Report from the
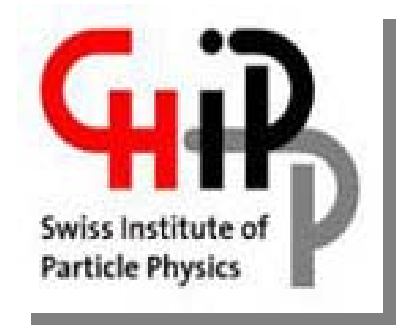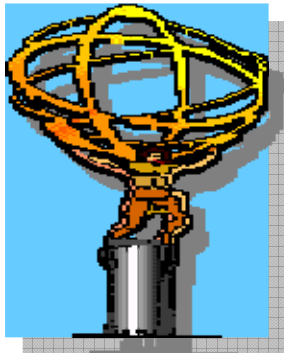# CHIPP Computing Board

Christoph Grab, ETH

Lausanne,   9.9.2008

Status of the Swiss Tier-2 PHOENIX cluster

- **The Swiss Tier-2** is a compute cluster "PHOENIX", with following tasks:

  - **Simulation for experiment's community**

  - **End-user analysis for Swiss community (mainly)**

  - **support (operation and data) for Swiss Tier-3 centres**

- **Incremental setup** in phases (0, A-C); final size be reached in 2009/10:

  - **Total: storage : ~ 1 PB disks ;   CPU : 2500 kSI2k (~1800 cores)**

- **Network traffic**:

  - **routing via SWITCH : two redundant lines to CERN → CERN to Europe**

  - typical numbers reached were 2-5 TB /day   to/from FZK and CERN

- **Financing** (HW and service, no manpower) :

  - **Financing mainly through SNF/FORCE (~90%),** with some contributions by Universities + ETH + PSI

  - **Costs: total investment ~ 3.7 MCHF; operational costs ~700 kCHF/a ≥ 2010**
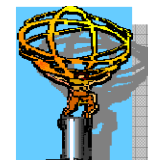
CHIPP ' s   *Swiss LHC Tier-2 at CSCS*

Phase 0

(36 TB, 220 kSI2k)

Phase-A installed in Dec 07

(225 TB, ~800 kSI2k, 400 cores)

PHOENIX
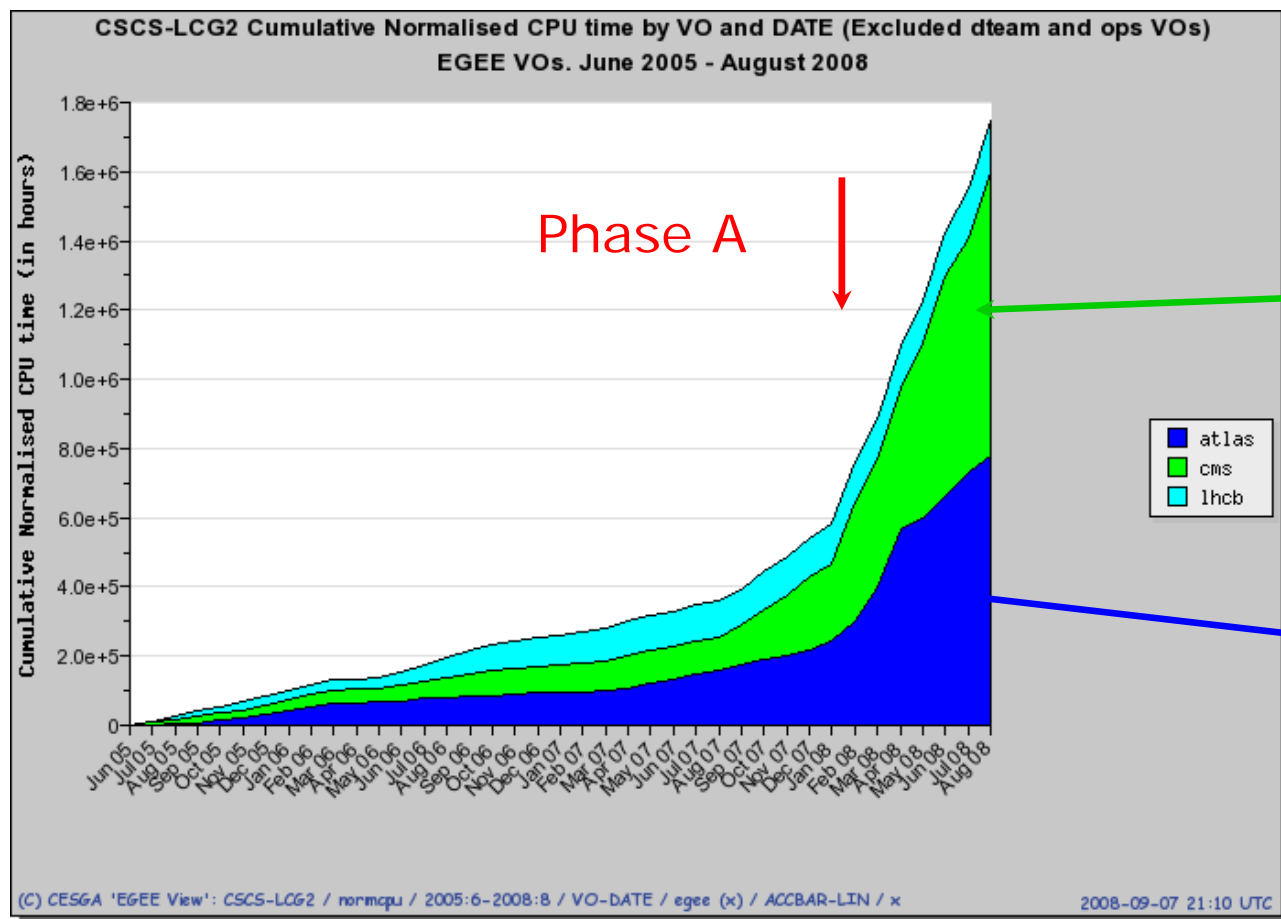
CPU

Storage

Dec '06 – Jan 08

Jan 08 - …

Complete System Phase A operational since Jan 2008

LCG-usage of CPU over last 3 years

CMS

Computing and Analysis Challenges

ATLAS

➔ **Stable operation over ~3 years**
steady contributions of resources to experiments

➢ https://twiki.cscs.ch/twiki/bin/view/LCGTier2/PhoenixMonOverview

Tier-2 Availability and Reliability Report

Federation Summary - Sorted by Name

August 2008

Critical SAM Tests - http://sam-docs.web.cern.ch/sam-docs/docs/htmldocs/MANUserManual/node22.html

Availability = % of successful tests
Reliability = Availability / Scheduled Availability
Reliability and Availability for federation - average of all sites in the federation

Colour coding : N/A | < 30% | < 60% | < 90% | >= 90%

| Federation	Reli-ability	Avail-ability	Federation	Reli-ability	Avail-ability
AT-HEPHY-VIENNA-UIBK	99 %	99 %	JP-Tokyo-ATLAS-T2	98 %	96 %
AU-ATLAS	64 %	59 %	KR-KISTI-T2	87 %	86 %
BE-TIER2	90 %	87 %	NO-NORDGRID-T2	N/A	0 %
CA-EAST-T2	69 %	67 %	PK-CMS-T2	79 %	78 %
CA-WEST-T2	94 %	89 %	PL-TIER2-WLCG	90 %	64 %
CH-CHIPP-CSCS	99 %	97 %	PT-LIP-LCG-Tier2	94 %	93 %
CN-IHEP	99 %	98 %	RO-LCG	96 %	95 %
CZ-Prague-T2	94 %	94 %	RU-RDIG	91 %	91 %
DE-DESY-ATLAS-T2	99 %	98 %	SE-SNIC-T2	N/A	N/A
DE-DESY-RWTH-CMS-T2	98 %	98 %	SI-SiGNET	87 %	84 %
DE-FREIBURGWUPPERTAL	96 %	96 %	T2_US_Caltech	88 %	90 %
DE-GSI	0 %	0 %	T2_US_Florida	92 %	92 %
DE-MCAT	92 %	91 %	T2_US_MIT	98 %	98 %
EE-NICPB	84 %	84 %	T2_US_Nebraska	94 %	94 %

Swiss Tier-2 : Cluster Evolution

Growth corresponds to Swiss commitment in terms of compute resources supplied to the expt's according to the signed MoU with WLCG.
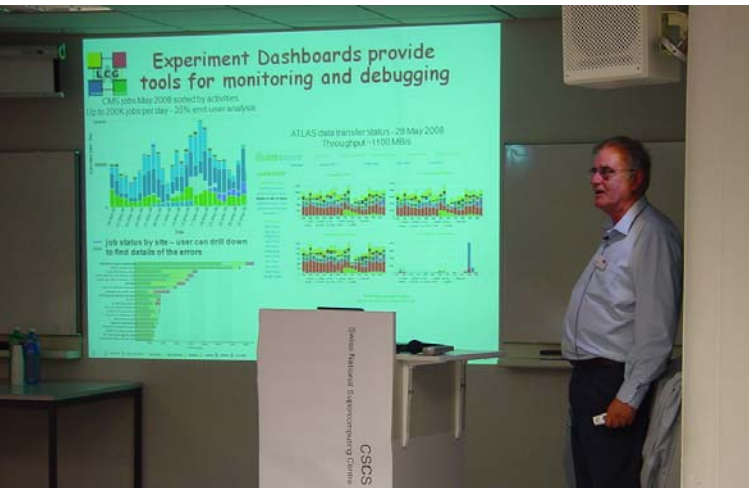
Phase A: operational since Jan.08 (following pre- phase-0)
Phase B: public tender done → order out on 2.9.08 → delivery ~Oct.08
Phase C: apply for funds on < 30.9.08

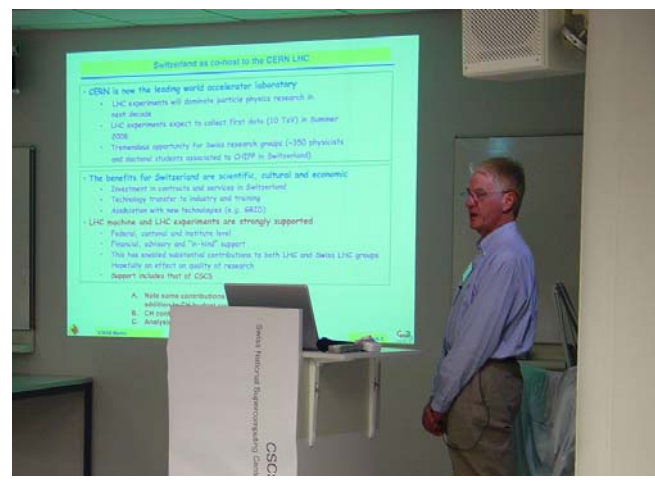Swiss Tier-2   Size Evolution
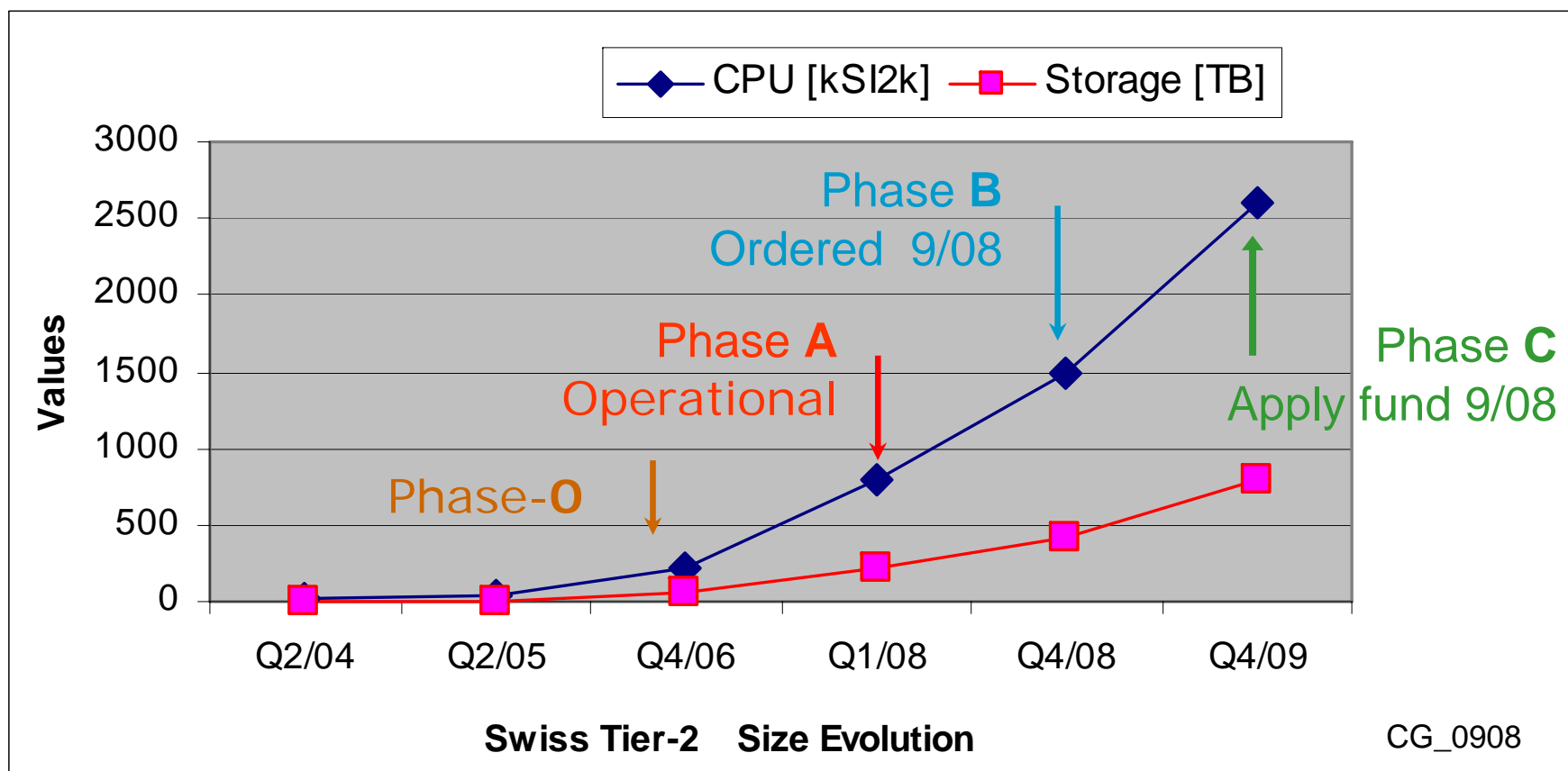
CG_0908

Grid : L.Robertson

LHC: D.Barney

Swiss contributions to LHC: A.Clark

Visit to Computer Room
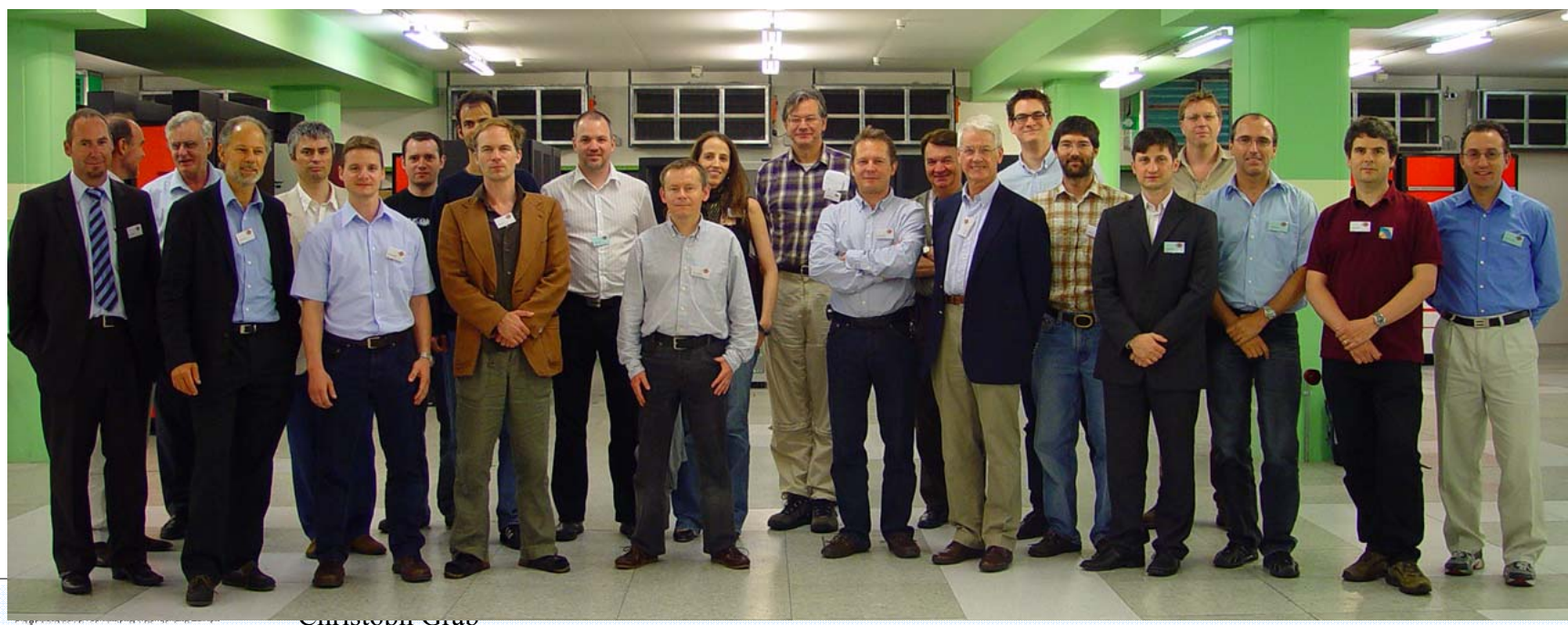
Status Swiss Tier-3 Efforts

- Large progress seen ...

- **ATLAS** : Swiss ATLAS Grid → operates a federation of clusters at
  - Bern uses local HEP + university resources
  - Geneva operates local cluster

- **CMS** :
  - uses local resources
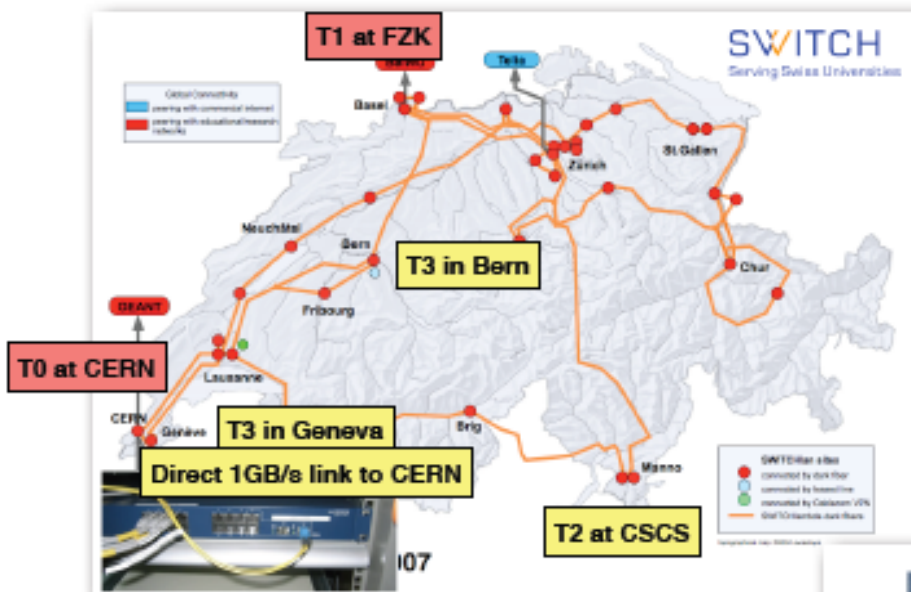 - setup a combined Tier-3 for ETHZ+PSI+UZH, done

- **LHCb**
  - UZH uses local HEP + university resources,
  - EPFL : operates local cluster

Covered by S.Haug

ATLAS Data from CERN (T0) via FZK (T1) and CSCS (T2) to Bern and Geneva

QUE - 2008-09-08

The Swiss ATLAS Grid
In 2008 ~400 cores and ~200 TB disk in four clusters

* For ATLAS data storage, simulation and analysis.
* Two T3s and one T2 (in cooperation with CHIPP CB,CSCS).
* Since Jan 2008 operation organized in SwiNG WG.
  S. Haug (present lead), S. Gadomski, C. Topfel.
  www.swing-grid.ch

See also CHIPP Computing Board report - C. Grab.

Expect use of Geneva T3 for commissioning and Data
Quality work on the ATLAS Trigger

* access to latest software via AFS (in place)
* access to latest data as at CERN (in progress)

If this use goes ahead, will be effectively extension of
the CERN Analysis Facility, no contradiction with the
Tier 3 role, which will continue planned flow of data:

* Commissioning: subsets of latest data from CERN.
* Analysis: data derived from large samples, from Grid.

Sigve Haug on behalf of DPNC Geneva and LHEP Bern - ATLAS/CH at First Beam - 7/11
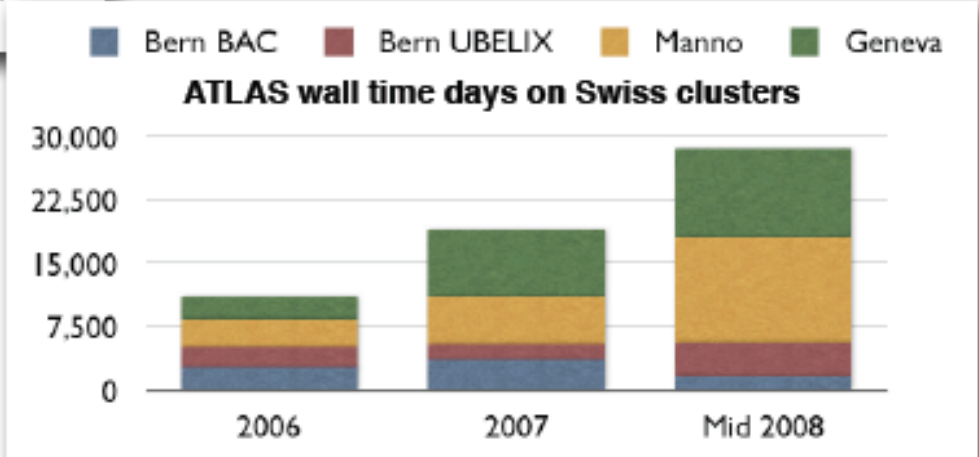
ATLAS Tier3 in Geneva

Hardware in production

- **61 computers**
  - 53 workers, 5 file servers, 3 service nodes
- **188 CPU cores in the workers**
- **75 TB of disk storage**

Upgrade plans

- **double the storage during next few months**

Advantages of the Geneva Tier3

- **environment like at CERN, latest ATLAS software via AFS**
- **direct line to CERN**

popular with ATLAS physicists (53 users)

S. Gadomski

Usage of the Geneva Tier3

- a grid site since 2005
  - mostly ATLAS Monte Carlo production
- a local resource for the group since 2007
  - development and testing of trigger code
  - final analysis
  - local batch
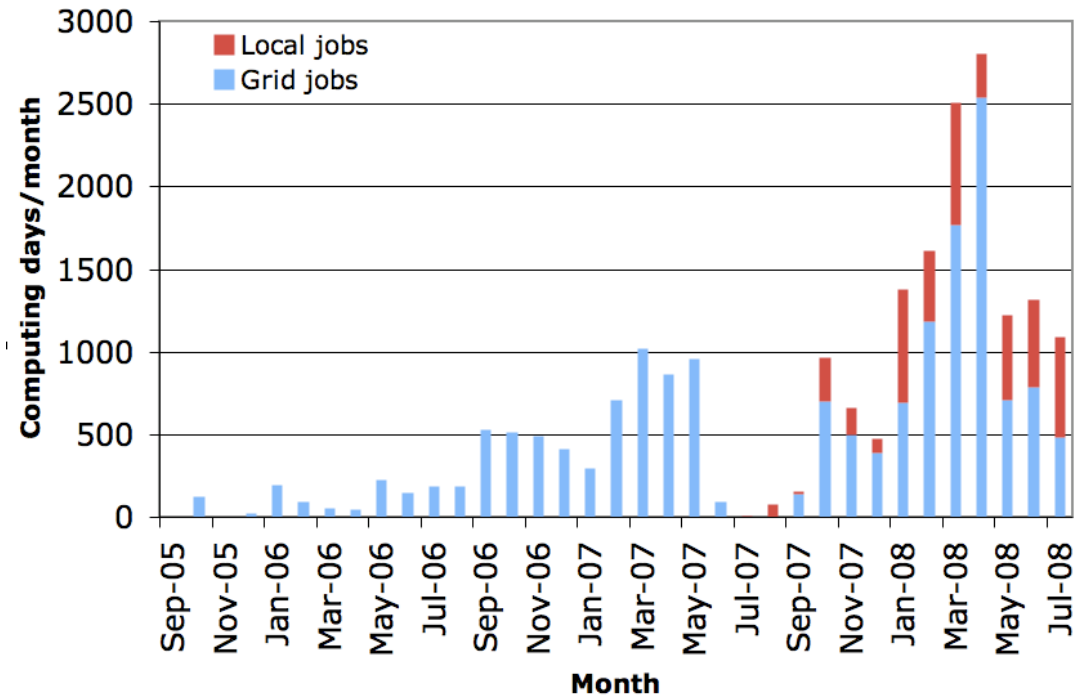
planned involvement in ATLAS commissioning and Data Quality work, particularly for the Trigger
- formal commitment to the Collaboration
- latest software, data access like at the CERN CAF

S. Gadomski

Status ATLAS - Bern T3

Two clusters with NorduGrid front ends in production since 2005. One cluster is shared and operated by "Informatik Dienste UniBE". For local physics analysis and simulation. Fills up with ATLAS central production jobs when not used by locals.

Hardware

~ 130 cores for ATLAS (~250 in 2009).
~ 33 TB disk (~ 50 TB in 2009).

Usage/Operation

~ 5000 Wall Time Days in 2006.
~ 7000 Wall Time Days in 2007.
~ 6000 Wall Time Days in Jan-Jun 2008.

S. Haug, Bern 2008-09-01

Status CMS – Tier 3 at PSI

- **Presently operating**: CMS-nodes at ETH (SD) [40 kSI2k, 5 TB].

- **Common CMS Tier-3:** **ETH, PSI, UZH**
  similar architecture as Tier-2
  HW is setup at PSI; users start mid Sep.

Year	2008	2009
CPU / kSI2k	**215**	500
Disk / TB	**105**	250
No of Worker Nodes	8	28
No of CPU Cores	64	224
No of Storage Nodes	6	14
Racks	1	3

Connection of 1 Gbps <-> PSI

D.Feichtinger+C.Grab

➢Tier-3 Wiki: https://twiki.cscs.ch/twiki/bin/view/CmsTier3/WebHome

CMS Tier-3 Hardware at PSI

Technical Details …

Common CMS Tier-3 for ETHZ, PSI + UNIZ

- **8 Worker Nodes**
  - SUN X4150, 2*Xeon E5410, 16 GB RAM, 2*146 GB SAS disk

- **6 Service nodes**
  - SUN X4150 servers adapted to services
  - CE, SE, Data Base, UI, NFS, Admin

- **6 File Servers**
  - SUN X4500, 2*Opt 290, 16 GB RAM, 48*500 GB SATA disk
  - Connected with 4*1Gb/s

No. of WNs	Processors	Cores/node	CINT2006/core	kCINT2000/core	No. of Cores	CINT2006	kCINT2000
8	2*Xeon E5410	8	18.8	3.34	64	1203.2	213.76

No. of Fileservers	Type	Space/Node (TB)	Total Space (TB)	
6	SUN X4500	17.5	105	(RAID5 space)

Tier-3 Wiki: https://twiki.cscs.ch/twiki/bin/view/CmsTier3/WebHome

Dcache Storage

Tier-3

Transfers (gsiftp)

Worker Nodes

Transfers (dcap, gsiftp)

The Grid

PhEDEx Data Set Transport

Grid Job

submits

submits

Batch master (SGE)

connects

User Interface

User
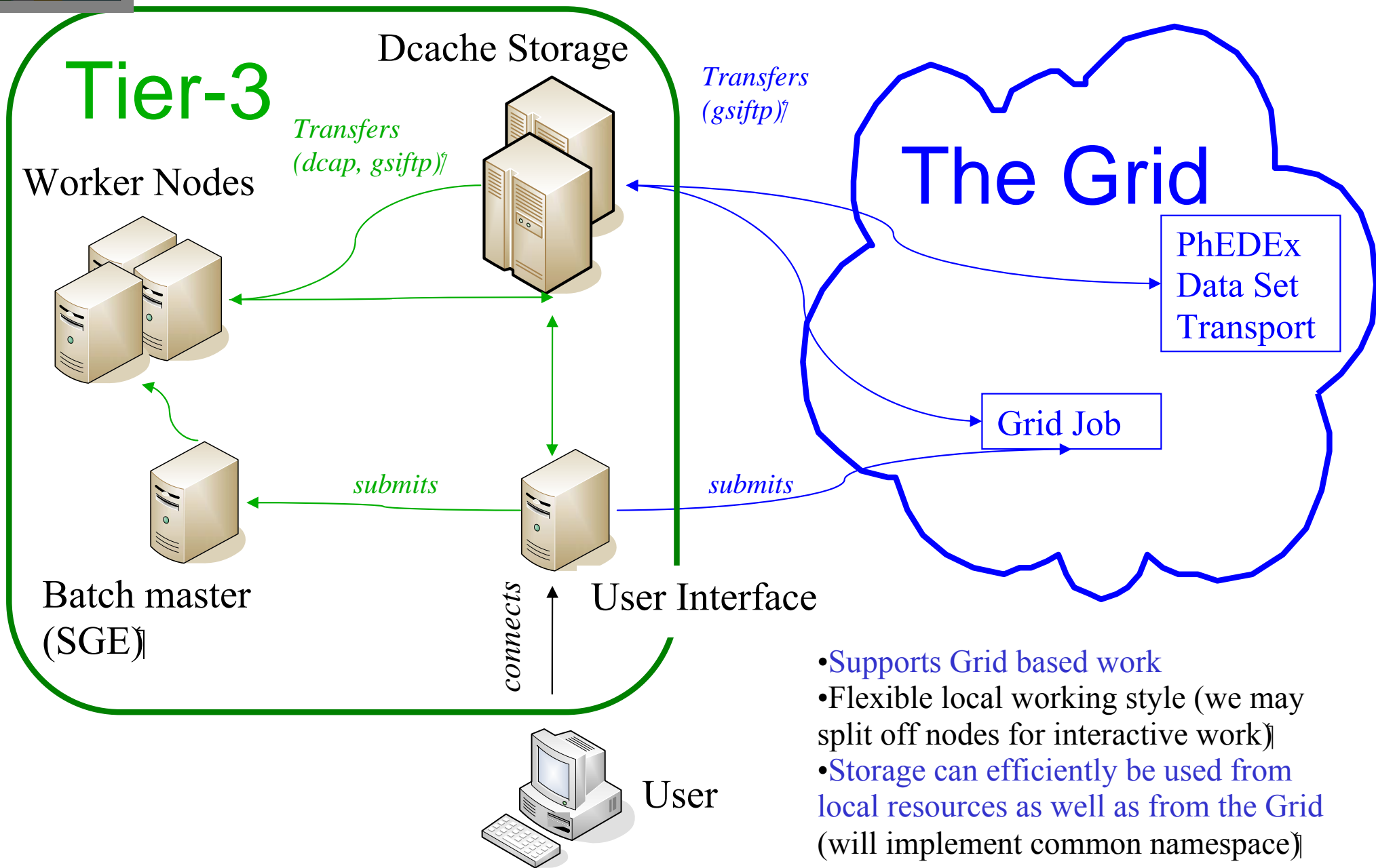
- Supports Grid based work
- Flexible local working style (we may split off nodes for interactive work)
- Storage can efficiently be used from local resources as well as from the Grid (will implement common namespace)

Tier-3 Status – Zurich

LHCb

- **Zurich HEP Cluster:**
  - Intel Cluster, 32-bit, openSUSE
  - DIRAC installation running SLC4 binaries
  - Local openSUSE LHCb Software Installation → local LHCb jobs
  - Hardware: CPU: 125 kSI2k,   Disk: ~15 TB
  - Concerns: File Access to RAID might be bottle neck

- **Zurich Matterhorn Cluster:**
  - AMD Opteron Cluster, 64-bit, SuSE
  - DIRAC installation running SLC4 binaries
  - Only used for Monte Carlo Production
  - Used by other institutes, load varies a lot
  - Currently no big disk allocation, would be available, if needed

Hardware:

- Intel Xeon Cluster, 64-bit, SL4 (Beryllium)
  - 60 worker nodes with $2 \times$ quad-core (2.5 GHz)
  - Current LHCb software installed
  - $6 + 30$ Tb of storage in RAID5 configuration
- Machines identical to those in the LHCb pit

Software:

- SLC4 binaries of recent LHCb software

Current Status:

- Intending to install SLC4 and DIRAC in the near future
- New cluster $\rightarrow$ few hardware teething problems (fixed)

Some Comments ...

- **User comments on the experiences with analysis on the Tier-2:**

  - Overall the users are very satisfied with usage of T2 at CSCS
    - high job success (80-100%), short queues, enough user disk space …

  - criticism of users (mostly CSCS independent) :
    - would like to have less downtimes (GRID enforced),
    - more adequate management of queues

- **Manpower situation : words of caution** …

  - cluster is growing in size and complexity;
    maintaining high quality operations requires skilled personnel.

  - **We NEED longterm manpower redundancy for operation at CSCS**
    - **Missing effectively ONE FTE at CSCS !!!**
    - **Just about to lose one of the experts** [only 1-year contract (S.M.)

  - We need to have more active support of CHIPP persons for Tier-2 operation:
    - we have ONLY <2 CHIPP persons actively supporting CSCS (DF, ZC) !

- ## Swiss Tier-2 cluster PHOENIX:

  - operates stably since Q2/2005 ☺

  - reliably delivers the Swiss pledges to the LHC experiments in terms of computing resources;

  - Growth in size as planned, final size reached end 2009/ early 2010; Phase-A in operation; Phase-B progressing.

    - Technically most challenging and intense "daily care" is needed for SW of middleware → Relies on dedicated long-term personnel !!!

  - This is a pivotal link of the Swiss HEP/LHC community to the LHC experiments and to CERN, and crucial for physics exploitation !

- ## Local Tier-3 centres made remarkable progress

  H→WW→2μ2ν

  We are prepared for PHYSICS !

  CMS

CHIPP Computing Board
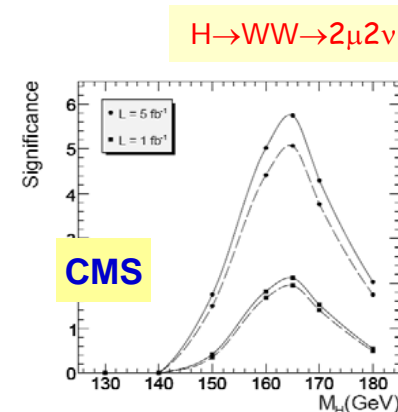
Coordinates the Tier-2 activities
representatives of all institutions and experiments

A.Clark, S.Gadomski (UNI Ge)
H.P.Beck, S.Haug (UNI Bern)

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

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

P. Kunszt (CSCS)