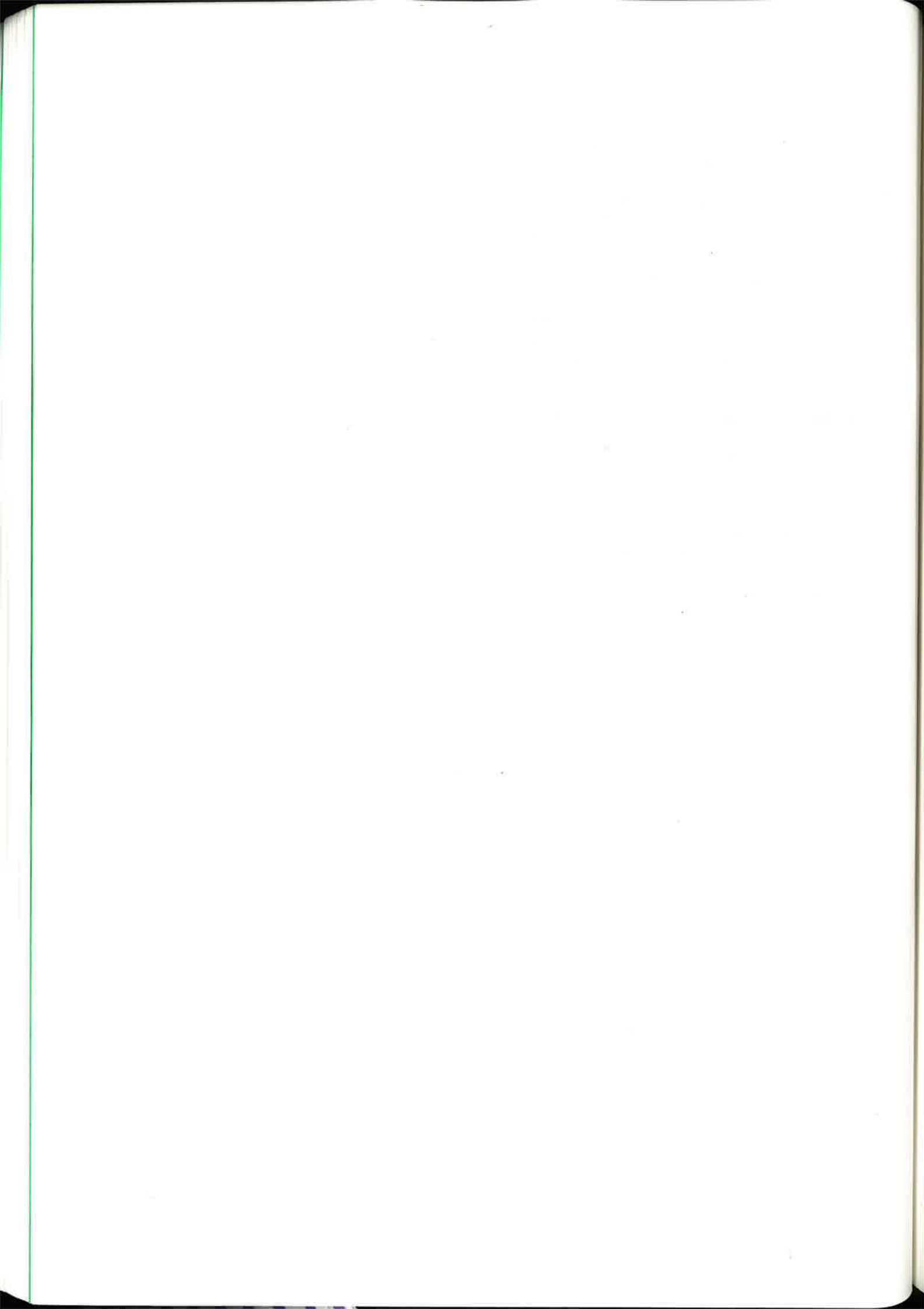
Part 1:	EXECUTIVE SUMMARY	53
Part 2:	SUMMARY OF RECOMMENDATIONS	55
1	RECOMMENDATIONS IN DATA ACQUISITION AND REAL-TIME COMPUTING FOR EXPERIMENTS	57
2	RECOMMENDATIONS IN GENERAL COMPUTING AND SOFTWARE SUPPORT FOR EXPERIMENTS	59
2.1	OVERALL MANAGEMENT OF COMPUTING FOR EUROPEAN HEP	59
2.2	WORKSTATIONS AND DISTRIBUTED COMPUTING	60
2.3	MASS STORAGE	60
2.4	CPU POWER	61
2.5	LOCAL AND WIDE AREA NETWORKS	62
2.6	DATA DISTRIBUTION TO OUTSIDE USERS	63
2.7	OPERATING SYSTEMS AND LANGUAGES	63
2.8	SOFTWARE SUPPORT	64
2.8.1	General Software Support Problems	64
2.8.2	Data Bases for Experiments	64
2.8.3	Software Engineering	65
2.8.4	Graphics Support Software	65
2.8.5	Education of Physicists	66
3	COSTS AND PRIORITIES	67
3.1	Costs	67
3.2	Priorities	69
3.2.1	Recommendations Concerning Mainframe-related Services	70
3.2.2	Recommendations Concerning Experiment-related Services	71
Part 3:	DATA ACQUISITION AND REAL-TIME COMPUTING	73
4	INTRODUCTION	75
4.1	Working Group Members	75
4.2	Overview	75
5	A REFERENCE MODEL OF EXPERIMENTAL ACTIVITY	77
5.1	Extrapolation of the Current Physics Programme	77
5.2	Preparation for a Future Hadron Collider	78
6	IMPLICATIONS FOR DATA-ACQUISITION SYSTEMS	81
6.1	Background	81
6.2	Synopsis of Areas Requiring Further Study	83

Part 4:	CENTRAL AND DISTRIBUTED COMPUTING, SOFTWARE SUPPORT	93
7	INTRODUCTION	95
7.1	Working Group Members	95
7.2	The Mandate of the Working Groups	96
8	THE COMPUTING NEEDS OF EXPERIMENTS	97
8.1	LEP Experiments	97
8.1.1	The MUSCLE Report	97
8.1.1.1	Processor power per LEP experiment.	97
8.1.1.2	Data storage per LEP experiment	98
8.1.1.3	Data manipulation	98
8.1.2	Beyond the MUSCLE Report	99
8.2	SPS Collider and LHC Experiments	100
8.3	Other CERN Experiments	102
8.3.1	LEAR Experiments	102
8.3.2	SPS Fixed-Target Experiments	102
8.3.3	Summary of 'Other Experiments'	103
8.4	HERA Experiments	103
9	THE COMPUTING ENVIRONMENT FOR THE CERN PHYSICS PROGRAMME	105
9.1	Location of Physicists and Computing Facilities	105
9.1.1	Off-site and On-site Computing	105
9.1.2	The Role of National/Regional Centres	106
9.1.3	Private Facilities at CERN	107
9.1.4	Data Flow in HEP Experiments of the 1990s	107
9.2	Elements of Computing for CERN Experiments	108
9.2.1	Total CPU	108
9.2.2	Type of CPU	109
9.2.2.1	Mainframes used as scalar processors	109
9.2.2.2	Mainframes used as vector processors	110
9.2.2.3	Parallel computing	110
9.2.3	The Move towards Personal Workstations	111
9.2.4	Data Handling	113
9.2.5	Networking	113
10	EVOLUTION OF CENTRALLY ORGANIZED COMPUTING AND NETWORKING AT CERN AND AT REGIONAL CENTRES	115
10.1	Introduction	115
10.2	Evolution of Batch Capacity	116
10.2.1	CERN Computer Centre Batch Facilities	116
10.2.1.1	Competitive and complementary services	116
10.2.1.2	Operating systems	116
10.2.1.3	Cost comparison of Cray and IBM-style computing	117
10.2.1.4	Manpower implications of the Cray service	118
10.2.1.5	The multi-purpose CERN VM service	118
10.2.1.6	IBM-compatible systems	118
10.2.1.7	DEC batch capacity in the CERN computer centre	119
10.2.1.8	Model for evolution of the central batch services	119
10.2.2	Large Regional Centres	121
10.2.3	Other Batch Capacity	121
10.2.3.1	Emulators and parallel processors	121
10.2.3.2	Private batch capacity	122
10.2.3.3	Batch capacity in university physics departments	122
10.2.3.4	Workstations as batch processors	122

10.3	Evolution of General-purpose Interactive Facilities	122
10.3.1	Workstations at CERN	123
10.3.2	The Future of Central Time-sharing Services	123
10.4	Interactive Data Analysis	124
10.5	Networking Implications	126
10.5.1	On-Site Networking Infrastructure	126
10.5.1.1	Short term evolution	126
10.5.1.2	Medium term	127
10.5.2	Wide-Area Networking	127
10.5.2.1	Short term evolution	127
10.5.2.2	Medium term	128
10.6	Data Storage	128
10.6.1	Mass Storage Evolution in the CERN and Regional Computer Centres	128
10.6.2	Distributing the Data	129
10.6.2.1	On-site distribution	129
10.6.2.2	Distribution of DST data to external institutes	129
10.6.3	Storage Management	130
11	EVOLUTION OF CENTRALLY ORGANISED SOFTWARE SUPPORT AND DEVELOPMENT	131
11.1	The Changing Landscape	131
11.1.1	Operating Systems	131
11.1.2	Languages	132
11.2	Software Support for Experiments and Packages/Libraries	132
11.2.1	Support for Experiments	132
11.2.2	Packages and Libraries	133
12	SOFTWARE SUPPORT: DATA BASE SYSTEMS	135
12.1	Data Base Applications in Present Experiments	136
12.1.1	Addresses of Collaboration Members	136
12.1.2	Electronic Mail Addresses	136
12.1.3	Experiment Bookkeeping	137
12.1.4	On-line Data Bases	137
12.1.5	Description of the Detector Geometry	137
12.1.6	Calibration Constants	138
12.1.7	Event Data (or Memory Management Systems)	138
12.1.8	Bookkeeping of Program Versions	139
12.1.9	Histograms and other Physics Results	139
12.1.10	Software Documentation	139
12.1.11	List of Publications	139
12.1.12	Other DB Applications	139
12.1.13	Present DB Applications in Experiments: Conclusions	140
12.2	Broad Areas for DB Applications and their Requirements	140
12.2.1	DB Applications where Generalized Packages are Possible	141
12.2.1.1	Experimental Administrative Database	141
12.2.1.2	Electronics and material Bookkeeping	141
12.2.1.3	Tape and Production Data Base	142
12.2.2	DB Applications for which Tools could be Specified and Written	142
12.2.2.1	Detector Description	142
12.2.2.2	Calibration Constants	143
12.3	Data Base Systems: Recommendations	143
12.3.1	Education and Training	143
12.3.2	Design Support Team	144
12.3.3	Data Model Software	144

12.3.4	Portability of the Data Base Information	144
12.3.5	Area of Experimental Administrative Data Bases	145
12.3.6	Area of Bookkeeping Data Bases	145
12.3.7	Area of Documentation Data Bases	145
12.3.8	Area of Detector Description and Calibration Constants Data Bases	145
12.3.9	Area of Interactive Data Analysis Data Bases	146
12.4	Data Base Systems: References for detailed reading	146
13	SOFTWARE SUPPORT: SOFTWARE ENGINEERING	147
13.1	Analysis and Design, Methods and Tools	147
13.1.1	Historical Background	148
13.1.1.1	Situation in the HEP community	148
13.2	Large Software Projects at CERN	150
13.2.1	Software Projects that have not Used Formal SE Methods	150
13.2.1.1	Organizational Tools	150
13.2.1.2	The Problems	150
13.2.2	Conclusions from Experiences with SE Methods	151
13.2.3	Requirements of Outside Institutes	153
13.3	Other Coding Support Tools	154
13.3.1	Source Code Management	154
13.3.2	Other Implementation Tools	155
13.4	Software Engineering: Recommendations	155
13.4.1	To Project Management	155
13.4.2	Training	156
13.4.3	Support and Consultancy	156
13.4.4	Software Tools	157
13.4.5	Licences	157
13.4.6	Requirements of the Outside Laboratories	157
14	SOFTWARE SUPPORT: GRAPHICS	159
14.1	Graphics in HEP	159
14.1.1	Graphics Applications	159
14.1.2	Implications for Software and Hardware	159
14.2	Graphics Software and Interface Standards	160
14.2.1	Basic Picture Generation	160
14.2.2	Basic Library Functions	161
14.2.3	Distributed Graphics	161
14.2.4	Document Processing	162
14.3	Graphics Hardware	162
14.4	Graphics: General Conclusions	163
15	SOFTWARE SUPPORT: USER TRAINING	165
15.1	Introduction	165
15.2	The Physicist's Demands	165
15.2.1	Documentation and Examples.	165
15.2.2	Formal Training Courses	166
15.3	New Working Practices for Physicists	167
15.3.1	Personal Workstations	167
15.3.2	PAW - The Physics Analysis Workstation.	168
15.3.3	Software Topics.	168
15.4	A Central Workstation Training Facility	169
15.5	Videa and Computer-based Training	169
16	TECHNOLOGY PREDICTIONS FOR THE 1990s	171
16.1	Introduction	171
16.2	CPU Technology	171
16.2.1	High-End Processors	171

16.2.2	Low-End Processors	172
16.3	Data STORAGE Technology	172
16.3.1	Disks	173
16.3.2	Solid State Devices	173
16.3.3	Tape/Cartridge	173
16.3.4	Optical Storage	174
16.3.5	Digital Audio Tape Technology	174
16.3.6	Exotic Technology	174
16.3.7	Summary of Data Storage	175
16.4	Data Transport	175
16.4.1	Local Area Network Capabilities	175
16.4.2	Wide Area Network Capabilities	176
16.5	Final Comments on Technology	177
Appendix A:	Current Mass Storage Usage in the CERN Computer Centre	179
A.1	VM/CMS	179
A.1.1	Disk Space	179
A.1.2	Users and Accounts	179
A.2	VAX/VMS	180
A.2.1	Disk Space	180
A.2.2	Users and Accounts	180
A.3	Cray	180
A.4	Summary	181



Part 1

EXECUTIVE SUMMARY

In the context of the report 'Computing at CERN in the 1990-s' various working groups were mandated to produce recommendations in the area of 'Computing for Experiments'. They have condensed their findings in two separate report parts, one dealing with questions of data acquisition problems (real time) the other containing all other aspects of data processing. The recommendations are based on the assumption that the CERN *physics programme* will be dominated by LEP experiments with 'Z⁰-like' data rates until 1995, that the SPS collider programme runs well into the early 1990-s, that other experiments like the LEAR or the fixed-target programme will continue, and that serious preparation for a new hadronic collider programme (e.g. LHC) will also be part of the future scenario.

The recommendations of the working groups can be grouped into the following three broad categories.

Evolution of Existing Services

The expected evolution of the *needs* of experiments and of the *market offerings* (i.e. new technologies) should result in a substantial boost of installed *CPU capacity* and *mass storage*, both at CERN and in member states. A major impact of modern *workstations* is predicted, resulting in serious needs for *interconnection* and *software support* for distributed systems.

The projections are high in capital expenditure and manpower. The working groups want to stress their opinion that *present levels* of funding and staffing are at a *marginal or inadequate* level. We believe that computing expenditures provide at least equally significant potential in *added physics returns* when compared to spending for other sectors such as accelerators or detector construction.

Buildup of New Services, Research & Development

Experiments have dramatically changed in scale over the last decade, and data handling methods used outside our community have developed significantly, too. The basic attitude to the use of computers in experiments has not evolved in a similar way. More *professionalism* in software design, and the move towards commercial software will have to translate into *new and more centralized support* for the HEP community. This concerns coordinated *product evaluation*, the conduct of *pilot projects*, HEP-wide *licensing* strategies in many areas, some HEP-specific *software developments* in areas not covered by the market, and a marked effort of user *education* and product *documentation*.

The reports also suggest several areas in which *research and development* work is thought necessary (the understanding of 'research' here is in the direction of technologies and methods existing in industry or academia, but largely unknown to the HEP community). Most notably this concerns 'parallel computing' and 'real-time data handling for high event rates'.

We believe strongly that the CERN management should provide clear guidelines for this critical technical area. There does today exist unwanted duplication and sometimes frustration in the collaborations, who have to tackle general problems with inadequate resources. While initiatives must come from the physicists, general development and research activities cannot be pursued uncoordinated without introducing inefficiencies. In some areas, lack of a general strategy may jeopardize our physics programme.

Relations Between CERN and the User Community

Most recommendations made in the working group reports concern the needs of the *community* of physicists participating in the CERN physics programme. In some areas such as mainframe computing aspects, present policies provide guidelines for the relative sharing between CERN-based and outside, where 'outside' includes regional centres and individual university institutes.

Such guidelines do not exist for many other recommendations, particularly in the area of application-oriented support such as data base systems, data modelling, software engineering tools, hardware description languages, or graphics.

The CERN management sees its role mostly under the angle of the host laboratory, and has not shown major initiatives to address community problems in the area of computing. The HEP community, on the other hand, has as only representative body in computing matters the 'High-energy Physics Computing Coordination Committee' (HEPCCC).

The working groups strongly suggest that a *mechanism* is created that permits the community to *take a decisive and responsible part* in shaping the future landscape of computing for HEP in Europe. This could be derived from a body such as HEPCCC or possibly ECFA. Until such a global solution is found, CERN should *take the lead* where necessary.

Part 2

SUMMARY OF RECOMMENDATIONS

Chapter 1

RECOMMENDATIONS IN DATA ACQUISITION AND REAL-TIME COMPUTING FOR EXPERIMENTS

The working group is of the opinion that CERN is entering a period of steady and predictable evolution as far as the real-time computing needs of the current generations of experiments are concerned. The major components of these experiments have been created, but will need running in. Continued support will be required to make the real-time systems of these experiments respond to the evolving demands of the physics programme.

The major challenge appearing on the longer-term horizon is that of constructing trigger and data acquisition systems for very high rate hadron colliders such as the LHC. If CERN chooses to pursue this ambitious program, a great deal of research and development work is required. This work is too complex to be left until the actual time of formation of experimental collaborations, and should be starting soon. We should clarify that R&D in this area relates to learning *techniques and methods in use outside HEP*, and will have to make use of all commercially available products and services.

The extreme environment of future hadron colliders requires major advances in almost all areas of detector construction as well as triggering and data acquisition. This forces a tight coupling of the issues of detector design, detector digitization, data compression and triggering. It seems imperative that CERN consider means of organizing and supporting the required long-term R&D in the widest possible context, integrating most or all existing expertise in the HEP community.

Some of the general features of an LHC experiment, and the implications for data acquisition and triggering, are listed here:

- Very high data rates with the associated massive high speed buffering and sophisticated data compression on very short time scales, will require integrated design of detectors and the trigger and data acquisition system to achieve the desired density and speed.
- Unprecedented complexity produced by a system containing many thousands of complex processing elements, both embedded in the system and isolated in processor arrays, will require the use of vastly improved software tools to create the necessary programs.
- Event rate reductions of almost 10^8 , which are several orders of magnitude greater than those achieved in the current generation of hadron collider experiments, will require the development of very sophisticated filtering algorithms.

There is a clear need for a better understanding of this environment, and of the methods to tackle the associated problems. Topical workshops as they have been or are being organized provide a useful starting point, but the next steps involve working with sophisticated detector and physics models, and

should then lead into a series of coordinated R&D and pilot projects.

Underlying all of this will be the need to design, construct and maintain a system of the required scale. The solution will require familiarity with methods applied in other areas of complex system design, as in industrial digital systems. The introduction of such new design methodologies will take a considerable amount of time and will not be inexpensive; the necessary high-performance workstations and the corresponding software packages are both costly. There will certainly be a long and painful learning period.

Serious investments will be needed in the training of engineers, programmers, and physicists, in order to take advantage of the computer aided design methods which must be used to construct such a system.

These methods are very powerful, but there will not be any instant results. The creation of pilot projects has proven to be an effective means for structuring such R&D. By pilot projects we understand projects which have some use beyond being strictly pedagogical, but which do not require the immediate mastery of advanced techniques, thereby leaving room to explore the new subject area.

Beyond the design methodologies mentioned above, pilot projects would also seem particularly important to acquire familiarity with VLSI technology, which we believe will be used extensively for both analogue and digital electronics of future experiments.

We recommend to set up coordinated pilot projects for introducing new high-level design methodologies in both software and hardware, and for acquiring familiarity with the application of VLSI techniques, both analogue and digital, to future experiments' real-time problems.

To achieve this goal, a better understanding is needed of the commercial market, of other applications, and of our own medium-term evolution. This would comprise familiarity with systems like Unix (trademark of AT&T) and associated real-time kernels, and future high-bandwidth busses, large systems with distributed functions, data base support, user interfaces, fault tolerance, etc.

We believe that improved mechanisms are required to evaluate existing commercial products and the real cost of in-house developments. We also think it extremely important and even urgent to find ways of taking community-wide decisions about the acceptance and general support of such 'standards', be they commercial products or complementary developments inside our community.

Chapter 2

RECOMMENDATIONS IN GENERAL COMPUTING AND SOFTWARE SUPPORT FOR EXPERIMENTS

An estimation of future computing needs for the CERN community requires predictions of the future of the CERN physics programme. It has been assumed that LEP runs mainly on the Z^0 until the end of 1991, after which running at higher energies is interspersed with additional running on the Z^0 for several more years. The SPS Collider and fixed-target programmes are expected to remain active well into the 1990s, and LEAR experiments have firm plans to take data until at least 1992. Finally and more generally, it is assumed that European physicists will pursue a vigorous, but as yet unknown, programme of physics at CERN beyond the currently planned limits of existing experiments.

By 1991-2 the physics programme of CERN will require within Europe at least 500-600 units of CPU power, the management of data amounting to hundreds of thousands of magnetic tapes, and well over 1000 Gigabytes of direct access storage. The number of workstations, each with about one unit of CPU, will approach the number of physicists in HEP. To ensure full university participation in this physics programme, users at home institutes will need access to both CERN and their regional centres at the speeds possible using modern fibre-optic technology.

The European HEP community has major experimental programmes at DESY and at the Gran Sasso Laboratory, in addition to those at CERN. The quantitative conclusions of this report cover only the requirements of the CERN-based programme. The qualitative conclusions, on networking and distributed computing, are almost certainly applicable to the other programmes.

2.1 OVERALL MANAGEMENT OF COMPUTING FOR EUROPEAN HEP

This report makes a number of recommendations which should be addressed to the management of European HEP rather than to CERN alone, or to laboratories, computing centres and universities independently. No existing body combines the managerial and detailed technical competence needed to implement the recommendations, although HEPCCC and/or Restricted ECFA may be in a position to initiate the creation of the necessary structures.

European HEP requires a body which can provide continuing managerial and technical co-ordination of all aspects of the computing environment.

2.2 WORKSTATIONS AND DISTRIBUTED COMPUTING

Workstations are single-user computers with high resolution graphics screens, and capable of running and debugging the largest programs used in HEP.

In the period up to 1995, the number of workstations used by European HEP will become comparable with the number of active physicists.

Full exploitation of workstations will require that each workstation can access data located at CERN or at regional centres at 40 to 400 kilobytes/second per CERN unit of computing power in the workstation. This requirement might well require a carefully organized hierarchy of servers.

A coordinated effort is required on the part of CERN and regional centres to propose, implement and continuously develop, a strategy for distributed data management for European HEP. This work should be supervised by a technical committee.

Management of workstation system software and communications will become a task comparable to the management of the mainframe computing service at CERN and at regional computer centres. As far as possible this management should be provided by the computer centres for the benefit of CERN-site and university users. Workstations applied to real-time tasks should be included in these considerations, wherever possible.

A service should be set up at CERN mandated to coordinate the management of systems software and communications for a defined but evolving range of workstations on the CERN site. Serious consideration should be given to offering a complete support service, rather than the coordination of manpower supplied by collaborations and other users.

Similar services supporting local and university users will be appropriate for many regional centres. A committee should coordinate the activities at CERN and regional centres.

Differing data representations reduce the transparency of any distributed computing environment. Conformance to the increasingly widely accepted IEEE data representation should be an important factor in the choice of hardware.

2.3 MASS STORAGE

Mass storage capabilities will be at least as important as CPU power in supporting flexible analysis of LEP physics, and will limit the scale of off-line analysis for hadron colliders.

By end 1991, CERN should have a total of at least 1000 Gigabytes of disk space and an automatic handling system for at least 40,000 tape cartridges.

By end 1991, any regional centre planning to contribute effectively to the analysis of major experiments should have at least 100 Gigabytes of disk space available to experiments together with automatic handling for at least 5,000 tape cartridges.

The cost of these mass storage facilities will be tens of millions of Swiss Francs, and even after this expenditure, resources will require very careful management.

An automatically managed storage hierarchy, from disk-files to data-tapes, is vital for CERN and will be important for regional centres. The financial value, in terms of more effective disk space use, is many millions of Swiss Francs. If such storage management remains unavailable commercially, it must be written by, or under contract from, CERN.

IBM 3480 tape cartridges will be the main mass storage medium for CERN physics in the next five years. However, even with the improvements in cartridge capacity now expected, there will still be hundreds of thousands of active tapes at CERN.

New mass storage devices with much higher capacity should be studied as alternatives in the longer term.

2.4 CPU POWER

Conventional mainframe computers will continue to be vital components of the HEP computing environment. However, in the period up to 1995, there will be a change in the way they are used. By 1995, most physicists will use workstations rather than terminals connected to mainframes. Physics analysis programs on the workstations will interact frequently with the mainframes, exploiting their unique combination of massive real-time data handling capacity and substantial CPU power. Workstations will certainly allow more productive and reliable physics analysis, but are unlikely to reduce the steady growth in the demand for mainframe services.

The CERN physics programme will require at least 500-600 units of CPU power by 1991-2. The physics would benefit considerably if much more power were available. Approximately one quarter of the 500-600 units should be in the form of conventional mainframe systems on the CERN site. One of the main functions of these systems will be to act as data servers and communications servers for off-site and on-site users. At the end of 1988 the mainframe scalar capacity at CERN will be about 80 units (CRAY, DEC and IBM). It is hoped that the vector capabilities of the CRAY and IBM machines will make additional power available for some Monte Carlo simulations, but vectorization is unlikely to have a major impact on other areas of physics analysis.

By 1991-2 CERN's mainframe scalar capacity should be increased to at least 150 units.

Major regional centres should also be upgraded so that their aggregate power available to CERN-based experiments exceeds that available on the CERN site.

It is assumed that parallel computing systems will be able to provide the balance of the 500-600 units at a total cost below that of conventional mainframes. If parallel computing proves to be much cheaper than using conventional mainframes, it would be beneficial to provide a total capacity substantially above 500-600 units. Conversely, if parallel computing is not successful, the capacity will have to be provided by other, conventional means.

A 'parallel computing project' should be set up as a coordinated effort involving CERN, major regional centres, interested university groups and industry. The project goals should be to do research in parallel computing relevant to HEP needs, and to offer, almost from the start, a parallel computing service to users. A system of commercially available processors adequately integrated with central data handling services should be given serious but not exclusive consideration. Physical implementations should be planned at CERN, and at some universities and regional centres.

Parallel computing systems are also important for on-line event selection. The off-line parallel computing project should be closely linked to the development of comparable on-line systems.

2.5 LOCAL AND WIDE AREA NETWORKS

The key to effective university involvement in today's collider experiments with high data volumes is fast wide area networking. Traffic between universities and 'data servers' at CERN or in regional centres is best carried by a dedicated network. The network should be distinct from a general purpose 'European Research Network', although its traffic might be multiplexed over the same low-level infrastructure.

A glass fibre has a capacity today of about 34 to 140 Megabits/second using widely available drivers. Off-the-shelf technology is expected to reach 565 Megabits/second and beyond in the early 1990s.

The short term target of 2 Megabits/second between CERN and each regional centre by 1991 should be seen as a step towards installation of at least one glass-fibre-equivalent as soon as possible. University involvement is fundamental to CERN's existence; a resolution of the funding and tariff problems associated with high speed international networking should be vigorously pursued at the highest level of management.

Countries where high speed lines are not yet available should be encouraged to speed the necessary installations to enable their full participation in the CERN physics programme. Progress in local area network technology and management is essential for the effective use of workstations.

Full exploitation of wide area links above 2 Megabits/second is beyond the capabilities of current commercial technology.

CERN and regional centres should continue, and where necessary expand, their efforts to work together with industry at the forefront of network technology.

Standards should be supported and adopted wherever possible, while recognizing that a blanket 'wait for the standards' policy is quite unacceptable.

2.6 DATA DISTRIBUTION TO OUTSIDE USERS

Where necessary, data distribution by wide area network should be supplemented by, and integrated with, the physical transport of magnetic tapes.

CERN should offer a tape copying and distribution service such that tapes can arrive at regional centres and universities within days of submission of an appropriately authorized electronic request. Regional centres should also have tape copying and distribution facilities.

2.7 OPERATING SYSTEMS AND LANGUAGES

Unix is already used on the majority of workstations, and is available for VAX and CRAY machines. It will soon be possible to install a version of Unix on IBM mainframes, offering, for the first time, the possibility to perform the full range of HEP computing tasks under a single operating system. The recent 'Open Systems Foundation' initiative increases the probability that these versions of Unix will appear very similar to the user.

The effectiveness of using this single operating system can only be measured by making it available to groups of users, or complete experiments, prepared to use it for all their computing.

A Unix service should be started on the central IBM system at CERN as soon as adequate software is available from IBM. An initial small-scale service should receive further resources according to user demand.

In spite of the many deficiencies of Fortran, there is no immediately obvious superior alternative. The CERN community should periodically review the developments in languages, both to add weight to its input to Fortran standardization committees, and to identify any serious alternative to Fortran.

The production of reliable analysis programs for complex detectors is increasingly slow and manpower intensive. The full exploitation of complex detectors at future colliders will require the use of techniques (or languages) to generate low-level code automatically from a higher level of abstraction. No existing language fulfils this need.

2.8 SOFTWARE SUPPORT

2.8.1 General Software Support Problems

HEP must organize its support for HEP-specific software more effectively.

Formal procedures are needed for determining and responding to user needs.

Some freedom should exist to pursue good ideas in the absence of clearly formulated demand, subject to periodic user-dominated review.

Physicists would benefit from overall coordination of the software support provided by CERN, DESY and regional centres. This coordination should be effected by the creation of a Software Support Committee.

A first task for this committee would be to assess the adequacy of current manpower levels.

There is no doubt that widely available and successful commercial software, meeting HEP needs, should be bought and not re-written within the HEP community. Experience with the purchase of less mature commercial software has not been uniformly encouraging. An additional task for the Software Support Committee should be to advise HEP management case-by-case on the relative benefits of buying commercial software compared with writing HEP-specific products.

2.8.2 Data Bases for Experiments

The use of data base techniques and similar access methods for managing large volumes of structured data has turned out to be a key factor in today's experiments. The basic support of Oracle, a commercial data base management system (originally called for by an ad hoc subcommittee of ECFA) has convinced experiments of the value of such general products, although this system is expensive.

However, commercial products do not normally come with a user interface of the desired level, or with the full functionality physicists request. Applications of data base techniques in experiments cover a wide range of areas: administrative directories of collaborators, calibration data, detector descrip-

tions, data models, and many others. The requirements of experiments, nevertheless, are quite similar or even identical. Users of the past have noticed systematic inadequacies of Oracle in different domains, particularly in access speed and in ease of learning. Substantial effort was invested by experiments to overcome these problems, without resulting in generally adequate solutions.

It is recommended to coordinate across the HEP community the definition, writing, installation, and support of data base-oriented software, for all applications common to large experiments. This concerns commercial products, physics-specific application layers to commercial products, and physics-specific products not available from outside our community.

For some applications, such work has to maintain close links with existing or proposed solutions found for data bases in a different context, e.g. administrative data bases in laboratories, or data models used as part of software engineering techniques.

2.8.3 Software Engineering

Some experiments of the generation now approaching the data taking phase have introduced software engineering methodologies and tools in the design of their experiments, mostly in the spirit of pilot projects. The overall experience with such methods is judged positive. Some aspects of software engineering, like *data modelling* and *structured design and analysis tools* are today considered vital in designing large experiments. This is particularly true for the frequent case that the collaborating individuals can not work in geographical proximity; 'vital' in this case has to be understood in its literal sense as: absence of adequate software planning may jeopardize the experiment.

It is recommended to coordinate, at the level of the HEP community, the selection, creation, and support of tools and methodologies in the area of software engineering. This includes the evaluation of commercial products, the conduct of pilot projects, the writing and/or installation of tools, and the ensuing support tasks.

2.8.4 Graphics Support Software

Graphical presentations of data have been recognized by all present experiments to be a critical element in modelling and understanding detectors, data reduction methods, and physics analysis. Despite large investments, the access to graphics remains limited by the available manpower: partly this may be ascribed to the notorious mismatch between visual and algorithmical concepts; more importantly, the rapid evolution of the market in both hardware and software makes the emergence of standards in this area difficult and slow. In addition, the structure of HEP collaborations with their independent decision mechanisms governed by quite different national constraints makes even the possible standardization difficult to impose.

The decision by HEPCCC to impose GKS as general portability standard is generally seen as a success, although the specific product supported by CERN is sometimes criticized. The community's need for CERN's support in this area is considered essential, new activities are suggested to become part of this support, and more user participation is desired. The present staffing level must be called

clearly insufficient.

Looking beyond the next few years, the success of GKS is possibly a partial and temporary one. The GKS interface is seen as too low-level, and standards can be expected to evolve towards more structured data (as in the PHIGS standard), better metafiles, better interactivity, introduction of surface rendering, etc. Experiments need HEP-specific high-level user interfaces not provided by GKS, and of longer lifetime than expected for GKS.

It is recommended to provide sufficient manpower with good technical expertise at CERN to support GKS, the chosen graphics standard. It is further recommended to coordinate much better the existing efforts in the graphics area across the HEP community, in view of harmonizing the non-standard graphics extensions, and with the aim of preparing today the phase of transition towards standards beyond GKS, e.g. by supporting an intermediate-level general interface for application programs.

2.8.5 Education of Physicists

Many common problems of software are solved independently today by different mostly small groups. This favors creative and novel solutions, but often is also seen as an undesirable duplication of efforts. The origin is not only inadequate coordination, but also a clear lack of communication. Existing solutions, even if implemented at a high level of professionalism, rely for their spreading across experiments on oral transmission and person-to-person tutoring more than on professionally written documentation and expertly organized training courses accessible to all.

We recommend that the problem of instructing the community in the use of generally accepted methods or tools be given very serious attention. We suggest as possible and adequate steps the hiring or redirection of qualified staff for technical writing (documentation) and for the organization of repeated technical training courses.

Training courses should be given by the best experts, from inside or outside the community. They should have access to training techniques novel for our community but well known in professional training elsewhere, like video presentations and computer-assisted instruction.

Chapter 3

COSTS AND PRIORITIES

3.1 Costs

The following points are relevant to the cost and manpower tables.

1. It was possible to make reasonably firm estimates for the years 1989, 1990 and 1991. In the years 1992 and 1993, the requirements for capital expenditure and new manpower are expected to be similar to those in 1991.
2. Capital costs are for CERN only. It is assumed that the total cost to universities and regional centres will exceed the cost to CERN. The cost of wide area network links to major regional centres is shown in the table for information only, while noting the current policy that none of these costs are borne by CERN. Line costs are for 2 Megabit/second links at current tariffs.
3. The tables show the capital expenditure and new staff required to implement recommendations which propose the expansion of existing services, or the creation of new ones. The working groups have not addressed the question of whether any of the manpower for new services could be obtained by the reduction or reorganization of other CERN activities.
4. It has been assumed that the CERN Computer Centre will receive continuing funding, on top of the operating budget, for the routine upgrading and end-of-life replacement of tape drives, printers, communications equipment etc. These continuing costs do not appear in the table.
5. No estimate has been made of the cost of buying workstations. The majority of workstations on the CERN site will probably be bought by experiments.
6. The manpower needs for workstations and distributed computing are additional to those required for the communications infrastructure described in the Communications Board report.
7. The new manpower needs for wide area networking are higher than those in the Communications Board report. The numbers in the table reflect our view of the extreme urgency of these developments.

Table 1: Approximate Capital Expenditures

These approximate estimates are for CERN only. The resources for the whole CERN community will be at least twice those for CERN alone. Wide-area line costs are included for information; it is not implied that they should be paid by CERN.

	Equipment Costs (MSF)		
	1989	1990	1991
New real-time projects	0.5	1	1
Mass storage - disks	9	11	11
Automatic cartridge handler	1.4	0.6	2
Mainframe CPU power	2	21	25
Parallel processing	1	3	6
Wide area networking			
Interfaces at CERN	1	1	1
Line costs	4	10	12
Tape copying service (excluding tapes or freight)	0.5	0.5	0.5
Infrastructure (buildings for uninterruptible power and tape archives)	1	1	-
TOTAL	20.4	49.1	58.5

Table 2: Approximate Needs in Manpower

These approximate estimates represent resources located at CERN and serving the whole of the CERN-connected community.

	New staff / year		
	1989	1990	1991
Pilot projects for real time processing	2	4	6
Workstations and distributed computing	10	4	4
Development of parallel processing service	6	4	2
Unix service on IBM	1	1	2
Wide area networking	5	-	-
Tape copying service	1	1	-
Application SW support:			
Graphics including X-Windows	3	3	-
Software Engineering, Data Bases	4	2	2
Training, documentation	2	1	1
TOTAL	34	20	17

3.2 Priorities

HEP computing needs cannot be expressed in absolute terms. The quality and quantity of the final physics output of an experimental programme changes rather smoothly as the total funding for computing, or its division between real-time, CPU, storage, networking, or application programs is varied. Ideally, this section should establish the way in which physics output would change as funding for each component was varied, leaving the management of European HEP to make appropriate informed decisions about resource allocation.

It is not possible to quantify physics output without getting into fruitless controversy. Nevertheless, it is the opinion of the working group that the computing needs described above lie well below the point where further resources would bring negligible physics gains. If the resources were either halved or dou-

bled there would be a significant (10 to 30%?) change in the potential for uncovering fundamental physics. To optimize funding for HEP computing it is necessary to compare these tradeoffs with other ways of improving physics output, such as running LEP for an extra year or even installing spin rotators. The cost of all the additional hardware (on and off the CERN site) needed to meet the stated computing needs of LEP experiments is of the order 10% of the hardware cost of LEP and the LEP detectors.

3.2.1 Recommendations Concerning Mainframe-related Services

The combination of CPU power, storage and networking described above form a balanced set of resources which have, in the longer term, equal priorities. If, after consideration of the likely impact on physics output, management decides that the resource allocation should be below (or above) that needed to meet the requirements described in this report, the reductions (or increases) in resources should, in the longer term, be spread evenly.

In the shorter term, two factors may lead to a need for uneven funding. The first factor is relatively trivial: hardware, particularly when bought at attractive discount levels, is usually offered in a highly quantized form. This working group cannot pre-judge whether pursuing a particularly attractive offer for a mainframe might warrant a delay in disk installations.

The second factor is related to the differing development times for the various facilities. Assuming that floor space is available, the acquisition of disks, or the enhancement of existing mainframes, can be delayed until just before the time when they are needed. Other facilities, which require new technological developments or imply new ways of doing physics analysis, cannot be delayed until the last moment. The following developments must be started very soon if they are to benefit the current CERN physics programme:

- *2 MEGABIT/SECOND LINKS*

It is technically possible to offer immediately a restricted range of services which use the full bandwidth of such links. To develop a full range of services, and to integrate these into the working methods of physicists in time for the most LEP physics analysis, will require that at least some 2 Megabit/second links are available very soon.

- *WORKSTATION INFRASTRUCTURE*

Workstations can only be effectively exploited for LEP physics if they are widely used from now onwards. The current level of support for workstation system software and communications is incompatible with wide use, and setting up improved services is itself a process which will take many months.

- *PARALLEL COMPUTING*

If parallel computing cannot be made widely available to the CERN community, additional mainframe CPU power will have to be installed to avoid serious restrictions on the CERN physics programme. Unlike mainframe CPU power, parallel processing will require a minimum time of one to two years before any generally available services can be offered.

• STORAGE MANAGEMENT

Automatic migration of files between disks and tapes will be needed to support LEP physics analysis. Powerful commercial systems may become available in the longer term, but a simpler HEP-specific system must almost certainly be developed by late 1989.

3.2.2 Recommendations Concerning Experiment-related Services

In addition to investments in mainframe-related services, several recommendations above advocate a more coordinated approach to problem areas common to many experiments. Many of these problems are new, due to the evolution of the market in software and hardware. Some arise because experiments have reached a level of complexity which cannot be handled by simple extrapolation from the past. The solutions of such problems today is largely left to experiments. Indeed, the suggestions to coordinate come predominantly from the physicists in the experiments trying to cope with general problems using inadequate support in staff and, sometimes, money.

The working group can not assign individual priorities to the proposed activities, be they relevant in a shorter term like the application-oriented software for graphics and data bases, or of longer-term impact like familiarization through an R&D program with digital VLSI technology for real-time data control and triggering, the introduction of methodologies in hardware and software design, or the European coordination of many of these activities. We believe that each recommendation, if followed, will make a *visible contribution to the physics programme*. Some of them may be *mandatory for CERN's future activities*: The dependence of a possible LHC program on mastering the real-time problems need not be underlined further.

Part 3

DATA ACQUISITION AND REAL-TIME COMPUTING

Chapter 4

INTRODUCTION

4.1 Working Group Members

The individuals who contributed to this working group at various stages are listed below :

K.Einsweiler	(CERN/EP,Convener)	A.Bogaerts	(CERN/DD)
R.Dobinson	(CERN/EP)	J.Dorenbosch	(NIKHEF)
J.Harvey	(RAL)	L.Levinson	(Weizmann)
A.Marchioro	(CERN/EP)	H.Muller	(CERN/EP)
D.Notz	(DESY)	M.Sendall	(CERN/DD)
H.Von der Schmitt	(Heidelberg)	—	—

4.2 Overview

This report discusses the perceived needs of the experimental physics community at CERN in the broad area of Real-Time Computing in the period up to 1995. This subject extends from the realm of front-end detector electronics where Digital Signal Processing using commercial processors is playing an increasing role in the data formatting and trigger decisions, up to the level of the cluster of Online Computers (at present, usually VAXes) which provide the user interface and other higher level functions in a modern experiment. On the lowest level, there is considerable overlap with electronics topics (special purpose processors, custom IC's, etc.), whereas on the higher levels, there are many concerns in common with offline subjects (software engineering, human interfaces, databases, etc.). In these overlapping areas, we try to emphasize those aspects of the problems which are of particular importance in Real-Time Computing.

The overall structure of the report divides naturally into two sections, one discussing the likely requirements of the expected physics programme, and one discussing the implications for the future evolution of data acquisition systems. It includes some speculations about what will be necessary to support physics in the 'post-LEP' era, which we assume to be dominated by a high-intensity hadron collider. For short, we will call it LHC. We have summarized our suggestions for the future in part 2 (chapter 1) above.

Forecasting the direction that Real-Time computing will take, needs multiple contacts, pilot projects, and a decision making process which includes the experiments. There was only limited time available for the preparation of this document. *This report, therefore, contains requests for further study in several areas rather than suggesting direct actions.*

Chapter 5

A REFERENCE MODEL OF EXPERIMENTAL ACTIVITY

The first requirement for this discussion is a Reference Model, containing our expectations for what experimental activity will be taking place at CERN in the period of 1989-1994.

5.1 Extrapolation of the Current Physics Programme

The following is a summary of the major areas of experimental activity.

- Fixed Target : in this area, we could imagine (as an optimistic estimate) a new large muon experiment, a large neutrino experiment, and a large heavy flavor spectrometer experiment. None of these experiments exist at the level of a proposal, so further discussion is difficult.
- LEAR : Four major experiments are under construction : Crystal Barrel, Obelix, CP LEAR and JETSET. They already have detailed plans for data acquisition systems, and preliminary versions working based on combinations of MODEL, VALET-Plus and OS9-based software. These systems may serve as a barometer for some of the current trends in Real-Time Computing.
- Heavy Ions : the possible approval of the Lead Injector Project would probably stimulate several new experimental proposals. Although the scale of these experiments has not been decided, they are unlikely to be significantly larger than the current experiments participating in this program.
- SPS Collider : both UA1 and UA2 have undergone, or are undergoing, major upgrades. Both experiments will probably continue running through the early 1990's, depending on the actual performance of the ACOL anti-proton source, but their data acquisition systems are already essentially complete.
- LEP : all four experiments will (presumably) start taking data in 1989, and continue well into the 1990's. For the moment, they are in the midst of completing their data acquisition systems, and there is no basis for discussing upgrades before they have experience with these systems under realistic conditions.

We know in principle how to solve most of the problems of these experiments. We can forecast some trends and extrapolate from present systems. It must be emphasized however, that a large effort is still needed on these "conventional" data acquisition systems. This must cover commissioning, running in and improving the data acquisition systems of the LEP experiments, and of those LEAR, fixed-target and collider experiments planning new or upgraded systems. The systems will need to track evolving technology and user requirements. In addition there will be new experiments and tests proposed, amongst them the activities in preparation of LHC, as recommended below.

5.2 Preparation for a Future Hadron Collider

Real-Time computing is driven by the needs of actual experiments. This makes it especially difficult to predict the long-term needs. The discussion here focuses on a future scenario involving a high luminosity hadron collider (LHC). It should be noted that this is a demanding experimental environment, with extreme needs, particularly in the areas of front-end electronics and trigger processing. The nature of the subsequent discussion thus depends very strongly on the detailed design of the experiments. The overview presented here will be very schematic, attempting to highlight some of the problem areas.

We start with a brief description of the characteristics of the machine and detectors, essentially summarizing the recommendations of the La Thuile study on Trigger and Data Acquisition (contained in the CERN Yellow Report 87-07). We restrict our discussion to the 10^{33} luminosity option. Although it appears that a high luminosity option (5×10^{34}) may be necessary to reach the full physics potential of the LHC, the first option is already at the limits of our technological imagination.

We strongly recommend that investigations of the many difficult problems of LHC data acquisition should commence immediately.

The standard machine parameters involve a crossing interval of 25 nsec at the design luminosity of 10^{33} . Combining this with an expected total inelastic cross-section of 100 mb gives about 10^8 interactions per second with an average of 2.5 interactions per crossing. The recommended trigger parameters for a three level system are summarized below (the division of the trigger into three levels is somewhat arbitrary – many sub-divisions can be easily imagined), limiting the discussion to the calorimeter trigger scenario which was emphasized in the La Thuile study. It should be noted that this scenario, while providing a useful starting point, is in serious need of further work starting from a more complete detector model, and more complex physics models.

1. The first level (LEVEL1) has a recommended decision time (latency) of 200 nsec in the La Thuile report. This implies that the data acquisition system must be capable of storing at least ten complete events in a deadline-free pipeline while the trigger operates. A calorimeter with about 5000 trigger cells (substantially reduced from 5×10^5 instrumented detector cells) would be a major source of input for the trigger. This can be compared with the performance of existing SPS Collider triggers (UA1 and UA2):
 - SPS Collider with a luminosity of 4×10^{30} : interaction rate of 2×10^5 Hz, with an output rate of 100 Hz, and a reduction of 2000 in a time of 1 μ sec (the crossing time is 3.9 μ sec so this generates no deadline).
 - LHC : interaction rate of 4×10^7 Hz, with an output rate of 10^5 Hz, and a reduction of 500 in a time of 200 nsec (the crossing time is 25 nsec so pipelined storage of complete events is required to eliminate deadline).
2. The second level (LEVEL2) might be a highly parallel processor system, possibly consisting of a network of commercial processor chips (such as Digital Signal Processors). It must reach a decision in 10 μ sec to keep up with LEVEL1 (assuming a multi-event buffer system between LEVEL1 and LEVEL2 to de-randomize the data). Experience at current hadron colliders has shown that it must use the full calorimeter granularity to achieve the desired reduction factor of 500 (the combined reduction factor in the SPS Collider experiments is roughly 2×10^4 , compared to the desired factor of 2×10^5 needed for LHC).
 - SPS Collider triggers using general purpose processors : input rate of 100 Hz, with a decision time of 1 msec using about 10^3 cells of calorimeter information.

- CDF trigger using programmable sequencers : input rate of 1000 Hz, with a decision time of 20 μ sec using about 10^3 cells of calorimeter information.
- LHC trigger : input rate of 10^5 Hz, with a decision time of 10 μ sec using about 10^5 cells of calorimeter information.

It is clear that these LHC requirements exceed the current state of the art by several orders of magnitude.

3. The third level (LEVEL3) is a large processor farm, receiving events at roughly 200 Hz. It should reduce the event rate to the canonical several Hz at the output. Presumably this reduction, after a previous reduction of 10^5 , will require very sophisticated analysis algorithms, and consequently a very powerful and sophisticated computing environment.

The performance requirements for the data acquisition system can be similarly summarized (again, the discussion is for a calorimetric detector – a more complex detector will surely introduce additional problems) :

1. The LEVEL1 pipeline for the proposed calorimeter must store 1 MByte of data every 10 nsec, for the duration of the trigger decision. This means that the required pipeline capacity is about 20 MBytes with an input rate of 10^{14} bytes per second. The other detectors will probably require even larger pipelines, depending on the details of their readout, and on the speed of any other LEVEL1 trigger decisions.
2. There is likely to be a data compaction phase in the front-end electronics, preceding the LEVEL2 decision (using a DSP per channel, or equivalent). This system should be capable of performing pulse finding and zero-suppression, with a reduction of more than 100 in data volume after 10 μ sec.
3. After the LEVEL2 decision, the data is finally transported from the front-end into the LEVEL3 processor farm. At this level, one expects 200 KBytes at a rate of 200 Hz, amounting to 40 MBytes per second. This rate is probably manageable with today's bus systems, but no-one has any experience feeding such a rate into a processor array.

A further comment should be made about the simple monolithic description of triggering and dataflow presented above. In an experiment operating in the complex physics environment of LHC, there will be many different physics analyses proceeding in parallel. Many of these will want to examine processes which occur at rates much higher than the output rate of several Hz normally discussed for permanent storage. It will be quite feasible to perform some of the DST generation in the LEVEL3 farm, vastly reducing the amount of information to be recorded for some classes of events. One would expect to see a migration of analysis from the offline environment (after permanent storage) to the online environment (before permanent storage), as experience accumulates, in order to cope with the data flow problems inherent in analyzing high cross-section physics processes. Thus, the LEVEL3 farm becomes a hybrid computing facility, containing elements of offline and online environments, which will need to support physics analysis as well as event filtering. This broad overlap with what is traditionally called offline computing should be examined in more detail. Since the bandwidth for transferring data to processor memories is much higher than that for permanent storage, this may be one of the most important tools in performing analyses on large event samples, avoiding the bottlenecks associated with the full offline computing environment.

Chapter 6

IMPLICATIONS FOR DATA-ACQUISITION SYSTEMS

6.1 Background

This group has not carried out a detailed analysis of the reference model, but it is clear that there will be a great deal of effort expended on making the major new systems work (especially those of the LEP experiments), and learning from their deficiencies. Much of this effort should be invested in such a way as to be a basis for future LHC systems, which in addition pose new and difficult problems. In discussing LHC, the natural bias of physicists is to concentrate on problems of data-flow and data reduction, which directly affect the deadtime or even the feasibility of the experiment. Coping with these problems already requires a substantial effort in programming intelligent devices.

In addition, we need to keep in mind that in a modern online system a huge effort goes into monitoring, control, and the organization and presentation of information. A typical large experiment may have hundreds of thousands of lines of code for detector monitoring and calibration, most of it detector dependent and written by experimental physicists for this experiment. A roughly equal amount of code provides a framework for the experiment-dependent programs. This latter code is often general, and may be used by different experiments. It complements the facilities, such as operating systems or compilers, which are provided by manufacturers. Together, the general (but HEP-produced) code and the manufacturer's software form an environment in which the experiment-specific applications are developed and run. The quality of both the application development environment and the final running system are crucial.

The following list of items gives some of the background to data acquisition systems currently in use or under development.

- Smaller experiments have needs which are frequently every bit as technically demanding as those of larger experiments. In particular, they have complex triggers and data acquisition systems which contain a mixture of Online computers, workstations, and microprocessors. They have less manpower than larger experiments and are therefore less able to provide for themselves and more interested in standard, off the shelf systems. If CERN wishes to continue its current broad program, it is important that these experiments are not neglected in the overall picture.
- Large experiments have very large data acquisition systems, with many specific detector and trigger requirements. In this environment, there is a trend towards experiment-specific solutions (and away from off-the-shelf solutions). Experiments are sufficiently large to feel that they can develop their own private standards, tuned to their specific needs, and taking advantage of the latest technology and software products. In addition they often find it difficult to reach agreement internally even without the added constraint of converging with other experiments or support groups. This makes centralized standardization difficult.
- All experiments want systems which allow independent testing and setup for each detector, as well as an integrated system for normal data taking. Ideally, these independent components of the final system should appear first in the home institute responsible for a particular detector component, and then migrate to CERN for integration, allowing a consistent environment to be

used throughout the commissioning of an experiment. This means that everyone wants a partitioned system, where the sub-systems are each capable of working independently as a complete data acquisition system or together as a component of the total system. It also means that the cost and complexity of a subsystem should be small enough that individual institutes can make useful contributions by buying their own system for home use.

- The functions originally supplied by the main Online Computer are being divided into two classes: 'front-end' functions such as data acquisition and monitoring (things which demand serious real-time response and need lots of local processing power) and 'back-end' functions such as user interfacing, system integration, information presentation, centralized databases, etc. This environment requires very modular software capable of providing basic services (message reporting, run control, information presentation) in a distributed processing environment with a mixture of Online Computers, workstations, and many microprocessors. In addition to the basic data acquisition services mentioned, there is a tremendous need for a coherent software development environment containing good compilers, debuggers, code management tools, software design tools, utility libraries and proper documentation.
- Local Area Networks play a vital role in these systems, often serving as the basic means of communication between the many embedded processors, workstations and Online Computers. These networks are almost universally based on Ethernet/IEEE 802.3 as the connection medium and support functions such as program loading, remote file/printer access, remote sampling of raw data and histograms, slow controls, and overall system coordination. The software support in this area is improving, but a great deal more work is required before the quality of support for a heterogeneous environment approaches that of a homogeneous one (such as that provided by DECNET and Local Area VAX Clusters).
- Many recent experiments have chosen VMEbus as their data acquisition system integration bus. The front-end electronics itself (ADC's, low level processing, etc.) is contained in a mixture of Fastbus, VMEbus, CAMAC, etc. The reason for this choice of VMEbus at the higher level is the commercial availability of a large, economically priced, selection of microprocessor and peripheral boards (memories, communications controllers, tape controllers, etc.). In addition, there is more extensive commercial software support available for VMEbus boards. It seems clear that the immediate future will see continued use of systems containing VMEbus and Fastbus, as well as CAMAC, to support the different environments required at different levels of experiments.
- Event recording in these experiments is not necessarily performed by the main Online Computer and there are strong desires for new data recording mechanisms which can provide cheaper, higher density storage. The SCSI intelligent mass storage bus seems to be emerging as a standard connection medium, and could provide a standard method for accessing data recording peripherals from either the main Online Computer, or from the Front End processors.

At the moment, the market for new recording devices/technologies is exploding, and we recommend some form of CERN- or HEP-wide coordination to rationalize the situation. There is a need to standardize the recording formats as well as the types of media in order to support the transport of data between the many online and offline systems used in a typical experiment.

- A large number of physicists and engineers are involved in the design, testing, integration and maintenance of particle detectors and their associated data acquisition and control electronics. These activities take place in the experimental areas at CERN, in the development laboratories of the collaborations both at CERN and in the home institutes, in support groups and in industry. There is a need for adequate hardware and software tools. Such systems should be based on commercial products as far as possible. They are likely to need complementing by hardware

and software specific to high-energy physics applications, as these systems must follow the trend in instrumentation typical for experiments. where they can be used by physicists and engineers alike.

6.2 Synopsis of Areas Requiring Further Study

This working group has not attempted to reduce the difficult problems described above to a simple list of recommendations. What we offer instead is a summary of some of the important areas, with indications of future directions. Before entering these details, it is worth noting that, while the raw performance required for LHC is rather awesome (bytes/sec, channels, decisions/sec, etc.), the real difficulties are likely to lie in the extremely intricate collection of software required to implement such a system. Such problems of complexity are already encountered in the present generation of experiments, and will be even more serious for LHC. Apart from the challenge of actually making such a complex system work, we need to make sure that it presents a coherent and comprehensible interface to the physicists who must operate the experiment.

There will be many thousands of powerful processors embedded in the experiment and appearing in arrays for higher level processing. It will require an unprecedented level of organization and discipline to create and maintain such a system. The design process must be more formal and sophisticated than anything attempted before, and must use the best computer aided design tools which are available. It will be necessary to spend a substantial amount of time defining, simulating, and optimizing the system at the architectural/protocol level before proceeding to the detailed implementation phase. In addition, the very high data flow rates will require very careful integration of the detector digitization electronics with data compression and trigger processing, thus coupling the design of the detectors and trigger/data acquisition system from the very beginning.

A more structured, 'top-down' design is vital for a system of this complexity, but it is especially difficult to carry out in the relatively unstructured and multi-institutional environment of high-energy physics. The use of modern design tools allows a group of designers (be they physicists, programmers, or engineers) to describe a design at the behavioral level, in terms of large blocks, and then proceed to specify the internal structure and detailed timing of each component. As more timing information is added, the corresponding simulations of the system become more accurate, allowing further refinements in the overall design. One can imagine that these kinds of detailed simulations of hardware and software will play as significant a role in the design of the data acquisition and trigger for an LHC experiment as the use of Monte Carlo simulations does today in the overall design and understanding of a detector. As part of this process, clear specifications are created for interfaces between the components of the system which are independent of the details of the implementation, allowing the distribution of design effort without losing control of the coherence of the system. This structured design process has been found to be essential for the commercial design of complex systems, and HEP has a great deal to learn in this area before embarking on the design of an LHC experiment.

We therefore recommend to support pilot projects of combined software and hardware engineering methodologies, in order to select or develop methods and tools that can be used to design the next generation of data acquisition systems.

1. APPLICATION SPECIFIC INTEGRATED CIRCUITS (ASIC) AND VLSI METHODS

The large channel counts and high performance demanded from the front-end electronics (pre-amplifiers and digitization circuits) needed for a typical LHC experiment will require extensive use of VLSI methods, with the additional complication of a high radiation environment. It may also be necessary to include actual processing elements in these front-end circuits to support the management of the large pipeline required for the LEVEL1 trigger, as

well as the data compression required (it may not be desirable to transport the vast quantities of data generated by the detector elements). This is an extremely complex area, involving the overlap of the detector design and the first levels of the trigger and data acquisition system. It is clearly an area requiring much more detailed discussion, such as that at the Fast Trigger meeting held at CERN in November 1988.

Given the need for very sophisticated design skills, it is strongly recommended that CERN invest significantly greater resources in the application of VLSI techniques to trigger and data acquisition problems than it has in the recent past.

We would think that one of the most promising ways to acquire new skills and knowledge, which has proven very useful in the past, is the pilot project. This allows one to gain experience with new tools and techniques, while solving a useful problem and avoiding the time pressure which often forces the use of conservative solutions.

2. *BUSSES*

In this area, we try to summarize the situation with VMEbus and Fastbus, and then attempt to gaze into the future to see what is coming next.

- *VMEBUS*

The major advantage of this bus is the commercial availability of a wide selection of processor and peripheral modules. There were initial difficulties because of the inadequate level of standardization — modules from different manufacturers didn't always coexist happily, but these have largely disappeared with time. The mechanics and cooling are more appropriate for single crate desktop systems, and the small board size makes it unattractive for front-end electronics. There are also difficulties because of the lack of multi-crate and multi-processor standards and concepts, i.e., the error handling and arbitration are awkward, there are no standard registers, configuration is generally by jumper instead of software, the interrupt support is primitive, there are no geographical addresses, and each large system must invent its own interconnect hardware. These problems make it very difficult to create any standard software for a multi-processor system. Nevertheless, despite its many technical limitations, VMEbus has been extremely successful because the large commercial market results in reduced costs and immediate availability of the latest technology, as well as the availability of commercial software. This availability means that upgrades in CPU power or memory density can be made fairly easily.

- *FASTBUS*

This bus, in contrast to VMEbus, is an HEP specific bus. It was tailored to HEP needs, and includes very powerful system integration concepts, and a quite clearly defined, standard specification. The specification has also been extended to include standard software for supporting data transfers, interrupts, etc. Although the specification has led to few problems of incompatibility between modules, it was created without regard for any commercial aspects — HEP is probably the only user. The result is that many products are expensive and require a large HEP design effort. A major problem in this respect has been the absence of a proper set of protocol chips to ease the burden of creating Fastbus interfaces which implement the complex protocol in a standard way. In addition to the hardware problems, there is an industry trend towards software packages for particular 'platforms', that is combinations of processors and peripherals. If one starts with a custom design, there is an additional burden of

porting even commercial software to a new environment (modifying device drivers, etc.).

On the other hand, the use of Fastbus in front-end electronics has been fairly successful. The large board size, mechanics and cooling have contributed to the implementation of high density, economical systems. Furthermore, the standard includes high speed synchronous transfers, and the terminated ECL backplane of Fastbus is not yet close to its bandwidth limitations, leaving more room for future expansion, and perhaps providing a better match to the needs of the next generation of experiments at a hadron collider.

The conclusion is that these two busses have somewhat complementary advantages, and are both likely to survive in the future. In current designs, there is a tendency to use Fastbus at the lower levels in the system (industry rarely makes the digitization modules that are needed, either in Fastbus or VMEbus, and Fastbus has many advantages here), with VMEbus at the higher, processor-oriented levels (where the commercial support is substantially better in VMEbus).

It is quite likely that new bus systems will emerge in the commercial world before the construction of any LHC experiment occurs. The success of VMEbus is a strong indication that the next bus system used in high energy physics should be an industry standard, but it seems unlikely that industry will find it important to implement large multi-crate systems with a distributed high bandwidth interconnect. Nevertheless, there are two short-term solutions that bear watching. VXI is a proposed extension to the VMEbus standard, oriented towards instrumentation, with larger boards, better cooling, and ECL support among other features. It should make VMEbus a much better place to build front-end electronics. Futurebus is a proposal incorporating much better support for multiple processor systems in a standard which could be strongly supported commercially, and provide a migration path from both VMEbus and Fastbus. For the longer term, there is the Superbus proposal (now known as SCI, for Scaleable Coherent Interconnect) which is a very high speed, transaction oriented, point to point system benefitting from advances in high speed network protocols. This could be of interest as the framework for building very powerful processor arrays.

3. NETWORKS AND COMMUNICATION

Everyone agrees that we should continue to follow emerging standards. Simple and relatively cheap hardware interfaces are available, but the integration of standard communications software into the environment of front-end processors is still difficult, and further improvements are certainly desirable. The use of network software to integrate the many microprocessors buried in the data acquisition system with the host computer is a prominent feature of many of the major present data acquisition systems at CERN and DESY. In general, the coherent integration of standard network software into the system software running on the microprocessors is a very important requirement for any microprocessor software packages to be considered in the future. Such systems should support a standard transport level, as many packages do now, but more complex layers of software are also required. Standard examples include remote file access (including directory modifications) and remote login. As a further example, online networking tends to be very transaction oriented, and high level packages such as Remote Procedure Call (RPC) are proving to be very useful for distributing functions within a multi-processor system.

4. *LANGUAGES AND OPERATING SYSTEMS*

- *LANGUAGES*

Data acquisition systems contain a mixture of time critical problems and high level problems, making it unlikely that a single language can provide the optimum solution for all problems. It is important for the future that we be able to use modern languages which are structured and well supported by software engineering tools. It is unlikely that FORTRAN will meet these requirements (although its use in some online applications is likely to be favoured by physicists for some time). There is already a strong trend to use the C language in the data acquisition environment. C is fast gaining importance in the outside world, as a reasonably portable and efficient systems programming language. Its links with the Unix operating system are another factor in its growing popularity.

The computer science community has placed significant emphasis on designing languages for large, multi-processor software projects. It is possible that we could benefit from their efforts, which have produced languages like ADA – such languages can encourage more modular, portable software and may support a measure of object-oriented design. ADA also has associated with it a powerful development environment. In addition, these languages can simplify programming in a multi-processor system because they already contain primitive concepts which are usually the domain of the operating system.

An example of an interesting combination of language and hardware is the INMOS Transputer supporting the OCCAM language. OCCAM is particularly powerful in its ability to allow users to develop multi-processor software on a single processor and then distribute the applications over many processors (although this language is not readily available for other hardware).

Taking a somewhat more futuristic approach, one can hope that we will enter an era where the use of behavioral descriptions and very high level specifications will start to supersede the actual writing of software in the current sense. Unfortunately, this objective is difficult and it may not even arrive in time for LHC. We conclude that physicists need much more experience with the use of modern languages, and in particular in their application to large software projects.

- *OPERATING SYSTEMS*

It is desirable to provide a coherent software environment, allowing users to work with a complete set of similar tools at all levels, from the Online Computer down through the general purpose processors to the application specific processors. This coherent environment will need to include support for centralized databases, resource management, run control, error and status reporting, information presentation, human interfacing, and all of the other services required for the organized collection of data. It should be integrated as far as possible with the corresponding facilities of the computer centre or distributed offline processing and interactive facilities.

At present, Unix seems to be the only candidate for a hardware-independent industry-supported operating system. It is already available on many workstations and widely used as a basis for microprocessor cross-software development systems. It is offered by DEC on the VAX family of computers, and is on its way to becoming more widely available on mainframes as well. Many efforts are underway to improve the current level of standardization, most notably the Open System Foundation (a

consortium of many industry giants) which is committed to creating a non-proprietary standard.

However, Unix is not uncontroversial. The VMS operating system is very popular, and has been chosen by the vast majority of present experiments in the online environment. This is probably due to its scope, user-friendliness, and manufacturer support. Second, Unix is known to have real-time problems, which may make it unsuitable for many applications on embedded microprocessors, without kernel modifications. There are efforts to make a 'Real-Time Unix', but it is not yet clear how successful this will be or how it will relate to attempts to standardize real-time kernels. Independent of any arguments about the merits of Unix, no experiment known to us today uses Unix in the online environment.

It is therefore by no means clear what will be the future role of Unix in the online environment. Nevertheless, in view of its potential in providing an integrated environment on a wide range of computers, the working group feels that Unix must be taken seriously as an option for the future.

We recommend that a study should be made of whether and how to introduce Unix in the data acquisition environment. This study should take into account the balance of technical and commercial arguments for Unix vis-a-vis other systems, the opinions of users, and the extent to which Unix can be used in real-time applications.

The area of real-time kernels is finally coming under scrutiny by a sub-committee of VITA (VMEbus International Trade Association), which contains representatives from all major producers of real-time kernels. Work is proceeding on calling conventions, driver interfaces, and the harder problem of a standard real-time kernel definition. The latter is known as ORKID (Open Real-Time Kernel Interface Definition). It should allow vendors to provide standard drivers with their peripheral boards, etc. (At present this work is oriented toward the Motorola 68K family of processors).

Nevertheless, in this area there seems to be a conflict between the real-time needs for front-end processors and the full operating system features which many find desirable. There are essentially two schools of thought, although one can hope that this situation will disappear in the future with the release of a new system combining the best features of both worlds.

- The embedded system approach relies on running cross-software on a host system, with only minimal services provided directly in the front-end processor. This can avoid the problems involved in creating a high-performance real-time Unix, and also the complication of creating a hardware environment to support the full operating system. The user would like to labor away on his favorite workstation, and manipulate programs in a remote processor as though they were running on his native machine.
- The full operating system approach today typically involves a fusion between a real-time kernel and something Unix like at the user level, providing compilers, debuggers, editors, file systems, etc. all running in the front-end processor itself.

The advantages of the first approach are a uniform high-level software development environment on a workstation with windows and graphics, combined with full real-time performance in the front-end processor. The disadvantage is the requirement of sufficient high-level CPU power to support fast response during software develop-

ment. On the other side, a system which provides a high quality set of compilers, editor, and multi-task debugger in a friendly way requires a complex hardware/software environment. It is not clear that we really want all of the complications of a full operating system in our front-end processors; there may also be penalties to pay in performance.

We recommend that standardization efforts on real-time kernels, and associated system software should be followed closely, and taken into account in planning future data acquisition systems.

5. DISTRIBUTED ONLINE SYSTEMS

A key technical question is: how can the tasks of an online system best be distributed to achieve efficiency and cost-effectiveness, whilst maintaining an effective user-friendly control, reporting, and information management system? There is no easy answer to this: it needs serious study based on experience with existing systems and the availability of modern techniques and tools. There are at least two kinds of distribution to be considered:

- Distributing functions between embedded microprocessors and higher level systems. This is done mainly in the interests of efficiency. Here there has already been a substantial amount of work, but there is a very wide spread of solutions: we should aim at convergence in the next generation of experiments.
- Distributing functions over workstations, conventional minis, (and perhaps the computer centre). Here the motivation tends to be user-friendliness, cost effectiveness, and the availability of attractive products. We need to track such products, and see how to make the best use of them in our environment.

The distribution of the functions performed up to now by the online computer, their modularization by sub-detector (with later integration into the overall production system), and the incorporation of "off-line" style tasks into the online environment, have usually been tackled in an ad hoc way.

More formal techniques for distributing functions will need to be studied, adapted, or developed. Some are more appropriate to loosely-coupled systems and others to tightly-coupled ones. They include facilities associated with languages such as ADA or OCCAM, techniques such as Remote Procedure Call (already widely used at CERN in online applications), object-oriented systems, distributed operating systems and many others.

The question of how best to distribute the tasks of an online system in an efficient and cost-effective way should be investigated. HEP needs to participate more actively in following the trends in research, and associated industry developments.

In addition, although functions may be distributed, it is essential to give a coherent picture of the experiment to the physicists. This requires attention to the management of information in the distributed environment and ease of access to it. Debugging in a distributed environment is also very important.

Both database-management techniques and user-friendly human interfaces are examples of areas where development is rapid outside our field. We need to keep abreast of these developments and learn to apply them in our environment, complementing them with in-house developments where necessary.

This whole subject is the domain of a separate group, but its significance, especially for LHC can hardly be over-emphasized. It is important to note that these tools serve a role both on the technical level, reducing design errors and optimizing performance, and also on the social level, by allowing the coherent development of a complex design by many autonomous groups.

- These methods are applicable to both hardware and software design. The use of modern tools enforces a systematic approach, improves integration, etc. In fact, our design problems are usually a question of system design, whether hardware or software, and many modern software design methods (SASD,...) are really system design methods that could be applied much more generally to trigger and data acquisition systems, not just to programs.
- In addition to the analysis of the structure of problems, provided by current software engineering tools, real-time systems involve a time element which requires the ability to simulate the operation of the system. This capability is only starting to appear in software design tools.
- We would like to benefit from the development of so-called high level description languages for circuit/logic design to allow us to specify and simulate hardware and software designs and to experiment with the distribution of functions within a system. Here, we can learn from the methodologies applied to VLSI design by first creating an abstract behavioral description, which can then be translated into a structural description, leading to the specifications for actual components.
- In addition to these formal approaches, a number of issues related to good engineering practices should be given some attention.
 - Modularity is important because it can provide a way to replace subsystems painlessly in order to track evolving technology or follow experiment upgrades. It also allows greater freedom of choice to experiments wishing to take advantage of common software packages (this was the main motivation for the MODEL project for LEP data acquisition).
 - Good layering and user interface design is also important. It protects the user's investment in monitoring code against changes inside the central data acquisition system. This is very important in view of the large investment in this area. One should not underestimate the long lifetime of software, which is desirable and inevitable given the size of the investment.
 - The investment in existing systems must be taken seriously, both in terms of code and experience in both user and support groups. The existence of a set of well-designed products allows an experiment to set up or make changes with a shorter lead time than if everything had to be redesigned from scratch. To make an analogy with hardware: no CERN experiment has thought it sensible to use Multibus; the investment in existing systems far outweighs any advantages it might have in a given situation.

Given these arguments, general-purpose packages are likely to be an important component of future data acquisition systems. Their planning, and combining their creation and maintenance with formal design methodologies, will be a central issue of future support.

7. RELIABILITY

The issue of reliability will take on much greater significance in the complex systems which we are proposing. It is associated with other issues such as error/exception detection and handling, fault tolerance and redundancy. Today's systems have been designed with very little consideration for these issues, and they survive because of their relative simplicity. A simple example is the transport of data in a contemporary data acquisition system. This data works its way through a series of memories, which generally have at best a parity check, and then proceeds via cable segments, which again contain at best a parity check, finally arriving at the 'Online Computer'. In the end, there are many possible failures which can remain undetected for long periods, requiring frequent hardware testing in order to stumble across the problem. This method of operating is inconsistent with the continuous use of the data acquisition system which will be required by the experiment itself.

In the future, it will probably be necessary for individual processing elements to be able to detect hardware faults and also for the system as a whole to isolate and tolerate faults of many kinds – if not, the mean time between failures is likely to compromise the operation of the system, as well as the user's confidence in the final results.

8. STANDARDIZATION

Although everyone expresses some fear that too much standardization can prevent the kind of innovation we need to keep up with technological advances, many reasons can be cited in favour of standardization:

- more efficient (faster, and with less effort) access to commercial products
- better 'survivability' across hardware and software upgrades
- improved productivity of designers, reduced duplication, etc.
- simpler collaboration, even within the same experiment

This is a very complex area, with many possible levels of standardization : international (IEEE, ISO, ESONE, etc.), groups of manufacturers (original case for VMEbus, 3480 cassette, etc.), and HEP (HBOOK, ZEBRA, etc...). A general conclusion of our discussions was that HEP should try to define areas where its needs are similar to those of industry, in which case we should try to standardize and use commercial products as often as possible (be they hardware or software), as opposed to those areas where we will continue to require unusual products, in which case HEP should try to reap the benefits of the revolution in computer aided design tools to ease the job of building our own products.

A major factor in evaluating whether to use a commercial versus local product should be a full cost estimate, including labor costs as well as proper estimates of the true maintenance requirements (especially for software).

Some areas where standardization is interesting are :

- commercial processors and their associated system software
- design tools for the creation of application specific processors
- recording media and the associated data formats, to support transparent exchange of data

- programming languages and their environments, families of compilers, allowing language mixing, symbolic debugging, etc.
- software packages and their interfaces to the user

Industry is making progress in many of these areas: we have pointed out some recent developments in an earlier section (Languages and Operating Systems). This issue of standardization is a general one and applies to other areas as much as to the online field. Careful consultation with users is required as well as attention to technical, support and commercial factors. The size and loose structure of modern collaborations make solutions quite difficult in our community.

We need to set up mechanisms to make sure that the user community participates in the selection and/or definition of standards, so that the standards selected are widely used and properly supported.

9. USE OF COMMERCIAL PRODUCTS/RELATIONS WITH INDUSTRY

Several general comments can be made, in addition to the points raised in the Standardization section.

- It is probable that HEP could benefit much more from the sophisticated software being commercially written for real-time applications, image and signal processing, user interfacing, etc., These are areas where there is likely to be an overlap between commercial needs and HEP needs, but there is a lack of further information. This is an area where CERN should make a more systematic evaluation of the possibilities.
- We should explore ways of using and supporting industrial expertise more intelligently. This is particularly true in electronics design and construction, and also in software development. It would be in CERN's interest to define a more elaborate tendering policy to foster contracts based on criteria other than minimum cost, in order to encourage companies with a high level of technical/engineering expertise. There is a tendency to rely on outside companies solely for production, and not for design and development. In addition, it would be helpful to simplify the procedures by which CERN may specify/prototype hardware and software for commercial production, distribution and maintenance. This is particularly true in the area of large software projects.
- HEP should attempt to alter its methods for cost estimation to properly account for the true expenses involved in designing our own electronics and writing our own software. Commercial products will always look impossibly expensive if their price is compared to HEP products whose 'cost' was evaluated without regard for the manpower required to create them. There is also a tendency to ignore maintenance costs — HEP should pay more attention to the full life cycle of hardware and software when estimating costs.
- In addition, we need to follow developments in other HEP laboratories, and in other fields which share some of our problems. Possibilities for collaboration should be explored. In particular, research projects carried out at CERN under the umbrella of the LAA Project must be expanded and fully integrated, and those supported by SSC R&D funds should be closely followed.

10. *MEDIUM-TERM EVOLUTION OF DATA ACQUISITION SYSTEMS*

In view of the arguments above, we need to plan for the evolution of data acquisition software both for conventional experiments and in preparation for LHC. We should develop a model of what the data acquisition system of a "conventional" experiment starting around 1991 should look like. An analysis of what is needed for such an experiment, and what is available both within CERN and outside, would allow us to make more serious estimates of what work is required. This study should be a joint effort between experienced users and online support and development staff. Such a study could be linked to the startup of a real experiment. It would be complementary to the research required for the special needs of LHC; liaison between these two activities should be as close as possible.

11. *SUPPORT FOR LONG TERM R&D*

It is also clear from the preceding sections that it will be important to embark on research and development work in many areas in order to create the hardware and software for an LHC trigger and data acquisition system. Some of this work is already starting in the form of workshops such as the 'Meeting on Fast Triggers, Silicon Detectors, and VLSI' held at CERN in November 1988, and the 'Workshop on Triggering and Data Acquisition for Experiments at the SSC' (Toronto, January 1989). Their main topic is the close relationship between detectors and the trigger and data acquisition system as necessary for LHC, as well as the overall system architectures which might be needed.

We should emphasize here that 'R&D' mostly relates to technologies or methods already used elsewhere, but largely unfamiliar to our community. Such R&D may be to a large extent a question of learning how to use, and adapt where necessary, the more sophisticated tools and products of the commercial world. For instance, an R&D project could involve learning to use structured design and simulation tools to create a complex processor module and to optimize its design inside a particular model for a data acquisition system. This does not involve true research, it is more a question of education in the use of certain tools and techniques, and an adaptation of those tools (and our way of thinking) in order to solve the difficult problems posed by the LHC environment. We emphasize the need to modernize the approach of physicists, engineers and programmers to solving problems in the field of data acquisition. This process is most effective if it occurs while solving realistic problems, but it should take place in an atmosphere which is somewhat free of the need to meet the deadlines which govern currently running experiments. Only in this sense we call it 'research' as opposed to 'support'.

This work will need longer-term support, and we strongly recommend that CERN trigger a discussion on the means for supporting long-term projects in the general area of trigger and data acquisition system development. It seems clear that CERN cannot tackle this problem alone, and that a broader European initiative will be needed.

Part 4

**CENTRAL AND DISTRIBUTED COMPUTING, SOFTWARE
SUPPORT**

Chapter 7

INTRODUCTION

7.1 Working Group Members

Central and Distributed Processing and Related Questions

Richard Mount	(Caltech/L3, convener)	Brian Carpenter	(CERN DD/CS)
Manuel Delfino	(Barcelona/Aleph)	Fridolin Dittus	(Caltech)
David Foster	(CERN DD/SW)	Chris Jones	(CERN DD)
Hans Hoffman	(DESY)	George Kalmus	(RAL)
Mike Metcalf	(CERN/DD)	Alan Norton	(CERN/UA1)
Dieter Notz	(DESY)	Alan Poppleton	(CERN/UA2)
Andre Rouge	(IN2P3)	Emilio Pagiola	(CERN/EP)
Les Robertson	(CERN/DD)	Julius Zoll	(CERN/EP)

Usage of Data Bases in Experiments

Luc Pape	(CERN/Delphi, convener)	Ken Knudsen	(CERN/Omega)
Andy Parker	(CERN/UA2)	Alois Putzer	(Heidelberg/Aleph)
Harry Renshall	(CERN/DD)	—	—

Software Engineering for Experiments

Tony Osborne	(CERN/DD, convener)	Nick Ellis	(CERN/UA1)
Stephen Fisher	(RAL/Aleph)	Fred James	(CERN/DD)
Gottfried Kellner	(CERN/Aleph)	Henri Kowalski	(DESY/ZEUS)
Karl Gather	(DESY)	Chris Onions	(CERN/DD)
Paolo Palazzi	(CERN/Aleph)	John Poole	(CERN/LEP)
Otto Schaile	(CERN/Opal)	Roman Tirlor	(Saclay)
Ian Wilkie	(CERN/SPS)	—	—

Graphics Support

Francois Etienne	(IN2P3-CCPM, convener)	Daniel Bertrand	(Brussels/Delphi)
Rene Brun	(CERN/DD)	Alan Grant	(CERN/Delphi)
Frank Harris	(Oxford)	David Myers	(CERN/DD)
Jean-Pierre Vialle	(IN2P3-LAPP)	Wojciech Wojcik	(IN2P3-CC)

Training

Stephe O'Neale	(CERN/Opal, convener)	M.Goossens	(CERN/DD)
----------------	-----------------------	------------	-----------

7.2 The Mandate of the Working Groups

A study of off-line computing and general software support for the CERN physics programme in the 1990s.

Including:

- Central and distributed batch, interactive and mass storage facilities.
- On-site and off-site data links (in collaboration with the communications board).
- Balance between on-site and off-site computing.
- Software support from systems to applications, and associated problems.

Chapter 8

THE COMPUTING NEEDS OF EXPERIMENTS

This chapter will examine the needs of each area of the CERN physics programme at the varying levels of detail allowed by existing studies and the plans of individual experiments. The total needs will be summarized in the following chapter, together with a more general consideration of the distributed computing environment which can best meet these needs.

8.1 LEP Experiments

8.1.1 The MUSCLE Report

The computing needs of the LEP experiments have been reviewed in some detail in the MUSCLE¹ report. The estimates were made per experiment, on a global (worldwide) basis, for the "years" "1989", "1990" and "1991" when the number of Z^0 events accumulated by each experiment reaches 1 million, 4 million and 10 million respectively.²

The following three sections are reproduced verbatim from the MUSCLE report.

8.1.1.1 Processor power per LEP experiment.

The following table summarizes the estimates for the different types of processor power that will be needed by each experiment. The attentive reader will notice the question mark attached to the Total? line. It is, indeed, rather unclear (and perhaps not even very meaningful to ask) how the different types of processor power should be added in order to come to a total. This is because there are many ways in which this power can be provided; some options permit sharing the power between different tasks, while for other options such sharing is effectively excluded. The major sharing that we believe might be effective would be to carry out some of the simulation work either using the interactive capacity (workstations etc.) during the middle of the night, or using the capacity needed to generate the master DST outside beam periods.

¹ The Computing Needs of the LEP Experiments, CERN/DD/88/1, January 15, 1988.

² On the assumption that these "years" will have 3000 hours of beam time, the corresponding average data taking rate for Z^0 events during beam periods is 330 events/hour in "1989", 1000 events/hour in "1990", and 2000 events/hour in "1991".

"Year"	"1989"	"1990"	"1991"
Monte Carlo generation	4	12	24
Processing of MC events	1	3	6
Generating the master DST	3	8	16
Accessing the DSTs	8	9	13
Extracting the physics	8	8	8
Total?	24	40	67

Processor Power per Experiment (CERN units)

Table 1

8.1.1.2 Data storage per LEP experiment

The following table summarizes the estimate for the total volume of data, in Gigabytes, that will have to be stored by each experiment. Note that the expansion needed for making duplicate copies is left to the appreciation of the reader. By the end of 1991, if the data are held on cartridges of 200 Megabyte capacity and no cartridge can be recycled, then each experiment will need a total of at least 35000 cartridges to hold these data.

"Year"	"1989"	"1990"	"1991"
Raw data	400	1600	4000
Master DST (real events)	20	80	200
Team DSTs	40	200	440
Personal DSTs	10	10	10
Simulated raw data	200	800	2000
Simulated DST events	20	80	200
Duplicate copies	?	?	?
Total	690	2770	6850

Accumulated Data Volume per Experiment (Gigabytes)

Table 2

8.1.1.3 Data manipulation

The following table gives the number of cartridges that each experiment will have to mount per hour in order to access the DSTs. The data are assumed to be stored on 200 Megabyte cartridges. One entry deals with the (hopefully hypothetical) case that no disk space is available, while the other assumes that each experiment has 100 Gigabyte of disk space available to hold the compressed team DSTs and personal DSTs. 20 cartridge mounts per hour corresponds roughly to an average data rate of 1 Megabyte/sec.

"Year"	"1989"	"1990"	"1991"
With no disk space	40	50	65
With 100 Gigabytes disk space	2	10	30

Cartridge Mounts per Hour per Experiment

Table 3

8.1.2 Beyond the MUSCLE Report

The MUSCLE report restricted its scope to the first 30 months of LEP operation. What happens beyond "1991" will depend on the physics uncovered by LEP in the Z^0 region, and by the progress in upgrading the machine to LEP-200. A major unknown is whether a programme of physics with polarized beams will become part of the plan for the exploitation of LEP. To give an indication of the LEP computing load beyond "1991" it has been assumed that the installation of LEP-200 cavities starts in 1992 and continues beyond 1995. It has also been assumed that energy scans above the Z^0 are interspersed with running on the Z^0 . Even if polarized beams are not available, additional Z^0 running is likely to be desirable to take full advantages of detector improvements made during the initial running; in some cases major detector components are scheduled for installation as late as 1991.

If LEP were to stop suddenly at the end of 1991, physics analysis would continue, initially needing most of the resources required during running, and reducing to minimal needs by about 1995. In the more likely scenario outlined above, the fall in Z^0 -related needs would be more gradual, and is very approximately estimated in Figure 1 on page 100.

The data rate generated by LEP at 200 GeV will be much lower than that generated by running on the Z^0 . The events of prime physics interest, e^+e^- annihilations, will occur at a rate of about 10,000 per year per experiment. However, two-photon and beam gas rates will be similar to those on the Z^0 , and any hint of a new particle signature in the annihilation events will require a careful analysis of many two-photon and beam-gas events to gain confidence that the new signature is not just part of the tail of the missing energy distribution of the background events. Thus although the computing needs of LEP physics above the Z^0 will thus be relatively low, they will be certainly above the 1% of the MUSCLE values which the event rate alone might suggest. An indication of plausible 'above Z^0 ' needs is shown in Figure 1 on page 100.

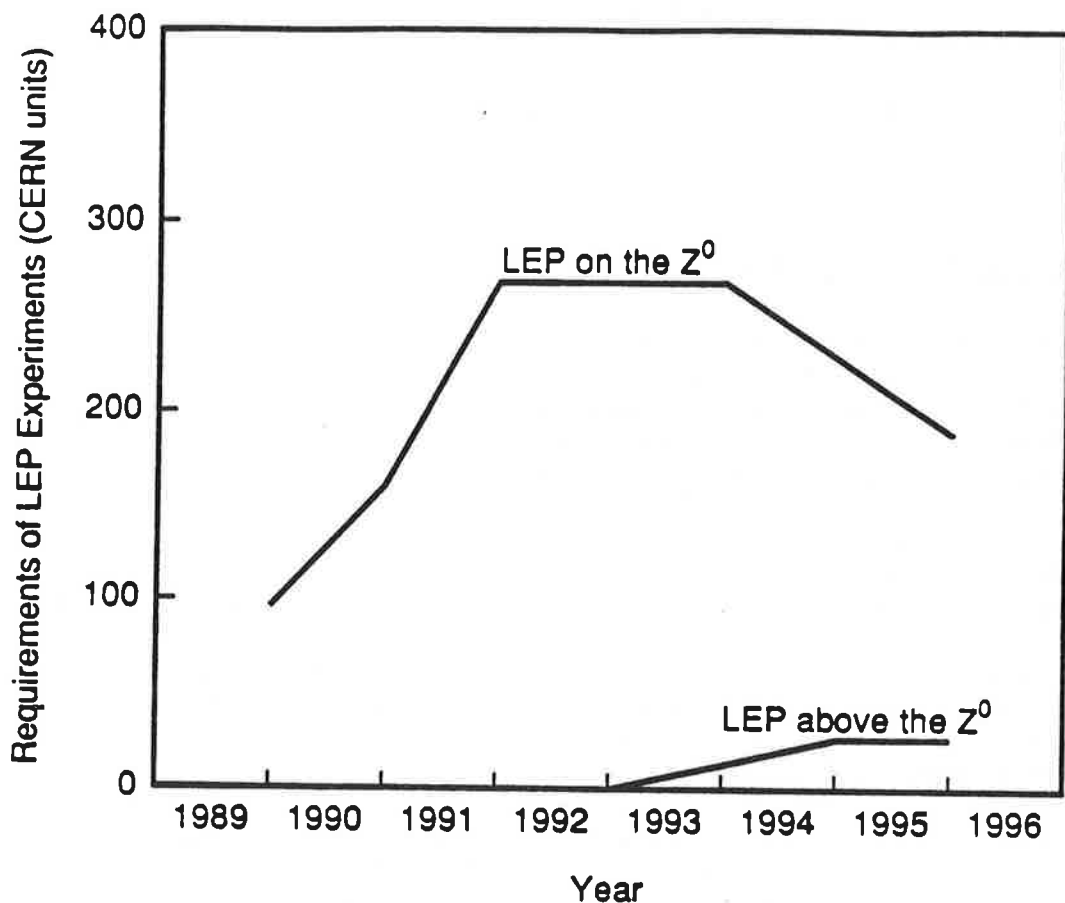


Figure 1: Total CPU Needs of the LEP Experiments

8.2 SPS Collider and LHC Experiments

The CERN SPS Collider programme is expected to remain active well into the 1990s. The associated computing requirements may be estimated by extrapolation from present experience, taking account of machine and detector upgrades and of the trends of the physics programme. This extrapolation also gives an indication of the likely scenario in the LHC era.

It is a fact that hadron collider experiments are limited by the available on-line and off-line computing resources. The on-line limitation comes from a dramatic mismatch between the total interaction rate (50K events per second at a luminosity of $10^{30} \text{cm}^{-2} \text{sec}^{-1}$) and the rate at which events can be written onto tape (typically five 200 kilobyte events per second with present technology). The goal of the data acquisition is to ensure that only "interesting" events are recorded. This is partly achieved by triggering on the presence of e.g. a high p_T lepton, missing energy and/or jets. Such triggers alone are adequate for W events, but not for more complicated topologies e.g. heavy flavor events where increasingly sophisticated higher order triggers, basically on-line processors, are used to reduce the rate

in successive steps. At expected ACOL luminosities ($3 \times 10^{30} \text{ cm}^{-2} \text{ sec}^{-1}$), UA1 will use at least 12 units of on-line computer power in a parallel processing configuration (3081E emulators) as the final trigger level.

On-line event rejection based on complex algorithms and uncertain calibrations is of course risky. Given faster data acquisition and sufficient off-line power, a safer procedure would be to write more events and be more selective off-line. For high cross-section processes this would also allow more useful events to be collected; for example, at a luminosity of 3×10^{30} , the rate of $b\bar{b}$ events is about 30 events per second (compared to one event every 5 minutes for W production)!

For off-line processing, the ultimate goal is to provide sufficient power to keep pace with the saturation capacity of the data acquisition. We assume that 50% of all events will be passed through something approaching full reconstruction; given detector and triggering improvements, events will either be usable for physics or will come from a background source which cannot be easily rejected at an early level. With the present data acquisition technology, this means that 13 million events could be written in an FME (Full Month Equivalent = one month of running at 100% efficiency); these events would occupy 8500 cassettes of the currently available 0.2 Gigabyte capacity. Taking UA1 as an example we can anticipate processing times of the order 30 seconds per event (standard CERN accounting seconds), and therefore about 7 CERN units would be needed to process the full FME (one CERN unit assumed to be 7000 hours/year). In a calendar year, a single experiment may expect to get two FMEs from four months of running time. Furthermore the first pass through the data should match the data acquisition speed. The requirement is therefore about 15 units, preferably in the form of a dedicated 45 unit processing facility available over four months.

We can expect that further data analysis and related Monte Carlo running will require at least twice the time for the first data pass, i.e. about 30 units. The Monte Carlo running can conveniently be run on dedicated facilities, while the remainder, say 15 units, should be in the form of general services.

CPU power is not the only parameter in the present scenario. The logistics of handling large volumes of data are already apparent today and are likely to become the limiting factor if denser storage media are not introduced, both in experiments and in computer centres. The above scenario implies 17,000 raw data cassettes per year, giving at least 34,000 total to cover Monte Carlo outputs, DSTs etc, again for a single experiment. An increase in density of at least a factor ten is highly desirable to ensure that the CPU power is accompanied by the appropriate ease of use.

UA1 has undergone a major upgrade to the muon detectors and to the data acquisition system. In the near future a sophisticated uranium-TMP calorimeter will be installed. The UA1 computing load can therefore be expected to rise from about 30 units in 1990 to about 60 units by 1992. No further upgrade is foreseen for the UA2 detector, and the throughput will be limited by the present data acquisition capacity. The UA2 requirements are therefore not expected to exceed 10 units over the period 1990 to 1992.

Qualitatively we may expect that the requirements of an LHC experiment will be similar to those discussed here for the SPS Collider. The on-line triggering will require a significant upgrade in order to operate in a many bunch environment with as little as 5 nanoseconds inter-bunch gap. The LHC will depend on high luminosities in order to reach new discovery thresholds, and $10^{33} \text{ cm}^{-2} \text{ sec}^{-1}$ is now discussed as "standard" with $5 \times 10^{34} \text{ cm}^{-2} \text{ sec}^{-1}$ in a special high luminosity intersection. The latter figure implies about 10^{10} interactions per second with, in the mean, 25 overlapping interactions per "event". Initially it is likely that highly selective triggers will be used, for example four muons from (hopefully!) the decay $H \rightarrow Z^0 Z^0$, producing a relatively low rate of especially clean events from low cross-section processes. As with the SPS Collider this phase may then be followed by studies of more complex topologies and higher cross-section processes.

8.3 Other CERN Experiments

There have been no systematic studies, comparable to the MUSCLE Report, of the needs of the other CERN experiments. A simple working hypothesis, that non-LEP computing needs would be similar to the total LEP computing needs, has been in use at CERN to make approximate predictions of the proportion of CERN's resources which would become available to LEP experiments.

It is probable that if CERN continues to support programmes of non-LEP, non-Collider physics, the experiments will also generate more and more data requiring more and more CPU cycles. More sophisticated electronics and computing techniques will be employed to increase the potential of detector hardware in an evolution typified by the pattern-unit to ADC/TDC to flash-ADC progression of recent years. This evolution is stimulated by the continuing increases in the cost-effectiveness of data-acquisition electronics. Since the same technologies also improve the cost-effectiveness of off-line computing, the fractional cost of computing (with respect to a complete experiment) need only increase very slowly, if at all.

8.3.1 LEAR Experiments

More than 25 LEAR experiments have now been approved involving over 500 physicists. The largest collaboration is CPLEAR with about 90 physicists. CPLEAR has made very preliminary estimates of their data volumes and CPU needs. They expect to take 2.5×10^8 good events in three years. With an event size of 2 kilobytes and guessing an equal volume of background, CPLEAR will accumulate 10,000 raw data tapes. The master DST volume is expected to be comparable. The necessary processing power is estimated to be 9 unit-years, assuming only one pass through the data.

Extrapolation from these figures to the total LEAR need is not easy, but it is probably safe to say that the total LEAR data handling requirement will be similar to that of a LEP experiment, while the CPU requirement may be somewhat smaller than that for a LEP experiment.

8.3.2 SPS Fixed-Target Experiments

There are currently more than ten SPS experiments taking data or in preparation. As in the LEAR programme several hundred physicists are involved. Although the long term future of SPS fixed-target physics is unclear, strong physics interest remains in making high statistics measurements to yield, for example, nucleon structure functions or the parameters of CP violation. Some collaborations, for example NMC, are preparing proposals for new experiments to last beyond 1992.

Typical current use of computing power by an SPS experiment is in the range 1 to 6 units, with some experiments expressing the view that they could make profitable use of much greater computing resources. Data volumes of several thousand tapes per year are also common. For example NMC wrote about 11,000 tapes in the last year, of which one third were raw data.

8.3.3 Summary of 'Other Experiments'

In addition to LEAR and the SPS programme, the SC programme and the many detector development and calibration activities cannot be ignored. As an example of the latter, the calibration of the L3 electromagnetic calorimeter has consumed 2 units for most of the last six months.

A further indication of the current needs of the non-LEP, non-Collider physics can be obtained by observing the recent use of the CERNVM service. The 'other experiments' now consume much more than half the total service. Most of this consumption is achieved in the face of considerable turn-around penalties (for exceeding the target allocation), showing that the physicists concerned consider that they have a real computing need.

In the absence of a systematic study, there can be no attempt to use the examples above to obtain quantitative estimates for the future. However, it seems likely that the 'other experiments' will not require less resources than one LEP experiment. At the other extreme, it seems very unlikely that their need would be greater than that of three LEP experiments.

8.4 HERA Experiments

HERA Experiments do not form an explicit part of the present study. However, it has been considered valuable to include a summary of HERA needs for comparison with the CERN estimates.

The data length for raw events is of the order of 100 kilobytes and for master DST-events 20 kilobytes. We assume on average 1 event/second and 0.5 years of Hera operation per year. This gives 1.5×10^7 raw events/year/experiment and 1.5×10^6 master DST-events/year/experiment with $Q^2 > 3 \text{ GeV}^2$. The total amount of data per year and experiment is 1500 Gigabytes for raw events and 30 Gigabytes for master DST-events. The computer time for reconstruction is assumed to be 20 seconds (on IBM 370/168) giving $1.5 \times 10^7 \times 20$ computer seconds per year or $30 \times 10^7 = 12$ Units/experiment (290 days/year = 2.5×10^7 sec).

For Monte Carlo event generation we assume to produce one simulated event for each master DST-event = 1.5×10^6 Monte Carlo events/year/experiment. 30% of Monte Carlo events have a full length of 100 kilobyte/event and 70% have a DST length of 20 kilobytes/event. For data storage one requires on average 44 kilobytes/event or $1.5 \times 10^6 \times 44$ kilobytes = 66 Gigabytes/year/experiment. The Computing time ranges from 60 to 900 seconds for 50% of generated events, varying from small to high Q^2 and from 5 to 10 seconds for 50% fast Monte Carlo events. This gives in total 1.5×10^6 events \times 60 seconds = 9×10^7 computer seconds per year or 3.6 units. 0.4 units are needed for reconstruction of Monte Carlo events. Assuming that Monte-Carlo generation and testing is performed 2.5 times, this would add up to 10 units.

To access DSTs and to extract physics we assume a CPU need for one hour/day/physicist. For 200 physicists/experiment this would give 200 hours/day or 10 Units (= 200 hours/24 hours/0.85 efficiency).

In Summary: Each Hera experiment needs 36 Units CPU power and 1600 Gigabytes of data storage. At Desy it is planned to install 15 Units CPU Power per Hera experiment plus 15 Units for infrastructure.

The needs for Hera experiments in terms of CPU and data storage are lower than those for LEP experiments. This comes mainly from two facts:

1. Both Hera experiments will use an on-line computer farm of about 20 units equivalent to select good events and to reduce the required storage. The experiments are connected directly to the DESY computer centre for data storage. Each link will have a bandwidth of only 1.5 Megabytes/second.
2. The physics event rate for $Q^2 > 3 \text{ GeV}^2$ is lower than the rate at LEP while taking data on the Z^0 . The rate at Hera is of the order of one event every ten seconds for high Q^2 .

Chapter 9

THE COMPUTING ENVIRONMENT FOR THE CERN PHYSICS PROGRAMME

9.1 Location of Physicists and Computing Facilities

We assume that in the future, as is the case now, HEP experiments will be undertaken by teams, usually comprising a large group of university physicists with smaller contributions from national laboratory and CERN physicists. Although this report does not concern itself on how HEP experiments should be done, the balance of the intellectual, technical and financial contributions of universities, regional Laboratories and CERN does have an impact on the overall organization of computing.

This mode of operating, with large and often geographically scattered, collaborations may not necessarily be the most efficient way to conduct any single experiment. However, this division of effort has great benefits; it ensures the continued inflow of talented young physicists throughout Europe and has the additional value of encouraging independent thought and lines of communication which may be stifled or discouraged if most resources are directly controlled by a single management.

Many of today's detectors, including the most complex, are designed, constructed and the data analyzed by university and national laboratory groups at their home institutes. However, several forces conspire to drag all but those with heavy teaching loads or little travel money to CERN. These forces should be minimized so that the physics largely emerges in the institutes rather than centrally at CERN. The forces are:

1. The shift rota for a complex experiment normally involves full-time-equivalents equal to about 20% of the nominal roll of a collaboration.
2. It is important for every physicist, and especially young physicists, to understand the experiment as a whole, which requires substantial participation in the work at the experimental site.
3. Personal contact and discussion has not yet been rendered unnecessary by electronic mail.
4. Access to computing facilities and access to the experimental data, is usually much easier for physicists at the experimental site than for physicists at their home institutes.

Forces (1) and (2) are not the concern of this report. Forces (3) and (4) should be neutralized as far as possible by improving electronic interpersonal communications, and by improving access to computing facilities and data from home institutes.

9.1.1 Off-site and On-site Computing

In the last ten years about half of the data processing and analysis of CERN experiments has been carried out at regional centres and at universities. An incentive to "export" physics analysis to the collaborating institutes was provided by the "one third/two thirds" rule under which CERN experiments

were expected to find two thirds of their computing needs outside CERN. In practice, this has not often been achieved, the average being closer to one half.

The 2:1 ratio was evaluated in terms of standardized CPU units. When the majority of computing was done on mainframes, these units were a reasonable basis of comparison. However, now mainframes, specialized processors, mass storage, large numbers of workstations and networking are all important and expensive components of the HEP computing environment. It therefore no longer makes sense to express support of physics analysis in terms of a simple CPU based rule. Nevertheless, considering that 1/2 to 2/3 of the physicists actively working on experiments do not need to be at CERN, and that somewhat greater resources are needed to support geographically dispersed physicists, it is reasonable to continue to aim to provide about 2/3 of the total computing resources outside of CERN. Every effort should be made to provide non-CERN based physicists with similar access to computing facilities and the experimental data to that available on the CERN site.

9.1.2 The Role of National/Regional Centres

National/Regional computing centres should play an important part in providing for the total computing requirements of the CERN community. This is important for many reasons:

1. Centres of computing and networking excellence are clearly beneficial to the region concerned. Supporting such centres brings much greater local benefits than exporting resources to CERN.
2. Regional communities will have easier access to data, particularly if lines are clogged. A distributed system is almost certainly less likely to have serious bottlenecks than a central one.
3. Distributed control of significant computing resources will give more flexibility to the user community, particularly those not resident at CERN.
4. There are economies of scale in centres offering facilities to many disciplines. HEP's ability to make effective use of any spare capacity is often welcome.
5. Regional centres can provide support for remote universities, including networking between the universities and the centre, and from the centre to CERN.

In the past these centres have played an important role in the overall HEP computing scene in Europe and there is every reason to expect this role to continue into the future. For this to be efficiently carried out good networking between the regional centres and both CERN and the local university community is a necessity, as is a mechanism for the exchange of bulk data between CERN and the centres.

A full exploitation of the technically possible 'European HEP analysis environment' with increasingly automated exchange of data between CERN, regional centres, and universities, will require an unprecedented level of cooperation between all parties, in which the role of regional centres will be particularly important.

More quantitative recommendations on the evolution of regional centres are given in section 10.2.2 on page 121.

9.1.3 Private Facilities at CERN

Private facilities are considered to be systems owned by a single group, or by a single experiment, which are large enough to provide a significant fraction of an experiment's computing needs.

It is recommended that the establishment of private general purpose computer centres at CERN be discouraged. If the facilities of the CERN computer centre are adequate, and the policy for their future development is firmly founded on the concept of allocating HEP's resources to optimize physics output, it makes no sense for individual experiments to set up and run general purpose computing centres at CERN.

The case of more specialized facilities is less clear. This report is not concerned with on-line event filters, but the processor farms which perform the filtering may also make excellent facilities for some of an experiment's production processing. Purely off-line private processor farms are being used effectively by several experiments, and it is expected that such facilities will continue to play a valuable role both on and off the CERN site.

9.1.4 Data Flow in HEP Experiments of the 1990s

To establish a reasonable working model, it is necessary to consider both the technology and the requirements. In this respect data volumes play a large role in deciding where CPU operations should take place.

Each LEP experiment will generate around 10,000 (200 Megabyte) raw data tapes in a year of running on the Z^0 . Volumes of simulated data are likely to be comparable. Planned volumes of the 'Master DST' (reconstruction program output) range from about 20% of the raw data to more than 100% of the raw data. If the Master DST is small, it is assumed that any detailed study of an event will require access to the raw data tapes. The high volume Master DST's will contain all the information likely to be needed for detailed analysis. In either case, physicists have a potential need for random access to data stored on tens of thousands of tapes. Clearly, the ability to handle such volumes of data is a resource which is at least as valuable as CPU power, and which must be regulated and allocated with at least as much care.

To reduce these data to the volume which can be handled by individual physicists and workstations (a few tapes at most), will require a hierarchy of DST's with progressively more stringent selections on either which events are present, or what data are present per event. It is extremely likely that substantial processing will be performed during the selection processes.

Given these large data volumes, how can off-site processing be maximized without introducing delays and difficulties in physics analysis? A possible approach is:

1. All data should be subject to thorough 'quality control' at CERN within hours of acquisition. Many experiments consider that the only adequate test is to run their reconstruction program, accepting that at least one re-run is inevitable after final calibrations have been obtained.
2. Similarly, during the early stages of running, during energy scans, or after any changes making new physics potentially accessible, very rapid physics analysis is essential, and a complete reconstruction must be done immediately at CERN.
3. Assuming (1) and (2) above, all the remaining 'production processing' of raw data can take place away from CERN. Transport delays for tapes will be small compared with the time required to get final calibrations and stable reconstruction code. Few collaborations would

be prepared to send off their only copy of the data, so the cost of media (SF 130,000 for 10,000 cartridges) will help enforce careful management of off-site production. Networks will be essential for maintaining coherence between software and database information on and off-site, but freight would be adequate for bulk data transfer.

4. Monte Carlo production can normally be run at any location with adequate network access and substantial computing capacity. An exception would be urgent Monte Carlo studies needed to understand time-critical data.
5. Wide distribution of physics analysis must be a prime objective. However, there clearly are problems associated with everyone having access to everything – remembering that "Master" DST's comprise 10,000 cartridges. In reality access to all data on 10,000 cartridges is going to be slow and expensive wherever a physicist is.

In order to try to overcome this problem, collaborations are likely to produce a hierarchy of DST's; large Master DST's, subset DST's, mini DST's as well as special physics topic DST's. How many copies of the Master DST will be required and where to locate them will clearly be experiment dependent. At a cost of 130 kSF the number is likely to be quite small. The smaller DST's will then be created and be transmitted either by network or freight to remote centres. The discussion above shows how the large data volumes in today's experimental HEP bring many problems. Data management will become an increasingly important activity for European physicists. Exploratory work by CERN and regional centres on distributed data management should be started.

Physicists are accustomed to the difficulties of exchanging data between computers with different data representations. While these difficulties are never insurmountable, they threaten to reduce seriously the transparency of a distributed data management system. The HEP community should treat an IEEE data representation as a significant advantage when considering future computer purchases.

9.2 Elements of Computing for CERN Experiments

9.2.1 Total CPU

Figure 2 on page 109 summarizes the evolution of the total batch CPU needs of CERN experiments up to 1992. It has been assumed that in providing the UA1 need of 45 units during the four months of data taking it is possible to make use of 50% of this capacity for other purposes when UA1 are not taking data. All the remaining non-LEP experiments, including other work at the SPS collider have been assumed to be equivalent to approximately two LEP experiments, which is an assumption allowed by the uncertainties described above.

It has been assumed that the components of the MUSCLE CPU estimates from Table 1 can simply be added. Any spare capacity outside data-taking periods will be more than offset by code development and calibration activities which were explicitly neglected by MUSCLE.

Figure 2 also shows an estimate of the CPU cycles used for networking, trivial interactive response, and various other system services. It is assumed that this can be kept down to 15% of the total needs. The estimated CPU needs of physics experiments invariably exclude such overheads, which must nevertheless be included when planning the size of systems to be installed.

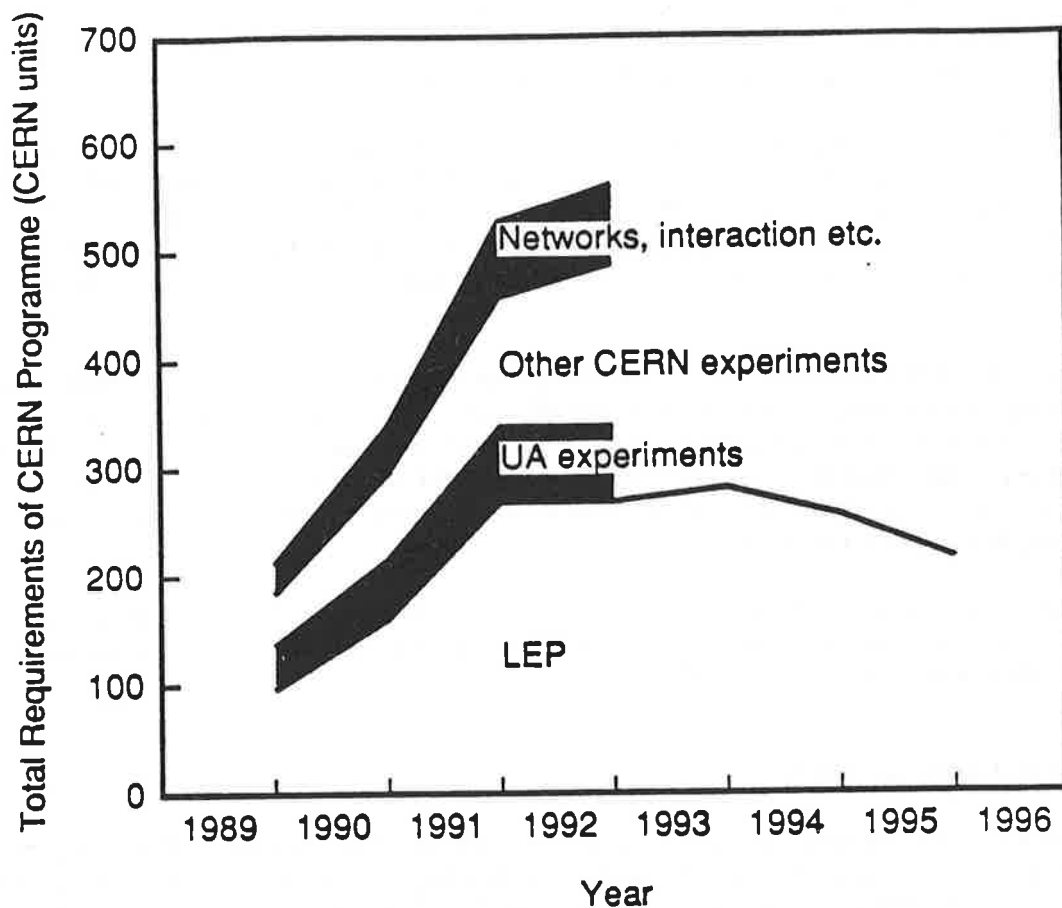


Figure 2: Total Batch CPU Needs for the CERN Physics Programme

9.2.2 Type of CPU

9.2.2.1 Mainframes used as scalar processors

This category includes such machines as IBM, IBM-compatible, Cray (used on unmodified code), and the larger VAXes. There is no doubt that almost all computing for HEP can be performed on such mainframes, and, as explained later, the differences between the (list)price/performance of the various machines is small. The large fraction of HEP computing which involves high volume data handling, linked to substantial CPU needs, is ideally suited for conventional mainframes.

9.2.2.2 Mainframes used as vector processors

On most modern mainframes vector processing hardware is either built in, or a relatively cheap optional extra. Thus only benefits can accrue from allowing the compiler and the vector hardware to do whatever they can to speed up unmodified HEP code. However, a substantial body of experience shows that the resultant increases in speed rarely exceed 10 to 20%.

Although it is too early to present a final verdict on the success of 'vectorization' through code modifications, it is now widely accepted that it would be a mistake to expect any substantial speed up for reconstruction and physics analysis code. By its nature this code has to be flexible, and therefore unstable. Vectorization of this code to achieve overall speed improvements of a factor two may well be possible, but at a cost in manpower and code maintainability which most physicists would consider unacceptable.

There is considerably more optimism about the prospects for speeding up Monte Carlo simulation through algorithm re-design rather than code modification. Work is now in progress to produce a vectorized version of GEANT3 although it is too early to make predictions of the benefits. It has already been demonstrated³ that the stable, experiment independent, components of a shower Monte Carlo can be re-written very effectively for a particular vector machine, but up to now, no complete Monte Carlo program has been vectorized.

Vectorization of Monte Carlo code may make it possible to generate simulated data at the level of detail, and with the statistical precision, merited by current and future detectors. Continued efforts in this area should be strongly encouraged.

9.2.2.3 Parallel computing

Most HEP computing exhibits inherent event-level parallelism. This parallelism offers opportunities to control the cost and expand the flexibility of HEP computing. To define the role of parallel computing for HEP in the 1990s, it is necessary to examine both the fraction of HEP computing which can be parallelized, and to specify which of many possible parallel architectures should be used.

Most experiments estimate that 40 to 70% of their minimum needs could be provided by parallel computing systems. Compared with the use of conventional mainframes, significant but tolerable inconvenience is expected, which is partially offset by the hope that cheap parallel systems could make it possible to exceed the minimum CPU requirements, with corresponding benefits for physics.

Exploiting HEP parallelism by allowing one experiment to run a job using all the CPUs of an IBM 3090 brings little or no benefit.⁴ Benefits arise when the processors used in parallel are much cheaper per unit of CPU power than mainframes. Price/performance can be improved by, for example, using mass-produced CPU chips, or by restricting I/O and inter-processor communications possibilities.

Processor 'farms' have been used in HEP for some years. Physicists who have replicated existing designs have found them cost-effective solutions for stable 'production' programs involving limited I/O. Whether the various emulator and 'ACP' programmes have been cost-effective as a whole is more difficult to assess. A major justification for the development of such processors has been their use as event filters in data-acquisition systems. Thus many would argue that the development costs should

³ e.g. K. Miura (Fujitsu USA), Lectures at the CERN Computing School, Oxford, 1988.

⁴ Some large programs may be required to use the CRAY processors in parallel because they might leave insufficient memory for other programs to exploit the remaining processors. This form of parallelism is not considered further in this report.

be ignored when evaluating the processors for off-line use. Whatever the true costs of their development, these processors certainly allow a cost-effective expansion of CPU resources for any experiment in which they are already heavily used.

The optimum future parallel architectures for HEP are not immediately obvious and this working group believes that the CERN (or even HEP) community should initiate a continuing and coordinated parallel computing project. The project should be rooted in reality, by requiring that it provide almost immediately a parallel computing service to users, while exploring the most promising approaches for the future.

The most promising approach for early investigation appears to be that of using commercially available processors adequately linked to the data handling capabilities of a centre's mainframe system. The major challenge would then be to make the parallel processing system as almost as easy to use as a mainframe of the same total power. Other approaches, such as building special processor boards, or buying complete parallel computers, should also be investigated.

9.2.3 The Move towards Personal Workstations

Central time-sharing services are still a cost-effective solution to the provision of general-purpose interactive computing for an organization of CERN's size,⁵ especially when measured in terms of cost per individual user of the service. Measured in terms of cost per active user at peak times, however, personal workstations of the VAXstation 2000 or Apollo DN3000 class are very competitive, and in addition provide a quality of service to the end-user which is incomparably better than that of a time-sharing service: human interface with bit-mapped screen, windows, pop-up menus and mouse; guaranteed availability — no delays, queues, interruptions outside the control of the user; personal choice of important details of the system. The evolution towards personal workstations as a vehicle for general-purpose interactive computing started early in 1987, following a substantial improvement in price/performance over most of the market, and is now well under way.

However, there are a number of points which must be considered.

1. *PERSONAL WORKSTATIONS CANNOT BE SHARED*

Personal workstations are personal — there must be enough of them to ensure that a convenient one is available when required. The result will inevitably be that these devices will often lie idle, but, as the above cost comparison shows, we should not worry about this if the workstation is used regularly at peak periods. If some batch work can be run on workstation clusters this should be treated as an added bonus.

⁵ A paper of A.Osborne and L.M.Robertson, prepared in February 1987, calculated the average cost of central computing, including all overheads (full costs of staff, management, energy, buildings, etc.), as 137 SF per standard CERN CPU hour. This is approximately double the cost calculated by taking into account only the cost of IBM hardware. Thus the cost of the current VM service (supporting 450 simultaneous users at peak times, and 2,300 different users each week, consuming a total of 800 CPU hours) might be estimated as: 110 kSF per week; 250 SF per peak user per week; 50 SF per user per week.

The cost of the basic hardware of a simple workstation today, including a minimal contribution to the cost of a "boot-host" or disk server, is 15 kSF plus 25% per annum maintenance. If we assume a three year lifetime for the station, this represents only 170 SF per week per workstation. The hidden costs for workstations are almost certainly higher than those of the Computer Centre, where economies of scale are possible, especially if manpower-intensive functions like disk backup are included. No realistic assessment of these costs has been carried out (to our knowledge), but it is clear that the cost per peak user is very similar for both types of service.

2. *USER PRODUCTIVITY*

It is clear that the end-user will be much more productive when using his personal workstation than he would be using a central interactive service. However, workstations are real computers, and require to be operated and managed: hardware installation and maintenance; software support; file-system maintenance; network monitoring; etc. It is important that the user does not exchange his new-found productivity for a part-time job as systems programmer cum operator cum network manager. To obtain the full benefits of the workstation a level of central service, staffed by specialists, equivalent to that provided for the central time-sharing services must be made available.

3. *INTEGRATION WITH CENTRAL SERVICES*

Workstations must be very well integrated with central services and with each other. While enjoying the advantages of his dedicated computer system, the user still requires to share data with his collaborators and to have access to central data storage facilities, batch services, printing services, networks, etc. Integration of workstations with the central IBM service has improved substantially at CERN during the past twelve months, but there remain many deficiencies in terms of both functionality and performance.

4. *NETWORKING*

The high bandwidth required between personal workstations and their file servers will have a significant impact on the general networking infrastructure of CERN and its community. While this in turn will necessarily place constraints on the topology of workstation clusters and services, it is essential that these are minimized, and the potential needs of workstations taken account of in the evolution of the network infrastructure, even at the expense of the installation of substantial over-capacity.

5. *THE FAST-MOVING WORKSTATION MARKET*

The technology of workstations continues to advance at breakneck speed, and shows little sign of slowing down. Apart from the problem of rapid obsolescence, this also poses serious support problems. At present CERN has effectively standardized on VAXstation/VMS and Apollo systems, in order to simplify support. This standardization has been the result of a judicious choice of products and the provision of (albeit limited) central support. While there are no immediate reasons to question the choice of these two systems, we must ensure that we benefit fully from developments in the market, and the selection of systems supported should be regularly reviewed. The de facto standardization of almost all suppliers on the use of Unix, ASCII and IEEE floating point (with DEC a notable exception) should limit to some extent the problems of support.

6. *LOW-END WORKSTATIONS/PC-S*

This study has assumed a definition of a workstation which includes a certain minimum CPU power, memory size, graphics screen and mouse, but also the provision of a conventional operating system with (Fortran) program development environment. It is the last factor which most clearly distinguishes personal *computers* from personal *workstations*. However, with the advent of machines like the Macintosh II this distinction is becoming very blurred. It is clear that the Mac II is an extremely attractive system, which is perfectly capable of running large Fortran programs, and has good networking facilities in addition to its excellent human interface. It will certainly be used in substantial numbers at CERN as a low-end general-purpose workstation.

9.2.4 Data Handling

The MUSCLE report made quantitative estimates of LEP data volumes, and assessed some of the benefits conveyed by keeping some of the LEP DST data on disk storage. The duplication of data needed for distribution to universities and regional centres was not included. Wider discussion of the MUSCLE data volumes has tended to show them as lower limits, particularly with respect to the 'Master DST's' which may be bigger than the MUSCLE estimates by large factors.

Noting that many non-LEP, non-collider, experiments have or expect raw data rate of thousands of tapes per year, it is reasonable predict total CERN data volumes of 8 to 10 times the MUSCLE single experiment estimate. This implies at least 300,000 active tape cartridges in 1992 assuming use of 200 Megabyte cartridges. In the absence of fibre optic networking, distribution of data to universities and regional centres will require at least a similar number of cartridges. It is reasonable to assume that raw data are copied once at most, but master DST's will have to be copied many times if off-site physicists are to have access to more than a fraction of their data.

The MUSCLE report set a target of 100 Gigabytes of disk space at CERN per LEP experiment by end 1991. This represents 1.4% of the expected data volume, and would suppress a large fraction of the tape mounts, in addition to facilitating physics analysis. It was also urged that this disk capacity was complemented by automated cartridge handling. Such a system would probably have a capacity of 10% to 20% of the total data volume. As argued above, it is reasonable to assume that an optimized storage hierarchy at CERN would have about 8 to 10 times the capacity needed for a single LEP experiment. Together with system needs estimated at 20% of the total, this leads to a target of over 1000 Gigabytes disk storage, and automated handling of at least 40,000 cartridges by end 1991.

Automatic migration of relatively inactive files from disk to tape-cartridge storage would be of considerable benefit in managing experiments' DSTs. For physicists' "personal" files, which can easily and justifiably reach hundreds of Megabytes per person, such automatic migration is even more important.

9.2.5 Networking

The existing CERN site network is only marginally able to support data access from workstations, and does not support bulk data transfer between computers. Physics analysis at CERN would benefit from improvements in both areas and efforts are underway to achieve this.

There is no physics reason why access to data from universities and regional centres should need a lower performance than CERN-site networking. This runs contrary to conventional networking philosophy which assumes that traffic falls off rapidly with increasing distance. Although it is financially unreasonable to expect that wide area networking should always offer local area network performance, the effectiveness of off-site physics analysis will increase as this goal is approached.

A recent study⁶ of how LEP experiments would use 2 Megabit/second links showed that the links would be used by a wide variety of services (paralleling the wide variety of services offered on the CERN site). For many activities remote users would have access similar to that on the CERN site, although a truly integrated environment for the location-independent analysis of HEP data will require much higher speeds.

Although 2 Megabit/second links are not sufficient to provide totally transparent access to arbitrarily large data samples in real time, their effectiveness can be enhanced by having sufficiently large data storage facilities at the remote end of the link. Frequently accessed data can then be stored at the uni-

⁶ E-A. Knabbe (CERN-DD) "What would the LEP experiments do with 2 Megabit/second link(s)?", September 1988.

versity or regional centre. The management of these distributed data stores could, in principle, become totally automatic.

A substantially integrated environment, supporting analysis activities requiring access to Gigabytes of data, could be created by using networks based on optical fibres between universities and regional centres and between the centres and CERN. The current capacity of such links is in the range 34 to 144 Megabits/sec., and the same physical fibres will be able to support at 4 or 16 times these rates within the next few years. The true commercial cost of such fibre-optic networking would be low enough to make its early use by HEP very attractive. In reality, its early use in Europe would require a significant change in PTT tariff policies.

Chapter 10

EVOLUTION OF CENTRALLY ORGANIZED COMPUTING AND NETWORKING AT CERN AND AT REGIONAL CENTRES

10.1 Introduction

The next five years will be a challenging period in the development of computing for offline High Energy Physics, as we progress towards a distributed computing environment, exploiting the fast-developing technologies of workstations and networks, at the same time ensuring that the needs of the early years of LEP are met effectively.

The main elements in the evolution of centrally organized computing, at CERN, in national and regional centres, and in the universities, are:

- The growth of batch capacity for LEP data analysis, involving the expansion of conventional general-purpose facilities at CERN and in major centres, the large-scale exploitation of cheap parallel computers, and the integration of these with cartridge tape and disk storage facilities on a scale new to HEP.
- The establishment of personal workstations as the principal tools for general-purpose interactive computing. The technological and economic conditions are ripe for this development, but the challenge will be to develop the support services, the network management and the tools to integrate workstations with each other and with central services for file management, batch processing, etc.
- The development of distributed co-operative processing for interactive data analysis, with different components of a single program executing simultaneously on a powerful workstation and on one or more central CPU and file servers.
- The introduction of high performance wide-area networking, which will enable the above developments to evolve on a European-wide basis and allow the individual physicist to work effectively from his home institute.

It is important that this evolution take place smoothly. While we must press ahead with the introduction of new technologies for distributed computing, we must also learn how to exploit these effectively, and be careful not to abandon prematurely existing proven services.

10.2 Evolution of Batch Capacity

There are three main sources of "production" batch capacity available to high-energy physicists: the CERN central computing services; large national physics institutes or computing centres; smaller systems such as departmental VAXes, emulator farms, off-duty personal workstations, etc.

10.2.1 CERN Computer Centre Batch Facilities

By the end of 1988, the CERN Computer Centre will provide two batch services, based on the following equipment.

IBM 3090-600E/VF – 39 CERN units (scalar);

Siemens 7890S – 12 units (no vector capability);

Cray X-MP/48 – 32 units (scalar).

These figures are the nominal capacities of the computers. In practice, a significant and growing proportion of the IBM service will be used for interactive computing, which consumes about 20% of the IBM-compatible capacity delivered at present. An additional 10% of the CPU cycles are consumed by general services, such as network file servers, mail handling, and printing. This percentage will increase significantly with the development of distributed computing. The requirement to increase the general-purpose batch capacity to over 150 CERN units by the end of 1992 may be met in a number of ways. This section discusses some of the general factors involved, and suggests a model which is reasonable today in order to enable costs to be estimated. In practice the route chosen will depend to a large extent on future product announcements and the results of financial negotiations with manufacturers.

10.2.1.1 Competitive and complementary services

Competitive batch services were first introduced into the Computer Centre at the end of 1976, when an IBM system was installed alongside the existing CDC 7600. The original motivations for installing a second manufacturer's equipment were to provide a wider service for users, allowing them to benefit from the different strengths of the different systems, and to provide greater possibilities for compatibility with services in outside institutes, avoiding placing the latter in a position where they were under pressure to buy from CERN's single supplier. These remain good reasons for continuing to operate complementary services today. However, the past twelve years have demonstrated that this competitive environment has additional advantageous effects: on the suppliers, in terms of quality of service as well as more imaginative commercial terms; on the staff providing the services; on the users, who are encouraged to avoid dependence on proprietary technology – software, operating systems and hardware. The scale of CERN central computing is sufficiently large for the operation of two services, and it is important for the flexibility and vitality of the central batch services that the competitive element remain.

10.2.1.2 Operating systems

Evolution of the capacity implies also evolution of the operating systems which can best provide it. There is clearly a strong incentive to provide a stable environment for the users, and operating system changes and the introduction of incompatible hardware must be kept to a strict minimum, while ensuring that obsolescence is avoided and a reasonable flexibility to exploit new hardware is maintained. These conflicting requirements are hard to reconcile. In preparation for LEP, the IBM service is being

moved to a new base on the VM/XA/SP operating system. This will ensure that the service will be able to exploit the hardware innovations which will appear during the next few years, particularly in terms of memory size and multi-processor configurations, at a minimal inconvenience to the user. Similarly, the Unicos operating system on the Cray will enable either of the Cray architectures (the X-MP/Y-MP and Cray-2/Cray-3 lines) to be used. There is therefore no need to make major changes in operating systems on these services during the next five years.

There is, however, considerable activity on the part of most computer manufacturers and in various standardization bodies in the area of Unix, which promises to provide a single standard user interface on all computer systems currently of interest to offline High Energy Physics (e.g. Cray, IBM 3090 and compatibles, DEC VAX, Gould, Apollo, IBM PC, Apple Macintosh). While there is ample scope for argument about the attractiveness of Unix as an operating system, we now have practical experience of good manufacturer-supported implementations (e.g. Unicos) which demonstrate that it has no fundamental flaws (and indeed has many advantages) in our environment. The benefits of a standard user interface are clear, but we must carefully weigh these against other factors like additional functionality, ease of use, and manufacturer commitment, before deciding to abandon excellent systems such as VMS and MAC OS. IBM's AIX system for the 3090 is particularly interesting in view of the company's declared commitment to it, and a number of deficiencies perceived in the current versions of alternative 3090 operating systems. However, the first release is not scheduled until 1989, and there are important features missing, such as tape and batch support. We should actively study the AIX system when it becomes available, and also investigate how we can best influence the various Unix standardization activities (POSIX, X/OPEN, the Open Software Foundation).

10.2.1.3 Cost comparison of Cray and IBM-style computing

The IBM-compatible and Cray batch services are very different in many respects, the most significant being that the Cray has a substantially higher potential vector performance, but has a limited disk configuration (due to the cost of the high-performance disks which alone are supported on Cray machines), and lacks virtual memory. Today, and probably during the five-year period being considered, the vector capability of these machines will not be of major significance in overall throughput, although individual programs may make dramatic gains, and in turn perform substantially better on the Cray. A brief analysis of comparative list price costs⁷ of IBM and Cray systems which will be available at the beginning of next year shows that these systems are rather competitively priced in terms of scalar computing capacity per Swiss franc. Thus, while Cray-style computing does not have to be financially justified in terms of vector usage, the inconvenience of restricted memory and disk capacity will limit the type of work for which these machines will be suitable.

⁷ The list price of a Cray Y-MP system with 8 processors, 32 Megawords of main memory and a 128 Megaword "solid state disk" is \$20M. 200 Gigabytes of disk space cost about \$10M. The total cost of a system with a capacity of about 96 CERN units is therefore \$30M, or about 500 kSF/unit (\$1 = 1.6 SF).

The list price, including standard academic discount at 20%, of an IBM 3090-600S/VF, which would deliver about 50 CERN units, together with 256 Megabytes of main memory, 256 Megabytes of extended store and 100 Gigabytes of disk space is about 27 MSF, or about 550 kSF/unit.

The list price for a four processor Amdahl 5990-1400, with 256 Megabytes of main memory, 512 Megabytes of extended memory, and 100 Gigabytes of disk space, but no vectors, and which would deliver about 50 CERN units is about 30 MSF. A discount would normally be expected.

10.2.1.4 Manpower implications of the Cray service

It is clear that the provision of two separate batch services using different equipment has direct manpower implications. These are most significant during the installation of the hardware and introduction of the service. However, the style of batch service offered on the Cray, aimed at a relatively small number of very large users, restricts the continuing support manpower considerably, particularly in comparison with the manpower required to run a general-purpose service like that offered on the IBM systems. About 5 staff are required to maintain the Cray service in addition to those which would be required to run a similarly restricted service provided by duplicating the hardware and software environment of the general-purpose IBM service. This may be considered as the cost of running complementary batch services. Although it is not possible to quantify the value of any single element of a financial negotiation, the competitive environment generated by the two-manufacturer policy for batch capacity has been an important factor in the achievement of keen commercial terms for computer hardware in recent years, largely, if not completely, off-setting the cost of the additional manpower involved.

10.2.1.5 The multi-purpose CERN VM service

The VM service operated on IBM and compatible computers by the CERN Computer Centre provides a number of quite different facilities, the most important of which are: major batch service; general-purpose interactive service; "front-end" support for the Cray; printing and text-processing facilities for many other computers and workstations, in particular the central VMS service; file server for workstations and small computers.

Operating this general-purpose service clearly requires much greater manpower than that needed for a straightforward batch service, due to the variety of facilities to be supported and the overall complexity of the system. It must also be noted that the ancillary services consume a non-negligible fraction of the CPU capacity of the configuration.

Over the next few years, as discussed in later sections, the VM service will play an increasingly important role as a network file server, which will require a substantial amount of high-priority CPU capacity. COCOTIME⁸ and the Computer Centre will find the task of scheduling resources increasingly difficult in this complex environment.

10.2.1.6 IBM-compatible systems

IBM-compatible systems have served the industry well, both in improving the price/performance of IBM-style computing but also in stimulating IBM to build powerful processors which perform well in the scientific and engineering markets. However, the competitive response of IBM has also included the provision of obscurely specified hardware and microcode enhancements which are essential for the efficient operation of the machine using IBM software. For example, the new VM/XA/SP operating system cannot operate on CERN's Siemens 7890S computer, installed less than four years ago. Great care must therefore be taken in assessing the overall financial benefits of IBM-compatible systems.

⁸ The CERN committee which allocates central computing resources to experiments.

10.2.1.7 DEC batch capacity in the CERN computer centre

The DEC services in the CERN Computer Centre have been set up in order to provide interactive computing facilities. However, due to the fact that an interactive service must cater for peak daytime loads, some 60% of the installed capacity is available for batch. At present this amounts to only about 4 units, including the capacity of the engineering service VAXes, but the recommendation below for the expansion of the interactive service could lead to somewhat greater capacity being available for batch. This will be of interest to small experiments which are thereby able to do all of their computing on VAX. The capacity available will not, however, exceed 10% of the total general-purpose batch capacity at CERN, and therefore does not justify the operation of a full-scale batch service.

DEC is not considered here as an alternative to IBM-compatible and Cray for the main central batch facilities, because of the large number of CPUs required to achieve a reasonable capacity, although the cost of an appropriately configured system is comparable with the others.⁹ While we can expect more powerful CPUs to be delivered by DEC over the next few years, the other manufacturers will also introduce new equipment, and it is not clear that the relative position will change significantly in the short term.

10.2.1.8 Model for evolution of the central batch services

1. REPLACEMENT OF THE SIEMENS 7890S

The first step is the replacement of the aging and obsolescent Siemens 7890S, which is not capable of running the latest version of the VM operating system, and which will provide less than 25% of the IBM-compatible capacity after the installation of the IBM 3090-600E system at the end of 1988. The Siemens should be replaced at the end of 1989, about six months after the startup of the LEP machine, by a system with at least 50 units of scalar capacity¹⁰ This could be purchased from IBM itself, or from a compatible supplier. Clearly, the chosen system must run the latest version of the VM operating system, with reasonable confidence that it will not become obsolete in this sense during its four-year lifetime. Possible candidates today would include a four processor Amdahl 5990-1400 or an IBM 3090-600S.

Assuming that a 40% discount over list price could be achieved, the cost of this operation will be around 18 MSF.

Note that the existing IBM computer, installed originally early in 1986, will itself have to be replaced by the end of the period being considered.

⁹ A cluster of four of DEC's largest systems, the four-processor VAX 8840 with 128 Megabytes of memory, together with 50 Gigabytes of disk space and 4 disk controllers, costs about 10.5 MSF with standard discount. This should deliver a *nominal* total of 24 units on *sixteen* processors. The multi-processing efficiency of the computers and our ability to schedule 16 processors efficiently for general-purpose batch are unknowns, but we might expect a degradation in throughput perhaps as high as 20%. The cost per CERN processing unit is therefore in the range of 440k SF to 550 kSF.

¹⁰ This will bring the total IBM-compatible capacity to 90 units, of which we assume 20% will be used for general-purpose interactive work.

2. *CRAY SYSTEM*

The Cray system was installed in the spring of 1988, and it is therefore too soon to assess fully the effectiveness of this style of service for the different types of processing required for LEP physics. It is clear that it will be better suited to work which is CPU-intensive, and that a very large disk capacity is excluded due to the high cost of the disk equipment. The I/O limitations are, however, sometimes over-emphasized – the system has excellent tape support and the 50 Gigabytes of disk space which should be available early next year will enable a good tape staging service to be provided to support most Monte Carlo and production programs.

Vector support will become available on almost all high-end processors over the next few years for relatively little additional cost. The ability of high-energy physics to benefit from such hardware would bring significant financial savings, even if the overall improvement in throughput is relatively low, due to the high total cost of the investment in such a large system. In order to stimulate the pioneers in vectorization a machine with an outstanding vector performance should be maintained beyond the life of the present Cray.

Subject to the satisfactory development of the Cray service, the current X-MP/48 should be replaced during 1991 by a system with at least 80 units of scalar capacity. The desire to minimize change will make the Y-MP system an obvious candidate (the eight processor model should deliver over 90 units). However, it may well be more attractive to install the Cray 3 system, if it is available at that time, providing a substantially better scalar performance with its GaAs technology, and a very large real memory. Clearly any other comparable supercomputers which may become available must be carefully considered, with due attention paid to the importance of the needs for a competitive environment and minimal change of operating system mentioned above. The system should include at least 100 Gigabytes of disk space, and we estimate a total cost of around 30 MSF, assuming a substantial discount.

3. *DISK CAPACITY*

The evolution of the disk capacity of the IBM-compatible service is covered in a later section on data storage.

Table 1: CERN Computing Centre – Evolution of Batch Capacity

service	end 88	end 89	end 90	end 91
IBM-compatible	36	60	57	54
Cray	32	32	32	90
DEC	6	10	10	10
Total Capacity	74	102	99	154

The nominal CPU capacity of the IBM service has been reduced by 30% in 1988, rising to 40% in 1991, to cover the part used by general-purpose interactive and distributed computing services.

10.2.2 Large Regional Centres

In order to make a major impact on LEP data analysis, regional computer facilities must include very large disk configurations. In turn, to make efficient use of this expensive investment and justify the personnel costs which will be necessary to manage the DST data and associated calibration files, substantial CPU capacity must also be available. Because of the practical management difficulties involved, an individual collaboration will be able to organize its large-scale computing at only a limited number of physical sites. We conclude that a centre which will make a significant contribution to several experiments must provide minimum resources of 100 Gigabytes of disk space, an automatic cartridge tape loader with a capacity of 5K tapes, and a CPU capacity of 20 units.¹¹ Increased CPU capacity would make increasingly effective use of the disk space, and, if adequate network connections to CERN are also available,¹² the centre could be a very attractive focal point for a substantial part of the analysis of an individual experiment.

We therefore suggest that the conventional batch computing capacity available to high energy physicists be concentrated in a number of large, adequately configured, regional or national centres, to provide a total processing capacity of at least 200 units.

10.2.3 Other Batch Capacity

10.2.3.1 Emulators and parallel processors

The experience with the 3081E and 370E emulators, and at Fermilab with the ACP, has already demonstrated that these systems can be used effectively in many cases where the program is relatively stable. In order to maximize the cost benefit of such systems, it is important that they are provided with an adequate level of operational support, including tape handling.

At the low-end of the CPU market, the cost of raw processor power is falling at an astonishing rate, which will continue to stimulate the development of cheap parallel processor systems. These range from straightforward assemblies of micro-computers which could be exploited in the same way as the existing emulator farms, to very specialized systems such as the Meiko Computing Surface.

A number of parallel processor "farms", established in universities, regional centres or in the CERN Computer Centre, well integrated with the institute's central computing services, to provide easy and cost-effective access to the main data storage facilities, could make a major contribution to the general computing capacity of HEP. Once established, these services would provide a framework which could later be very rapidly expanded in capacity, for relatively small additional cost.

We therefore make the following recommendations.

1. Commercially available parallel computers, or systems which can be assembled from commercially available processors, should be studied actively, as a coordinated effort between CERN and interested universities and regional computing centres. An immediate aim would be to identify commercial systems which could be used to provide parallel processing services in HEP computing centres. A longer term goal would be to do research in parallel computing relevant to high-energy physics needs.

¹¹ This represents an investment of about 13.5 MSF (100 Gigabytes cost 3 MSF; a one third share in a tape loader costs 0.5 MSF; 20 units of IBM CPU power costs about 10 MSF at academic list price).

¹² See a later section for discussion.

2. During 1989/90 parallel processing farms, providing a minimum of 20 units of capacity each, should be installed in several of the larger regional centres, and at CERN. These should be integrated with the central services of the computing centres, to provide operations and software support, and access to the central data storage facilities.
3. The pilot service at CERN should provide about 25 units during 1990, rising to 150 units by the end of 1992.

10.2.3.2 Private batch capacity

We note that substantial batch processing capacity will exist in the form of private computing facilities, owned and operated by individual collaborations. Much of this will be in the form of emulators and other parallel processors, often purchased for a specific task associated with data collection but capable of being used for more general offline analysis at other times. These facilities are outside the scope of this section of the report.

10.2.3.3 Batch capacity in university physics departments

The aggregate capacity of the small computers installed in university physics departments will grow substantially over the next few years. Some fraction of this will be available for data analysis, depending on the ability of the physics departments to supply data to the computers and of the collaborations to organize the distribution of work to the many small units. The main contribution of these is likely to be as tools for a department's own staff, thereby enabling them to perform more of their work while at home. See also the discussion in the later section on interactive data analysis.

10.2.3.4 Workstations as batch processors

The number of personal workstations installed at CERN will grow dramatically over the next few years, and may reach a total of 1000 within four years, representing a nominal CPU capacity of several hundred CERN units. Since the workstations will be under the control of the individual users it is unreasonable to imagine that any sensible fraction of this capacity could be used by a central service. However, individual experiments should be encouraged to take advantage of spare capacity on their workstations for suitable tasks such as Monte Carlo.

10.3 Evolution of General-purpose Interactive Facilities

General-purpose interactive facilities include a program-development environment, text-processing and document handling facilities, electronic mail and database services, together with good connections to the main local and wide-area networks, to ensure that the user is able to control effectively his activities on central services at CERN and in regional centres. To an increasing extent, interactive services and facilities will be used for data analysis which, because of the rather special requirements for CPU power, data storage and network bandwidth, is discussed separately in the next section.

The following discussion refers specifically to CERN, but it applies also, with appropriate scaling, to regional centres.

10.3.1 Workstations at CERN

The benefit to physics analysis of a rapidly expanding use of personal workstations has already been discussed. In order to ensure that the evolution towards a personal workstation for every physicist and engineer at CERN may proceed as efficiently as possible, we make the following recommendations.

1. Centralized support for VAX/VMS and Apollo workstations should be strengthened, with the ultimate aim of making the purchase, installation and management of a personal workstation as easy as using the central VMS service.
2. The bandwidth, reliability and facilities available between workstations and the central services should be improved substantially. Planning should aim to provide an instantaneous 30 kilobytes/second for file transfer between an individual workstation used for general-purpose work (see the next section for the needs of workstations used for physics analysis) and the central IBM service, adequate capacity being installed to meet the aggregate demand at peak times.¹³ The central VAX/VMS service will form an important focal point for users of VAXstations, and sufficient capacity should be provided to enable them to take advantage of this facility for file storage and other services.
3. The investment in the networking infrastructure of the laboratory should take account of the projected growth in the use of workstations (see discussion in the section on interactive data analysis).

10.3.2 The Future of Central Time-sharing Services

While it is clear that there will be a steady and rapid growth in the use of personal workstations, the life of the central time-sharing service is far from over. As we have noted above, workstations are today cost-effective only for intensive users. The demand for the central VM and VMS services continues to grow, and shows no sign of diminishing. We expect that the move of users from the central services to workstations over the next two years will be more than balanced by the increase in the number of new users of the central services, and that the demand at peak times will double during the same period, which will see the growth in the management information and engineering services, and the startup of LEP. Thereafter, we expect to see the rate of increase diminish, and demand eventually tail off.

The central interactive services are at present provided by the VM service (at the level of about 20% of the installed IBM-compatible capacity), and by the VMS service (which has been installed specifically for this purpose). The popularity of both of these services indicates that they are effective and complementary, and this and the advantages of a competitive environment enumerated above lead us to recommend that they should both be maintained and developed. The current load problems of the VM service will be largely resolved with the introduction of the VM/XA/SP operating system which will enable the service to expand to take advantage of additional processors and memory which are already installed. It is assumed that the VM interactive service will continue to use up to 20% of the installed IBM-compatible capacity, and that no specific recommendation for expansion is required.

The VMS service is, however, badly overloaded at present, and an effective increase of around 60% in its capacity is required merely in order to provide an acceptable service. This should be installed without delay. In addition, the capacity of the service should be doubled before the end of 1989. The total cost of the initial upgrade, including a modest 10 Gigabytes of disk space will be 1.4 MSF. The cost of the subsequent doubling in capacity, replacement of obsolete equipment and growth of the disk capacity to a total of 100 Gigabytes, is likely to cost about 7 MSF at standard prices, over

¹³ Note that this implies high-priority CPU capacity as well as cables and communications controllers.

two years.

The growing use of personal workstations implies a growing demand for centralized file and CPU server facilities (increased productivity will lead to increased consumption of resources as well as increased output). The VM and VMS central time-sharing services are the natural vehicles for such support, and as users move towards workstations they will in any case go through a phase where their working environment is split between the central service and the workstation. We can be sure that any new investments in the central VM and VMS services will be of long term value for workstation support, even if the time-sharing demand drops faster than we predict.

10.4 Interactive Data Analysis

The use of more or less powerful workstations for interactive data analysis gives rise to the following requirements, which are in addition to the general needs of workstations, discussed in the previous section. These requirements are stated in terms of CERN, but it is clear that they can be extrapolated to apply to the larger national and regional centres used by HEP.

1. *ACCESS TO PHYSICS DATA*

All current models for LEP data analysis assume that the master "data summary tape" (DST) will be stored on cartridge tapes managed by the central VM service, with successive levels of subsets (team and personal DSTs) stored on disks on the same system. In some cases it may be possible to pre-select a relatively small subset of the personal DST (say 5-10 Megabytes) and transfer that to the workstation for processing during a working session of a few hours. This mode of working can be accommodated within the relatively modest data rates required for basic workstation support. However, in cases where this pre-selection is not practical all of the external data consumed by the analysis program must come from the IBM file server. This gives rise to a requirement for very much greater data rates than can be achieved by any existing product. The minimum requirement is binary file access at an aggregate data rate of 500 kilobytes/second per cluster of 10 workstations of the VAXstation 2000 or Apollo DN3000 class, rising to 1 Megabyte/second by the end of 1989. The demand will continue to rise faster than the technology available.

2. *VERY HIGH PERFORMANCE NETWORKING*

The processing power of the workstations will continue to grow at a rate far outstripping developments in deliverable network bandwidth. It will be necessary to envisage the installation of the highest performance network facilities available, specifically for the support of the data analysis workstations. It is important that the development of distributed physics computing is not restricted by the considerations of a general-purpose network infrastructure. It is likely that the coming 100 Megabit/second FDDI (Fibre Distributed Data Interface) standard will be an appropriate choice in the medium term.

3. *CO-OPERATIVE PROCESSING*

In order to benefit from the different characteristics of central mainframes and workstations, and to minimize the effects of practical constraints such as insufficient network bandwidth, considerable investment should be made in the development of co-operative processing tech-

niques. Already the Physics Analysis Workstation (PAW)¹⁴ developments permit the separation of the user interface (on a workstation) and the Zebra data server (on a mainframe). Future plans envisage the extension of these techniques to provide "object servers" which may be distributed with some freedom between the workstation and specialized central services such as the Cray (for CPU-intensive tasks) or the IBM (for data storage manipulation). The effects of this style of working on the scheduling of the central resources may be profound, but the problems raised must be tackled with vigour both by COCOTIME and by the technical staff involved. Co-operative processing will become a major factor in the analysis of LEP data.

4. *DISTRIBUTED DST FILE SERVERS*

We have seen that the centralization of DST data poses a major problem for distributed computing. It may be necessary to consider the use of distributed file servers, in order to relieve the bottleneck of the central IBM file server. We must note, however, that the task of handling the DST data for LEP requires considerable CPU capacity¹⁵, and so a distributed file server will be itself a substantial system, resembling more a mainframe than a workstation. The decision to install distributed file servers is largely a financial one, and at present the economics favour IBM-compatible systems, particularly due to the price of disk storage. In the immediate future, therefore, we recommend that the task of providing a central file server facility should be the responsibility of the IBM/VM service.

It is important that collaborations include estimates of their needs for such facilities in their computing requests, in order to ensure that appropriate funding and planning may be organized. Note that the figure of 20% of the IBM service which was allocated above to general-purpose distributed computing support does *not*, by a very substantial amount, include the costs of support for interactive data analysis. It is assumed that this is included in the quoted batch CPU requirement of 160 units.

5. *COSTS AND STAFFING*

There are at present about 100 Apollo and 100 VAX workstations installed at CERN. We have been experiencing exponential growth, but it is expected that the rate of increase will settle down to about 200 new systems per year, leading to a ceiling of some 1000 stations. The *useful* life-time of a station might be expected to rise from about three years today to five years by 1992. We expect about 10% of the stations to be "high-end" systems costing an average of 80 kSF each, while the others will cost on average 30 kSF each, giving a total annual investment of 4.3 MSF. Central support costs, including special servers, gateways, software licences, etc. will bring the total annual cost to around 5 MSF. Maintenance charges for these systems are at present very high, amounting to about 25% of the (heavily discounted) purchase price per annum.

The most efficient way to staff workstation support activities will be to provide a well-managed and strongly coordinated central service, rather than using large numbers of part-time people with a variety of employers, interests and priorities. Estimates for central staffing are as follows, assuming that half of the workstations are Unix-based, and the other half use VMS.

¹⁴ PAW - Physics Analysis Workstation, Comp.Phys.Comm. 45, 1987; The PAW Users' Guide, CERN Program Library Writeup Q121.

¹⁵ The MUSCLE Report estimated that an IBM 3090-600E will be required to perform this task for the four LEP experiments in 1991.

- a. Management and basic systems support: 5 engineers;
- b. Development of distributed computing support: 3 engineers;
- c. User support: 2 engineers;
- d. Systems management and operation: 1 technician for each 50 nodes.

6. *WORKSTATIONS IN REGIONAL CENTRES AND UNIVERSITIES*

As the use of personal workstations for interactive data analysis evolves, they will become a requirement at all institutes participating in the analysis. Large centres with substantial disk capacity will therefore require to provide services to support these workstations and integrate them with their central processing and file server facilities. The larger institutes and CERN should also provide software and systems management support services for the smaller institutes.

Once distributed co-operative processing is established as the normal way of computing for offline physics at CERN and in the major regional centres, there will be a strong demand for improved integration between these centres and CERN, and with smaller institutes and individual physics departments. We should be aiming for the development of distributed data analysis on a European-wide basis, with the individual physicist able to work effectively from his personal workstation, using databases and DST files located at a number of different institutes. This will require not only a high level of coordination between the main computing centres, but also high-performance reliable networks. The requirements for the latter are covered in the following section.

7. *DATA REPRESENTATION*

The distributed computing environment which we envisage will highlight the inefficiencies of using different floating point number representations on the various computers involved. Although most workstations now use the IEEE standard format (with the exception of VAXstations), each of the mainframe types commonly used within high energy physics uses its private format.

10.5 Networking Implications

This report 'Computing at CERN in the 1990s' includes a part dedicated to general questions of networking. This section is therefore restricted to a brief summary of the requirements of centrally organized computing.

10.5.1 On-Site Networking Infrastructure

10.5.1.1 Short term evolution

The on-site Ethernet Service must continue to grow and be improved in a number of areas.

1. The planning, installation, management, monitoring and maintenance of the Ethernet infrastructure must be provided as a central service.
2. Installation planning must take account of workstations, and their requirements for connections to central services. Ample capacity must be installed in areas where the growth in the use of workstations by physicists and engineers will take place.

3. The reliability and availability of network services is crucial to the success of distributed computing. The central operation must ensure that major high-level services are monitored and appropriate corrective action initiated when necessary. It should be noted that the complexity of networks and network services is such that problem determination requires highly skilled staff and sophisticated network management systems.

The Domain token ring infrastructure for Apollo workstations must be maintained, with appropriate bridges to the Ethernet infrastructure and central services.

10.5.1.2 Medium term

The medium term (the next three years) will see the migration towards 100 Megabit/second FDDI token ring, initially as a "backbone" interconnecting the Ethernet and Domain infrastructures and the central services. An early priority in this development, which may well dictate the topology of the initial network, must be to ensure high bandwidth connectivity of workstations used for physics data analysis and the central IBM/VM service. It should be assumed that within three years, most new physics analysis workstations will be of sufficient power to require connection directly to the FDDI network.¹⁶

We should also expect to see FDDI coming into use to connect remote peripherals to the central mainframes (e.g. magnetic tape units installed in experimental areas), and as a means of interconnecting mainframes from different manufacturers within the Computer Centre.

In preparation for the longer term, CERN should acquire experience with the 100 Megabytes/second High Speed Channel (HSC),¹⁷ which will initially become available as a means of connecting very high performance graphics workstations to supercomputers. It should also, however, develop as a standard inter-processor channel suitable for the Computer Centre.

10.5.2 Wide-Area Networking

10.5.2.1 Short term evolution

The EARN and HEP-DECnet networks which have developed over the past few years with emphasis on connectivity and guaranteed delivery provide good general-purpose facilities suitable for remote job entry and mail, and for the transfer of small to medium files. The European-USA links provided are very important features of these services. Both DECnet and the HEP X.25 network also allow elementary remote login. The ability of EARN to increase its services to include full-screen remote login and inter-process communication facilities will be major factors in its continued success, as will the improvement of the capacity planning for its backbone network. In the event that a pan-European network for research and development emerges from the Eureka COSINE project, an adiabatic transition from the above services, including transatlantic gateways, would be vital.

Effective high-bandwidth connections between CERN and the major national and regional centres are required by the end of 1989, in order to enable the efficient management of physics production and data analysis work at these centres and at CERN. These links will provide database access and file

¹⁶ By that time, the cost of an FDDI connection should be comparable with that of an Ethernet connection, making it the network of choice if the cable is available.

¹⁷ ANSI Committee X3 T9.3.

transfer (of calibration data and programs, and perhaps small samples of physics data), in addition to remote job entry and other facilities. Since much of the communication would be automated, such as direct access from programs to controlling databases at CERN, reliability is vital. Efficiently operated 2 Megabit/second links will be required at least between CERN and Aachen, RAL, IN2P3, Saclay, and an Italian site (Bologna?) and in due course also CIEMAT, DESY, NIKHEF and the USA. Experience with such links should be obtained as soon as possible.

It is likely that neither HEP, nor the wider European research community would benefit from a completely shared network for all activities. The high traffic rate needed by HEP for distributed data analysis is inappropriate for general purpose networks. However, low-level multiplexing of HEP and general purpose traffic over the same fibres does seem a promising approach.

10.5.2.2 Medium term

The use of personal workstations will extend the user's normal computing environment to his office desk, in a form which he will be able to duplicate conveniently in his home institute and at CERN. The degree of integration with centralized file and CPU servers, which the CERN or regional computing centre offers him locally via LANs, will become an increasingly important factor in his productivity, giving rise to a demand for general-purpose communication links between CERN and the major computing centres in the member states, supporting 10-100 Megabit/second data rates. The current proposed experiment to connect CERN to several German laboratories using 10 Megabit/second channels on a 140 Megabit/second PTT fibre will give useful experience in this area and should be actively supported by CERN.

Personal workstations, and the development of distributed computing which they will stimulate, offer the first real opportunity for physicists to work effectively away from the CERN Computer Centre; the real cost of high-performance networking is now sufficiently low to enable them to do this from their office at CERN, or from their desk in their home institute; HEP must ensure that the tariff policy of the European PTT monopolies does not prevent European physicists from taking advantage of this situation.

10.6 Data Storage

10.6.1 Mass Storage Evolution in the CERN and Regional Computer Centres

This section summarizes the requirements described earlier in this report.

1. Growth from 200 Gigabytes at CERN today to 1000 Gigabytes by the end of 1991, implying a growth rate of 300 Gigabytes/year at a cost of 11 MSF¹⁸ per year.
2. Automatic cartridge loader at CERN capable of handling 200 mounts per hour by the end of 1991. The system should be able to handle cartridge tape drives connected to Cray, IBM and DEC, using a single cartridge store.
3. Minimum of 100 Gigabytes at any regional or national centre playing a major role in the processing of any large experiment.

¹⁸ Assumes 200 Gigabytes of IBM-compatible storage at 30 kSF per Gigabyte, and 100 Gigabytes of other (VAX, Cray) storage at 50 kSF per Gigabyte.

10.6.2 Distributing the Data

10.6.2.1 On-site distribution

The DST data required for LEP physics analysis, as estimated in the MUSCLE Report, poses two major problems for storage and distribution.

1. There will be a great deal of it: a minimum of 800 Gigabytes for each experiment by the end of 1991 (some estimates are a factor of four higher). On that timescale, economics will dictate that all but 10-20% of this data must be held on cartridge tapes, but even then the tape mounting rate will be very high.
2. The computing power required to maintain the master DSTs and prepare the subset DSTs for subsequent analysis will be twice the computing load needed for the final physics analysis, requiring the power of an IBM 3090-600E for the fraction of the work done at CERN.

Thus, managing the LEP DSTs will require a lot of disk space, serviced by a high-capacity cartridge tape robot, and supported by powerful CPUs. Economic (the cost of disk space and CPU power, and the possibility of sharing robots and tape drives) and practical (ease of management and operation) arguments will favour a centralized solution, at least for the first phase of LEP. The principal drawback to such a solution is the difficulty in achieving sufficient network bandwidth to feed the data to the workstations where the final processing will take place. This is mainly due to the current relatively poor throughput of heterogeneous network connections to IBM, rather than to any fundamental weakness in the concept of a centralized file server.¹⁹

We must study carefully early experience with LEP data analysis using workstations, in order to understand better the issues of data distribution. In the meantime, strategies for file servers should remain flexible and be reviewed regularly in order to be able to benefit from changes in storage costs and technology.

10.6.2.2 Distribution of DST data to external institutes

During the next five years, improved wide-area networking will enable individual events, or even occasional complete tapes, to be transferred between CERN and external institutes with some convenience, but until optical fibre bandwidths become available, almost all of the DST data will be distributed on cartridge tapes. In order to assist in this massive operation the following central services should be provided.

1. Tape copying service, including the allocation and labelling of the new cartridges.
2. Direct mailing service, whereby a cartridge can be allocated, labelled, written and then packaged and sent by express carrier to an external institute.

¹⁹ DEC LAVCs achieve sustained aggregate network data rates of around 250 kilobytes/second, more than twice that available with a single Ethernet connection to the IBM. Even higher rates can be achieved between some 80386 based PC-s. This is due to the use of light-weight protocols and efficient implementations of them. While the former may not be easily applicable in a heterogeneous situation (where standard protocols must be used), we can expect to see the development of more efficient implementations, and the use of "protocol engines" in the hardware of the communications interfaces, in a 2-4 year timescale. In the meantime, it seems reasonable to connect multiple Ethernets to a central file server, using several physical controllers, in order to achieve higher aggregate throughput.

3. Centralized cartridge tape database holding details of tape history and location, implemented using Oracle and available to collaborations locally or remotely, for inclusion in their book-keeping systems.

The cost of tapes and freight should be charged to the requestor of the data, probably through a CERN team account.

10.6.3 Storage Management

This report has already described the need for a data management for the whole CERN community. This problem is too complex, and depends too greatly on the situation at regional centres and universities, for a technical solution to be proposed here. Technical solutions, almost certainly starting as small scale trials rather than global services, should be pursued by a committee reporting to the management of European HEP.

A more specific recommendations can be made about a more limited goal. The need for automatic disk-to-tape migration of user files and experiments' DSTs has been described in section 9.2.4 on page 113. This facility must be made available on the central IBM system at CERN by early 1990 at the latest. If suitable commercial software will not become available in time, it should be written at CERN or, possibly, be commissioned from a competent software house.

Chapter 11

EVOLUTION OF CENTRALLY ORGANISED SOFTWARE SUPPORT AND DEVELOPMENT

In this chapter we address the questions of software support. We concentrate on system-related software and on general questions of software support. The following chapters will go into more detail on specific areas where software support of an application-related nature is seen as an existing or potential problem.

11.1 The Changing Landscape

The distributed computing environment described in this report could revolutionize physics analysis. However, the increasing complexity of the hardware infrastructure must not be allowed to produce increasing complexity for the user. Every effort must be made to render the appearance of the operating system, and of the data storage, as location-independent as possible.

At the same time as hardware complexity is increasing, so is the complexity of HEP software. The software for any current large experiment is well beyond the point at which any single person could understand all the details. A later chapter will describe the promise of software engineering techniques for the design of HEP software. In addition, it seems inevitable that coding will eventually have to be done at a far higher level of abstraction than that of current 'high-level' languages such as Fortran. The creation of an HEP-specific 'even-higher-level' language is an option meriting further study.

11.1.1 Operating Systems

An environment in which physicists must use only one viable, but unexciting, operating system may well be more productive than one in which they must use several better, but different, operating systems. This is the context within which HEP should give serious consideration to an increased use of Unix.

For all workstations except those from DEC, Unix is now the default operating system, and on many workstations there is no other possibility. At CERN the CRAY runs under Unix, and VAXes and VAXstations could (but usually don't) run under Unix. However, until now, Unix has not been available for the IBM systems which are the centre of data-handling and computing for most experiments. With the announcement that AIX, (IBM-Unix) will become available for IBM mainframes during 1989, it becomes possible for physicists to contemplate using only one operating system. The "Open Systems Foundation" initiative, involving Apollo, DEC and IBM among others, offers the hope that the different implementations of Unix will really appear very similar to users.

It would certainly be premature to recommend that HEP convert rapidly and totally to Unix. However, it would be unrealistic to recommend that no Unix service should be offered on central IBM systems until its success was assured. Few physicists have any experience of being able to do all the analysis of a large experiment within one operating system, and an abstract study of its advantages would

almost certainly be judged inadequate.

It is recommended that an AIX service be started on the CERN IBM system as soon as the AIX specification becomes adequate to meet HEP needs. This service should be started on a small scale and its expansion should be driven by user demand. Some flexibility will be necessary in assessing whether AIX has reached an adequate state of development. It is not recommended that CERN write batch and tape handling systems to provide essential facilities missing from early versions of AIX, but smaller deficiencies might be repaired by CERN effort. This recommendation implies that some new manpower must be found, or that there will be some degradation in the support for the existing CERN services.

It should also be noted that, even though great benefits are expected from the use of a single operating system, exchange of data between systems will still be hindered by their different data representations, except between systems using the IEEE/ASCII form

11.1.2 Languages

The working group has not attempted an evaluation of alternatives to Fortran, or even of the impact of the likely evolution of Fortran. For all its inelegance, and lack of safety features, it seems certain that Fortran will remain the main language for HEP code well into the 1990s.

No existing or planned language, including, of course, Fortran, provides machine independent tape or network I/O. HEP has overcome this problem with utilities such as the ZEBRA-FZ package, which are effective, although naturally less convenient than if they were integral parts of the language.

Fortran is notoriously lacking in data-structure concepts. Many more modern languages have some support for data-structures, although it is not clear that they could support the complexity required for handling HEP data. Fortran at least allows the freedom to provide sophisticated data-structure support through additional packages such as ZEBRA-MZ. The ensemble of HEP packages which run on top of Fortran, together with Fortran itself, form a powerful environment for HEP data analysis. It would not be easy to guarantee that a move to another language would improve the overall environment.

Most physicists do not have the time or background knowledge to perform a thorough review of languages before starting an experiment. The CERN or HEP community should periodically review developments in languages, both to add weight to their input to Fortran standardization committees, and to identify any serious alternatives to Fortran.

11.2 Software Support for Experiments and Packages/Libraries

11.2.1 Support for Experiments

Ten years ago, most experiments at CERN expected and received support from CERN, or a major laboratory, to write the experiment-specific software. This help normally took the form of programmers assigned to the experiment, and was especially important for real-time programming and data-acquisition. Today, CERN and the major laboratories are not the sole repositories of computing expertise, and the case for this type of software support has weakened. However, there remain some good arguments for the provision of some experiment-specific support. Even in the largest experiments continuity is a major problem; the efficiency of the software effort can be greatly enhanced if at least some people can work full-time for many years. It is very difficult for universities to provide these

people, especially if they must work largely at CERN. In the case of smaller experiments with shorter timescales, it is inefficient to require that, for example, university physicists become experts in a standard data-acquisition system which they will only use for a few months.

The current level of experiment-specific support is widely criticised by physicists. This criticism may reflect the unclear (or non-existent) algorithms used to allocate support personnel, and the lack of coordination between CERN and other centres, rather than truly inadequate resources. CERN, and the regional centres and laboratories which can also provide such support, should review urgently how support is organized, and how this organization should be improved.

11.2.2 Packages and Libraries

Central support for widely usable packages and libraries has been extremely successful for many years. The CERN libraries, and packages like GEANT3, HBOOK and ZEBRA are used not only within the CERN community, but also for many HEP experiments worldwide. A relatively small centralized effort has prevented much wasteful duplication, and has made high quality software available to experiments which could not afford to write it themselves. It is very unlikely that a similar success could have been achieved merely by hoping for cooperation between experiments.

In spite of this evident success, there are still some problems:

- Existing central support is overstretched;
 - widely used products are sometimes not sufficiently tested and documented before release,
 - there is little or no room for additional projects.
- Although informal relations with users have been mainly excellent, there is no formal mechanism for making requests and assigning work. As a result, some users, particularly in smaller experiments, feel the system is unfair, and that only the needs of large and vociferous experiments will be considered.
- As with experiment-specific support, there is no mechanism for coordinating these activities on a European scale.

An additional, and even more complex problem, is the determination of which software should be bought commercially. There can be no doubt that widely available commercial software, meeting the needs of HEP, should not be duplicated. The success of less mature commercial products, where HEP is, perhaps, a major fraction of the market, is much less clear, and existing experience does not make it easy to formulate a simple policy about what should be bought and what should be written. These decisions should be made on a case-by-case basis by a technically competent body representing the CERN community.

The existing coordination problems, particularly in relation to CERN and major centres, make it impossible to say whether manpower levels are grossly insufficient or not. It is recommended that a software support committee be set up by European HEP with the brief to:

1. Examine support manpower at CERN and at other European laboratories and centres. Assess whether manpower levels are adequate to provide the coordinated software support needed to minimize wasteful duplication.
2. Coordinate support activities among all laboratories and centres willing to participate. This implies that software support personnel be instructed by their local laboratory to perform at

least some of their work under the direction of the software support committee.

3. Advise, case-by-case, on the relative merits of buying commercial software as opposed to writing packages within HEP. This advice should take into account the extent to which commercial products meet HEP needs, the support which the HEP community would have to give to either HEP-written or commercial products, and the purely financial implications for (at least) the whole of the CERN community.

The committee should not limit its consideration to simple extrapolations from existing packages and libraries. The distributed computing environment, and the increasing complexity of HEP software make it necessary to explore new areas for HEP software support. Examples of such areas are support for a location and device independent user-interface to data, and possible HEP-specific very high level languages which could simplify the coding of data processing and physics analysis.

Chapter 12

SOFTWARE SUPPORT: DATA BASE SYSTEMS

The questions specific to data base systems have been looked at by a separate group, with the following guidelines:

- Emphasis should be put on the user needs, and thus most contacts have been with users, rather than designers of data bases.
- Adequate commercial data base systems are to be given preference over homemade developments, wherever available. Limits do, however, exist due to specific requirements of users like access speed or simplicity of use.
- Due to the specific structure of our collaborations, the transfer of data base information between institutes is a vital constraint as important as portability of code and event data.

Our group has used as means of contact with different groups a widely circulated questionnaire, followed by detailed contacts with selected users, and all our conclusions are based on this input. L.Adiels, F.Bruyant, G.Gopal, A.Osborne, P.Palazzi, E.Rimmer, and M.Sendall have contributed substantially.

We have observed a large amount of duplication of work in this area. We therefore present clear recommendations in an attempt to limit the development of data base applications to a small number of acceptable solutions. We are conscious of the fact that the present recommendations still need further discussions with both users and the designers of data bases, before final solutions can be adopted.

Whenever specific products (e.g. Oracle or ADAMO) are mentioned in the recommendations this is not to be taken literally. They have been chosen since they are used at present and define the minimum requirement we have for any future product. We are fully aware of the rapid development in this field.

In this report, the word "Data base" is used in a broad sense and matches with the definition given in the ECFA report:²⁰

"A collection of stored operational data used by the application system of some particular enterprise."

In principle a data base need be no more than a set of computer-based records, but the term is generally used to signify that the data are

- integrated, combining records that would otherwise be stored in separate files
- structured, with some inter-relationship between various types of record
- shared by a number of users, perhaps sharing it for different purposes
- organized so that rapid access to any subset of the data is possible.

²⁰ Databases and bookkeeping for HEP experiments, ECFA Working Group 11, ECFA/83/78 (Sept 83).

12.1 Data Base Applications in Present Experiments

12.1.1 Addresses of Collaboration Members

The purpose of this information is to keep the name, address, telephone number both at CERN and in the institute for administrative purposes. The integration of this information with electronic mail addresses has been made by a few experiments.

Among the solutions adopted, some experiments have no mailing list, others use card image files with command files to use them, later experiments like Aleph and Delphi use the NADIR system. NADIR makes use of the Oracle relational data base system to store experiment members names, addresses, telephone numbers and electronic mail addresses. Interactive updates and queries are performed via Oracle full-screen forms. Various ADAMO tools are used to extract the Oracle information into portable files which are then exported to member institutes which do not have the Oracle system. ADAMO-based tools are then used to reconstruct and access local direct-access files and include an electronic mail distribution system called AMSEND. One experiment uses HYPERCARD on a Mac-Intosh.

Most experiments express their satisfaction with the adopted solution, even card image files when the number of participants is sufficiently small. It should be noted, however, that these data are purely for internal use. A single data base collecting all CERN users could be beneficial to all. But the latter question was not raised in the questionnaire and is thus the interpretation of our group only.

12.1.2 Electronic Mail Addresses

The aim is to allow to send mail electronically to a user's family name, or to lists of people, without the need to know his account(s). This mail is automatically converted to his preferred mail address and optimally routed. This tool is heavily used to inform the members of a collaboration rapidly and easily.

The presently most frequently used method to retrieve electronic mail addresses seems to be NAMES under VM. Two LEP experiments (Aleph and Delphi) are using information stored via NADIR and used in AMSEND, developed in Aleph. NADIR stores, as well as geographical addresses, the collaboration members home computer, his login id(s), networks and mail servers. Other experiments rely on simple files. All experiments found their system adequate and preferred that the updating of both the mailing lists and the electronic mail addresses be made directly by the experiments (e.g. by the group secretaries).

The DD division has set up an Oracle based system called EMDIR (electronic mail directory) where each user is invited to enter his preferred electronic mail address. This system is unpopular (it has a poor user interface), many users are not keeping it up-to-date and because of this it cannot be relied upon by collaborations. None of the experiments which replied mentioned that they are using it, although the authors report "thousands of queries per month". They intend this to be "the basis of future mail systems".

In summary, as is the case for the geographical mailing lists, the information is used exclusively within the experiment itself.

The DD division has recently implemented an Oracle data base of the login identifiers on the central systems of all registered central computer users and this data base includes the CERN telephone directory. This perhaps could become one part of a common geographical and electronic mail addressing system for all CERN experiments where the experiments themselves would furnish the other infor-

mation about their collaboration members.

12.1.3 Experiment Bookkeeping

The purpose of bookkeeping is to keep track of data taking runs, experimental data, storage media (disks, tapes) and program versions used in the different phases of event reconstruction and analysis.

Amongst the current experiments, several use simple systems, like card images, others use random access files with KAPACK or the Oracle RDBMS and some experiments have just started with Oracle or will start soon and are still undecided. Opal intends to use Oracle to hold "log-book" information. It seems that this is a place where a common effort could lead to a general solution.

12.1.4 On-line Data Bases

There is no single set of data base tools today to span the whole range of on-line applications such as storing detector constants and calibration data, histogramming, error logging and run parameter logging. This has led to a large variety of individual solutions ranging from simple files to KAPACK based applications or even the usage of a RDBMS such as Oracle. The CP-LEAR experiment will use Oracle online on a VAX but only as an indexing system pointing to other logical and physical entities (histograms, tables and files).

12.1.5 Description of the Detector Geometry

The aim is to provide a detailed geometrical description of the detector components and their evolution in time. It is needed in the off-line programs for the simulation and the reconstruction of the events and for graphics display. The system must be able to cope with a simple description, as used in reconstruction, and a very detailed one, as used in the simulation. A fast access time is crucial, given that the simulation or reconstruction will frequently access it to get answers of the type "where am I" and "how far can I go".

Again, a wealth of homemade solutions have been developed. Some experiments use random access files based on ZBOOK, KAPACK or ZEBRA. The manpower investment in this area is very large (several man-years per experiment, on average). The main difficulty resides in the extremely fast access required. This explains why different experiments have tried to optimize the data storage to their needs by using different solutions. It also excludes the direct usage of Oracle for the data access at execution time.

The need for a DBMS to provide fast access to Detector Description data was identified very early on in Delphi. At first (1984) they developed a system based on KAPACK, tailored specifically to this particular application. So long as the originally designed Data_Structure was adhered to the system worked very well. However, Delphi were not happy to freeze the design of the Detector Description Data_Structure so far ahead of data-taking.

So, a generalized DataBase Management System - CARGO, also based on KAPACK, was developed, care being taken not to tailor it to the needs of any specific application. The Delphi Detector Description - Geometry Data at first (1986), Calibration Data later (1988) - has been successfully structured according to the conventions of CARGO. Fast access to this data for Simulation, Analysis and Graphics programs is provided via an application specific access package (DDAP) which structures the data requested by the User Program into a ZEBRA division.

The Geometry File contains the detailed Geometry for each of the 18 or so sub-detectors including the beam pipe, coil and the cable duct areas.

A first version of the calibration constants have now been installed in the calibration file for almost all the sub-detectors and programs are now beginning to use this data.

The L3 experiment has built a detector description data base using the ZEBRA (CERN data structure management package) direct-access package RZ. The RZ package allows a stored ZEBRA bank to be identified by up to 100 4 character long keys. L3 has written a Fortran data base package (called DBL3) on top of Zebra RZ to provide the user interface to their RZ data base. Their system stores changes to data as update records rather than changing base records and they also store pointers (in the form of file names) to their multiple RZ files. The user of DBL3 is normally returned a ZEBRA pointer to his requested data, the system having read it from disk into ZEBRA memory. DBL3 took about one man-year to build and they feel it will provide the functionality they need for many years. In addition they have built a package for interactive access to the data base using the KUIP (Kit for an User Interface) component of the PAW (Physics Analysis Workstation) system. Since PAW is itself based on ZEBRA and RZ this was a natural extension.

The Opal experiment will also use an RZ based system for detector description but make the general remark that parts of their data would be suitable for a relational data base.

A general solution for experiments needing the functionality of a commercial data base such as Oracle but having constraints of performance or portability may consist in a hybrid solution of maintaining the master information in Oracle and providing translate programs that generate normal Fortran files.

This is the solution adopted by the Aleph experiment who store their detector description constants in an Oracle data base including validity time-stamps. The data definition is made using data modelling with tools from the ADAMO package. The data are then extracted into RZ or BOS (the DESY data structure management system) banks which themselves can be stored in Fortran direct-access files. Each Oracle table becomes a BOS/RZ bank with the numbers of rows and columns stored as the first words of the bank (this amounts to extracting a single "view" from the relational data base). The Oracle data base contains data from the different detector components (including the beam pipe, cables etc.) and updating is done by 1-2 people for each of them.

12.1.6 Calibration Constants

The purpose of this package is to store the calibration constants and to keep track of their changes as function of time.

Similar solutions are adopted as for the detector geometry, since the two need to be linked together, the main difference being that the volume of calibration data is normally much larger than that for detector description. Many of the same comments apply as for the previous section. Opal will use RZ direct-access files, while Aleph will include in their Oracle data base time stamps and pointers to their calibration data (e.g. file names or tape numbers).

12.1.7 Event Data (or Memory Management Systems)

All the answers received so far were related to the management of the event data in memory. All existing memory management packages are mentioned (BOS, HYDRA, ZBOOK, ZEBRA), though most of the new experiments tend to standardize on ZEBRA. None of these applications involve a real data base. In particular no experiment mentioned the need of a data base for direct access of event

data. Opal envisages an event server which can recall events from specific blocks on 3480 cartridge tapes and expressed a wish for a tape management system at the computer centre.

12.1.8 Bookkeeping of Program Versions

The intention was to find out whether people feel the need to keep track of program versions in time, in order to allow the reprocessing of events in the same conditions (and with the same bugs) as they were a long time ago. Also here, the question was not clear enough and all answers were related to source code management.

12.1.9 Histograms and other Physics Results

The purpose is to keep a catalogue of results/presentations for on- and off-line. In the on-line context, this may be monitoring histograms, and in the off-line context, reconstruction results or physics data.

In the answers, a majority expresses support for PAW.

12.1.10 Software Documentation

A data base could keep the whole software documentation and the description of the data.

Several experiments use simple files. The only applications using data bases are from Aleph using the data model tools from ADAMO to provide software documentation (in addition to write-ups on paper).

Opal uses an internal system called OpalDOC but would merge this with the DD CERNDOC system when some bugs in it had been fixed and when it accepts TEX formatted files.

12.1.11 List of Publications

The purpose is to have an up-to-date list of all publications for the experiment. This should include the title, the list of authors, the reference number, the data and subject keywords.

Currently, the experiments use simple homemade solutions. They are satisfied with their solutions, but some wonder whether in the long term a more elaborate system may not be preferable, provided the access remain sufficiently simple. The Particle Data Group uses Oracle. The EP Electronics are using ISIS and find it hardly adequate.

12.1.12 Other DB Applications

A need expressed by experiments and mechanics groups is to have a data base application which can handle the registration of parts of the equipment and their life history. The Opal experiment uses a "4-th Dimension" data base on a MacIntosh for the inventory of their electronic modules and find this very convenient with the drawback of no easy access for the general members of the collaboration.

The DD Online Computing group use Oracle to maintain the pool of online equipment (several thousand items). This is now used by a wide range of people for arbitrary queries and has been extended to include experiment hardware configurations of equipment not owned by the group and also to

include information on maintenance contracts. This runs on the IBM VM system and makes use of the Oracle report generator. They did not consider using a PC based system because of lack of connectivity with other central data bases. They feel there is a gap between the very competent technical support offered by the DD data base section and the end user community (indeed the User Support group feels they would need dedicated Oracle experts to do other than refer users to this section).

An interesting data base application is the one set up by the DD Fastbus System Management team which allows experiments to generate code which is used to initialize a Fastbus configuration (this contains full configuration and routing tables, arbitration levels and so on). Fastbus is a multi-level triggering system where the hardware is fully addressable and functionally defined. By using an Entity-Relationship model they showed the data describing a Fastbus configuration to be relational in nature. They thus use an Oracle data base on the central VAX cluster to store the basic descriptions of the different Fastbus modules and the relations between them. A given experiment uses a set of Oracle SQL forms (some tens of thousands of lines of SQL*FORMS containing many inter-table checks) to build the description of their particular configuration. A Pascal program then reads the resulting data base and generates the required configuration initialization code in an ordinary VAX file which can be down-loaded to the experiment concerned. Use of Oracle gave them the benefits of a backed-up data base with a high-level of integrity, possibilities of roll-back and multiple concurrent access.

12.1.13 Present DB Applications in Experiments: Conclusions

The most striking fact appearing in the answers from the experiments is the large variety of solutions adopted. Several experiments expressed their dissatisfaction about the chosen solution. Such a situation seems to have originated from the regrettable lack of general purpose data base applications software adapted to their problems. The design and development of such general purpose applications needs expertise with a given DBMS and manpower. A simple ad hoc solution is often "cheaper" for the individual experiments, but, summed over the experiments, it is certainly more manpower consuming than to provide once a general package. Some exceptions should be noted, for instance the NADIR system developed in the Aleph collaboration. It is also striking that this package was adopted by other collaborations, like Delphi, although it needed some "customizing" before it could be used. NADIR is a good example of an application very close to being a general solution for mailing lists, which is now being worked on by the MIS unit. This example shows that when a sufficiently general application is provided, the experiments will tend to use it, independent of the complication of the underlying DBMS.

A general observation is that experiments are using both commercial relational data bases and the traditional HEP data management packages, depending on the nature of the data concerned. Some translate the relational data base into a HEP package for reasons of portability.

12.2 Broad Areas for DB Applications and their Requirements

Several areas can be identified where the use of a DBMS is clearly justified. and is quite similar for the different experiments. Each of these areas has specific requirements and needs different application software for access and update by the users. Some applications already exist, most of them need, however, further development. Even where this may require a sizeable effort, overall harmonization seems to be the only way to avoid the present wasteful approach of multiple ad-hoc solutions, each of which is manpower-intensive.

We will first discuss the fields where a HEP-wide solution can easily be envisaged. For these applications a complete tool package should be written by some DB "professionals" in consultation with an advisory team" from different experiments.

This is then followed by a short discussion of data base applications where a general purpose package seems to be – at least at the moment – not in sight. We believe, however, that at least for the modelling phase and for the transportation of data bases general tools can and should be provided.

12.2.1 DB Applications where Generalized Packages are Possible

12.2.1.1 Experimental Administrative Database

As mentioned above this type of data base should contain all data necessary for the information exchange (mail AND electronics mail) between all members of a collaboration such as address, phone number, office, function, responsibility, affiliation, network addresses, etc. It will contain only few, well defined, tables. To avoid unnecessary duplications and work this data base should be coupled to already existing ones like: user registration, telephone directory and personal data.

All members of a collaboration need read access to (part of) the data base via panels and/or via coupling to high level language programs. However, since probably only reduced functionality is needed at some places the access can be realized on an extracted d/a file or even a sequential file. All other access might be (should be?) limited to the secretaries of each collaboration. At least part of the data base should be accessible HEP wide.

It would be desirable if each member of a collaboration could initiate the changes to the DB as far as his own data are concerned. In practice this would essentially mean that a panel system has to be provided which creates the necessary update corrections sent then to the Data Base Administrators (e.g. the group secretary). Since probably only reduced functionality is needed at some places the access can then be realized on an extracted d/a file or even a sequential file.

Since the requirements are almost the same for all experiments, a general data base package including all necessary tools for the communication with the data base and for the transport to/from external institutes should be provided by CERN. As a starting point the NADIR system could be considered.

This application can best be realized on top of a relational data base. If done so it will be quite easy to increase the functionality with time, whenever needed, without having to redesign the already existing parts of the data base.

Tools are needed for:

Panel systems (SQL*FORMS)

Coupling to higher level languages

Exchange of the data base between different computer systems

Exchange of the information between d/a files and the master data base on e.g. Oracle.

12.2.1.2 Electronics and material Bookkeeping

With this database we want to keep track of all electronics modules and materials used inside a collaboration. Therefore data such as module description, address of manufacturer, purchase information, inventory and location are entered.

The requirements are quite similar to the ones mentioned for the experimental administrative database. In addition coupling to a general data base describing the internal functionality of modules should be envisaged.

This data base will be accessed only by few members of a collaboration. Therefore it can be kept centrally at CERN with some access from outside institutes. The package itself should, however, be made available to all collaborating institutes to have their local bookkeeping system identical to the one used centrally.

As for the experiment directory this is an important area where a general product should be written and maintained.

12.2.1.3 Tape and Production Data Base

In a large experiment it is difficult to follow the history of a single event throughout the different phases in the off-line analysis. Therefore a database is needed which keeps track of files and program versions used for real and MC-events. In addition a database which contains information on storage resources (tapes, optical disks etc.) is needed.

This data base has a simple and well defined structure and could be realized on almost all DBMSs. What is important here is that read and write access is needed from terminals and from higher level programs (e.g. reconstruction programs) in real time. It has to be accessible – may be not with the same functionality – from all institutions of a collaboration.

As for the previous application areas the structure is well defined and a general package can easily be designed. If realized on top of a RDMS, experiment dependent additional information can easily be integrated.

12.2.2 DB Applications for which Tools could be Specified and Written

12.2.2.1 Detector Description

In a complete detector description database the geometrical description of the detector components, the description of the electronics path to each sensing device, the deficiencies, the alignment corrections and their variation with time have to be entered.

In general, this leads to a complicated structure with many different tables (100-1000). To cope with this complexity of the data structure and to allow for easy navigation through it a RDBMS is needed. In particular, there are basic records and, separately, update records which are merged with the basic records only at the time of access.

Access to the data base is needed by many different applications and packages in read and write mode. Access time is crucial, hence the data base is probably read in once and all access is made to a structure in memory as long as there are no changes in the time dependent constants.

In this complex system updates to the data base have to be done under the control of the Data Base Administrator of the collaboration.

The data base has to be available (with reduced functionality) on all computer systems at CERN and in the outside institutions.

Unless one introduces strong restrictions a general package cannot easily be provided. However, it would be desirable to model an example data base which could be used as a starting point for experiments which have a small number of (simple) detector components.

In addition to the tools already mentioned for the experiment directory, more and probably application dependent tools are needed, especially if the commercial DBMS cannot be accessed directly at all places.

12.2.2.2 Calibration Constants

Calibration constants used for (online) corrections of experimental data present generally a huge amount of unstructured data. Therefore, for most experiments there is no strong need to store them directly into a data base. However, the data have to be accessible via the database. It would be sufficient to enter the access information into the data base and keep the data on separate files outside (but nevertheless as part of) the data base especially since some coupling to the detector description data base is needed.

The access to the data base is almost exclusively done from higher level language programs.

12.3 Data Base Systems: Recommendations

General recommendations are presented first. They are followed by further recommendations specific to certain areas.

As mentioned at several occasions in this report, a large duplication of effort has gone in developing real or pseudo "data base applications" inside the experiments. It frequently happened that a simple ad-hoc approach was chosen, which in the long term proved to be too restrictive, once the volume of data increased or the data base was needed for applications it was not originally designed for. It also happened that commercial systems were tried and found too slow to be acceptable. The most frequent arguments used by opponents of data bases remain:

- Slowness of access, especially in the area of the detector description and calibration constants. This slowness now seems to have been partly overcome.
- High learning threshold, which is undoubtedly a strong argument in case of simple problems like lists of publications. This is, however, purely a question of user interface and it could be overcome by providing general purpose applications with a suitable (graphics?) interaction with the user. The intricacies of the underlying DBMS can be completely hidden to the user.

Specific applications should be written as part of the general support software. This should be the result of a concerted effort between the DD division and the users. Suggestions are given in the specific recommendations below.

12.3.1 Education and Training

Although DBMS have found a wide acceptance in the engineering world (e.g. LEP), there is still a reluctance for their use by physicists, who often prefer their "homemade" solution. It should however be realized that, in the future, experimental data are more likely to be managed by a DBMS than by our existing packages, which remain tightly linked to our prejudices on Fortran as the main language used for our programs. We recommend that:

An effort should be made to make physicists in experiments more aware of the potentialities of commercial DBMS for their applications. This could be achieved by intensifying training in the areas of data models (software engineering) and DBMS.

12.3.2 Design Support Team

It should be avoided that the effort to write General Purpose or applications code be duplicated by the different collaborations, as is often the case today. To help overcoming the initial difficulties and the long term maintenance, we recommend that:

Manpower should be made available to support centrally the experiments, starting with the design of the data base and continuing during the whole life cycle, including the implementation of the application dependent code. This support team should also ensure the long term maintenance of the General Purpose applications described below.

As the design of a data base application involves software methodologies and as software methodologies make implicit use of data bases, the same support team could cover both areas. This will help to speed up the implementation and to ensure that the resulting application can be used by several experiments.

12.3.3 Data Model Software

It should be realized that the usage of a DBMS for a specific application requires one to several man-years of work to design the structure of the data and to write the necessary SQL and/or SQL*FORMS code. Several areas described below will need such application code to be written. It has been stressed at several occasions that this task is greatly helped by the use of a formal Data Model. This immediately provides a link with the "Software engineering group", but we independently recommend:

A package should be provided to design interactively a Data Model and to store this definition in the form of a dictionary in ZEBRA files. The Entity-Relationship Model and related software from ADAMO should be considered as a first step in this direction. This would allow to profit from the experience gained and possibly from existing tools, including commercial ones.

12.3.4 Portability of the Data Base Information

The Oracle Relational Data Base Management System has found a fairly general acceptance in the user community as the standard data base system. It is already being used and will be used even more in many different applications. Some features which were stressed by users at several occasions are its handling of concurrent write access, its user interface via SQL*FORMS and the possibility to implement tools on SQL to guarantee the data integrity. A serious problem with Oracle is its price, which prevents it from being used throughout the HEP community. As an indicative value, it costs approximately 25 kSF for a VAX780 and 10% per year for maintenance. Other tools have to be paid for in addition. A solution must be found which allows also the smaller institutes to use the data.

A package should be provided to map data from Oracle to a ZEBRA (RZ) data structure. The reverse could also be implemented, provided a data model describes the structure of the data in the DB. A decent user interface should be written on top of these files to allow the users to inquire about the information contained in this structure and to update it. Tools provided with the ADAMO package could be used to learn from the existing experience and could possibly be used directly as part of the proposed package.

This approach guarantees that the information can be used and updated in the collaborating institutes free of charge.

12.3.5 Area of Experimental Administrative Data Bases

Several data base applications, all using Oracle, have been developed, like the User Registration DB, the EMDIR system, several mailing lists from experiments using NADIR and the AMSEND electronic mail system. Data are often duplicated and contradictory, because some are out of date. It is recommended that:

A data base should be set up covering all CERN (or HEP?) users and other people related to experiments. It should link with information from existing data bases. It should include the functionality required for experiment mailing lists and experiment specific data. Control of the data, i.e. entering and updating the information, should stay within the experiment concerned. We further recommend that a study be made on existing tools and their performance, in order to coordinate any future efforts, such as those being made around NADIR and EMDIR. The functionality should cover at least the one of the NADIR and AMSEND systems.

12.3.6 Area of Bookkeeping Data Bases

LEP experiments seem not to have invested much manpower in this area yet and a common solution still seems possible. Therefore, we recommend:

A solution should be researched and developed urgently, in common between the LEP experiments, in the area of tape bookkeeping, to avoid duplication of effort.

12.3.7 Area of Documentation Data Bases

The part of the software description related to program code and data description may come automatically with the Data Model tools and is considered as part of recommendation 3.

In addition, "paper documentation" will still be required and makes this an important area to be looked into. CERNDOC seems to have basically the right functionality, but it should be coupled to other data bases, like the CERN Library data base and experiments administrative data bases. It is also essential that the data can be transported and printed throughout the HEP community. We recommend that:

The redesign of existing documentation data bases (CERNDOC, HEPPI, ISIS etc.) into a common data base system (e.g. Oracle) should be envisaged.

12.3.8 Area of Detector Description and Calibration Constants Data Bases

This is probably the area with the largest investment of manpower and the largest savings if a common solution could be found. What makes it particularly difficult is the large data volumes encountered and the high speed needed for accessing the data. It seems unlikely that the existing choices for the detector description and calibration constants data bases could be modified, given the imminence of the data taking for the experiments involved. It also seems difficult to provide a completely general solution, given the large differences between the detectors.

It is at least expected that the experiments will store the information in a DBMS and in ZEBRA files. The availability of software to handle a data model and to convert between a DBMS and ZEBRA could give some help. Some guidelines and assistance for structuring the data would also be helpful. The experience gained with the usage of the large data bases set up by the LEP (and other)

experiments should be exploited to try to define further support tools and minimize duplication for post-LEP experiments. A study group should be set up around the end of 1990, after a sufficient experience is accumulated.

12.3.9 Area of Interactive Data Analysis Data Bases

PAW datasets are expected to play this role. PAW is already widely accepted by the user community, and no new software needs to be developed. The design of PAW explicitly does foresee to use PAW datasets as data base; the present PAW implementation, however, seems to be lacking the built-in possibilities for identifying the contents of the various directories. In-line identification seems necessary to make PAW data exchange inside and between experiments largely self-documenting.

12.4 Data Base Systems: References for detailed reading

1. *Databases and bookkeeping for HEP experiments*, ECFA Working Group 11, ECFA/83/78 (Sept. 83).
2. P. Chen, *The entity-relationship model – Toward a unified view of data*, ACM Transactions on Database Systems, Vol 1, NO 1, March 1976.
3. S.M.Fisher, P.Palazzi, *The entity-relationship model of ADAMO*, Aleph report ADAMO Note 3, March 25, 1988.
4. R.Mount, *Database systems for HEP experiments*, Computer Physics Communications 45 (1987) 299.
5. A.Putzer, *Data-base system in High Energy physics*, Proceedings of CERN School of Computing Troia, September 1987 and Heidelberg University report HD-IHEP-88-02.
6. Z.X.Qian, F.Blin, W.G.Moorhead, P.Palazzi, *NADIR, the New Aleph DIRECTORY*, Aleph report ADAMO Application Note 10, November 11, 1986.
7. Z.Qian et al., *Use of the ADAMO data management system within Aleph*, Computer Physics Communications 45 (1987) 283.
8. E.M.Rimmer, *Programmed initialization of complex data collection and control systems*, CERN-DD/88/6, March 88.

Chapter 13

SOFTWARE SUPPORT: SOFTWARE ENGINEERING

This subject was considered by a separate working group. To define what is meant by software engineering: 'the application of engineering-like methods for large scale software production'. In this report we consider the methods and tools which are likely to become a necessity for such complex projects as will be experiments (and accelerator control) at a future hadron collider.

The major recent trend in HEP experiments and one that promises to continue, is that experiments have become very large, in terms of equipment, numbers of physicists and data volumes. Experiments, in addition, have lifetimes measured more in decades than years. Necessarily, the software, both on-line and off-line, required to analyse and control these projects has also grown in complexity and volume. The issues of how to build, manage and maintain large software projects are addressed by modern software engineering (SE) methods and techniques. Furthermore, SE methods will undoubtedly be essential to enable the integration of *both hardware and software* in future real-time systems.

The question of which programming language(s) should be used is not discussed by this working group. It should be stressed however, that interesting developments in such areas as automatic code generation and the like are not likely to be commercially provided for a Fortran environment with typical HEP memory managers and associated data structures.

We have no intention of recommending the precise tools and methods that should be used. This is a field which is rapidly evolving and it would be presumptuous to presuppose what might exist in 5 years time. We do, however, base many of our conclusions on the Structured Analysis/ Structured Design (SASD) methodology that, thanks to pioneering work by the Aleph collaboration, has gained a strong foothold in the HEP world, both at CERN and at other HEP laboratories.

In the following, we consider in some detail analysis and design methods and related tools which constitute the relatively recent major innovation in software engineering. We also discuss briefly the tools required at the coding and implementation level. This includes code management which, traditionally, has been the most widely used implementation tool.

13.1 Analysis and Design, Methods and Tools

Modern analysis and design methodologies provide systematic techniques for the analysis of a problem or project, together with design strategies for translation into software. The basic components of a typical method are:

- The notion of the life-cycle of software, or its existence between the inception and the end of its utilization. In analogy with a manufacturing process the following approximate steps exist ; specification of requirements, analysis of requirements, design, implementation, testing and maintenance.

- Abstraction or the concept of construction of a logical model without regard to physical implementation. This enables a clearer understanding of the problem at hand, followed by consideration of the practicalities of implementation at a later stage. Complex problems are broken down into subproblems and the latter subdivided until the resulting problems become intellectually manageable.
- Graphical methods to enable a thorough analysis and design of the project concurrently with the production of the documentation. Diagrammatical techniques, with the goal of partitioning the problem into manageable portions, are vastly superior to monolithic tomes of written prose.
- Data modelling techniques provide a clear understanding of the data, and may provide a path to physical data structures.
- A Data Dictionary consisting of specification of the data elements, their inter-relationships and descriptions. The data dictionary provides a central backbone of data definition and documentation for all the data throughout all phases of the software.
- Automated tools to enable manipulation, editing and verification of the various diagrams, specifications and data descriptions and their inter-relationships. Integration of all tools via a common data base is essential to be able to go back and forth between information relating to various phases of the project.
- Design criteria aimed at producing specifications for independent modules which may then be built by smaller teams. The final refinement of the design diagrams results in specifications not far removed from actual code, and indeed there are developments in the area of automatic code production.
- Management procedures to both ensure the quality and monitor the progress of the work in hand. At each stage of the life-cycle real deliverables enable a proper assessment of the project. Improved quality of the product, especially in the critical analysis and design phases, is a built in, rather than added on, feature.

In addition to the above, 'reverse' engineering tools are used to analyse code and produce diagrammatical descriptions in order to verify what the code really does. Given the addition of HEP to Fortran and other concepts such as memory managers, these tools have, in general, to be home grown.

13.1.1 Historical Background

The term 'software crisis' was coined in the 1960s as recognition that large software and system development was, in many instances, out of control. The various studies made of large projects indicated that a major contributor to this problem was the lack of a systematic approach to software design and development. The concept of 'software engineering' emerged with the idea of applying engineering-like techniques to the various stages of the software life cycle. With the advent of significant new ideas on Structured Design and Structured Analysis²¹ in the early 1970s together with the concept of the software life cycle, it became possible to formulate frameworks that incorporate the known techniques for software production and the appropriate management practices.

13.1.1.1 Situation in the HEP community

Whilst the concepts of software engineering have been embraced in particular by the aerospace and oil industries over that past decade, the start of any significant involvement in the HEP world can be traced to 1984. The workshop on "Software in High Energy Physics" in 1982 provided the HEP com-

²¹ Some literature: W.P. Stevens, G.J. Meyers, L.L. Constantine, *Structured Design*, IBM Systems Journal 1974; T. DeMarco, *Structured Analysis and System Specification*, Yourdon 1978; C. Gane, T. Sarson, *Structured Systems Analysis*, Prentice-Hall, 1979; G.J. Myers, *Composite/Structured Design*, Van Nostrand Reinhold, 1978; E. Yourdon, L.L. Constantine, *Structured Design*, Prentice-Hall, 1979.

munity with a chance to meet and discuss with people who not only advocated but had actually used these techniques. Several groups, notably the Aleph software group, expressed further interest and, with the aid of the results of several comparative studies Aleph selected SASD as a suitable methodology. This selection was based on the criteria that SASD supported all phases of the life-cycle, was widely used in industry, was relatively easy to learn and use and was itself supported by (at that time rather limited) automated computer tools. This view was confirmed by a tutorial on software design techniques in 1983.

Subsequently courses in SASD were given at CERN and towards the end of 1984 SASD was formally adopted by Aleph for both the on-line and off-line software.

In the following summary of the existing situation, SA refers to the analysis part of SASD and SD to the design part.

- Aleph : SA is used extensively for off-line and on-line data acquisition software and SD is used for certain subsets. Many extensions have been made by Aleph, sometimes because the early versions of the tools did not support required features such as, for instance, data modelling for which the ADAMO toolset²² was built. Support of the methodology and tools, including 'reverse' engineering tools such as program analysers (for example PRODES²³ is done by the group.
- SPS Control System : SASD has been used to analyse the existing controls software and design modifications for the SPS to become a "multicycling" machine. This work²⁴ comprises a major revision and upgrade (essentially a complete rewriting) of the large body of applications programs used for controlling the SPS.
- LEP Control system : Following the lead of Aleph and SPS Operations group, SASD techniques are used for creation of the applications programs to control LEP. The work is split into an analysis section (using SA), comprised of accelerator physicists and engineers, plus a design section (using SD) consisting of computer professionals.
- Fastbus system management : SASD methods have been used to make a system to automate the management and control of large Fastbus systems. This work²⁵ has been a collaboration between DD, Aleph and Delphi and is relevant to any Fastbus setup.
- Opal : SA is used for the off-line reconstruction program, although, as noted below, there is no official commitment by the collaboration to use these techniques.
- OBELIX : SA is used for a pilot project, with the intent to use it for off-line and data acquisition software.

There is currently no official CERN support for software engineering analysis and design methods and tools.

Partly as a result of the Aleph experiences, the FNAL experiment D0 adopted the same methodology for its software development. Subsequently FNAL has negotiated a license for use of a modern commercial system based on SASD. The first users of this tool are D0 and the FNAL data acquisition group. Consultancy support is eventually foreseen.

²² Reference: Qian Z. et al, *Use of the ADAMO Data Management System within Aleph*, Comp. Phys. Comm. 45 (1987), 283 - 298.

²³ Reference: J. Bunn, *PRODES, a tool to analyze and describe a Fortran Program*, ADAMO Application Note 3.

²⁴ Reference: R. Bailey, J. Ulander, I. Wilkie, *Experience with using the SASD Methodology for Production of Accelerator Control Software*, SPS/AOP/Note 87 - 16

²⁵ Described in: E.M. Rimmer, *A Database for Fastbus DD/87/23*, Oct. 1987

Towards the end of 1986 and following into 1987 courses on SASD were given to experimenters from the two HERA experiments, the HERA machine group and the DESY computer centre. It was decided to adopt SASD for both on-line (ZEUS) and off-line (ZEUS and H1). In November 1987, after evaluation and comparison of existing products, a common analysis and design tool²⁶ was selected, with licences negotiated by DESY and provided for the experimental groups. The ADAMO toolset is used.

13.2 Large Software Projects at CERN

In this chapter we attempt to summarize the experiences of large scale software production both with and without the use of software engineering methods. For the former we have looked at major experiments including Delphi, EMC, L3, Opal, UA1 and UA2. The software for the older experiments (EMC, UA1 and UA2) was written well before SE techniques became fashionable. Delphi and Opal use SE techniques to some degree, but the methods have not been officially adopted by the collaborations. In addition we reflect on the SPS controls software prior to 1985.

The conclusions on the use of SE techniques are based on the experiences of Aleph and, to a lesser extent those of ZEUS and H1 at HERA and D0 at FNAL, together with those of the SPS and LEP control systems. Experiences at the European Space Agency have provided specific insights into project management possibilities.

13.2.1 Software Projects that have not Used Formal SE Methods

The conclusions and experiences below are an average and not necessarily applicable to all projects.

13.2.1.1 Organizational Tools

Reliance is made on dynamic memory and data managers to organize data structures. The latter are in general well documented, both with text and (often) diagrams, and contain some elements of data dictionaries. Access to data banks is sometimes provided via a higher level package, which thus acts as a sort of data model. There are clear analogies with the SE principles outlined above.

13.2.1.2 The Problems

There exist general problems that do not specifically result from the lack of adoption of SE methods ; rather they reflect the attitude of the management of the project that software building is a second rate activity compared to, say, building detectors. This often results in no formal managerial structure for the software team which itself is often largely composed of young physicists who have little or no experience working on large software projects. Furthermore there is usually no means of obtaining suitable training. Reliance is made on consultancy provided by in-house experts who often themselves have received no formal training.

One of the main reasons why SE has not had a bigger impact on the more recent experiments is the lack of money for automated tools and, until recently, lack of good tools themselves.

²⁶ *CADRE/Teamwork*, by Cadre Technologies Inc, Providence, R.I. This tool has been selected by (amongst others) : FNAL, DESY for HERA, Cornell for CLEO and CERN for SPS controls.

The following is a summary of other problems which are addressed by SE methods and techniques.

- The lack of precise design specifications leads to wasted effort particularly when code writing is not centralized.
- Documentation is often poor and sometimes non-existent. Often the only way to know what is going on is to read the code.
- Due to a lack of suitable communication methods, it is often the case that software specialists are unable to get feedback and input from physicists not directly involved with the software. These same physicists feel cutoff from the software activities.
- Project management and progress assessment is difficult due to the lack of any management tools.
- Frequently, the people who write the code have left the experiment before the first real data has been processed. The resulting combination of poorly designed software with (often) minimal documentation results in serious maintenance problems.
- At the SPS, the previous approach to writing accelerator controls software consisted of providing a powerful interpretive environment where users were free to write their own programs. A decade of such practices resulted in diverse packages, often undocumented, that constituted a system which was very hard to maintain. A major upgrade, affecting nearly all sub-systems, was essentially unthinkable given the then current practices. An alternative approach to writing control systems software has been to rely on the accelerator expert explaining the requirements to programmers who translate directly into code. Experience has shown this is not the most efficient way to work.

13.2.2 Conclusions from Experiences with SE Methods

It should be emphasized from the outset that SE techniques are not a panacea. Rather, the use of these methods brings with it an awareness of the necessity for proper managerial structure and proper training. The adoption of a particular computer aided tool, however versatile and powerful, will do little to solve the problems of large scale software production without the necessary commitment to apply the systematic engineering-like concepts inherent in SE methodologies.

For SE to be effective, the project management must understand the principles, support the effort and be prepared to make serious commitments in money, for the necessary hardware and software tools, in time, for training and learning, and in manpower at the start of the project.

The methodologies comprise a wide selection of techniques and practices, and it is up to each project team to select the sub-set they feel comfortable with. Though the HEP environment differs from that in industry in terms of management structures, it is clear that it should be as unthinkable to work on a complex software system without following guidelines and having some level of expertise as it would for the non-expert to take a soldering iron to the inside of a Fastbus module in the hope of improving it.

To cope with the problem of high learning restrictions before physicists can start to work on large software projects, it should be recognized that not every one needs to be equally skilled. Alternatively, in this domain of specialized skills, a certain division of labour is required.

In the SPS controls software system upgrade the results have been quite positive with relatively straight forward design and coding following a proper analysis phase. The system performs well with very few changes needed in the requirements specification.

Use of SE methods works both in a traditional HEP environment, where there is a broad overlap between analysts/designers and those writing the code, as well as a more structured situation, such as

LEP controls, where accelerator physicists analyse the requirements and professional programmers perform program design and implementation.

The following is a summary of more specific conclusions:

- It is necessary to have formal training by experts who themselves should preferably be knowledgeable in the context of the particular project. After training, there has been a good acceptance of the methods, in particular by younger members of the community.
- Graphical analysis methods make systems easy to visualize and have been found to be very useful for communication and documentation between workers in the field and interested parties such as the project management or newcomers, who themselves have had no formal instruction. Data Flow Diagrams of Structured Analysis have proved to be a very good analysis tool and are often used in the broad design as well. Although the analysis and design is self documenting, discipline is required to keep the diagrams up to date in the light of subsequent modifications.
- Data modelling is of paramount importance, as it extends from the analysis phase down to code writing and data presentation thus providing continuity throughout the life cycle. For HEP the Entity Relationship (ER) model²⁷ has been found to be very useful. The mapping onto the relational (or tabular) model is very easy, making it convenient for data base design work. With suitable tools, such as those developed within Aleph, physical data structures can be built and documented automatically from a data definition language. Data may also be manipulated and validated. The corresponding graphical depictions of the data model are termed Entity-Relationship diagrams (ERD). For certain applications in information analysis and data base engineering, other data models are preferred.
- In the context of a complex multi-processor, multi-tasking, on-line environment, SA tools which enable modelling of control and real-time aspects have proved to be very useful in providing a clear understanding of the required functionality of the system²⁸. The advantages of problem analysis, partitioning into components and early documentation are as valid in this context as for the off-line. Of particular importance in this area are the newer additions to SASD i.e. State Transition diagrams, for modelling system response to stimuli, and Data Flow diagrams for modelling the data and flow of control.
- Automated computer tools for diagram manipulation, editing and verification are essential. Aleph has been restricted by the lack of widely available tools. An important feature in any such tool is that it be of an open nature rather than a closed system, thus permitting extensions to be built as required and to allow different tools to be used on the same project. Aleph have made important extensions, particularly in the areas of data modelling and 'reverse' SE products that analyse written code.
- Today the relevant hardware environment for use of SE tools is a workstation with its large graphics screen, window capabilities, adequate CPU power, text editor, compiler, etc.
- Although the analysis and design features have been successful, there are still gaps, or less precise methods, to go from design to the actual coding.
- The 'walk-through' or suggested method of analysis, design and code review has proved to be useful. The reviews of the analysis were usually done by groups, with each person in the group standing up in turn to show his work, and the rest looking for deficiencies. The temptation to correct the work rather than assess it was largely resisted, as this would waste time and offend the person who had done the work. Code reviews are time consuming and were carried out by assigning an individual to review a set of subroutines. Such reviews, plus the necessity to merge input from different sub-projects, require relatively frequent meetings of representatives from the

²⁷ P.P. Chen, *The Entity-Relationship Model - Towards a Unified View of Data*, ACM Transactions on Database Systems, Vol.1, No.1(1976), 9 - 39

²⁸ Reference: T. Charity, R. McClatchey, J. Harvey, *Use of Software Engineering Techniques in the Design of the Aleph Data Acquisition System*, Comp. Phys. Comm. 45 (1987) 433 - 441

entire software team.

- Fairly simple but effective techniques may be used to monitor the progress of the overall project. Time estimates are made for each sub-project and for the relevant milestones (whose definitions are agreed in advance). The study of the evolution of regular reassessments of these time estimates enables the project managers to spot delays and problems before they impact seriously on the overall project. Commercially available tools like MacProject (APPLE) are useful in this context.

To summarize, software engineering techniques have been successfully used in Aleph, accelerator controls and other CERN projects. The methods have been very helpful in providing a clear understanding of the problem at hand and as a means of communication of ideas within the groups both between experts and non-experts. Similar conclusions have been drawn by groups from FNAL, DESY and other HEP laboratories. In the past, the lack of suitable automated tools has been a restriction. This situation has now changed. Good tools are available and further improvements can be expected. The majority, if not all, of the groups mentioned in this report have selected the same SASD based tool (Teamwork) which is therefore becoming the de facto HEP standard.

13.2.3 Requirements of Outside Institutes

Software development does not require full-time presence at CERN and is therefore an area in which physicists in the outside institutes contribute to experiments. This is especially true of small groups in universities which lack the technical resources to make contributions to detector hardware. Some of these groups have local computing resources, such as medium sized mainframe computers or workstations, others rely on network access to remote computer centres at CERN or elsewhere. Some institutes have quite limited resources for buying new computer equipment and software.

Within the above constraints, the users in institutes outside CERN have much to gain from SE techniques. The tools for breaking large problems down into well defined parts which can be tackled independently are of particular benefit to the community outside CERN. Such subproblems can be taken on by individuals or small groups of users in the outside institutes. While the work is done independently, the common use of SE tools ensures consistency, so that the code which is produced should fit together correctly. At the level of a common data dictionary the different groups may maintain close communication. The diagrammatic tools of SE, such as data flow diagrams may also improve communication between groups at remote sites. However, as already noted above, SE tools will not remove the need for discussion, and regular meetings between contributors to a software project will continue to be important.

Training in SE methods and tools must be provided for users from outside CERN. Where possible this is probably best done nationally to save on substantial travel costs; this is easiest in countries with a large national laboratory, e.g. Rutherford Appleton Laboratory in the U.K.. In any case, courses generally only cater for a limited number of people (15 or so) for maximum benefit. Good documentation is clearly a necessity — both pedagogical and reference. The special problems of training students need to be considered. In general it would be sufficient that limited training be given so that they could use the results (e.g. of analysis and design) of SE techniques.

Standardization between the different HEP laboratories would be highly desirable since many institutes are involved in collaborations with more than one laboratory. Standardization with other users of national laboratories would also be useful, although probably harder to arrange. The benefit of this to the outside institutes is that costs of hardware, software licenses and expert support may be shared. A more general benefit is that physicists will not have to learn new tools when they change experiments.

At present, some of the best SE tools are available only on workstations which are still relatively expensive items; the software licenses can also be very expensive. Before committing the CERN community to a particular computer aided software engineering (CASE) package, the financial implications for the outside institutes must be fully considered. It would not be acceptable for HEP groups outside CERN to be excluded from contributing to the software effort due to the choice of an unnecessarily expensive set of tools. In view of this, it may be worth considering systems which are available at relatively low cost and which run on cheap or existing hardware.

The problems of maintaining a central database for SE tools will have to be faced. One of the key ideas in (some) SE methodologies is the data dictionary. This at least must be maintained centrally, although copies may be made for reference. The problems of maintaining such a database, with updates coming over wide area networks, must be addressed.

13.3 Other Coding Support Tools

The software engineering environment is completed by implementation and coding tools. In this area tools have already been widely used in the past, and experience exists. We present here a brief review of the more important tools and their functionality. No mention is made of standard compilers, and we also exclude Documentation and Data Base system tools, discussed in other parts of this report.

13.3.1 Source Code Management

Source code management means the use of a tool which helps to maintain different versions of a program (or other text file). Various tools have been in use by most experiments over many years. Several exist today and have shaped the habits of their respective user communities. No tool exists which clearly responds to the modern interactive workstation-oriented development environment.

An overview of the situation in the HEP community is the following:

- Patchy and Historian are currently the only available tools for those who need a portable code manager. Patchy will be used as long as there is someone to maintain it. Historian continues to be offered for those who need the features not available in Patchy. The list of such features could include binary program management in the near future.
- BASE is a recent portable product written to specifications of the HERA experiments, by a commercial software house. It targets a completely new type of source code manager. The product is presently being installed on DESY's IBM system. BASE clearly bears watching.
- CMZ is an interesting new extension of the conventional source code manager concept. It offers convenient interfaces to the operating system via a Unix-like command language. Though a private initiative of CERN-connected people and currently unfinished, it has attracted significant interest in the CERN community. The terms of availability of this product have not yet been defined.

There is no consensus concerning source code management requirements in the community. In view of the rapidly changing environment, new projects are not recommended in the area of code management tools. In the long term, it is possible that Unix will offer a sufficiently portable and friendly environment which may obviate the need for portable tools. Opinions differ considerably as to when and if this will happen.

13.3.2 Other Implementation Tools

In addition to the functionality traditionally offered by source code managers, other important tools exist that help in the development and maintenance of software projects. The following is a partial list of some of the more useful categories.

- *Syntax-Directed Editors* "understand" the syntax of the language. In addition they provide coding templates and (sometimes) partial compilation/execution capabilities. The best known such example is LSE (DEC) which, although not so useful for Fortran code generation via expansion, is very useful for ADA. It can be customized for completely different applications such as entry of DDL in the ADAMO syntax.
- *Static Analyzers* take source code as input and perform checks such as searching for undefined variable, inconsistent common block lengths and conventions not being followed. Other functions include the production of program calling trees with plots and checking for use of recommended library routines. Currently such tools have been developed at CERN and have been widely distributed.²⁹
- *Program Testing Tools* of which the most common and essential variety is (symbolic) debuggers, enable the monitoring and modified execution of programs with possibilities to display the source code as it is executed. The best example is probably the VAX/VMS debugger although IBM's IAD is functionally equivalent. *Path Coverage analysers*, such as PCA (DEC), show which program paths have been executed or not. Such tools are useful for performance analysis as well. *Regression testers*, such as DTM (DEC), attempt to ensure that programs under development run correctly by establishing benchmarks and then comparing with runs made with new developments or modifications.
- *Optimization tools*, such as PCA (VAX/VMS), SPY (CDC,IBM/MVS) and IAD (IBM/VM), enable the identification of parts of programs that consume the most CPU time. Optimization should, however, be done late in the development of a program, bearing in mind the probable loss of clarity of the code.

An important aspect in the choice of support tools is the selection of a coherent set within a common framework. There exist several ambitious projects³⁰ that attempt to create a global software environment consisting of an integrated suite of software utilities.

In order to improve efficiency of software production, programmers must change their habits to make increased use of such tools. This will require both learning time and financial investment. Effort is required to keep abreast of new developments in this rapidly changing field and also to provide local support for selected products.

13.4 Software Engineering: Recommendations

13.4.1 To Project Management

Large software production is a serious activity and deserves, from the start of the project, as much attention as detector building. Adequate software tools are required as well as the necessary computer hardware.

²⁹ Documentation: J. Bunn, *FLOPPY User's Guide*, CERN/DD/US/112; H. Grote, *FLOP : a Fortran Oriented Parser*, CERN/DD/US/13.

³⁰ The Toolpack Project from Numerical Algorithms group and PCTE : Portable Common Tool Environment (ESPRIT).

Software engineering tools and methodologies have been shown to be effective provided sufficient support is given by the project management. Use of such tools alone will solve little, unless the required managerial structures and practices are adopted such that very large collaborative efforts may be partitioned into smaller teams working in parallel.

Our choice is "pay now or pay lots more later".³¹ In this context later could mean never or at best only partially fulfilled physics goals from experiments employing vast manpower and financial resources.

13.4.2 Training

Training in the concepts and ideas of software engineering is essential and should be provided by (repeated) courses given by experts from HEP.

Realistically it should be recognized that, given the range of skills needed to design and construct large scale software, there has to be some division of labour within the project team. Different skills are required by the project managers, say, than by those working more at the implementation end.

A central compilation of the documented experiences from past and existing projects would provide the necessary complement and back-up to formal training.

13.4.3 Support and Consultancy

Support and consultancy is required on several, not necessarily independent levels.

A central support group should be set up to evaluate European HEP requirements and negotiate licences for the necessary tools. This same group would also act to provide feedback to and from the vendor for bugs, problems and improvement, and help with product installation.

This group should coordinate the use of SE tools across different projects and different HEP laboratories, and build and maintain extensions required by the HEP community.

A special interest group should be formed consisting of users of SE tools both at CERN and the outside institutes.

The functions of this body would be to provide communications between the groups involved, to decide on the common tools required and to make sure that where possible, progress may be made on common problems.

Day to day consulting on conceptual problems is best dealt with by expertise from the project itself. This requires an establishment of a base of expertise in each project.

³¹ From: J. Manzo, *On Managing Large Scale Projects : Some Simple Principles for Developing Complex Systems*, Comp. Phys. Comm. 45 (1987), 215 - 228

13.4.4 Software Tools

Automated computer tools are essential.

A powerful set of tools capable of supporting development of a variety of projects (batch oriented systems, real-time applications and data base modelling) should be selected with a view to standardization across the CERN and HEP community. The relevant hardware environment is the workstation.

Adequate, though not perfect, tools exist today. It should be remembered that like any other tool, they will evolve, and new ones will have to be obtained as requirements change.

13.4.5 Licences

CERN should negotiate site-wide or HEP-wide licences for groups working on experiments at CERN, and for other required applications.

13.4.6 Requirements of the Outside Laboratories

Software development will remain a strong, major activity of the outside institutions. When negotiating licences it will be important to keep in mind the limitations of hardware and financial resources of these institutes. Standardization of SE tools, both CERN-wide and HEP-wide is very relevant in this context. Training for outside users should be provided at CERN although not exclusively.

Chapter 14

SOFTWARE SUPPORT: GRAPHICS

The subject of graphics support was considered by a separate working group. We gratefully acknowledge the kind collaboration of P.Schilling and J.Olsson from DESY.

14.1 Graphics in HEP

14.1.1 Graphics Applications

Within the last years the range of graphics applications within HEP has become extremely large, from the simple display of histogrammed data to highly interactive dynamic event viewing. This growth in the kinds of application programs has been paralleled by the very fast development of graphics devices and workstations. The types of applications considered in this document are as follows:

- Detector design: CAD and physics simulation in order to optimize detector performance and save beam test time.
- Software development: Detector representation, interactive display of reconstructed events, debugging pattern recognition programs, etc.
- Experiment analysis: Event and detector viewing, physics parameter representation, multi-dimensional statistical analysis, histogramming. (Both in 'batch' as well as 'interactive' modes.)
- Publications: use of graphics displays with a graphics editor to produce figures for output onto high-quality hardcopy devices.
- Human interface for control applications.

With respect to the above list of applications we will concentrate essentially on software development and experimental analysis which will be, together with the publication and human interface area, the main use of graphics during the coming years of the LEP era.

14.1.2 Implications for Software and Hardware

The large range of devices available from numerous manufacturers (all with different native graphics interfaces), the large number of graphics applications mentioned above, the need to run programs on several types of computers (workstations and mainframes) inter-connected through networks, and the large number of physicists participating in the current and future experiments, are all factors which lead to the following requirements:

- transportability,
- device independence,
- common interfaces for on-line and off-line applications,

- metafiles to exchange graphics information at any level,
- access to multi-window environments,
- fast response times,
- good availability.

It is essential that a common approach to graphics and user interfaces is taken by programmers in different areas so that modules of code will work together in a compatible fashion.

The different layers of HEP graphics applications can be classified in three levels:

- Basic picture generation using primitives for lines, characters etc.
- Library functions necessary for physics analysis and event viewing e.g. Histogramming, jet drawing etc.
- Physics application interfaces.

The first 2 of these levels will be discussed at some length as the fundamental building blocks of all graphics applications systems.

Not only must applications be able to work at any one time on a range of graphics devices, but over the time-scale of a LEP experiment it will also be necessary to adapt to several generations of equipment.

Whatever choices are made for hardware and software standards for the first generation of LEP experiments, they must take into account the need for a smooth evolution of the applications programs.

14.2 Graphics Software and Interface Standards

14.2.1 Basic Picture Generation

The large number of graphics packages (Application Program Interfaces, or APIs) like WAND, GPR, GSR, Template, DI3000 etc. cover the full range of possible hardware in a non-transportable way. Thus, since it is not possible to standardize application programs to only one type of graphics machine, the 'natural' solution has to be provided by graphics standards. (Ref: Report on the Status of Graphics Standards, D.R. Myers, DD 88-14.) Of course, if use is made of features in a standard, which are not supported in hardware and hence have to be emulated, then the performance will suffer. Nevertheless, as an example, the software emulation of 3D displays, although slow on low-end graphics machines, does allow the development of 3D code on cheap hardware, and may also be sufficient for applications which require 3D but do not need real-time transformations.

As a minimum, the features required of a standard include the use of hardware primitives, characters, segments and real time transformations (when available). However, at the top end of the scale, a particular device may contain features not supported by the standards. It should be noted that the current standards, such as GKS and PHIGS, provide inquiry functions to allow one to find out at run time which features are available, and thus to decide which algorithm is best suited to the occasion.

Whilst using a graphics standard, such as GKS, for all new graphics devices can result in a significant loss of performance compared with what could be attainable utilizing the hardware features directly from the application program, to a large extent performance depends on how much effort is available when writing the device drivers. Performance is certainly the reason why some people using workstations prefer the native graphics package to any standard. However, there have been several

notable cases where poor performance was found to be due to the use of a standard in a completely inappropriate way.

Training and experience in the use of GKS (or any other standard) is essential to produce well written and efficient graphics applications, and CERN support in this area is necessary.

In the more complicated examples of HEP application programs, for example event viewing and analysis programs, the complex data structures within the application program are represented in the structures within the application code, for example in Zebra or Bos banks. Since a dynamic hierarchy of coordinate systems is not required it is sufficient to map the physics data structure to a flat graphics structure, such as a set of GKS segments. This also avoids the complexity of programming graphics data structure editing functions.

GKS does not provide explicit support for workstation 'window' systems. However, assuming that the workstation operating system provides a window manager to manipulate the windows, GKS does allow multiple windows to be used, with some limitations, and this feature has been implemented at the driver level for several existing workstations.

From the experience gained so far, GKS provides a very good standard low-level graphics application interface, giving a reasonable compromise between portability and performance.

14.2.2 Basic Library Functions

However, as GKS is by definition a 'kernel' system, it is recommended that a graphics library of higher-level functions should be developed, based on what has been done in various experiments over the previous two years. The working party believes that the supported basic library should be layered on top of GKS, and this can be used as a starting point for those who wish to adapt the library to non-standard picture generation software packages. This library will include dedicated HEP primitives, some of which are described as graphic macro-primitives in the Higz manual, as well as features such as histogram production, helix drawing, jet drawing, etc. This would help to prepare a smooth transition of the existing code toward new hardware and graphics standards which might be used in future experiments.

We recommend that a set of high-level graphics tools and functions based on GKS of particular use in HEP should be provided as part of the CERN Program Library. They may be written by groups or individuals throughout HEP, but the work should be coordinated at CERN, which should also provide long-term support.

14.2.3 Distributed Graphics

Interactive graphics will run increasingly in a distributed environment, hence it is essential that we gain practical experience in X-Windows implementations, and thus be able to give sensible advice to those in the community wishing to acquire software products to support distributed graphics. As X-Windows is a very low-level protocol, it will also be necessary to evaluate the various tools available layered on top (DECWindows, Open Dialog, etc.) and to choose one, or at most two of these as the CERN supported HEP interface to window systems. Moreover, X-Windows is not a graphics package, so an important area to explore will be the co-existence of X-Windows with layered graphics packages (perhaps using the X-Windows extension called PEX).

Also we have to solve the problem of communication from local interactive graphics programs with remote databases containing thousands of histograms, DSTs and raw data, as well as to define a framework open to the evolution of the hardware to new systems (Unix) and to new languages (C, PROLOG...).

We believe that networked graphics, where all or part of a graphics application runs in a host remote from the graphics workstation, will be of increasing importance during the period addressed by this report. We recommend that CERN effort should be put into gaining experience of windowing implementations, with a view to selecting one or two as supported interfaces for the HEP community.

With the introduction of the CRAY running UNICOS, Unix has made a definite appearance in HEP, and we expect to see expanding use of this operating system. The CERN libraries are supported already on Unix, and one can foresee the use of other languages which complement Fortran in the graphics area. The use of Unix would certainly ease the use and support of a wider range of workstations, including those from SUN, Silicon Graphics, Ardent, etc. Our goal is to provide more choice via the use of open systems.

14.2.4 Document Processing

Another area which is still developing is the integration of graphics with documents. The signs are that CGM will emerge as the accepted picture storage medium for such applications. However as many hard-copy device manufacturers are providing built-in Postscript interfaces, we should not ignore this de-facto standard. In addition, the 'appendix E 2D metafile format' of GKS is now an acceptable standard, but the lack of useful 3D Metafile standard must be corrected.

14.3 Graphics Hardware

It is acknowledged that many of the major institutes in the large collaborations plan to use at least one 3D workstation and several 2D workstations in addition to normal terminal utilization for software development and experimental analysis.

The range of possible graphics hardware is very wide, and is evolving at a rate much faster than the progress in the preparation and analysis of experiments and the lifetime of software packages. Applications exist needing high-resolution, 2D and 3D, colours or grey levels (sometimes with a palette of 256 colours for smooth shading), and high-resolution hard copy facilities. For the coming LEP experiments wire-frame representations seem to be adequate, but we believe that for future experiments (e.g. at LHC) we may well be interested in solid modeling techniques. Nevertheless, there are many less-sophisticated applications, and so hardware with a range of performance levels is required. In most cases we need full interaction with the screen, both to change dynamically the aspect of the graphical representation, and to select physics quantities by 'picking' their graphical representations.

At both ends of the performance range there is a need to use simple graphics terminals on mainframes for program development, as well as 3D graphics machines with hardware processors to do real time transformations. Although the trend is clearly to replace terminals by workstations, due to the life-time of the range of machines currently existing or being bought, terminals will still be in use during the period covered by this report. We have also found that workstations currently available with a typical performance of one MIPS are not sufficiently powerful, either for compilation and linking of large programs, or for good graphics performance. Fortunately, the trend continues to be towards faster machines, and affordable devices are appearing with 4 to 10 MIPS. This is already sufficient for reasonable 3D performance for many applications which do not need smooth transformations in

real-time.

At the high-end, CPUs are approaching the speed necessary to perform dynamic rotations of 3D wire frame objects without special transformation units. This has interesting implications for the utility of display lists, as their usefulness is less obvious for machines at this level. Although ultra high-level graphics performance is becoming available in the some workstations, we are more interested in enhanced CPU power, since our work concerns CPU intensive interactive computing using complex data structures rather than computer design applications.

Already in the 'MUSCLE' report, as well as in another chapter of the present work, networking between mainframes and workstations or clusters of workstations has been identified as being an important issue. We certainly endorse that conclusion. Tools which optimize file transfer and the transmission of graphics commands over heterogeneous networks will be of the utmost importance for graphics applications. Whilst it might be necessary to group workstations (perhaps of a similar type) in clusters with a database server, the ideal is clearly to be able to work in a heterogeneous environment with the minimum number of constraints.

We recommend that CERN coordinate the definition of application-oriented benchmarks for 2D and 3D devices, these being based on GKS 2D and 3D code for the basic picture generation. In addition CERN should, in cooperation with the institutes, perform benchmark evaluations.

14.4 Graphics: General Conclusions

We endorse the choice of GKS as the graphics standard for HEP for the coming LEP experiments. GKS will exist and be developed for more than 5 years, and new hardware capable of running PHIGS can certainly support drivers for GKS-3D, although these might need to be written or commissioned. The use of GKS-3D has the important advantage that 2D modules of code may be used within 3D applications without modification, and that application programmers do not need to learn a second interface standard. However, during the period under consideration it will be necessary to monitor developments in the field of graphics standards, especially as a new generation of standards is scheduled to appear during this time.

CERN should provide support for the basic graphics packages and be responsible to ensure that they run on a range of equipment, as well as for distribution and documentation. CERN should also evaluate and test new terminals and drivers. It is essential that sufficient manpower with the necessary technical knowledge be made available to fulfil this task.

Application programs used in big collaborations are written in several places using a range of hardware. Once developed, these programs are used by physicists who are not software experts. The support to maintain the applications and test them in new environments must be provided by named people within the experimental groups. If these people can be located at CERN, so much the better.

In order to facilitate the communication necessary for decision making we recommend that an HEP Graphics Steering Group be formed with representation from CERN, collaborations, institutes and major computer centres associated with HEP graphics processing. It may be sensible to look for institute representation from country working groups where they exist, since such groups cut across boundaries between collaborations in a positive manner.

Chapter 15

SOFTWARE SUPPORT: USER TRAINING

15.1 Introduction

This topic was considered by a separate subgroup, the work of which has proceeded primarily via discussions with working experimental physicists and programmers. There have been no questionnaires distributed to the HEP community and no response has been received from the implicit invitation to contribute in the CERN computer newsletter.

CERN provides major central computing facilities for experimental physics. There is, in general, a small overlap between the activities of the professional staff running the central computing facility so their training should come from commercial sources. Often CERN staff are involved in the development of system software in collaboration with the manufacturers and this contact takes on the role of training. Similarly, clerical and technical staff are encouraged to attend courses organized by the Technical Training Committee. Having ascertained that the Education Service has no surplus manpower, terminals, workstations or conference rooms, we have made no in depth study of this area.

15.2 The Physicist's Demands

We found that the majority of physicists rely heavily on the documentation and working examples of the libraries and program products which they use. Many have acquired new skills through reading books, guides and manuals.

There was a wide range of responses to the idea of receiving formal training. Some physicists were prepared to attend a course which suited their needs; most supported the idea that training should be available but found it difficult to fit it into their own work schedules and a few suggested that if they needed to attend a course to use a package then this alone was proof that the package and its documentation were inadequate.

15.2.1 Documentation and Examples.

Most physicists drew attention to their need for greater clarity, simplification and general improvements of the user guides, reference manuals and other documentation of the libraries and program packages they use. Two detailed points were that working examples should work and there should be more simple examples in the documentation of the major packages currently in use at CERN.

The online FIND and HELP facilities were generally appreciated as was the assistance which may be obtained via the User Consultancy Office (immediate help with simple problems and requests for information, distribution of user manuals and access to reference manuals, and the transfer of difficult problems to the best available expert). There was however some concern, expressed particularly by physicists making occasional visits to CERN, that even online documentation may be out of date and examples do not work because they are not fully maintained in the evolving CERN computing envi-

ronment.

In general, physicists working outside CERN were content to receive printed copies of documentation from their experiment and from major CERN products by air mail, and corrections by telefax. A few expressed a preference for a service, to deliver documents and updated pages over the electronic mail networks. This is difficult to achieve because the printing and publishing facilities on the Central CERN computers are incompatible with those in the HEP community as a whole.

Everybody, from the worst hackers to the most talented and prolific producers of the routines and packages used by experimental physicists found it difficult and time consuming to produce and maintain documentation in English. The experts, with their familiarity and detailed knowledge of their working environment, also find it difficult to appreciate the needs of the end user.

To respond to the expressed needs of producers and consumers we recommend that a qualified and experienced technical writer should be recruited.

There are a large number of creative people at CERN it is not too surprising that it has often been a long and arduous process to reach agreement on a standard such as common procedures for the production of technical notes, papers and other forms of documentation. However, given the considerable dissatisfaction with the documentation within experiments and the manuals for many CERN packages, we would expect that the authors and managers of code should approach any new initiatives on the structuring of documentation with an open mind and a spirit of flexible co-operation on their part.

15.2.2 Formal Training Courses

Training courses for physics-computing lasting two to five days are given approximately once a month at CERN. The requests to attend these courses exceeded the number of places, especially for those courses with a large component of practical work using terminals and workstations.

We found several cases where physicists had not reached the point of applying to attend a course. For example, there were very important activities at CERN which took place during the course and the course was given some months before the new skills would be put into use and the knowledge would not be retained. The pressure to produce results, for the design and installation of an experiment or the publication of results at conferences, inhibits attempts to take the time to learn new techniques which might lead to more efficient working practices.

The demand to attend courses already exceeds the capacity and may be assumed to rise if there were better access and a choice of dates and extra subjects were introduced.

The studies of data-bases, software engineering and graphics in this report identify several areas where (200) experts within experiments should receive training and, with the introduction of the personal workstation, many of the (2000) physicists at CERN will also wish to learn new techniques.

Therefore we recommend that training courses and tutorials should be given repeatedly

For subjects given by visitors to CERN, or for physicists attending commercially run courses suited to CERN's requirements then the number of people receiving training will, most likely, be determined by financial constraints.

For topics in which there is already expertise at CERN, or a CERN based package, the people writing or maintaining the package often prepare and present the courses. This does give them some useful feedback on the package itself.

Once a course on a topic becomes stable and mature, or the level of demand requires the course is given many times a year, then the author's time is not spent most effectively in repeating the course.

Resources should be provided to maintain an appropriate repetition rate.

The courses may be given by trained demonstrators, the course may be recorded and video training introduced, and eventually a large component of the training may be given through computer based learning techniques.

As courses are given more frequently they may also become more modular.

The tutor does not have to cram in all topics of interest to people with differing levels of ability into the only one week course which will be given on that topic for two to three years.

Pre-requisite skills may be defined (for example the ability to use a workstation, a knowledge of the languages used in practical work and units of related courses) and units suitable for the technical expert in a collaboration may be missed by the end user physicist.

HEP regional centres should consider providing host facilities for their user community to attend live courses given on CERN based products by visiting CERN staff.

15.3 New Working Practices for Physicists

There was a strong feeling within the working groups of this report that resources should be allocated to educate the HEP community to exploit the advances of software engineering and computing technology.

It has been difficult to persuade busy, self-motivated physicists that the investment of their time in learning new techniques may lead to improved productivity (or faster physics). Indeed, there are physicists who have made almost daily use of the keyboards of teletypes, punch card machines and computer terminals for over a decade have not found the time to learn to type efficiently.

We have seen the start of the widespread introduction of personal workstations into offices at CERN, we expect 1000's by 1992. Thousands of physicists will use the CERN based Physics Analysis Workstation (PAW) package for interactive data analysis.

15.3.1 Personal Workstations

The estimated rate of the introduction of workstations leads us to believe that there will be one or two new users of workstations per day at CERN.

The new user, and visitors to CERN, should be able to sit with a relative expert for 2 to 3 hours and learn the basic use of the workstation, for example screen management, the help system, the file system and editing techniques. Other material could be divided into modules such as an introduction to the use of interactive debuggers and the available software engineering tools; electronic mail, text processing and graphical data presentation; and remote login (terminal emulation), job submission and file transfer.

We recommend that there should be an introductory session for new users of workstations.

There may be a demand for several such sessions each week. This has proved to be a very efficient way for a physicist getting started within the workstation clusters of large experiments but it is a workload which should not be carried by other physicists.

15.3.2 PAW – The Physics Analysis Workstation.

The Physics Analysis Workstation project (PAW) is an attempt to provide interactive data analysis tools across a wide range of mainframe computers, terminals and workstations. The project is strongly supported by CERN based experiments and the software has recently been generally released for user feedback. In short, PAW may become the interactive data analysis tool for thousands of physicist over the next few years.

Physicists outside CERN given the manuals of the PAW suite of programs (and told by their local system manager that PAW is now available) have reported that it is difficult to get started. Although the product itself could be easier to use, there is clearly scope for improvements to the manuals. We repeat that the introduction of more simple examples is a frequent request and the manuals of the PAW suite would benefit from the attention of a technical writer.

Courses given at CERN on the use of PAW have been oversubscribed and well attended. The excess requests from one course may completely fill the next course in three to six months time. The demand for an introduction to PAW is expected to rise as the LEP physicists move from installation to data analysis and as PAW becomes more mature.

The policies and priorities for funding within many institutes make it difficult for many physicists involved in software development and data analysis to come to CERN purely to attend a PAW course.

PAW is an ideal candidate for the production of video training cassettes at CERN and for the courses with personal contact to be given by "demonstrators" rather than the authors of the component sub-systems.

15.3.3 Software Topics.

The experts within experiments are aware of the pioneering work of Aleph in the use of software engineering techniques. Many recognize the shortcomings of traditional methods but now wish to concentrate on the physics analysis aspects of their experiment rather than start to make the significant investment to adopt formal methods within a collaboration after a few hundred thousand lines of code have been written.

It should not be left to the "next" experiments before some training in the use of software engineering techniques is introduced at CERN. Young physicists should have the opportunity to be trained to use these techniques as they will be used in industry or future experiments where they may eventually find employment. New projects could be undertaking using the techniques and packages which are available today; for example the data processing experts in the computer centre and the Lep experiments could co-operate in the design and implementation of their tape handling utilities.

We considered the possibility that the adoption of formal specification methods would allow physicists to concentrate on the development of algorithms and analysis and use trained programmers to implement the code. This proved to be very unpopular when suggested to working physicists.

Training in data modelling would also lead to an awareness of the potential use of data bases. We have found that there is a significant acceptance of SQL as a language with which to make data-base enquiries for book-keeping activities (during detector construction).

High level physics analysis languages are emerging as the LEP experiments draw near to data taking. These languages reduce the work to be done by a physicist analysing data but an expert may be needed when features are missing from the language. They give many advantages when comparisons of analyses are made as many cuts and definitions are defined within the package. At present their use seems likely to be private to one experiment and its own training resources. In the past such packages (like macro SUMX) did not gain a widespread user acceptance. It is considered to be a strength of PAW that user actions can be taken via interpreted Fortran files (via COMIS).

15.4 A Central Workstation Training Facility

We are entering an era where modern software working practices may be introduced into HEP work. The techniques which physicists and programmers should learn are discussed earlier in this report. Some subjects may be taught to a large audience in the CERN lecture theatres using traditional classroom techniques. However the bit mapped personal workstation provides the best user interface to the software tools which will become part of everyday working life. A large proportion of training activities should take place where the physicists may learn through relevant exercises on workstations.

We recommend that the provision and operation of workstations and terminals for hands on training and demonstrations should be a central responsibility.

A 'divisional workstation training office' in the computer centre has recently been established. This has five Apollo workstations, three VaxStations and two Pericom Graphic terminals. This gives a reasonable probability that a person attending a course may choose to use the same type of terminal as they use in their own work. A class size is limited to around twenty by the number of available terminals with two people sharing one keyboard. This also seems a reasonable limit for the personal attention of a tutor and the size and air conditioning of the office.

In round numbers, if this facility were to be used for training courses for five days each week and fifty weeks of the year then 2500 members of the CERN community, working on the frontiers of modern technologies, could each receive 16 hours training per year.

There may be a demand to enlarge this facility or to provide further units of a similar size.

The workstations have also been made available to experimental groups for private demonstrations of their own analysis packages and for evaluation of workstations in general. The processing power of the workstations may be exploited by appropriate processing tasks overnight and as a low priority activity during the hands-on sessions.

15.5 Video and Computer-based Training

Video and computer based training have the advantage that the student may work at his/her own pace, at any convenient time, without the presence of an expert or teacher.

There is already some experience of the use of video training at CERN and in the production of video training cassettes. CERN computer centre staff have received video training in some basic VM operations and many find this an acceptable learning process. Some of these videos were recorded

specifically for HEP use³² and videos have been recorded at CERN for the safety training of the members of each of the Lep experiments.

There should be a library of video and computer based training courses at CERN. There should further be a public facility for a single user, or a small group, to use this material.

Video and computer based training courses are available commercially for operating systems, programming languages, data bases, word processing and so on. A selection of subjects from a catalogue of 80 courses³³ is tabulated. The videos runs for 20 to 45 minutes and the list price is about 500 sfr (200 ukl).

Table 4: Selection of video course modules

Title of Module	cassettes
Unix Executive Perspective	1
Unix Overview	6
Unix Fundamentals	15
Unix Shell	14
Unix Vi Editor	4
Unix System V internals	8
Using VAX/VMS	7
VAX/VMS for programmers	11
'C' language programming	16
Advanced 'C' language programming	3
Understanding database management	3
Introduction to ORACLE (2hrs)	2
Informix SQL (2hrs)	2

These courses are typically available without charge for one or two weeks and an evaluation exercise may be mounted using video recorders and monitors already at CERN. An opportunity is to evaluate a Unix course occurs when the CERN Apollo domain support moves from Aegis to a the Unix version of the operating system.

Video cassettes and computer based training material may be exchanged between CERN, the regional centres and individual educational institutes of the HEP community.

A multi-standard video recorder and monitor might complement the workstation clusters of major experiments.

³² B.White visiting CERN from SLAC

³³ Computer Technology Group, Fox & Elder, U.K.

Chapter 16

TECHNOLOGY PREDICTIONS FOR THE 1990s

16.1 Introduction

This chapter covers, from a technology standpoint, the three major aspects of the computational process, namely: the CPU, the data storage and the provision for data transport (networking). It is not intended to cover all research activities, but to investigate those which appear relevant to high energy physics needs in the 1990-1995 timescale.

16.2 CPU Technology

This section looks at the CPU technology that will be available in the next few years. It appears that there will not be a clear division of labour as in the traditional data processing centre. Rather, a functional co-operation between many different classes of CPU will be the mesh on which physicists will perform their data processing.

The following sections broadly cut the CPU spectrum into 2 classes. The "high end" meaning high performance, expensive CPUs in relatively small numbers, and "low end" meaning cheap, plentiful machines. The reason for these broad categories stems from the over use and now hazy distinction between PC-s, workstations, minis, mainframes and even supercomputers.

When considering a given class of CPU, it is worth remembering that scalar, vector memory and I/O must all be in balance. They must be scaled to each other to deliver the highest possible speed without system bottlenecks.

16.2.1 High-End Processors

Mainframes and supercomputers fall into this category. There seems to be no problem in doubling the basic power of 1 CPU every 2-3 years for these machines. The emergent gallium arsenide technology promises to make provision for smaller cycle times and higher gate switching speeds, although commercial exploitation of this technology is only now just beginning. This should eventually lead to higher basic CPU engines. For the future (>10years), warm superconductors promise a further increase in switching speeds.

Manufacturers are also exploiting some parallelism in the current high end machines at the operating system level. The IBM 3090 with six processors, the VAX 8840 with four and the CRAY Y-MP with eight. It seems likely that this trend will continue as a way of providing ever more powerful general purpose machines from tightly coupled multi processors.

Vector processing is now available on some IBM and CRAY machines to provide performance enhancement at the user algorithm level. It is to be expected that both IBM and CRAY will upgrade their machines, a 16 gigaflop CRAY-3 and a 64 processor CRAY-4 are rumoured, whilst IBM have

several parallel supercomputers under development.

Despite disagreement over architectures, there is no doubt that the design and manufacturing of parallel systems is a growing industry. Much of this growth has been spurred by the demand for cheaper supercomputers which has created the "crayette" or mini-supercomputer market. The "preferred" architecture and methodology for exploiting such architectures has not emerged to date.

For HEP, exploitation of parallelism at the algorithm level would not seem to be a general solution to the computing needs. Although applicable, and highly beneficial in some cases, it would appear that running the same algorithm many times in parallel would map better onto the bulk processing requirements.

16.2.2 Low-End Processors

The evolution and development of VLSI technology, together with advances in computer architecture have led in recent years to an upsurge in popularity of the RISC architecture. As well as specialized architectures such as the INMOS transputer. The potential here is for large numbers, of cheap, high performance, self contained processing elements that can either co-operate in the processing of data, or operate independently on data. The latter, massively parallel case, is of great interest to HEP. The emergent Motorola 88000 microprocessor claims a performance equivalent to a single processor of a VAX 8800 style machine.

The workstation market has already adopted at its top end high performance RISC machines capable of providing substantial CPU for the analysis of physics events. For example, the APOLLO DN10000, recently announced, has a multiprocessor RISC architecture that claims a performance of 50-100 times that of a VAX 11/780 style machine. There are indications that probably all workstation manufacturers will have machines of comparable power by 1991. The concept of the emulator farm for processing IBM compatible object code images in parallel has been well accepted. Similarly, there is a growing interest in using many workstations, networked as a parallel computing engine.

The growth in higher speed VLSI processors is expected to continue for the foreseeable future. There will continue to emerge machines based on multiple processor architectures that will attempt to exploit the parallel processing functionality with more or less success. Of immediate interest to HEP during the next 5-10 years, will be the availability of such processors to create specialized systems for parallel data analysis.

16.3 Data STORAGE Technology

This section looks at the current technology trends in mass storage devices and to relate that to the future expectations of the HEP community. Firstly, it is appropriate to review the current situation *vis-a-vis* mass storage usage in the CERN central computer centre today. In particular, appendix A gives the current and expected figures for the end of 1988.

The data processing stream begins at the experiment. It is here raw data will be captured. Processing of these data will produce the summary data that will subsequently be reworked many times. Hence, decisions concerning technologies for mass storage begin at the site of the experiment. We will consider in detail, however, the treatment of these summary data. The model that is adopted here concerning future directions in mass storage management will be the centralized one. That is to say a relatively small number of mass storage data centres providing support for bulk mass storage media. However, any long term mass subsystem must be able to provide connectivity to a wide variety of hosts (including workstations) in a heterogeneous environment. The following points indicate some of

the key facilities demanded of centralized mass storage facilities:

- Simultaneous access to shared storage by multiple users.
- Transparency of storage capacity to the end application.
- Transparent file and record level access to online and offline data.
- Automatic, system managed storage to minimize overheads.
- Ability to store all types of digital information (bitmaps, images etc)

Despite the strategic goals required to satisfy the mass storage access requirements in the 1990s, one of the problems facing the HEP community is that of storage device density. In this context, the following sections review briefly the current status and expectations of the storage devices likely to play a role in the following 5 and in some cases 10 years.

16.3.1 Disks

Evidence from research laboratories indicate that the magnetic storage technology still has a long potential growth. Magnetic storage devices have, on average, increased density at a rate of 25% per year. Higher density, higher bandwidth to the CPUs and higher rotation speeds in order to improve seek/read/write time are essential in order to keep pace with the increasing CPU bandwidth. Experimental magnetic disks have been made with a track width of 20 millionths of an inch. This would potentially give a 3 1/2 inch diameter disk a storage capacity of 10 gigabits, around 50 times the current density.

16.3.2 Solid State Devices

The solid state disk is a device that appears to the CPU as a real disk but with a greatly improved performance. Today, the unit cost of solid state disks is a factor of 10-100 greater than disk technology. However, recent solid state subsystems use the 1 Megabit chip and development labs have 4 and 16 Megabit chips available. In 5 years time, with 16 Megabit chips at prices close to today's 1 Megabit devices the price gap between solid state and magnetic disks may close considerably. The advantages of such devices in terms of improved bandwidth, reduced latency and high reliability would make them attractive, as magnetic media replacements in some applications, in the 5 to 10 year timescale.

16.3.3 Tape/Cartridge

Today the density of a cartridge tape unit as attached to the IBM systems at CERN achieves capacity of approximately 0.2 Gigabytes. Given the data rates proposed by LEP, and collider experiments, a major logistical problem in cartridge handling will result unless a significant factor in storage density be achieved in the next 5 years. One might already expect to see a factor 4 increase forthcoming in the short 1-2 year timescale. From the technological point of view, the helical storage devices based on the video cassette technology can already today achieve a storage capacity of roughly 2 Gigabytes. An experiment taking roughly 35000 cartridges/year at today's densities might (hopefully) expect this to reduce to 9000 in a short timescale (1-2 years) and down to 3000 in a longer timescale (5+ years)

There seems to be enough life left in the cartridge tape technology to achieve a substantial increase in density. However this will not help the record level access required to some datasets, as might be expected in a distributed processing environment. This will lead to an increasing demand for tape stag-

ing and online disk storage to aid this serial access technology.

16.3.4 Optical Storage

The promise of the optical storage media, namely, high density, random access, multiple read/write capability and (potentially) high bandwidth has not been fulfilled in recent years. However the write once/read many (WORM) technology has become very popular in the text storage and retrieval business. Typically this offers today a storage density of 0.3-0.4 Gigabytes/side, data transfer speeds of 0.5-0.7 Megabytes/second and access times of 100-500 milliseconds. The next generation of devices (available in the 1-2year timescale) should offer 1 Gigabyte/side, transfer speeds of 1-1.5 Megabytes/second and access times reduced to 50 milliseconds. If such performance and density becomes cost-effective, this would appear to be a natural choice for the raw data taking which is only written once. Subsequent processing of these raw data could profit from the random access capabilities of the optical disk.

By far the most interesting developments will in the erasable optical disk units. There are today many companies working on this technology and although some units will probably become available in the 1year timescale, it will take some time before these units have the price/performance/reliability that will make them interesting. The initial units will probably offer densities comparable with that of today's cartridges (0.2 Gigabytes/side) and rather slow transfer speeds of 0.2-0.4 Megabytes/sec. However if in the 5 year timescale the technology matures to a density of 2-5 Gigabytes/side and transfer speeds in the 3-5 Megabytes/second (equivalent to the maximum throughput of today's IBM channel) then they may well become the preferred storage media for HEP.

16.3.5 Digital Audio Tape Technology

It is perhaps appropriate to inject some quantitative comments on the 'low end' of the new technology coming. The digital audio tape, (DAT), like the compact disk, has also found use in the PC and workstation market. A 120 minute DAT offers a capacity of 1 Gigabyte with a maximum access time of around 80 seconds.

For the transportation and subsequent analysis of data by PC-s this medium will probably become cost-effective in the 1-5 year timescale for the individual, especially considering the falling cost expected of the DAT units. This may open up new opportunities for data analysis. With the increasing power of PC-s and workstations and the large volumes of data transportable by DAT, 'home analysis' could become very attractive especially considering that running reconstruction programs on a PC should not be a problem in this timescale.

16.3.6 Exotic Technology

Finally, in the research labs there are some developments that bode well for the future 5-10 year timescale. One such idea is the 'digital paper' concept. This is a flexible optical recording medium with the sensitive surface coated onto a flexible polyester substrate. This is today equivalent to the WORM media. The physical characteristics make it suitable to be wound on to cartridges as does not have the rigid structure of today's optical disk.

Drives employ the standard 780-830nm lasers used for the rigid disk technology. The projected capacity would be 600 Gigabyte to 1 terabyte, for a 2400 foot tape, and the life of such a tape is quoted today at 15 years. Experimental tape drive units, have apparently achieved 3 Megabyte/second transfer speeds and an average access time of 28seconds to any byte on the tape. The packaging was a standard 12 inch tape reel. Although it may be many years before commercial products based on this medium

become available, the data density, and transfer speeds seem to be encouraging for the future. Such technologies are being driven by the industries that, like us, have the requirement to store huge volumes of data for long periods (10 years or more).

16.3.7 Summary of Data Storage

From the above technology review, and bearing in mind the volumes of data to be taken in the LEP, post-LEP and possibly LHC era a major problem of data management exists. Since 1978, according to IBM, user demand for MIPS and associated storage have grown at an exponential rate. From 1.5 Gigabyte/MIPS in 1978, through 4 Gigabyte/MIPS in 1984 to a projected 12 Gigabyte/MIPS in 1990. With the volumes of data expected for the future generations of experiments, automated handling of mass storage resources is a necessity. Indeed, the era of distributed processing attempts to relieve the user of the knowledge of specific devices and hardware configuration. We are starting to see the building blocks of automated storage management such as the cartridge robot and the optical disk juke box, with software tools such as the IBM Data Facility Storage Management Subsystem (DFSMS). Such tools, aimed at a global system strategy have been slow to come, therefore we must be careful to invest in system managed storage tools as they become available in order to keep pace with the storage requirements of the HEP community.

A totally heterogeneous environment of system managed storage does not seem feasible until overall system integration is tackled at a co-operative level between all manufacturers.

16.4 Data Transport

The transport of data to and from the processing units and storage in a distributed environment becomes one of the mechanisms that must have a bandwidth matched to the rate of data consumption and production. The sections that follow review the technology both for local "campus" style networking, and for wide area networking. are referred to explicitly where relevant.

16.4.1 Local Area Network Capabilities

Ethernet is the preferred local area medium today at CERN. Its wide acceptance in the industry has made it attractive for connecting a large spectrum of heterogeneous equipment. However, by 1995, Ethernet (10 Megabits/second) will be regarded as a well-established, old-fashioned technology. Many Ethernet segments will be saturated by distributed computing traffic. Many segments will be unreliable due to aging problems. The IBM Token Ring will most likely be in a similar situation.

The "normal" LAN for serious distributed computing will by then be FDDI (100 Megabits/second). A typical workstation or mainframe will be able to achieve a burst throughput at least 1 Megabyte/second on FDDI and FDDI costs will be comparable to today's Ethernet costs.

On the CERN site, the bottlenecks will be

- the interconnections between different FDDI networks (say limited to 20000 packets/second or 20 Megabits/second).
- the mainframe interface (say 1 Megabyte/second as above).

1 gigabit/second LAN technology will be within sight but not yet available. As today's unsolved problem is Ethernet-FDDI interconnection, the 1995 problem will be connection between FDDI and a 1

gigabit/second backbone.

On-site ISDN (at 64 kilobits/second) will be available if we want it, but will be too slow to be interesting for scientific users. It is unclear whether office automation users will still be interested by such low speeds at relatively high cost.

The number of network nodes will be so high (4000+ at CERN) that configuration control and operational management will be a very major activity, comparable to running today's computer centre.

On-line data recording, if required, will not be able to escape the mainframe bottlenecks. Hence, as today, transmission bandwidth will not be a blocking factor.

16.4.2 Wide Area Network Capabilities

By 1995, 64 kilobits/second lines will be regarded as slow and their cost will not be an issue. Basic rate ISDN (64 kilobits/second) will be in use to supplement traditional public X.25 as a transport mechanism between small sites and for small sites to access major sites.

By 1995, 2 Megabits/second lines between major centres will be regarded as normal and probably inadequate (indeed, NSFnet in the USA is already installing a mesh of 1.5 Megabits/second lines). At present international tariffs, HEP collectively would have to pay at least 11 MSF/year for a set of ten 2 Megabits/second lines in Europe, not including a transatlantic line. An approximate calculation shows that a comparable set of lines in the USA, where prices are set by market forces, would cost only about 1.5 MSF/year.

In the USA, there is reasonable hope that by 1995, trans-continental capacity in the 1 gigabit/second range will be available to the research community in general, with Federal funding, and presumably American HEP will be using it.

In Europe, there is reasonable hope that in some countries capacity in the 140 Megabits/second range will be available to the research community on an experimental basis. The prospects of providing such capacity on a multi-national basis at economic cost look poor at present, essentially for political reasons. However, fibres with capacity of 565 Megabits/second or above will be widely installed by 1995 and HEP should aim to exploit them on a trans-continental basis. After the establishment of the single European market in 1992, this may become economically realistic (since real costs are a small fraction, at most 10%, of current PTT tariffs). A political effort will be needed to achieve this.

It is completely unclear what will be the organization of international research networking in Europe by 1995. It is quite unpredictable whether EARN will survive as a separate entity and become a high-performance network. It is quite unpredictable whether the Eureka COSINE project will come to fruition and produce a working international OSI network for the research community, and if so whether the capacity of COSINE will approach the 1 gigabit/second, 140 Megabits/second, or even 2 Megabits/second level.

It is therefore not certain that HEP will still need to operate its own international networks by 1995, but it seems highly probable. The only safe assumption is that we will be on our own as far as high capacity links are concerned.

By 1995, the OSI protocols over X.25 will be in general use in Europe for routine applications (mail, file transfer, remote full-screen, remote job entry). In America, the academic networks will have replaced TCP/IP by OSI, but not over X.25. The HEP world will as a result be engaged in phasing out proprietary and ad hoc protocols for routine purposes. Whether this will be a complete or partial

process will depend critically on whether HEP is still dominated by two proprietary operating systems (in which case DECnet and SNA will remain of great importance), or whether Unix comes to dominate. However, "modern" applications (particularly remote use of high-performance workstations), and the need for bulk file transfer at very high speed, will mean that proprietary and ad hoc protocols will also be in use. It is entirely unclear that these will be based on X.25.

Assuming a continued trend towards campus networks, we can foresee two major performance bottlenecks in 1995 or earlier:

- the traditional bottleneck at each interface to a mainframe, probably around 1 Megabyte/second (10 Megabits/second) by then.
- a new bottleneck at the interface between the campus LAN and the WAN, probably around 5000 packets/second (say 5 Megabits/second for general-purpose traffic) by then.

Thus one can assume that bandwidths significantly above 2 Megabits/second where available, will be multiplexed between mainframe-mainframe channels and LAN-LAN channels, rather than being shared by packet-switching.

16.5 Final Comments on Technology

CERN is not alone in pushing the frontiers of technology for scientific data processing. A recent report³⁴ on the NASA Ames research center included their requirements for scientific computing in 1990:

1. A Super computer capable of 10 gigaflops and 1 gigaword real memory
2. Software to aid programming multiprocessor supercomputers
3. On site local networks of 1 gigabit/second transfer speeds
4. 3-D Graphics hardware and software
5. Hierarchical mass storage systems with gigabit/second transfer rates.
6. Wide area networks in the hundred gigabit/second range

Given the current technology status, this would seem optimistic for 1990, but compatible with HEP needs for the 1990-1995-2000 time frame.

³⁴ Datamation, April 15th 1987, page 60.

Appendix A

Current Mass Storage Usage in the CERN Computer Centre

In what follows the numbers are in Gigabytes. On VM a single density 3380 realistically holds .53 Gigabytes.

A.1 VM/CMS

A.1.1 Disk Space

- Minidisks are used for source programs and small data files.
- Maxidisks are used for intermediate and large data files.
- There is another 8.5 Gigabytes of VM disk space (expanding to 25.5 Gigabytes by the end of 1988) allocated for CRAY use. This is considered in the CRAY section of this note.
- The above scenario for expansion to the end of 1988 assumes minimal growth outside the LEP experiments. Since LEP accounts for only 20% of the accounts, this may be rather optimistic.
- Approximately 8 to 10 Gigabytes are needed per year to keep up with normal user minidisk growth. Any maxidisk expansion is in addition to this.

Looking at 1988 as a whole, the CPU capacity will have been doubled but the disk space increased by less than 35%.

A.1.2 Users and Accounts

- There are currently some 5000 VM/CMS accounts (belonging to 4000 users) of which 20% are in the 4 LEP groups.
- 3000 users log on to the 'IBM' service per month.
- The number of different VM accounts used each week has increased by 50 % in the last 7 months.
- 10% of the accounts were registered in the last 3 months and 18% in the last 6 months. The LEP experiments account for about 20% of this growth.

A.2 VAX/VMS

A.2.1 Disk Space

The current total of 12.5 Gigabytes will be increased to 26.9 Gigabytes by the end of 1988. The distribution is :

Use	Current Total	Expected End 1988 (but see below)
System + Scratch + Stage	3.7	5.5
LEP User disks	4.0	4.0
Non LEP User disks	4.8	5.4
Total	12.5	14.9

- The 2.4 Gigabytes increase indicated above will occur in June.
- Another 2.4 Gigabytes will arrive in July/August for use as User disks.
- Of the 9.6 Gigabytes (RA90s) that will arrive in October, 2.4 Gigabytes has been foreseen for stage and the rest not yet specified.
- If the pilot batch service is a success, it is estimated that a total of 10 Gigabytes will be required for stage and 10 Gigabytes for user data files. Of this 20 Gigabytes only about 5 Gigabytes can be accommodated from the known increases.
- It has been estimated that, at the current rate of expansion, approximately 5 Gigabytes per year are required for normal user growth.

A.2.2 Users and Accounts

- There are currently some 2100 VAX/VMS accounts (belonging to 2000 users) of which 50% are in the 4 LEP groups.
- 1400 users log on to the VAX cluster service per month.
- 10% of the accounts were registered in the last 3 months and 20% in the last 6 months. The LEP experiments account for about 45% of this growth.

A.3 Cray

For the disk space for CRAY on VM the real figure of .53 Gigabytes per 3380 single density equivalent has once again been used.

By September 1988 there will be 25.5 Gigabytes VM/CRAY (on VM for CRAY) and 27 Gigabytes on the CRAY itself. In the table below only the expected end 1988 situation is considered.

The likely scenario is :

Use	Expected End 1988
System,Scratch (on CRAY)	3.5
User files (on CRAY, non-LEP)	1.5
User files (on CRAY, LEP)	2.0
Staged tapes/files (on CRAY)	20.0
User files (on VM, non-LEP)	9.5
User files (on VM, LEP)	16.0
Total	52.5

- The LEP, Non LEP splits above are only suggestions.
- Another 20 Gigabytes (on CRAY) might be obtained over the next two years, and if so, LEP could be the prime beneficiary.

A.4 Summary

Given the known increases as outlined above, by the end of 1988 the LEP experiments could have approximately :

- **VM/CMS** : 7 Gigabytes of minidisk space and 6 Gigabytes of maxidisks. They would be major users of 25 Gigabytes of staging space.
- **VAX/VMS** : 4 Gigabytes user disks plus perhaps another 6 to 7 Gigabytes of the increases not yet allocated.
- **CRAY** : 2 Gigabytes of user files on CRAY, 16 Gigabytes of user files (mostly maxi-like) on VM/CMS. They would be major users of 20 Gigabytes of staging space on the CRAY.

It is clear that these numbers fall far short of the 200 Gigabytes as called for by MUSCLE for the end of 1989.

CERN Data Networking Requirements in the Nineties

Technical Board on Communications

9 November 1988

Chapter 1

Preamble

This report has been prepared by the CERN Technical Board on Communications (Telecom)¹ in response to a request from the Steering Committee for the report "*Computing at CERN in the 1990's*" chaired by J.J.Thresher.

Due to the very tight timescale, this report was first drafted on the basis of recent and current studies, and on the basis of formal and informal contacts with the CERN user community. Input from the LEP experiments was solicited during the summer concerning high-speed links. The interim version has been freely circulated as a draft since 16 June both inside and outside CERN, and all comments received have been taken into consideration.

However, this report is not and should not be read as a statement of requirements for HEP² *in general*, nor for on-line control and data acquisition systems. It refers only to the needs and priorities *for general purpose networking at CERN*. Neither is this document intended to prescribe a technical implementation plan; it is limited to stating the requirements.

The general theme is one of continuity and expansion. Since the studies by Working Group 2 of the LEP Computing Planning Group [1-2], and the studies by ECFA SG5, around 1982, CERN has been pursuing a consistent data communications strategy [3-7], which we do not propose to change, but rather to reinforce, to meet future needs.

Furthermore, input was solicited from members of ECFA Sub-Group 5 on Links and Networks, and the draft report was extensively discussed at the SG5 Meeting in September where three Telecom Board members were present. These sources and other basic reference documents are listed in an Appendix. Finally, an open presentation and very lively discussion took place in the overcrowded CERN DD auditorium in October.

As a result of these and other interactions we feel mandated to *stress* the following points. Networking between HEP laboratories and CERN can offer the large collaborations unique possibilities in running and evaluating experiments, to monitor experiments from home institutes and to promote distributed computing. Bandwidths of 2 Mbits/s at the time of the LEP data taking are considered extremely useful.

A variety of initiatives are being undertaken in several countries in order not to miss these opportunities. One of the most basic problems of experimental institutes, how to concentrate enough effort at CERN inside the limitations of travel budgets and teaching obligations, may be eased considerably using high quality network connections.

These institutes rely on CERN to fulfill its role not only to provide functioning accelerators but also to provide services in the field of computing and, in particular, of networking. Computer networks need planning, management, development, and transfer of knowledge. CERN is the most important HEP institution in Europe from which the external HEP institutes expect to receive such services.

¹ J.V.Allaby (Chairman), W.Blum, B.Carpenter (Editor), F.Fluckiger, D.Lord, C.R.Parker, G.Shering (Secretary).

² High Energy Physics

Chapter 2

Background and History

By 1990, CERN will have been at the forefront of data networking for twenty years. One can cite OMNET (1972), CERNET (1978), and CERN becoming a test customer of the Swiss X.25 service (1981) as older examples. More recently, CERN has become one of the world's largest Ethernet users (probably the first to have a site-wide bridged Ethernet), a major hub for the use of X.25 and DECnet in HEP, and the Swiss national node of the EARN and EUNET networks. CERN operates a unique set of gateways for electronic mail and file transfer. The reasons for this continual lead are those of fundamental necessity: the needs of a large, increasingly computerised, site, the increasing size and internationalism of the experimental collaborations, and the increasing size of experimental data samples. Our experience is that however much capacity is provided, it will rapidly be filled, and that there is always a suppressed demand for extra capacity. The *typical* annual growth rate for the traffic in each data communications service operated at CERN is now between 20% and 100%. The current pressure to decentralise data analysis (see the section on Central & Distributed Processing in the report on Computing for Experiments) will only emphasise this trend.

The current illustration of this is of course the LEP experiments, whose preparation has depended critically on the availability of networking (principally for electronic mail in terms of human transactions, and principally for file transfer in terms of bulk network capacity). It is unlikely that the collaborations could have advanced as they have without these facilities. The MUSCLE report [8] has shown, furthermore, that today's network capacity is certainly inadequate for the data taking and analysis phases of the four experiments.³

We cannot see this demand for more and more network capacity going away in the next few years. Indeed the availability of high bandwidth WAN⁴ capacity from European PTTs, the imminence of very high bandwidth LANs⁵, and the approaching deregulation of European telecommunications, will encourage and intensify this demand.

In addition, the range of people using data networking on the CERN site will increase precisely as the use of computers for non-traditional purposes increases. Thus the demand for more capacity per user will be multiplied by the increase in the user base.

At the same time, the organisation of networking for European research in general will improve, moving from the present somewhat fragmented situation to a more integrated one (catalysed by the eventual arrival of standards). CERN will benefit from this process, but will continue to have some requirements which go beyond those offered by the emerging standards.

A final background remark is that current networks can be described as "third generation" (industrialised and largely standardised). A fourth generation of networks, providing truly distributed computing rather than bilateral communication between computers, is now coming to maturity. In addition to continuing to provide adequate capacity and basic services, CERN must not miss the boat of distributed computing.

³ An equivalent analysis of the needs for the new generation of American experiments leads to a similar result. ECFA SG5 believes that these two analyses can legitimately be extrapolated to HERA.

⁴ Wide Area Network

⁵ Local Area Networks

Chapter 3

General model of computing activity

This chapter gives our view of the computing activity at CERN and in its external partners' institutes, as a scenario within which to state requirements for general-purpose networking.

We assume that there will continue to be one very powerful Computer Centre at CERN. There, one will find most of the batch capacity and non-volatile storage, an operations centre, and the focal point for networks. We assume that roughly half of the interactive computing activity involves this Centre (either by direct connection, or indirectly through intelligent workstations).

We additionally assume that there will be several powerful independent computing centres elsewhere on the CERN sites, which will provide some batch capacity but principally be used interactively for physics analysis. These will rely on the Centre for reliable non-volatile storage. Examples of these centres are the Work-station Clusters being set up by the LEP experiments. Each such centre would have a single well-defined connection to the general-purpose network.

We also assume the existence of several networks for on-line systems (control and data acquisition systems), which should each appear as a single, self-contained, entity for purposes of security and management. These networks would each be connected to the general-purpose network by an appropriate gateway system. (Although the requirements for such on-line networks are outside the scope of this report, we do not mean to imply that they necessarily require different technology, but only that they should be logically separated and could be separately managed.)

We assume that there will be several thousand intelligent desk-top devices around the sites (personal computers and workstations). At the beginning of the 1990's there will be more than 3000 dumb terminals, but we expect that this number will decline as the number of intelligent desk-top devices increases. Many PCs (or terminals) will be in people's homes.

External communications will be required with several comparable laboratories in Europe and the USA, and with hundreds of smaller institutes and university departments, mainly in Europe and the USA, but potentially anywhere in the world. Each of these external sites will be subject even more than CERN to its own pseudo-political pressures from governments, PTTs, and manufacturers, so may not be free to choose communications techniques at the behest of CERN.

Connections required to some countries may be forbidden for purely political reasons connected with restrictions on technology transfer. We regard this as being *prima facie* incompatible with the scientific openness of CERN.

A quantitative prediction of network traffic from these assumptions is difficult, especially since all past experience is that traffic expands to fill all available bandwidth, with compound growth rates of at least 100% per year. Also, all previous studies of data traffic at CERN have shown an atypical pattern in which the highest throughput is often required over the greatest distances. For the early 1990s we believe that the aggregate cross-site bandwidth should be about 100 Mbps (including safety factors), as foreseen in earlier studies. Technological trends in LAN technology suggest that this could become inadequate from about 1993, but it is too early to predict how great an increase will be required.

For off-site networking, the potential demand is effectively infinite and traffic predictions as such are probably meaningless. Certainly in the period under study, the bandwidth will be limited by cost or technology, not by needs.⁶ In this connection, the recent very detailed bandwidth requirements devel-

⁶ The technological limit today is at around 2 Mbps for mainframe-mainframe connections and perhaps double that for Ether-

oped by the American HEPnet Review Committee [10] clearly show that even mundane use of networks requires high bandwidths. We believe that these estimates can readily be extrapolated to European HEP. The 2 Mbps links between CERN and the major external centres recommended by the MUSCLE report are certainly inadequate to allow true distributed computing over a wide area.

Our estimates of the real requirements by 1993 are collected in Table 1, together with the current figures for this year and the extrapolated numbers for the end of 1989. It should be noted that the large increases in peak backbone traffic correspond in 1989 to the start-up of 4 LEP experiments with links from the 4 pits to the computer centre as well as the start of LEP data analysis, and in 1993 to the transition to fully distributed computing. It is useful to point out that even if the bandwidth and connectivity requirements are met by the communications infrastructure, performance bottlenecks may arise at the interface to central computers. (There is anecdotal evidence of 9600 bps connections providing one character per second of throughput!)

Table 1: Estimated Requirements by 1993

	Current 1988	Extrapolated 1989	Required by 1993
Nodes on LANS (computers,workstations,PCs)	800	1500	4000
Dumb terminals	3000	3500	2000
Peak No of internal logins to comp. centre, full screen	700	1000	1500
Peak backbone traffic	1Mbps	10Mbps	50Mbps
Raw backbone bandwidth	10Mbps	> 20Mbps	> 100Mbps
No of external leased lines (up to 64 kbps)	18	20	15
No of external leased lines (high speed)	0	1	10
Peak No of external logins (in- or out-, full screen)	100	200	400
Integr. external bandwidth	0.5Mbps	> 2Mbps	> 100Mbps

The above forms the context for our view of CERN data networking. Scenarios for networking must not pre-suppose any particular choice of operating system etc., as networking solutions must function

net-Ethernet connections. Both of these limits will evolve rapidly in the next two years. At least one proposal for use of 140 Mbps capacity is being prepared at the time of writing.

in a diverse environment over many years. However, it must be noted that some companies have proved better than others at providing "peer-to-peer" networking which requires a minimum of central management. This has clearly driven many purchasing decisions in the HEP community in recent years, except in the pure Computer Centre environment where the cost of batch and storage capacity takes priority in the choice of suppliers.

Any networking scenario must provide mechanisms for user relations, both on a day-to-day and on a strategic basis. We believe that daily user support tasks are best integrated with the general user support mechanisms of the CERN Computer Centre. Advisory bodies which are representative of the user community are also required. These bodies already exist, in the form of the Telecom Board for CERN's internal affairs, and ECFA SG5 for the European HEP community as a whole.

Chapter 4

Services Required

This chapter gives a brief summary of the services ideally to be provided by internal and external networks. It should be noted that user groups can and should build specific application-oriented services on top of these basic services.

The services are listed in no particular order of priority, although the six more "traditional" services are listed first.

1. Remote login in both line mode and full screen mode. Full screen mode is now predominant on LANs, and should become so on WANs, with important effects on bandwidth needs.
2. Inter-personal messaging i.e. electronic mail, augmented by telex, and telefax.
3. File transfer, including transfer of documents, and programs, and down-line loading.
4. File backup
5. Remote job entry and retrieval.
6. Remote printing.
7. Sharing of files (including documents, and programs) and databases between systems. Includes remote file access and remote directory manipulation.
8. Distributed transaction processing, including task-to-task communication, and remote procedure call.
9. Access to public telematic services such as data bases, travel reservation services, and electronic data exchange for commercial business.
10. Remote monitoring of equipment or experiments.
11. Visualization i.e. remote graphics capability.
12. Remote windowing systems such as X-windows.

13. Workstation cluster services.
14. Ancillary services such as distributed directories, name servers, date-and-time servers.

Chapter 5

Technical Requirements

We note that technical evolution in areas such as fibre optics and ISDN⁷ continues without pause. We therefore expect that current 10 MHz LAN technology will be replaced by 100 MHz technology in the coming five years. Similarly, the current 64 kbit/second WAN technology will be replaced by 2 Mbit/second technology (indeed, this is already today largely a financial issue), and later by substantially greater speeds. As mentioned above, PTT regulations in Europe are rapidly becoming more liberal, and this process should be completed by 1992 with the establishment of the full Common Market. HEP must, if necessary, insist at the political level on its need for freedom from arbitrary PTT rules. We also note the rapid evolution in the area of standard protocols, and in particular OSI.⁸ All technical requirements in this report must therefore be kept under continuous review and updated as conditions change.

5.1 Review of 1983 requirements

Firstly we support the analysis of requirements made in 1983 in the LEP Computing Planning Group report. The resulting recommendations have been followed as far as possible by DD and still remain valid. We give here a summary of these recommendations:

Implementation of a general-purpose high-speed backbone network interconnecting LANs.

Use of industrial products for the backbone, and integration with other facilities including digital telephone.

Use of CCITT transmission standards across the site [i.e. G.703 series TDM (Time Division Multiplex) equipment and optical fibres].

⁷ The Integrated Services Digital Network to be offered in the 1990's by the PTTs.

⁸ Open Systems Interconnection standards defined by ISO and CCITT.

These recommendations have still not been fully implemented due to the lack of suitable (100 MHz) industrial products for the backbone, but an intermediate backbone network has been created using slower products. G.703 equipment has been widely installed, mainly by SPS/ACC for the LEP site, and it is used both for the digital telephone and for the intermediate backbone.

Adoption of an agreed set of high-level protocols.

(See below for more specific requirements on this point.)

Expand wide-area network services.

This expansion has been a continuous process since 1983, notably concerning EARN, DECnet, and X.25, and shows no sign of slowing down.

Choose and install a standard LAN.

ISO 8802.3 (Ethernet) has been chosen and installation is well advanced, and should of course continue.

Adopt an integrated approach to terminal connections.

This recommendation is valid, but has proved difficult to implement, and does not allow for the impact of personal computers. However, the Telecom Board has recommended use of Ethernet for terminal and PC connections, with the digital telephone as a backup for remote buildings or special cases. The impact of ISDN in this area remains to be seen.

5.2 Protocol requirements

A CERN policy on communications protocols was adopted in 1985 [3]. Particular requirements in this area are:

Strategically, the protocols used should be independent of any particular computer manufacturer and should be able to communicate with the widest possible range of external sites with the minimum diversity of protocols.

This requirement would be best matched by a migration to OSI protocols. This strategy matches that of the academic and research networks in the member states and even the USA. The subset of OSI protocols chosen by CERN should be as far as possible the subset chosen by the member state academic communities as a whole.

While awaiting OSI products, or in cases where the performance penalty caused by a standard protocol is too great, CERN requires to use interim (*de facto* standard, or manufacturer's) protocols. However, diversity of such protocols should be resisted and no development effort should be devoted to such protocols. Home-designed protocols should be avoided.

Sadly, numerous interim protocols are in use today due to slow industrial progress in realising OSI protocols. We wish to emphasise the very substantial manpower cost of this diversity. All moves to introduce new protocols should be evaluated with great care.

Protocol gateways continue to be required as a major tool in providing connectivity despite diversity of protocols, and as an aid for migration to OSI protocols.

We support the continuation of the existing gateway programmes (GIFT⁹, MINT¹⁰, and several industrial gateway products). Additional gateways should be implemented if the need becomes apparent.

5.3 Specific requirements

This section gives a number of more specific requirements.

Access to CERN at speeds of at least 2 Mbps is required from major HEP sites in Europe, to allow decentralised analysis of the LEP data, as soon as this analysis starts in earnest. Higher speeds (equivalent to those attainable on a LAN) will be required as soon as practicable, to allow effective use of workstations over geographical distances.

This requirement was initially raised by the MUSCLE report, and is justified in detail in the report on Computing for Experiments. The links will in all probability be based on PTT fibre optic cables and on the same G.703 TDM standards as used on the LEP site. The potential demand for bandwidth on such links is potentially infinite, with 2 Mbps being the minimum level at which the connection ceases to be frustratingly slow at peak hours. The limiting factors will be technical availability (for some

⁹ General Internet File Transfer, a file-transfer protocol convertor implemented by a CERN-INFN-RAL-Oxford-NIKHEF-SARA collaboration.

¹⁰ Mail INTerchange, a set of mail gateways implemented at CERN

countries), tariffs, and bottlenecks at the computer interface (for some computers).

Until now, CERN has always viewed its WAN connections as a logical star with CERN at its centre, and CERN effort has officially been confined to support of connections to CERN. We have informally stepped outside this role in several cases, such as by acting as the hub for DECnet, X.25, and electronic mail in the HEP community, and by acting as host for the Swiss national nodes of EARN¹¹ and EUNET¹². CERN and ECFA SG5 have jointly taken the role of representing HEP in both the RARE¹³ Association and in discussion with the Eureka project COSINE¹⁴. These activities have been informally known as 'HEPnet'. At the initiative of ECFA SG5, a proposal¹⁵ has been prepared for the the HEP-CCC to put HEPnet on a more formal basis.

In summary, the proposal is that HEPnet should firstly represent HEP interests in the international networking community (RARE, COSINE, EARN, EUNET, etc.), and secondly coordinate the management of whatever HEP-specific network facilities are required to complement the general-purpose networks. It is proposed to create a HEPnet Requirements Committee as the successor of ECFA SG5, and a new HEPnet Executive Committee to coordinate the technical work.

It should be noted that the relationships between RARE, COSINE, EARN, EUNET, and various national and regional networks are very complex and in constant evolution. Nevertheless, to avoid duplication of effort and wastage of money, it is desirable for HEP to use these networks as much as possible. Only by taking an active role in HEPnet can CERN strike the correct balance between use of the general-purpose networks and use of CERN-managed networks.

CERN should take an active role in HEPnet activities, without losing sight of the priorities of the CERN experimental programme. CERN should also, in conjunction with HEPnet, collaborate with general-purpose network activities (RARE, COSINE, EARN, EUNET) as long as this is beneficial to the HEP community.

The dramatic growth in services, traffic, and number of users in recent years has caused a great increase in the load of operational management of internal and external networks. This would become even more acute if the creation of HEPnet led to an increased commitment by CERN to support transit traffic. Today, no operational staff are assigned specifically to data communications and many routine jobs are carried out by systems programmers. It is planned to assign one or two operational staff (on day-shift only) to alleviate this situation. Corresponding technical tools must also be provided, and the use of appropriate service contracts should be investigated.

¹¹ European Academic and Research Network

¹² European Unix Network

¹³ Réseaux Associés pour la Recherche Européenne

¹⁴ Cooperation for OSI Networking in Europe

¹⁵ Authors: R.Blokzijl (SG5 chairman), B.Carpenter (CERN), H.Hoffmann (DESY).

Organisation, staff, and tools for network operational management should be boosted.

For historical reasons, some users still depend on CERN-designed and -implemented protocols. Pressure on support staff, and the availability today of adequate protocols from industry, requires correction of this anomaly.

Dependency on CERN-designed protocols should be eliminated as rapidly as possible, in favour of industrially supported protocols.

Thus CERNET should be phased out in the coming years, to avoid future maintenance problems. (The same argument clearly applies to the old TITN networks for the PS and SPS control systems.)

The general-purpose LAN infrastructure should take into account specific needs of the large experiments, to avoid their needing to implement "private" extended LANs across the sites.

For example, this specifically concerns fast links to the LEP pits, and high-speed connections between workstation clusters and mainframe computers. The detailed justifications are given in the MUSCLE report. We would in general discourage any solutions requiring special (or worse, home-made) hardware or software.

Network services must be fully integrated between LANs and WANs.

CERN users commute regularly between CERN and their home institute; as far as possible they should see a 'seamless' network in which the difference between LANs and WANs is hidden. In principle, there should be a single LAN/WAN gateway at CERN for each type of service, to simplify security and accounting procedures, and to optimise resources.

The old telephone exchange will be replaced by expansion of the digital exchange within two years. This will then clear the way for use of the next generation of PTT services, namely ISDN, when this becomes a financially and technically attractive alternative to X.25 and/or LANs (probably not before 1992).

The use of ISDN services on and off the CERN sites should be planned.

Chapter 6

Resources Required

It is clear that the need for data communications facilities will continue to grow dramatically, as in recent years, certainly until about 1992. Whether we then reach a period of relative stability, or whether growth continues, depends on the future of the CERN experimental programme. (For example, a trend away from high-statistics experiments might significantly alter the traffic patterns implied by the MUSCLE report). However, even in a period of stable demand, there would be a need for a rolling replacement programme for equipment purchased in the 1980's. We prefer to assume that the resources required throughout the 1990's can be extrapolated from the well-understood needs of the next four years.

Our financial model is therefore built out of five items:

- The financial model developed in the report by the Telecom Board of October 1987 on resources required for communications throughout CERN.
- An allowance for the areas which were consciously omitted from that report (especially wide area networking, protocol gateways, and network management).
- The financial model for the STK digital telephone exchange.
- An allowance for the new needs revealed by the MUSCLE report.
- Current operational costs.

The expenditure on Ethernet and the backbone should tail off, but will presumably need to be replaced by other expenditures (such as the replacement of INDEX as it becomes unreliable, or the introduction of the next generation of LANs and of ISDN equipment). To allow for this assumption, a flat extrapolation of current rates of capital spending has therefore been adopted in Table 2. However, two capital items are not extrapolated beyond 1992, namely the STK digital exchange and the MUSCLE items. It is not excluded that by 1991, other major capital items will prove to be necessary, so a contingency line has been added. This financial model is coarsely rounded; it would be misleading to imply greater precision than is possible by giving more precise figures.

Table 2: Financial Profile of Capital Expenditure (MSF)

	1989	1990	1991	1992	1993
Internal data network	2	2	2	2	2
STK exchange (voice)	1	6	1	-	-
Ext. network, gateways	0.5	0.5	0.5	0.5	0.5
MUSCLE (pits, workst.)	1	1	-	-	-
Contingency	-	-	1	1	1
Totals	4.5	9.5	4.5	3.5	3.5

Table 3: Financial Profile of Operations Expenditure (MSF)

	1989	1990	1991	1992	1993
Maintenance and ops (all data services)	1.5	1.5	1.5	1.5	1.5
Maintenance and ops (telephone & telex)	0.7 ¹	0.7 ¹	0.7 ¹	0.7 ¹	0.7 ¹
Calls ('phone, telex)	3	3	3	3	3
[estimate: paid by DA]					
X.25, EARN, EUNET	0.3	0.3	0.3	0.3	0.3
[estimate: paid by DA]					
Totals	5.5²	5.5²	5.5²	5.5²	5.5²

1. including 0.4 MSF paid by DA
2. including 3.7 MSF paid by DA

Two major items are not included in table Table 3. The first of these is the cost of the high-speed external links required by the MUSCLE report. It is not clear whether these lines would be funded by CERN or by the external institutes. At normal PTT tariffs, a total annual operations budget rising from 500 KSF in 1989 to 12 MSF from 1992 would be required. Costs of conventional direct links to CERN are currently paid by the Institutes requesting them. The second item is the cost to CERN of establishing and operating its share of HEPnet, currently unknown.

It is very difficult to give a true picture of staff requirements, since almost everybody involved in computing support is involved for some fraction of his or her time in communications support. The estimates here are only for people involved for a significant fraction of their time in communications. The Telecom Board gave a detailed staff scenario in its October 1987 report on resources [7], showing a total of about 20 FTE staff¹⁶ supporting *internal* networking, and about 15 for telephony. About 28 of these FTEs are in DD today, and several more should be transferred to DD from SPS in due course. Several other Divisions have a few staff involved part-time in data communications, as do all large collaborations.

To this one must add staff needed to support *external* networking (currently ≈ 10 FTE). This gives a total of about 30 staff providing central data communications support in DD today. These staff are overloaded by a burden of installation work, service support, and user support, almost to the exclusion of any work in preparation for the future.

It is necessary for the other Divisions to make available to DD substantial technician support for the planning, installation, and commissioning of their Ethernets and for ongoing operational support. The direct participation of the large collaborations in the planning and operation of their on-line Ethernets has been assumed and is fortunately happening. It will be necessary for this participation to be extended to cover also the *first-line* intervention in the extensive local area networks planned for the analysis centres, where the changing topology demanded by the physics requirements will require *local* expertise, supplemented by the minimum of consultation with the central DD support group. For this purpose, the DD support group will have to organize suitable training courses in network management and operations, to allow physicists, engineers, and technicians to acquire the appropriate expertise.

¹⁶ Staff figures are quoted as full-time equivalents, not as individuals.

Table 4: Current Staffing Levels (FTE) for Communications

	DD ¹	PS	SPS	LEP	EP/EF	All Divs
Software	14	0.5	0	0	1	15.5
Hardware	2	0.5	2	0	0	4.5
Technicians	15 ²	0.5	4	0	0.5	20
Operators	15	0	0	0	0	15
Totals	46	1.5	6	0	1.5	55

1. all DD groups, not only CS
2. includes one ST telephone technician

Furthermore, extensions of our wide area networking possibilities can only be made possible by increased collaboration and *participation* by the major European HEP institutes in their maintenance and operation.

Table 5: Future Staffing Needs (FTE) for Communications

	DD ¹	PS	SPS	LEP	EP/EF	All Divs
Software	17	0.5	0.5	0.5	1	19.5
Hardware	4	0.5	0	0	0	4.5
Technicians	18 ²	0.5	0.5	0.5	1	20.5
Operators	13	0	0	0	0	13
Totals	52	1.5	1	1	2	57.5

1. all DD groups, not only CS
2. includes one ST telephone technician

Continued priority should be given to Communications, within the general staffing policy of the Organisation. In particular, we believe that the Communications Systems group needs several additional staff to help with daily administrative or operational tasks, today wastefully carried out by software engineers, and an additional boost in software effort.

Table 4 shows current approximate FTE staff, and Table 5 the numbers we consider are required to cope with future needs. *They exclude on-line networks, TV, audio, and safety communications.* Table 5 on page 195 only represents a survival level of staffing only; an increase of about ten posts could readily be justified, for example to provide the level of support that many users would like to see.

Table 5 on page 195 also implies that all departures must be compensated, with the exception of some telephonists, and that operational and software manpower should be boosted by a few posts, as an absolute minimum for survival. It will be noticed that the future staffing needs have been slightly increased from those estimated in [7], since we now believe, as a result of our extensive consultations, that we were too modest in the former estimate.

Appendix A

Networking references, sources, and acknowledgements

1. Report of the LEP Computing Planning Group, ed. D.O.Williams, DD/83/8
2. Working Group 2 Final report, ed. J.M.Gerard, December 1982 (and at least 50 associated working papers)
3. CERN Data Communications Policy, ed. B.Carpenter, DD/85/14
4. General recommendations on LANs at CERN, B.Carpenter, J.M.Gerard, J.Joosten, C.Piney, DD/CS - R18, November 1986
5. Recommendations for the Connection of Terminals at CERN, B.Carpenter, C.Piney, I.Willers, DD/CS - R22, December 1986
6. Recommendations for Terminal Connections and the Use of Local Area Networks, Technical Board on Communications, April 1987
7. Resources Required for Communications Throughout CERN, Technical Board on Communications, October 1987
8. The MUSCLE Report - the Computing Needs of the LEP Experiments, ed. D.O.Williams, DD/88/1
9. A research and Development Strategy for High Performance Computing, Executive Office of the President, OSTP, Washington, November 1987.
10. High Energy Physics Computer Networking - Report of the HEPNET [USA] Review Committee, in preparation.
11. What would the LEP experiments do with 2Mbit/sec links? Memorandum by E.-A.Knabbe, DD/CS. (October, 1988)
12. ECFA/SG5 Working Papers (more than 150 documents produced or circulated by ECFA SG5 since 1981)

Plus individual input or comments from: *R.Dolin (L3)*, *P.Frenkiel (Paris)*, *F.Harris (Oxford)*, *M.G.N.Hine*, *P.Ponting (CERN)*, *E.Pagiola (CERN)*, *A.Rouge (IN2P3)*, *W. von Räden (ALEPH)*, *B.Segal (CERN)*, and several members of the DD/CS Group. Apologies to all those who have been overlooked.

Accelerator Computing Requirements in the Nineties

Working Group on Accelerator Computing

Final Version, March 1989

Chapter 1

Introduction

This report has been prepared in response to a request from the Steering Committee for the report "Computing at CERN in the 1990's" chaired by J.J. Thresher. To prepare the report a subcommittee of six members was formed with as chairman Roy Billinge who is the member of the editorial committee responsible for accelerator computing.

The composition of the subcommittee was as follows. Roy Billinge as chairman and providing the input on MIS activities as chairman of SCAIP. George Shering as secretary and providing the input on networking as a member of TELECOM. Lyn Evans as specialist on SPS matters, and as a user of controls rather than as a builder. Claude Hauviller as liaison with CAE, in particular the mechanical design and engineering aspects. Fabien Perriollat as the expert in controls and member of TEBOCO. John Poole as the expert on databases and to bring input from LEP/CLIC on the requirements for simulations.

The members were charged to reflect their own views, those of their divisions, and those of the whole of CERN, not only in their specialist fields but also over all the subjects discussed. The subcommittee met six times, and this report summarises the discussions and decisions made.

Chapter 2

Main Conclusions

Networked workstations will become the generalised computing tools, with about 1000 general purpose office workstations and 100 special purpose workstations in the accelerator divisions.

Local and special purpose servers and networking support will continue to be supplied from within the accelerator divisions, but the accelerator divisions will increasingly rely on DD for central network support, for the main computing facilities, for the accelerator database engine, and for backup facilities. For all this they will expect as good support as the research divisions.

It is agreed that CERN should remain a multi-vendor site as a whole. The accelerator divisions, however, retain the right to opt for a single line, either as a group or each division separately. Support for open systems such as UNIX, and products that will run on a variety of support engines, should be continued and even given preference.

The accelerator divisions have a substantial commitment to IBM compatible office workstations (PCs), mainly of European manufacture. Network facilities and interconnections are being set up to capitalise on their use. DD support for the supply and networking of this line should continue, and all CERN-wide facilities made available on such networks. Coordination should be provided for the choice of software to be used, and its relationship with other CERN software options.

Networking in the accelerator divisions is seen on three levels, the office LAN, the main controls LAN, and regional LANs for equipment. A reduction in diversity of networking technologies is desired. DD should provide strong central networking support, and should handle all aspects concerning licences, distribution, and grouped ordering. DD should also do the site network management, including allocation of names and IP numbers. A load of up to 1000 packets per second between the accelerator LANs, and to the centre is envisaged. About 100 home terminal connections will be required in the accelerator divisions.

There should be a single central database management system for all accelerators, preferably running on the same computers. Over the period in question this should be ORACLE, which will be increasingly relied on for operational requirements. Thus an ORACLE piquet, and a piquet for its support engine, will be required.

Accelerator design and modelling techniques will move in part to workstation networks, but many large jobs will still require the main computing facility. This could require between 0.2 and 1 CRAY XMP/48 year equivalent spread over the five years 1990 to 1995. With the increasing power of workstations on the one hand, and the better performance of networks on the other, the accelerator control systems would be interested in replacing their local computation facilities by on-line access to the central facility provided a suitable agreement on availability could be achieved.

CAE will continue to be important with Euclid on VAX being the mainstay of the mechanical design, Autocad on the PC networks for mechanical drawing, and PCAD on the PC networks for electrical design. It is foreseen that that over the period the main priority for mechanical CAE will be the integration around a central database of design, drawing, computations, manufacturing, and documentation packages. PRIAM has become vital to the accelerator divisions. The support for embedded microprocessor systems should increase, and DD should support an integrated software workshop including tools for the specification, design, development and exploitation phases. This should work in liason with ORACLE.

The central documentation service should be expanded to cope with technical and engineering documents, and should work in liason with ORACLE.

Increased manpower will be required in the accelerator divisions for computing other than the control systems. In particular up to 15 people will be required for local informatics and networking support, working in liason with the central support services. Database application design will require another 15 to 20 people, which may become less if one team could do all accelerator applications.

Chapter 3

Background and History

The Accelerator divisions are involved in computing in two specialist areas, namely Accelerator Design and Accelerator Controls, and in three more general areas of Computer Aided Engineering, Database usage, and Management Information Systems. They were early into computing with programs for lattice design, particle tracking, and field calculations. With the advent of computer control of the accelerators, the accelerator divisions had to build up a very considerable expertise in computer activities, including things like busses, language design, networks, databases. This also overflowed into other uses of the computers such as CAE and MIS so that the accelerator divisions often played a leading role. With the increasing penetration of the computer into all areas of CERN's activities, however, many of the accelerator control based activities became of general application, especially as the experiments grew to almost accelerator size. Also there has been considerable pressure on human resources in this area. Thus we expect to see a major shift in the nineties to the accelerator divisions becoming users of more generally supplied services, albeit very knowledgeable and discerning users.

Chapter 4

General Model of Accelerator Computing

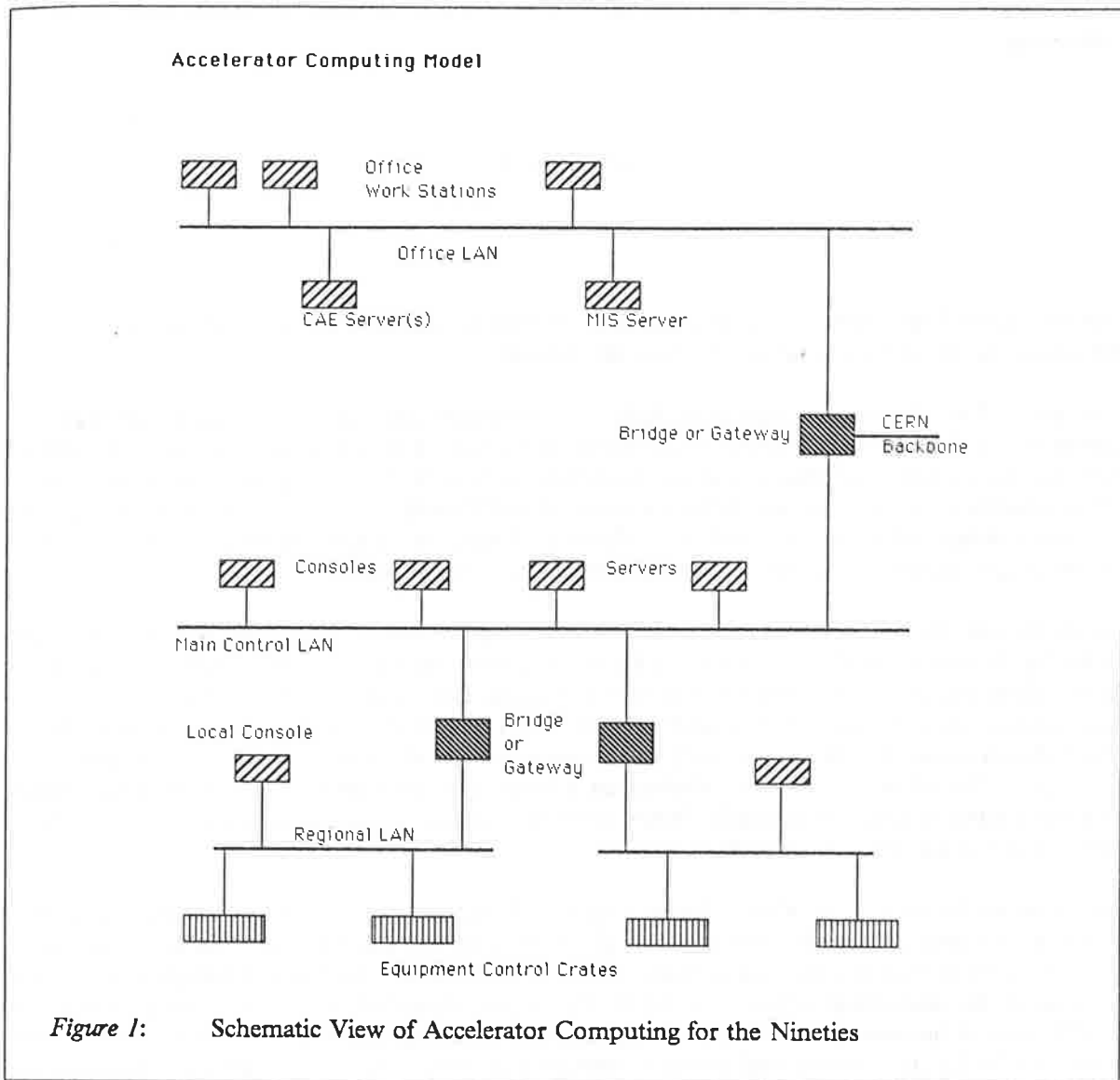


Figure 1 shows a schematic view of accelerator computing as foreseen for the nineties. Office workstations will progressively replace terminals, and will be networked together and to various servers. A bridge to the CERN Backbone will provide access to the central facilities and to other networks.

It is reasonable to suppose that the accelerator divisions will remain grouped geographically into the PS, LEP and SPS areas, each with a local area network infrastructure. As the personnel distribution and political organisation of the accelerator divisions is not invariant with time, the division into geographical areas must not hinder logical communication between areas.

The Main Control LAN will be connected to the CERN backbone for access to central services, but also to the office LAN as most of the accelerator division staff have responsibilities for accelerator equipment or operation and will want to have some interaction through the office workstation. Also

many will be involved in control software development.

Regional LANs will increasingly replace the more traditional equipment control highways such as serial CAMAC and MPX. These will be connected through the gateways to the other LANs for control commands, of course, but also for more general purposes such as database access and microprocessor development. Some direct access from the office workstations may also be allowed for equipment monitoring.

Chapter 5

Workstations

There is essential agreement on the trend towards networked workstations, together with local servers, and increasing use of the central services over the network.

It is agreed that a "workstation on every desk" is a reasonable aim. In the Accelerator divisions it is estimated that 1000 general purpose office workstations will be in use in the Nineties. In addition there will be a number of special purpose workstations, Vax Station or Apollo. Approximately, 30 will be used for mechanical design, 40 for controls (20 in PS controls, 20 in SPS/LEP controls), 20 for accelerator design and modelling, and 10 for database design. An attempt is being made to minimise the use of special purpose workstations, so a total of 100 seems reasonable.

The accelerator divisions make extensive use of a wide range of industrial software, and so tend to prefer the IBM compatible office workstations for which there is the widest range of software and add on cards. There are, of course, other reasons such as history, European element. The PS, SPS, and ST have a policy on this matter, but in LEP both the IBM and Macintosh streams are present. It was agreed that divisions should have a policy in this area, and an adequate mechanism for implementing the policy. The advantages in concentrating on a single type of workstation include better support from the limited manpower available, better network facilities, easier communication, and a better spread of knowledge from other users.

For office workstation application software a two level approach is recommended. The "high" level should be the best package for the specialist application (often, unfortunately, the most expensive). The "low" level should be a popular package running on the office workstation. Examples are found in many areas. In mechanical design it could be Euclid and AutoCad, in electrical design Daisy and PCAD, in word processing NORTEXT and a popular PC package, in databases ORACLE and Symphony or EXCEL, in software engineering Teamwork and PCSA. At the popular level there is a danger of diversification due to the individualism of divisions and groups. This might be controlled by recommendations from the appropriate Technical Boards, by selective training and support, and by providing interchange between the approved high and low end systems. Interchange is made easier if the low end system is the same as the high end, eg. Euclid or Oracle on a PC. This may not be the best answer, however, if the PC version is not a sufficiently popular one. To avoid diversification the low level package must be popular and easy to use. Management support is essential if a coherent advance in these areas is to be made.

Chapter 6

Networking

As seen from the figure, there are three levels of networking in the accelerator areas: the office LAN; the Main Control LAN; and the Regional control LAN.

For the office LAN, the accelerator divisions are following the TELECOM recommendations and soon all offices will be wired for Ethernet. It is thought that this will be adequate for the nineties.

On the traffic this may generate, on average only 20% of the users may be using their workstation at any given time. However at peak periods, eg. morning, lunch, evening, up to 80% of the workstations may be in use for mail, looking at machine performance, or checking up on equipment. Thus the network should be configured for almost maximum utilisation at say one transaction per 10 seconds per workstation, maybe 10 packets per transaction, say 1000 packets per second over the three networks with a substantial fraction, say 50%, going to the central machines, mainly VM.

The Main Control LAN is a subject which has preoccupied control system developers since the first successful introduction of the TITN network in the early 1970's. LEP have chosen the IBM Token Ring as it has a number of advantages for control system use. Networking now has a very wide application, however, especially within the large experiments. Much of the SPS/LEP networking effort on telephones and TDM will be combined with the DD effort in these areas in the future. The PS has only limited effort to devote to this subject. It was agreed that the accelerator divisions would be happy that a single entity, eg. DD/CS, should provide most of the LAN directions and support, with only support for local problems provided within the divisions. DD should also provide support for controls protocols, eg. TCP/IP moving later to OSI. This should include some measure of standardisation and support for Remote Procedure Call protocols of which at present there are three, the original one developed in PS plus two others, all different, one in DD and one in SPS.

Connectivity between the Main Control LAN and the office LAN should be provided. This can take the form of MAC level bridges if the LANs are of the same type (Ethernet), or IP gateways if the LANs are different but use the TCP/IP protocols.

DD should handle all matters concerning licenses and distribution of new versions, and provide grouped ordering. DD should also manage the main network file servers, especially from the point of view of backup. DD should provide the network management including names and IP numbers. Some critical servers may be local, managed by the controls groups, but others may be central and managed by DD.

Increasingly it is believed that a regional control LAN should replace the serial control highways more traditionally used. Connection between the Regional LAN, through controlled gateways of course, should be possible to the central database and microprocessor development facilities, and also to the office workstations for monitoring and maintenance. Although it is agreed that Token Ring (a standard in its own right), Uti - Net, Apollo Domain, and Apple Talk will be around for some time yet, future decisions should move towards reducing LAN diversity where possible. Here again the Technical Board recommendations should be given management support, in particular the TELECOM recommendation for Ethernet.

A number of home terminals will be required in the accelerator divisions, in connection with the more or less well specified piquet system. Although some engineers have expressed as desire to work from home, the mainline feeling in the accelerator divisions is that the official allocation should be restricted to those with on-line responsibilities. This will be about 30 per division, say 100 in total.

Chapter 7

Databases

The accelerator divisions are heavy users of databases of all types and the first main application for Oracle was the LEP planning and installation. It is agreed that the primary repository of accelerator information should be a large central relational database. This is particularly important as the lifetimes of the accelerators and their associated equipment are increasingly outliving a given person in a given post.

The lifetime of the database system should be at least ten years for reasonable stability. After an intensive inquiry ORACLE was chosen and used for the LEP planning and installation. It has grown until it now uses two big Vaxes, VXLDB1 and VXLDB2, running the cluster version of ORACLE. There is a support effort for ORACLE in DD of about 5 persons. Many users in the accelerator divisions are just now getting their applications running on ORACLE. It is agreed that the support effort is such that only one database management system can be handled at any given time. ORACLE Corporation seem to be keeping up well with the technology, and indeed are the leaders in certain fields such as SQL standardisation. It is therefore agreed that unless there is a major change in the market place, ORACLE should be the database system through most of the nineties.

For accelerator data there was a consensus that there should be one single central database. This could be quite a big load. This would imply some changes in DD operations. All major upgrades should be made during accelerator shut-downs, ie. January-February. A 24 hour service with on-line backup capability and an agreed maintenance service are required together with a piquet service of Oracle and system expertise. The support required in the accelerator divisions should only need to be applications expertise.

It is important that ORACLE be interfaced in a user friendly way to the office work station. Two approaches seem possible here. One is to use distributed database access such as SQL*net which essentially make the database available on the PC. The other is to use spread sheet programs on the PC and provide easy automatic transfer between the spread sheet and the database. ORACLE are working in both of these areas and the choice may be up to the user.

Accelerator controls have been reliant on on-line databases, often in the form of files in the control network. It was agreed that the primary source of control data should also be the single central accelerator database, along with all the other accelerator data. The PS are currently moving their main control data base to ORACLE on VM. This raises two problems, speed of access and availability. The availability problem should be solved in any case as mentioned above. Several approaches to the on-line problem are being investigated. One is to retain control system files and provide programs to swap data between ORACLE and the files when changes are made. A second is to use a fast on-line database such as C-Tree and again have programs to interchange data. A third is to use an on-line relational database so that the links between the on-line and central database will be provided as part of the manufacturer supported distributed database. It is not yet known which of these approaches will be best, but the third will put more demands on the central database support group.

A solid and fast network connection is required, possibly based on an Oracle product. There is a Remote Procedure Call interface to the Oracle VAX, but this is seen as a backup solution. Priority for small accelerator jobs could be required. It is important that Oracle be supported on a subset of agreed work stations with full networking, in particular the workstations used for accelerator control.

A tool to interface with Teamwork would be required if teamwork becomes the standard SASD tool. DD should provide the interface between the standard tools to get the best results.

Currently there are limitations on what can be held in and manipulated by Oracle. In the engineering database it is required to store the "geometry" of objects, not just numbers and strings. The whole concept has to be developed to handle this. At present work is going on to interface Euclid to Oracle. There is a need for unification of accelerator data and this will require concepts to handle 3D objects. At present dimensions are entered by hand. In 5 years or so we may have to replace Euclid, and it would be good if the eventual replacement product could be Oracle compatible. Graphics capability is also needed for applications in documentation, also for archiving and comparing waveforms to check if a signal looks as it should.

Oracle and Euclid should be on the same network, so that if, say, we move a magnet using MAD we can see the new layout directly using Euclid. LEP and SPS have a machine element database which is used by the machine, survey and installation people.

In the accelerator divisions, it is estimated that about 15 people will be involved in database design, and up to 30 software engineers, expert in database applications, will do applications programming. The number of designers may be reduced if there is a strong central nucleus of database engineers giving support to all divisions.

On user feedback, Oracle has three user groups, the VM users, Accelerator controls, and the application developers (mainly LEP so far). This arrangement seems to be working well. When the PC users build up in numbers a separate group could be formed. User presentations will become more important as Oracle expertise begins to spread.

It would be very convenient to have the same database for ADP and accelerators. Often in the design process it is important to have access to lists of manufacturers, and in operation to know the manufacturer, how many exist at CERN, contract numbers, etc.

Chapter 8

Accelerator Design and Modelling

Although not so demanding as physics event analysis, accelerator design is increasingly a heavy user of CPU cycles on the big central machines. Requirements include lattice design, wakefield calculations, field calculations, particle tracking, 3-D structure design. Linac design can be particularly demanding in these areas. At present the main client is LEP with MAD, but as LEP moves into operation and more resources go into the design of new machines, the computing requirements will grow. Some indication can be had from the American labs. These range from 1000 hours per year of Cray time (SLAC), to one Cray XMP/48 year equivalent (SSC). Much of the work is still done on IBMs, however, so the load will be split.

Accelerator design is becoming increasingly an interactive task requiring good graphics and a fast turn round time. There is a demand for the use of networked workstations, for example up to 20 Apollos for LEP, to meet this need.

The control systems are increasingly using modelling techniques in the on-line control of the accelerators. At present these programs are run off-line, or on-line on number crunchers attached to the control system. All this involves extra work and complication which would be removed if the modelling could be done on-line on the central machines. This will require improved and more solid networking connections, faster turn around for remotely submitted jobs, and some declared availability. The availability need not be 100% as the modelling need not be used in the primary control loop, but will be required for operation mode changes, and optimisation studies. The availability maximum requirement might be similar to that requested for the database machines. Some statement on the

on-line networked use of the IBM and CRAY are required from DD to enable the accelerator controls to plan correctly.

Chapter 9

Computer Aided Design

The accelerator divisions have a high concentration of engineers, and have contributed heavily to the report on Computing for Engineering. Indeed more than half of the representatives on the committee for Computing for Engineering were from the Accelerator Divisions. Despite this, or perhaps because of this, some discussion and representation from CAD is also included in this report on Accelerator Computing. This is to ensure that the CAD influence on general computing requirements is not neglected. This is limited, however, to mechanical CAD, databases, and the most common requirements of the control system. To redo all the CAD requirements of the Accelerator Divisions would be to duplicate the work of the Computing for Engineering committee. Not to mention this area nor to consider its impact on workstations, networking, and the control system would unbalance this report on Accelerator Computing. Therefore any mention of CAD requirements in this chapter and elsewhere in this report should be considered as inclusive and not exclusive. For a view of the requirements of CAD the specific report on Computing for Engineering should be taken as the basis for requirements in the Accelerator Divisions.

The rest of this chapter is concerned with mechanical design. As mentioned above, the subject of Computer Aided Mechanical Engineering is also treated in parts 3 and 4 of the report on Computing for Engineering.

CAD started at CERN in 1982 with the introduction of EUCLID on a VAX780, and has expanded till in 1988 we have EUCLID-IS on a VAX8650 plus a 785 connected to about 30 dedicated work stations. This gives a total of 7 - 8 MIPS computing power. In the near future we will introduce VAX stations in a Local Area Vax Cluster (LAVC) with a 3500 file server connected to the cluster. DEC's new 6210 multi processor number cruncher and database server will replace the 785, and 7 new VAX stations will soon be put into service.

Drawing started in 1984 with the introduction of AutoCad on PCs. At present there are about 50 Autocad licenses at CERN, bought at prices ranging from 7000 francs down to 1000 francs, clearly a case for coordination. Of these 25 are in the SPS, 3 in LEP, 3 in the PS, 1 in DD, and the rest in EP and EF.

Work is in progress to link Euclid and Autocad. Next year there will be a mini-Euclid on PCs which will be the rough equivalent of Autocad, so Matra-Datavision themselves are not interested in links to Autocad.

More than 60% of the LEP drawings have been produced with Euclid. The layouts for the whole machine are available on the database. In the SPS only 7 people, none engineers, use Euclid. This could be because the SPS are doing less in-house design and using more sub-contracts for design and development. The main use in the SPS, as in the PS, is for maintenance type changes in layouts, and for detailed layouts of tight systems such as target areas. There is only a little use of Euclid for CAM and links to the workshops, mainly in ST.

Within 5 years CAM could be vital to CERN, with the use of CAD in-house for all design, and the diskettes sent to outside industry for manufacture. There will probably never be more than 5 or 6 numerically controlled machine tools at CERN.

Communication is vital to be able to send designs around from system to system, even or especially to different systems. There is an IGES standard which Matra - Datavision are supporting which may help with this problem. Autocad was originally stand alone in use, but now many of the AUTOCAD systems are linked over Ethernet using NOVELL in the PS and SPS.

The support required from DD could be quite different for the high and low level systems. The high level system should be staffed and run by DD, whereas for the low level system DD should organise licences, distribute software, provide courses, etc.

There is not yet a link between Euclid and the finite element packages used at CERN. Euclid is a 3-D package, however, and can calculate weights, inertias, and other geometrical attributes. It is too early to have links with the finite element packages for stress analysis, etc, but this would be very desirable in the future. This might mean moving the finite element packages from the IBM to the VAX but this would put too big a load on the VAXES. At present we use ANSYS which is an American package and CASTEM from the CEA in France. There is a possibility of buying a CRAY vectorisation package for CASTEM.

To reduce the cost of using expensive packages for limited applications, it may be possible to use them across the network in our sister institutes. For example a lot of the AA design work was done using programs (TOSCA) on the Rutherford computers. This could add to the wide area networking load.

The current 7.5 MIPS of CAE VAX power will suffice in the medium term, going up to perhaps 10 MIPS, but with increasingly powerful work stations at the user end.

A multi-purpose mid-range work station (PC with Autocad typically) will be required per engineer, about 30 in total for mechanical engineers in the accelerator divisions. A dedicated work station (Euclid type, probably VAXstation) will be required per 1.5 to 2 designers, again about 30 in total.

Software licensing will become important as this could become an expensive item. Site licensing would be ideal, or perhaps a license for N simultaneous users with no restriction on the number of potential users.

Personnel utilisation for mechanical CAD is at present 3.5 specific staff in DD plus another 2 in supporting activities. Note that 2 of the 3.5 are leaving at the end of 1988. Applications support in the divisions is about 4.5 full time equivalent, mostly on a part time basis. In the future it may be necessary to use fellows and people on shorter term contracts for CAD work. A new person can become productive after about six months training, which is a reasonable investment for 1.5 to 2.5 years productive work. People with a few years experience in CAD at CERN have no trouble finding jobs elsewhere.

Chapter 10

Local Computing Centres, Servers, Central Support.

In the past the accelerator control groups have set up fairly powerful local computing centres because their needs were new and specialised. These centres grew to handle jobs such as accelerator modelling, controls program development, microprocessor software development, controls documentation, MIS applications for the controls and operations personnel. This will change, however, as CERN is setting up new central units to handle things like MIS and CAE.

The PS have two NORD - 500s, one for modelling and one for general applications. Also the PS have a local centre with two NORD - 100s serving the secretariats. Running these computer centres takes

time and money and is something the PS feels could better be done centrally. Originally the microprocessor work was done locally on the NORD-500 but now this has been shifted to the PRIAM VAX. A lot of the PS Controls group MIS work used to be done on the local computers but with the divisional emphasis on VM this is shifting. The controls documentation is still on the NORD but this could move to a CERN-wide scheme. For controls accelerator modelling, advantages are seen in doing this on the same central machines as are used for the accelerator design, using as far as possible the same programs. This will require, however, some decisions on the central machines' availability for networked real time job submission. Thus the PS see the eventual replacement of the local controls computing centre replaced by general central facilities. On the other hand the SPS are happy with a local facility for accelerator modelling (currently a NORD-500, possibly an Apollo DN-10000 in the future), as they find it easier to get computing time and a good interactive service. Also it is easier to connect to the control system. LEP would like to move to workstations for both accelerator design and modelling.

As mentioned in several other parts of this report, the new model of accelerator computing is the use of personal workstations connected through a network to a variety of resources. This applies to both the office workstations and the control system workstations. The network resources take the form of servers. The control systems have file servers for their data and programs. For office servers, both the SPS and PS use NOVELL servers over Ethernet. These are used for PCAD, AUTOCAD, and general MIS applications. LEP is likely to go in this direction as well. This concept is replacing the traditional concept of the local computer centre run by the controls groups in the accelerator divisions. The servers can be located close to the people providing the support for that particular service, yet be available to anyone on the network, either in the particular division, or even throughout CERN. Thus the NORDs will see their role reduced to becoming servers on the network. In the case of the NORD-500s this will be to provide a real-time on-line computational facility. As mentioned above, this function could even be moved to the central machines using the network. The evolution of the administrative support NORDs will be in the direction of becoming a server on the network for the PS secretariats.

A particular issue relating to local and central computing is backup. It is agreed that backup facilities across the network to centrally provided and run facilities is essential. This would apply to local computing centres, local servers, and even office workstations. For the servers a daily incremental backup is required, already done partially at the PS.

Chapter 11

Control System Specific Issues

An important item is the cross software support for the Motorola 68000 family. This includes the RT kernel, at present RMS68K, and the cross compilers. at present C, Pascal, Modula II, and FORTRAN. Also the linker/loaders, symbolic debugger, and executive libraries. The aim should be to move to industrially supported products. The industrial cross development facility for OS9, UNIBRIDGE, is urgently needed for large project work in controls. The use of cross software must be carefully evaluated. It was generally agreed that cross development was about an order of magnitude more difficult than native development. Native development, however, requires provision for disks etc. at least in the development phase. One of the advantages of OS9 is that it can easily allow the development environment to be removed later. This was one of the driving forces for the 68000 in G64. An advantage of the cross software is that the central machine can provide better facilities for project management. In any case the accelerator divisions will require the PRIAM support for the 68000 and RMS68K to continue well into the nineties until a replacement has been found, run in, and a reasonable changeover period has elapsed.

In accelerator control systems and equipment there is indeed a requirement for embedded microprocessor systems, say on a single card, for which a cross development environment is called for. However, much of the higher intelligence could be placed up a level where native development systems are used. Anyway DD should support a native system, for example native OS9 for the 68000 family.

An Computer Aided Software Engineering (CASE) toolkit should be supported by DD with full support for the accelerator divisions as well as the research divisions. This should include tools for the specification, design, development, and exploitation phases of big software projects. It should work in liason with ORACLE. It should in particular cover the requirements of cross software projects on the CERN standard microprocessors. The choice of tools is a difficult matter. Often they are first chosen by user groups, eg. SASD, NOVELL, DV - DRAW. This is normal as the user is usually the first to perceive a problem and see a solution in the literature. Once it becomes established technology, however, it can be taken up and supported centrally. Who then should drive the evolution? The user who still best perceives the requirements, or the central support team who risk putting effort into things nobody wants? There was consensus that the initiative should normally come from the users. Note that the specification tools for databases are evolving on different but parallel paths with the SASD tools for program specification. The hardware, work station and networking aspects of the tools support should be set up, or at least specified, by DD.

Accelerator controls are big users of electronic CAD and CAE, but with no particularly difficult requirements. The main uses will be for digital printed circuit board design. Services will be required for programmable circuits such as EPROMS, PALs, FPALs. This should include support for the compilers and provision of equipment for burning, equivalent or better than is provided at present.

The use of artificial intelligence and expert systems is expected to have a big impact on accelerator controls and operation. At present a limited amount of manpower is being put into exploratory work involving th KEE system. When this reaches the stage of generalised application, considerable extra effort will be required to realise its promise. At this time a central support for these new facilities will be required.

Controls requires a lot of documentation with the associated classification and printing. Much of this is done at present using control system specific facilities. In the future these should be provided on a general basis. Good quality (Laser) printing should be available on the corridor level for all applications including control. This raises a question of support for the corridor printers. An overall support operation could be quite expensive in manpower, so continuing with an ad hoc combination of secretarial and controls group support might be the only way to spread the effort.

Chapter 12

N - manufacturer policy

There is agreement that CERN as a whole should not be dependent on any single manufacturer, and that competition is vital for good services and prices.

Accelerator controls have had a 15 year involvement with Norsk Data which has illustrated the good and bad points of a single manufacturer policy. The SPS became very conscious of the bad points, and after a period of in house designed computers, have settled for an "open systems" approach. This involves the UNIX operating system and TCP/IP protocols, with hardware available from a number of sources. To retain some of the benefits of a single manufacturer policy the SPS has formed a close relationship with Apollo for workstations. The PS have also felt the need to move to workstations and have chosen the DEC Vaxstation. The PS have also decided to adopt the "open systems" approach with UNIX, and will run ULTRIX (DEC's version of UNIX) on these computers. This

provides a common system across both controls groups, and will facilitate any future changes in hardware. DD already has a number of support activities in the UNIX, TCP/IP, Apollo, and Vaxstation areas, and clearly the Accelerator divisions expect this expertise and support to be continued and strengthened in the future.

For the general purpose office workstations the accelerator divisions tend to prefer the IBM compatible PC types. Partly this is because of the wider range of general purpose software available. Also, however, the accelerator divisions are more stable and must take a longer term view on the re-usability of hardware and software which favours the use of the most widely spread material. This may result in a different choice from the research divisions which could prefer more state of the art material. Nevertheless it is essential that all CERN services be accessible from a single office workstation, so if there are "N-manufacturers" of office workstations at CERN, many central services will have to be made available on all approved types of workstation.

Chapter 13

Support and User Feedback

There has been a feeling in the accelerator divisions that DD interprets its mission as too much in favour of the research divisions and not enough in terms of support for the laboratory as a whole. Hopefully this is changing with the establishment of the MIS unit and Computing Support for Engineers. Also the Accelerator Control Groups have been a bit "go it alone" in terms of computing technology, and again this is changing due to pressure on manpower. Hopefully in the nineties all divisions will work together to provide a range of agreed CERN-wide options.

DD should see to all matters of purchasing, licensing and registration for the approved range of products. An important role of DD will be the networking and communication sitewide. It should be ensured that all equipment can communicate well, and with the outside world. This could be a main guideline in the selection of which material should be used.

Technical support should be supplied by DD, in particular for the central services such as the central interactive service, central batch service, and central database facilities. However a local informatics support is also needed, and in several examples from industry, and in the accelerator divisions experience, a figure of 1% seems to be required, making about a dozen for the accelerator divisions. This local support effort should work closely with the centrally provided DD support activities.

User Feedback can be provided on two levels. CERN-wide boards such as SCAIP and TELECOM with the chairman outside DD seem to provide good guidance at the top level. At the lower working level, boards coordinated by DD for special application areas such as data acquisition, MIS, CSE, PRIAM, DEC user coordination groups, etc., are appropriate.

Chapter 14

Resources

The resources allocated to computing within the Accelerator divisions can be divided into three areas; Controls, Databases, and General Informatics. Controls are the dominant consumer of resources at present, but Databases and General Informatics are increasingly demanding.

The accelerator control groups currently employ 90 persons for SPS/LEP and 44 persons for PS. Only a small fraction of this effort is directed to end user accelerator control applications. These applications are, of course, essential and should obtain more effort in the future. The small percentage of the total effort going to these end user applications is due in part to the old SPS philosophy that end user applications should be done by the end users themselves, ie machine physicists, equipment specialists, and operations personnel. PS and LEP controls have, however, officially adopted the professional applications programmer approach. The bulk of the controls effort has gone into hardware, system software, and applications support structures. There is some hope that rationalisation and the use of industrially supplied products will liberate more effort for end user applications in the future.

Database applications were pioneered by LEP with ORACLE and are still limited in the PS and SPS divisions. As discussed in chapter 7, it is estimated that about 15 people will have to be allocated within the accelerator divisions to program applications if the database benefits are to be fully realised.

General Informatics support has grown up in a rather ad hoc way, partly within and partly outside the controls groups. About 10 people can currently be identified as providing informatics support in some way, but it is estimated that a more coherent effort of about 15 people will be required in the future to provide adequate local support, including local network support as requested by the TELECOM recommendations.

Provision of strong central support in DD is essential if the above local support is to be most effective. Also good central support is vital to achieve integration and economy with the rest of CERN's computing effort. The accelerator divisions therefore support the requirements of DD for effort in the MIS, CSE and Database areas. It is pointed out, however, that a correct balance between central and local support must be achieved as no amount of central effort can replace support close to the application.

**Computing for Engineering
at CERN in the 1990s**

9th December 1988

Journal of the
Royal Society of Medicine

Computing for Engineering

Composition of the Working Group.

The main Working Group was composed as follows:

- *Convenor: David Jacobs/DD*
- *Digital CAE/CAD: Ludwig Pregernig/EF*
- *Analog CAE: Guy Baribaud/LEP*
- *Mechanical CAD/CAM: Claude Hawiller/LEP*
- *Structural Analysis: Detmar Wiskott/DD*
- *Field Calculations: Hans-Horst Umstätter/PS*
- *Microprocessor support: Fabien Perriollat/PS*
- *Database Support: Josi Schinzel/LEP*
- *General Advisor: Pier Giorgio Innocenti/SPS*

Appendix A

Table of Contents

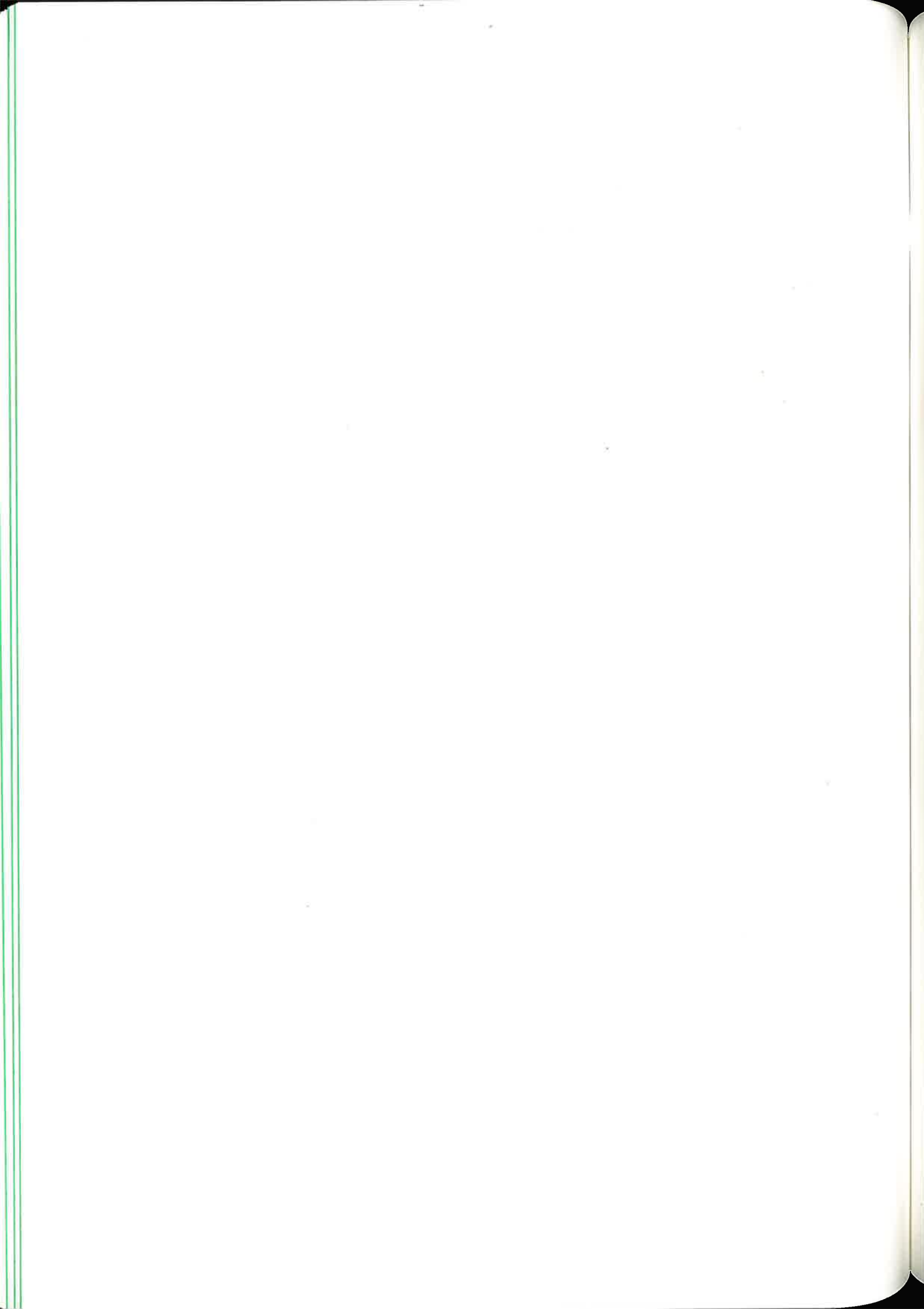
1. Introduction	1
2. Methodology	2
3. Results	3
4. Discussion	4
5. Conclusion	5
6. References	6
7. Appendix	7
8. Bibliography	8
9. Glossary	9
10. Index	10

Contents

Computing for Engineering	iii
Composition of the Working Group.	iii
PART 1: Introduction.	213
PART 2: Support for Microprocessor Users.	215
2.1 Introduction.	215
2.2 Hardware Support.	215
2.3 Software Support.	216
2.4 Training	217
2.5 Resources.	217
2.6 Summary of Other Recommendations.	217
PART 3: Computer Aided Mechanical Engineering and Related Fields.	219
3.1 Foreword.	219
3.2 Introduction.	219
3.3 Present Situation.	220
3.4 Near and Far Future.	220
3.4.1 Integration.	220
3.4.2 Population.	220
3.4.3 Hard- and Software.	221
3.4.4 Structures and Resources.	222
PART 4: Structural Analysis.	223
4.1 History.	223
4.2 Present situation.	223
4.3 Development into the 1990s.	224
4.3.1 User population.	224
4.3.2 Programs for structural analysis.	224
4.3.3 Computing.	225
4.3.4 User support.	225
4.4 Estimates of resources.	225
PART 5: Software Support for Field Computations.	227
5.1 Introduction.	227
5.2 Existing Field Computation Programs in use at CERN.	227

5.2.1	Static fields in 2 dimensions (2D).	227
5.2.2	Static fields in 3 dimensions (3D).	227
5.2.3	Electromagnetic cavity fields and others.	228
5.3	Present and Future.	228
5.3.1	TOSCA.	228
5.3.2	ANSYS.	228
5.4	Budget.	228
5.5	Human resources.	229
5.6	Hardware.	229
PART 6:	Analog Electronics CAE/CAD.	231
6.1	Introduction.	231
6.2	Fields of Application.	231
6.2.1	Integrated Circuit Design.	231
6.2.2	Discrete Component Analog Circuit design.	232
6.2.3	System Design.	232
6.2.4	RF and Microwave design.	232
6.3	User Population.	233
6.4	Conclusions and Recommendations.	233
6.4.1	Private versus collective access to CAE/CAD.	233
6.4.2	Organisation of activities.	234
6.4.3	Some hints for foreseeable investments.	234
6.4.4	Education and human resources.	235
PART 7:	Digital Electronics CAE/CAD.	237
7.1	Present Situation.	237
7.2	Future Trends.	238
7.3	Recommended Approach.	238
7.3.1	Resources for Digital CAE/CAD at CERN for the 1990s.	238
7.3.2	Management Issues.	241
7.3.2.1	Training.	241
7.3.2.2	Infrastructure.	241
7.3.2.3	Design Management.	242
7.4	Benefits.	242
PART 8:	Database support for Engineering.	243
8.1	Present Situation.	243
8.2	Advances in database technology.	244
8.3	Development of ORACLE.	245
8.4	Engineering Database Service Hardware.	245
8.5	Database exploitation support.	246
8.5.1	Central Database Support.	246
8.5.2	Local Application Development Support.	247
8.6	Budget for the database service and exploitation.	248

PART 9: Training	249
9.1 Topics for training.	249
9.2 Organisation of training.	249
9.3 The duty to train.	250
9.4 Summary of recommendations.	250
Bibliography	251



PART 1

Introduction.

In the first Green Book "Computing in the LEP era", Computing for Engineering was not recognised as a separate heading and the associated topics were barely touched on. That five years later it should merit a chapter of its own, is a measure of the extent to which the engineering community at CERN has come to use computers as tools in recent years and the vastly increased sophistication of their needs. Indeed in several areas, such as the use of formal software engineering techniques, they play a leading rôle in the laboratory. As a corollary, strong requests have built up from the technical divisions for increased support from DD Division.

The working group which prepared this report was composed of the chairman of the former Advisory Committee on Computing Support for Engineering (ACCSE) [3] along with experts from each of the fields of engineering computing which ACCSE had identified. The only exception was Computer Aided Software Engineering where the working group on Computing for Experiments was left to handle the topic for engineering as well. There is unavoidably some overlap with other chapters, especially those on Computing for Experiments and Computing for the Accelerator Divisions. This overlap occurs more or less at random and the fact that a topic is repeated in several chapters should not be taken as evidence of any special emphasis.

Each expert was left to prepare a section on his field as he saw fit. Most organised formal subgroups which met several times to discuss the contents. Others simply prepared a draft and circulated it to relevant people for comment. The drafts were then merged by the convenor. In cases where they were significantly shortened, the original papers are referred to in the bibliography.

PART 2

Support for Microprocessor Users.

L. Casalegno, R. Dobinson, C. Eck, H. Muller, F. Perriollat (chairman), P. Van Der Stok

2.1 Introduction.

The community of microprocessor users at CERN and collaborating institutes is very large, including both physicists and engineers, and it can be assumed that the impact of this technology will continue to grow during the next decade. Some differences exist between the engineer and physicist communities. The engineers are looking for established industrial standards and products and they expect a system lifetime of 5 to 10 years. The physicists are more looking for frontier technology and for highest possible real-time performance, to be able to cope with the expected data rate in the experiments of the 90's. The lifetime of their systems is normally shorter. Nevertheless, it is to be expected that the requirements of physicists and engineers can largely be met by the same set of tools.

2.2 Hardware Support.

The current activity is a good model and should continue in the same spirit in the future.

The low end microprocessors (8 bits) should not be included in the central support. Support must be concentrated on a limited number of microprocessor families; the M68K, which is currently widely used, probably a reduced instruction set computer (RISC) family (for example the M88K), and likely also the Transputer, if the current projects testing its use will lead to a request for support to the User Committee. The Intel 386 family, even if it will dominate the personal computer office automation market, will probably not have a strong impact in real-time applications at CERN. This is due to the current very strong engagement in the Motorola family and the fact that the market of VME boards with the Intel 386 is not very active. This assumption must certainly be reviewed periodically.

At the board and bus level, activity must continue in the areas of evaluation, consultancy and participation in standardisation committees. A key function, however, must be to maintain close relations with industry to negotiate the best commercial and technical support to CERN. As well as giving continued support for the VMEbus family, the evolution of next-generation commercial bus systems should be monitored.

For the difficult problem of system integration, which will continue to be the responsibility of the user, some expertise and advice is expected from the central support organisation. It is not, however, foreseen that the central support organisation will provide much support for naked processors, nor will it

be providing support for FASTBUS and CAMAC equipment.

2.3 Software Support.

In the software area, which addresses both the development facility and the run-time installation, two classes of environment need to be supported; a microprocessor operating system with native development tools and a host-target configuration.

For a native operating system like OS-9, which is currently very widely used at CERN for the M68K family, the support should be mainly concentrated on logistics: licences, distribution of software and documentation and the organisation of user fora. The future evolution of this type of system, e.g. bridges to UNIX and VMS, will be followed.

The next years will see important new industrial products emerging in the area of host-target development systems. A high quality and powerful environment must be provided for this class of systems. The chosen environment must be available for DEC's VAX-VMS, for UNIX and probably also for workstations using the IBM/Microsoft OS2 system. It must be well integrated with the selected Computer Aided Software Engineering (CASE) tools. This environment must be able, for new applications, to replace the currently supported Real Time kernel RMS68K in the near future. Important criteria for selection are the quality of debugging facilities for programs and systems, and good integration of communication. The DD central support group should continue to make this development environment available on a computer (or a cluster of computers) used as a server by user workstations as well as by users who only have access to a simple terminal. The existing system, a VAX 8530 running ULTRIX, will have to be regularly adapted to growing demands on its capacity. It is assumed that over the next three years most of the additional computing power required for this development style will be provided by workstations, bought by the users and networked to the central system. It will thus probably be sufficient during this time to expand only the disk capacity of this system but a major upgrade still has to be foreseen in the first half of the 90's. It should be noted that the PRIAM VAX supports a number of general services which are not connected to microprocessor support and it will continue to do so in the future.

The programming languages will continue to be C and FORTRAN, with perhaps PASCAL in addition. Modula 2 is not much favoured. There is already pressure from the technical divisions for ADA support in the 90's and it can be assumed that the physics community will follow this lead. Interpreted languages, like NODAL and PILS, are well accepted as testing tools in both engineering and physics applications. Based on lower level tools provided by the central microprocessor support, these languages have been developed in CERN groups working directly for end-user application support. It is assumed that this remains so in the future. It is clear that the implications of attempting to offer support for a RISC processor and/or Transputers must be evaluated. It seems to be unavoidable that, during a trial phase, a number of incompatible processor architectures will be used at CERN. Nevertheless, based on the experience gained in this field, this number should be reduced as soon as possible to the smallest value compatible with the needs of the user community, in order to be able to give a reasonable level of support at all.

A centralised CERN-wide support for microprocessors is strongly recommended in order to reduce duplication and to increase the quality of the service provided. The effort must be concentrated on integration of well established industrial software components. No more in-house development of basic components seems to be necessary. Expert advice to the user remains an important aspect. Unless,

however, the central service is properly funded and staffed, it risks to be ineffective and alternative ad hoc solutions will inevitably appear.

2.4 Training

It is recognised that the microprocessor user community at CERN will only be able to cope with the continuous evolution in this field, if sufficient training possibilities for all levels of usage of the supported tools are offered. It will be the task of the central support group, in collaboration with CERN's academic and technical training services, to provide the necessary courses for this.

2.5 Resources.

- As recommended by ACCSE [3], the manpower of the DD central support group should be increased by 2.5 to the level of 7 (including one Fellow).
- The annual operating budget required is estimated to be 250 kSFr.. The expected level of regular central investment will be about 250 kSFr. per year.
- A major upgrade of the central support computer (500-700 kSFr.) will have to be financed in addition to the above regular budget.

2.6 Summary of Other Recommendations.

- A central support service must continue to be provided and user demand for such a service will continue to grow.
- Hardware support must continue according to the current policy with an open eye on new technology: RISC, Transputer and new buses.
- A substantial effort must be invested on the software side to renew the facilities, using industrial solutions. A good software development environment which includes CASE and management tools is strongly requested.
- In addition to User Information Meetings, a User Committee, composed of representatives from the different parts of CERN, the collaborating institutes, and from the DD central support group, should meet regularly to advise the support team on policy and the short and medium term work plans.

PART 3

Computer Aided Mechanical Engineering and Related Fields.

C. Hauviller (Chairman)

3.1 Foreword.

This part is intended to be a summary of a series of papers on the future for Computer Aided Mechanical Engineering published during the last two years [1-9]. The draft document has been forwarded for comments to the following people: G. Bachy, O. Bayard, G. Cavallari, D. Jacobs, R. Mackenzie, M. Mathieu, R. Messerli, S. Oligier, A. Poncet, M. Price, P-L. Riboni, D. Wiskott.

3.2 Introduction.

Computer Aided Design (CAD) has been considered for long as a stand-alone subject and this still remains often the case at CERN, but the general evolution in firms of any size is towards as complete an integration as possible of the very diverse mechanical and electrical engineering activities. To extend this notion of integration to CERN does not seem unrealistic, even if it must be adapted to the organisation's specific activities and environment. For this reason, the present section, although mainly dealing with CAD, also emphasises the larger notion of Computer Aided Engineering (CAE).

After quite a slow start in computing for mechanical engineering, software adapted to accelerator and detector technology is now widely used at CERN, especially for the present very large projects. In the mid 70's, the first finite element method (FEM) packages were installed on the CDC 7600. Ten years later, packages provided by software firms are resident on the central computers, together with smaller ones installed on PC's. This subject of Structural Analysis is dealt with in Part 4 on page 223.

The next step in the introduction of computer intelligence in mechanical engineering has been numerically controlled machining. This has, however, been limited to a small number of machine tools by the fact that only a very small amount of equipment is and probably will be actually produced in-house.

The major breakthrough has been CAD. The decision to install a CAD system in the technical divisions was made in 1981 and one year later, after a detailed technical evaluation, a true 3-D solid modelling system, EUCLID from Matra Datavision (MDTV), was selected as an interdivisional project. The choice of a 3-D package oriented towards design instead of a 2-D (drawing) one was difficult. The market trend was still towards the simplest packages and solid modelling software was in its infancy. It was considered, however, that the potential usefulness of solid modelling was much greater, in particular the assembly capabilities allowing the integration of various parts of accelerators and detectors. EUCLID was installed on a VAX 780 with workstations in the form of simple storage tube graphic terminals.

3.3 Present Situation.

Six years later, the new user-friendly version, EUCLID-IS, is running on more than thirty dedicated workstations connected to two clustered VAXs, a 8650 and a 785. More than 60'000 geometrical entities are stored in the site-wide database. The computing power has been multiplied by 8 since 1982.

Besides this powerful design system, a simpler drawing package, the very widespread commercial product AUTOCAD, has been introduced recently. It runs on IBM-PC compatibles, usually in a stand-alone configuration. About 35 licences of this software are in use at CERN.

Both packages are complementary and have their own advantages: EUCLID, built around a central infrastructure, is more powerful but AUTOCAD is cheaper and easier to learn. The former is adapted to large projects which have to be co-ordinated and require a major database whereas the latter is preferred by individuals working on little stand-alone projects. Moreover, both are in use in the High Energy Physics community.

Mechanics is the main application of these packages but EUCLID is also used for infrastructure services such as civil engineering, electrical engineering, survey (in addition to the ESPACE and LILIAN packages), etc..

3.4 Near and Far Future.

3.4.1 Integration.

Communication is a must in large engineering projects like accelerators or detectors. It is, therefore, unthinkable that CAD should live in autarky. The general industrial trend is towards as complete as possible an integration of the various engineering tasks. The aim is to gather around a unique database, under the term Computer Aided Engineering (CAE), diverse activities such as design, computations, drawings, schematics, simulation, manufacturing, maintenance, etc..

Only design and drawing are considered here. All the administrative tasks such as budget, planning, etc. are and will be done using tools like ORACLE, available on a CERN-wide basis accessible through a local or general network.

3.4.2 Population.

The population involved in mechanical engineering is presently about fifty engineers and more than hundred full-time designers [2]. To these numbers, one should add engineers and designers in related fields (infrastructures) and many occasional users preparing (usually small) drawings and sketches.

It can be assumed that this population will remain constant or decrease slightly during the next decade. However, many of them are not yet very familiar with computers and so it can be expected that the CAE user population will grow. In addition, subcontracting will take more importance and this stresses the need for good means of computer communications towards the outside of CERN.

3.4.3 Hard- and Software.

After an initial period of investment, only a maintenance budget was awarded to the EUCLID project during two years. The consequence of that situation was a lack of continuous hardware upgrading. This upgrading has only been re-started during the last year and the expansion of the system (number of users) will only resume by the end of 1988. The total number of dedicated workstations is likely to be about 50 by 1990.

All the future upgrades are expected to be based upon a mixed (CI/LAVC) VAX cluster configuration, the workstations being VAXstations (VS315 and successors) and lower price stations supported by MDTV (MAC II, SUN,...) and the central computer, used as number-cruncher and central database, a multi-processor VAX.

This upgrade is planned in four steps:

- create a LAVC with six VS315 VAXstations and a server (1988).
- replace the old VAX 785 by a VAX 6210 running under VMS 5 (1988).
- study running VS315s in mixed cluster configuration with special attention to the communication problems (data transfer); add VAXstations (1989).
- upgrade the VAX 6210 by adding more processors and phase out the VAX 8650; add VAXstations (1990).

The EUCLID software has been continuously upgraded toward a more user-friendly product. A very powerful 2-D package, available at the beginning of 1989, will provide functionality equivalent to AUTOCAD.

Some 3D features have appeared in the latest versions of AUTOCAD. In addition, the power of the IBM-PC compatible platforms is increasing. Budget constraints presently limit the number of users of EUCLID and so AUTOCAD use at CERN has doubled during the last twelve months. Since AUTOCAD is considered to be a good introduction to CAD and a standard in collaborating institutes, this trend will probably continue for a short time. However, the difference in investment between the two products decreases rapidly and will become negligible in less than two years. This will lead to an equilibrium between the two packages, the choice depending of the type of work, its aim and the environment of the user.

The 2-D drawing packages (AUTOCAD and EUCLID-IS-2D) should be usable on the multi-purpose mid-range workstation which will be found on everybody's desk (the last release of AUTOCAD is available on μ VAX, MAC II, Apollo,...). The two packages should be fully interfaced to each other.

To support more than two packages (EUCLID and AUTOCAD) will be too expensive (direct and indirect costs) and inefficient. Introduction of new programs must be avoided except if it can clearly be demonstrated that their features are much better than the existing ones. However the official choice should be re-examined periodically in terms of technique and finance, especially if a new large project starts.

All the numerical control applications have been integrated in the CAD environment through EUCLID under the same user interface. The numerical control language available is APT. The present capabilities satisfy all the needs met until now and it is recommended to proceed in this direction. Future Computer Aided Manufacturing (CAM) subcontracting will also be based on the same principle.

Integration should be the main priority to avoid duplication of work and stored data, especially for large projects. The 3-dimensional geometric description provided by EUCLID already allows partial integration but a complete object definition encompasses more than geometric data. Technical (specifications, tolerances,...) and administrative (drawing number, supplier,...) data should be added. The open structure of a Database Management System (DBMS) like ORACLE is the obvious core for such an integration but standard databases must evolve to handle correctly the combination of geometric, text and numerical data found in engineering.

3.4.4 Structures and Resources.

The ACCSE report [3] has already made a series of recommendations which should serve as basis for the next years:

- a user representation body : Computer Aided Engineering Committee (CAEC),
- a central support group in DD division, tackling informatics support and, together with engineers in user divisions, applications support (this includes a request for two extra staff),
- an operating budget for mechanical CAE (including structural analysis software) of 700 kSFr./year and an average annual investment of about 700 kSFr. in order to keep pace with the evolution of the CAE market both in hardware and software.

PART 4

Structural Analysis.

F. Bertinelli, M. Bona, D. Jacobs, A. Lefrancois, R. Maleyran, A. Poncet, M.J. Price, D. Wiskott(chairman).

4.1 History.

During the last 15 years a number of engineers and designers have gained access to programs that permitted them to analyse structures for their mechanical and thermal properties. These structural analysis (S.A.) programs were obtained either through purchase or with the help of other institutes. They perform calculations on models that approximate the structures with an assembly (mesh) of finite elements. In the beginning, pre-processing (mesh generation) and post-processing (result evaluation) was done in a more or less ad-hoc manner. The matter was an affair of specialists who wanted to solve their specific problems. These programs run on the CERN mainframes, originally CDC computers. Some of them (e.g. SAP-V, DOT, BOSOR) were later also transferred to the IBM under VM or to VAX under VMS (often without the interactive facilities that meanwhile had grown around them).

Beside the older programs, three products are used at present: ANSYS, sold by Swanson Analysis Systems Inc.(USA), CASTEM, designed by CEA-Saclay (F), and CASTOR, written by CETIM (F). For ANSYS, CERN profits from very advantageous (educational) commercial conditions. With the CEA, CERN has a development contract for CASTEM (and therefore has also access to the source code). ANSYS and CASTEM run on IBM under VM, and a reduced version of ANSYS can also run on PCs under MS-DOS. CASTOR runs on VAX and PCs. For the first two programs updates are regularly received from the suppliers of the code. Entirely new code for CASTEM ("CASTEM 2000") has been received but only partly been installed yet. CASTOR, limited to 2D problems, has been acquired as a simple tool to serve the needs of users in drawing offices.

4.2 Present situation.

The number of users of these programs, in particular CASTEM and ANSYS, is increasing slowly but steadily. This increase corresponds on the one hand to the growing complexity of the engineering tasks, where performance is often pushed to the technical limits, on the other to the desire to have more insight into the behaviour of structures conceived in day-to-day work, to reduce the necessity for prototyping and to optimise the final design. Increased acceptance by the users due to improved program performance plays an essential rôle as well.

The number of users, excluding occasional ones, is around 30 engineers, designers and applied physicists. Fellows and visitors (who carry a substantial part of the work load) are counted in these figures.

Support for the user population is given either by experienced users or through them (mainly on an ad-hoc basis and for special informatics problems) by the DD user support (DD-US) group. A num-

ber of courses have been given on the theoretical basis of structural analysis and on the use of ANSYS and CASTEM.

Considering a given user, work with S.A. programs is not a continuous activity. Very often, however, solutions of some problems require substantial resources in computing power and disk space for a limited time. It is not always easy to find these resources readily.

4.3 Development into the 1990s.

4.3.1 User population.

Extrapolating from the present population and assuming improvements in all kinds of user support, the number of regular users, during the next five years, probably will not go beyond 45. It is probable, however, that their access to S.A. programs will become more intensive and more demanding on resources, as they will have to face the very involved design problems of future accelerators and detectors. Simultaneously, with better user interfaces and improved support, the community of occasional users will grow, but this activity is difficult to quantify.

4.3.2 Programs for structural analysis.

It is not the intention of the working group to recommend here explicitly and on the basis of the actually installed versions, one or other of the programs that are presently available. They all have their merits, are known to their users and cannot be abandoned within short time.

The working group intends to study the technical features of a range of programs. It is agreed that, for the selection of general purpose programs to be maintained or introduced and supported at CERN, particular emphasis, apart from powerful algorithms and needs for resources, should be put on user friendliness and interactiveness. Versions with identical user interface (syntax, man-machine dialogue) should be available for different levels of problem solving power (e.g. running on PCs, workstations and large mainframes). Even for problems that demand mainframe computing power, pre- and post-processing should as far as possible be shifted to workstations or PCs.

Structural analysis in the above defined sense should not be seen as an isolated domain. Integration with other aspects of CAE is becoming increasingly desirable. Magnetic and electrostatic fields as well as thermal effects have repercussions on the behaviour of mechanical structures.

Bridges should exist between the domains of structural analysis and field calculations. Standardised data formats used for description of the problem (e.g. meshing) and for the output of results would permit the transport of data from one aspect of the problem to the other. The availability of source code for "peripheral" programs could help the mutual adaptation.

Modern CAD systems offer a natural method to describe the geometrical properties of a structure to be analysed.

Efforts must be made to profit from the interactive modelling facilities offered by CAD systems (e.g. EUCLID) and of the possibility to implement a CERN-wide design database.

4.3.3 Computing.

With the tendency towards distributed computing and with the increasing computing power of independent units, it can be estimated that a large fraction of the problems can be solved on workstations or even on personal computers. Fast network connections should provide easy transfer of the program units needed from a server station. For extended problems the same network should provide access to the large mainframes.

It is expected that the availability of large computers with vector processing facilities opens up new possibilities for S.A. calculations. Larger problem sizes will become accessible and the present day problems, especially where iterations are involved, will be solved in much less real time.

4.3.4 User support.

The present situation suffers from the fact that on the informatics side no explicit user support is available.

The central support group in DD division should be given the necessary resources and should then be made responsible for installation and updating of a set of "standard" and general-purpose S.A. programs, for documentation and for basic user information. This work must be co-ordinated with that of staff members in user divisions who should be appointed to train new users and to give technical advice.

4.4 Estimates of resources.

It may be supposed that the needs of the regular users for workstations or powerful personal computers will be satisfied by their divisions. Computing on mainframes would be done at the computer centre. There the work should be facilitated by permitting temporary access to large amounts of disk space during calculations (up to 300 Mbytes or even more).

On average each of the estimated 45 regular users will use some 40 hours of 168-equivalent computing time per year (= 1800 h/y in total) and will need a permanent storage allocation on disk of about 15

Mbytes (= 675 Mbytes in total).

As hardware is assumed to be provided by the user divisions and the computer centre, and expenses for purchase and maintenance of programs and documentation are covered by Part 3 on page 219, the specific central budget request can be limited to 30 kSFr/year for training of the users and the members of the support team.

As was requested in the ACCSE report [3], one post (programmer/engineer) should be provided for the support work to be done in the DD central support group.

PART 5

Software Support for Field Computations.

D. Cornuet, D. Dell'Orco, K. Freudenreich, R. Frühwirth, H. Henke, A. Ijspeert, R. Perin, M.J. Price, T. Tortschanoff, J. Tuckmantel, H.-H. Umstätter (chairman), J. Vlogaert

5.1 Introduction.

A fuller version of this part is to be found in [10]. Field calculation software has been in use at CERN for some 25 years. Only recently, however, has there been a move towards commercial rather than user-written packages. The 46 replies to the questionnaire of ACCSE [9] made it clear that there was an urgent need to catch up in two directions: graphics pre- and post-processors and programs for 3D fields. Since then the TOSCA 3D program has been acquired and installed on VM-CMS, and a general purpose FEM program, ANSYS (see Part 4 on page 223), has also been found useful for computation of 3D fields. RF-cavity designers are moving to a 3D package, MAFIA.

5.2 Existing Field Computation Programs in use at CERN.

5.2.1 Static fields in 2 dimensions (2D).

The most widely used packages are POISCR (a CERN version of POISSON) and MAGNET. These employ two different methods: MAGNET the older finite difference method with square meshes, POISSON the finite element method. Unlike MAGNET, POISSON admits cylindrical co-ordinates, plots fields, computes magnetic forces and is also useful for high voltage engineering, electrostatic separators, septum magnets and wire chambers since potentials can be specified on mesh-points. Its main advantage is inherent in the finite element method with triangular meshes of variable size. Its main drawback is the lack of a graphics preprocessor. This makes it necessary to enter the mesh by hand as a list. MAGNET is very precise, easy to use due to good documentation and, despite its limitations, is still in use because data input takes less time than for POISSON. MAGNET has no graphics output at all.

5.2.2 Static fields in 3 dimensions (3D).

The 3D finite difference program PROFI is available at CERN. Although difficult to use, it allowed precise 3D calculations whereas the earlier GFUN was more flexible but less precise. CERN's versions of both packages are old and lack graphical input.

5.2.3 Electromagnetic cavity fields and others.

There is no commercial offering in this domain. We list here the main packages in use at CERN:

1. RF-fields in resonant cavities are computed with SUPERFISH, URMEL, URMEL T, TBCI and the more recent 3D code MAFIA and MASK. The RF problem in modern accelerator design is (coupling) impedances.
2. Wire chamber designers compute electric fields with POISSON, EFIELD or GARFIELD.
3. Linear accelerator designers use electron gun programs like SLAC226 by W. Herrmannsfeldt.

In view of the lack of commercial products, central informatics support is especially important for this class of users, mostly for installation, updating, help with graphics and workstation selection.

5.3 Present and Future.

The two main problems of the old software available at CERN; lack of pre- and post-processing and lack of good 3D programs, are being solved by the introduction of TOSCA and ANSYS.

5.3.1 TOSCA.

TOSCA from the Vector Field (VF) company, Oxford was the most popular choice in the ACCSE user inquiry and is now installed on VM-CMS. It is well known among accelerator magnet builders and much experience has been accumulated worldwide and at CERN. In addition, it has modern graphics pre- and post-processors, can solve both 2D and 3D problems and can make use of the CRAY for the main calculations. In the future it may be interesting to acquire additional software packages (e.g. CARMEN or PE2D) for eddy current prediction in pulsed magnets. These use the same pre-processor as TOSCA.

5.3.2 ANSYS.

ANSYS (see Part 4 on page 223) is a good complementary solution, in particular for mechanical engineers who use it anyway to compute mechanical stresses. An advantage, especially important for superconducting magnets, is that the same pre-processor and finite element mesh can be used for both stress and field computations. In addition, the pre-processor can be run on a PC and the problem then submitted for solution on VM-CMS. ANSYS can treat time varying fields, stationary or transient and is well documented. Time varying fields in magnets have also been computed with MAFIA.

5.4 Budget.

It is estimated that an annual central budget of about 100 kSFr. will be required for software maintenance and for the purchase of new packages.

5.5 Human resources.

The 50 or more staff active in the field require central informatics support at a level which they estimate as involving two persons.

One staff member in the DD central support group should be assigned to look after package installation and maintenance, advise on workstation acquisition and disseminate information on new developments. He should work closely with the staff member assigned to support of structural analysis packages so that some level of mutual backup is achieved. The DD central support group should of course negotiate licence agreements on advantageous terms and organise training in an efficient and cost effective way. Modern sophisticated tools are complex to learn and their diversity should thus be limited.

5.6 Hardware.

The software makes use of the normal computing infrastructure at CERN, including the computer centre and networking. In addition it will be necessary for divisions to invest in colour-graphics terminals, PC's and workstations according to needs. Some feel it to be desirable that newcomers to the field should make their first experience with PC versions of the programs, since large amounts of central computing time can be wasted in the hands of novices. It is clear that demands on the computer centre will increase, both in terms of CPU time and disk space.

PART 6

Analog Electronics CAE/CAD.

G. Baribaud (chairman), A. Fowler, B.I. Hallgren, E. Heijne, P. Jarron, K.D. Lohmann, C. Schieblich, E. Vossenber, R. Zurbuchen.

6.1 Introduction.

In the area of accelerator development and design there is already a long tradition for the use of CAE. In the field of analog electronics, on the other hand, CERN has to make a considerable effort if it wants to keep pace with the introduction of CAE/CAD-based methods already in use in industry. It is felt that a global introduction of these methods, applied to the total engineering cycle from conception and design to acceptance testing and documentation, would necessitate a reorganisation of all related activities, and require investments and the hiring of specialists. Time did not allow us to develop alternative well-documented scenarios for the transition of CERN's analog electronic activities into the age of computer aided engineering but this is clearly a high-priority task for a future working group composed of representative users and support personnel. It is anticipated that a major trend will consist of a drastic shift from conventional analog electronics towards Application Specific Integrated Circuits (ASICs). In view of this situation, and the very limited time spent in assessing future requirements, it is difficult to make forecasts beyond the next 3-4 years.

6.2 Fields of Application.

At CERN, Analog CAE/CAD covers a wide range of application areas [13]. The design of electrical and electronic equipment involves a broad spectrum of disciplines, ranging from high-power equipment for the accelerators to front-end electronics for particle detectors. An implicit consequence of this is that the needs of designers with regards to CAE/CAD are also likely to be diverse. For example, contrast the short lifetime of integrated circuit (IC) development tools, due to fast technological changes, with the necessity to retain stable design tools for long-lived accelerator equipment.

Some important application areas, such as for instance signal processing, have not been covered by the working group because of lack of time and expertise. Four main areas of CAE/CAD, each characterised by specific requirements, have been identified, and are listed in no particular order.

6.2.1 Integrated Circuit Design.

The complete cycle for the design and manufacturing of ICs depends entirely on computer aided methods. The manufacturer normally provides the IC designer with an integrated set of tools, tuned and tailored to the fast-evolving technology, together with a process-specific design database. The designer needs a workstation-based, high quality working environment. During the course of design optimisation, a great number of simulation and layout iterations have to be performed. Hence adequate com-

puting power for simulation is of prime importance and access to mainframe-based simulation is very essential. The simulator used to this end must guarantee to keep pace with the fast evolution of IC technology and provide for up-to-date device models. For layout work, high resolution colour graphic displays and the use of a high quality colour plotting facility is also a requirement. Another area of CAE-based activity related to microelectronics is the characterisation of a large number of semiconductor devices, both on wafers and in packaged form. This requires a laboratory equipped with suitable hardware and software. Furthermore, assembly and test equipment in a clean room environment must be available.

6.2.2 Discrete Component Analog Circuit design.

For analog electrical and electronic circuits made of discrete components assembled and mounted on printed circuit boards, breadboard prototypes are still widely used during the design phase. Analog simulation can favourably replace the breadboard, provided the designer can find his component models in a component library. At present, no adequate component library is available for that purpose at CERN. The usefulness of a library depends strongly on the availability of component models and their quality. The procurement of semiconductor component parameters and other coordination tasks requires a highly qualified and experienced engineer. For efficient discrete component analog circuit design, circuit simulation must be complemented by schematic entry and printed circuit layout CAD tools. Analog circuit layout is often very critical and requires a close collaboration between the designer and the layout specialist. Their proximity and the integration of the design and layout tools is therefore of prime importance.

6.2.3 System Design.

In this context, a system can include elements of very different nature and involve several disciplines; electrical, mechanical, thermal, etc.. Discrete component circuits, sensors, actuators, high power switching devices, superconductors, plasmas, high power conversion and distribution equipment, etc., can be parts of the system. The use of a physical prototype is normally impossible or impractical and analog simulation must be used as the means for design optimisation. The emphasis for the analog simulator must clearly lie on flexible and versatile modelling capabilities, allowing behavioural and functional descriptions, along with the possibility to optimise for specific design criteria. The fact that a large part of system design at CERN concerns equipment for accelerators, calls for a long lifetime and good stability of these design tools.

6.2.4 RF and Microwave design.

RF and Microwave design differs from more conventional circuit design in several ways. The influence of component interconnections cannot be dissociated from the overall behaviour of the circuit, and has to be modelled accurately by using measuring equipment coupled with the simulation tools. Tools for circuit layout and mask generation are also needed, and should be linked to a mask cutter. The use of high frequency component models, and their strong dependance on circuit layout, means that special device libraries are required. Finally, frequency considerations impose special circuit descriptions in terms of wave parameters. This leads to highly specialised and integrated RF and microwave CAE/CAD tools, running on workstations in conjunction with special laboratory equipment.

6.3 User Population.

Table 1 shows the volume of present analog activity in the divisions. These figures are only indicative and will evolve. They should be used with precaution. Two categories of analog circuit designers can be distinguished: full-time designers and part-time designers. They have different needs regarding user interfaces and complexity of their design tools.

Table 1: Analog electronics activities by Division

Division	Number of Designers	Average % time spent in analog electronics design
EP (specialists)	10 - 20	25 - 50
EP (others)	30 - 50	5 - 10
EF	10 - 15	50 - 80
PS	40 - 50	10 - 15
SPS	20 - 25	20 - 30
LEP	20 - 25	20 - 30

6.4 Conclusions and Recommendations.

6.4.1 Private versus collective access to CAE/CAD.

Designers feel that for designs of small to medium size, or for preliminary studies of a concept, they need a private CAE/CAD working environment based on a PC or graphic terminal, with locally multiplexed moderate quality printing and plotting. For more comprehensive design work, a high quality workstation can serve 2 to 5 intensive users. Apart from this, highly specific local working environments, composed of integrated measuring or handling equipment and associated driving software is increasingly needed, particularly for IC and RF design. In all cases, where compatibility allows, access to common central facilities should be provided by networking.

6.4.2 Organisation of activities.

In the context of the present organisation of CERN, support activities are proposed at three levels:

1. Facilities existing at one place in CERN only:
 - printed circuit workshop (CAM facility for prototypes)
 - networking
 - software licensing administration
 - installation, maintenance and technical support for CAE packages of general interest, including central general-purpose plotting option and component library(s) for discrete analog circuit and system simulation, maintenance of existing tools until discontinued. The support of these activities by DD will require the urgent introduction of new software and associated libraries. In addition to the present resources, two further staff are required: one for informatics support and one qualified electronics engineer for modeling and library support.
 - colour plotting facility for high quality output, interfacing to industry standards, (indispensable for IC design, e.g. colour VERSATEC)
 - mask tape generation software with interfacing and archiving facilities (for IC design).
 - communication of expertise by way of a user group to ensure coordination and compatibility, and introduction of a newsletter (already done, in the form of the CSE Newsletter - *Ed.*).
2. departmental/regional level:
 - printed circuit design (CAD)
 - device characterisation and assembly laboratory
3. group/project level:
 - advanced, group- or project-specific techniques and products
 - evaluation of new tools for transfer to central or regional level within the framework of the user group.

6.4.3 Some hints for foreseeable investments.

- central CAE packages of general interest, (e.g. SABER, SPICE for IC design), including library updates and extensions: 150 kSFr./year
- colour plotting facilities for IC design: 300 kSFr.
- ASIC design equipment: 200-400 kSFr. per working place.

6.4.4 Education and human resources.

Last, but by far not least, there is a consensus that training on the subject, both at the technical and the managerial level, has a capital role to play in introducing and applying CAE/CAD successfully. Joint activities with outside lecturers and the CERN training services should be organised.

PART 7

Digital Electronics CAE/CAD.

*P. Bähler, B. Flockhart, L. van Koningsveld, G. Leroy, M. Letheren, A. Marchioro,
L. Pregernig (chairman)*

This part will (7.1) describe the present situation of digital electronic design at CERN, (7.2 on page 238) look at future trends, (7.3 on page 238) recommend an approach for CERN, and (7.4 on page 242) discuss some benefits of the recommended approach.

7.1 Present Situation.

At CERN, engineers and technicians design electronic circuits covering a wide range of applications, complexities, and technologies. Fairly detailed knowledge exists about the different categories of design work at CERN. In 1987 designers ordered several hundred new board designs from CERN's printed circuit board (PCB) workshop and 30% of them were "high density" jobs. The high density jobs may account for about 70% of the total time spent on designing electronic circuits. It is expected that in-house design efforts will continue at more or less their present level. Designers are lacking skills in new design methods and tools to establish and to verify system descriptions, such as functional descriptions, which can be used for subcontracting.

CERN employs about 450 people in staff categories 204 and 304. Maybe one third of them are hardware designers. Often, they work at the group or section level of the CERN hierarchy and few official communication channels exist for them. Some designers have access to modern computer aided engineering (CAE) tools which CERN has purchased in the past. But so far, less than 10 people have received formal training to the level where they are proficient in all the capabilities offered by the tools classified as Class 2 in Table 2 on page 239.

At present, CERN's layout experts are reasonably well equipped with board layout tools and they meet regularly to coordinate their efforts. However most of our system and circuit designers lack (i) modern design tools, (ii) skills in modern design methods, and (iii) the infrastructure to help them to apply such tools and methods. It will require a co-ordinated effort from designers and management to keep up with the evolution of electronics.

Recommendations about CAE/CAD tools for the 1990s also need to consider trends in the electronic industry. The next section describes those that are likely to influence electronic design at CERN.

7.2 Future Trends.

Ten years is a very long time in electronics and therefore it is difficult to predict accurately how design methods and tools will change beyond the next 3-4 years. However, one can identify 3 major trends in industry that are likely to influence electronic design at CERN in the near future.

Firstly, more designers will use Application Specific Integrated Circuits (ASICs), like complex programmable logic chips, gate arrays, or custom chips. Predictions are that 80% of all designs will contain at least 1 ASIC by the year 2000.

Secondly, more designers will apply "Top Down" design methods, i.e. they will use hardware description languages to describe architectures and functions of complete systems before starting any detailed design work. For instance, VHDL (Very-High-Speed-Integrated-Circuits Hardware Description Language) has become a standard (IEEE-STD-1076-1987). Unlike text specifications, such functional specifications are "executable". One can verify them by simulation. Furthermore, major research efforts aim at developing programs which convert functional specifications (descriptions) into detailed circuit diagrams and/or layout (logic synthesis and hardware synthesis). Today, some companies already use synthesis tools successfully. Major CAE vendors are integrating new classes of tools (e.g. hardware description languages or silicon compilers) with their existing products.

Thirdly, components housed in new types of package, e.g. surface mount devices, etc., will become more widely used. This trend, together with increasing device speeds, will make board layout a more difficult task.

7.3 Recommended Approach.

The recommendations in this section suggest how CERN can cover the needs of its designers, keeping up with the evolution of electronics. The recommendations consist of 2 parts, a resource and a management issues section. The working group mainly focussed on design methods and did not consider in detail matters such as computer aided testing, computer aided manufacturing, networking and the present generation of programmable logic tools.

7.3.1 Resources for Digital CAE/CAD at CERN for the 1990s.

At present, no single tool can cover all requirements for CERN. Therefore, the proposal is to provide and support 4 categories of tools. Table 2 on page 239 describes the tools and Table 3 on page 240 the related resource requirements.

One way to estimate the number of tools required is to assume that:

1. the number of designers at CERN will remain constant,
2. 30% of electronics staff are active designers,
3. about 50% of these will require access to "Class 2" tools,
4. one "Class 2" workstation can serve 2-3 designers.

Under these assumptions CERN should gradually introduce approximately 20 additional "Class 2" workstations (13 already exist). Ideally, "Class 1" and "Class 2" tools should be for personal use, i.e. designers need access to them without sign-up procedures.

It is also assumed that improved design methods will help to reduce (or to maintain at the present level) the workload for Class 3 tools. For instance, formal design methods (behavioral descriptions, top-down approach) are a means to avoid redesigns due to poor initial system specification. Furthermore, simulation helps to eliminate design errors and therefore reduces the number of iterations in hardware prototyping. Finally, shifting design capture tasks from Class 3 tools to other classes would effectively increase the resources available for layout tasks.

Table 2: Recommended Tools (Descriptions)

Category	Application Areas	Recommended System Configuration	Recommended Applications Software
Class 1.	Drafting (e.g. schematic capture, rack layouts, etc.) netlist generation and simple integrated PCB layout.	PCs in a network configuration. Operating system DOS (OS2?). System usable for other applications, e.g. text processing, etc..	Integrated software packages like PCAD.
Class 2.	Design specification (functional description), schematic capture, design verification (simulation), netlist generation, test pattern generation.	Workstations that have an open architecture, networking capabilities and an assured growth path (e.g. SUN, Apollo). Operating system UNIX.	Integrated software packages like DAISY.
Class 3.	Layout of complex PCBs and multiwire boards.	Workstations like those for Class 2.	Layout packages that are integrated with Class 2 tools.
Class 4.	Conceptual design and layout of large ASICs (> 20k gates).	Workstations like those for Class 2.	Project-specific packages.

Table 3: Resource Requirements

Tool Category	Capital Investment (per station, if purchased now)	Annual Expenses (per system/person)	Number of Systems (see assumptions)
Class 1.	Appr. 15 kSFr. Includes a PC, software, networking resources, plotter, 25% share of PCB layout software, 4% share of server.	Maintenance: 1,500 SFr. Depreciation: 4,000 SFr. System support: 1 person for 30 stations. Training: 500 SFr./person. Library coordination and support: 1/2 my plus 10 kSFr. for class.	100 (50 exist)
Class 2.	Appr. 125 kSFr. Includes the hardware platform, networking resources, digital-design software packages, 20% share of a server and a plotter.	Maintenance: 12 kSFr. Depreciation: 30 kSFr. System support: 1 person for 10 stations. Training: 1.5 kSFr./person. Library coordination and support: 1 my plus 50 kSFr. for class.	30 (13 exist but need upgrading)
Class 3.	Appr. 140 kSFr. Includes the hardware platform, networking resources, layout software, 20% share of a server (to run autoplace and/or autoroute) and a plotter.	Maintenance: 14 kSFr. Depreciation: 35 kSFr. System support: 1 person for 10 stations. Training: 1.5 kSFr./person. Library coordination and support: 1 my.	< 10 (7 exist)
Class 4.	Appr. 200-400 kSFr. for the hardware platform and project dependent software packages.	Project dependent, in general included in project costs.	Project dependent, but small.

7.3.2 Management Issues.

Investments in modern tools will only be efficient if:

1. all designers receive proper training in modern design methods and tools,
2. an appropriate support infrastructure is provided,
3. design management methods appropriate to the tools are adopted.

7.3.2.1 Training.

After an initial training (e.g. 1 week), designers will go through a learning curve of 3-6 months, before attending advanced training (e.g. 1 week). After that, designers should attend advanced training courses once per year (1 week). The recommendation is that industry provides the training.

7.3.2.2 Infrastructure.

The essential elements of an appropriate support infrastructure are informatics and operational support, co-ordinated library development and application assistance.

Informatics and operational support involves keeping installed systems running, installing new systems, upgrading systems, developing and maintaining interfaces, tracking down bugs, running an electronic bulletin board (for information exchange), etc.. Library incompatibilities cause major problems porting jobs from one system to another. It requires considerable effort and ongoing support to solve such problems. One should therefore avoid interfacing different systems, whenever it is possible. If commercial interface standards become available (e.g. EDIF), the recommendation is to adopt them. CAE vendors (and third-party vendors) offer libraries but they are incomplete. Therefore, a team of experts, e.g. cluster leaders and library coordinators, should assume responsibility for CERN's component libraries, (i) establishing strict rules for component creation and (ii) monitoring their application.

Applications support means to assist designers in using CAE tools and to ensure that they follow guidelines (e.g. for component creation etc.) in applying the tools.

7.3.2.3 Design Management.

Modern design tools are complex and a designer who works in isolation is less likely to use them as efficiently as a designer working in a team.

It is recommended that workstations be organised in logical clusters with experienced CAE engineers designated as leaders of these "design clusters". The leaders, besides working on a specific project, provide the applications support for designers who (temporarily) join a cluster to perform their design tasks.

Such a structure is modular and therefore expandable. For instance, after an initial pilot project, one can gradually increase the number of design clusters according to the design requirements. Furthermore, when new tools become available, one could test them on one cluster, before approving them for general use.

It is suggested to expand the use of CAE tools and methods in phases. The first phase should see (i) the upgrade of existing Class-2 tools, (ii) the organisation of a pilot design cluster, (iii) the co-ordination of the use of Class-3 tools and (iv) the implementation of the basic infrastructure.

If the pilot design cluster is successful, one could expand according to requirements (e.g. one design cluster per year). The capital requirements would be some 500 kSFr. for the first phase and about 750 kSFr. for a subsequent six-workstation cluster (at November 1988 prices). This suggestion assumes that groups will purchase Class 1 tools on their individual budgets.

A final recommendation is to introduce a cost accounting scheme for design and layout work. Such a scheme is a powerful tool for management to control in-house design efforts through pricing, when services are transferred between different units. The opinion of the working group is not unanimous on this last recommendation.

7.4 Benefits.

The recommended approach aims at systematically introducing modern design tools and methods at CERN. Implementing it will generate benefits for CERN, such as:

1. Improving reliability and cost effectiveness of CERN's electronic systems,
2. Access to leading edge technologies (essential in the post-LEP era!),
3. Improving skills, motivation and morale and, therefore, productivity,
4. Providing powerful project management tools for supervisors,
5. Providing efficient means to subcontract jobs to industry or collaborating institutions.

PART 8

Database support for Engineering.

J. Schinzel (chairman)

The number of people involved in database application development for engineering is fairly extensive and many of these were very much occupied with the LEP Injection Test at the time when this part was being prepared. Therefore rather than forming a committee, the draft documents were circulated to the people known to be involved or interested in the engineering application software using databases. The many comments were gratefully received.

During the 1980s the usefulness of Database Management Systems (DBMS) has been generally accepted by engineers. The rapidly expanding user community of the Relational Database Management System (RDBMS) ORACLE both on the VAX/VMS and IBM VM systems and the growing number of data management applications suggests that database technology will be increasingly exploited in the 1990's. One can assume that during the next decade ORACLE will continue to be offered to engineers at CERN, since there is at present no significantly better product on the RDBMS market. However, in the database field, technology is evolving so rapidly that it is dangerous to predict too far ahead. Increased development support will be required for all fields where data must be managed, whether engineering, accelerator control, physics experiments or Management Information Services (MIS). CERN's investment in application development and data entry is growing rapidly. The full benefit of this investment will only be realised if there is CERN-wide coordination and management of all data and software. More emphasis will be placed on reliability and ease-of-use. Compromises may be necessary between stability and the introduction of new tools. Workstations are already being used for database application exploitation. As the number of remote processors accessing databases expands, so will the need for more sophisticated network management tools and support. Around-the-clock access to data will have to be considered, as essential services depend more and more on database technology.

8.1 Present Situation.

ORACLE was installed on a dedicated VAX in 1982 for the management of LEP construction data. Applications include project planning, installation logistics tools and systems for managing inventories, cables, drawings, documents, magnet data, etc.. Many facilities share common data such as lists of CERN personnel, machine elements and locations. The control database presently in development will access much of the data collected for equipment installation and testing. Approximately 400 accounts have been requested on the LEP DB (LEP database) VAX, including 'public' multi-user accounts. About 40 LEP and SPS staff are active in developing applications of which approximately 30 are in production. The production database contains well over half a gigabyte of data.

ORACLE was introduced on the public VMS service in 1984 and on the IBM VM system for general CERN use in Autumn 1985. At present there are about 200 accounts, approximately 40 of which are in constant use. The applications treat both administrative and technical data.

The Database Section in DD-SW Group is responsible for database administration, software distribution, installation and support to application developers in VMS, VM and workstation environments. Support includes advising users on technical ORACLE problems and database design, as well as exploring new software and reporting bugs. DD Division also provides operation and system support for the central database hardware.

Engineering applications are developed by the groups concerned. The LEP experience has shown the advantage of using a dedicated machine and of storing data centrally where it can be shared. Coordination of application development helped to minimise the inevitable redundancy of both application software and data caused by decentralised development. LEP Database service management also proved to be a non-negligible task. Although tools for database design at present are primitive, the advantage of structured analysis and of an efficient database design are becoming recognised.

PC "database" products cannot be ignored. They are cheap, easy to use and in many cases adequate for the users' needs. The migration of these databases to ORACLE, once the volume of data becomes significant or the data is of public interest, must be considered.

The portability of ORACLE, which runs under VMS, VM and other mainframe systems as well as on a large number of workstations and PCs has led to a rapid growth in the number of remote systems accessing a central database. The need for management tools for distributed database systems is becoming apparent.

8.2 Advances in database technology.

The breakthrough of the 1980's is the exploitation of DBMS based on the Relational Model which greatly simplified the task of application development. It is predicted that in the 1990's there will be extensions to the present Relational Model with enhancements in the direction of 'logic databases' and 'object-oriented databases', two different approaches proposed to include the semantics or meaning of the data in a relational database rather than programming them into the application software or storing them in the mind of the developer.

Distributed databases, where data is physically stored in many different places, but is managed as though it were contained in a single database, are still a subject of research, but systems which allow the interconnection of databases across heterogeneous networks are emerging. Distributed databases could ease the load on a central database, but must be carefully planned to avoid other bottlenecks.

Development in the database field is moving away from the DBMS software and moving towards providing user-friendly interfaces and more powerful development tools. Workstation packages for database application development are beginning to appear on the market. These make use of the powerful graphics and windowing tools as well as accessing local or remote DBMS and guide the end-user by displaying icons (pictures) and/or menus which he selects with a mouse. Expert system software to solve complex queries accessing relational databases is emerging.

Database design requires a more rigorous method of analysis than the entity-relationship model offered by the analysis tools on the market today. TEAMWORK, adopted as the CERN standard, is a powerful tool for SASD (Structured Analysis, Structured Design). Software for using NIAM (Nijssen Information Analysis Method) for database design is beginning to appear, but as yet, no package spans both SASD and rigorous database design.

As database technology becomes more widespread, the availability of data becomes increasingly important. It is no longer acceptable to close the database down during several hours for backup or for data reorganisation. DBMS software designers are working on ways of executing these tasks while the database is active.

The comments of Bachy et al. [11] and of Andriopoulos [12] provide further useful reference material on advances in technology.

8.3 Development of ORACLE.

ORACLE is evolving with the technological trends. A procedural language to describe the integrity constraints between 'objects' is partially implemented. Procedures will be stored in the database where they may be accessed by applications. SQL*NET (a component of the planned fully distributed RDBMS, SQL*STAR) allows users to query distributed data and to modify remote data even if no local database exists. ORACLE Corporation along with other RDBMS vendors is working on a solution to the integrity problem of updating distributed data. In order to satisfy the PC market, ORACLE has now been grafted to LOTUS 1-2-3, a popular spread-sheet facility. ORACLE is also moving towards the workstation environment with an increasing number of ports. Development is continuing on their design tool, CASE, to provide tools for structured analysis, database design, software documentation and maintenance. A version for the VAX workstation is expected soon. Natural language support for multi-lingual environments is in development. Within the coming year, on-line backup of databases should be possible.

Since there is no significantly better product on the market and because of the considerable ORACLE experience acquired at CERN, ORACLE will be the RDBMS offered to engineers, at least in the foreseeable future. However, research into new DBMS products and related software is essential.

8.4 Engineering Database Service Hardware.

During the next decade both applications and data will increase significantly, consequently demanding more computer resources, both power and disk storage. Database transactions make heavy use both of computing power and of I/O processing, an important consideration when aiming at providing an efficient database service. End-users, unlike application developers, are intolerant of system failures and bad response times. An increasing number of users will be accessing ORACLE from applications running on workstations, which must be able to communicate with a database machine. The complexities of the hardware connecting a terminal or workstation to the database computer must be transparent. Application software development will also move to the workstation environment.

A public database service is provided on the general purpose VM and VMS systems. The LEP database is installed on dedicated VAXs which are used exclusively for database exploitation. The dedicated hardware solution adopted for the LEP database project gave LEP Division the freedom to provide the service that the project required. Initially the LEP machines ran independently. Today these systems are integrated into the Computer Centre VAX cluster with the public service VAXs as well as the EUCLID CAD/CAM and other service machines. The LEP DB machines are shared by all LEP users both for development and production. The advantage of sharing data and applications is offset by the

disadvantage caused by competition for resources between the users themselves. The need for new software and enhancements must be carefully weighed against the need for stability. As members of a large VAX cluster, there is inevitable interference from other Computer Centre services and from changes and upgrades which are not necessary to the database service and which may introduce interruptions.

There is a strong request from the machine divisions to provide a database service 24 hours a day, seven days a week. Applications concerning radiation protection, LEP beam-monitoring equipment management and the LEP alarm system use the LEPDB for storing data which is assumed to be permanently available. The LEP control system is less sensitive to database availability, since although control data must be available for loading into local storage before the start up of the machine, immediate access to the database during a run is not essential. Down-times of under 30 minutes are believed to be acceptable.

Decoupling the database service from the public services and introducing a 'production' VAX cluster may improve availability. Further measures could be taken such as scheduling service interventions outside critical periods, extending the hardware maintenance coverage and providing redundant hardware. Interference between users as well as unpredicted hardware, software and network failures may, however, still occur. It may be more efficient to build local safeguards against database unavailability rather than to make significant investment in improving a service which, because of its generality, cannot be guaranteed. Reliability and availability could be improved by isolating all or part of the service. However, this may not satisfy the needs of the application.

The present LEP database service should be extended to the CERN engineering community on dedicated VAX computers. Additional resources are necessary to provide the computing capacity necessary for the larger user community. Possible ways of improving availability should be investigated such as decoupling the database service from the busy general purpose VMS service to reduce interference from non-database activities. The need for a 24 hour service should be studied and, if the database service availability is inadequate, alternative solutions proposed.

8.5 Database exploitation support.

Two levels of support can be identified for database application development.

- Central Database Support
- Local Application Development Support

8.5.1 Central Database Support.

With the growth in complexity of database technology and the rapidly increasing number of users, applications and data, the demands on Central Support for database administration, consultancy, ORACLE expertise and exploration of new trends will continue to grow. Consultancy includes recommending appropriate development tools, database design software and optimisation methods. New trends such as storing "pictures", graphics and documents must be explored as well as expert system technology and workstation interfaces to ORACLE. The organisation or recommendation of training courses for database design and application development using ORACLE is an essential part of the service.

The present Central Database Support team of 6 provides expertise and consultancy not only to the engineering community but also to the MIS and physics database users. To continue providing this service they should be reinforced by at least two posts to meet the demands due to the rapid increase in database exploitation foreseen over the next five years.

8.5.2 Local Application Development Support.

Complementary support to that provided by DD Division is already provided by a small team (3-4 people) in LEP Division who work in close collaboration with Central Database Support. The support covers

- **SERVICE MANAGEMENT**
Management of the database service includes the supervision of user accounts, databases and exploitation of the service, prediction of growth rates for data storage and CPU usage, estimation of the hardware additions necessary to meet the expansion, and collaboration with DD Division to ensure optimum working conditions for database users.
- **CONSULTANCY FOR ENGINEERING APPLICATION DEVELOPMENT**
By following the evolution of engineering application software development both at CERN and commercially and by knowing the nature of the data accessed, the local development support team can minimise development effort by advising users on existing software and data. The support team is also responsible for following trends and new techniques in other engineering disciplines, for example, Structured Analysis Structured Design (SASD) methods, and for investigating the possibility of using new methods or techniques in database application development.
- **APPLICATION DEVELOPMENT**
The support team undertakes to develop general engineering applications as well as specific applications when programming manpower cannot be provided by the end-user.
- **USER MEETINGS FOR APPLICATION DEVELOPMENT**
Regular meetings are organised to inform application developers of new ORACLE software and enhancements, and to provide a forum for the discussion of the service management and the presentation of user applications.

The development of engineering applications is open-ended. Since most applications are developed within the User Divisions, it is logical that the development support be situated in the User Divisions where priorities for development, manpower and service expenditure can be directly determined.

Until the future organisational structure of the Machine Divisions is known, the support presently provided under LEP Division management should be extended to include the whole engineering community. To cover the larger number of users, the present team of 3 to 4 should be increased to 5 in the near future and to 10 over the next five years, to cope with the growing number of requests for application development. (At least 40 people within LEP Division users groups are responsible for application development.)

8.6 Budget for the database service and exploitation.

The LEP Database hardware has evolved from one VAX/11-780 to two VAX 8700s in 1988 for database application development and exploitation. An average of 720 KSFr. has been spent per year on hardware, software and maintenance. The yearly expenditure can be roughly broken down into 300 KSFr. for maintaining the present service and 420 KSFr. for configuration evolution. Over the past six years, the capacity of the LEP database machines has been doubled every two years. Although the emphasis has moved from the development of planning and installation applications to control and maintenance, the growth rate of approximately 3 to 4 account creation requests per week is continuing at a steady rate. There were over 400 user accounts registered in mid-1988. Development expenses have been kept to a minimum. However, with the rapidly increasing number of workstations and PCs in the laboratory, a development budget must be included in the future cost of the service.

Growth trends are difficult to predict. Extending the scope of the present LEP database service to serve CERN's engineering community would clearly increase costs. A minimum of 1.2 MSFr. per year is estimated to provide this service. approximately 300 kSFr. for maintenance, 100 kSFr. for small developments, 200 kSFr. for disk and cluster hardware, and 600 kSFr. for configuration evolution. Major configuration changes tend to be biannual. It is considered appropriate that the cost of workstations is covered by the users.

PART 9

Training.

D. Jacobs

9.1 Topics for training.

All of the tools and methods described in this report require a high level of expertise on the part of the users if maximum benefit is to be drawn from them. The required knowledge falls into three broad categories:

1. Theoretical and practical knowledge of the engineering field in question.
2. Knowledge of the design methodology necessary in order to use the tools effectively.
3. Detailed operating knowledge of the tools themselves.

With some reason it can be expected that staff are already equipped with the first kind of knowledge by virtue of their education and experience. Design methodology, on the other hand, requires special attention in the training program. Colleges and universities are now training their students in effective use of CAE tools but at CERN, where new hiring is rare and most of the engineering population graduated before CAE tools were introduced, there is a great need to organise this kind of training. This also applies to Management who supervise the work and must understand the new methods in order to see that they are applied. Training in the use of particular tools is an obvious need, recurrent with each new tool or tool version, although the learning time will tend to decrease as staff become more familiar with CAE culture.

Underlying all this is a need, in many cases, to provide foundation courses in computer use. For many, the application of a CAE tool is their first contact with a computer and basic training can help to avoid the kind of hurt surprise that greets the first crash of a disk for whose files no backup has been made.

9.2 Organisation of training.

Training must be adapted to the needs of the students and to the constraints of the subject being taught and so there is no universal formula.

The training service is well equipped to organise courses but, apart from those on basic computer use and perhaps some on the operation of specific tools, it is probably better that instructors should come from outside, thus ensuring a flow of new ideas into the organisation.

The more general courses may be organised at regular intervals but for the use of specific tools it is necessary to respond rapidly to the needs of designers. Video-taped courses are worth investigating for this purpose.

There is a general feeling that it is most effective to follow concentrated courses given on a number of consecutive days, during which the staff are relieved of their normal work responsibilities. This permits them to invest the necessary time in revision and preparation, essential to reinforce the course work.

In the same vein, although running the courses at CERN may be the most economic financial solution, there are benefits in taking the students out of their normal working environment, which does not necessarily mean that they have to travel long distances. A successful recent example has been the practice of running management courses at Ferney Voltaire or Cartigny.

Most CAE training will require hands-on use of the tools in question. This can pose a particular problem for the training service in the case of workstation-based or individually licenced products. All available copies of the tool will normally be heavily committed to the laboratory's development programme and it will not be easy to release and move them to a central location for the course. In such cases there is no alternative to the expensive option of sending students to attend courses given by the supplier's training organisation. The situation is obviously easier for software which is accessible on the central computers or which runs on standard workstation or PC platforms with which the training service can be equipped.

9.3 The duty to train.

Training is not free and training in the use of sophisticated tools can be quite expensive. Supervisors and Management in general must come to realise, however, that to deny their staff systematic training in the use of CAE tools is a false economy. It is just not good enough to hope that they will pick things up as they go along. A typical course might cost SFr. 2500. A week's wasted design time due to the use of inadequate methods will cost the organisation more than the price of the course.

9.4 Summary of recommendations.

The organisation should make available to its engineering staff training courses of a number of kinds:

- **Basic training in computer use.**
- **Design methodology for each engineering field.**
- **Use of specific CAE tools.**

The courses should be concentrated and it may, in some cases, be advantageous to organise them off-site. Supervisors should be aware of their duty to make sure that their staff follow such courses, not just for their own good but also for that of the organisation.

Bibliography.

1. Co-ordination de l'IAO et proposition budgétaire - Mechanical Board paper - February 1987
2. Interim report - Mechanical Board - September 1987
3. Report of the ACCSE - P. G. Innocenti - September 1987
4. Engineering computing in the SPS/SME group - M. Möller - March 1988
5. Mechanical activities in the research divisions - G. Petrucci - April 1988
6. Computing for engineering needs in the mechanical field. The case of the PS/ML group - A. Pace - May 1988
7. Future CAD computer and work station configuration - C. Hauviller, C. Letertre, R. Messerli, D. Wiskott - May 1988
8. Computer Aided Engineering in High Energy Physics - G. Bachy, C. Hauviller, M. Mottier, R. Messerli - June 1988
9. Questionnaire on Computer Aided Engineering in Mechanics and Electricity, C. Hauviller, H. Umstätter, D. Wiskott, 7.8.87.
10. Software Support for Field Calculations, H.H. Umstätter, June 1987.
11. Comments on the Technical Note "Database Support for Engineering in the '90s by J. Schinzel/LEP 25 May, 1988", G. Bachy, C. Genier and M. Mottier - June 1988
12. Database Support for Engineering in the '90s, X. Andriopoulos - June 1988
13. CAE for analog electronics, LEP/BI-CS/mb(09841)

MIS Computing in the 90's

M. Benot²⁾, R. Cailliau²⁾, J. De Groot²⁾, J. Ferguson²⁾, E. Freytag¹⁾, L. Griffiths²⁾, S. Lauper²⁾, A. Lecomte²⁾, J. Mandica²⁾, R. Martens²⁾, F. Ovett²⁾, B. Sagnell²⁾, J. Schinzel²⁾, G. Shering²⁾, S. Schwarz²⁾, R. Többsicke²⁾, E. van Herwijnen²⁾, and P. von Handel¹⁾

14 September 1988

1) DESY
2) CERN

Chapter 1

Introduction

In the context of the studies relating to Computing in the Nineties it was agreed that the MIS Board, with representatives from each Division, was the obvious mechanism for coordinating the MIS contributions.

It was decided to form three sub-groups for the purposes of specialist discussion and provision of papers:

ADP G.V. Frigo, R. Martens, J-M. Saint-Viteux, M. Baboulaz, D. Duret, J-L. Valaud, E. Dheur, R. Mäder

OCS R. Cailliau, R. Többicke, E. van Herwijnen, F. Ovet

Administration G. Lindecker, JD. Mandica, S. Lauper, J. de Groot, V. Attarian

In addition each Divisional representative was requested to provide an individual contribution for their Division.

An extended MIS Board (with P. von Handel and E. Freytag as participants from DESY) was then asked to consider the specialist input and in particular to reach consensus where possible on the recommendations.

The following papers were produced by the working groups and are referenced in the rest of this text:

1. *ADP Long Term Planning 1988 - 1993*, Aug. 29, G. V. Frigo et al.
2. *Computing in the 90's: MIS Board Papers*, Aug. 31, MIS board members
3. *MIS: Central Applications the next 5 years*, July 27, R. S. Martens et al.
4. *Office Computing in the 90's*, Aug.31, R. Cailliau et al.

The following papers are indispensable input to the above:

5. *Management Information Systems at Cern*, Dec. 86, J. Ferguson
6. *Future Office System Requirements*, June 88, E. van Herwijnen

This paper aims at extracting as briefly as possible the essential elements from the above papers. Detailed explanations and justifications are given in the original papers.

The general CERN MIS context is not very different from that of any other enterprise: corporate data are generated at the base, in a fashion which should be reflected in distributed electronic data entry at the location where the data are generated. These data should then be collected in large corporate data bases for correlation. Subsequently, managers and office workers should have access to these data from their office workstation to obtain decision information, to present them and to use them in models and reports at a local level. The current situation at CERN, its problems and the implications of recommended solutions are discussed below.

Chapter 2

Assessment

2.1 Current Situation and problems

There are two main areas to be considered in CERN MIS activities: applications concerning corporate data and office systems. The data must be available to many at the same time, and must be a correct representation of reality. Office systems are used by MIS users in their daily work, combining corporate data and other information for management decisions. Following is an inventory of the current situation and its problems.

2.1.1 Corporate applications

The CERNADP computer at the Administrative Data Processing center is a small IBM mainframe [1]. Over 100 IBM terminals and a number of ASCII terminals are connected to it. The computer is saturated despite the fact that most of the batch production is run outside prime shift. Response times are often unacceptable.

For reasons explained in [1] we are running an heterogeneous mixture of software both at the system and at the user level: two operating systems (DOS/VSE and VM/SP), three data base systems (IDMS/R, DL/I and GIP file system) and three transaction subsystems (IDMS/DC, CICS and GIP-TP). Users are forced to switch context when passing from an application to another. The Financial data base system (IDMS/R) and the Purchasing/Receiving system (DL/I) have logical ties between them but run in two different contexts.

2.1.2 Office systems

A limited number of office type applications have been used for a long while on the central mainframe computers. For many visiting scientists, the central machines are the only tool they have to help in office type tasks. Central text processing, mail and document storage is important.

The oldest application of computers in the office is for text processing, and while substantial progress has been made since the introduction of stand-alone word processors, the greatest needs remain located in this area: coherent laboratory-wide use, document interchange, the integration of graphics, document storage and retrieval, mathematical Text processing is a tool used daily in almost all MIS areas as well in physics.

Many vendors offer more than text processing functions on their equipment and in the absence of standards, the situation is chaotic: the survey [5] found systems from no fewer than 11 different vendors on site, all with completely incompatible systems. This situation has eased somewhat, as many of these 11 have either stagnated or been substantially reduced, but so far not a single one has been removed entirely, because of capital investment amortising considerations or because of continued important usage. The financial investment [5] is of the order of 12 MCHF. There is no expansion, but there is a recurring annual maintenance expenditure of the order of 2 MCHF. These figures do not include use of central computer facilities, communications equipment, PCs, consumables, and other mini-computers primarily used for other purposes. On the personal workstation front, at present there is an installed base worth in the order of 10 MCHF, which is expanding at the rate of 5 MCHF per year (over all uses). Maintenance costs are still relatively low (because the machines are new) or unknown.

Fortunately, at present, no new applications are undertaken for the obsolete systems, and over time they will be gradually phased out. Four contexts will however clearly remain important at CERN in the next years and all effort is concentrated in these "Four Operating Environments" (or FOEs): IBM with VM/CMS, DEC VAX with VMS, Olivetti IBM PC compatible with MS-DOS, Macintosh with MacOS. These are linked by "the fifth force": the Ethernet local area network.

It should be pointed out here that there is a growing number of high-end workstations with physics applications on which some MIS type work is done: for example, there are about 100 Apollos running a Unix-like operating system, which are partly used for document preparation. (note: during the next five years, we may see an increasing number of presentations on top of Unix as the underlying operating system kernel. This will be true for all powerful workstations but is also likely for the PC compatibles and Macintosh. The OS/2 for PC compatibles will probably make an impact in the MS-DOS world.)

As to personal computers, they came into CERN from mainly two directions: the experiments (Macintoshes) and accelerator control & engineering support (IBM PC compatibles). We observe that the middle ground, where users are not constrained by the needs of either of those environments, has been taken largely by Macintoshes, which are the most popular micros at CERN.

Finally, there is the corporate publishing unit. CERN has at all times had the need to distribute information about itself, and this unit has provided two functions: (a) the preparation of documents and the stipulation of their formats and (b) the actual production (bulletin, Annual Report, "Yellow" reports, ...). Around 20 document types are currently produced by Publications service, and 70 million A4 pages printed per year. At present there is no integration between the system used in the Publications unit and the others.

2.2 Requirements

This section describes present user requirements and industry trends.

2.2.1 Corporate applications

One of the most urgent problems is that of reducing the maintenance activities which now make up the overwhelming part of the day-to-day program of work, to allow progress towards unification and improvements. It is extremely desirable (mandatory?) that all major corporate data bases be unified in the next two to five years by replacing the current incoherent applications by a coherent and extended set (including for example the stores catalog, stores management and the inventory). Additionally, divisional systems (buildings in ST and the User Office in EP) should be developed in coordination with MIS. The consequence of the heterogeneous state of affairs described in the previous chapter is a tremendous overhead in programming interfaces between systems that were not made for coexisting. The ADP application programmers spend most of their development time in programming new interfaces when the need for new applications arises.

A most important part of the requirements for the corporate data base engine is the set of tools available for software analysis, development, implementation and maintenance. A truly integrated data dictionary is essential [3]. Broad endorsement both by management and by the user base is essential in ensuring that the data definitions are both appropriate and enforced. Applications that cross-reference databases will increase (eg. some application will use the stores catalogue data together with the suppliers data in the purchasing data base and accounting data from the financial data base).

The applications should be easily available within the MIS user community, while the specialist functions should be available to Finance and Personnel divisions and well-protected for security reasons. Security and discretion is imperative, but where possible should not inconvenience users. Present security levels should not decrease.

A number of services external to CERN should be accessible: external commercial data bases, the data bases available through Minitel, a computer aided translation service, data bases useful to the CERN legal service.

A friendly system is needed at all levels. The user interface should tend to the graphical, iconic interface [6]. The applications should interact intelligibly to help the user by providing intelligent default values and hiding irrelevant information. User input should be verified as early as possible to avoid cumbersome correction of typing errors [1], and automatic validation of updates in general should be available. Expert systems and artificial intelligence should be supported as a user-friendly means to access data bases.

The various applications should be available from within the same context (no switching into subsystems with overall different philosophy) and should themselves provide a uniform environment. The user should be able to seek help without disrupting the context. These requirements are only fully met by hiding all environments except that of the user's workstation, which is now becoming achievable within emerging local area network environments. Thus the user should no longer have to be aware of the operating environment on which the corporate data bases reside, nor of how the data actually get to the screen of his workstation.

The system must be reliably available during office hours, but preferably also outside these hours. The loss of data is not acceptable. Its response should be fast. For "trivial" tasks, the response should be particularly quick. A trivial task is a task that the user considers trivial. The system should allow back-ups to be made without affecting the functionality. The hardware architecture of the data base engine should be extendible to accommodate growing loads. It should also have some (limited) non-stop operation features eg. disc mirroring.

2.2.2 Office systems

A coherent office system should be available throughout the Laboratory. Two major facts will make this very difficult to achieve: (a) there is a community of physicists with needs different from those of the administrative MIS user, (b) the four operating environments will be here for at least another 5 years. For document preparation, the visiting physicists want compatibility with whatever hardware and software they are running in their institute while also wanting to communicate with resident CERN staff. The large installed base of the FOEs excludes the introduction of a new, panacea solution except at very high cost.

Among our users there is a great variety of educational level with respect to computing. This in itself creates great stress on the unit that has to provide service for both the sophisticated systems programmer (who knows little about administration) and the sophisticated administrator (who knows little about computing).

In the FOEs, the central VM/CMS system is unsuited for interactive work of the type now common on workstations, it is incompatible in its ways with all other systems. Basing the office system on mainframe computers makes insufficient use of the power of the workstations installed, and indeed makes them into very expensive terminals. However, workstations and even workgroup servers do not provide today adequate data storage technology for the volume, speed and reliability needed for a large shared document store which must be at the basis of the office system.

In any case, there is not enough manpower available in the present phase of CERN history to provide a good office environment on all of the FOEs. Thus it is indispensable to find a correct rôle for each of the FOEs without wasting resources on duplicating facilities, especially as, depending on the environment, this cannot be done effectively and efficiently.

Document production will increase dramatically in sophistication: users today want excellent text processing tools, but only a few systems currently available are up to the task, whereby each usually is strong in a few areas and weak in all the others. The search for an acceptable set of tools for scientific document preparation is one of the major tasks of OCS [6]. Full integration of text and graphics and good support for the production of tables is mandatory. Use of graphics will increase dramatically: scanned images, drawings, graphically presented results from computations, even animated sequences will not only be produced in large numbers, but they will be shipped across the networks and availability from a central repository will be requested. Document production will increasingly evolve towards the Hypertext paradigm. Document history (changes, versions) needs to be kept, storage and retrieval is important. Long-term archiving, e.g. on optical discs, must be provided, including friendly retrieving over long periods.

It is highly desirable that coherent links between the office systems and the CERN Publications unit are brought into service and that the unit uses techniques compatible with those of the office systems.

The increases in interactivity of computing implied by the above list demand high bandwidths available everywhere, from the central data repository (which will have to handle many transaction-type requests simultaneously), through networks to the desktop workstation which will have to be a powerful graphics machine. It is in this bandwidth boosting activity that we see a potential major rôle of departmental systems.

2.2.3 Required Tools

All the MIS services should be accessible from the user's single workstation on his desk. This requires a high-quality, fast multi-tasking machine. Such machines are now available at affordable cost (eg. Macintosh II or Olivetti M380), the objective should be to select a minimal number of types and provide a maximum number of different services on the workstation. The availability of the workstation has brought us the problem of how to use it efficiently:

- how to use it better than as an extremely expensive terminal,
- how to organise training, help, maintenance,
- how to avoid that the user spends more of his time on experimenting with the workstation and its software than on actually using it for his MIS job at hand.

The following patterns are important to MIS work in the coming years:

- a single corporate data base system,
- a single data dictionary to describe the corporate data,
- coherence between the different corporate applications,
- Workgroup servers to increase the bandwidth and the services offered,
- SQL as the data base query method, irrespective of the underlying data bases,
- SGML as the underlying document format,
- PostScript as the display format on printers and screens,
- a set of recommended & supported hardware and software packages to increase the usefulness of the workstations,
- Electronic forms processing and electronic signature for transactions in all procedures,
- EDI (electronic data interchange) for communication with outside firms,
- Hypertext to structure the top layer of access to information,

- Unix as the only conceivable standard operating system,
- Ethernet as the network carrier,
- extensive use of electronic mail.

2.2.4 Objectives

MIS wants to

- provide a coherent set of services accessible from any office in CERN,
- reduce the number of different and incompatible office systems and administrative data bases and procedures,
- reduce paper consumption and paper shuffling,
- introduce modern techniques and keep up to date.

Thus the following services are necessary:

- automated administrative activities.
- comprehensive electronic forms handling system with well-accepted electronic signature, including a data-entry interface to corporate data bases.
- long-term document archiving.
- corporate applications for:
 - personnel management (staff & users),
 - financial applications (accounting & budgeting),
 - asset management (inventory & buildings),
 - stores & internal sales (catalog & stock management),
 - purchasing & suppliers management.,
- easy access to the corporate data bases,
- good support for text processing activities (administrative and scientific),
- good shop-like service for purchasing of hardware and software by individuals and divisions, in the style of the present PC Shop,
- good end-user training and support, backup services and other day-to- day help,
- a reliable MIS computing platform.

Chapter 3

Recommendations

3.1 General

MIS should integrate the access to corporate data, transaction processing and electronic document handling with the use of personal workstations, which have text processing, spreadsheet and graphics applications locally. As much as possible, the local processing power should be used, processing should move towards more central systems only where necessary, for reasons of reliability and availability. A general guideline should be the use of standards wherever possible and acceptable. We propose therefore the following recommendations, which are discussed in more detail below (each explanation also has details of the resource estimates given in the table at the end). They are listed here in three categories, but all are equally vital.

-
- A. Introduce coherence and modernisation
1. audit, modernise and subsequently automate the organisation's procedures,
 2. establish the CERN corporate data model,
 3. use a unique DBMS if possible,
 4. replace all corporate applications with a coherent and extended set,
 5. integrate the existing platforms (FOEs),
 6. introduce a comprehensive electronic forms handling system,
- B. Evaluate future choices / developments:
1. evaluate the use of a front-end system for ADP,
 2. study EDI and introduce it for contacts with suppliers,
 3. study the potential rôle of workgroup servers & departmental computers,
 4. recommend an environment for desk-top office use,
- C. Operational activities:
1. upgrade the ADP computer,
 2. move the ADP computer to the computer centre,

3. run a computer shop activity,
4. increase and organise user support functions,
5. contract out the repair and maintenance service for workstations,
6. phase out all non-recommended systems over the next 3 years,
7. replace the stores management system.

A.1 Audit, modernise and subsequently automate the organisation's procedures

The document [5] which is at the basis of the CERN MIS unit specifically states some goals and pre-conditions for their successful achievement, among which is cited:

audit of the laboratory's administrative procedural practices should be carried out with a view to increasing efficiency and the introduction of automation. This will require external expertise in order to perform a comprehensive study.

A.2. Establish the CERN corporate data model

The purpose is to ensure a coherent approach to the creation of new data bases wherever needed. This is considered to be a prerequisite for the audit of A.1., and a refreshed model will be the result of the audit.

A.3. Use a unique DBMS if possible

This could be Oracle, which is supported by a vendor-independent company on all of the FOEs, but it could turn out to be different, as Oracle is not the only DBMS with this characteristic. Oracle is however already firmly implanted at CERN in other domains.

A.4. Replace all corporate applications with a coherent and extended set

This will have two advantages: it will diminish the effort now needed to run a heterogeneous mixture of systems and it will make it possible to move towards a truly homogeneous interface for the users. It is the equivalent of the reduction of the number of systems in office computing.

A.5. Integrate the existing platforms (FOEs)

Servers will be introduced to act as link between the Macintoshes and the PC compatibles. The central system is to be used as a document store, a backup store and a gateway node for some of the electronic mail exchange. It is not to be used as an interactive office system. The networks are Ethernet and Appletalk, as is already the case.

A.6. Introduce a comprehensive electronic forms handling system

A rudiment now exists and is used extensively in PS division. We should look for a commercial product or replace the existing system by a CERN-wide one. Its exact form again depends on the outcome of the audit. It may not be necessary to use electronic mail (or equivalent) if distributed applications can gain automatic access to a central data base engine.

B.1. Evaluate the use of a front-end system for ADP

Data would be put there for public use, which will give easy access to a wide range of users via a common DBMS, while leaving the back-end repository of corporate data as the secure system accessed only by the specialist departments. A study should be carried out in '89, at the end of which a decision should be taken on extending the approach.

B.2. Study EDI and introduce for contacts with suppliers

Commercial products are becoming available to help with electronic interchange of business documents. The body for standardisation of the forms has however not yet published more than a few of them, so that we have a breathing space before we have to handle EDI. Evolutions have to be followed closely.

B.3. Study the potential rôle of workgroup servers & departmental computers

Among the more obvious uses of the servers are: printing, backup, mail, file store cache. The workgroup servers can in principle be any make, but for reasons of familiarity, uniformity and compatibility machines from the VAX range are strongly preferred. DEC clearly are ahead in integration of services in multi-vendor environments.

B.4. Recommend an environment for desk-top office use

The desktop workstations are Apple Macintosh II and Olivetti M380 machines, with, where needed, A3-size or colour screens, and local or group laser printers.

However, Macintoshes will certainly remain the most popular machines for the next 2 years: the Mac OS achieves a level of integration of the applications sold by different software houses that is not matched by any other system today. They are preferred by almost anyone not constrained by technical

issues. Both Macintoshes and IBM compatible machines will be purchased by divisions with specific technical needs. Both will be supported, leaving us free to benefit from a substantial technological leap on either side.

C.1. Upgrade the ADP computer

An upgrade of the present computer centre is undeferrable. the machine needs to be replaced by a 5 to 10 Mips unit in order to be able to continue to provide the present services.

C.2. Move the ADP computer to the computer centre

The move to the computer centre has already been accepted, it will enable the ADP group to concentrate on its key functions of analysis and applications development. It should be done quickly, for economy reasons since it will reduce the overheads now incurred in operation, backup, power supplies and air conditioning, as well as making it easier to link to other machines. Extensions will be easier.

C.3. Run a computer shop activity

There already is a shop providing the service of channeling the purchases of desktop machines of both PC compatible type and Apple Macintoshes. The shop also provides counseling, runs a few public services, and deals with the technical contacts with our suppliers. It is an important part of the OCS activities and must be extended.

C.4. Increase and organise user support functions

User training must be organised. A day-to-day help consisting of "hand- holding", question answering and help is necessary. As an example, if every PC/Mac user has one question per month, and if s/he knows there is a help service, then with the present level of 2000 stations, that help service would get 109 questions per day. This rate is probably exceeded today: we not do see a catastrophe only because users help each other and because the reliance on workstations is still very low. Increased user support is also needed for the corporate ADP applications which will be more widely used than today. ADP users need to be supported with analysis of their needs for applications.

C.5. Contract out the repair and maintenance service for workstations

The repair and maintenance activities should be contracted out. With an installed base of close to 6000 units in 1993, each with a useful life of 5 years, showing 1 important failure per year during this lifetime, that implies 6000 failures per year or approximately 27 per working day. There will also be upgrades and modifications, probably of the same order of magnitude.

C.6. Phase out all non-recommended systems over the next 3 years

The existing secretarial systems will be phased out and replaced by workstations. Some of the data on those systems must be transported to the new stations, taking a non-negligible amount of time. There is however no guarantee that the functions of existing applications will be preserved in all cases.

C.7. Replace the stores management system

The now obsolete computers used in the CERN stores should be replaced by a new centrally sited and operated system.

Chapter 4

Resource estimates

The tables below summarise our estimates for each of the points cited in the recommendations. Estimates are very approximate.

A short explanation for each item can be found in 4.3, after these tables.

4.1 Money

Table 1: A. Introduce coherence and modernisation

Ref	Description	1989	1990	1991	1992	1993	Forever
A.1	procedures(ext.consult.)	250	250				
A.2	data model						
A.3	unique DBMS	100					
A.4	replace corp. appl.	100	100	100	100		
A.5	integr. platforms						
A.7	elec. forms system						
	Total KCHF	450	350	100	100	0	0

Table 2: B. Evaluation of future choices / developments

Ref	Description	1989	1990	1991	1992	1993	Forever
B.1	ADP Front End	100	1050?				
B.2	Study EDI						
B.3	dept. computers	500?*	500?*	500?*	500?*	500?*	500?*
B.4	desk top env.	500 *	500 *	500 *	100 *	100 *	100 *
	Total KCHF	1100	2050	1000	600	600	600

Table 3: C. Operational activities

Ref	Description	1989	1990	1991	1992	1993	Forever
C.1	ADP upgrade	2000					
C.2	ADP move	300					
C.4	user support	50	5	5	5	5	5
C.5	repair workstations	900*	900*	900*	900*	900*	900*
C.6	phase out non-rec	50					
C.7	replace stores system	1000*					
	Total in KCHF	3400	5	5	5	5	5

? = conditional on outcome of study

Table 4: Special projects

	1989	1990	1991	1992	1993	Forever
Sub total projects in 90's:	4950	2405	1105	705	605	605
Reserve						
Free to allocate to new projects						

Table 5: Base-level services

	1989	1990	1991	1992	1993	Forever
ADP running cost	1000	1000	1000	1000	1000	1000
OCS running cost	740	740	740	740	740	740
C.3 computer shop	50	50	50	50	50	50
Total running costs	1790	1790	1790	1790	1790	1790

Table 6: Global Total

1989	1990	1991	1992	1993	Forever
6740	4195	2895	2495	2395	2395

* = financed by the user division(s), not by MIS.

Estimates are very approximate.

4.2 Manpower

Table 7: A. Introduce coherence and modernisation

Ref.	Description	1989	1990	1991	1992	1993	Forever
A.1	procedures (ext. consult.)	2	2	1	1	1	
A.2	data model	2	2	1	1	1	1
A.3	unique DBMS						
A.4	replace corporate appl.	5	5	5	5		
A.5	integrate platforms	2	2				
A.6	electronic forms system	2	2	2	0.5	0.5	0.5
	Total in person years	11	11	7	7	2	1

Table 8: B. Evaluation of future choices / developments

Ref.	Description	1989	1990	1991	1992	1993	Forever
B.1	eval.front end	1	2?	2?	2?	2?	2?
B.2	study EDI	0.5	0.5	0.5			
B.3	dept. computers	2?	2 ?	2?	2?	2?	2?
B.4	office environment	2	2	2	2	2	2
	Total in person years (MIS)	3.5	4.5	4.5	4	4	4

Table 9: C. Operational activities

Ref.	Description	1989	1990	1991	1992	1993	Forever
C.1	ADP upgrade						
C.2	ADP move						
C.4	user support	5	5	5	5	5	5
C.5	contract repair WS	0.5	0.5	0.5	0.5	0.5	0.5
C.6	phase out non-rec sys	2	2	2			
C.7	replace stores sys	2	2				
	Total in person years	9.5	9.5	7.5	5.5	5.5	5.5

Table 10: Special projects

	1989	1990	1991	1992	1993	Forever
Sub total projects in 90's:	24	25	19	16.5	11.5	10.5
Reserve	25	25	25	25	25	25
Free to allocate to new projects	1	0	6	8.5	13.5	14.5

Table 11: Base level services

	1989	1990	1991	1992	1993	Forever
ADP running cost	15	15	15	15	15	15
OCS running cost	14	14	14	14	14	14
C.3 computer shop	5	5	5	5	5	5
Total in person years	34	34	34	34	34	34

Table 12: Global total (MIS only):

1989	1990	1991	1992	1993	Forever
59	59	59	59	59	59

4.3 Short explanations of resource levels

A.1 Audit, modernise and subsequently automate the organisation's procedures

The cost of external consultancies is estimated at 250 KCHF for two years. During the analysis phase of two years, two MIS analysts must assist, and during implementation of three years one analyst.

A.2. Establish the CERN corporate data model

First analysis and implementation needs 2 analysts for two years, maintenance goes on for ever with one analyst.

A.3. Use a unique DBMS if possible

If Oracle is not chosen, a license for software will cost around 100KCHF. A site license for Oracle exists.

A.4. Replace all corporate applications with a coherent and extended set

Costs for replacements of the four major corporate applications is estimated at 100KCHF per application, adapting them to the CERN environment depends on the willingness of CERN to change procedures, but is estimated at 5 manyears per application migration.

A.5. Integrate the existing platforms (FOEs)

Two systems programmers for two years are needed to provide a set of systems solutions.

A.6. Introduce a comprehensive electronic forms handling system

Two systems programmers are needed for three years to implement or adapt a solution that is sufficiently comprehensive and elegant. Maintenance is at 0.5 manyears per year.

B.1. Evaluate the use of a front-end system for ADP

In the case of choosing a VAX, a 10Mips/10Gb system is evaluated at 1MCHF in 1990. Evaluation is one year by an analyst. Implementation and maintenance, independent of the choice, is estimated at 2 analysts forever.

B.2. Study EDI and introduce for contacts with suppliers

This probably needs half an analyst in the first three years, because of integration problems.

B.3. Study the potential rôle of workgroup servers & departmental computers

If the study results in a recommendation of departmental computers, they could be introduced at a sufficient rate as corridor machines, about 10 per year, and needing CERN-wide two support staff (technicians). The machines would have to be purchased by the user groups at approx. 50KCHF per unit (based on today's approx. price for a μ VAX).

B.4. Recommend an environment for desk-top office use

They need CERN-wide two support staff (technicians). The machines have to be purchased by the user groups at approx. 7.5KCHF per unit. For Apollo's and Unix, we estimate of the order of two software engineering posts to support these users with their MIS needs, but currently there is no manpower available.

C.1. Upgrade the ADP computer

This urgent upgrade to a more powerful IBM machine is estimated at 2MCHF, including 15Gb discs and cartridge-based backup and archiving unit (replacement of old equipment).

C.2. Move the ADP computer to the computer centre

The move will cost 300KCHF.

C.3. Run a computer shop activity

The activity costs three staff (technicians and data aid) not counting user support (see elsewhere), 50KCHF is estimated for risks and overheads.

C.4. Increase and organise user support functions

Five manyears per year are required, from both ADP or OCS staff, and the consultancy office will need to be equipped to handle the questions (50KCHF initially, 5KCHF later on).

C.5. Contract out the repair and maintenance service for workstations

This supplies work for 6 to 12 technicians, or approx. 900 KCHF per year in wages. However, it represents an extra cost of only 200 CHF per unit over its entire lifetime (not counting the price of the replaced or installed parts). One contact person is needed half-time. The cost of repairs is for the users.

C.6. Phase out all non-recommended systems over the next 3 years

Manpower is 2 programmers for three years, there is probably some small cost in special equipment and software during the changeover, in the order of 50KCHF.

C.7. Replace the stores management system

The system costs (estimated at 1MCHF) will be carried by ST division, MIS needs two analysts for two years to interface the new system to the other corporate systems.

**Computing at CERN in the 1990's
Theory Computing in the 1990's
Final Report**

T.E.O. Ericson
G. Martinelli
H. Satz
S. Schwarz

November 1988

Contents

Chapter 1: Structure of TH computing	271
Chapter 2: Electronic mail	272
Chapter 3: Databases and networking	272
Chapter 4: Personal computers	272
Chapter 5: Reference model	273
Chapter 6: Software	273
Chapter 7: Training	274
Chapter 8: Documentation	274
Chapter 9: Library	274
Chapter 10: Economic Summary of local equipment	275

Monday

- 12.
- 13.
- 14.
- 15.
- 16.
- 17.
- 18.
- 19.
- 20.
- 21.
- 22.
- 23.
- 24.
- 25.
- 26.
- 27.
- 28.
- 29.
- 30.

Chapter 1

Structure of TH computing

The scientific part of the division consists, including visitors, of approximately 120 individuals at any given time; they are located in 60 offices. The many visitors have a very diverse background and stay in close contact with home laboratories. As a consequence they are involved with all major computer systems in Europe and USA; in particular short term visitors cannot be expected to invest time and effort into learning unfamiliar new systems which will not be useful at home, however excellent such systems may be.

In the past the research activities of the Theory Division only depended in a minor way on computing. The exceptions mainly concerned a few individuals associated with phenomenological work, and, otherwise, small programs. This pattern is now changing in different ways.

1. The division has always been engaged in a number of collaborations between individuals with the partners located elsewhere in Europe, United States and Japan. Nearly 400 scientific reports are produced yearly. These collaborations have always suffered from the slowness and difficulty of information transfer by mail and phone both in discussions and in the elaboration of reports. We now meet a strong demand for
 - a. electronic mail not only for usual correspondence but for transfer of files and manuscripts to distant laboratories,
 - b. scientific text processing (VAX/IBM and PC's) which now has reached a readily used standard accepted by a large international community [e.g., the TEX system].

As a consequence of this generalized use which affects all members of the division independently of detailed research profiles we attach a high priority to the equipment of all offices with terminals, PC's and workstations (e.g., Mac II's) as well as ready access to printing facilities. At present, the TH equipment in this respect is far below normal standard in member states and the US.

2. Conventional calculations for theorists concern mainly symbolic manipulations, in particular evaluations of Feynman graphs on the one hand, phenomenological Monte Carlo simulations (as for the Lund model) and evaluation of equations or formulas with subsequent plotting. These activities require adequate PC's and low end workstations, easy access to a VAX cluster, and to some extent to main frames. They concern only a moderate size group of users; the use of the VAX is increasing, since it is very user friendly.
3. Recently developed techniques for solving basic non-perturbative problems of QCD by use of Monte Carlo simulation on the lattice require powerful computing facilities of the CRAY type and/or parallel processors. They require large amounts of computing time. Such computing requirements arise in particular also in the studies of the phase transition to the quark-gluon plasma, closely linked to the experimental heavy ion program. Moreover, CERN theorists are involved in large international computer simulation collaborations; this is possible only if a non-negligible part of the work can be carried using CERN facilities. There are at present eight theorists in the Division engaged in these activities. Based on present needs, we estimate that about one quarter of the CRAY computing power is a reasonable goal for the early 1990. In addition, any idle time could be put to very effective use here.
4. A special part of the TH division is the library, which presently is being computerized. This is a general public service, the needs of which have been stated in the MIS report. In particu-

lar, the preprints will become available electronically after transfer to and distribution from an optical disc. They will be read by the users on high resolution screens. This new mode of operation will eventually have an important impact also on the scientific part of the division. Theoreticians as other physicists at CERN will have access to preprints via such special screens. In the short term these will be in public areas, in the long term when costs are less prohibitive at least some individual offices will have high resolution screens. In many cases the preprints will also be printed into hard copies.

The implications of these general considerations will now be quantified.

Chapter 2

Electronic mail

The use of electronic mail presently concerns less than 50% of the division and visitors. We expect this to grow to 100% in the next few years. The electronic mail in TH is predominantly international. Local electronic mail is mainly of collective nature (seminar notices, etc.); in addition a strongly extended information on preprints, etc., from the library is developing. Individual mail messages from the TH go predominantly to other laboratories. At present only a few modems at home exist. We believe the use of modems will increase substantially in the future, in particular since electronic correspondence is conveniently done after normal working hours from home.

Chapter 3

Databases and networking

Systematic use of databases and networking only concerns the division to a minor degree. The main exception is the library public service, with the distribution of information from the SLAC, ISIS etc. database on preprints, books and so on, both inside CERN and to other laboratories.

Chapter 4

Personal computers

Our plan is to equip all of the 60 scientific offices with terminals in addition to the public terminal rooms. A special problem arises from the rapid turnover of visitors; only the simplest operational terminal system is easily used by short term visitors. As a consequence, personal computers will only be installed for TH staff and long term visitors. Furthermore, we must provide visitors with access to personal computers of varying types including printing facilities, so that they can use their home programs readily. Low end workstations will mainly be installed in public areas. We estimate needs for personal computers to be mainly Mac's, with a few Olivetti's and some Mac2's. The demand for workstations may be underestimated.

Estimated cost: a mix of 20 units at an average price of SFr 6.000: SFr 120.000 (including programs; in particular TEX). Printers: 2 laser printers: SFr 12.000 + operating cost SFr1.000/year.

For the choice of other terminals we would ideally like to have cheap graphics screens, in particular in view of the WYSIWYG demand for scientific text processing with TEX. For the same reasons there will be pressure for a few units with big screens for page editing. Present compromises are Falcon's and Atari's. Estimated cost: 35 offices SFr 2.500 = SFr 87.500.

For its internal needs, the library will install about 10 Mac's: cost $10 \times \text{SFr } 3.500 = \text{SFr } 35.000$.

The TH divisional secretariat has presently an AES system as a networked multipurpose workstation. This system is poorly compatible with all the other divisional activities and should be replaced by a modern, suitable system in the coming years. This requires in particular special screens, but can probably otherwise be based on a medium size workstation suitable for text processing.

Estimated replacement cost: SFr 50.000 + 1 printer SFr 14.000.

Chapter 5

Reference model

In the TH division we do not expect to put workstations on everyone's desk even if the money is available. This is not suitable for short-term visitors who ask for the simplest mode of operation first.

Low end workstations in the TH division will be excellent for public areas, for the divisional secretariat and for some of the staff members and long-term visitors. We expect "dumb terminals" to be a non-negligible part of the standard equipment even in the future, complemented with good personal computers and cheap graphic screens.

Chapter 6

Software

On the software side, TEX is essential to TH both on VAX and on IBM; we will certainly use it also on Mac's and PC's when available. This is presently an unofficial international standard, that cannot be ignored for practical work in particular in collaborations. Any other standards must coexist with TEX for quite a while.

For drawing, plotting, etc., the general system like PAW is an overkill and too sophisticated for most TH work. We need quick, simple, standard systems (Mongo, Topdrawer). People who want to use such systems once in a while desire to invest only the briefest time possible before using them. For the average visitor, plotting is less trivial than DD thinks.

Of the presently available symbolic manipulation programs REDUCE is presently partially OK, but too slow (SCHOONSCHIP on ATARI's is faster, specialized and in most cases more suitable for theorists). The program MAXIMA now on SYMBOLICS is inadequate.

UNIX is presently of minor interest to the division although we have one happy user.

Chapter 7

Training

The problem with visitors is that nothing is ever achieved: they change rapidly. There is probably a need for some introductory training of visitors on arrival; one may consider a 3–5 hour practical introduction to mail and text processing. The courses should be given when needed in groups of, say, 6 to 8 after signing up on arrival.

Chapter 8

Documentation

We need the availability of good, simple, practical guides with down-to-earth instructions. This concerns in particular basic use of terminals, printers, etc., access to various systems, mail and text processing. Practical examples are very important. We need in addition better availability of documentation on TEX and plotting. These should be readily available via the divisional secretariat.

Chapter 9

Library

The library has a special status both inside the division as well as inside CERN and the member-state laboratories as a public service. Its operations presently are profoundly changed by the installation of an electronic integrated library system foreseen for 1989 at an approximate cost of 200–250 kSFr for computer hardware and another 75 kSFr for peripherals (input/retrieval stations, printer, barcode readers, etc.). The accompanying software will cost 100–150 kSFr. The optical disc project for preprints will be implemented in late '89 or early '90; the cost for the central equipment is foreseen of 250 kSFr in budget discussions. The exact impact of these electronic systems on the use by internal and external users is presently difficult to evaluate with any precision. However, any remote access station implies 15 kSFr per high resolution screen with PC. A minimum of 5–6 such screens must be installed in public areas. Individual groups at various places inside CERN may ask for private screens. To this should be added eventual network costs, so that image transfer of the optically stored articles can be made with little delay inside CERN. A substantial amount of printing will accompany this service (say 500.000 to 600.000 pages/year).

In the past, access to external databases via the CERN networks has sometimes been difficult. The goal should be that such consultations can be made directly by the R & D personnel without the intervention of intermediaries in SIS.

Chapter 10

Economic Summary of local equipment

Equipping TH with terminals:

20 Mac's, PC's, Mac II's:	120 kSFr
35 graphic screens, etc.:	87.5 kSFr
2 printers:	24 kSFr

Secretariat: Replacement of AES System:

Workstations, screens, printer:	64 kSFr
---------------------------------	---------

Library: Computerized library system:

Hardware:	200-250 kSFr
Software:	100-150 kSFr
Peripherals:	75 kSFr
Optical disc storage system:	250 kSFr
High resolution large screens (5 to 10)	75-150 kSFr

Normal replacement and maintenance at 15%/year should be added as well as operating costs.

Appendix A

Letters from Member State Representatives

In September 1988, the Director of Research responsible for Computing, J.J. Thresher, wrote to members of Restricted ECFA in order to solicit input from outside users in the Member States. His letter is attached.

The replies received are reproduced in full in this Appendix rather than summarized out of context.



ORGANISATION EUROPÉENNE POUR LA RECHERCHE NUCLÉAIRE
EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH

Laboratoire Européen pour la Physique des Particules
European Laboratory for Particle Physics

CERN
CH 1211 GENEVE 23
CERN Site de Prévessin
F 01631 CERN CEDEX

Tele. Telex 619000 CER CH
Telegrams (CERN) GENEVE
Telephone (077)
Direct (077)
Central Exchange (077) 61 11

Publications: The Institute
Publications: The Institute

DC-R/JJT/am 5th September 1988

Dear

Computing at CERN in the 1990's

CERN is currently undertaking a review of its computing requirements which will include an assessment of the expected evolution of high energy physics computing in the near to long term. Although the emphasis is being placed on computing at CERN, the report will take into account the increasing interdependence between the computing facilities at CERN and those available in the home institutes of its users. It is intended to embrace all facets of the computing scene for the next five years (until the mid 1990's), although longer-term issues may also be addressed as appropriate.

The report is being prepared by working groups who report to an Editorial Board that is itself responsible to a Steering Committee which I chair. The target date for the completion of a first draft of the report for discussion at the Editorial Board is mid-October 1988, and already several preliminary drafts from the constituent working groups have been received.

The following topics will be treated in detail in the report:

- Computing for accelerator divisions
- Computing for engineering
- Computing for experiments
- Computing for theorists
- Management information services
- Networking

We are eager to obtain the views of outside users of the CERN computing facilities in the preparation of the report; a number of such users are members of the working groups. Nevertheless we believe that it would be useful in addition to have user input from each Member State, since needs may well vary from country to country.

I am therefore writing to you as a member of the RECFA to ask if you would obtain this input from your country and pass it on to us. We would particularly like to have the views of users on the computing services they would like CERN to provide in the future, with some indication of priorities if possible. Comments or criticism relating to the existing services provided by CERN would also be welcome as well as any other points that you may wish to raise.

Thank you in advance for your co-operation.

With kind regards,

Yours sincerely,

J.J. Thresher
Director of Research

...

DEUTSCHES ELEKTRONEN - SYNCHROTRON DESY

NOTRESTR 65 2000 HAMBURG 52 . TEL. 0410/00 90-0 . TELEFAX 00 90-32 82 . TELEX 2 15 124 4

Dr. H. F. Hoffmann

10 MMS 1989

Dr. J. Thresher
CERN

CH-1211 Geneve 23
Switzerland

Hamburg, March 7, 1989

Dear John,

The query in German HEP institutes concerning the quality of the CERN computing services has shown no striking new points. Generally the services offered are appreciated and found excellent. There are some comments (some multiple) which I just enumerate without further comment:

- Reading, writing video tapes (Exabyte 8mm)
- Better documentation of new software products
- Access to Cray and IBM should be provided for East Europeans
- Batch: Go to Unix-like operating systems; extend VMS services.
- Interactiv: Extend VMS, offer Unix (Ulrix); in the future: replace VM by AIX
- Networking: Experiment-Computer Center: FDDI-bandwidth Universities - Cern: at least 2 Mbit/sec
- more help to outside institutions in networking
- Quality control of DD Software should be improved
- Commercial software centrally acquired at CERN should be free of charge to outside institutions
- Some network services are not very reliable
- Streamline software support, no expensive commercial software competing with CERN products.
- Which software is used in which HEP-Institut world-wide (HEPLib). Offer more coordination.
- CERN program libraries very interesting and helpful for outside institutes
- Better coordination of software HEP-wide
- Development of new software products: stop trying to be compatibel with old versions
- Improve distribution and updates to outside institutes
- More support for outside users by adding DD personnel to experimental teams
- Improve decentralized analysis

Sincerely,

H. F. Hoffmann

- Keep up user support, servies and continue to increase the computing power
- Improve networking with outside institutes:
- Access to commercial packages like Oracle should be easier
- Access to HEP data; literature search, preprints without user ID for outside HEP-people
- Central VAX-cluster inadequate with respect to offered CPU-power and interactive response
- Disk-space in VAX cluster and IBM not sufficient.



Roma 4th Oct. 1988
BIB. Roma (ITALIA) - Fisica Alta Mare 3

Dr. J.J. Thresher
Director of Research
CERN

Your reference DG-R/JJT/sm

dear Dr. Thresher,
as representatives of Italian HEP community in CCC, we answer your letter 5th September 1988 on Computing at CERN in the 90's.

As a general remark, computing in the years 1990's will be more and more distributed on a wide geographic scale involving the exploitation of regional or local computing centers having power a not negligible fraction of the computing power installed at CERN. In order to make an efficient use of the CERN computing center and of the regional and local centers already existing or which are expected to come into operation, a consulting body with membership from high level managements of regional centers should be immediately formed, to establish the coordination needed for their efficient use by part of high energy physicists. Rather than giving political answers, this committee should give technical advice to be followed by centers willing to contribute to the computing needs of High Energy Community in the years 1990's. That will raise a number of explicit technical issues which cannot be covered into a letter, but should be rather discussed within working groups.

We understand that a Green Book Study has been established at CERN, well before your 5th September letter; at our surprise, it does not include any member from Italy, while that would be the appropriate place to illustrate and discuss our views, requirements and priorities.

Here follow comments on specific issues.

Workstations:

the use of workstations is essential for physics analysis, CAD/CAP and graphics applications. CERN should play a leading role in the guidance of local physics communities to choose a restricted number of producers and to adopt software that meets standard requirements in collaboration with HEP community and coordinated between the regional centers. A working group should be established between the regional technical competence and provide a forum for confrontation.

Mass Storage:

this is a main issue for experimentalists. Coordination and agreements should be established between regional centers, including CERN, to ease the transport of large volumes of data from any Laboratory as producer (CERN, DESY, GRAN SASSO, FNAL?) to regional computing centers as users. Suitable device (robots) should be soon installed at CERN, providing a copying facility for transfer and use of data elsewhere. Again, compatibility with the other large centers is essential: labelling of IBM and VAX cassettes must be standardized.

SEDE UNIVERSITA' LA SAPIENZA - DIPARTIMENTO DI FISICA
TELEF. 813233-813235 TELEFAX 4957887 TELEFONO 4716

Enzo Va. Penbe
Chairman of INFN
Computing Comm. HEP

Best regards
Enzo Va. Penbe
Chairman of INFN
Computing Comm. HEP

Enzo Rem. 88i
Director of INFN-CNAF

CPU Power:
the demand on mainframes and links will peak during working hours. This consideration calls for high CPU power mainframes with a high performance on the I/O. From the point of view of the physicist sitting at CERN, regional computing centers will be mainly exploited for processing batch jobs of consolidated software analysis. This computing effort spread between CERN and regional centres makes sense only if each of them will have a CPU and mass storage capacity close to CERN, i.e. the aggregated powers should largely exceed that of CERN. Parallel processing should be implemented on any center willing to contribute in this field. The CERN VAX cluster must become a computer system providing real CPU resources to the users. As is it almost useless. Not only the cluster has to be upgraded, but also it has to be considered the possibility of splitting in two clusters with more homogeneous classes of users. Ease of access to all CERN resources, as well good support for their use, should be provided to all theorists, in particular to short term visitors, (there are at present strong limitations, for instance, in the use of VAXes, of printers supporting TeX, etc.); concerning large scale calculations, CERN should somewhat decide whether to support or just to tolerate them, but, in any case, make any effort to ensure compatibility with other existing centers for intensive computing.

Operating Systems:

any kind of experimental service on UNIX or other operating systems should be established at CERN with care. Right now, it doesn't seem that HEP community has a strong demand for UNIX. In any case a political decision should be taken by the CCC. A strong support, in terms of CERN manpower, has to be kept to support the operating systems used in physics analysis, like HEP-VM and VMS, adopted by the LEP collaborations.

Networking:

it is the central and most crucial point of all the computing policy.
CERN Local Area Network:
At present it has a too low performance.

Wide Area Networks:

On this point we should see a complete change of attitude from the part of CERN. Until now CERN has only accepted data links with external laboratories without taking any responsibility in their management and installation nor contributing to their cost. In spite of this, we found in a preliminary version of the report "CERN Data Networking Requirements in the Nineties", paragraph 5.3, 'Specific requirements', an incredible serie of statements describing the role played by CERN in the past on networking matters. CERN, similarly to other laboratories like DESY and Gran Sasso, should support links with regional centers providing the necessary equipment and manpower for management of connections at the CERN computers. Sharing the cost of the lines between CERN and other laboratories should be also envisaged. All protocols in use, even non OSI, should be supported. Moreover, we believe that the only efficient management of links is within the users community. Therefore we should have a dedicated REPNET network, at the condition that its setting up and management be agreed by the HEP community.

Rutherford Appleton Laboratory

PARTICLE PHYSICS DEPARTMENT

17th October 1988

Chilton, DIDCOI, Oxon
OX11 0QX
Tel Abingdon (0235) 21900 Ext
Direct Line (0235)
Telex 83159 RUTHLB G
Fax (0235) 44 5608

Dr. J.J. Thresher,
EP Division,
CERN,
1211 - Geneva 23,
Switzerland.

Dear John,

Computing at CERN in the 1990's

I am responding to your request of 5th September for my views on the above topic. I have discussed this within the laboratory in particular with Paul Williams and Brian Davies and outside the laboratory with Fred Bullock. Fred, who you wrote to separately, will obtain and send to you the views of the University community. I will try to give you the views of RAL, however, I am sure that you know that these take into account the needs of the UK user community.

Needless to say, I also bear very much in mind the fact that we all want CERN to flourish for the foreseeable future, it is our laboratory.

As you know, I am a member of the Greenbook II study, working in the group convened by Richard Mount. Also as you know I have not been happy with several aspects of the work of this group or indeed with its composition. I believe that the report which will emanate from it has improved compared to its earlier versions and its conclusions I generally accept. I have, however, misgivings on the premises, arguments, presentation and lack of proper priority list. I therefore would like to take this opportunity to give my (our) views.

1. The CERN computer centre is pivotal in providing facilities of all types to the whole CERN community.
2. CERN also plays a dominant role in providing systems and user software for much of the community.
3. Both these functions are essential and should be continued, however, they raise two questions: firstly, should the computer service and systems etc. be part of the same division as the software/hardware support for the physics programme, and secondly, although CERN should continue to be "pivotal" and "dominant" it must for everyone's health become part of an integrated approach to computing in the 1990's.

4. CERN should therefore have as its first priority areas where it is unique because of its location i.e. data storage (both fast and slow), data handling, data copying and the ability to analyse subsets of data quickly. It must be networked by as fast links as possible to its community and in particular to regional centres which would act as nodes in many cases.

5. Having said this, the CERN computer centre also needs to have the power to satisfy the large number of physicists (both CERN staff and visitors) who will wish to work interactively from terminals and work stations. It must also provide a significant batch power, but not necessarily at an astronomical level.

6. The national/regional laboratories must play an important role in the European HEP scene. This is for the good of all including CERN. They should coordinate their efforts in hardware and software development with CERN, act as centres of excellence in their regions and cater for the needs of their local communities. For this they must provide significant data storage and computer power and excellent links both to CERN and to the Universities. They must be involved in critical areas of the work and not just as monte carlo factories. Effective ways must be found for coordinating their activity with CERN and with each other.

A vital aim is for the physics analysis to be carried out largely at home universities and institutes. Care must be taken not to prejudice this by assuming that all problems will be solved by increased line speed to CERN. It is unlikely that this will be true in the next few years and certainly has not been demonstrated to be so.

7. Development of user software and specialised hardware (e.g. J081/E, CAVIAR etc.) should be in general (but not exclusively) demand led and accountable to the "customers".

CERN must make more efforts genuinely to collaborate with outside institutes as peers, the image of "if its not invented here we are not interested" is very damaging and getting any work done jointly with outside institutes has greatly suffered. Maybe CERN needs to spend more money on travel for its own staff to spend significant periods of time in other institutes?

8. I am strongly opposed to private computer centres being set up at CERN. Although the definition of what is and what isn't such a centre is both time dependent and vague, I believe that requests from experimental teams for money to set up computing facilities for bulk processing at CERN is damaging. I also believe that setting up such centres distorts the peer review systems which is used for allocating central CERN resources. This argument is strengthened when these centres are provided wholly or partly by non-member states who do not pay for the infrastructure costs of CERN.

9. The management of large networks of workstations and large volumes and flows of data will pose very challenging problems. CERN must take the lead in solving these problems. Without this much of the investment in links, storage and computing power will be lost and the new generation of experiments will be endangered.

10. With the advent of the completion of LEP I construction and the start of data taking, a new look as to be resources devoted to computing is clearly in order. It is certainly true that there are shortfalls in hardware in some areas and I have indicated the main ones. I believe that just as significantly there may well be (probably are) shortages in manpower. However, a clear spitting of the computer operations from the rest and a clearer understanding by the community of what manpower is available, what their responsibilities are and how these can be influenced, would make the task of judging the level of support much easier, and would be of benefit to computing at CERN and to all of us.

I hope that these comments will be of use to you.
Sincerely,



George Kaimus

cc Paul Williams RAL
 Brian Davies RAL
 Fred Bullock UCL
 John Barlow RAL

1 Introduction

What follows is a summary of the views held at the Institut National de Physique Nucléaire et de Physique des Particules (IN2P3) concerning the future developments in computing. The scope of this note is limited to those aspects relating to the CERN computing program.

Among the 17 Laboratories of IN2P3, 10 are involved in High Energy Physics (HEP), with a total number of physicists which is close to 300. All of these Laboratories are equipped with minicomputers (VAXes) with Ethernet LANs, and are connected to an IBM mainframe, located in Lyon. These computing resources can play and should play the role of one of the "regional centers" considered in the MUSCLE report. Note that the Lyon computer center has two distinctive features :

1. it is entirely dedicated to Particle and Nuclear Physics,
2. it is located not far from CERN

Since we find it difficult to foresee with some accuracy what kind of hardware will be available beyond 1995 and at what price, no attempt has been made to make plans beyond that year. Until then, the main CERN programs in which IN2P3 will be involved are : UA1, UA2, experiments with relativistic ions, LEAR and LEP (for other experiments like GAMS or Hyperons in the Ω , computing does not look like a major issue). We have considered all of these programs, but clearly LEP will be the most demanding one, since it will produce a very large number of events worth analysing. Even at the start of LEP 200, computing needs will increase if the luminosity of LEP 100 is substantially increased.

Within the large collaborations which are usual in our field of research, there are many groups. For the future of HEP, it is crucial that these groups of physicists stay in close contact with their universities (would it be just to have students joining them) and that they also stay in close contact with the technicians of their home Laboratories. This implies that computing developments maintain a right balance between CERN and the regional centers, and that all of these resources be developed in such a way as to ensure maximum efficiency of their operation. In this respect, proper attention should be paid to the fact that an optimisation for short-term goals could jeopardize the long-term future. It may be that the shortest way to get the first physics results on the Z^0 is to centralise all the computing at CERN, but such an attitude would have very damaging effects in the long term : the HEP community would sever its links with the universities and thereby risk being sacrificed for political/economical reasons without an ally to speak for it. Indeed, computing is one of the major tasks for

the event rate is much smaller than at the Z^0 peak. A few 10^4 events per year are expected instead of the 10^7 events per year for LEP 100. One might therefore prognose a drop in the computing needs for LEP 200, before the hadron colliders era (LHC, SSC). Nevertheless the load due to the analysis of the LEP 100 data will continue for some time, and data from HERA will still be taken and analyzed in the regional centers.

The prospect of having to analyze heavy ions collisions at the SPS, with Pb beams, will increase the needs of fixed-target experiments.

In fact, some other scenarios than running only LEP 200 are presently envisaged. A great interest among physicists exists about b physics (b-b mixing, CP violation, ...). Different schemes for getting a b factory by 1995 are presently considered, like a dedicated e^+e^- machine in Europe, or an increase by a factor of 10 or more of the luminosity of LEP at the Z^0 energy. In such schemes, the experiments should accumulate 10^8 events/yr or more, which represents an increase of an order of magnitude compared to LEP 100.

For regional centers, the arrival of any b factory will represent an increase of the load. For CERN, if one envisages an interleaved mode of LEP 200 runs with LEP 100 high-luminosity runs, the turn-on of LEP 200 will not mean a decrease of the overall computing needs.

2.3 Management

It is essential for such an effort to be closely coordinated between CERN and Lyon.

1. Compatibility/transparency between the two IBMs is necessary to allow easy checks between centers. Ideally, one would like to be able to check and update either in situ or remotely from CERN the contents of the databases and of the programs used for processing and Monte Carlo. This point is crucial to avoid the risk of error inherent to decentralized processing. Here a unified system like HEPVM appears to be necessary.

2. Very large numbers of cartridges, a few thousands per experiment in the active years, with raw data and DSTs, will be transferred/copied between the two centers :

- copying facilities and manpower have to be foreseen on both sites,
- shipping of cartridges seems possible on a daily basis, given the vicinity of Lyon, but one should avoid the usual bottlenecks (customs, ...),
- automatic cartridge handling equipment, operating on several thousands of cartridges, is needed in Lyon.

a physicist which can be carried out indiscriminately at home and at CERN, if care is taken of making it at all possible.

As this note will make clear, a regional center like the one of IN2P3 can contribute a very substantial amount to the total computing capability in the three major off line areas : DST production, analysis, and Monte Carlo simulation. For this to work, there are some constraints, mainly organizational, which have to be satisfied. It is the main object of this note to try and analyze these constraints.

2 Data processing and physics analysis

2.1 Present needs

The CERN physics program will require a very large computing effort by 1991-1992. For LEP 100 experiments alone, present estimates are of the order of 80 CERN units per experiment in 1991. This power includes data processing, Monte Carlo, and physics analysis. French experimentalists believe that, even taking into account private computing facilities based at CERN, the corresponding effort is well beyond CERN resources.

IN2P3 physicists alone, of Aleph, Delphi and L3 will need 5 to 10 CPU units¹ per experiment in 1991 to participate in the off-line analysis of their experiments. For non-LEP experiments — UA1, UA2, LEAR, heavy ions, ... — it may be reckoned that up to 10 more units will also be needed.

Still another share will be devoted to non-CERN accelerator experiments, and to non-accelerator experiments, but it may be neglected at this level of accuracy. As far as theorists are concerned, there are two kinds : regular phenomenologists and adepts of lattice gauge theories. The needs of the former can be considered as negligible, while those of the latter are essentially infinite, and heavily oriented towards special vector processors, which are best used on dedicated machines (like APE in Rome).

Accelerator physicists' needs are more unpredictable still, but they may come up some day with large demands.

2.2 Needs to come

The prospect, presently discussed, of increasing the LEP 100 luminosity by an order of magnitude reinforces the need for large computing capabilities in regional centers. The energy upgrade foreseen for LEP from 100 GeV to 200 GeV is assumed to occur between 1992 and 1995. At 200 GeV (above the WW threshold),

¹ CERN CPU unit = one 370/168

4 Networking

4.1 Needs

Physics analysis requires a fast access to a large sample of data stored on DSTs, and this work should be possible to do for physicists at home and in CERN without any substantial change of environment :

1. Very reliable links allowing 64 kbits/s per LEP experiment are necessary to allow frequent updates of constants and programs, and remote running between CERN and Lyon.
2. Physicists foresee to have large samples, if not the whole DSTs, in Lyon. A large disk space — 10 to 20 Go per experiment — will be used to store DST samples on which physics analysis is performed remotely either from CERN or from the IN2P3 Laboratories.
3. Very high rate links (2 Mbits/s) would be needed to transfer large data samples for immediate work (even using very condensed formats like "n-tuples") from Lyon to the workstations operating in the Laboratories. On the other hand, heavy interactive processing in Lyon, e.g. with virtual machines emulating PAWs, would seriously endanger the overall efficiency of the center.

4.2 Resources

Links to IN2P3 Laboratories

- Up to now, the links (leased lines) had a standard rate of 9,600 bits/s, or 14,400 bits/s. Some Laboratories equipped with several links totalled up to about 50,000 bits/s.
- The new standard for IN2P3 will be the 64 kbits/s digital link. All HEP Laboratories in the Paris area will have such a link before the end of 1988. The others will follow in the course of 1989. We do not consider as realistic, in terms of cost/efficiency ratio, to generalize links with a rate higher by an order of magnitude.
- This implies in turn that data for physics analysis should be transmitted through tape, or over the night hours, to the various Laboratories. These methods are somewhat cumbersome and one is eagerly calling for a breakthrough in technology or cost in high-rate links. Here a top-level political action towards PTT would be called for.

A permanent effort from the collaborations and from the centers staffs will be needed to ensure compatible and decentralized management methods for programs, libraries, systems, databases or experimental data. Each collaboration should have its own production office, in permanent contact with the Laboratories and with the centers.

3 Evolution of the computing resources of IN2P3 as a regional center

3.1 Central processing power

IN2P3 is presently equipped with an IBM 3090-200 VF with 64 Mo of main memory, 64 Mo of expanded memory and 32 channels. This represents roughly 10 CERN units. It is to be expected that by 1989, this processing power will be increased by a factor of 2 to 3.

3.2 Peripheral storage

Disks The present space of 80 Go will be extended as soon as 1989 to 120 Go, and it is foreseen to reach 200 Go by 1991-1992.

Tapes

Manual handling. There are presently 10 conventional 3420/8 model tape units and 12 3480 cartridge units. The trend for these equipments is certainly stability or slight decrease, due to the introduction of automated cartridge units.

Automated tape storage. Funds are available since 1987, and a storage unit of at least 1 To (6,000 cartridges of 200 Mo each), presently under study, will be installed in the first months of 1989.

5 Conclusions

Computing in the LEP era will be a demanding task. In order to make the most of all available resources — at CERN, in the regional centers and in the Laboratories — a serious effort of coordination should be done in several domains, so that work is not duplicated, and so that physicists and computer experts can continue their work indifferently, whether they happen to be at home or at CERN.

A severe obstacle for this program is the price of high-rate links, which is an order of magnitude above their real cost, or above their US price, for example. This is largely due to the fact that a large part of our traffic is classified, and priced, as international. A top-level action should be undertaken to grant our research more realistic tariffs. Until this action succeeds, we are all the more urged to use with a maximum efficiency the means we have at our disposal everywhere.

As an example of the feasibility of such an enterprise, we have already reached a unified system on all IBMs, HEPVIM. This is not a job done once for all, but it is a living organism, with regular meetings, issuing updates to keep pace with the changing environment. The effectiveness of this teamwork is due to the choice of the right partners, at the right level, and with the right complement of expertise.

We advocate the creation along these lines of several coordination groups with specific domains of activity :

- Networking — This group should not be a substitute to ECFA SG5, which has more a prospective calling.
- Data management and exchange — This group should be set up before the community becomes flooded with data, the lines clogged and the mass storages mess storages.
- Workstations — Will it be possible for a physicist to work at home and at CERN on a workstation without having to read each time a new user's manual and to change his finger reflexes ?

• ...

Again, special care should be taken to choose the right people for these groups, people who are really active in the field, and who enjoy a sufficient status, so as to be collectively of influence. Of course, physicists, especially from the LEP collaborations, should be in fair numbers.

There is the need for some existing body to call these groups to life, and to follow and encourage their action : LEPC, SPC, HEPCCC, ...

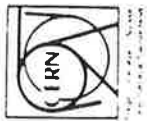
- The task of producing and shipping cartridges to the Laboratories for physics analysis, which represents a large load for Lyon, will correspondingly reduce the burden on CERN.

Links to outside networks

- the link to EARN is done through a 9,600 bits/s line between Lyon and Montpellier,
- a link between Saclay (DPHE) and Lyon is being established. This triangle structure with a node in Saclay, a node in Lyon and a node at CERN, will provide all necessary back-up and safety for the Paris area HEP community,
- a direct link between DESY and Lyon is under study, unless access to DESY is possible via CERN.

Links to CERN

- A leased line exists between the computing center of IN2P3 and CERN since about 1976. Its rate has grown from 9.6 kbits/s then to 64 now. Virtual needs have been, and still are, much larger. The basic limitation is economical : the cost of international lines. It seems very desirable to increase towards a 2 Mbits/s rate, but it looks unrealistic to reach this order of magnitude under the present tariff conditions (1 MSF/yr). Large amounts of data will have to be transferred in the form of carloads of tapes.
- The main uses of the link are now remote login and file transfer. The increase of the line rate should allow an increase of the X25 capacity in order to reach the private X25 networks within CERN.
- Until the various manufacturers supply homogeneous OSI-based protocols, we compromise between standardization and efficiency, using in different contexts the Coloured Books, TCP/IP, Decnet and RSCS.



ORGANISATION EUROPÉENNE POUR LA RECHERCHE NUCLÉAIRE
EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH

Laboratoire Européen pour la Physique des Particules
European Laboratory for Particle Physics

CERN
CH 1211 GENEVE 23
CERN Site de Prévessin
F 01631 CERN CEDEX

Télé. Telex 419000 CER CH
Télégramme CERNLAB GENEVE
Téléphone (022) 767101
Divers (022) 767101
Central exchange (022) 767101

Computing at CERN in the 1990's
Bureau informatique - Dienst Informatica

Computing at CERN in the 1990's, DG-R/JJT/sm

23rd September 1988

Dear Dr. Thresher,

This week we have had a discussion at NIKHEF-H with the relevant experts of the various physics groups and the computer group itself.

1. The main concern is whether the organisation of the computing needs for the LEP experiments takes into account sufficiently the needs and interests of the home institutes (universities and dedicated institutes).
After the considerable and significant contributions of the home institutes to the hardware and software in the construction stage of the LEP experiments, they should get an equivalent return in the exploitation stage of the detectors.
This is not only of importance for the scientific life and viability of the home institutes, but also as a justification of the large investments approved by scientific committees and granted by funding agencies, from which we also have to secure the future, large running costs of the experiments.
The computing activities contribute largely to this return. Therefore, the home institutes should have a major function in the data processing and physics analysis of the experiments.

In this respect, the dominant position of CERN should not lead to the creation of two classes of physicists, of which the least equipped are those in the home institutes. Hence, it is desirable that other large centres outside CERN, e.g. RAL, Lyon, etc., are involved from the very beginning in the large scale data processing. These large centres should serve as distribution points to the smaller institutes and groups to draw maximum benefit from the available potential of scientific manpower in the member states. The optimum potential will not be reached by installing all computing capacity at CERN only. In this context, the "private" main frames installed at CERN are a cause of concern as well. It should be avoided that participating groups have to provide funds and are charged for computing via many competing channels drawing highly needed funds away from investment in the home institutes.

2. Present estimates for computing capacity, disk space, etc., need better founded, more realistic and in particular physics arguments as justification for large investments than up to now. Considerations of competitiveness between experiments are an insufficient guideline and justification for the large requirements in capacity which are brought forward.

3. It is desirable that CERN provides a UNIX service, as this system is gaining field on the world market and is implemented at various home institutes.

4. A coordinating function of CERN is needed in the field of R&D on parallel facilities.

5. The coordination of software for work stations is insufficient.

I hope that these few points collected on short notice do contribute to the discussion on Computing at CERN in the 1990's.

Yours sincerely

C. Daum, RECFA member for the Netherlands



ORGANISATION EUROPÉENNE POUR LA RECHERCHE NUCLÉAIRE
EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH

Laboratoire Européen pour la Physique des Particules
European Laboratory for Particle Physics

Telex: 419000 CER CH
Telegram: CERNLAB GENEVE
Telephone: Direct: (022) 83 30 25
Exchange: (022) 83 61 11

Your reference: DG R/JTT/m
Our reference: EP/PS 195/FV

Prof. J.J. Thresher
CERN/DG
1211 Geneva 23

5 December 1988

Dear Prof. Thresher,

After I had received your letter of September 5, concerning computing at CERN in the 1990's, I investigated by letter (attaching a copy of your letter) to all Swiss High-Energy Physics institutes about their views. None of the institutes involved in big experiments at CERN (LEP, UA2) replied. These institutes have efficient communication channels with CERN via their experiments. On the other hand, institutes involved in smaller experiments at CERN or using CERN facilities only occasionally have responded. The answers may be summarized as follows:

The need in CPU-time at CERN will increase by at least a factor of 2 for those smaller experiments.

Software at CERN should be developed and the software group at DD strengthened in future. In particular, software should be made compatible for use at machines like VAX, IBM, CRAY II and CRAY-XMP.

CERN should be connected to the future SWITCH network.

CERN should reinforce software and hardware support for smaller groups or experiments.

In choosing ways of development or extensions, CERN should seek for solutions, which are affordable also for those smaller institutes, which would like to maintain compatibility with CERN facilities.

I hope this letter is of some use for your report.

Yours sincerely,

L. Tauscher

(Swiss Representative in restricted ECFA)

Audi Dr. Molner-30
CERN, Verspion (VALENCIA)
ESPANA

(86) 341870
60119 ICE

Mr. J.J. Thresher
Director of Research
CERN
1211 Geneva 23
SUIZA

Valencia, 27th October 1988

Dear Mr. J.J. Thresher,

Thank you very much for your letter of last 5th September 1988 concerning the Computing at CERN in the 1990's document.

I have forwarded this letter to all high energy physics groups responsible in Spain.

Here follow the main comments from spanish physicists community.

1) At the moment in Spain we have a HEP network based on 7 VAXes, working in DECNET, connected to CERN. This is an important achievement and we give it as information. The network is called FAENET.

2) Three largest groups (eg. Barcelona, Madrid and Valencia) participating at three LEP experiments have at their disposal three IBM 3090, well equipped, ready to be connected each other, to FAENET and to EARN.

3) The most important request from Spanish physicists is

- to urge CERN to let the 3 spanish IBMs to be connected directly to CERN IDMs complex.

- ask CERN to push for high bandwidth communication line (64 Kbits) between CERN and Spain (IBM to IBM).

I believe that CERN will take into account this requirement for a better efficiency of LEP event analysis.

Sincerely yours,

A. Ferrer
HEP Group Leader
RECFA member.



universitaire instelling antwerpen

Departement Natuurkunde

datum October 11, 1988
NAT/FV/1d
11/88-366

Prof. J. J. THRESHER
Director of Research
CERN
CH-1211 GENEVE 23

Re: Computing at CERN in the 1990's

Dear Dr. Thresher,

As you state in your letter of September 5, 1988, some Belgian users with responsibilities are members of working groups which prepare the report and therefore have their input to the document. After consultation of the group leaders I would like to summarize the general remarks as follows:

- 1) As far as hardware is concerned at CERN, people are rather confident that most needs will be covered but they still want to stress the need for large amounts of storage capacity on one hand and active support for high bandwidth transmission of data to the laboratories in member states who prepare for it.
- 2) On the software side there is serious worry that support is degrading considerably e.g. concerning software delivered by third parties, where often information and flexibility is lacking. Taking graphics as an example, it is felt that the rapid evolutions of the needs and the standards are delaying some developments and the fact that software is delivered from third parties does not ease the situation.

- 3) The most important remark is the following.

During the last years, numerous university groups have made very large efforts to raise funds, find manpower and build detector parts for LEP experiments and the time of exploitation is coming soon. These university groups also realized that their resources on computing should be adapted to these large scale experiments and have either already invested considerably or will do so in the near future, because they realize that A SERIOUS FRACTION OF THE PHYSICS ANALYSIS MUST BE DONE AT THE UNIVERSITIES.

Many people fear that there is too strong a movement towards concentrating all the physics analysis at CERN and; although this might be the most efficient way of working at first sight, this would

.../...

Universiteitplein 1 - B-2610 Antwerpen (Witrijk) België - Tel (03) 820 24 41 & 42 - Telex 33646
Telefax (03) 820 20 48 & 49 - BITNET (EARN) : SECR@BANUJAS

.../...

NAT/FV/1d
11/88-366

- 1) leave an important fraction of the university groups physicists "unused" for the analysis and
- 1i) would quickly lead to a loss of support for HEP by university and other funding authorities.

We really fear that not enough responsible people at CERN realize this (and act accordingly!).

Hoping that these few general comments are helpful to you, I remain

yours sincerely,

Prof. F. VERBEURE
Belgian representative at R.ECF.

Goran Larlskog
Representative for Sweden

J.J. Thresher



Delegação de Coimbra
Departamento de Física
Universidade de Coimbra
3000 Coimbra
Telex 23671-23675 29252
Fax 52801 DEFOUC P

LABORATÓRIO DE INSTRUMENTAÇÃO E FÍSICA EXPERIMENTAL DE PARTÍCULAS

Contribuição Nº 501694-650

Dr. J.J. Thresher
Director of Research
CERN

CH - 1211 GENÈVE 23
Suíça

number of
copies of
your ref
date
date

DG-R/JJT/sm of
5 th September 1988
27 th September 1988

number
subject

Computing at CERN in the
1990's

5th. December 1988

Dear Dr. Thresher,

From Sweden I have the following comments and wishes:

1. TCP/IP support should be given for remote access of files and computers at CERN (is now missing)
2. Upgrading of all links to CERN to ultimately 2 MB important
3. The CERNVAX system is congested. The groups used to Vaxes are eager/anxious to have good VAX facilities also at CERN (strong wish from the Stockholm team)
4. The groups used to Apollon also wish CERN to give continued support to Apollo systems, especially interesting now that the 10000 is out! (strong wish from the Lund team)
5. Upgrade of old CERN terminals to graphic terminals, also easier access to terminals for visitors
6. Continued availability of CERN library updates to outside users

Best regards,

Goran Larlskog

Dear Dr. J.J. Thresher:

I send you a report "Computing facilities for High Energy Physics in Portugal" that was prepared by a colleague of mine about the situation and expected evolution of high energy physics computing facilities in my country.

I hope that this report can help you, although I realize it does not correspond exactly to the needs you refer. It is anyway the best we can do now.

With kind regards.

Yours sincerely

Armando Policarpo
(Prof. Armando Policarpo)

C. Kourkouvelis
UNIVERSITY OF ATHENS
PHYSICS DEPARTMENT
NUCLEAR AND PARTICLE PHYSICS SECTION
ΠΑΝΕΠΙΣΤΗΜΙΟΥΠΟΛΙΣ - ΚΟΥΦΟΝΙΑ
157 71, ATHENS - GREECE . TEL. 7243 382
106 60 ATHENS - GREECE . TEL. 3637 438

Dr. J. Thresher,
Director of Research
CERN,
1211 Geneva,
Switzerland

15 January 1989

Dear Doctor Thresher,

I apologize for the delay of this answer concerning my country's input for the computing at CERN in the 1990's. I have been away of the country for half of the time and for about a month I did not have access to my office due to student demonstrations!

I tried to gather the HEP community views on the computing facilities, and I succeeded to do so for most of the groups based in Athens. I also used some of the data from joint requests all the groups made, less than a year ago, for computing to the Greek secretariat of Research and Technology.

As far as the computing services we would like CERN to provide, I can state the following:

- 1) For the use of the Greek groups (80-100 users) we need a power of at least 3 CERN units, spread over the mainframes (VAX cluster, IBM 390, Cray). Also we need a disk space of at least 10 Gbytes.
- 2) Tap mounts (3480 cassettes or 6250 bpi) 1000-2000 per year. We would also like to see CERN providing facilities for reading and handling the new exabyte tapes, and more access to writing/reading TK50's.
- 3) Given the number of users we would like to have access for a large percentage of time to at least 12 work stations of > 3000 series with disk storage > 150 Mbyte. Greek users find difficulties in finding terminals in common rooms (RIOS like). We would like to see more, general use, terminals. Another difficulty encountered was the long delays for logging in the systems. We would like to see an improvement there.
- 4) Since most of the time we are based in Greece and we have to transfer files and data we need a DECNET type line of > 64 Kbit/

sec speed which we expect the government to provide. We would like CERN to provide access for remote login for a least 15 simultaneous users.

5) Lastly, a general comment: since the new large experiments make necessary the use of parallel processing we would like to see CERN supporting and developing farms of microprocessors for off-line reconstruction and lattice calculations.

I hope that this information reaches you in time and is of some use to you. I remain.

Yours sincerely,



C. Kourkouvelis,
CERN representative in RECFA



UNIVERSITY COLLEGE LONDON
DEPARTMENT OF PHYSICS AND ASTRONOMY

including
UNIVERSITY OF LONDON OBSERVATORY
and
MILLARD SPACE SCIENCE LABORATORY

Telephone 01 887 7050
Fax 01 887 7050
Telex 98722 UCLPHYS G
Facsimile 01 887 7145

Department of Physics
and Astronomy
University College London
Gower Street
London WC1E 6BT

25 October 1988

Dr J J Thresher
CERN
CH-1211 Geneva 23
Switzerland

Dear John

Computing at CERN in the 1990's

I have now had some responses to your letter concerning the UK attitudes to the above subject. I am enclosing copies of the written responses received so far, at the risk of adding yet more paper to what must be a considerable pile already. My justification for this is that, the responses contain expressions of disquiet in one or two areas. The conversations I have had with various users of CERN tend to support these worries. There are three themes which recur:-

1. The belief that CERN is too fond of inventing software (and hardware) solutions when perfectly adequate solutions are available commercially. The result is that outside users are either unable to make use of the CERN solution at all, or they must commit scarce resources to support the solution in order to be able effectively to participate in the research. People would prefer to see more effort going into the negotiation of favourable community-wide deals on commercially supported products.
2. The inadequacy of user support. It is clear that users expect, rightly or wrongly, that there will always be available at CERN someone to help them with problems in using the CERN facilities or CERN supported software. There is obviously some concern that CERN does not fully understand the requirements of UK University users. The suggestion of CERN staff taking sabbaticals in Universities seems interesting in this respect.
3. The fear that history will repeat itself and only those who can afford to be on-site at CERN will be able fully to participate in the analysing of LEP data. This is at the heart of the remarks in every letter (and conversation) concerning the present inadequacy of networking. However, I believe the problem is deeper than just networking. It is not difficult to perceive the growth of 'private' computer centres at CERN. The possibility of being excluded thereby from the most exciting topics has been highlighted for UK users by the severe reduction in their travel possibilities.

- 2 -

The above three points need to be addressed in the new Green Book exercise and such criticisms answered effectively.

As for the rest of the points contained in the responses, they are much as might be expected and reflect the particular interests of the writers. In some cases there is unawareness of what the current situation is but this may also be evidence that point 2, above must be taken seriously.

I imagine that more responses will have been received by the time of the HEPCCC so I can convey any further points to you then.

Yours sincerely

Professor F W Bullock