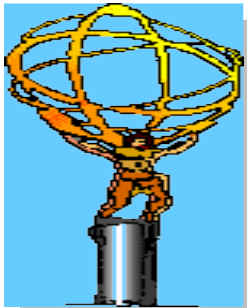


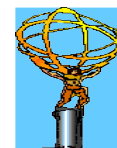
# Computing Board Report



Christoph Grab



# Status of the Swiss Tier-2 PHOENIX cluster



CHIPP runs its own **Swiss LHC Tier-2 at CSCS**

## Phase 0

(36 TB, 220 kSI2k)



Dec '06 – Jan 08

## Phase-A installed in Dec 07

(225 TB, ~800 kSI2k, 400 cores)



Jan 08 - ...

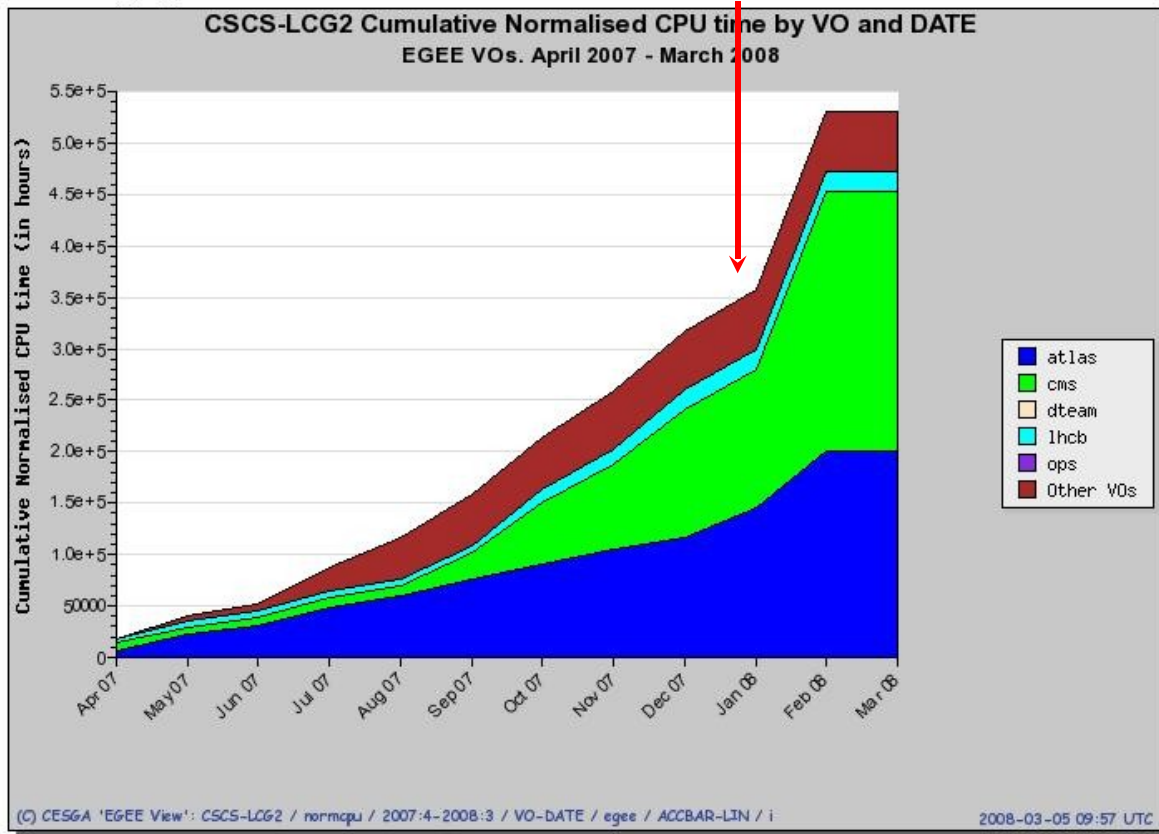
**PHOENIX**



**Complete System Phase A operational since Jan 2008**

shown is **LCG-usage over last 1 years**

Phase A



Rest (not shown) are

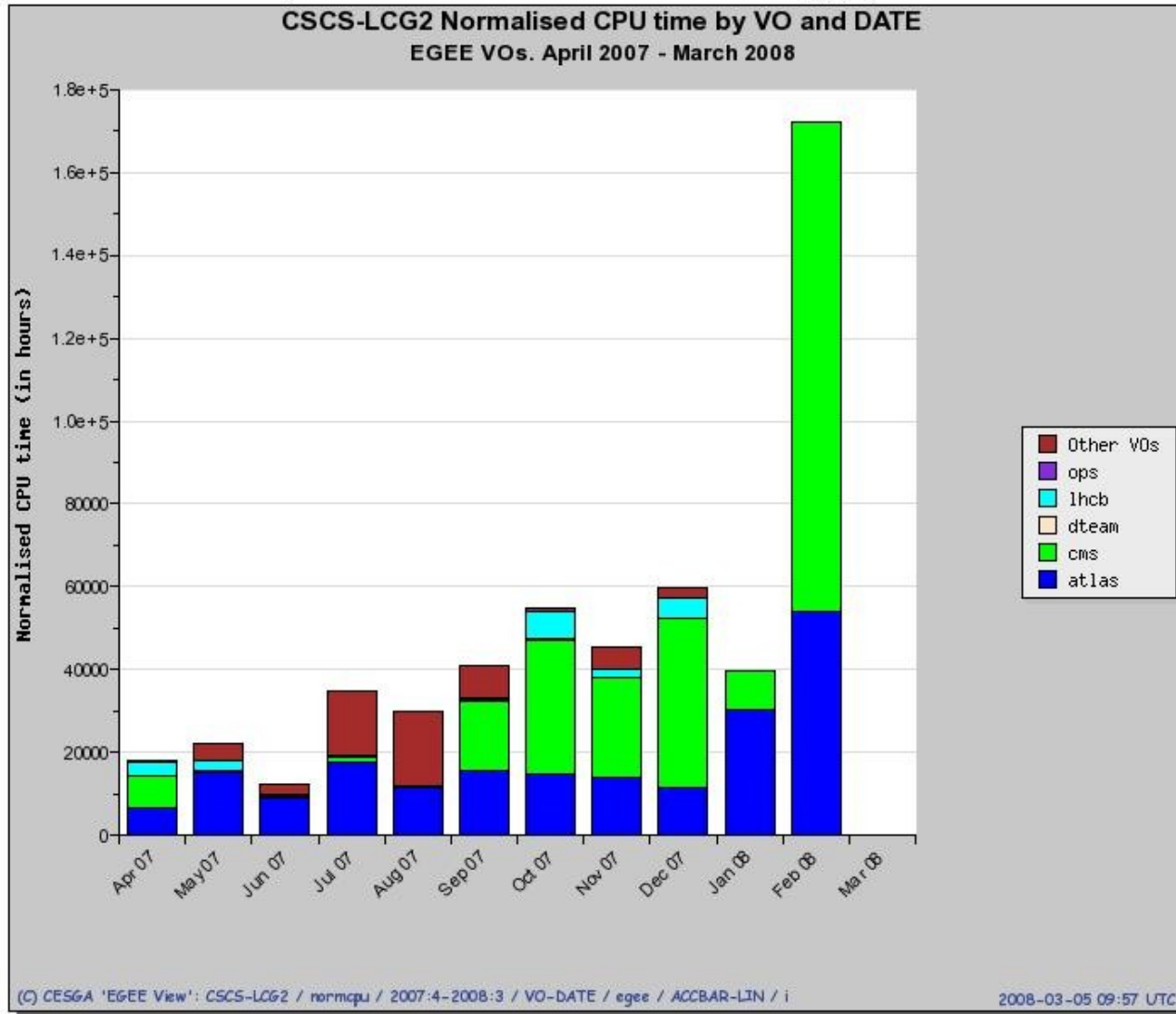
- NorduGrid + user jobs

- High efficiency for user analysis jobs

- plots available via Phoenix Wiki

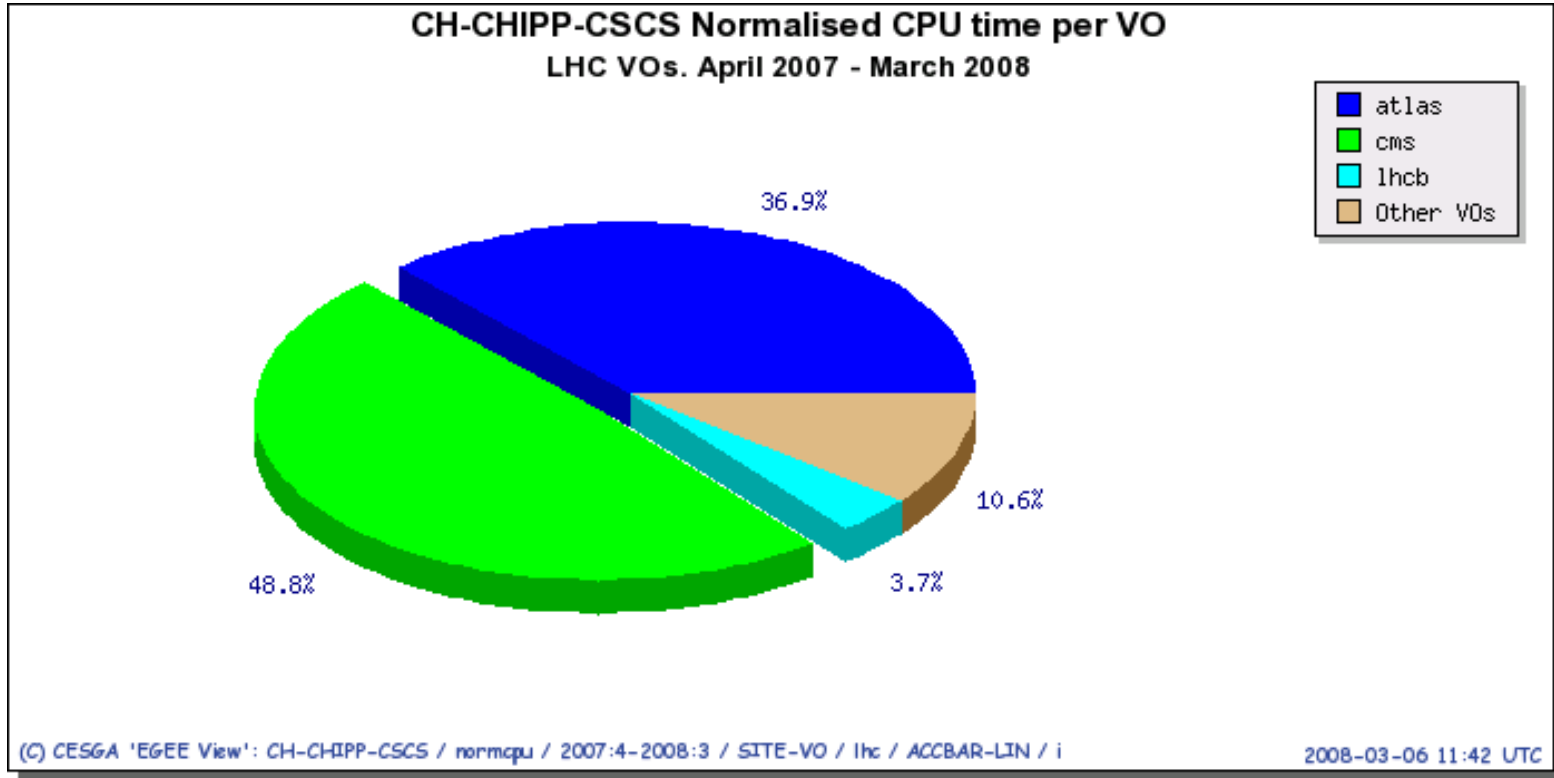
➔ **Stable operation - high efficiency**

# PHOENIX LCG usage (4/07-3/08)



LCG-usage  
per VO

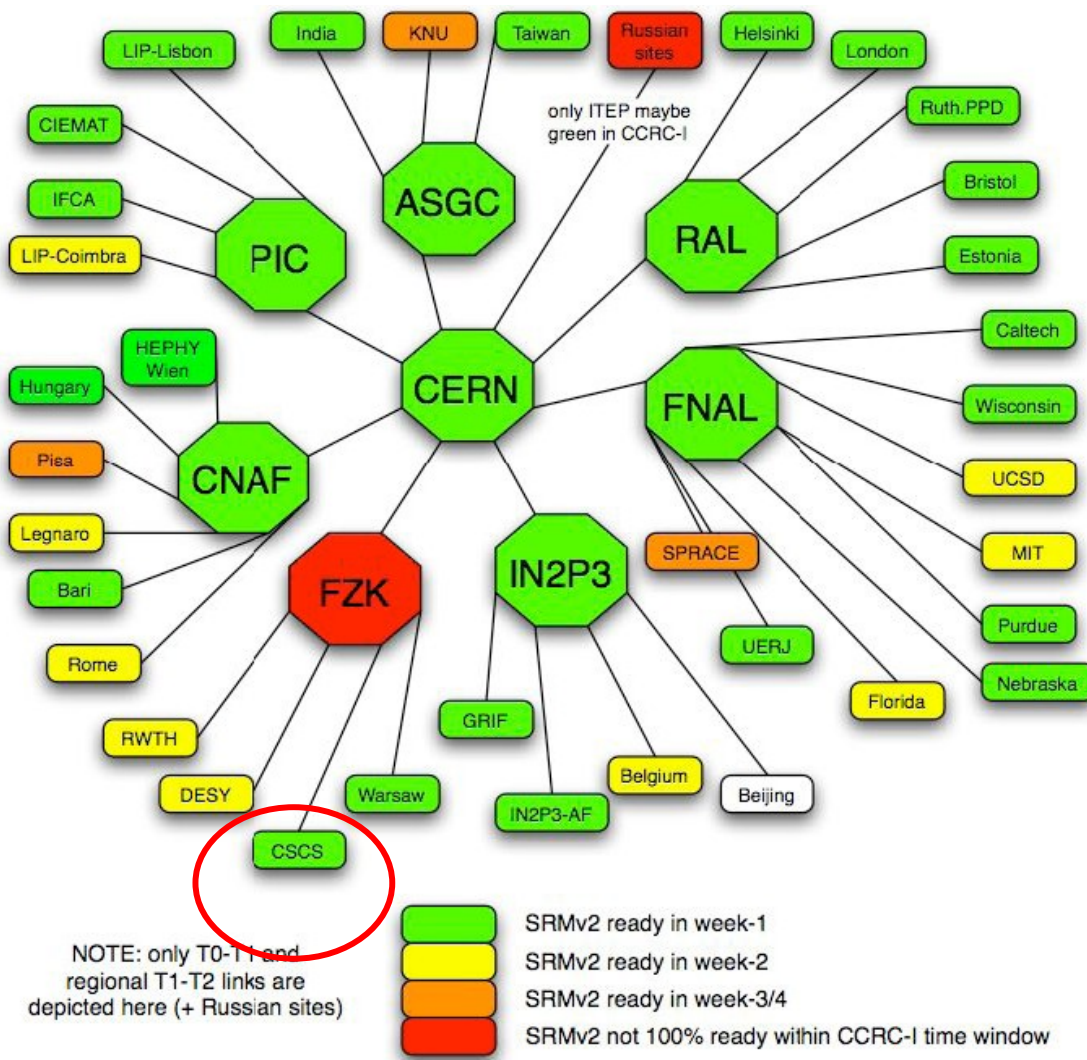
Visible are:  
CMS CSA07  
and CCRC08



**CH Tier-2 was used mainly by CMS and Atlas**

# CCRC08 : SRM2 adoption

At start of CCRC08



**CH Tier-2 :  
SRM (storage resource manager)  
was operational early !**

## Evolution of PHOENIX cluster - next phases

# Ramp-up Plan CH-Tier-2: status Q1/08

- Actual **ramp-up schedule** ; calculations based on Q2/07 pricing, installed in collaboration with SUN (phase-A completed 1/08)

Phases	Latest installation time	Minimum <i>aggregate</i> compute capacity [kSi2000]	Minimum <i>available</i> disc space [TB]	Cost estimates in kCHF
existing + CSCS GRID cluster	2006/07	214	52	250
Phase A	Operational: 1/08	~800	225	1120
Phase B	End 2008	1500	420	~ 1300
Phase C	End 2009	2600	800	~ 1300
Operation+repl. /a	>= 2010	2600	800	~700

\* 1 XEON 3 GHz  $\approx$  1.5 kSi2000

- WAN: need > 3 Gbps in 2008/09**  
have already now 1 Gbps, and can get 2x10 Gbps anytime 😊



## Assumptions:

- **CPUs:** Opteron 2.6 GHz CPU with 1.5 kSi2k per core.
  - ❖ **Phase A** Dual Core CPUs
  - ❖ **Phase B** Quad Core with Phase A duals upgraded to Quads
  - ❖ **Phase C** Quad Core
- **Disk: X4500 :** net capacity of 17.85 TB (incl. raid, mirror, spares, overhead)



	Aggr. CPU Performance in Ksi	Additional Cores needed	Cores delivered	Aggr. Net Storage Capacity in TB	Storage needed	Storage delivered
Phase A	680 **	453	400 (5 Blade Centre)	225	225TB (12.6 X4500)	178,5TB (10 X4500) +1 CSCS
Phase B	1440	+453 + 54 (507)	+(400 + 160) (+ a 6 <sup>th</sup> Blade Centre)	490	+265TB (+14.8 X4500)	+303.4TB (+17 X4500)
Phase C	2640	800	+800 (5 Blade Chassis)	910	+420TB (+23.5 X4500)	+410TB (+23 X4500)
Total	2640	1760	1760 (11 Blade Centre)	910	910TB (50.9 X4500)	910.3TB ( 51 X4500)

\*\* new processors (available on delivery in 2007) boosted this 680 to about 800 kSi2k

# Status Financing

# Financing Tier-2 - Timeline

Phase A is Operational

- ◆ Prototype financing by institutes: in 1.2004 for CSCS-LCG cluster [10 nodes ]; granted 50 kFr
- ◆ **1st FORCE grant** : requested 128 kFr. in 3.2004; → for “Phoenix cluster” [ Prototype 15 nodes ] granted 128 kCHF
- ◆ **2nd FORCE grant** : requested 670 kFr in 9.2005; → “Phoenix cluster” [ **Phase A/1** ] in 3.2006 granted 300 kCHF
- ◆ **2<sup>nd</sup> +**: addendum in 9.2006 → “Phoenix cluster” [ **Phase A/1+** ] granted 190 kCHF
- ◆ **3rd FORCE grant** : requested 670 kFr in 9.2006; → “Phoenix cluster” [ **Phase A/2** ] in 3.2007 granted 500 kCHF
- ◆ **4th FORCE grant** : requested 1300 kFr in 9.2007; → “Phoenix cluster” [ **Phase B** ]
- ◆ Planned 5th FORCE grant : ask for ~1300 kFr in 9.2008; → “Phoenix cluster” [Phase C ]
- ◆ from then on... rolling replacement; order 700 kFr / year

asked for phase-B



## Additional contributions so far:

- **Unis+ETH:** granted in 1.2004 50 kFr.  
invested in prototypes CSCS-LCG cluster [ 10 nodes ]
- **ETH-IPP :** granted in 12.2006 50 kCHF  
➔ invested in thumper addition for PHOENIX [ 24 TB ]
- **ETH-PSI :** granted in 2007: 50 kCHF  
➔ invested in PHOENIX phase A
- **UNIZ :** granted in 2007: 30 kCHF  
➔ invested in PHOENIX phase A

**Note:** also in the future:  
contributions by Unis and ETH are strongly suggested !

# Tier-3 Issues ...



- a system in production since 2005, mainly as a grid batch job facility
  - ◆ ~150'000 CPU hours (70'000 jobs) in 2007 for ATLAS
- Evolving towards 1st data
  - ◆ More hardware in Summer 2007
  - ◆ More local use, batch and interactive
  - ◆ More security/redundancy
  - ◆ Work on data transfers between Swiss and other sites
- size
  - ◆ 61 machines, 188 CPU, 75 TB
  - ◆ This is what we will use for 1<sup>st</sup> data thanks to S.Gadomski

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.

## Size

- ~130 cores for ATLAS (~200 in 2009).
- ~ 33 TB disk (end of 2008 44 TB).

## Usage

- ~ 120 000 Wall Time Hours in 2006.
- ~ 170 000 Wall Time Hours in 2007.
- ~ 40 000 Wall Time Hours in Jan+Feb 2008.



S. Haug

# Status CMS - Tier 3

- **Presently operating:** CMS-nodes at ETH (Trueb, Dambach):  
 5 Servers (each: 2 Dual-Xeons 3.2 GHz + eff. 4.5 TB) :
- **Planned: common CMS Tier-3** for ETHZ, PSI + UNIZ;  
 located and operated at PSI; choose similar architecture as Tier-2  
 (planning D.F. + C.G.)

Year	2008	2009
<b>CPU / kSI2k</b>	<b>180</b>	<b>500</b>
<b>Disk / TB</b>	<b>75</b>	<b>250</b>
No of Worker Nodes	10	28
No of CPU Cores	80	224
No of Storage Nodes	4	14
No of Racks	1	3

→ emphasis on storage

→ to become operational in 08

**Network:** aim for 6 MB/s read access per job and kSI2k;  
 Connection of 1 Gb between PSI and CSCS.

D.Feichtinger+C.Grab

# Personnell ...

## ● Personnel at CSCS for GRID computing :

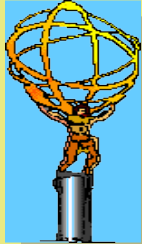
- Peter Kunszt; S.Maffioletti, M.Ricardo, +
- total: 2 GRID-FTE + about 2 FTE for operational/technical support

## ● CHIPP personell for experiment's related operations

- **CMS:** D.Feichtinger (80%, PSI); crucial for T2- operation – thanks ! support by Z.Chen (ETH)  
Tier-3: DF for operation of planned T3-cluster; Dambach/Trueb now;
- **ATLAS:** Szymon Gadomski (Geneva) (100%), S.Haug (Bern)
- **LHCb:** Roland Bernet (UniZ), parttime
- **CHIPP:** C.Grab parttime for all expts./CHIPP:  
coordination Tier-2 (+ CMS-T3)

● **Note:** we will need at least a second person to support middleware for each experiment at Tier-2 and Tier-3 !

- Status technical operations:
  - PHASE-A installation of PHOENIX cluster completed !  
hardware and software operate stably, with the usual minor problems 😊
  - previously borrowed phase-0 nodes returned to CSCS
  - Participation in CCRC08, the combined computing readiness challenge of all experiments:
    - CH-CSCS early and successful participation 😊
    - data transfers: continuously exercised by CMS,  
→ reached 50 MBytes/s sustained !
    - Technically most challenging and intense “daily care” is needed for SW-stack of storage middleware
  - preparations for PHASE-B in progress
  
- Inauguration of the PHOENIX cluster at CSCS  
will be on 29.May at CSCS



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

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

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

**P. Kunszt (CSCS Swiss Grid Initiative)**