ORACLE
EDUCATION AND RESEARCH

ORACLE

EDUCATION & RESEARCH Training

CERN and the Oracle Value to Research

February 3rd , 2014

- Eva Dafonte Perez, Deputy Head of Database Services, CERN
- Monica Marinucci, Director for Research, Global Education & Research Business Unit

INTERNAL WEBINARS

Jan 27th

Oracle on Oracle Win: Peoplesoft on Exadata at INRA

Oracle Speaker: Philippe LEMERLE, Sale Rep Tech France, Education & Research
Recording available here **http://oukc.oracle.com/static12/opn/login/?t=checkusercookies%7Cr=-1%7Cc=1451669392**

Feb 3rd

CERN and Oracle value in Research

CERN Speaker:
Eva DAFONTE PEREZ, Deputy Head of Database Services Group, CERN

Feb 10th

Research Project Portfolio Management on Fusion at Pacific Northwest National Lab

PNNL Speaker:
Rich Davies, Division Manager, PNNL - Jeff Deal, Battelle Memorial Institute

Feb 24th

Exadata in Lifescience: a cost-effective and scalable Research solution for the Swiss Bioinformatics Institute

Swiss Bioinformatics Institute (SIB) Speaker:
prof. Ioannis Xenarios, Director, SIB

EXTERNAL WEBINARS

Feb 11th

CERN and Oracle value in Research

CERN Speaker:
Erich GRANCHER, Head of Database Services Group, CERN

Invite your customers

All Webinars take place at 5pm CET / 11am ET / 8am PDT
Further Information: monica.marinucci@oracle.com

3 February 2014, Eva Dafonte Pérez

CERN, deputy head of database services

CERN and Oracle, a 30-year collaboration

Outlook

- CERN
- History of using Oracle
- Current usage
- Collaboration
- Why using Oracle in our research environment?

CERN

- European Organization for Nuclear Research
  - Founded in 1954
  - Research: Seeking and finding answers to questions about the Universe
  - Technology, International collaboration, Education

Twenty Member States

Austria, Belgium, Bulgaria, Czech Republic, Denmark, Finland, France, Germany, Greece, Italy, Hungary, Netherlands, Norway, Poland, Portugal, Slovakia, Spain, Sweden, Switzerland, United Kingdom

Seven Observer States

European Commission, USA, Russian Federation, India, Japan, Turkey, UNESCO

Associate Member States

Israel, Serbia

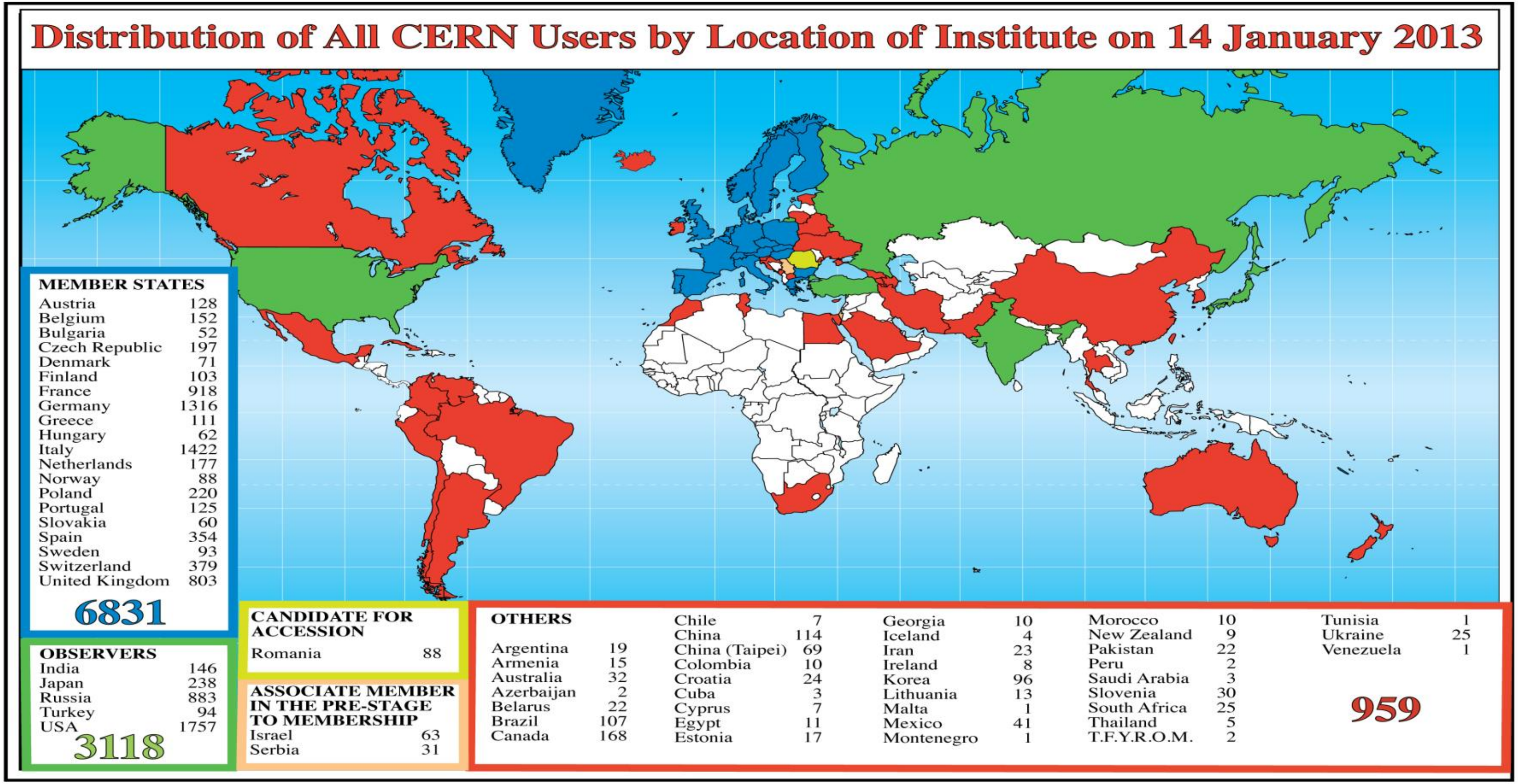
Candidate State

Romania

People

~2400 Staff, ~900 Students, post-docs and undergraduates, ~9000 Users, ~2000 Contractors

A European Laboratory with Global reach



Distribution of All CERN Users by Location of Institute on 14 January 2013

MEMBER STATES

Austria	128
Belgium	152
Bulgaria	52
Czech Republic	197
Denmark	71
Finland	103
France	918
Germany	1316
Greece	111
Hungary	62
Italy	1422
Netherlands	177
Norway	88
Poland	220
Portugal	125
Slovakia	60
Spain	354
Sweden	93
Switzerland	379
United Kingdom	803

6831

OBSERVERS

India	146
Japan	238
Russia	883
Turkey	94
USA	1757

3118

CANDIDATE FOR ACCESSION

Romania	88

ASSOCIATE MEMBER IN THE PRE-STAGE TO MEMBERSHIP

Israel	63
Serbia	31

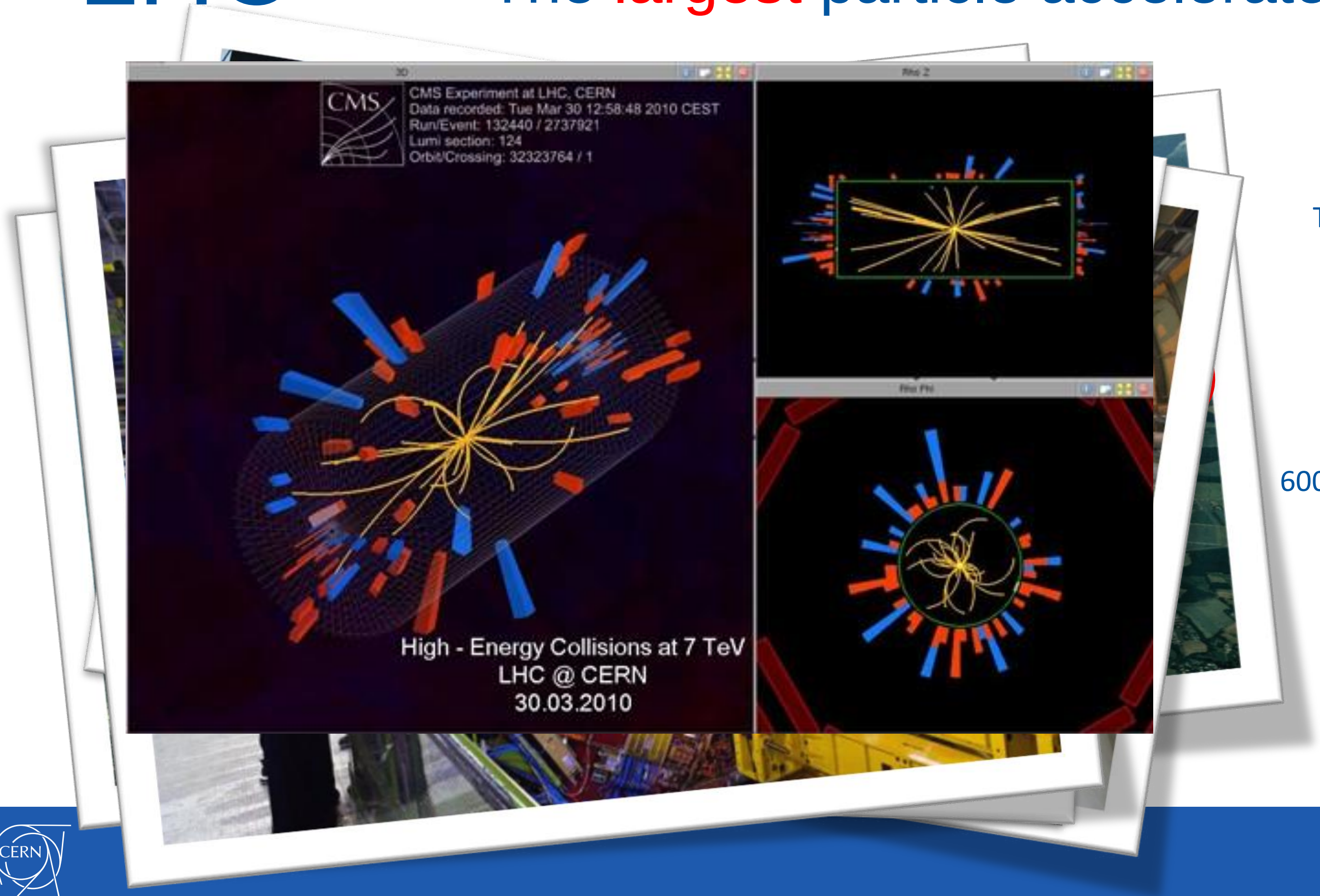
OTHERS

Argentina	19
Armenia	15
Australia	32
Azerbaijan	2
Belarus	22
Brazil	107
Canada	168
Chile	7
China	114
China (Taipei)	69
Colombia	10
Croatia	24
Cuba	3
Cyprus	7
Egypt	11
Estonia	17
Georgia	10
Iceland	4
Iran	23
Ireland	8
Korea	96
Lithuania	13
Malta	1
Mexico	41
Montenegro	1
Morocco	10
New Zealand	9
Pakistan	22
Peru	2
Saudi Arabia	3
Slovenia	30
South Africa	25
Thailand	5
T.F.Y.R.O.M.	2
Tunisia	1
Ukraine	25
Venezuela	1

959

LHC

The largest particle accelerator & detectors



CMS Experiment at LHC, CERN
Data recorded: Tue Mar 30 12:58:48 2010 CEST
Run/Event: 132440 / 2737921
Lumi section: 124
Orbit/Crossing: 32323764 / 1

High - Energy Collisions at 7 TeV
LHC @ CERN
30.03.2010

17 miles (27km) long tunnel

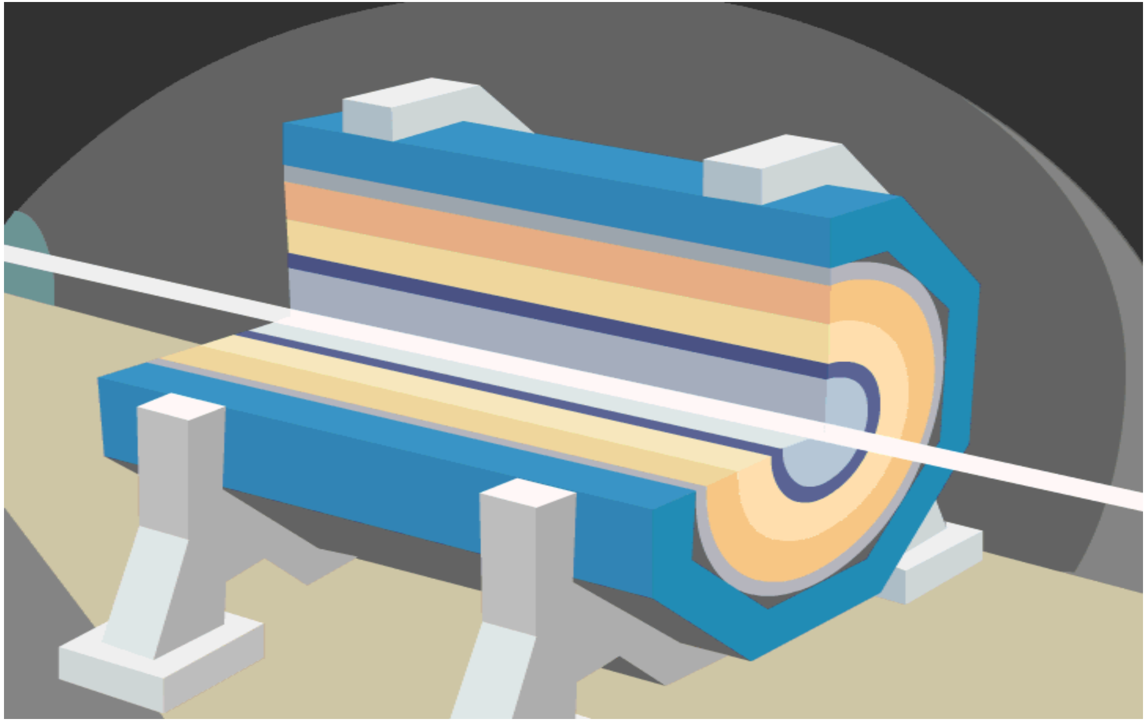
Thousands of superconducting magnets

Coldest place in the Universe: *1.9 K*

Ultra vacuum: 10x emptier than on the Moon

600 million collisions per second / analysis is like finding a needle in 20 million haystacks

Events at LHC

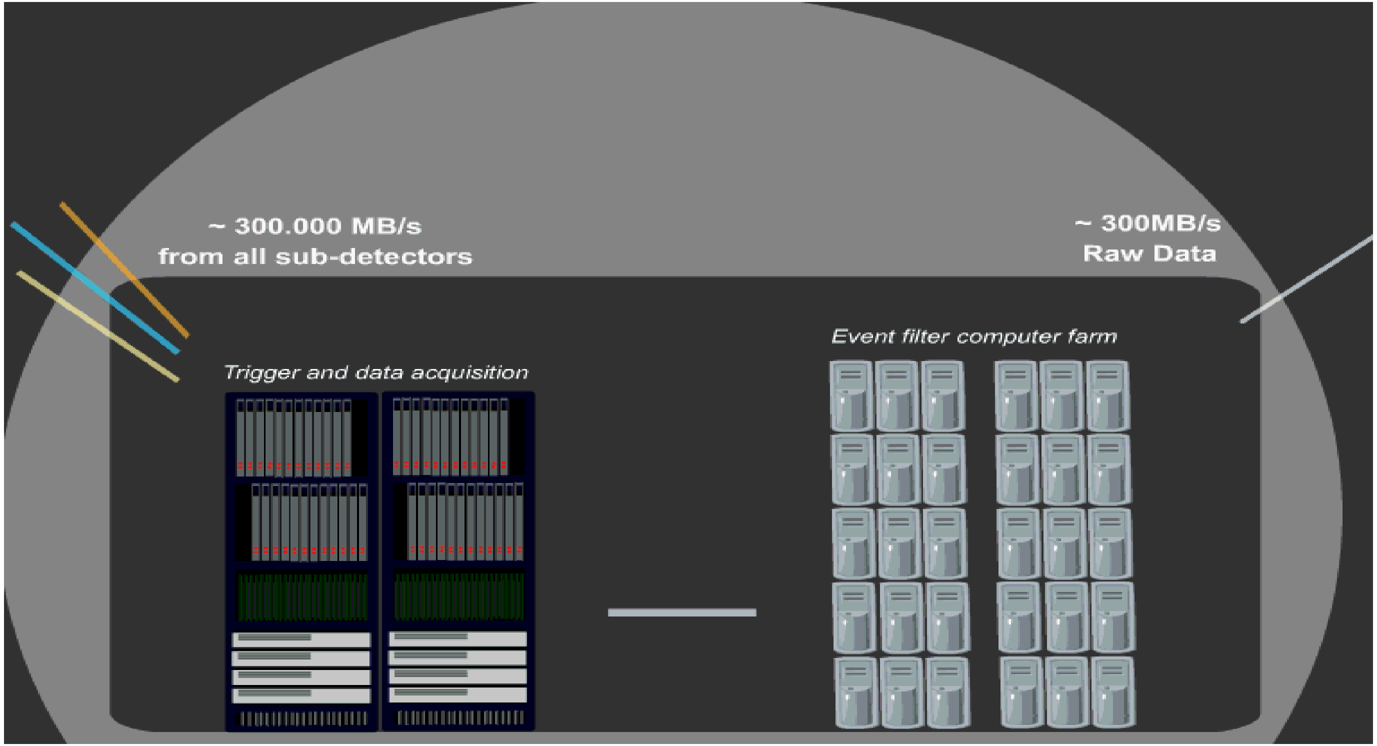


Luminosity :
10^{34}cm^{-2} s^{-1}

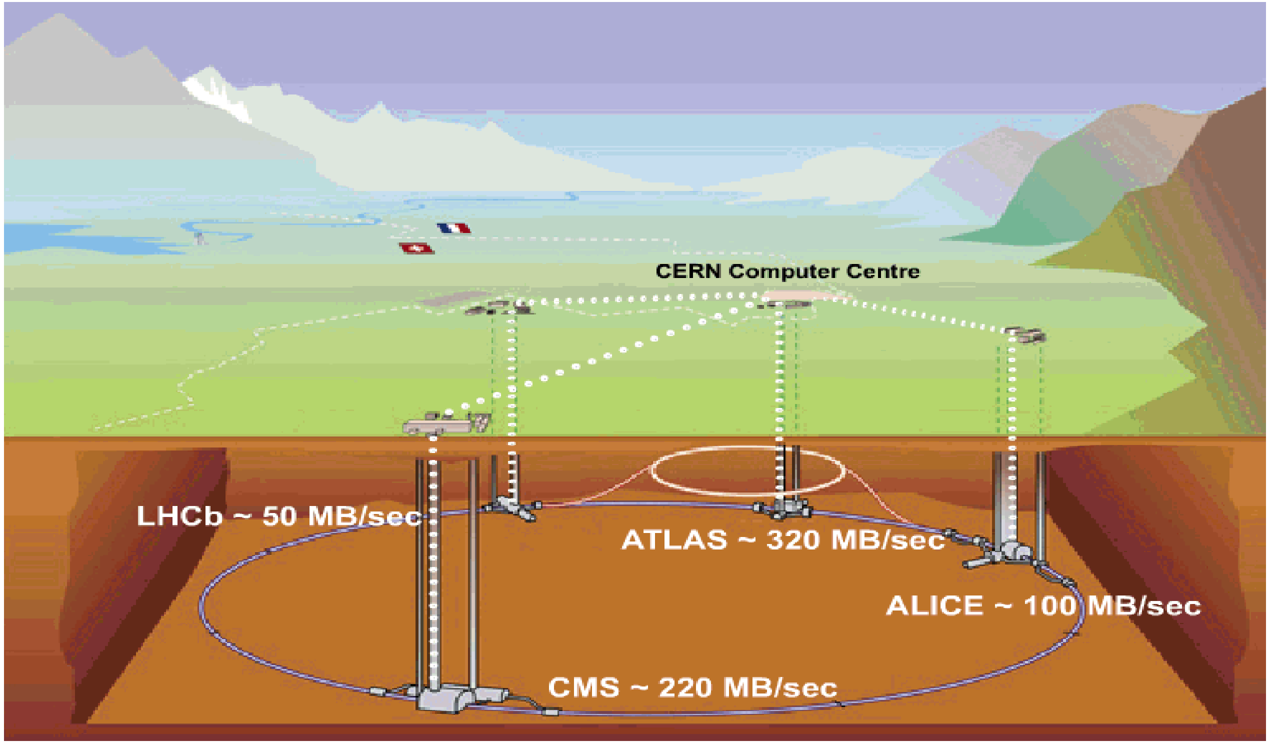
40 MHz – every 25 ns

20 events overlaying

Trigger & Data Acquisition



Data Recording



World’s largest computing grid - WLCG

1 PB raw data per second before filtering

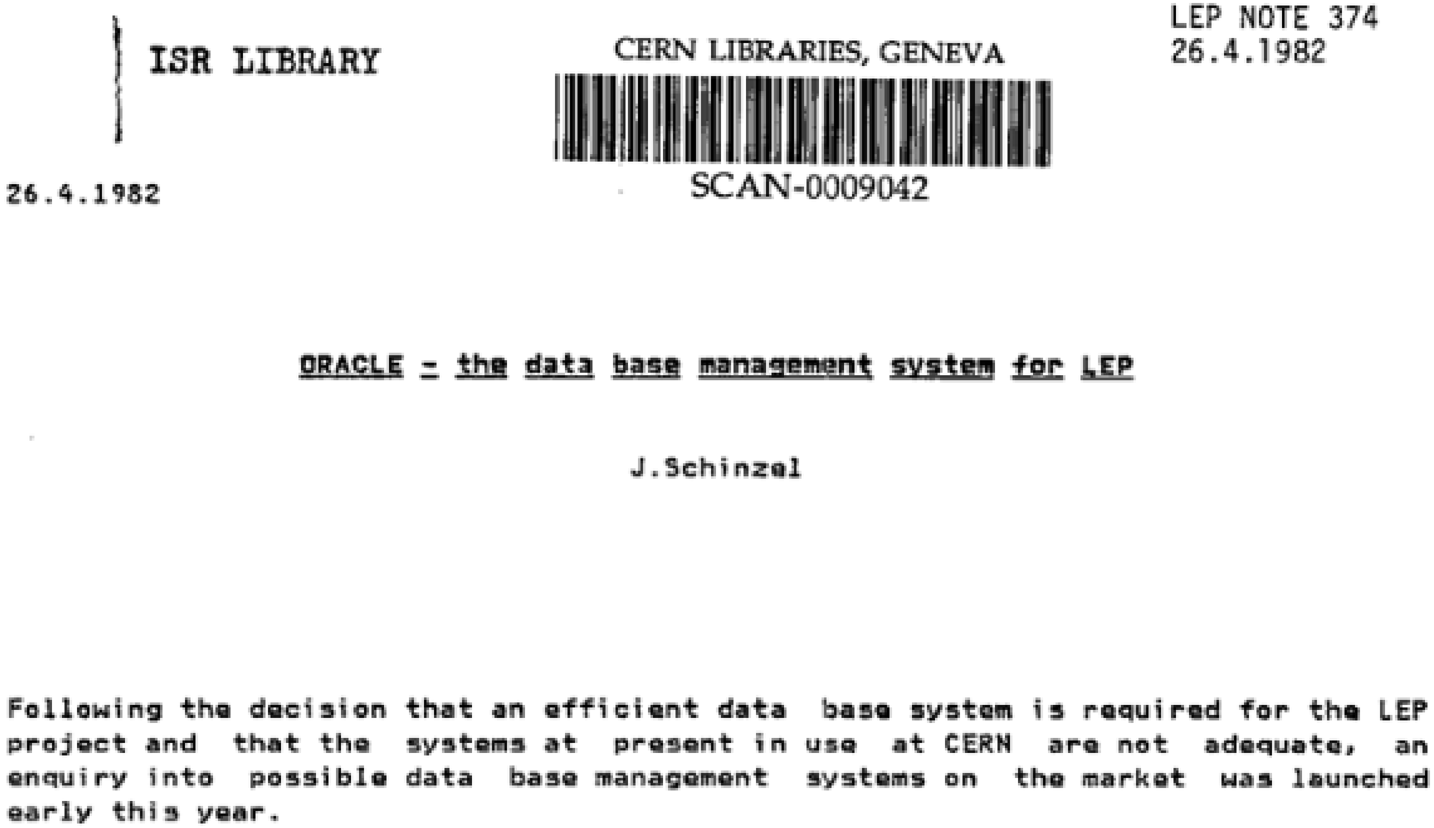
>20 PB of new data annually

68,889 physical CPUs / 305,935 logical CPUS

157 computer centres around the world

Oracle at CERN, 1982 accelerator control

http://cds.cern.ch/record/443114?ln=en



LEP NOTE 374
26.4.1982

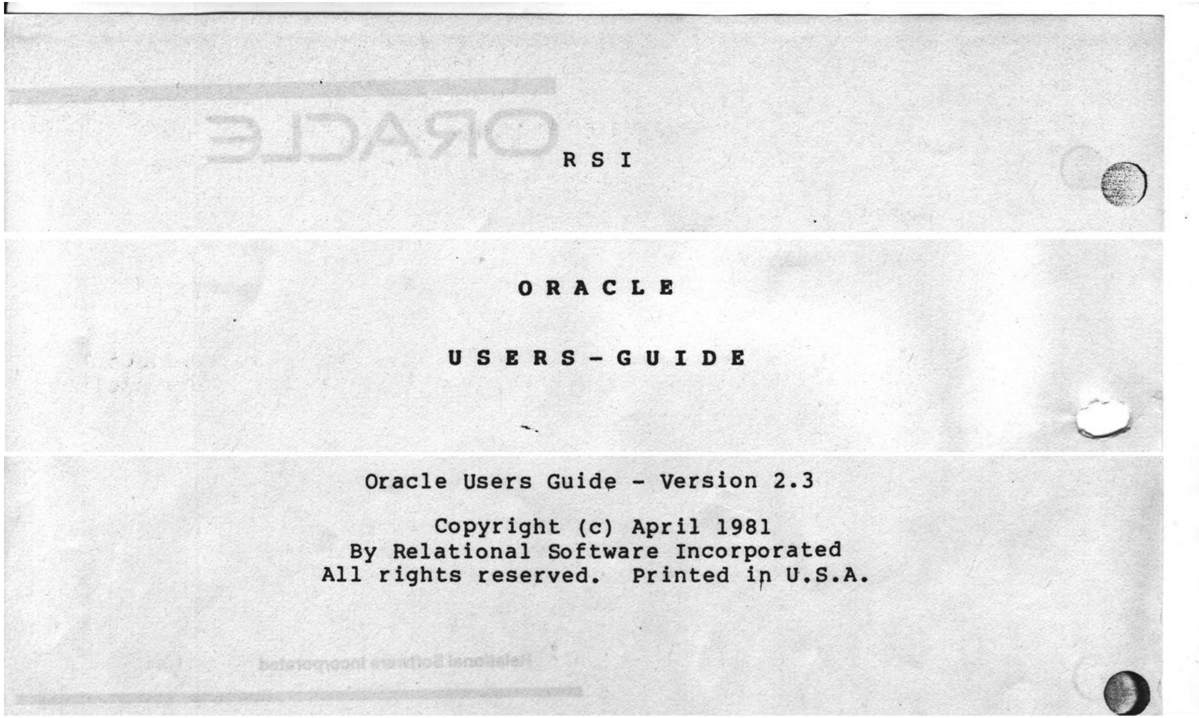
26.4.1982

ORACLE - the data base management system for LEP

J.Schinzel

Following the decision that an efficient data base system is required for the LEP project and that the systems at present in use at CERN are not adequate, an enquiry into possible data base management systems on the market was launched early this year.

Oracle at CERN, version 2.3



RSI

ORACLE

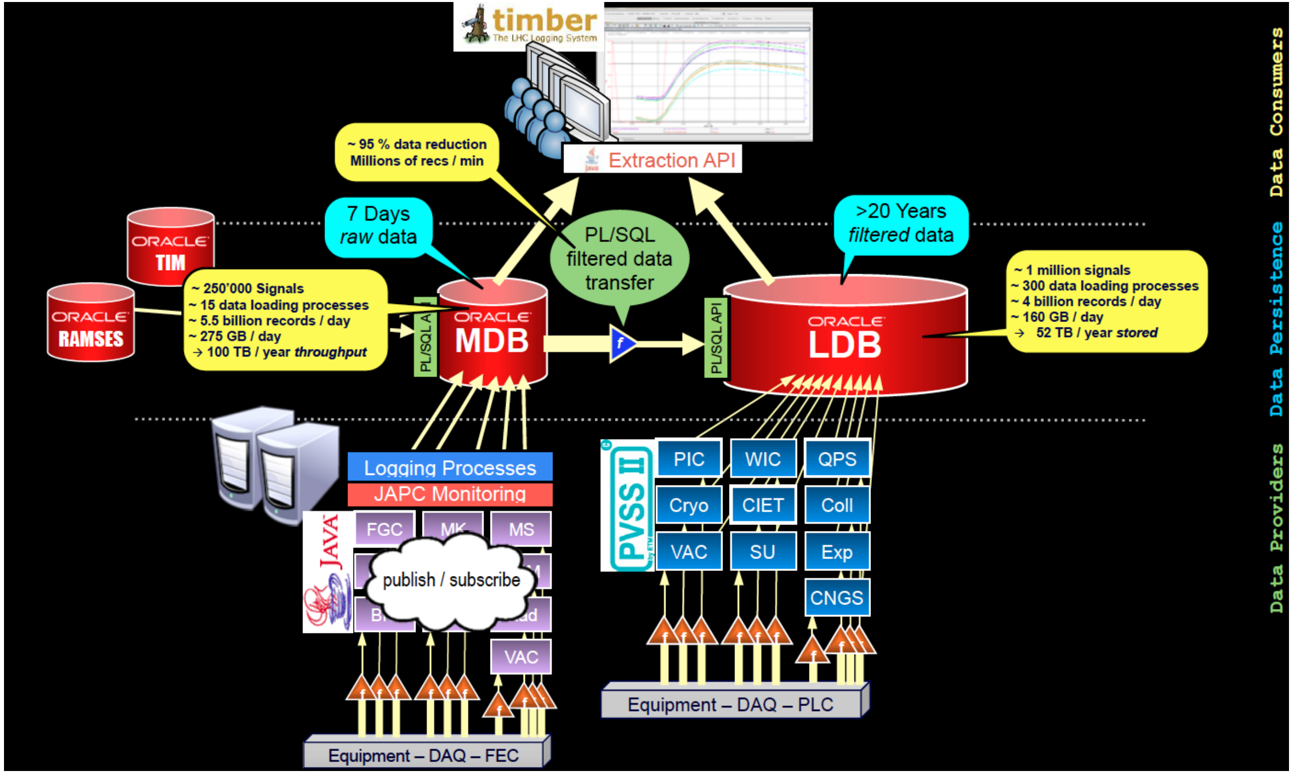
USERS-GUIDE

Oracle Users Guide - Version 2.3

Copyright (c) April 1981
By Relational Software Incorporated
All rights reserved. Printed in U.S.A.

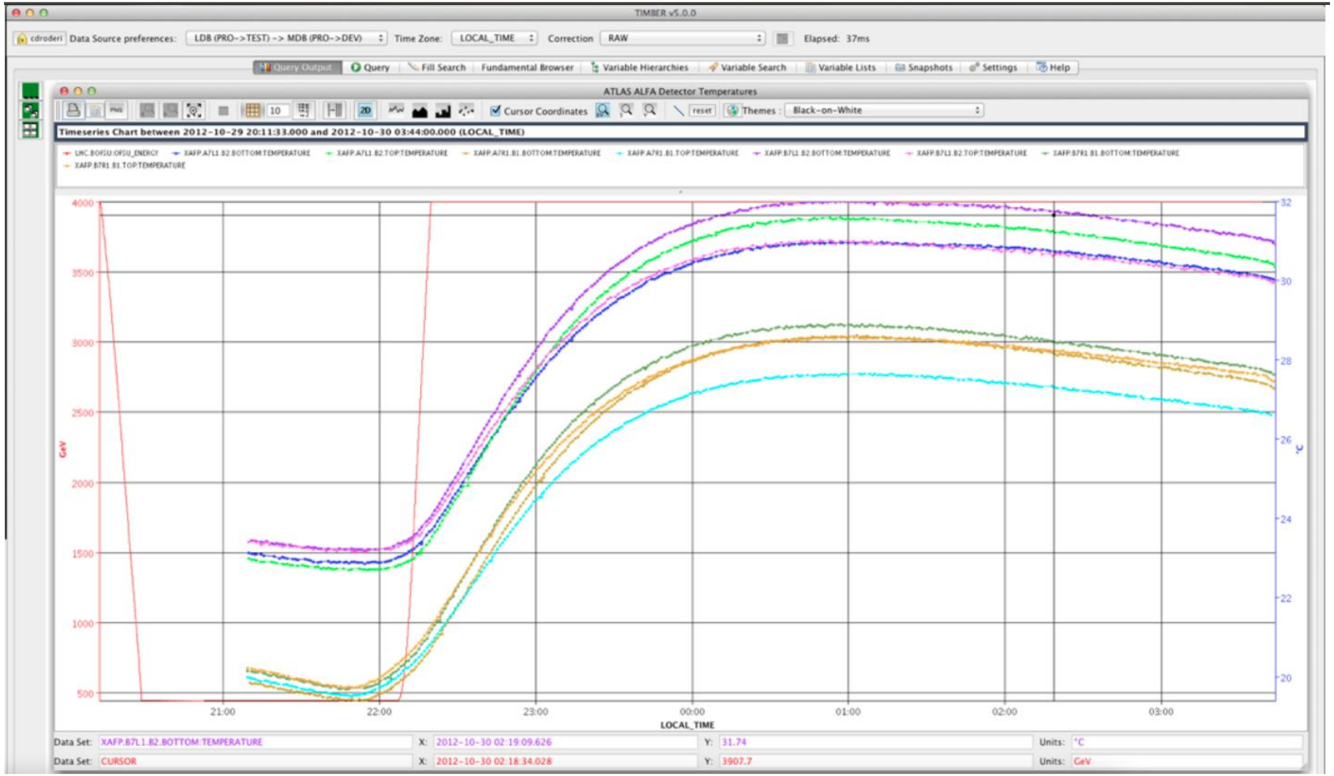
Credit: N. Segura Chinchilla

Accelerator logging



Credit: C. Roderick

Accelerator logging


TIMBER v5.0.0
cdroderi
Query Output
Query
Fill Search
Fundamental Browser
Variable Hierarchies
Variable Search
Variable Lists
Snapshots
Settings
Help
ATLAS ALFA Detector Temperatures
10
2D
reset
LHC.BOFSU:OFSU_ENERGY
XAFP.A7L1.B2.BOTTOM:TEMPERATURE
XAFP.A7L1.B2.TOP:TEMPERATURE
XAFP.A7R1.B1.BOTTOM:TEMPERATURE
XAFP.A7R1.B1.TOP:TEMPERATURE
XAFP.B7L1.B2.BOTTOM:TEMPERATURE
XAFP.B7L1.B2.TOP:TEMPERATURE
XAFP.B7R1.B1.BOTTOM:TEMPERATURE
XAFP.B7R1.B1.TOP:TEMPERATURE
4000
3500
3000
2500
2000
1500
1000
500
GeV
32
30
28
26
24
22
20
°C
21:00
22:00
23:00
00:00
01:00
02:00
03:00
LOCAL_TIME

Data Source preferences: LDB (PRO->TEST) -> MDB (PRO->DEV)
Time Zone: LOCAL_TIME
Correction RAW
Elapsed: 37ms

☑ Cursor Coordinates
Themes : Black-on-White

Timeseries Chart between 2012-10-29 20:11:33.000 and 2012-10-30 03:44:00.000 (LOCAL_TIME)

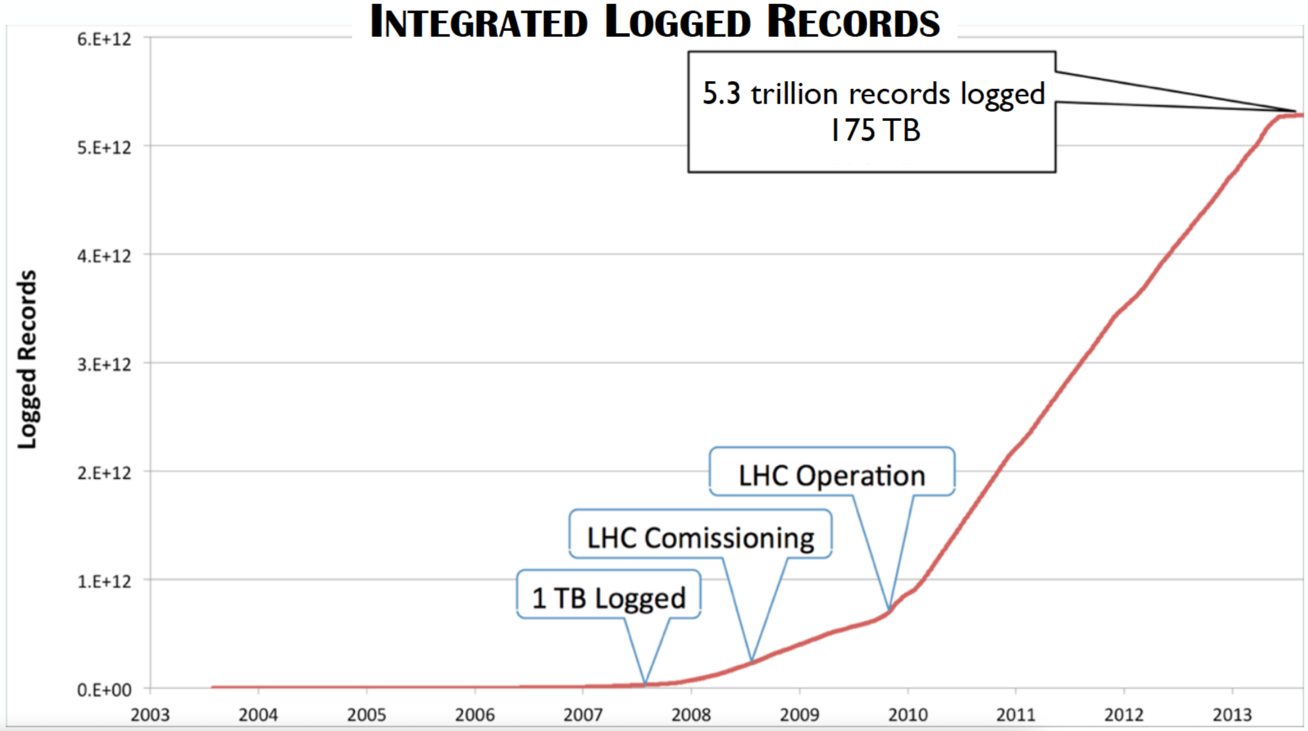
Data Set: XAFP.B7L1.B2.BOTTOM:TEMPERATURE X: 2012-10-30 02:19:09.626 Y: 31.74 Units: °C
Data Set: CURSOR X: 2012-10-30 02:18:34.028 Y: 3907.7 Units: GeV

Credit: C. Roderick

Accelerator logging

50TB/year, rate to increase to 100 – 150 TB in 2014

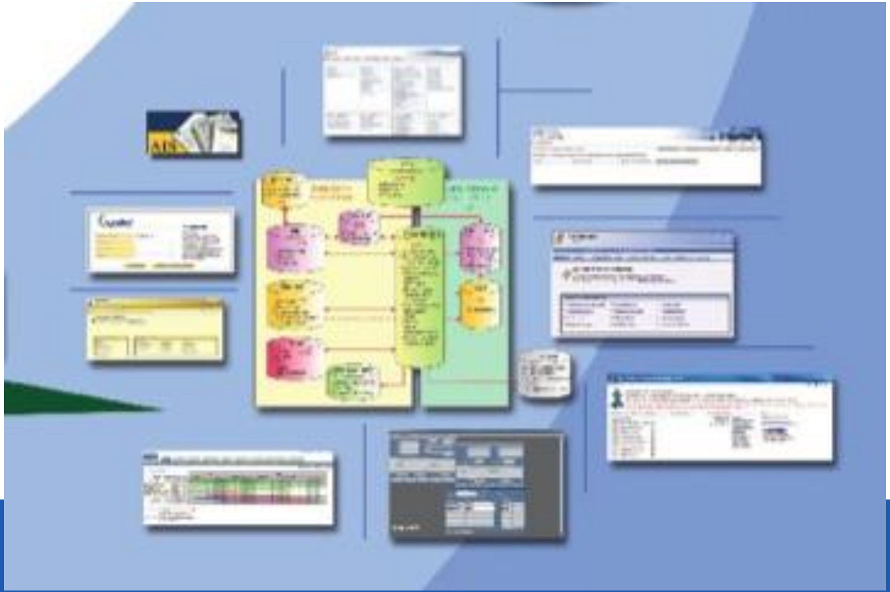
(Quench Protection System)



Credit: C. Roderick

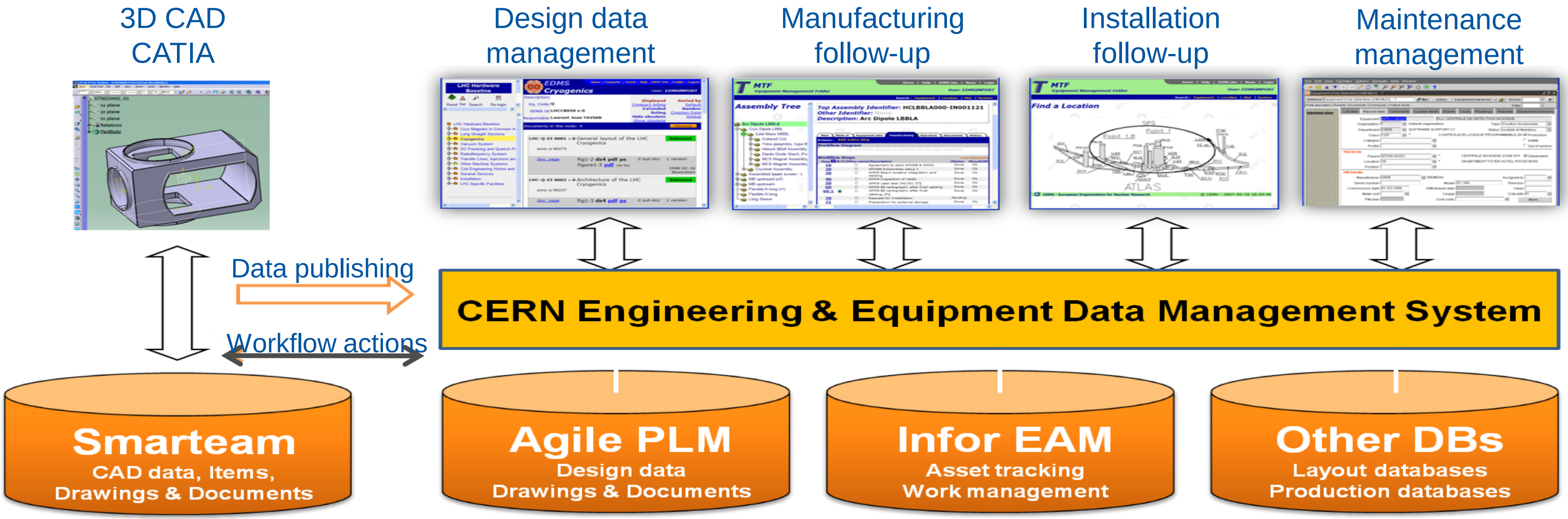
Administrative systems

- AIS has standardized on Oracle as database and uses it as interface between the tools
- Java EE and Apex, deployment with Weblogic
- Oracle E-Business HR



Engineering applications

- An integrated PLM platform based on commercial tools
- Simplified web interfaces for precise tasks



3D CAD
CATIA

Design data
management

Manufacturing
follow-up

Installation
follow-up

Maintenance
management

Data publishing

Workflow actions

CERN Engineering & Equipment Data Management System

Smarteam
CAD data, Items,
Drawings & Documents

Agile PLM
Design data
Drawings & Documents

Infor EAM
Asset tracking
Work management

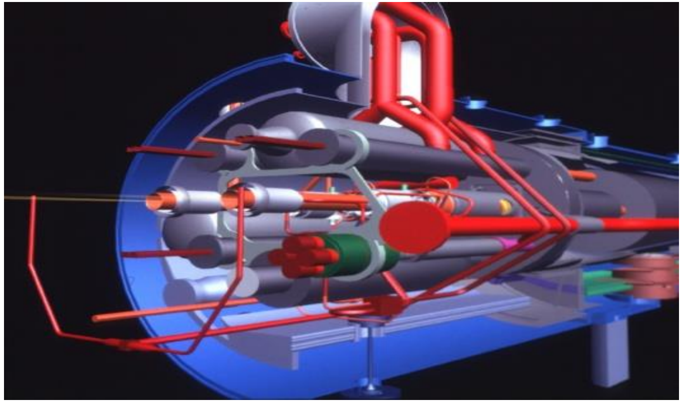
Other DBs
Layout databases
Production databases

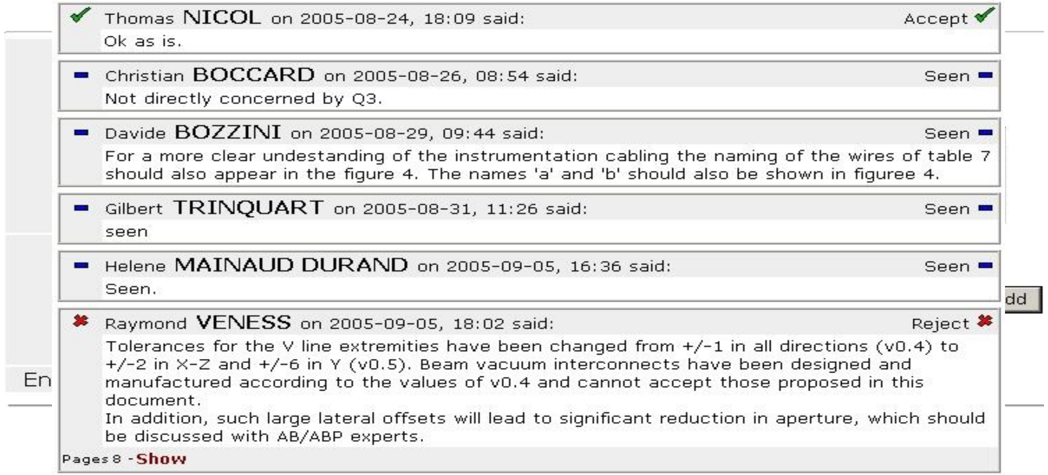
Credit: D. Widegren

Design data management

Design baseline with full configuration management

- Workflows, versioning rules and access control based on project dependent contexts
- Fully web-based and distributed approval processes





✓ Thomas NICOL on 2005-08-24, 18:09 said: Accept ✓
Ok as is.

▬ Christian BOCCARD on 2005-08-26, 08:54 said: Seen ▬
Not directly concerned by Q3.

▬ Davide BOZZINI on 2005-08-29, 09:44 said: Seen ▬
For a more clear undestanding of the instrumentation cabling the naming of the wires of table 7 should also appear in the figure 4. The names 'a' and 'b' should also be shown in figuree 4.

▬ Gilbert TRINQUART on 2005-08-31, 11:26 said: Seen ▬
seen

▬ Helene MAINAUD DURAND on 2005-09-05, 16:36 said: Seen ▬
Seen.

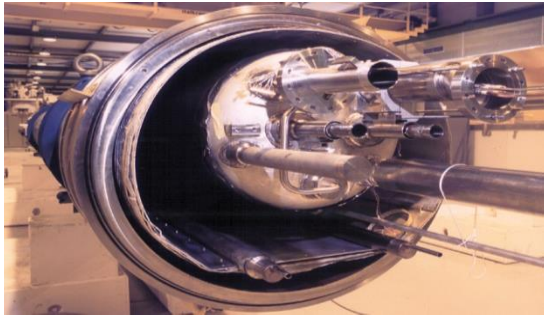
✖ Raymond VENESS on 2005-09-05, 18:02 said: Reject ✖
Tolerances for the V line extremities have been changed from +/-1 in all directions (v0.4) to +/-2 in X-Z and +/-6 in Y (v0.5). Beam vacuum interconnects have been designed and manufactured according to the values of v0.4 and cannot accept those proposed in this document.
In addition, such large lateral offsets will lead to significant reduction in aperture, which should be discussed with AB/ABP experts.

Credit: D. Widegren

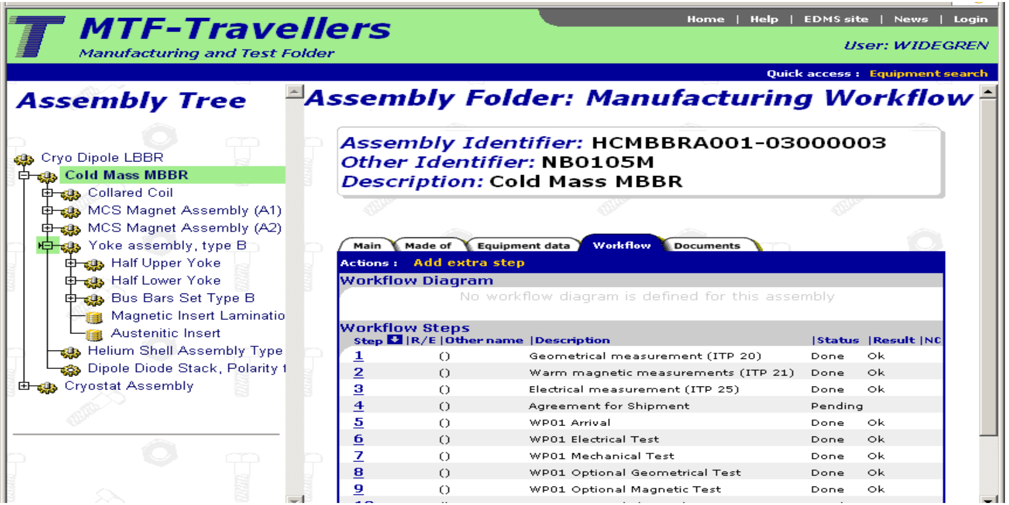
Manufacturing follow-up

Follow-up of each manufactured component

- Manufacturing & test data captured at manufacturing sites
- Predefined manufacturing workflows for each equipment type



HCMCDOC024-CG000718



MTF-Travellers
Manufacturing and Test Folder
Home | Help | EDMS site | News | Login
Quick access : Equipment search
Main
Made of
Equipment data
Workflow
Documents
Actions : Add extra step

User: WIDEGREN

Assembly Tree

- Cryo Dipole LBBR
  - Cold Mass MBBR
    - Collared Coil
    - MCS Magnet Assembly (A1)
    - MCS Magnet Assembly (A2)
    - Yoke assembly, type B
      - Half Upper Yoke
      - Half Lower Yoke
      - Bus Bars Set Type B
      - Magnetic Insert Laminatio
      - Austenitic Insert
    - Helium Shell Assembly Type
    - Dipole Diode Stack, Polarity t
  - Cryostat Assembly

Assembly Folder: Manufacturing Workflow

Assembly Identifier: HCMBBRA001-03000003
Other Identifier: NB0105M
Description: Cold Mass MBBR

Workflow Diagram

No workflow diagram is defined for this assembly

Workflow Steps

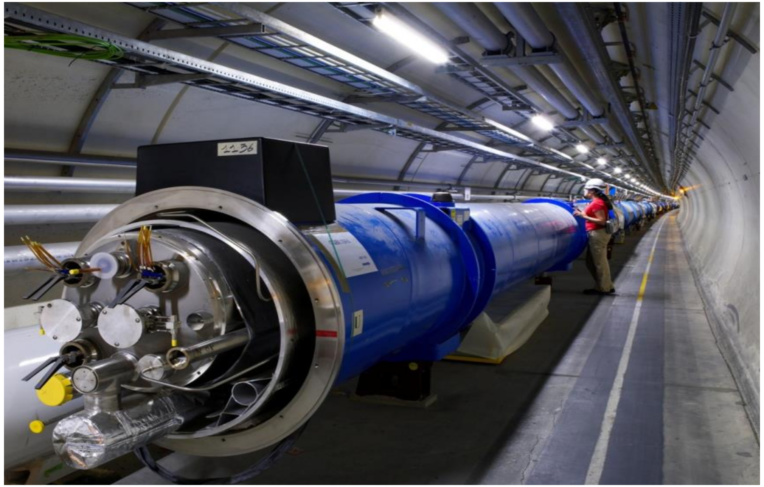
Step	R/E	Other name	Description	Status	Result	NC
1		()	Geometrical measurement (ITP 20)	Done	Ok	
2		()	Warm magnetic measurements (ITP 21)	Done	Ok	
3		()	Electrical measurement (ITP 25)	Done	Ok	
4		()	Agreement for Shipment	Pending		
5		()	WP01 Arrival	Done	Ok	
6		()	WP01 Electrical Test	Done	Ok	
7		()	WP01 Mechanical Test	Done	Ok	
8		()	WP01 Optional Geometrical Test	Done	Ok	
9		()	WP01 Optional Magnetic Test	Done	Ok	

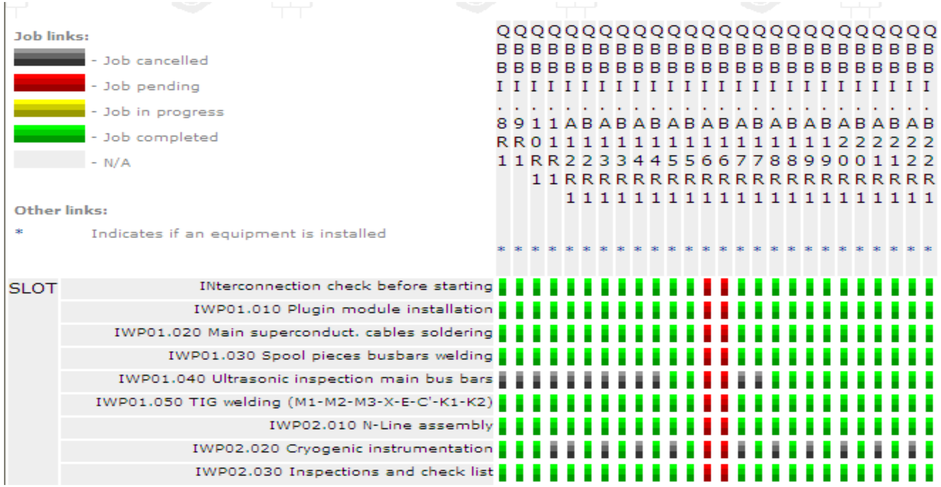
Credit: D. Widegren

Installation follow-up

Detailed logging of Installation & Commissioning tasks

- Over 150.000 jobs logged – allows detailed progress reporting
- Resolution of non-conformities with distributed approval processes.





Job links:

- ⬛ - Job cancelled
- 🟥 - Job pending
- 🟨 - Job in progress
- 🟩 - Job completed
- ⬜ - N/A

Other links:

* Indicates if an equipment is installed

SLOT	QBBI.8R1	QBBI.9R1	QBBI.10R1	QBBI.11R1	QBBI.A12R1	QBBI.B12R1	QBBI.A13R1	QBBI.B13R1	QBBI.A14R1	QBBI.B14R1	QBBI.A15R1	QBBI.B15R1	QBBI.A16R1	QBBI.B16R1	QBBI.A17R1	QBBI.B17R1	QBBI.A18R1	QBBI.B18R1	QBBI.A19R1	QBBI.B19R1	QBBI.A20R1	QBBI.B20R1	QBBI.A21R1	QBBI.B21R1	QBBI.A22R1	QBBI.B22R1
	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
INterconnection check before starting	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟥	🟥	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩
IWP01.010 Plugin module installation	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟥	🟥	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩
IWP01.020 Main superconduct. cables soldering	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟥	🟥	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩
IWP01.030 Spool pieces busbars welding	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟥	🟥	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩
IWP01.040 Ultrasonic inspection main bus bars	⬛	⬛	⬛	⬛	⬛	⬛	⬛	⬛	⬛	⬛	🟩	🟩	🟥	🟥	⬛	⬛	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩
IWP01.050 TIG welding (M1-M2-M3-X-E-C'-K1-K2)	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟥	🟥	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩
IWP02.010 N-Line assembly	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟥	🟥	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩
IWP02.020 Cryogenic instrumentation	🟩	🟩	🟩	⬛	⬛	🟩	⬛	🟩	⬛	🟩	⬛	🟩	🟥	🟥	⬛	🟩	⬛	🟩	⬛	🟩	⬛	🟩	⬛	🟩	⬛	🟩
IWP02.030 Inspections and check list	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟥	🟥	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩

Credit: D. Widegren

PLM @ CERN in numbers

Document & Drawings (incl. CAD):

~1,500.000 documents & drawings

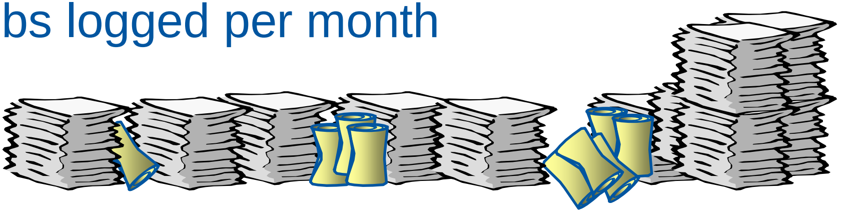
~7,000 new documents & drawings created per month

Components:

~1,300,000 registered individually followed equipment

~3,000,000 equipment interventions/jobs logged

~ 15,000 equipment interventions/jobs logged per month



Credit: D. Widegren

CASTOR and Oracle, tapes

- Home made mass storage system, relies on Oracle databases for name server, request handling and staging

- 4 libraries, SL8500
- 10088x4 = 40K slots (4500 free)
- Occupancy: 65PB worth of data
- Drives: 20 T10KB legacy drives; 40 T10KC drives (to be replaced by T10KD's)

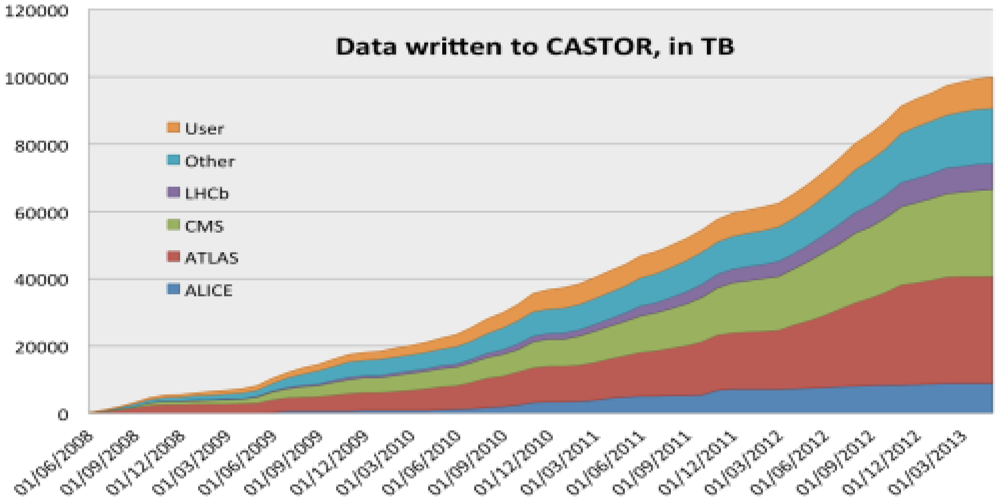
Credit: German Cancio Melia

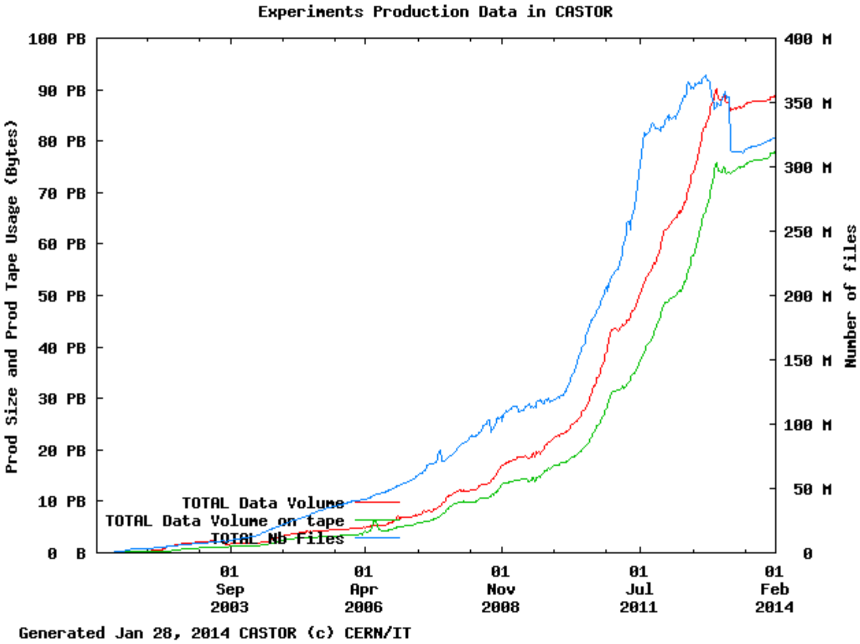
CASTOR Archive in Numbers

Credit: German Cancio Melia

Data:

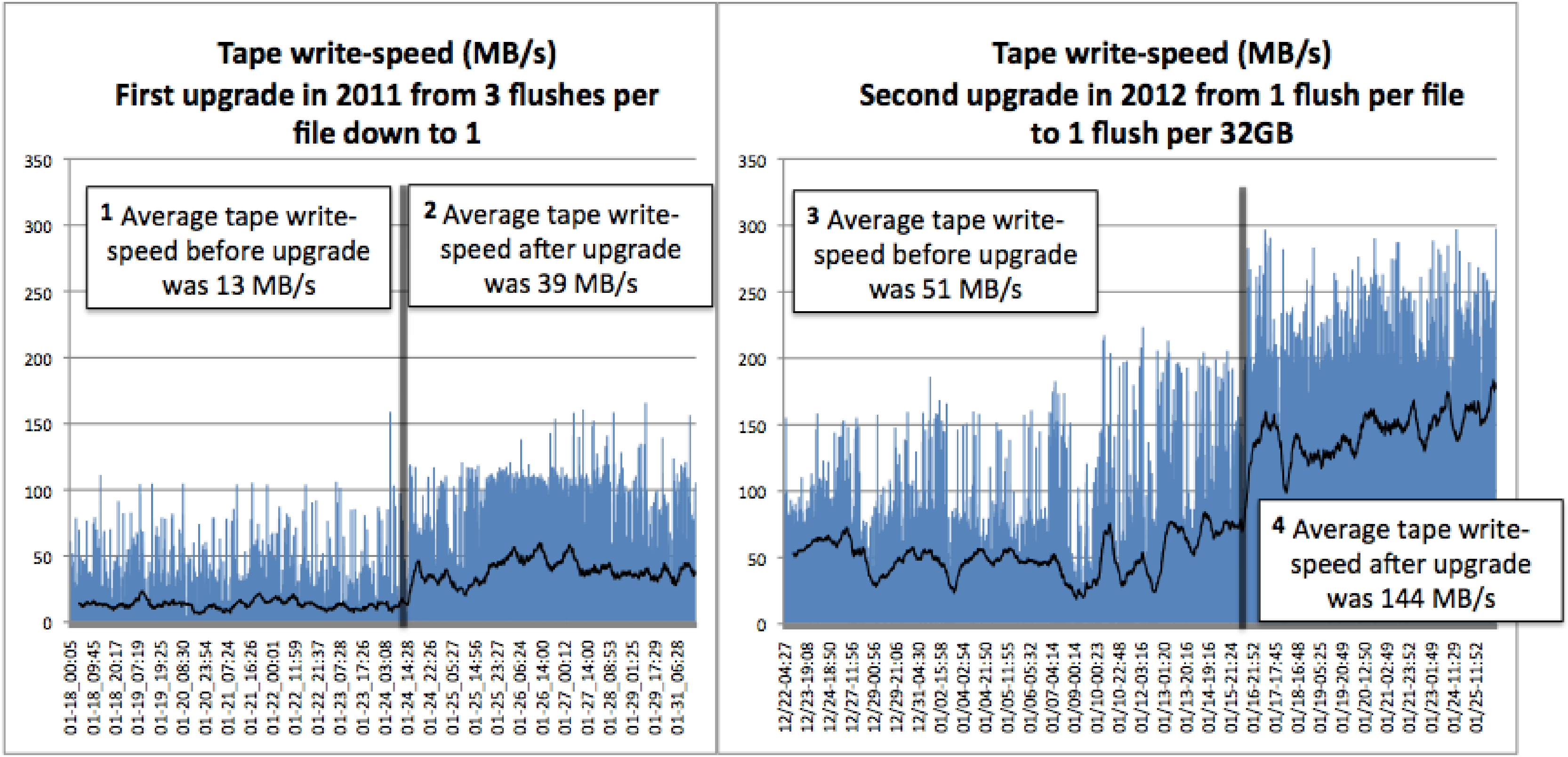
- ~90PB of data on tape; 250M files
- Up to 4.5 PB new data per month
- Over 10GB/s (R+W) peaks





Tape Writing

Credit: German Cancio Melia



Tape write-speed (MB/s)
First upgrade in 2011 from 3 flushes per file down to 1

1 Average tape write-speed before upgrade was 13 MB/s

2 Average tape write-speed after upgrade was 39 MB/s

Tape write-speed (MB/s)
Second upgrade in 2012 from 1 flush per file to 1 flush per 32GB

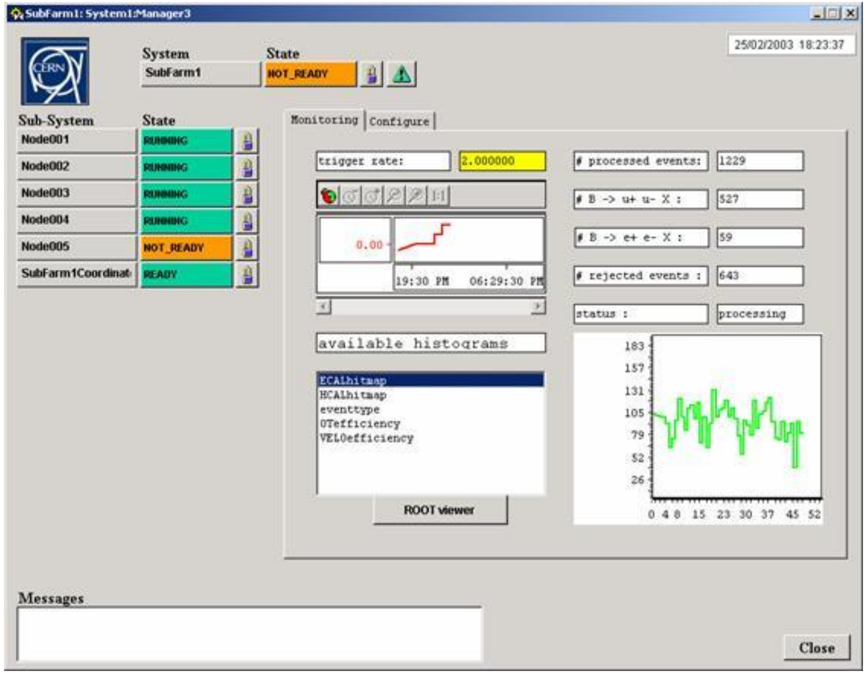
3 Average tape write-speed before upgrade was 51 MB/s

4 Average tape write-speed after upgrade was 144 MB/s

Experiment online systems

- Experiments rely on a SCADA system for their control

- Up to 150,000 changes / second stored in Oracle databases



SubFarm1: System1:Manager3

25/02/2003 18:23:37

System	State
SubFarm1	NOT_READY

Sub-System	State
Node001	RUNNING
Node002	RUNNING
Node003	RUNNING
Node004	RUNNING
Node005	NOT_READY
SubFarm1Coordinat	READY

Monitoring | Configure

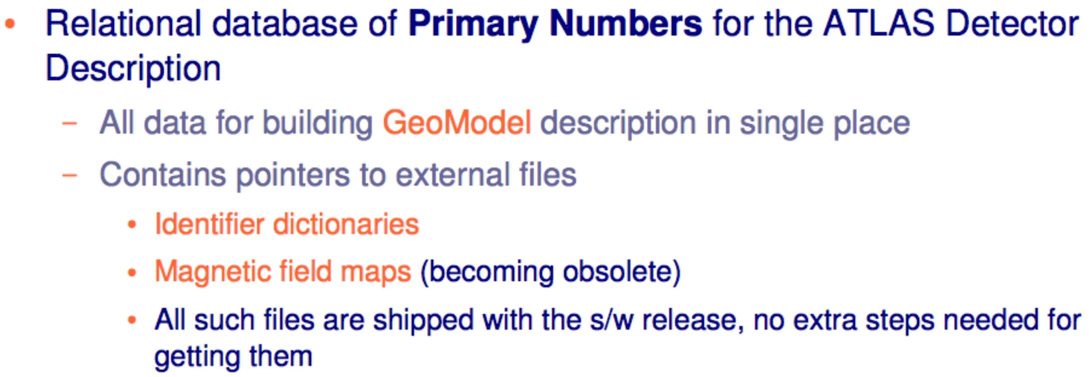
trigger rate:	2.000000
# processed events:	1229
# B -> u+ u- X :	527
# B -> e+ e- X :	59
# rejected events :	643
status :	processing

available histograms

ECALhitmap
HCALhitmap
eventtype
OTefficiency
VELOefficiency

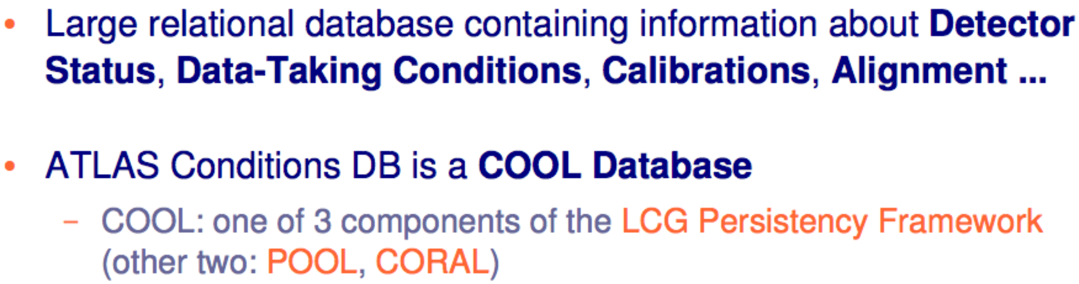
Experiment offline systems

- Geometry DB



- Relational database of **Primary Numbers** for the ATLAS Detector Description
  - All data for building GeoModel description in single place
  - Contains pointers to external files
    - Identifier dictionaries
    - Magnetic field maps (becoming obsolete)
    - All such files are shipped with the s/w release, no extra steps needed for getting them

- Conditions DB

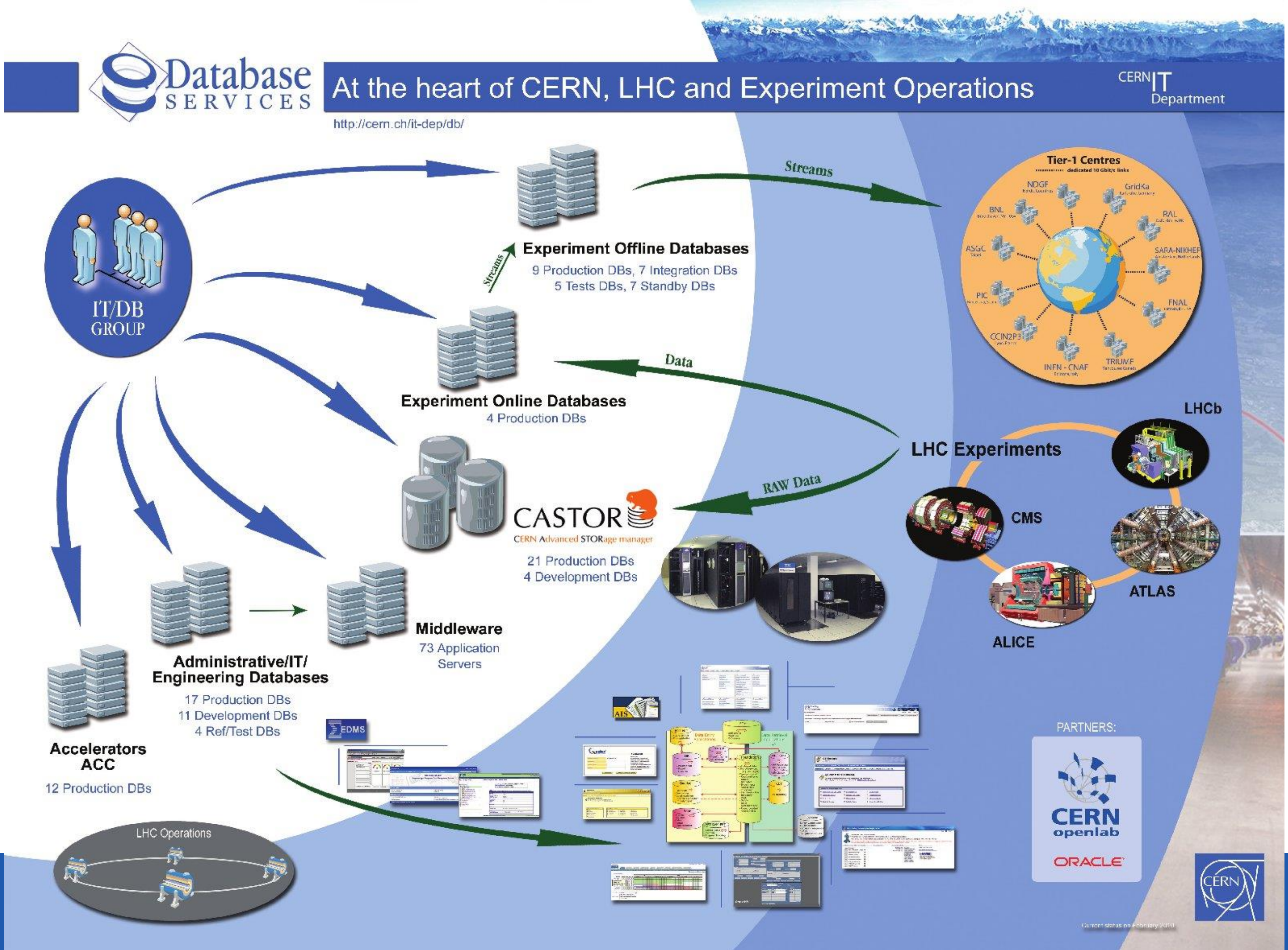


- Large relational database containing information about **Detector Status**, **Data-Taking Conditions**, **Calibrations**, **Alignment ...**
- ATLAS Conditions DB is a **COOL Database**
  - COOL: one of 3 components of the LCG Persistency Framework (other two: POOL, CORAL)

Credit: Vakho Tsulaia

Oracle at CERN

- From accelerator control to
  - accelerator logging,
  - administration,
  - engineering systems,
  - access control,
  - laboratory infrastructure (cabling, network configuration, etc.),
  - mass storage system,
  - experiment online systems,
  - experiment offline systems,
  - Etc.



Database SERVICES

At the heart of CERN, LHC and Experiment Operations

CERN IT Department

http://cern.ch/it-dep/db/

IT/DB GROUP

Experiment Offline Databases

9 Production DBs, 7 Integration DBs
5 Tests DBs, 7 Standby DBs

Experiment Online Databases

4 Production DBs

CASTOR
CERN Advanced STORage manager

21 Production DBs
4 Development DBs

Middleware

73 Application Servers

Administrative/IT/ Engineering Databases

17 Production DBs
11 Development DBs
4 Ref/Test DBs

Accelerators ACC

12 Production DBs

Tier-1 Centres

dedicated 10 Gbit/s links

LHC Experiments

PARTNERS:

CERN openlab

ORACLE

Current status on February 2010

Credit: M. Piorkowski

openlab (1/3)

- Public-private partnership between CERN and leading ICT companies, currently in fourth phase (started in 2003)
- Its mission is to accelerate the development of cutting-edge solutions to be used by the worldwide LHC community
- Innovative ideas aligned between CERN and the partners, for products “you make it, we break it”

Partners

hp

HUAWEI

intel

ORACLE

SIEMENS

openlab (2/3)

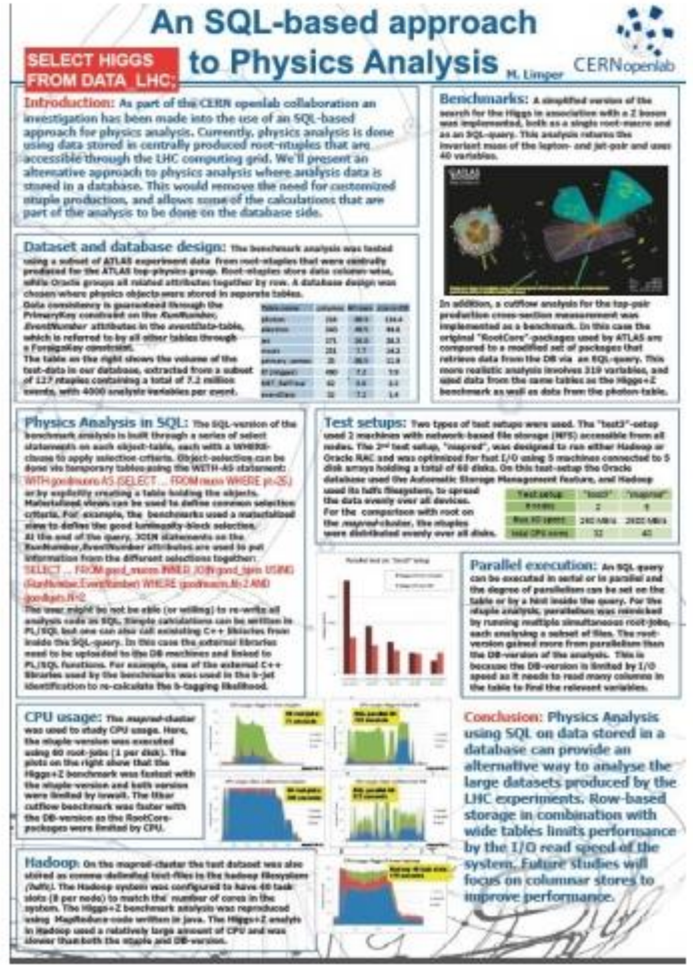
- Many successes:
  - RAC on Linux x86 (9.2 PoC and 10.1 production with ASM),
  - Additional required functionality (IEEE numbers, OCCI, instant client, etc.),
  - PVSS and RAC scalability,
  - Monitoring with Grid Control,
  - Streams world wide distribution,
  - Active DG, GoldenGate,
  - Analytics for accelerator, experiment and IT,
  - Etc.

- Regular feedback with joint selection of topics, some of the projects are common with more than one partner

openlab (3/3)

- Publications (web, paper) and presentations of results, visitors

- Maaike Limper, best poster award at The International Conference on Computing in High Energy and Nuclear Physics 2013



SELECT HIGGS FROM DATA_LHC;

An SQL-based approach to Physics Analysis

M. Limper

CERNopenlab

Introduction: As part of the CERN openlab collaboration an investigation has been made into the use of an SQL-based approach for physics analysis. Currently, physics analysis is done using data stored in centrally produced root-ntuples that are accessible through the LHC computing grid. We'll present an alternative approach to physics analysis where analysis data is stored in a database. This would remove the need for customized ntuple production, and allows some of the calculations that are part of the analysis to be done on the database side.

Benchmarks: A simplified version of the search for the Higgs in association with a Z boson was implemented, both as a single root-macro and as an SQL-query. This analysis returns the invariant mass of the lepton- and jet-pair and uses 40 variables.

In addition, a cutflow analysis for the top-pair production cross-section measurement was implemented as a benchmark. In this case the original "RootCore"-packages used by ATLAS are compared to a modified set of packages that retrieve data from the DB via an SQL-query. This more realistic analysis involves 319 variables, and used data from the same tables as the Higgs+Z benchmark as well as data from the photon-table.

Dataset and database design: The benchmark analysis was tested using a subset of ATLAS experiment data from root-ntuples that were centrally produced for the ATLAS top-physics group. Root-ntuples store data column-wise, while Oracle groups all related attributes together by row. A database design was chosen where physics objects were stored in separate tables.

Data consistency is guaranteed through the PrimaryKey constraint on the *RunNumber, EventNumber* attributes in the *eventData*-table, which is referred to by all other tables through a ForeignKey constraint.

The table on the right shows the volume of the test-data in our database, extracted from a subset of 127 ntuples containing a total of 7.2 million events, with 4000 analysis variables per event.

[illegible]	[illegible]	[illegible]	[illegible]
photon	216	80.5	114.4
electron	340	46.5	84.8
jet	171	26.8	38.3
muon	231	7.7	24.2
primary vertex	25	86.5	11.8
EF (trigger)	490	7.2	7.9
MET_RefFinal	62	3.6	1.1
eventData	52	7.2	1.4

Physics Analysis in SQL: The SQL-version of the benchmark analysis is built through a series of select statements on each object-table, each with a WHERE-clause to apply selection criteria. Object-selection can be done via temporary tables using the WITH-AS statement:

WITH goodmuons AS (SELECT … FROM muon WHERE pt>25)

or by explicitly creating a table holding the objects. Materialized views can be used to define common selection criteria. For example, the benchmarks used a materialized view to define the good luminosity-block selection.

At the end of the query, JOIN statements on the RunNumber,EventNumber attributes are used to put information from the different selections together:

SELECT … FROM good_muons INNER JOIN good_jets USING (RunNumber,EventNumber) WHERE goodmuons.N=2 AND goodjets.N=2

The user might be not be able (or willing) to re-write all analysis code in SQL. Simple calculations can be written in PL/SQL but one can also call existing C++ libraries from inside the SQL-query. In this case the external libraries need to be uploaded to the DB machines and linked to PL/SQL functions. For example, one of the external C++ libraries used by the benchmarks was used in the b-jet identification to re-calculate the b-tagging likelihood.

Test setups: Two types of test setups were used. The "test3"-setup used 2 machines with network-based file storage (NFS) accessible from all nodes. The 2nd test setup, "mapred", was designed to run either Hadoop or Oracle RAC and was optimized for fast I/O using 5 machines connected to 5 disk arrays holding a total of 60 disks. On this test-setup the Oracle database used the Automatic Storage Management feature, and Hadoop used its hdfs filesystem, to spread the data evenly over all devices.

For the comparison with root on the *mapred*-cluster, the ntuples were distributed evenly over all disks.

Test setup	"test3"	"mapred"
# nodes	2	5
Max IO speed	240 MB/s	2500 MB/s
total CPU cores	32	40

Parallel execution: An SQL query can be executed in serial or in parallel and the degree of parallelism can be set on the table or by a hint inside the query. For the ntuple analysis, parallelism was mimicked by running multiple simultaneous root-jobs, each analysing a subset of files. The root-version gained more from parallelism than the DB-version of the analysis. This is because the DB-version is limited by I/O speed as it needs to read many columns in the table to find the relevant variables.

CPU usage: The *mapred*-cluster was used to study CPU usage. Here, the ntuple-version was executed using 60 root-jobs (1 per disk). The plots on the right show that the Higgs+Z benchmark was fastest with the ntuple-version and both version were limited by iowait. The ttbar cutflow benchmark was faster with the DB-version as the RootCore-packages were limited by CPU.

Hadoop: On the mapred-cluster the test dataset was also stored as comma-delimited text-files in the hadoop filesystem (hdfs). The Hadoop system was configured to have 40 task slots (8 per node) to match the number of cores in the system. The Higgs+Z benchmark analysis was reproduced using MapReduce-code written in java. The Higgs+Z analysis in Hadoop used a relatively large amount of CPU and was slower than both the ntuple and DB-version.

Conclusion: Physics Analysis using SQL on data stored in a database can provide an alternative way to analyse the large datasets produced by the LHC experiments. Row-based storage in combination with wide tables limits performance by the I/O read speed of the system. Future studies will focus on columnar stores to improve performance.

Oracle in our research environment

- Even if computing is critical for HEP, it is not the goal, there is a lot to do using solutions from commercial vendors which are industry supported and scalable
- Oracle has provided solutions along the years
- We have worked with Oracle to improve the tools to our (and others') needs with success
- Good for staff to work on industry standards for their future career

Conclusion

- Not every day you build a 30+ years collaboration
- A long way since 1982, now very wide usage with applications, tape and database
- Oracle has proven to be reliable partner who cares and supports research
- Provide feedback and ideas for enhancements
- Helps focus on our core challenges
- A collaboration which works!

CERN

www.cern.ch