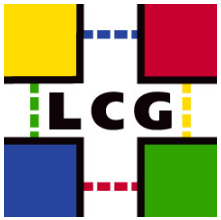
Swiss LCG Tier-2 Center at CSCS



Derek Feichtinger, PSI

on behalf of the

Swiss Institute of Particle Physics (CHIPP)

Personnel

People directly involved with Tier2 administration:

- CSCS
  - Peter Kunszt (SM Swiss Grid Initiative)
    - *joined Feb 06* from EGEE JRA1
  - Vincenzo Annaloro (Hardware + OS)
  - Gian-Luca Volpato
    - our principal LCG administrator *left CSCS in Nov 05*
  - searching for an additional Grid Expert to work on cluster and SW
- ATLAS
  - Frederik Orellana, Uni Ge and LCG/ARDA
- CMS
  - Derek Feichtinger, PSI
    - *returned in Dec 05* to PSI after 3 years in LCG (ARDA)
- LHCb
  - Roland Bernet, ETHZ

Many mutations

Current LCG Cluster

- 1 **CE**
- 1 **SE** (classical SE, 8TB RAID5)
- 15 **WN**s (30 CPUs)
- 1 **UI**
- 1 CMS **VO-Box**
- 1 Installation server / experim. SW server
- 1 Nordugrid Front End

Node Architecture:

Intel dual-Xeon 3.0 GHz, 4 GB RAM

Intel Pentium4 3.0 GHz, 1 GB RAM

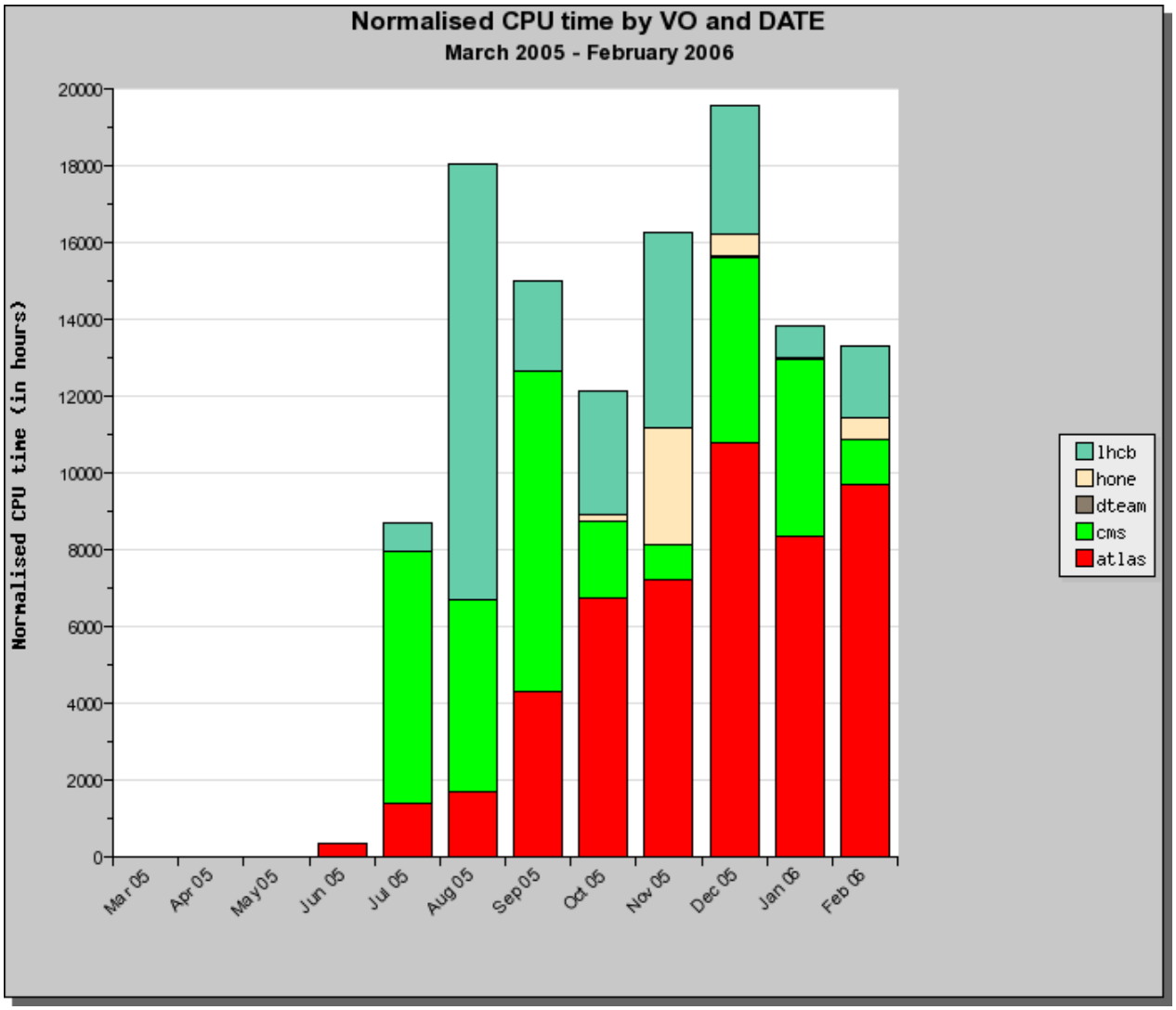
45 kSI2000

(1 Xeon 3.0 GHz ≈ 1500 kSI2000)

CMS VO-Box

- since January 2006 running PubDB and PhEDEx with the same local MySQL backend.
- PubDB
  - currently only publishing Data received by PhEDEx
  - publishing step needed lots of tailoring and is error prone – still done interactively
  - CRAB jobs work fine
- PhEDEx
  - receive data directly from CERN (castorgrid)
  - issue with certain zipped data sets and POOL limit of 250 *char* PFNs

Cluster usage (LCG)



- Atlas
- CMS
- H1
- LHCb

- idle cycles were opportunistically used by H1 (lesser priority queue)

Nordugrid

- Nordugrid front end mainly used by Swiss ATLAS community
  - ease of deployment (less manpower needed)
  - non intrusiveness (sites need not change their Linux distribution)
  - functionality sufficient for current user requirements

2006-03-01 CET 11:05:40

Processes: Grid Local

Country	Site	CPUs	Load (processes: Grid+local)	Queueing
Switzerland	Bern ATLAS Cluster	16	15+0	4+0
	Geneva-DINF/DPNC	25	0+0	0+0
	Geneva-DPNC	10	0+0	0+0
	PHOENIX (CSCS)	36	0+30	7+0
	UBELIX (University of>	288	0+216	0+4
TOTAL	*5 sites*	*375*	*15 + 246*	*11 + 4*

ALL

Software upgrades

- Highest priority
  - **DPM**
    - need to support SRM protocol
    - need to support disk pools to handle large storage space
  - **LCG-2_7_0**
- desired
  - Linux Kernel 2.6.x (desired for cluster management software)
  - **SCL4**
- interested in
  - **Xrootd**
    - SRM interface is in the line
    - good experiences from ARDA/ALICE
    - we currently offer on site access to SE via xrootd
  - **PROOF** – parallel ROOT facility

Cluster Ramp Up Plan

- **Gradual extension of available sources with goal to buy as late as possible.**
- **RFP (Request for Precurment) Document for this year's extension currently in preparation.**
- **No Grid enabled tape storage foreseen! (but local users may use CSCS tape resources**

Ramp up table:

	2005	2006	2007	**2008**
CPU (kSI2000)	45	231	692	**2307**
Disk (TB)	9	79	236	**787**
CPU fraction w.r.t. 2008 (%)	2	10	30	**100**

WAN connectivity to CERN ramp up (Gbit/s):

2005	2006	2007	**2008**
1	1	1-10	**10**

Cluster Design: Storage connectivity

- Inter-WN connectivity not an issue for HEP jobs, but connectivity to storage nodes is important

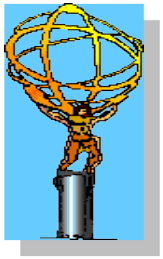
- **What is the needed bandwidth per running job?**
  - ***CMS computing TDR***: ≥32 Mbit/(s·kSI2K) ≈ 6 MB/s for a single Xeon 3GHz CPU
  - questions to experiments people running actual jobs brought up numbers ≈ 2 MB/s per CPU core on current hardware.

- If all parts of the storage need to be accessible from every WN with the stated bandwidth, an expensive network setup is required for big systems.
  - If the middleware can deal with "sub-clusters" one can save a lot!

Who will be the Tier1 for CSCS?

- Had over the last two years oral confirmation of CERN representatives, that CERN will figure as T1 for CSCS
  - good connectivity using the internal Swiss networking infrastructure from SWITCH
  - We get PhEDEx transfers directly from CERN
- However: Many official documents do not mention T1 functionality of CERN, but only the external T1s
  - We cannot clearly see all the consequences of the T1 choice yet.
    - also experiment dependent?
  - Will a T2 only be able to receive data via its T1? Hopefully not. Opportunistic use of routes with free bandwidth should be possible.
  - What persistent Data will a T1 store on behalf of a T2?
  - Already organizationally well integrated in ROC DE/CH.

LCG Contacts



- A.Clark (UNI Ge)
- H.P.Beck (UNI Bern)
- F.Orellana (vice-chair CCB, Uni Ge)

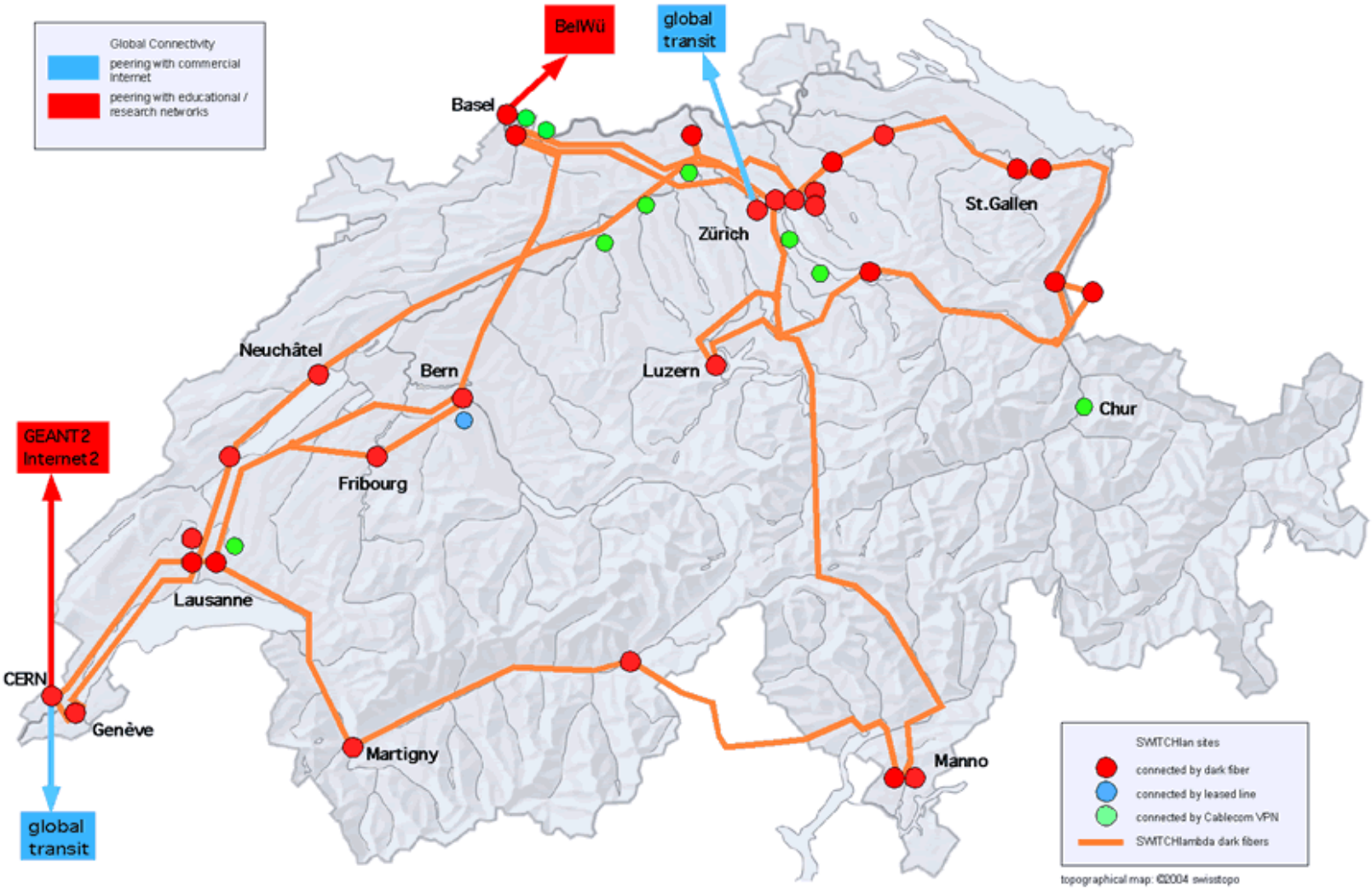
- C.Grab (chair CCB, ETHZ)
- U. Langenegger (ETHZ)
- D.Feichtinger (PSI)

- R.Bernet (UNIZH)
- J. Van Hunen (EPFL)

- M-C. Sawley (CSCS general manager)
- P. Kunszt (CSCS Swiss Grid Initiative)

■ main contacts for T2 related technical issues

EMail Addresses and WWW Links

- CSCS Tier2 Wiki: http://minusio.dmz.cscs.ch/twiki/bin/view/LCGTier2/WebHome
  - Main entry point for information
- CSCS Tier2 mailing list: project-lcg-cscsgrid@cern.ch
- CHIPP (Swiss Institute of Particle Physics): http://www.chipp.ch
  - contains addresses of all members
- CHIPP Computing Board Chair (CCB) C. Grab grab@phys.ethz.ch
- Tier2 CSCS
  - P. Kunszt (CSCS) peter.kunszt@cscs.ch
- Tier2 experimental contacts and admins
  - R. Bernet (LHCb) bernet@physik.unizh.ch
  - D. Feichtinger (CMS) derek.feichtinger@psi.ch
  - F. Orellana (ATLAS) frederik.orellanar@cern.ch

SWITCHlan connectivity