LCG Project Status

LHCC Open session
CERN – 12 October 2005

Frédéric Hemmer on behalf of
Les Robertson – LCG Project Leader

LHC Computing Grid Project - a Collaboration

Developing & supporting the applications development environment, common tools and frameworks

Building and operating the LHC Grid

A global collaboration involving –

The physicists and computing specialists from the LHC experiments

- The national and regional projects in Europe and the US that have been developing Grid middleware

- The regional and national computing centres that provide resources for LHC

- The research networks

Researchers

Computer Scientists & Software Engineers
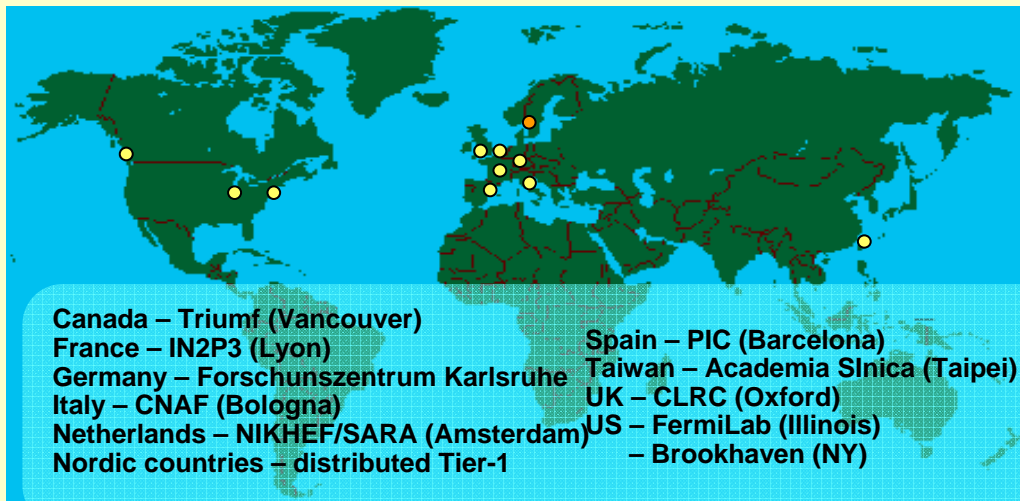
Service Providers

LCG Service Hierarchy

Tier-0 – the accelerator centre

- Data acquisition & initial processing
- Long-term data curation
- Distribution of data → Tier-1 centres

Tier-1 – "online" to the data acquisition process → high availability

- Managed Mass Storage –
  → grid-enabled data service
- Data-heavy analysis
- National, regional support

Canada – Triumf (Vancouver)
France – IN2P3 (Lyon)
Germany – Forschunszentrum Karlsruhe
Italy – CNAF (Bologna)
Netherlands – NIKHEF/SARA (Amsterdam)
Nordic countries – distributed Tier-1

Spain – PIC (Barcelona)
Taiwan – Academia SInica (Taipei)
UK – CLRC (Oxford)
US – FermiLab (Illinois)
– Brookhaven (NY)

Tier-2 –   ~100 centres in 30+ countries

- Simulation
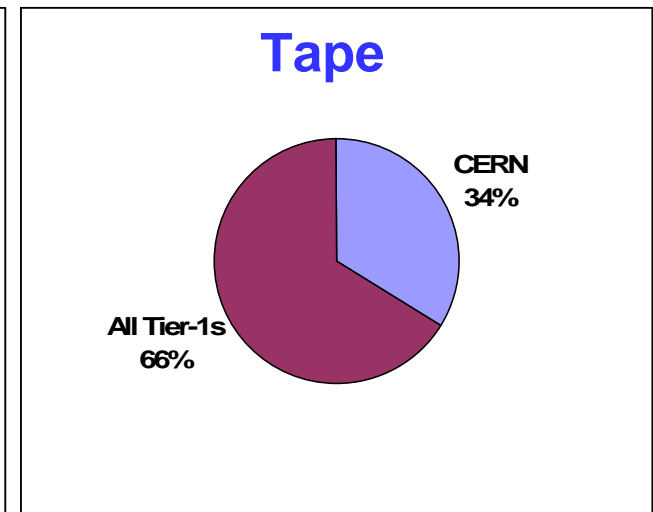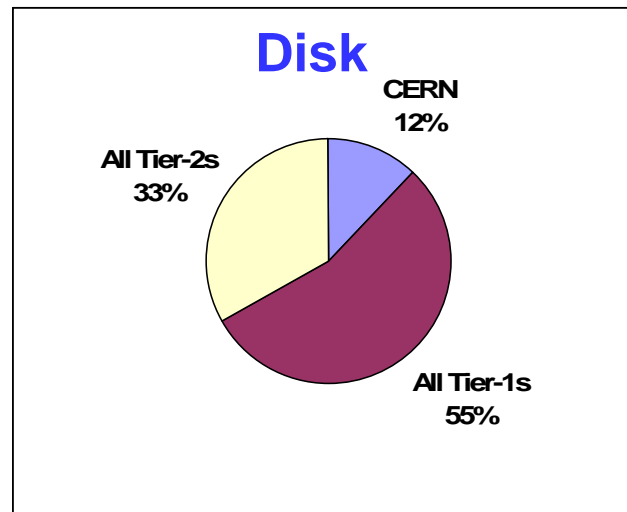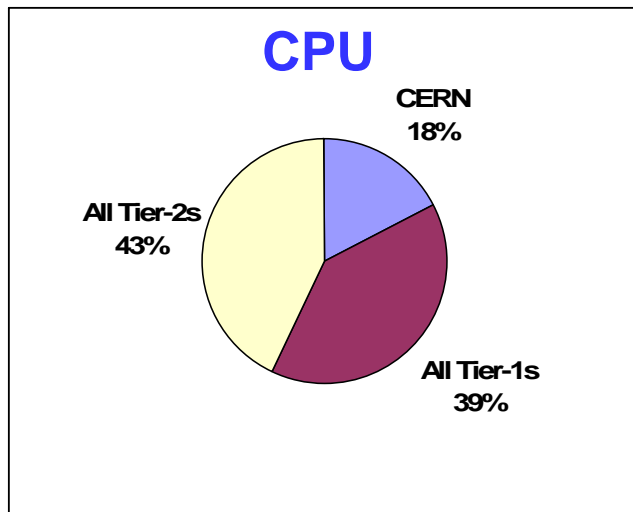- End-user analysis – batch and interactive

LCG

> 100K of today's fastest processors
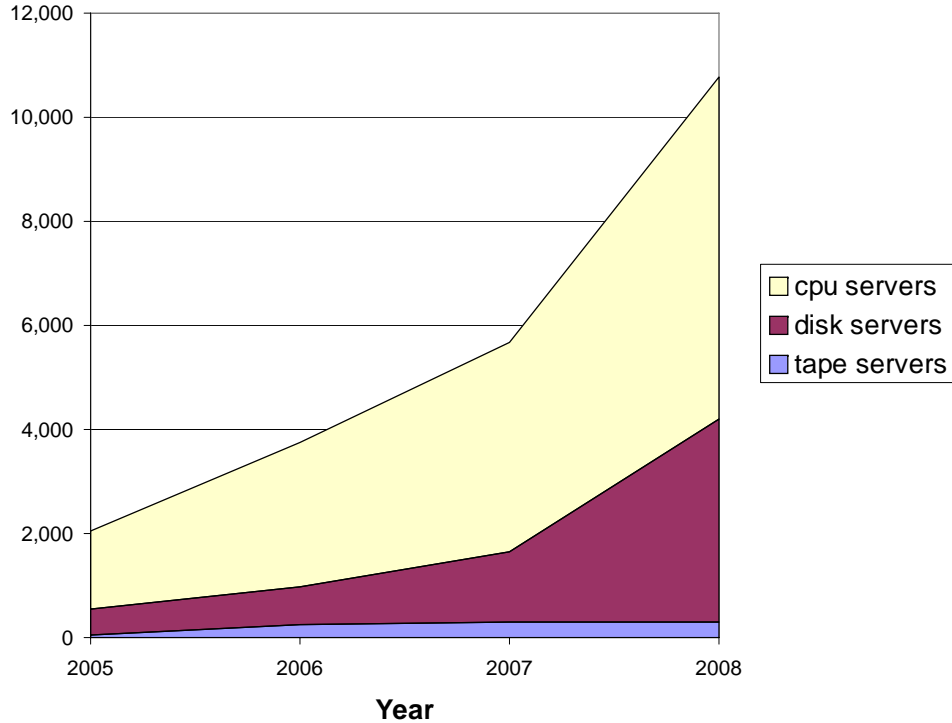
Summary of Computing Resource Requirements

All experiments - 2008

From LCG TDR - June 2005

	CERN	*All Tier-1s*	*All Tier-2s*	*Total*
CPU (MSPECint2000s)	25	56	61	142
Disk (PetaBytes)	7	31	19	57
Tape (PetaBytes)	18	35		53

CPU

CERN 18%
All Tier-2s 43%
All Tier-1s 39%

Disk

CERN 12%
All Tier-2s 33%
All Tier-1s 55%

Tape

CERN 34%
All Tier-1s 66%

How big is this?

LCG

Growth in Boxes in CERN Computer Centre

Legend:
- cpu servers
- disk servers
- tape servers

Evolution of Capacity at CERN

Legend:
- Processing Capacity (MSI2K)
- Disk Capacity (PB)
- Tape Capacity (PB)
- Tape Drives
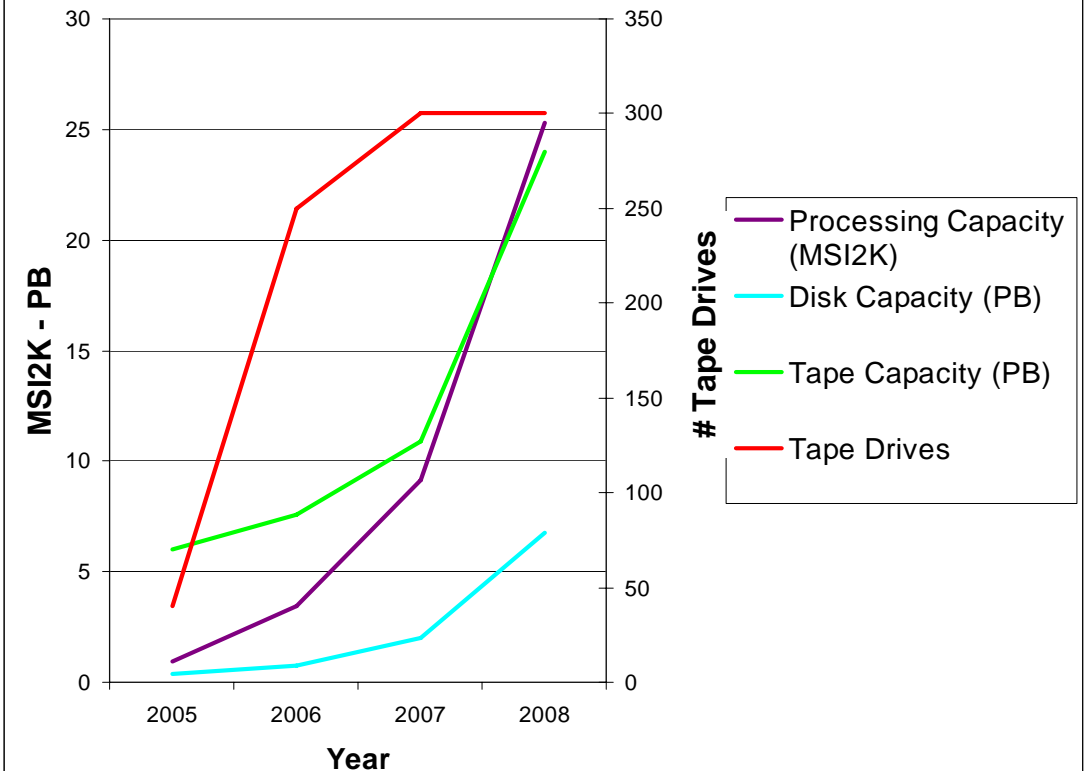
At CERN by 2008:
- 5-fold increase in number of servers to be managed
- 7-fold increase in number of tape drives

Similar growth rates at other sites

last update 26/01/2006 12:48

CERN Computer Center Preparations

CERN Computer Centre Management

Capacity Growth from now to 2008

LCG

Tier-1 Evolution
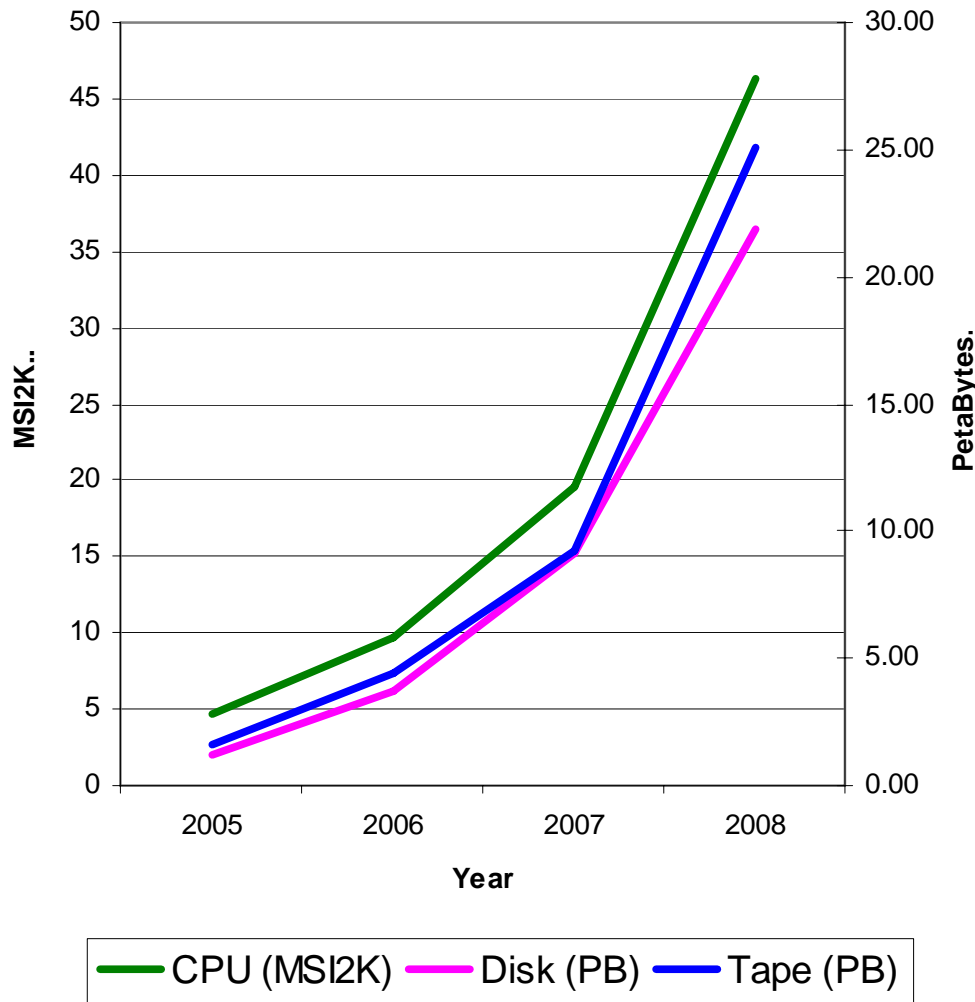
Tier-2 Growth

Tier-2s that have provided planning data sep05
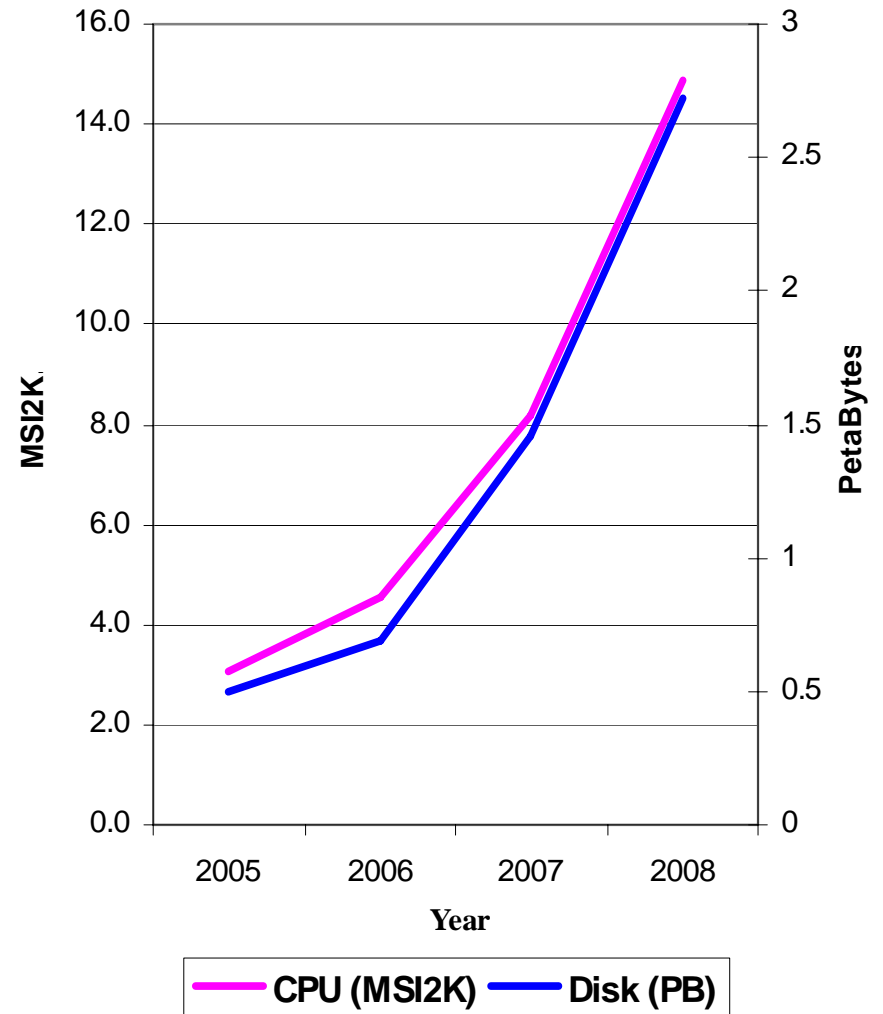
Tier-1s

Tier-1 Centre	Primary Experiments			
	ALICE	ATLAS	CMS	LHCb
TRIUMF, Canada		X		
GridKA, Germany	X	X	X	X
CC, IN2P3, France	X	X	X	X
CNAF, Italy	X	X	X	X
SARA/NIKHEF, NL	X	X		X
Nordic Data Grid Facility	X	X	X	
ASCC, Taipei		X	X	
RAL, UK	X	X	X	X
BNL, US		X		
FNAL, US			X	
PIC, Spain		X	X	X

-- 70% of sites support >1 experiment
-- almost all sites support other (non-LHC) communities

Preliminary Tier-1 Resource Planning Capacity at all Tier-1s in 2008

Update expected at October C-RRB

Tier-1 Planning for 2008			ALICE	ATLAS	CMS	LHCb	SUM 2008
CPU - MSI2K		Offered	6.7	22.7	12.5	4.4	46.3
		TDR Requirements	12.3	24.0	15.2	4.4	55.9
		Balance	**-46%**	**-5%**	**-18%**	**-0%**	**-17%**
Disk - PBytes		Offered	2.8	12.5	5.7	2.2	23.2
		TDR Requirements	7.4	14.4	7.0	2.4	31.2
		Balance	**-62%**	**-13%**	**-18%**	**-10%**	**-25%**
Tape - PBytes		Offered	3.2	9.1	8.1	1.9	22.3
		TDR Requirements	6.9	9.0	16.7	2.1	34.7
		Balance	**-54%**	**1%**	**-51%**	**-9%**	**-36%**

Includes current planning for all Tier-1 centres

Tier-2s
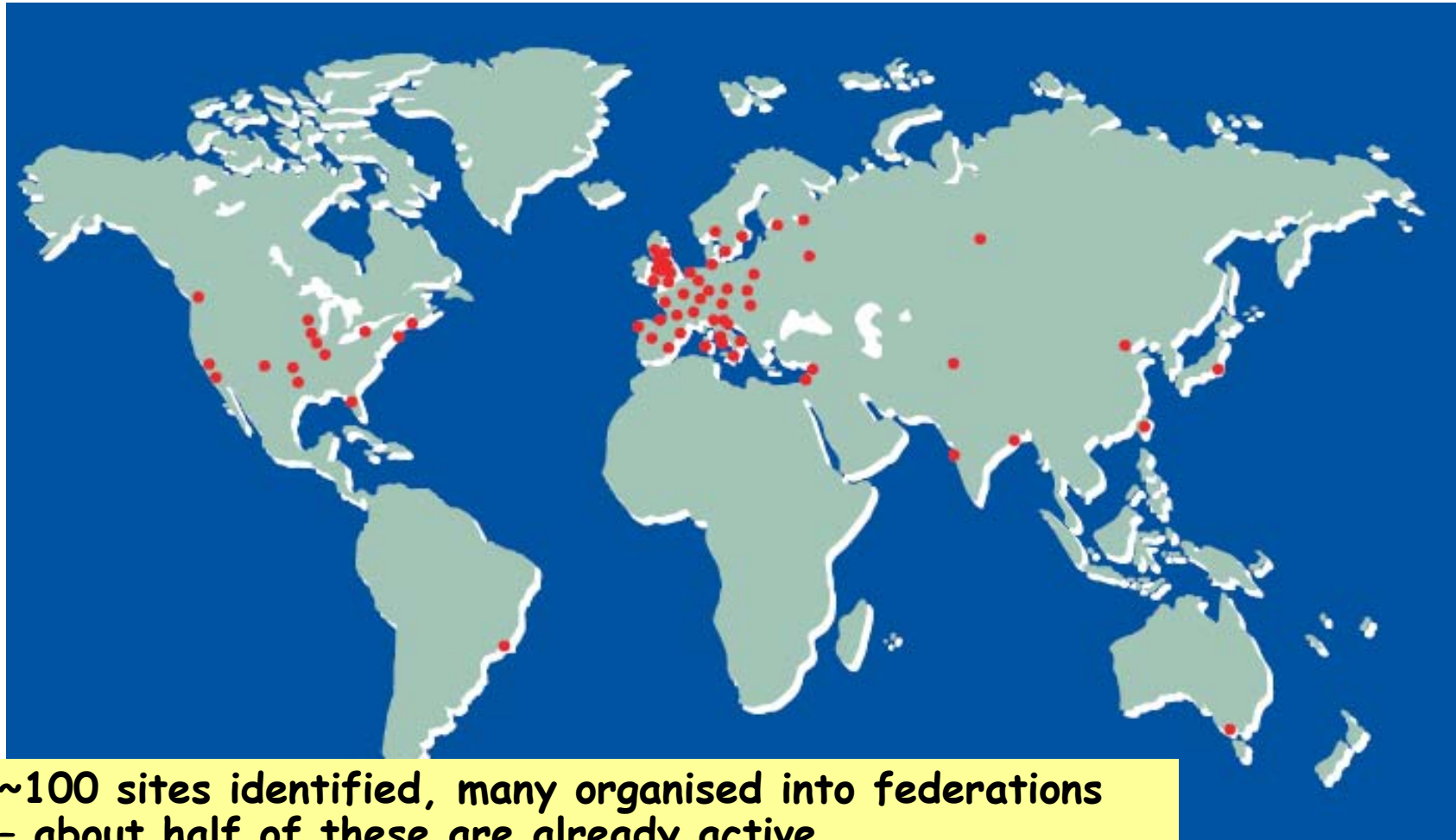
~100 sites identified, many organised into federations
– about half of these are already active

Preliminary Tier-2 Resource Planning Capacity at all Tier-2s in 2008
Update expected at October C-RRB

Tier-2 Planning for 2008			ALICE	ATLAS	CMS	LHCb	SUM 2008
CPU -	MSI2K	Offered	5.0	19.8	17.4	4.4	46.6
		TDR Requirements	14.4	19.9	19.3	7.7	61.3
		Balance	**-65%**	**-1%**	**-10%**	**-42%**	**-24%**
Disk -	PBytes	Offered	1.4	6.2	4.5	0.8	12.9
		TDR Requirements	3.5	8.7	4.9	0.023	17.1
		Balance	**-59%**	**-29%**	**-8%**	**n/a**	**-24%**
# Tier-2 federations - included(expected)			12 (13)	21 (28)	17 (19)	11 (12)	29 (37)

- 29 Tier2 centres have their data included in above table.
- 8 more centres plan to join as soon as possible.

Ramping up to the LHC Service

- The services for Phase 2 will be ramped-up through two **Service Challenges** SC3 this year and SC4 next year

- These will include CERN, the Tier-1s and the major Tier-2s

- Each service Challenge includes –

  -- a set-up period

    - check out the infrastructure/service to iron out the problems *before the experiments get fully involved*

    - schedule allows *time to provide permanent fixes* for problems encountered

    - A throughput test

  -- followed by a long stable period for experiments to check out their computing model and software chain

Service Challenge 3

- ~15 centres active in setup phase
- First time that mass storage and tape involved
- Reliability and performance issues (being tackled)
- Questions of readiness of some centres to sustain service for extended period (staffing issues?)
- Need more detailed joint planning, more visibility of site planning
- Service phase started September – now 20 sites
- Task Forces starting up now → integrate experiment, LCG and EGEE effort in getting experiments working on SC3 (ALICE – Carminati; ATLAS – Perini; CMS – Belforte; LHCb - Tsaragorodtsev)

Analysis of GridFTP Log for Service Challenge 3

Hourly Averaged Throughput on 17-07-2005
From CERNCI to ALL SITES

Legend: BNL, CNAF, DESY, FZK, IN2PCC, NDGF, OTHERS, PIC, RAL, SARA, TRIUMF
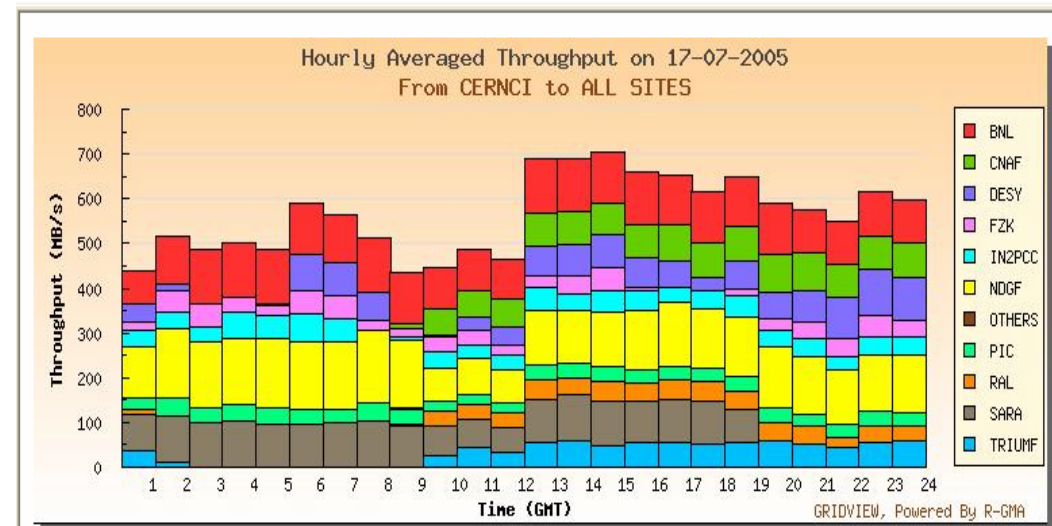
(OTHERS: Sites giving throughput less than 5% of max, viz. BUDPST, INFCNA, LIP, TAIWAN, UIBK, UNREGD,)

GRIDVIEW, Powered By R-GMA

ATLAS Rome Simulation on EGEE

Number of jobs/day

~10,000 concurrent jobs in the system
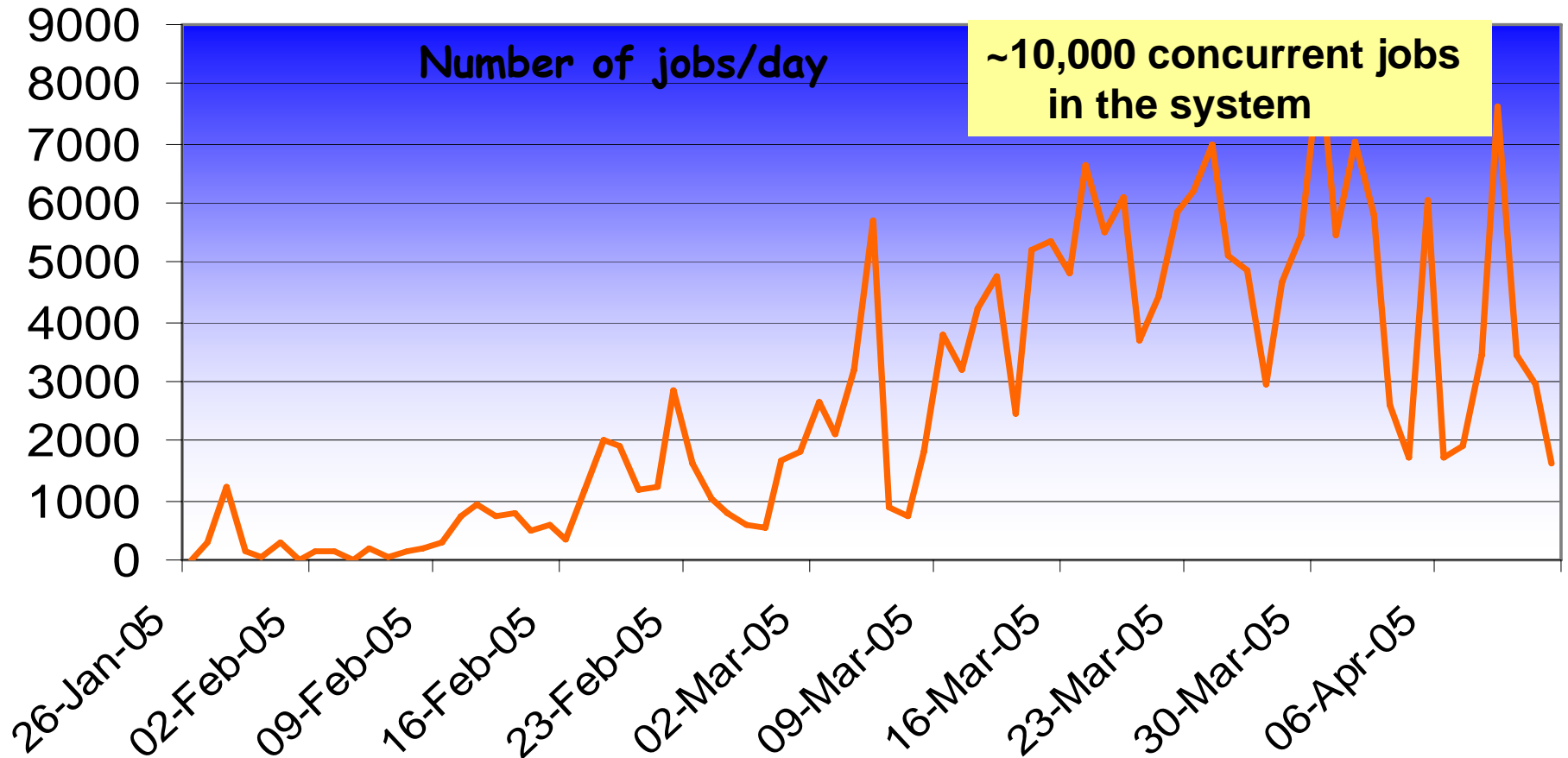
- In latest period up to 8K jobs/day
- Several times the current capacity for ATLAS at CERN alone
- Shows the reality of the grid solution
  -- but still far to go in reliability

Integration LCG and EGEE

- **Goal**
  Create a European-wide production quality multi-science grid infrastructure on top of national & regional grid programs

- **Scale**
  70 partners in 27 countries
  Initial funding (€32M) for 2 years
  -- proposal for a 2-year extension

- **Activities**
  Grid operations and support (joint LCG/EGEE operations team - includes resources at many LHC sites)
  Middleware re-engineering (close attention to LHC data analysis requirements)
  Training, support for applications groups (inc. contribution to the ARDA activity)

- **Builds on**
  LCG grid deployment
  Experience gained in HEP
  LHC experiments → pilot applications

EGEE Grid Status

October 2005

- running at ~13,000 jobs in the system
- many more sites and processors than we anticipated at this stage
  → ~150 sites, ~15,000 processors
  scalability is *already close to that needed for the full LHC grid*
- Grid operation now working, sharing responsibility between operations centres in *Taiwan, Italy, Germany, France, UK, Russia and CERN*
- Funds significant resources for operations at many LCG sites, including CERN
- Reliability is still an issue – major improvements in the past year
- Middleware evolution → aim for a solid, though basic functionality by end 2005

34 countries working together in a consensus based organisation

LCG

NORDUGRID
Grid Solution for Wide Area
Computing and Data Handling

30 sites
3200 cpus

egee
Enabling Grids
for E-sciencE

Inter-operation EGEE, Open Science Grid in the US and NorduGrid:
→ Very early days for standards – still getting basic experience
→ Focus on baseline services to meet specific experiment requirements

September 2005

150 Grid sites

34 countries

15,000 CPUs

2800 CPUs

Open Science Grid

South Korea

Baseline services

LCG

- Storage management services
  - Based on SRM as the interface
- Basic transfer services
  - gridFTP, srmCopy
- Reliable file transfer service
- Grid catalogue services
- Catalogue and data management tools
- Database services
- Compute Resource Services
- Workload management

- VO management services
  - "VOMS" system with support for user roles, groups, subgroups
- POSIX-like I/O service
  - local files, and include links to catalogues
- Grid monitoring tools and services
- VO agent framework
- Applications software installation service
- Reliable messaging service
- Information system

- Each site agrees to implement baseline services appropriate for the "Tier"
- Depending on the service - agree on standard interface/protocol, choice of implementations, or specify specific implementation,

last update 26/01/2006 12:48

Summary – Grids and Technology

Early days –

- Middleware: basic functionality, still with some reliability and performance issues

- Outside HEP there are few grids in operation, and none on this scale

- Multiple grid infrastructures bring additional complications
  - but the grid projects bring important resources for operation
  - and may in the longer term enable HEP to benefit from general science infrastructure

- The Worldwide LCG Collaboration is straddling Grid collaborations and experiment collaborations –
  - each with its own goals, priorities
    and need for visibility and differentiation

- With data taking in sight –
    - we must keep things simple
    - realistic (modest) expectations
    - willingness to compromise

Wide Area Networking

- Working group set up by the Grid Deployment Board to bring together the Tier-1 centres, the national and regional research networks to plan for LHC

- On-going work
  - Overall architecture and implementation schedule evolving
  - Most Tier-1 centres should be connected with light paths by end 2005

- Active participation of NRENs
  - Abilene/ESNET/LHCnet (USA), ASnet (TW), Canarie (CDN), DFN (D), GARR (I), NorduGrid, RedIris (E), RENATER (F), SURFNET (NL), UKERNA,

  and the European backbone network GEANT

Service Challenge 3 – T0-T1 Network

CERN AS513 - 128.142.224.0/20

CERN Geneva

Foundry switch

cernh7

4x1Gb

b513-c-rca86-4

cernh5

SURFnet ONS @CERN

6x1Gb

SURFnet ONS @AMS

2x10Gb

2x1Gb

2x1Gb

2x1Gb

Netherlight — SARA 145.100.3.80/28

UKlight — RAL 193.63.181.128/25

ASnet — ASCC AS9264 - 140.109.248.0/28

Canarie — TRIUMF 206.12.1.0/24

ar5-chicago @STARLIGHT

ESnet

BNL AS
130.199.185.0.0/24
130.199.48.0/23

FNAL AS 3152
131.225.206.0/23

Geant

Renater — IN2P3 AS789 – 193.48.99.64/26

DFN — FZK AS680 – 192.108.45.0/24 192.108.46.0/24
DESY AS1754 – 131.169.5.0/24 131.169.80.0/24 131.169.98.0/24 131.169.223.0/24

GARR — CNAF AS137 - 192.135.23.0/24

2.5Gb

RedIRIS — PIC AS766 - 193.146.196.0/22

2.5Gb

NORDUnet — NORDUgrid 130.226.158.112/28 130.236.102.0/24 192.36.236.0/24

— 10G link
— 1G link

edoardo.martelli@cern.ch 20050722

last update 20/01/2006 12:48

GÉANT 2 research network backbone

strong correlation with European Tier-1s

Initial Backbone Topology

GÉANT2 is operated by DANTE on behalf of Europe's NRENs.

last update 26/01/2006 12:48

Wide Area Network

T2s and T1s are inter-connected by the general purpose research networks

Dedicated 10 Gbit links

Any Tier-2 may access data at any Tier-1

IN2P3

GridKa

TRIUMF

Brookhaven

ASCC

Nordic

Fermilab

CNAF

RAL

SARA

PIC

CERN

LCG

T2 (repeated around the diagram)
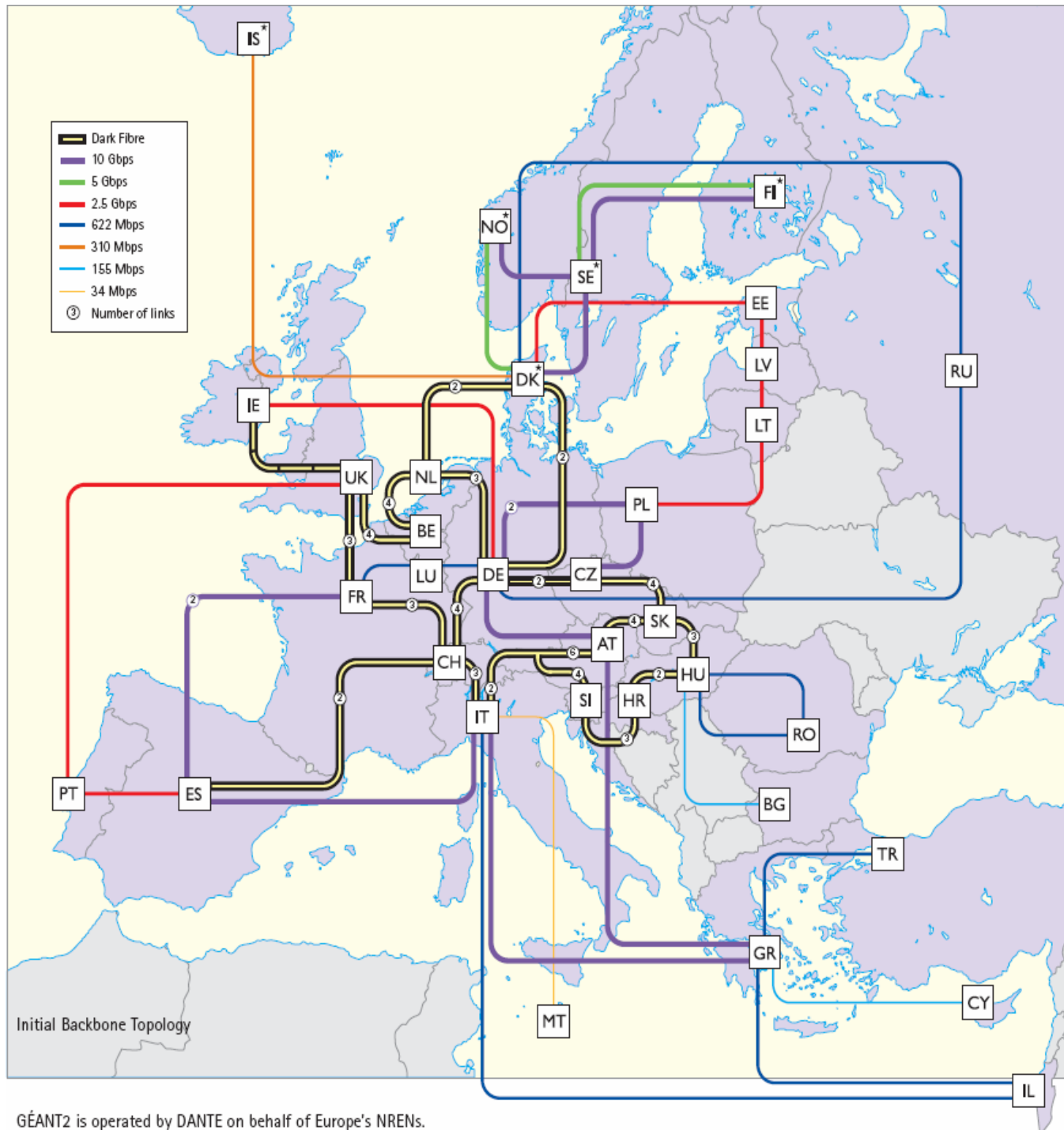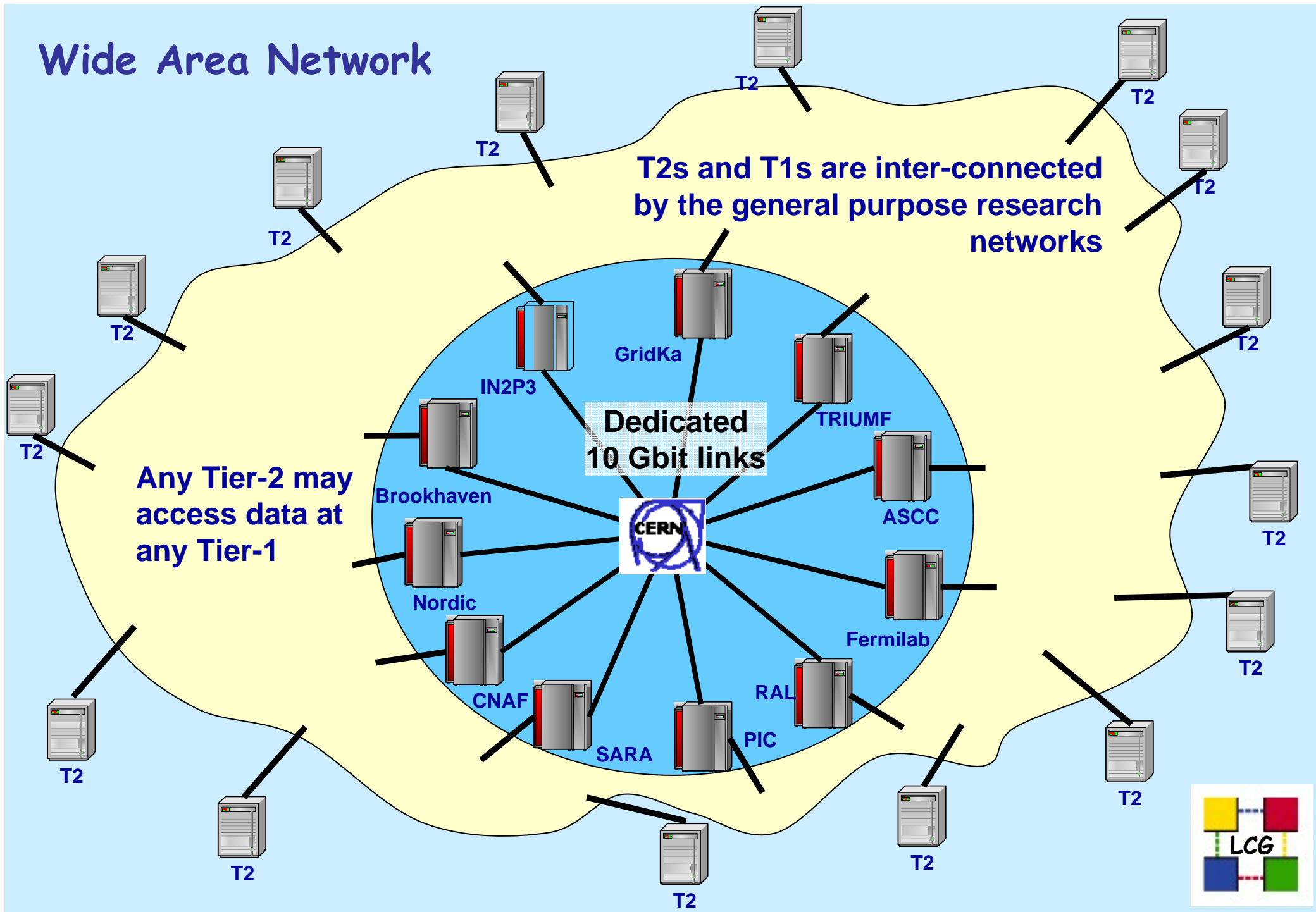
Applications Support

LCG

- All products now in production use by experiments

- Plan for Phase 2 agreed
  - Agreement on the fusion and evolution of SEAL and ROOT
  - Preparing for longer term support of common products

- Persistency Framework
  - POOL
  - Conditions Database
- Core libraries and services – a merge of
  - SEAL (components for experiment frameworks) with
  - ROOT (analysis framework)
  - → Common maths library, dictionary, ..
- Simulation
  - Simulation framework
  - GEANT4
  - Fluka
  - Physics validation
  - Garfield
  - MC generator services
- Software Process & Infrastructure

Planning for Phase 2

Phase 2: service commissioning, initial operation – 2006-08

- Technical Design Report – 20 June 2005
  - Review of the LCG and experiment Computing TDRs on 7-8 October

- Phase 2 planning
  - Phase 2 Service challenge schedule
  - Applications Area plan
  - Service Challenge 4 (March-September 2006) plan being elaborated
  - CERN fabric
    - acquisition plan completed
    - CASTOR 2 testing & migration plan

  - Detailed plan for next 18 months ($\rightarrow$ service commissioned)
    - Regional centre milestones
    - Service level metrics
    - DAQ – Tier-0 – Tier-1 testing
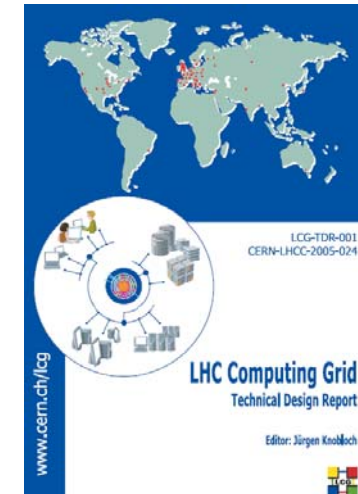
Building the Worldwide LHC Grid

LCG

2005

2006

today →

2007

2008

cosmics

first beams

first physics

full physics run

SC1 - *Nov04-Jan05* - data transfer between CERN and three Tier-1s (FNAL, NIKHEF, FZK)

DRC1 – *Mar05* - data recording at CERN sustained at 450 MB/sec for one week

SC2 – *Apr05* - data distribution from CERN to 7 Tier-1s – 600 MB/sec sustained for 10 days (one third of final nominal rate)

SC3 *service phase - Sep-Dec05* - demonstrate reliable basic service – most Tier-1s, some Tier-2s; CERN→Tier-1 data rates 150 MB/sec (60 MB/sec to tape)

DRC2 – *Dec05* - data recording sustained at 750 MB/sec for one week

DRC3 – *Apr06* - data recording sustained at 1 GB/sec for one week

SC4 *service phase - May-Aug06* - demonstrate full service – all Tier-1s, major Tier-2s; full set of baseline services; data distribution and recording at nominal LHC rate (1.6 GB/sec)

LHC Service in operation – *Sep06* – over following six months ramp up to full operational capacity & performance

DRC4 – *Sep06* - data recording sustained at 1.6 GB/sec

LHC service commissioned – *Apr07*

LCG

SC -   Service Challenge
DRC -Data Recording Challenge

last update 26/01/2006 12:48

Summary – Current Status

- **Scale of underlying grids already at/beyond target level**
- **Basic operational environment established**
  - good and growing collaboration between operations centres
- **Baseline services agreed**
  - implementation by start of next year realistic
- **Service challenges have progressed**
  - from 4 sites last November
  - through 7 sites in April
  - to ~20 sites in SC3 - *including all major centres*
  - lots of *problems being identified – and solved*
- **Applications medium term plan agreed**
- **Database services still to be decided and deployed**
  - Workshop end of October

First data in less than 2 years

- CERN + Tier-1s must provide an *integrated* and *reliable* service for the bulk data from first beams

- *NOT an option to get things going later*

- Priority must be to concentrate on getting the basic service going
    - modest goals
    - pragmatic solutions
    - collaboration

last update 26/01/2006 12:48