



LHC Computing Grid

A few notes about IPHC T2

Yannick Patois



IN2P3

INSTITUT NATIONAL DE PHYSIQUE NUCLÉAIRE
ET DE PHYSIQUE DES PARTICULES



- **IPHC as a T2**
- **Machine room**
 - Infrastructure
 - Security
- **Software**
 - Quattor
 -
- **Network QoS**
- **Storage solution**
- **Virtualization**
- **Regional VO**

Became a T2 in 2008

CMS and Alice VO

Capacity history:

2005:	4 cores,	0,2 To
2006:	48 cores,	14 To
2007:	48 cores,	54 To
2008:	680 cores,	180 To
2009:	1024 cores,	300 To
2010:	1216 cores,	550 To
2011:	1216 cores,	700 To

Other supported VO:

- ILC
 - Biomed
 - Renabi
 - AGATA (+EGEODE)
- Local VO users VO: Ramses, imabio, spiral2
 Regional VO : vo.grand-est.fr



Network

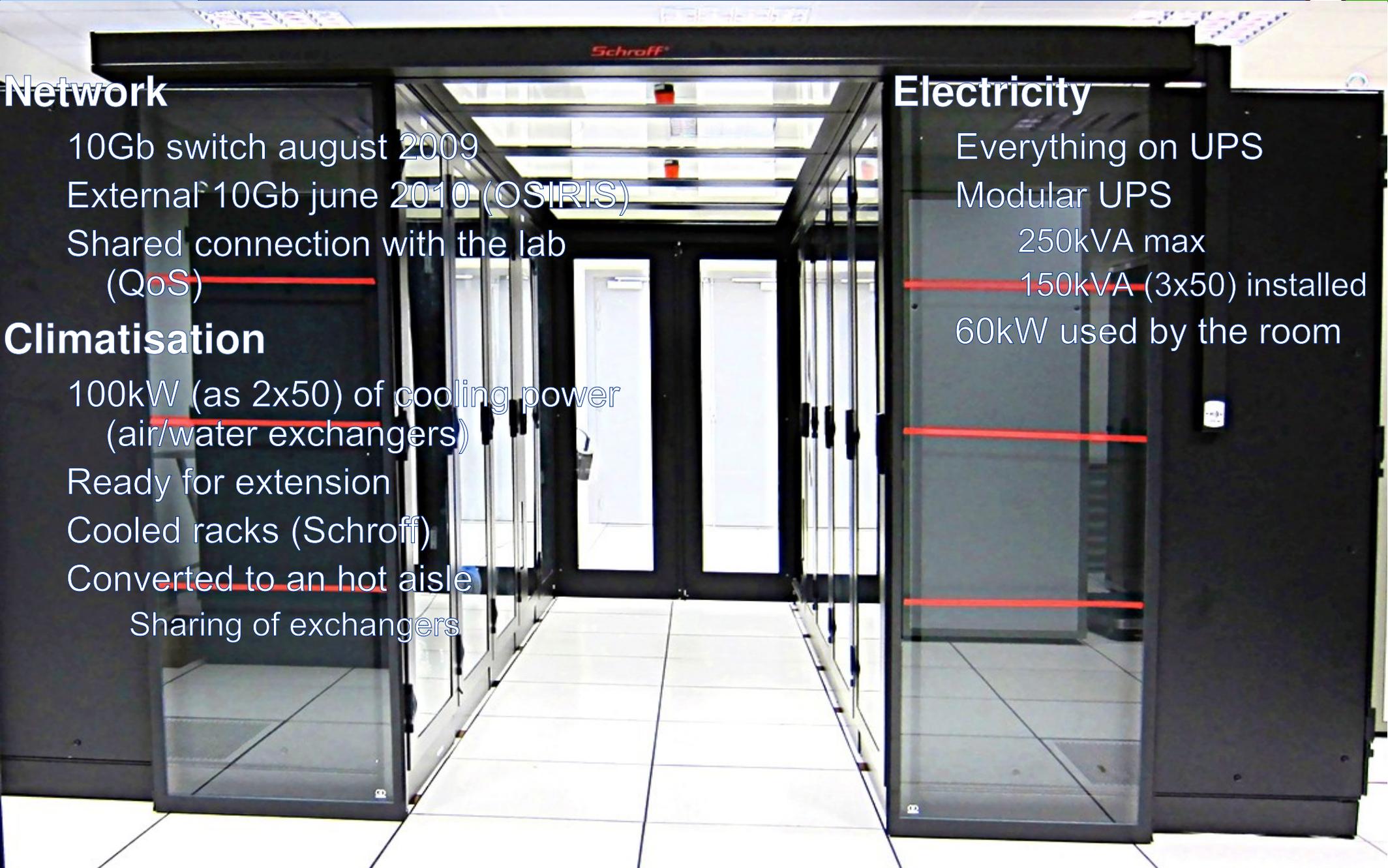
- 10Gb switch august 2009
- External 10Gb june 2010 (OSIRIS)
- Shared connection with the lab (QoS)

Climatisation

- 100kW (as 2x50) of cooling power (air/water exchangers)
- Ready for extension
- Cooled racks (Schroff)
- Converted to an hot aisle
- Sharing of exchangers

Electricity

- Everything on UPS
- Modular UPS
- 250kVA max
- 150kVA (3x50) installed
- 60kW used by the room



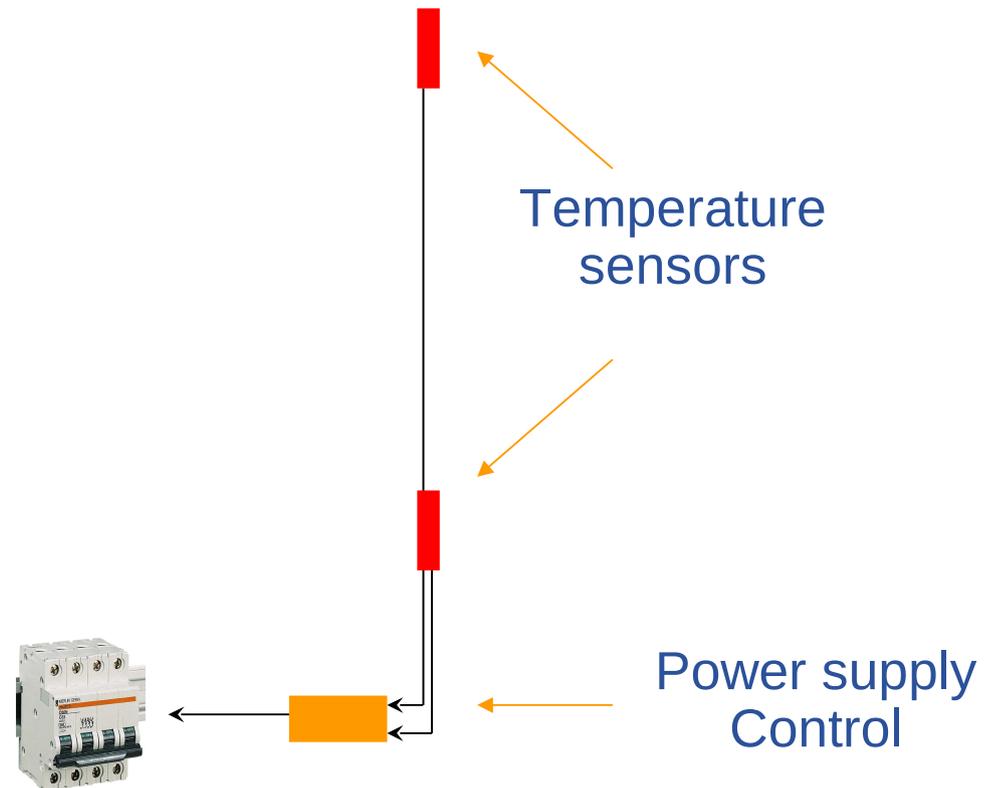
Automatic shutdown

- Two Temp sensor by rack
- Threshold detection (confirmation)
- Steps:
 - DNS warning
 - Pooled by a python script
 - Command shutdown
 - Hardware shutdown

Fire extinction

Inert gaz (installed 2011)

Baie





logged in as ypatois | [Logout](#) | [Preferences](#) | [Help/Guide](#) | [About Trac](#)

- Wiki
- Timeline
- Roadmap
- Browse Source**
- View Tickets
- New Ticket
- Search

[Last Change](#) | [Revision Log](#)

source: @ **4649**

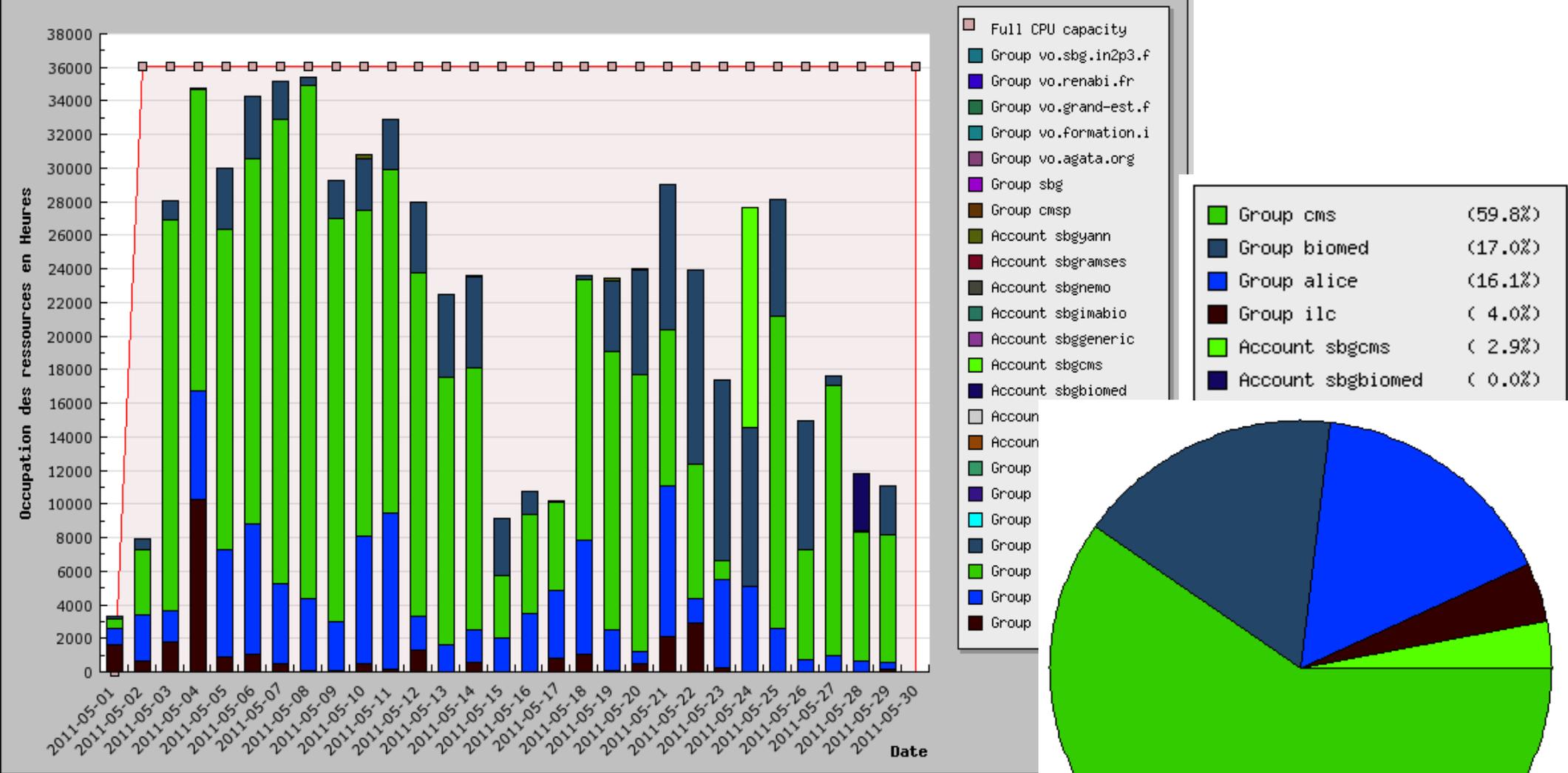
View revision:

Name ▲	Size	Rev	Age	Author	Last Change
▼ scdb		4649	3 days	pansanel	ant tag
▶ branches		2135	2 years	ypatois	Fix for NFS
▶ scdb		2139	2 years	ekieffer	
▶ tags		4649	3 days	pansanel	ant tag
▼ trunk		4648	3 days	pansanel	fix drbd config
▶ clusters		4648	3 days	pansanel	fix drbd config
▶ grid		4630	3 weeks	pansanel	Minor spelling fix
▶ os		4611	3 weeks	pansanel	Update to the las

Très content de ce choix

Jérôme Pansanel fortement impliqué dans le développement de plusieurs templates Quattor.

Occupation des ressources par jour et par groupe entre le 2011-05-01 et le 2011-05-30



Name	Ratio %	For 1Gb	Renater	DSCP	CoS
Supervision	1%	10	0	48 (CS6)	7
Téléphonie	2%	20	0	46 (EF)	5
Visioconf.	3%	30	34 (BBE)	34 (AF41)	4
Défaut	50%	500	0	0 (CS0)	0
Grille Ctrl.	4%	40	8 (LBE)	13 (AF12)	2
Grille Data	40%	400	8 (LBE)	14 (AF13)	1

- **QoS is really useful**

- First introduced to solve local problems:
During heavy WN to SE connections, service machines (including CE) couldn't access the WN.
- Very useful when having only 1Gb to the outside: even during full Phedex download, labs activities were unaffected.
- It helps to make full use of the available bandwidth.

Our historical setup:

Sun X4500 given by CC

Low density

Reliable (after fixing XFS issues on SL5)

AIC RSC-5D / Supermicro

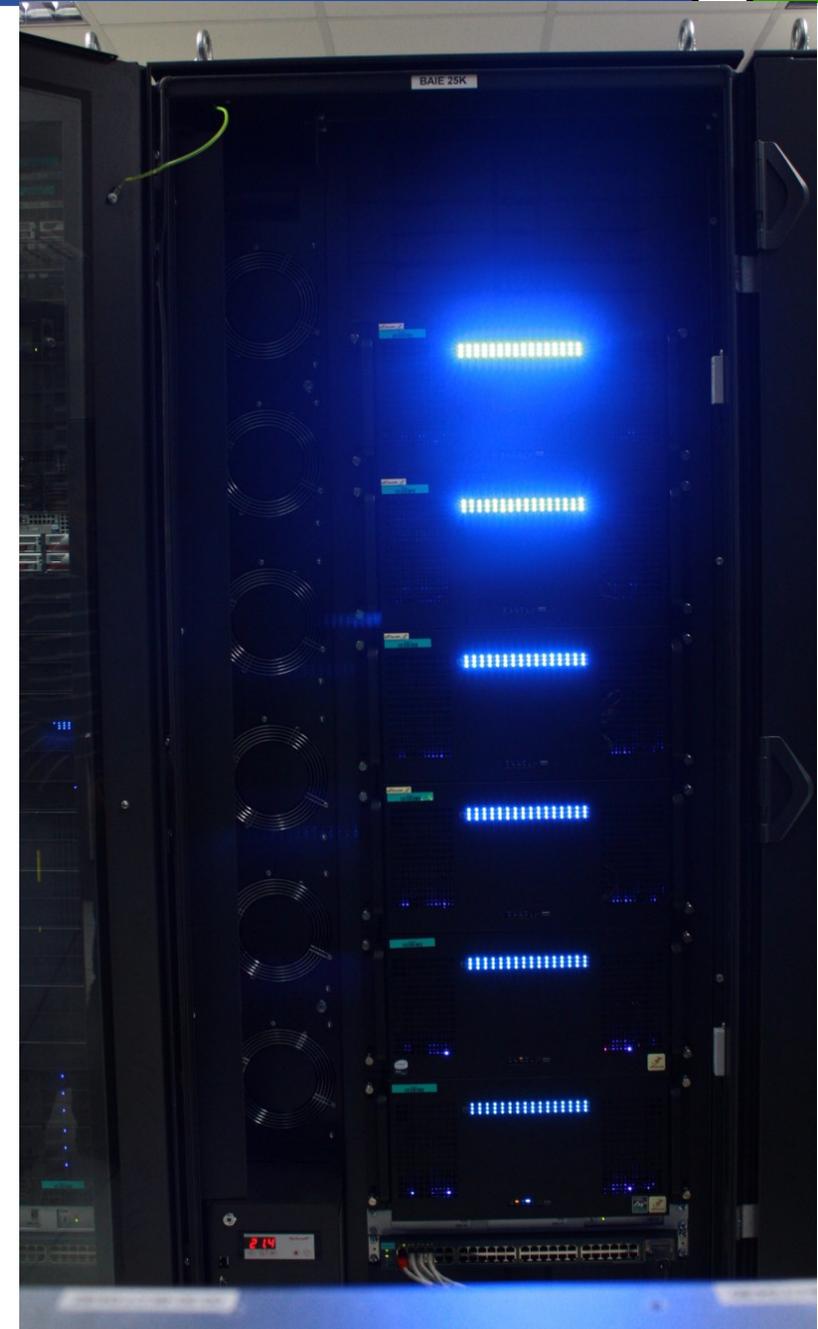
Very low price (0.35€ / Go in 2008)

But unreliable dealer

Cheap desktop quality disks

But pretty fast solution (each added machine comes with it's interfaces)

Availability problem: as storage is spread across machines, a single machine down break the whole system. Unavoidable statistical effet.



c7000 HP enclosure
BL460c HP blades

10Gb network switch



10Gb network switch



3Gb SAS switch



3Gb SAS switch



2x35 disks

Jérôme Pansanel

- **Performances (per logical unit)**
 - Read: 355 MB/s (100 MB/s before optimization)
 - Write: 185 MB/s (185 MB/s before optimization)
- **Optimization**
 - Small 5 disks RAID5 units wisely spread over the MDS600
 - Kernel optimization
- **Specs**
 - HP c7000 enclosure
 - 5 x HP BL460c blades
 - 3 MDS600 (68x2 TB MDL SAS + 2x500 GB S-ATA)
 - Up to 6x10Gb uplinks
- **Advantages**
 - Reliable
 - Easy to manage (iLO, HP tools for Linux)
 - Easy to setup (once a first disk configured)
 - Mac Address Flex Mapping
 - Affordable price (less than 0.50 €/GB)
 - Modular

sbgse1
sbgce1
sbggridnfs1

sbgwms1	sbgwms3
sbglfc1	sbgce2
sbgvoboxcms	sbgvoboxalice
sbggridsrv1	sbggridsrv2

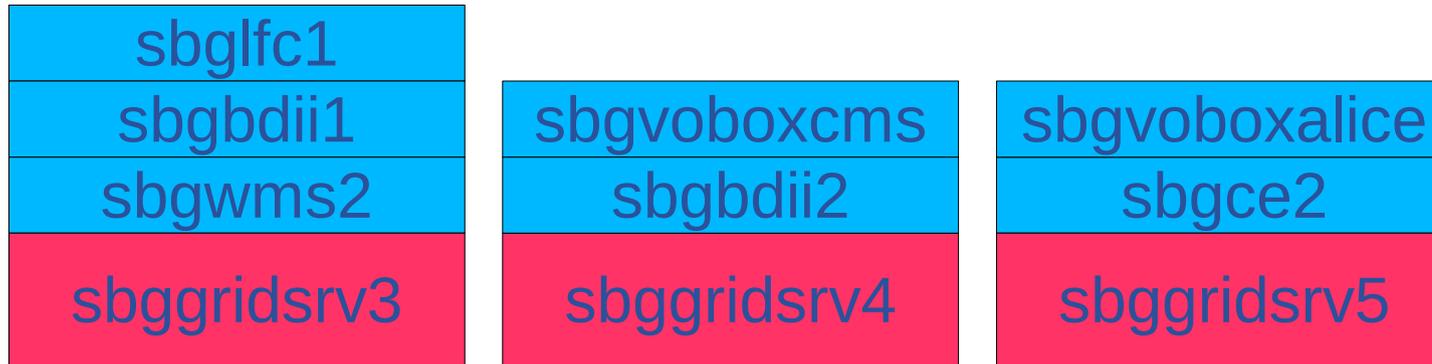
sbgwms2

Physical machines

Xen virtualization

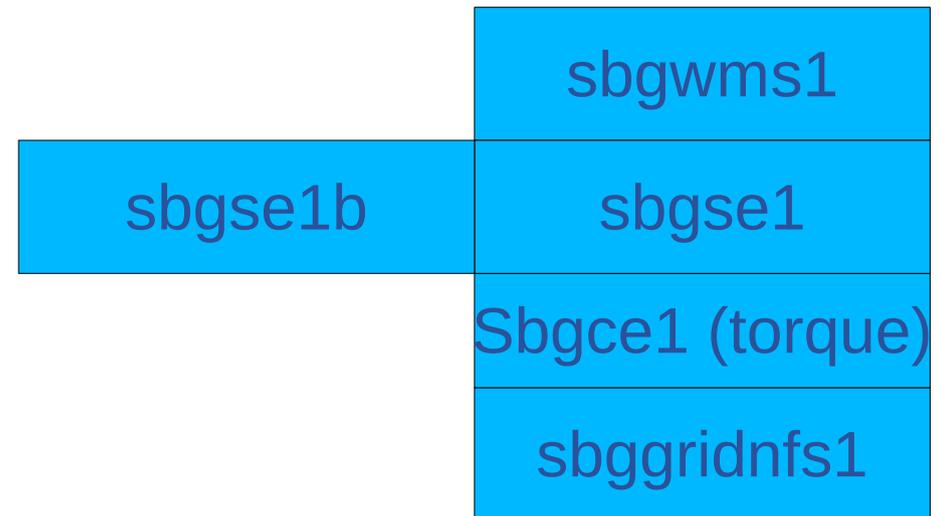
...see next slide

- **Physical machines**
 - Most of the time not needed
- **Xen virtualization**
 - Based on two 2U machines
 - Works well but not flexible enough



KVM virtualization

- **KVM**
 - No need for a dedicated kernel
 - Native support by RH (and derivatives)
- **DRBD**
 - CE, WMS
- **MySQL replication**
 - LFC, SE1 (on se1b)



Physical machines

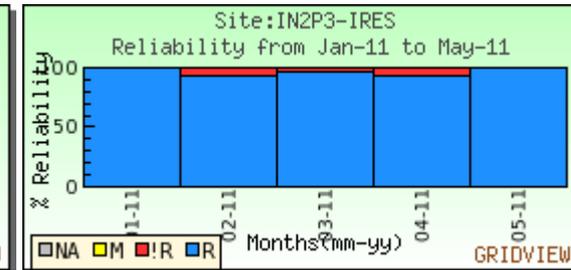
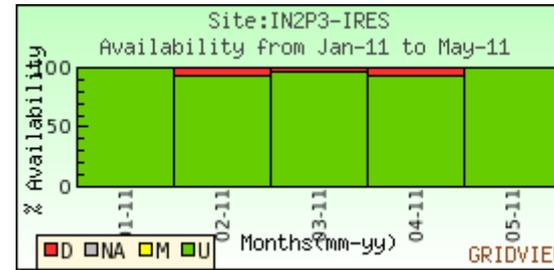
sbgwms3

Elsewhere (see next slide)

- **Several laboratories from Strasbourg and area participates:**
 - IBMP, IBMC, Hopital de Strasbourg (Genetic, molecular biology)
 - Institut de Physique du Globe (CGG Veritas)
 - Faculté de Chimie and Faculté de Pharmacie (modélisation moléculaire et criblage virtuel)
 - LSIIT (cloud computing)
 - IPHC (Imabio, Plateforme Protéomique)
 - Université de Franche-Comté – Laboratoire d'Informatique – Mesocentre (radioprotection, and more to come)
- **Sharing of ressources**
 - Computer centers (hardware & softwares)
 - IPHC
 - Mesocentre Université de Franche-Comté (maybe a grid node in the future).
 - IGBMC (small cluster/storage – in project – at least some services (wms))
 - Skills
 - Administration
 - development
 - User support
 - Access to the world grids
 - <http://www.grand-est.fr> (documentation and resources)

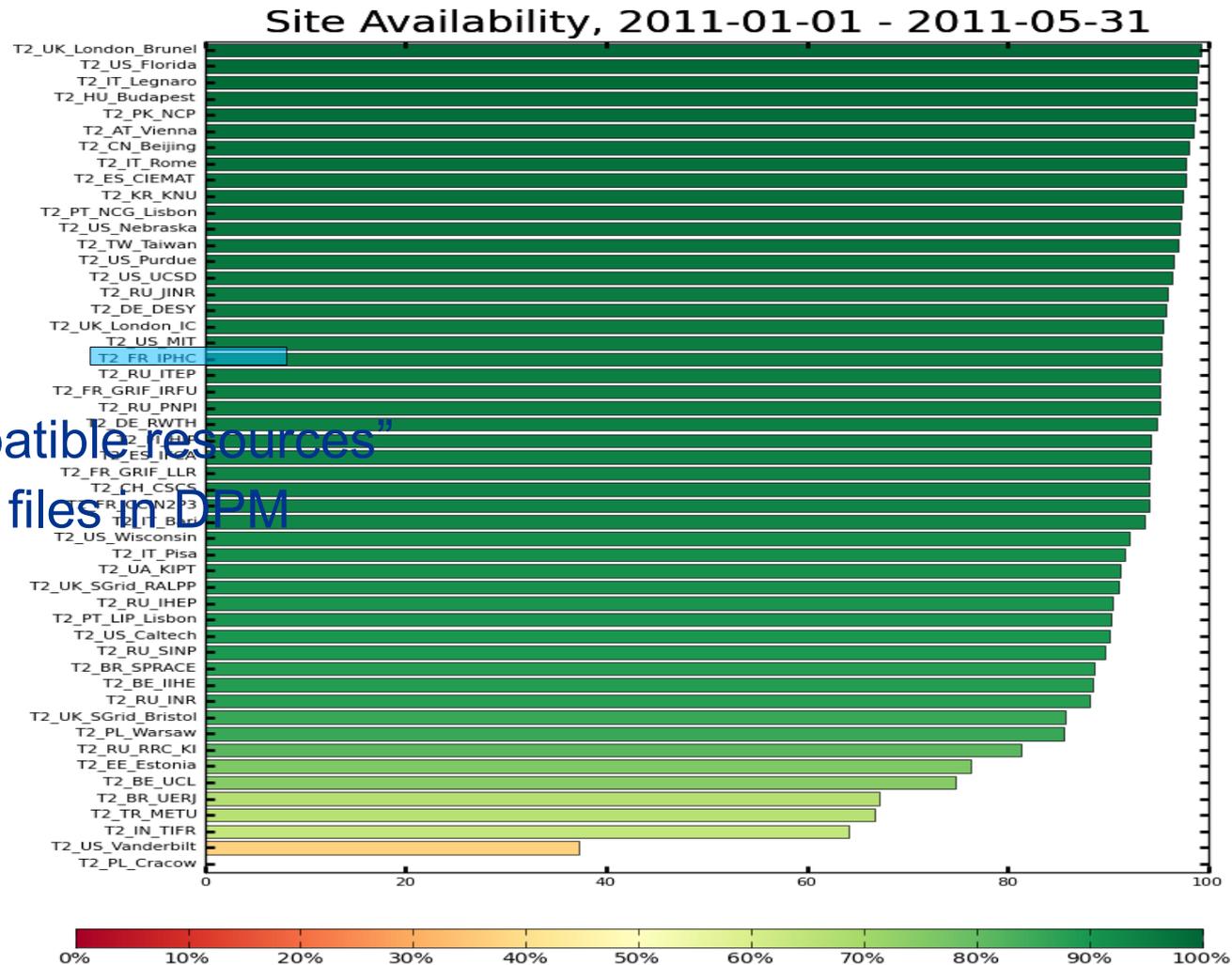
- **Present**

- The site is working well
 - Disponibility
 - Reliability
 - Usage



- **Problems**

- Storage (solved)
- BDII / WMS “No compatible resources”
- RFCP creates volatile files in DPM



- **Future**

- Build on sound foundations
 - Technical infrastructure
 - *Climatisation upgrade*
 - Efficient grid team
- Continue with LCG
- Improve diversification (from 10-15% today to up to 50%)