

Site report of the Grid Computing Centre Karlsruhe

Forschungszentrum Karlsruhe GmbH
Institute for Scientific Computing
PO Box 3640
D-76021 Karlsruhe

Holger Marten

<http://www.gridka.de>

The Grid Computing Centre Karlsruhe

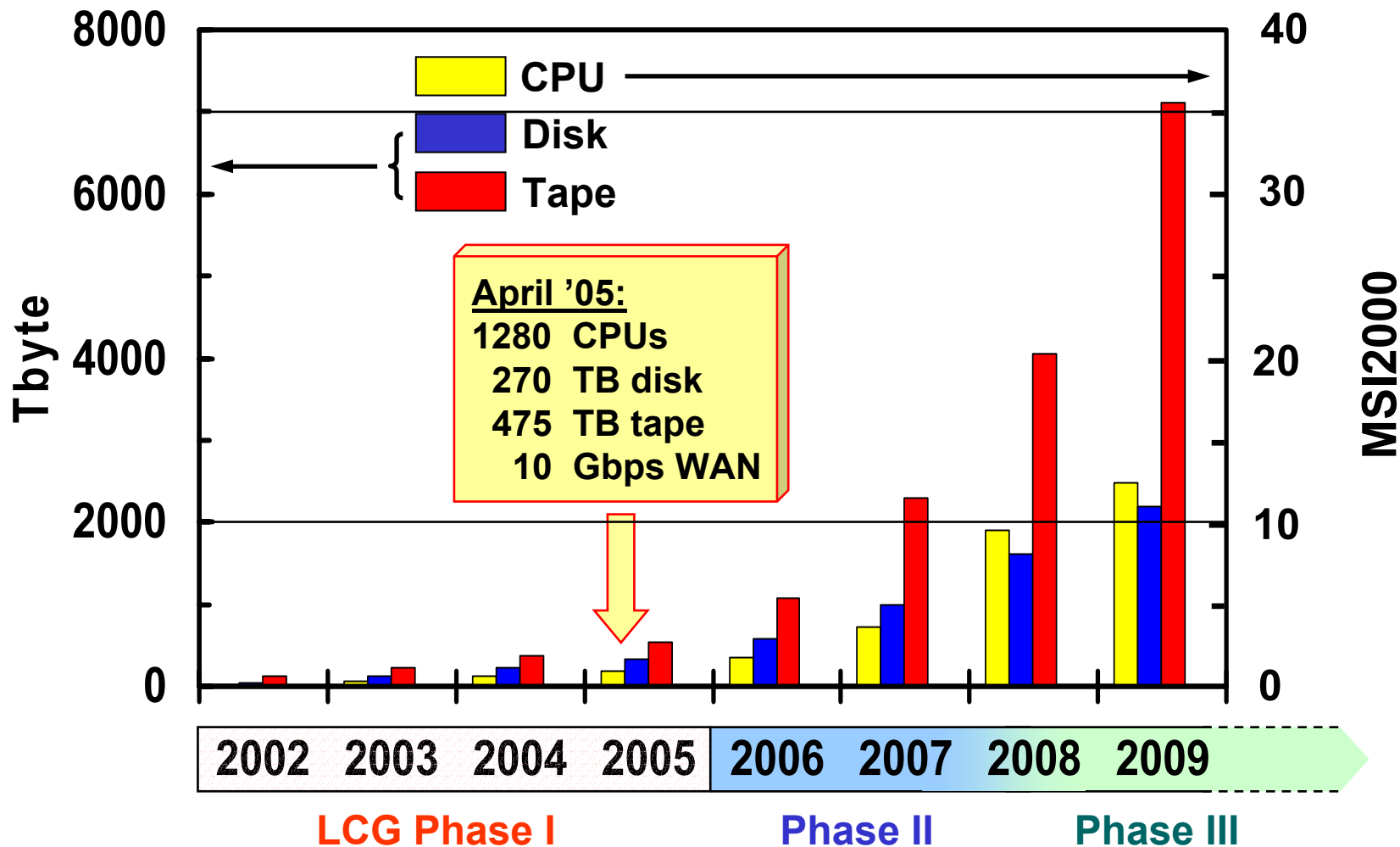
Requested as a Regional Data and Computing Centre by 41 HEP user groups in 19 German universities and research institutions.

Founded in 2001 as part of the computing centre of Forschungszentrum Karlsruhe.

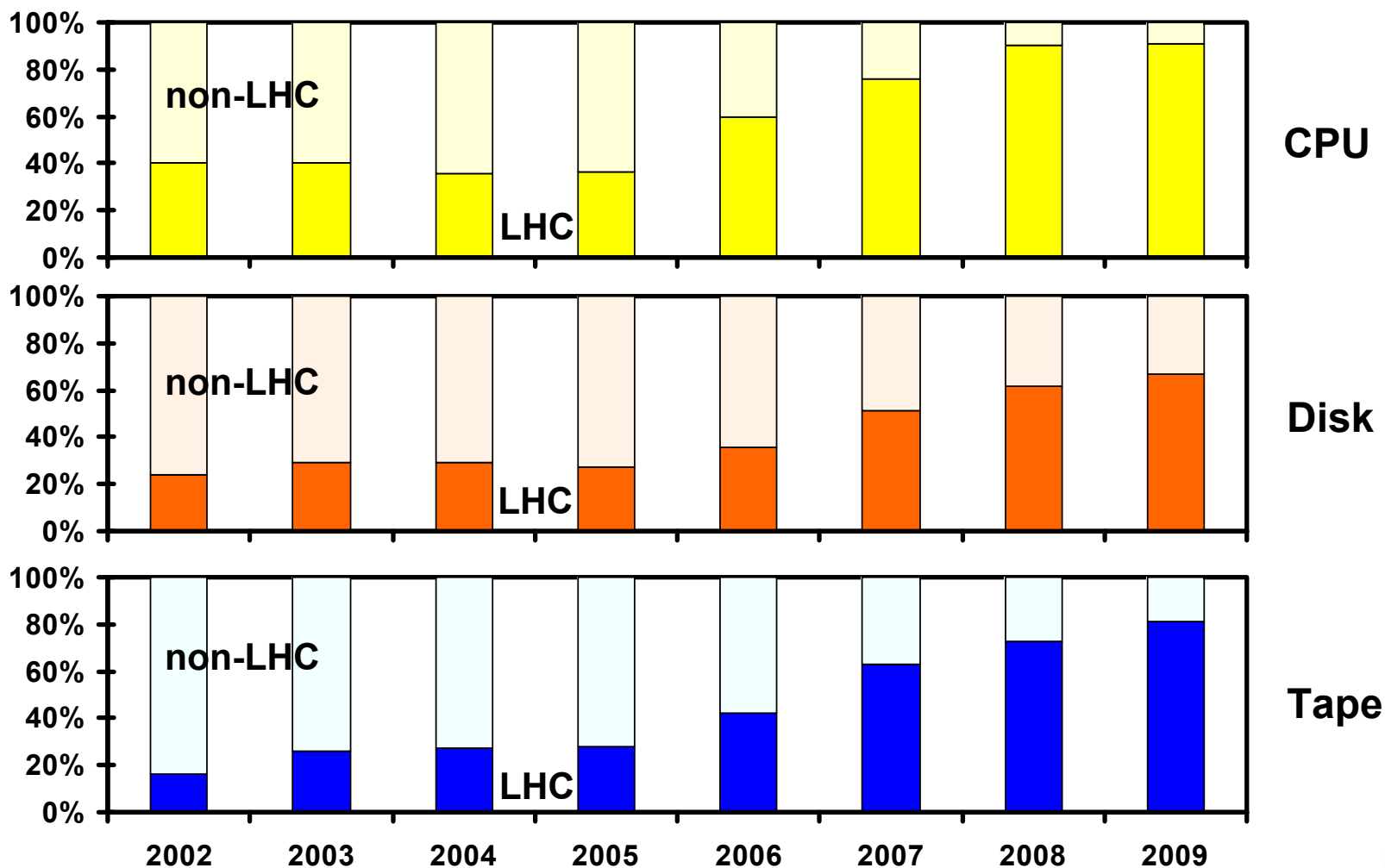
Main goals

- test environment for LHC (**ALICE, ATLAS, CMS, LHCb**)
- LHC Tier-1 in 2007+
- production environment for non-LHC (**BaBar, CDF, D0, Compass**)
- environment for Grid R&D (**CrossGrid, LCG, EGEE, ...**)
- user support, services, education & training
- grid environment for other sciences (astrophysics, bio-informatics...)

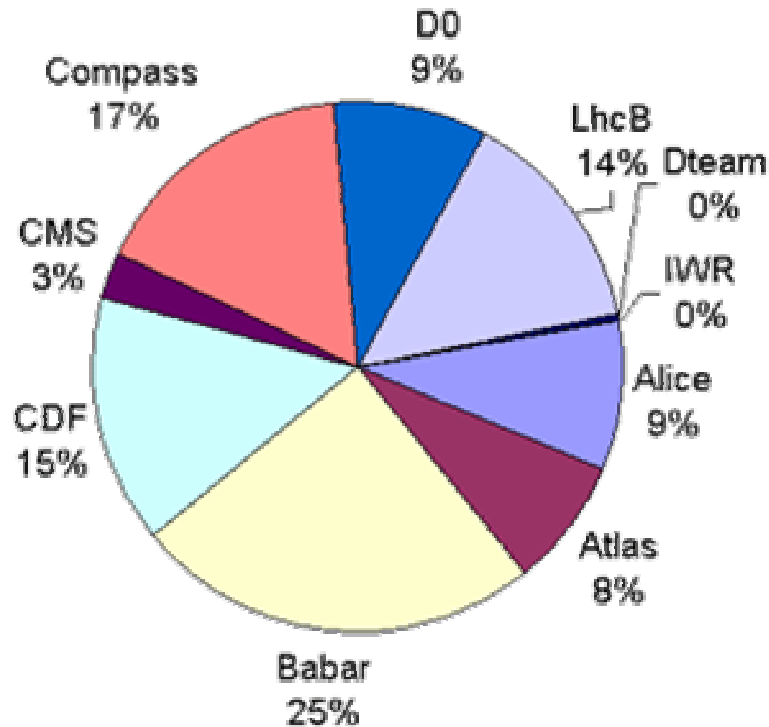
GridKa – Planned hardware resources (10/2004)



Distribution of planned resources at GridKa



Usage statistics 2004



2004

LHC 34%

nLHC 66%

Number of CPUs

01/2003	380
06/2003	460
04/2004	680
06/2004	790
10/2004	1070

	2003	2004
Processor usage [h]	1.980.000	4.183.000
Number of jobs	332.000	1.442.000

Infrastructure

Use existing infrastructure of FZK (buildings, electricity, ...) Special Technical Infrastructure Division at FZK

Floor space

- ~600 m² floor space reserved for GridKa

Electricity

- 20kV supply network
- several transformers with 2 MW for computing centre available

Cooling

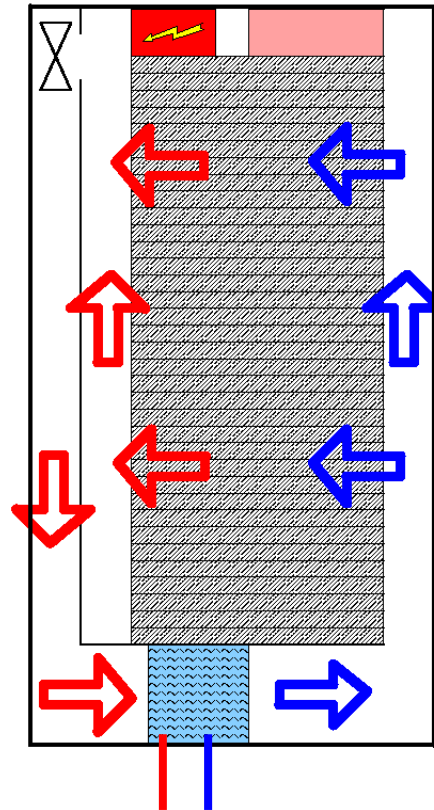
- 3(+1) external heat exchangers with 600 kW each
- however, air condition was an issue !
(reduced cross sections of air flow channels)

Power rails at the ceiling



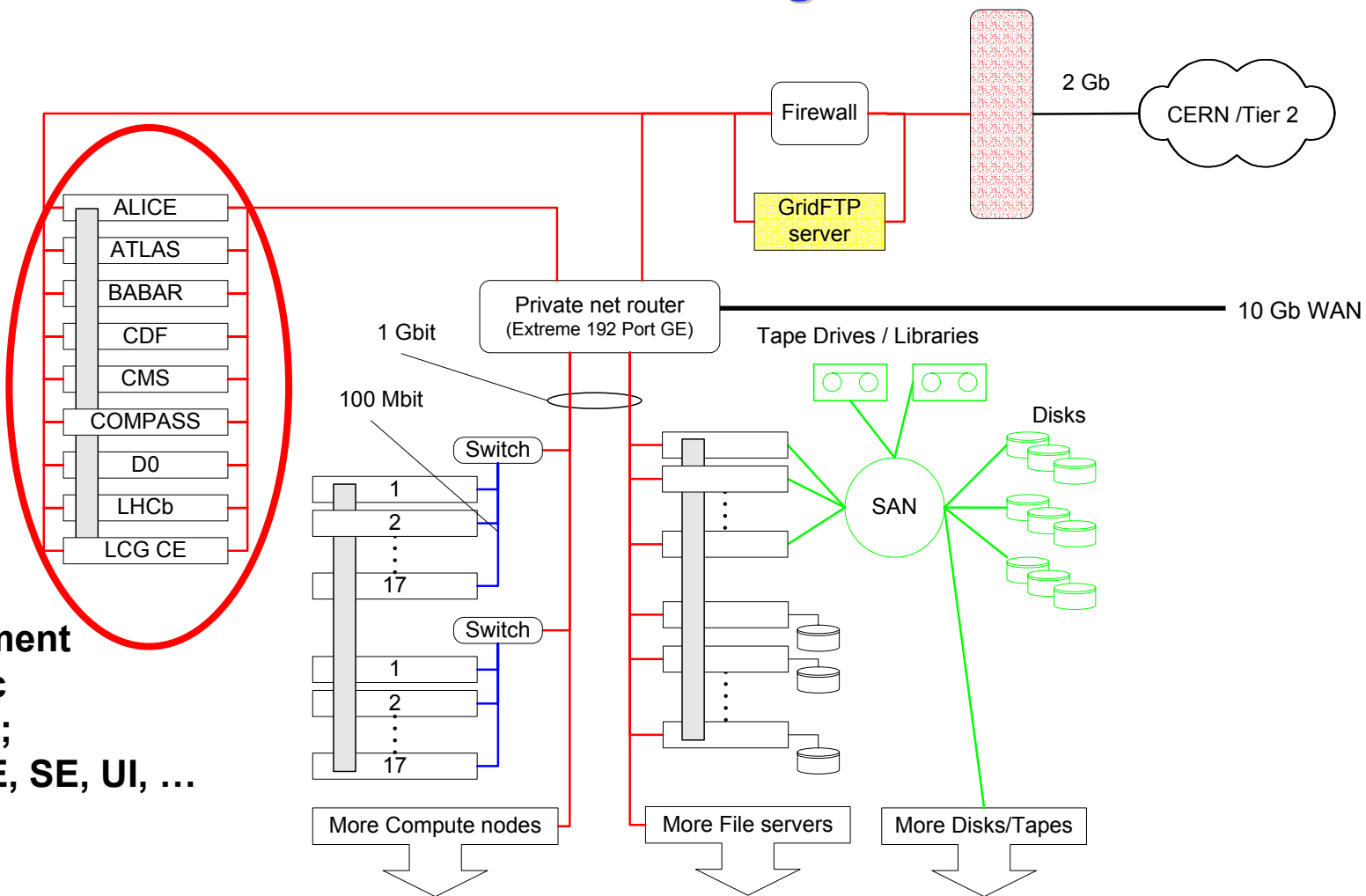
- separate well defined rails for “normal” power and UPS
- very flexible and scalable
- prevents from (un)plugging wrong cables
- prevents from electrical hazards
- separates electricity from water

Equipment cabinet with water cooling



- 19" technique
- 37 units usable height
- 70x 120 cm floor space
- 10 kW cooling
- redundant DC fans
- temperature regulated fans
- temperature controlled
 - 22° warning
 - 26° critical
 - 30° power off
- internal smoke detector
- SNMP monitored
- **manual reset after power off !**

Network diagram



**Experiment
specific
servers;
LCG CE, SE, UI, ...**

Worker Nodes “Production”

97x dual Intel PIII 1.26 GHz
64x dual Intel PIV Xeon 2.2 GHz
72x dual Intel PIV 2.667 GHz
267x dual Intel PIV 3.06 GHz
144x dual AMD Opteron 246, 2.2 GHz

- 1 or 2 GB ECC RAM (~0.75 TB RAM)
- 40 or 80 GB HDD IDE (ca. 35 TB local disk!)
- LAN groups: 18x 100 Mbit + 1 Gbit uplink or 18x 1 Gbit + 2x 1 Gbit uplink
- completely with SL3.0.4 + LCG 2.3
- moving to LCG 2.4 during the next days

Batch system:

- started with OpenPBS (didn't scale), now PBSPro (package with 5000 licences)
- dynamic queue limits & fair share for optimal resource usage
- problem “cooling” solved (Knürr cabinets), but high cost for electricity !

Online Storage

Total ~365 TB gross,
~ 2300 disk platters

DAS

- ~ 2.6 TB gross, 7.2k SCSI 20x 120 GB, attached to dual Intel Linux; runs since 2001 !
- ~ 3.8 TB gross, 2 system, 12x 5.4k IDE 160 GB, attached to dual Intel Linux

NAS

- ca. 50 TB gross, 22x dual PIII 1/1.26 GHz, 16x 5.4k IDE 100 / 120 / 160 GB
dual 3Ware controller
 - many problems with IDE-disks & controllers; unfavourable mechanical assembly?
 - poor software for error analysis
 - inflexible; fixed file system sizes, fixed performance

SAN

- ~ 12.5 TB gross, 1x STK D280, 6x 14 10k fc 150 GB
- ~ 252 TB gross, 12x IBM FastT700, 10x 14 10k fc 150 GB
- ~ 43 TB gross, 1 Data Direct , 16x 9 S-ATA 250 GB replaced with 12x 9 S-ATA 400 GB
- driven by 3 gpfs clusters, 12 machines each, single / dual Intel, Linux RHE3.0
 - more scalable than NAS; file systems & performance extendable
 - load balancing through file server cluster

Tape Storage



IBM 3584

- ~ 4000 slots LTO
- 100 GB and 200 GB /tape
- 475 TB native available
- (8 drives **LTO-1**, 15 MB/s FZK)
- 8 drives **LTO-2**, 30 MB/s
- backup/archive with **Tivoli Storage Manager**
- combination of **dCache with TSM**

- **LTO-2: ~ 4000 mounts per week**
- **different qualities of cassettes from different vendors**

Installation and Management

Software Installation

- NPACI Rocks with own extensions since 2001
- <http://rocks.npaci.edu>

System Monitoring

- cluster usage for users available on the web
- Ganglia Cluster Toolkit <http://ganglia.sourceforge.net>

System Management

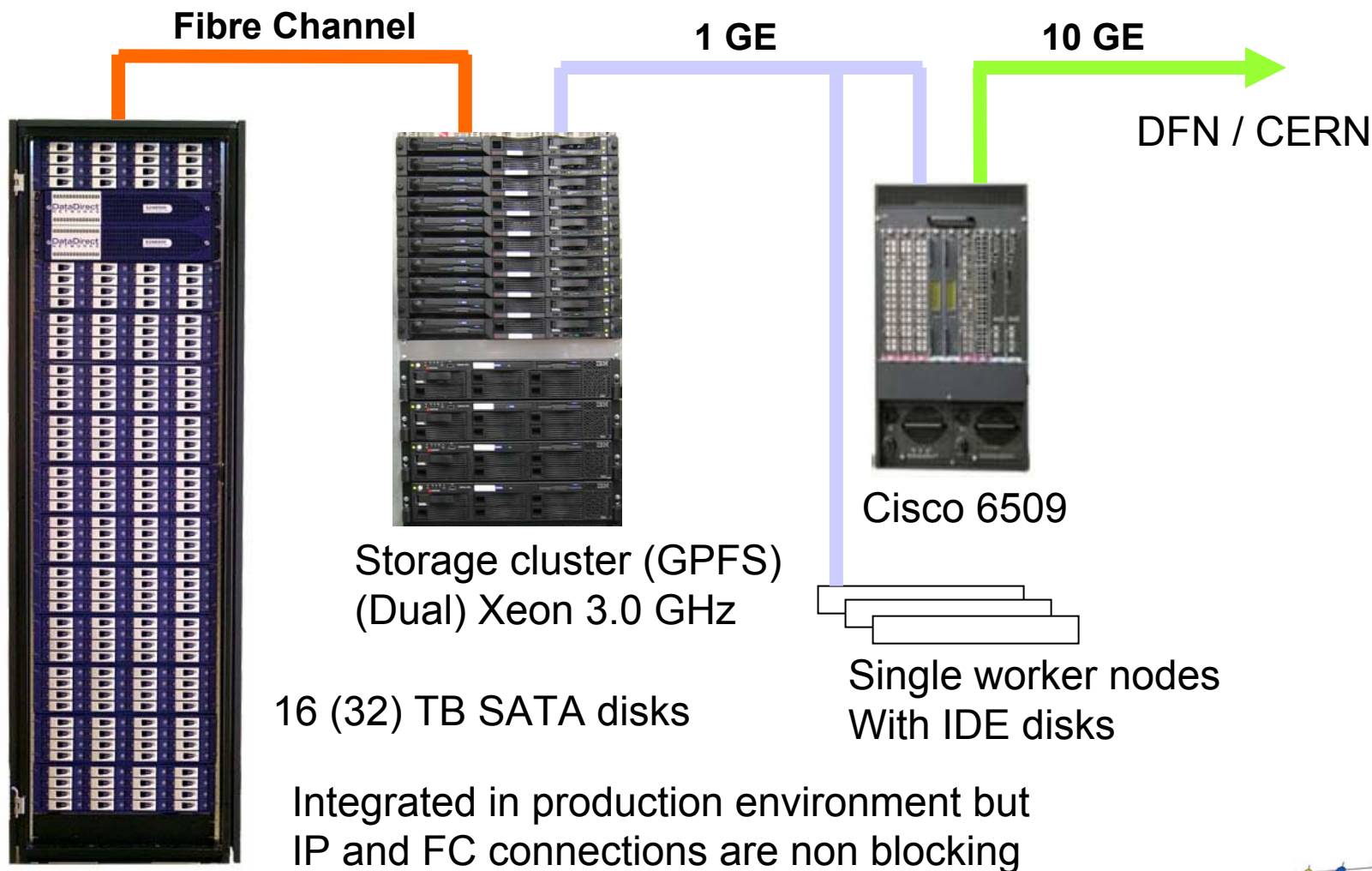
- “health monitoring” with NAGIOS <http://www.nagios.org>

User statistics, cluster & cable data bases

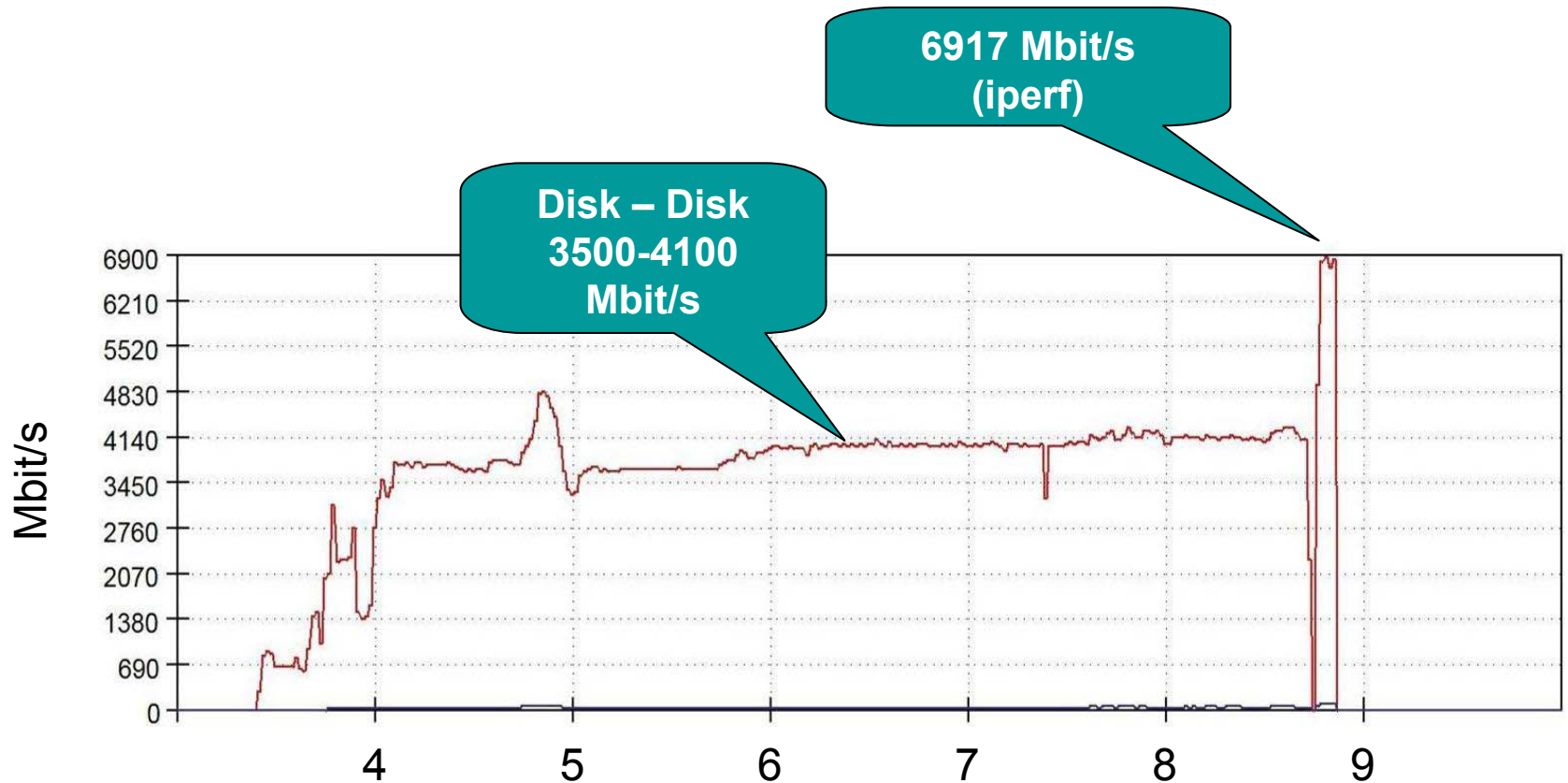
- home made

External monitoring & accounting via EGEE GOC

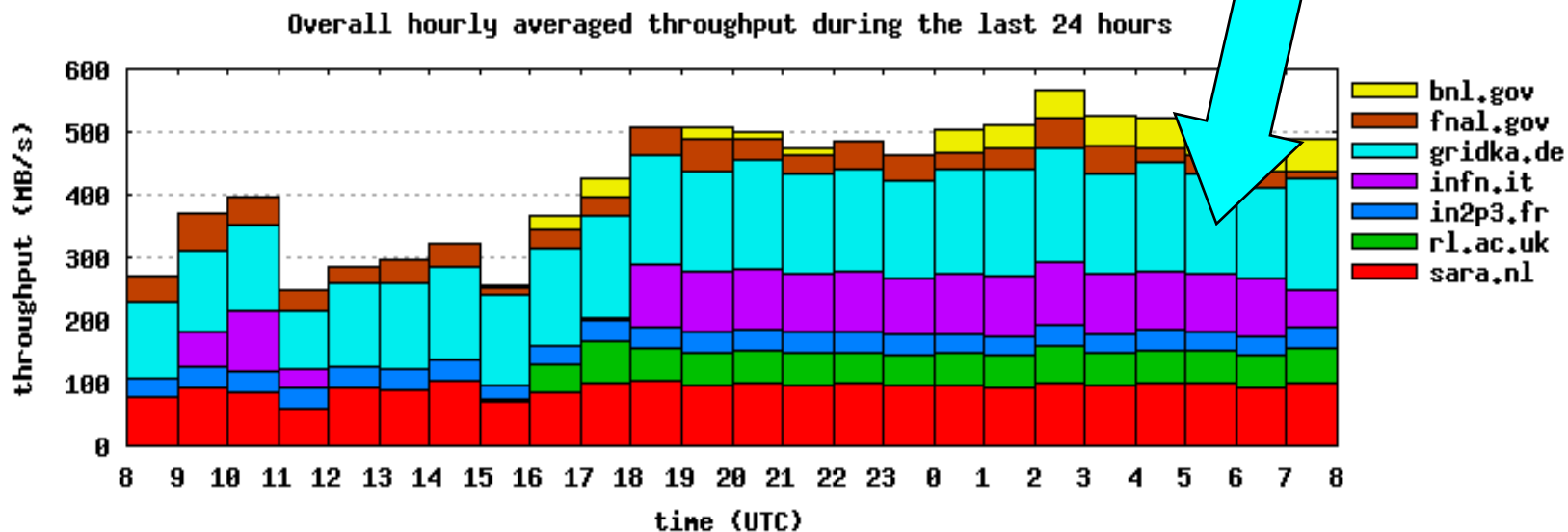
Service Challenge Setup @ GridKa



Service Challenge @ GridKa Transfer Rates 3.-9. Feb'2005



Service Challenge @ GridKa Live Monitor 22-23.3.2005



- **10 sites in the German-Swiss federation**
 - CSCS, DESY, FhG/SCAI, FhG/ITWM, FZK (GridKa), GSI, HU Berlin, Univs. Aachen, Karlsruhe (IEKP) & Wuppertal,
 - Supporting HEP, BioMed, Earth Science & Astrophysics
- **Pre-Production Service at FZK (preparation for gLite)**
- **Distributed Regional Operation Centre (ROC) & rotating operation support together with DESY, FhG and GSI**
- **Global Grid User Support at FZK**
- **Several mw & installation training events (NA2)**
 - Preparing GridKa School 2005 in September

Forschungszentrum Karlsruhe in der Helmholtz-Gemeinschaft



No equipment without people. Thanks !

**We appreciate the continuous interest and support by the
Federal Ministry of Education and Research, BMBF.**



Bundesministerium
für Bildung
und Forschung