CMS Tier-1 & AF at CC-IN2P3

Overview - visit of CMS computing

Fabio Hernandez
fabio@in2p3.fr

Lyon, October 23rd 2009

Contents

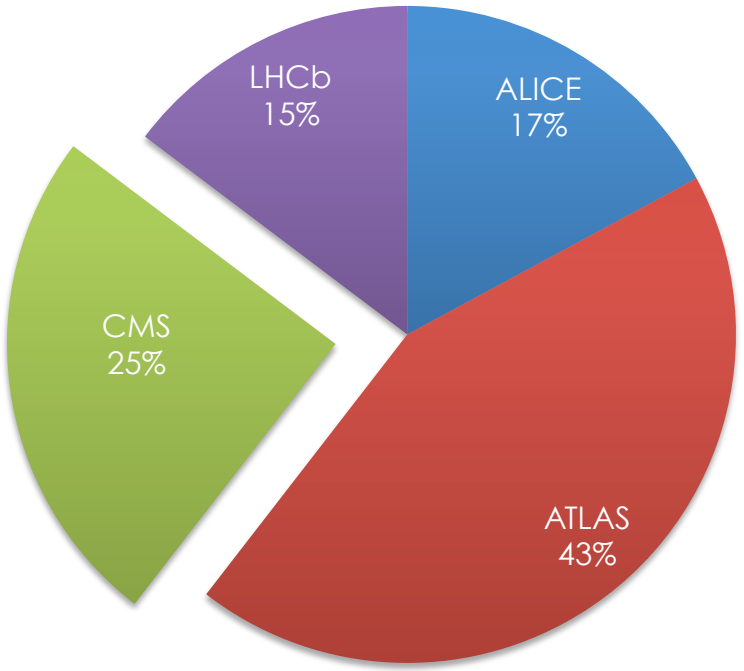
- Budget
- Site overview
- Functional organization
- Resource utilization
- Conclusion
- Questions & comments

Budget: all LHC Experiments

- Budget requested on a pluriannual basis
  - Approval on a yearly basis
  - Impact on hardware procurement process
- Equipment and running costs for all LHC experiments at CC-IN2P3 (2005-2010)
  - Total required: 23 M€
    - *the refurbishment of current machine room and the construction of a second one are NOT included*
    - *Salaries are NOT included*

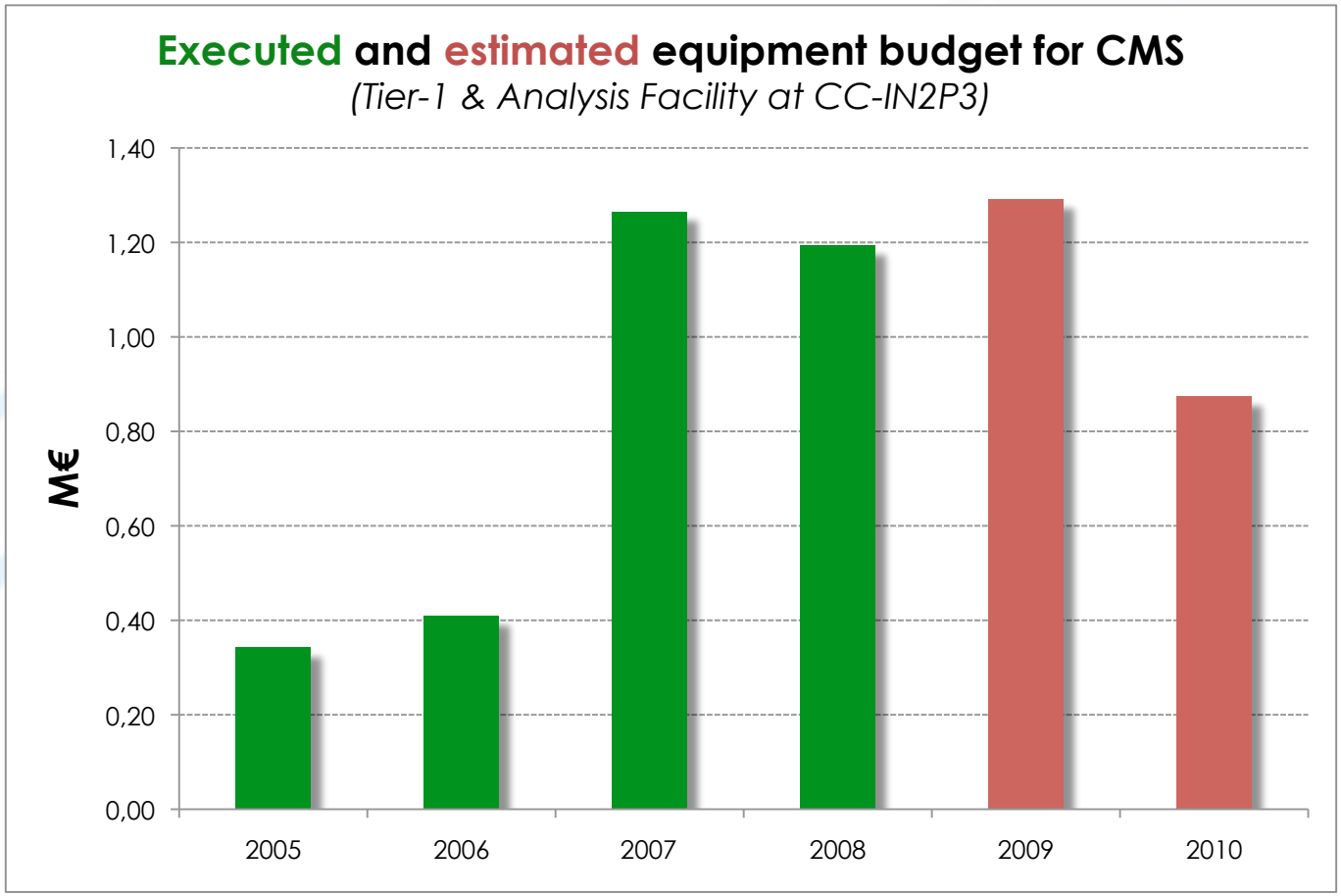
Equipment budget share for LHC experiments -- 2005-2010

(Tier-1 & analysis facility at CC-IN2P3)



Budget: CMS

- Equipment and running costs for CMS needs over the period 2005-2010
  - 5.4 M€



Pledged Resources for CMS

	Tier-1		Tier-2	
	2009	*2010*	*2009*	*2010*
CPU [HEP-SPEC06]	5 147	11 055	3 384	6 825
Disk [TB]	7 15	1 536	228	322
MSS [TB]	1 666	2 563	n/a	n/a

~11% of the required resources for all CMS tier-1s

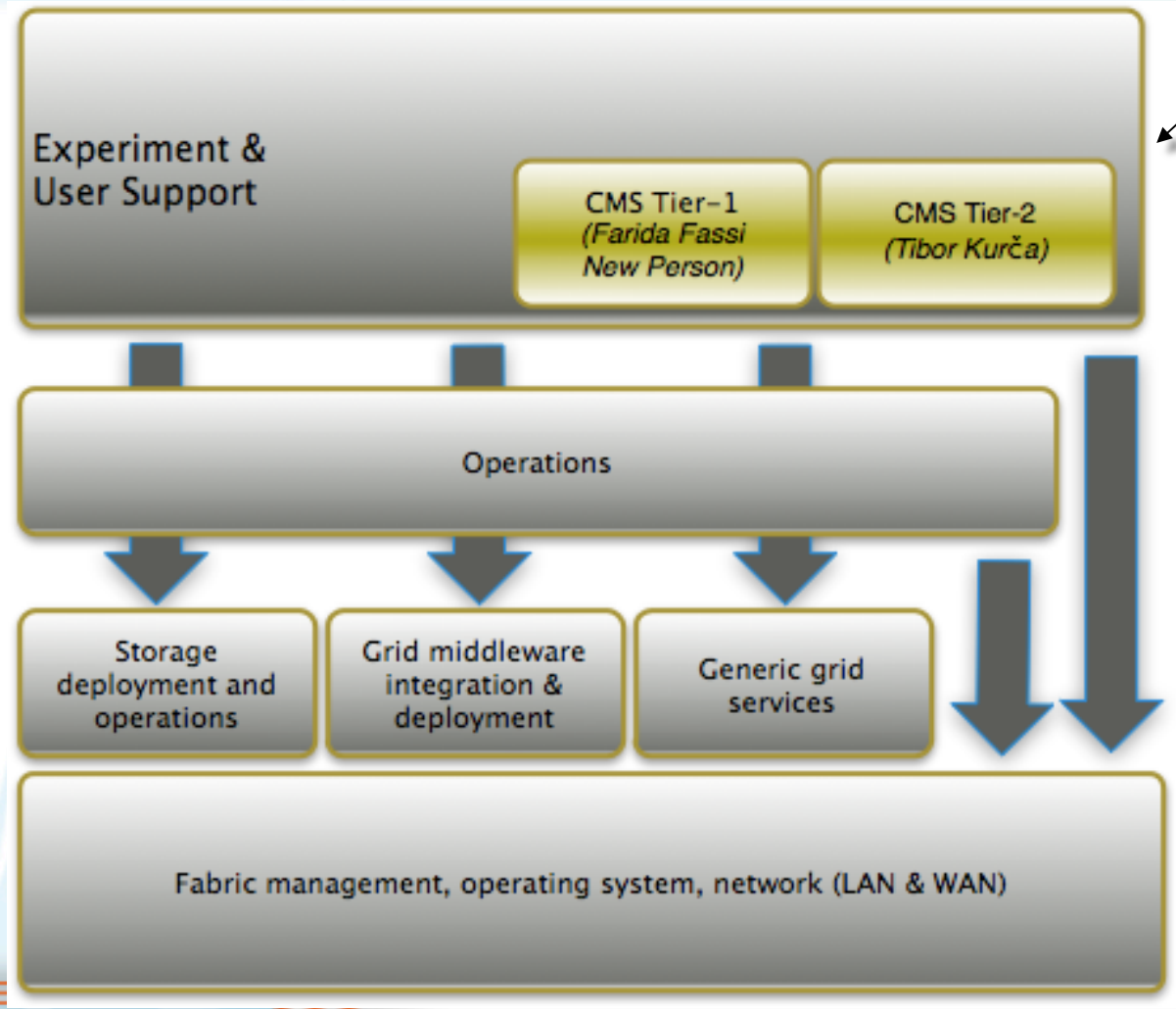
~8% of the required resources for all CMS tier-2s

Site overview

Introduction to CC-IN2P3

- Data repository and processing facility **shared** by several experiments
  - Operates a WLCG tier-1, a tier-2 and an Analysis Facility for the 4 LHC experiments
- The main compute farm used by both grid and local users
  - Grid middleware is "just another" interface for using our services
- Data storage infrastructure (disk and mass storage) accessible to all jobs running in the site
  - Although not all storage spaces have a gridified interface

Experiment-specific support



Selection process of new person for reinforcing the tier-1 dedicated support for CMS finished.

Expected start on Dec 1st. 2009

Grid
Internet &
LHCOPN
Compute Elements
Tier-1
Tier-2
BQS
Batch farm
Analysis Farm
xrootd
PROOF
LAN
Various services
Software
Area
(AFS)
Nagios
Monbox
VO Box
User
Interface
FTS
LFC
Information
System
Oracle
Storage Element
dCache
xrootd
HPSS
Author: Fabio Hernandez
Last Updated: 2009-06-09
CCIN2P3

CE configuration

Tier Level	CE hostname [.in2p3.fr]	ALICE	ATLAS	CMS	LHCb
Tier-1	cclcgceli01	✓	✓		
	cclcgceli02	✓	✓		
	cclcgceli03			✓	✓
	cclcgceli04			✓	✓
	cclcgceli07	✓	✓	✓	✓
	cclcgceli08	✓	✓	✓	✓
Tier-2	cclcgceli05	✓	✓	✓	✓
	cclcgceli06	✓	✓	✓	✓

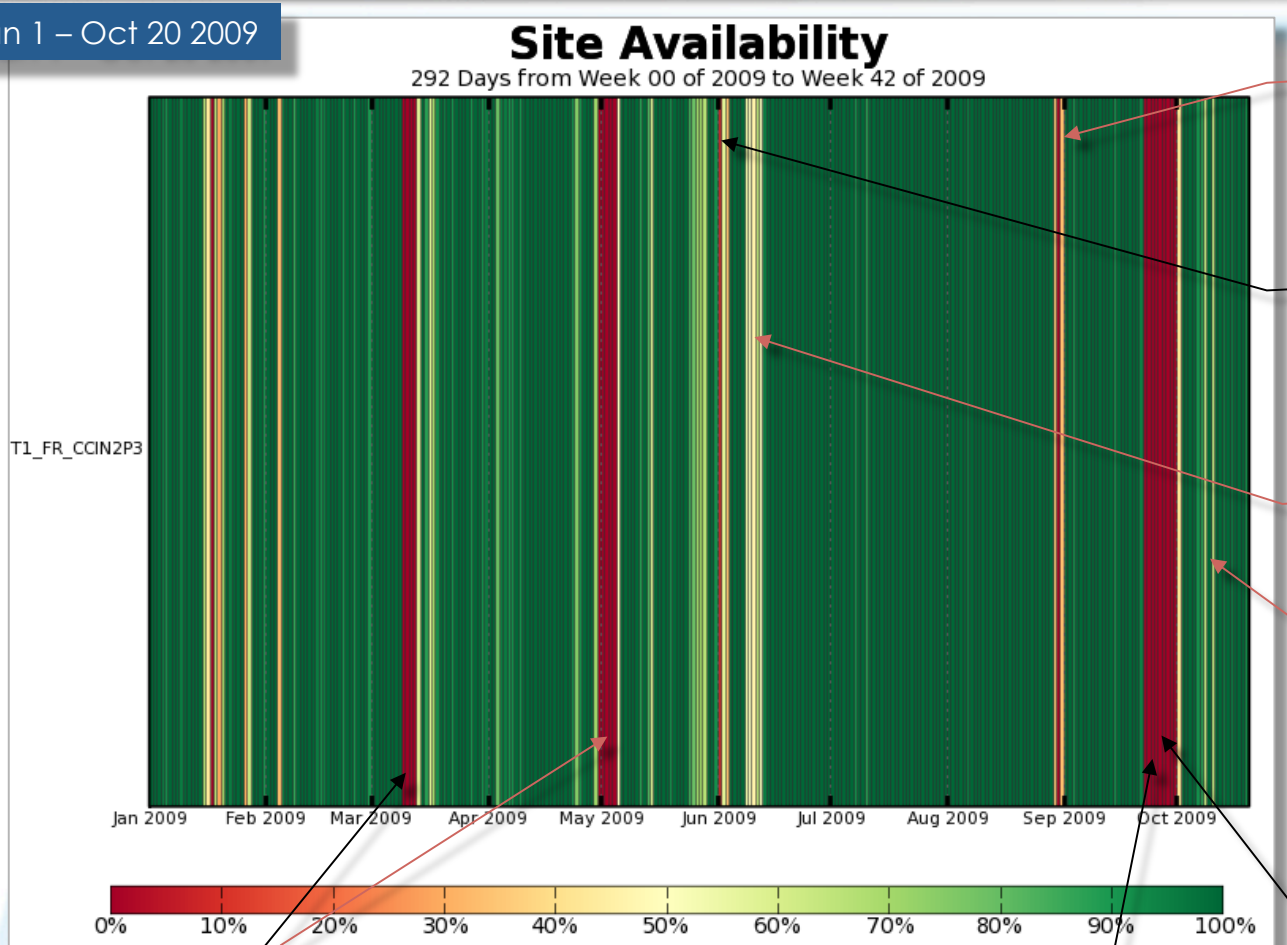
For submission to the WNs running SL5

VOMS roles configured according to CMS needs. Details in Farida's talk

This configuration will be modified for decommissioning SL4 for 3 experiments, including CMS (planned for 29/10/2009)

CMS tier-1s – site availability

Jan 1 – Oct 20 2009



Low reactivity to tranfer-related problems due to lack of CMS-specific support on site during this period

Scheduled interruption of service, Jun 1-4. (upgrade to HPSS v6.2)

CMS-specific SAM tests failed due to human errors (local TFC changes not committed to CVS) and some gridFTP errors

SRM instabilities (8 and 10 Oct.) in dCache v1.9.4-3 corrected since then

Scheduled downtime 10-11 March (electrical power works)

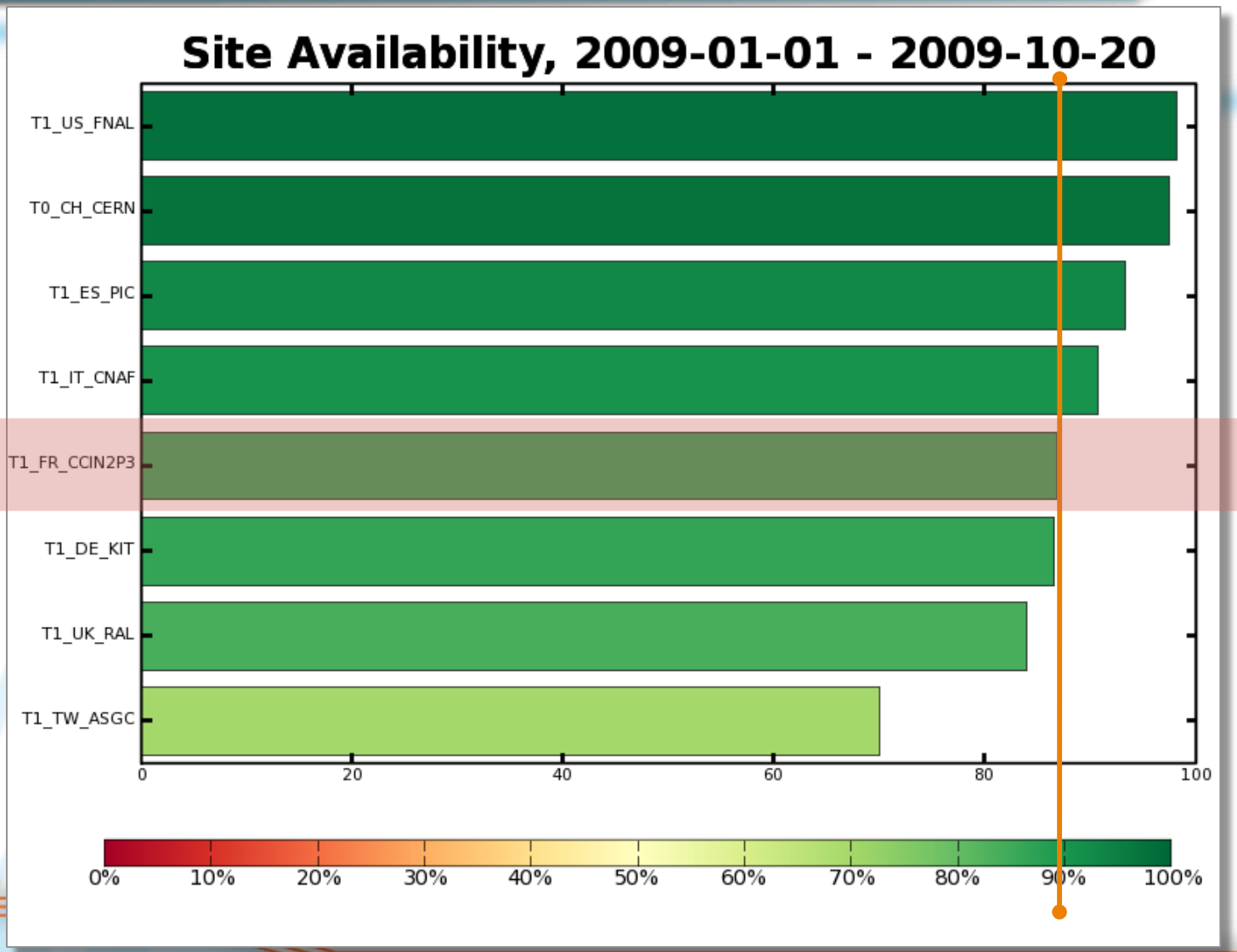
CMS-specific SAM tests failed due to human errors (local TFC changes not committed to CVS)

Scheduled downtime 21-25 September (eletrical power works, deployment of dCache v1.9.4)

Scheduled interruption of service, 28 Sep. - Oct. 1 (deployment of Chimera)

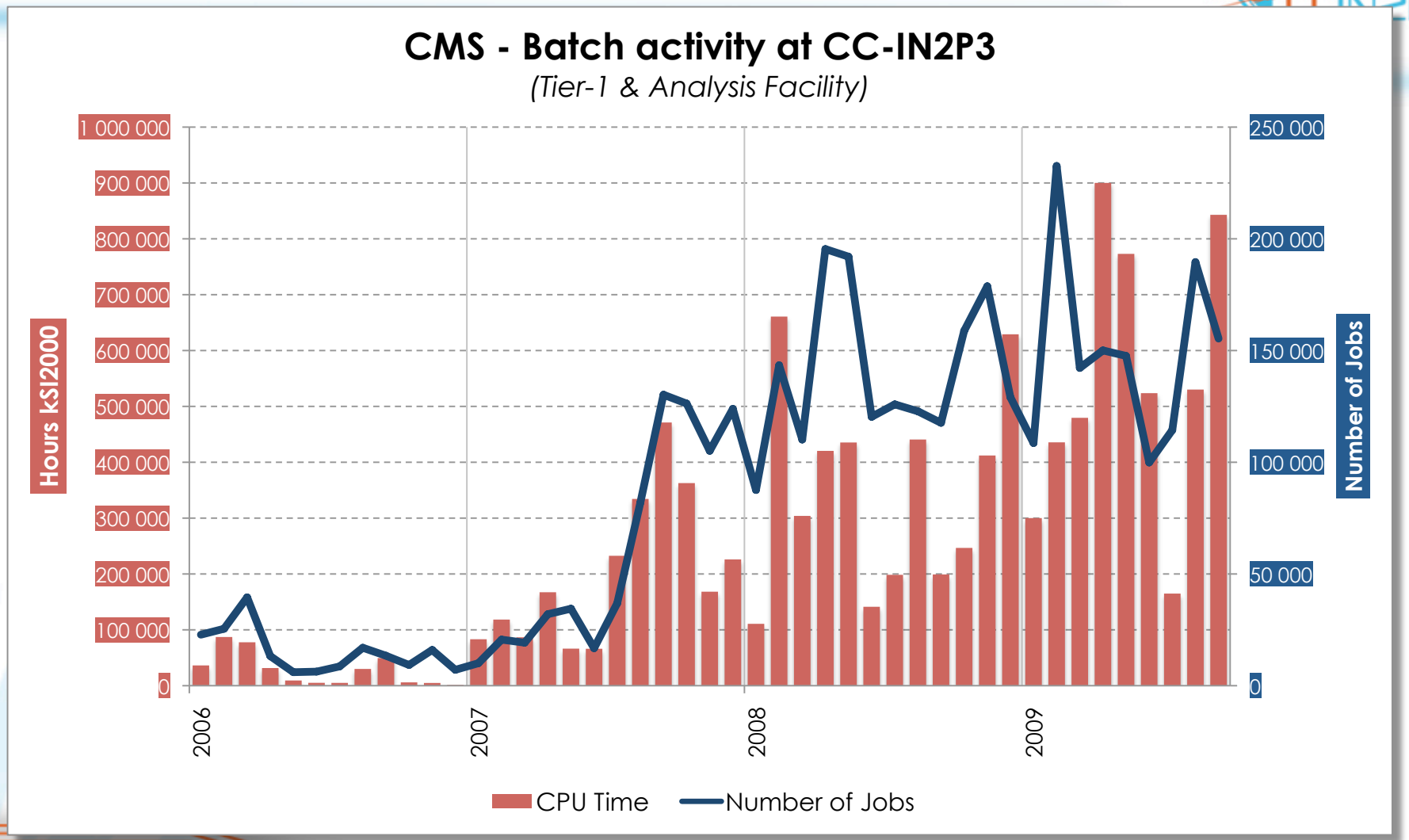
Source: http://dashboard.cern.ch/cms/

CMS tier-1s – site availability (cont.)

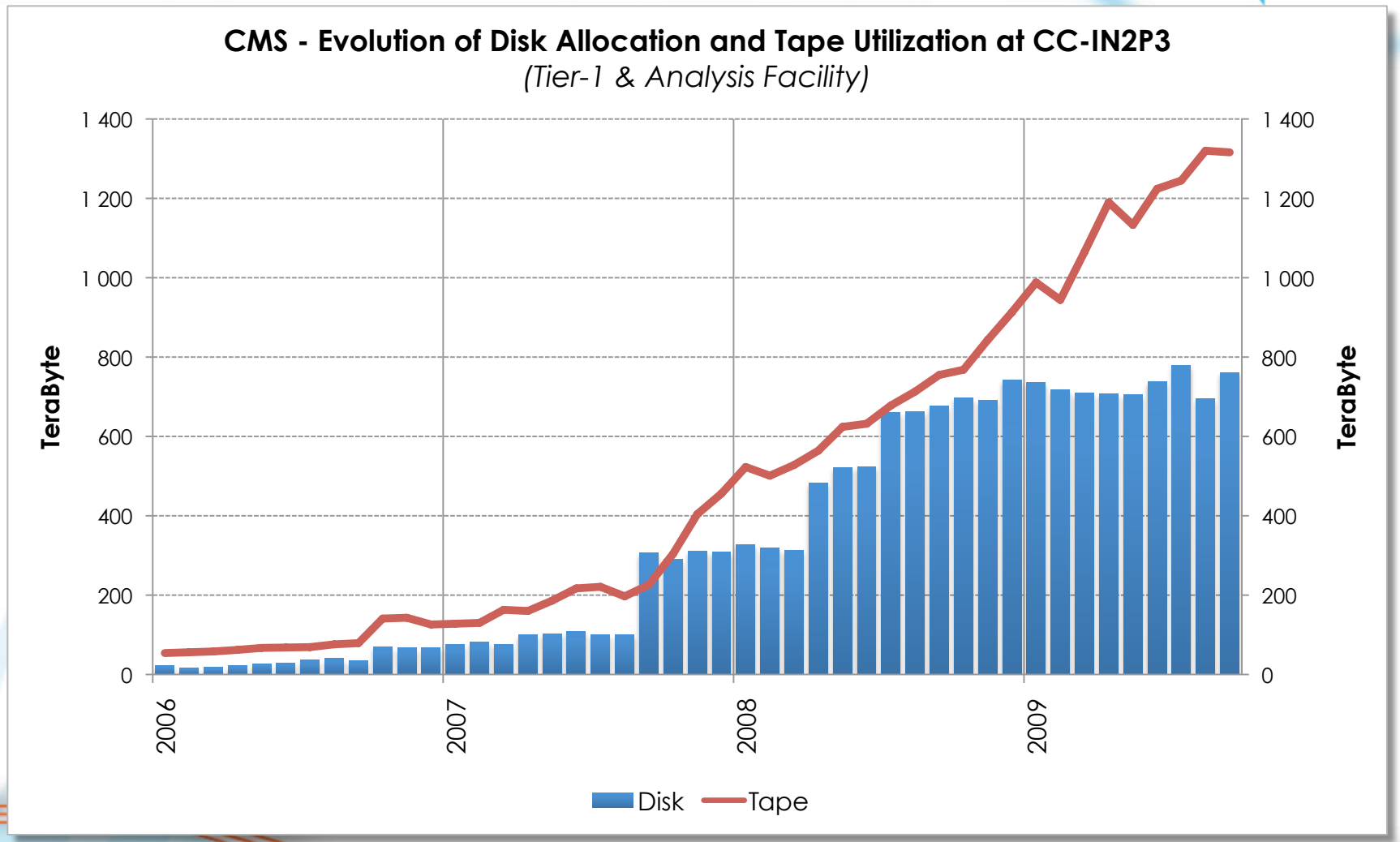


Source: http://dashboard.cern.ch/cms/

Resource Utilization -- CPU



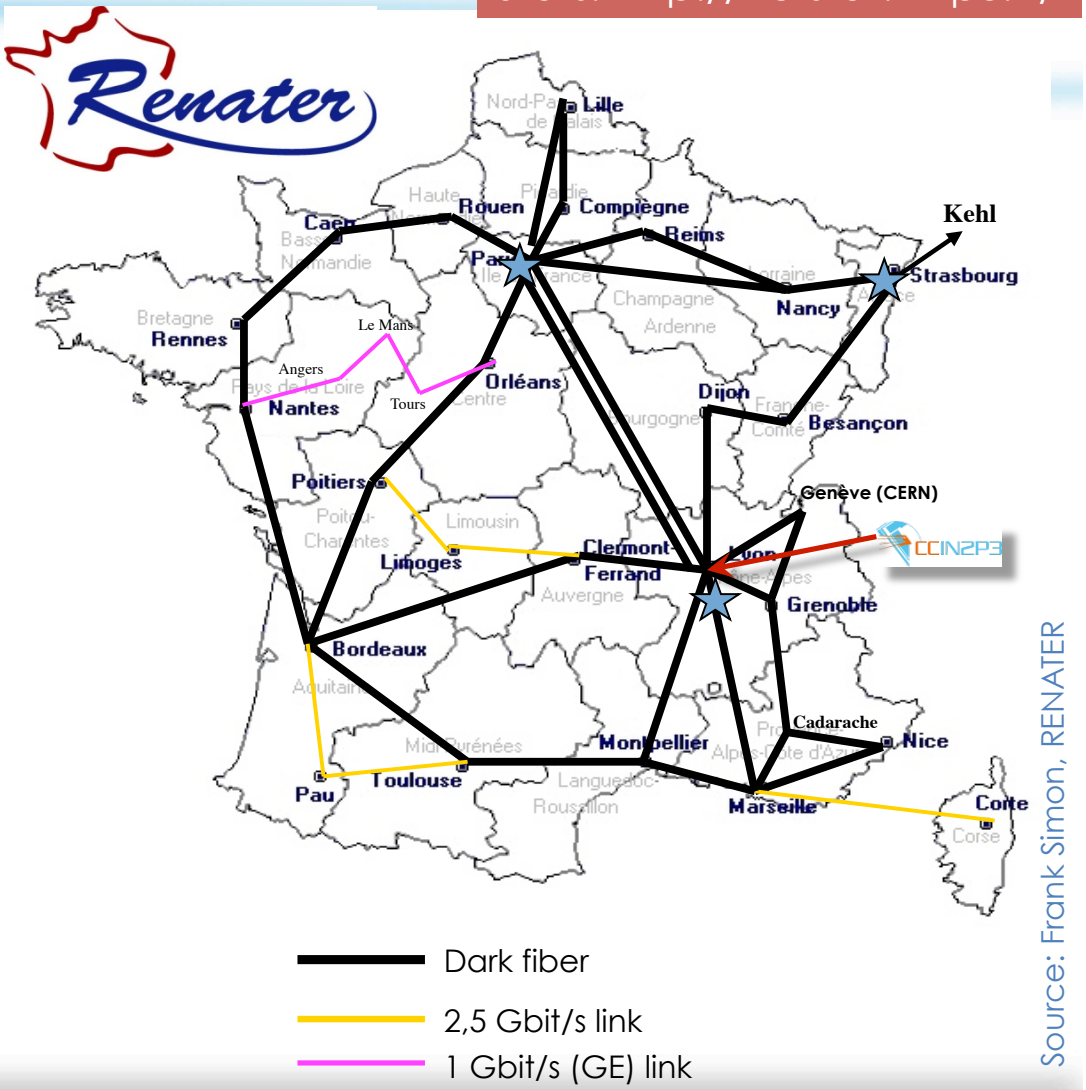
Resource Utilization -- Storage



Connectivity

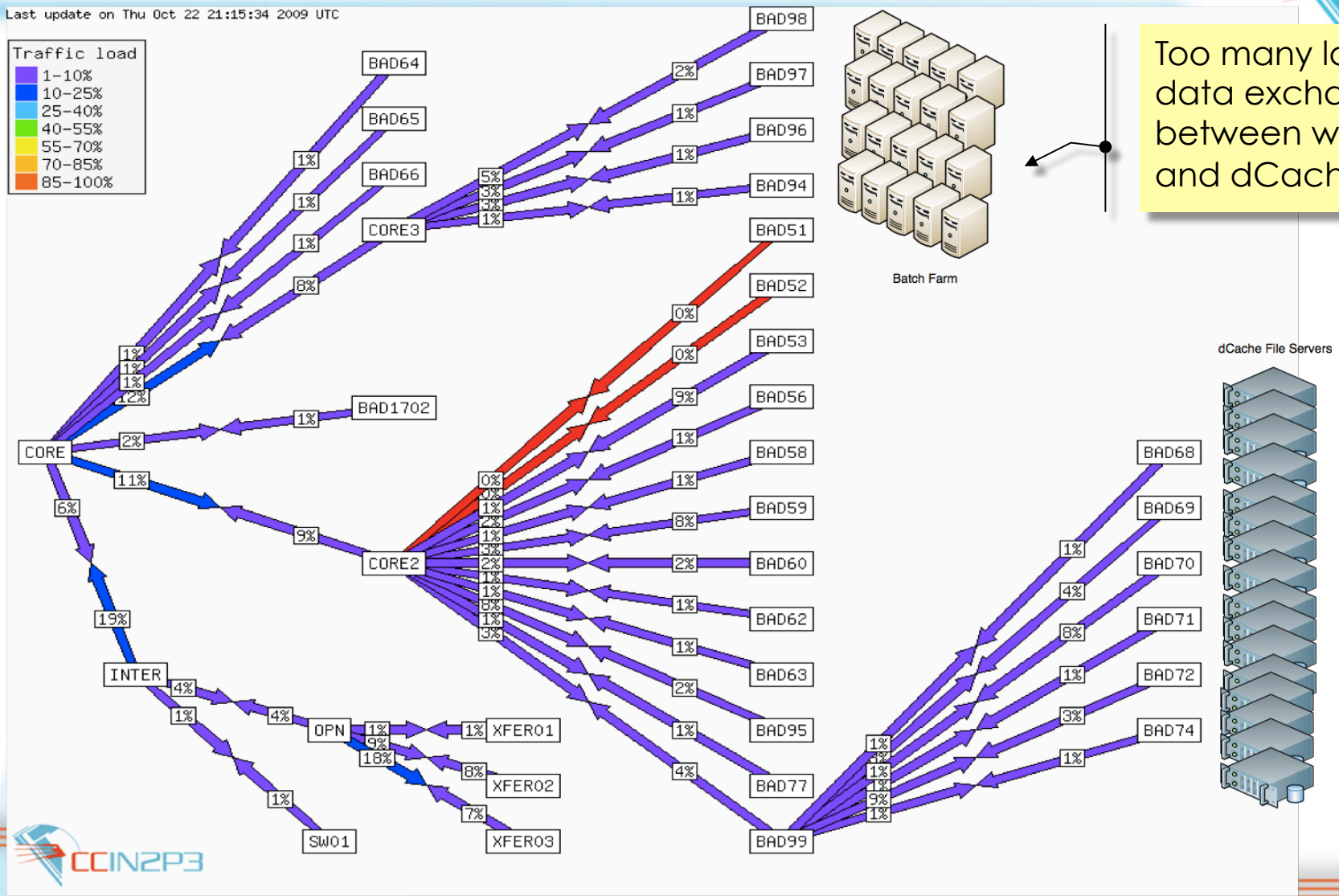
Stats: http://netstat.in2p3.fr/

- Tier-0 and tier-1s
  - LHCOPN links (10 Gbps):
    - *CCIN2P3 ↔ CERN*
    - *CCIN2P3 ↔ KIT ↔ CERN*
- Domestic CMS tier-2s and tier-3s
  - GRIF: 10 Gbps
  - IPHC: 1 Gbps
    - *Limited by the metropolitan network*
  - IPNL: 1 Gbps
- Foreign CMS tier-2s and tier-3s
  - Connection to GEANT routers at 10 Gbps



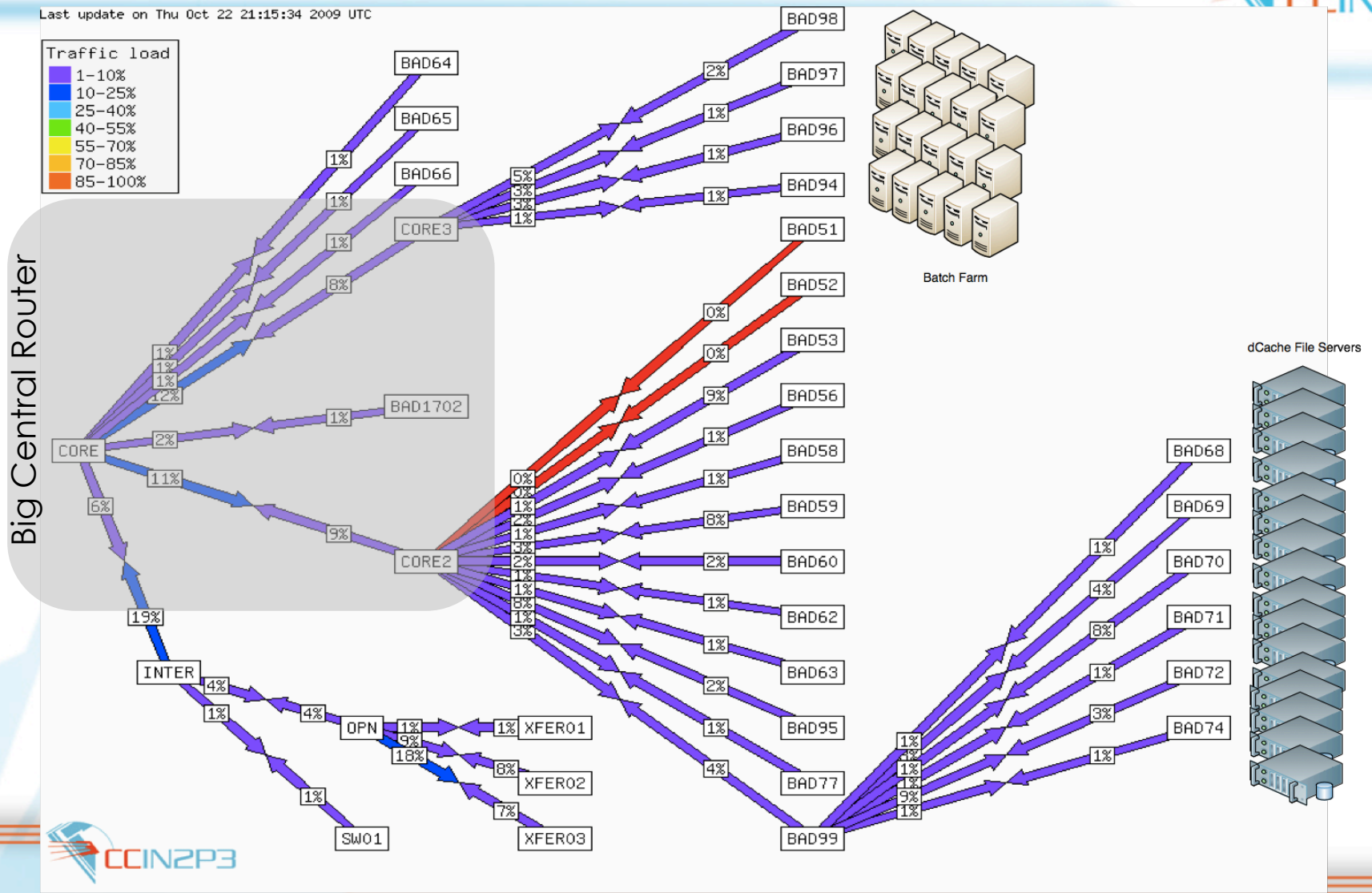
Source: Frank Simon, RENATER

Network: LAN



Too many layers for data exchange between worker nodes and dCache file servers

Network: LAN (cont.)



Network: LAN

- Perspectives
  - Replace a set of central routers/switches by a single core router/switch
    - *Increase bandwidth between worker nodes and file servers. In some cases, possibility of using non-blocking ports for guaranteeing bandwidth*
    - *Decrease latency and simplify routing*
    - *Increase number of ports for interconnecting more machines in the future*
  - Timeframe for production: 2nd quarter 2010

What's next today

- *Farida*: CMS-specific activities and results
- *Benoit*: Storage infrastructure used by CMS
- Lunch
- *Dominique*: Visit of the machine room
- *Suzanne*: Overview of day-to-day operations

- CMS facilities and data operations reports

Conclusions

- Quite a lot of activity during this year for preparing the start of data taking
  - Power and cooling infrastructure
  - Storage chain
  - Operations tools
  - Wide area network
- CPU and storage resources provided on time according to pledges
- Several improvements already planned for next year

- I'm confident we are in a far better position for data taking than 12 months ago

Questions/Commentaires