Report from the
CHIPP Computing Board
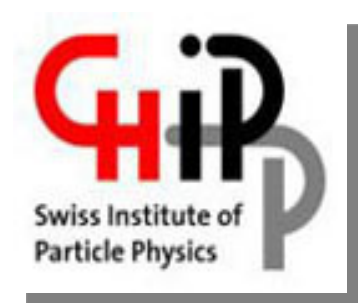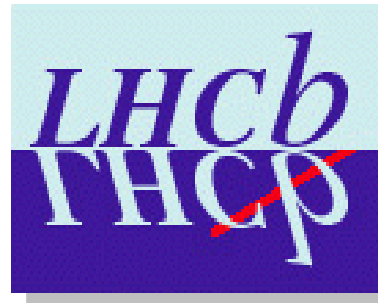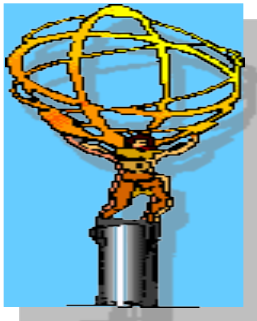
Christoph Grab

- The Swiss Tier-2 Cluster – called "**PHOENIX**"
  *"located at CSCS,
  operates in production mode" -
  delivering the Swiss contribution to WLCG –
  was and continues to be part of the LHC physics analysis tool.*

- Cluster was physically moved successfully to Lugano in May 2012 - *all planned hardware changes/replacements successfully completed*

CSCS moved from Manno
to Lugano in May '12

Inaugurated 31.8.2012

F.Schiesser, A.Berset, R.Eichler

Our Swiss Tier-2,
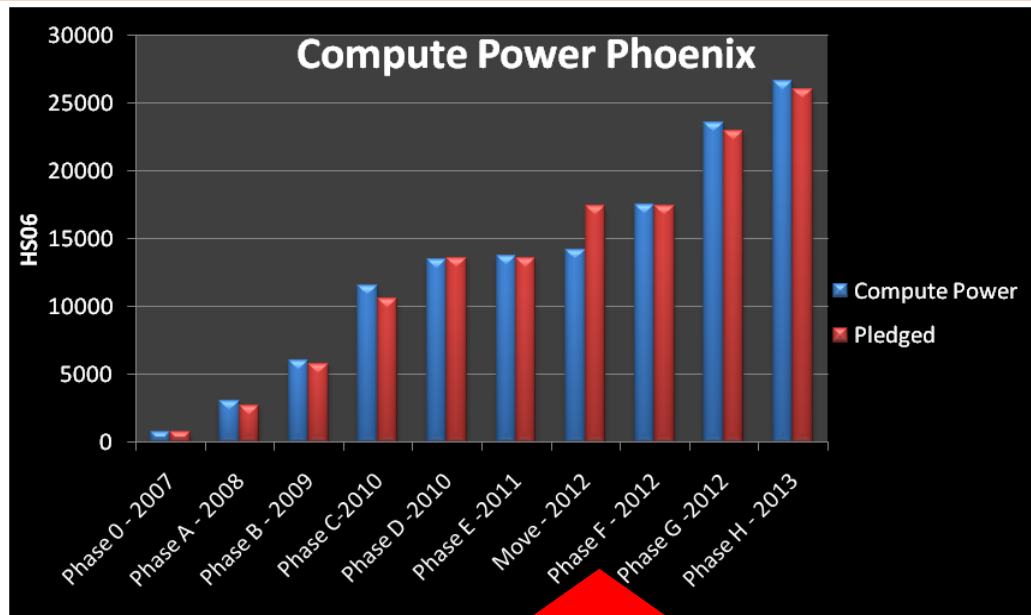the Phoenix cluster

is ONE ISLAND of clusters

Phoenix Evolution
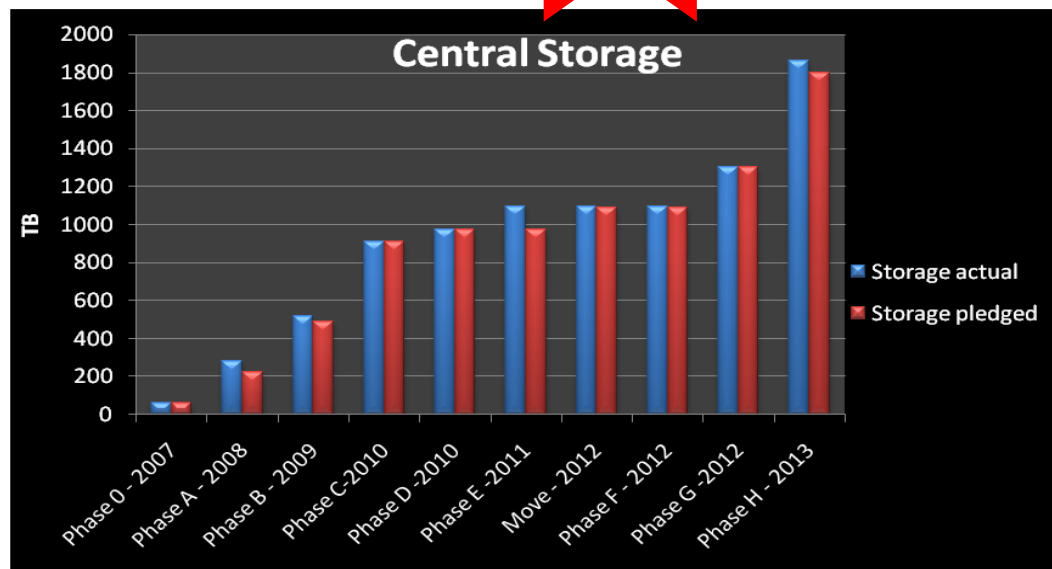in 2007 – 2012

(2013 planned)

- **CPU Power in HS06**

- **Storage in TB**

= we are here

❖ Phase E = move to Lugano - completed in May;

❖ Phase F = to fullfil the 2012 pledges
  ❖ storage ok in Aug; CPU not before Q1/2013
  ❖ Replace old thors storage

❖ Phase G = preparation for 2013 pledges (ready by 1.4.2013)
  ❖ Expand CPU in early 2013 (after tendering…)

❖ Phase H = prepare FORCE request for fulfilling 2014 pledges

In more detail …

	Computing (HS06) Have / Pledged	Storage (TB) Have / Pledged	Scratch (GB/s) Have / Desired
March 2012	13600 / 13550	1100 / 976	8 / 7.5
After move / Phase E	14400 / 17400	1100 / 1090	8 / 7.5
Phase F (Q1/'13')	17500 / 17400	1100 / 1095	8 / 9.5
Phase G (pledge '13)	P: 23532 / 23000	1300 / 1300	12 / 11

❖ Note: power consumption dropped from 87 kW to 57 kW
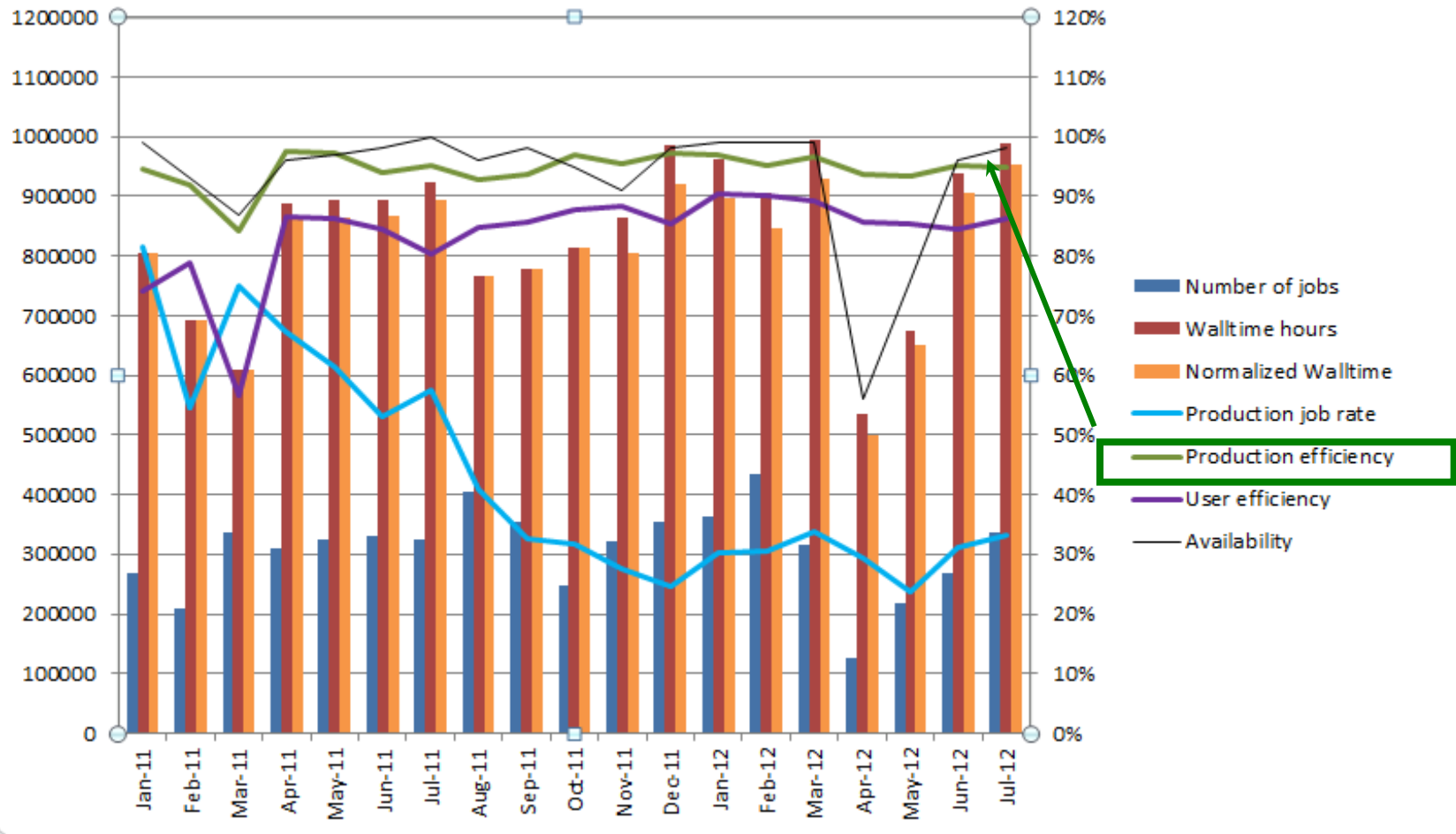
Overall very high efficiency and availability achieved !
        Note: move in Apr/May 12; …

CSCS-LCG2 Normalized Elapsed time (HepSpec06)
January 2011 – July 2012

Reasonable distribution of CPU to experiments
(note: ARC contributions NOT visible)

Pledges to WLCG and Capacity (8/2012)

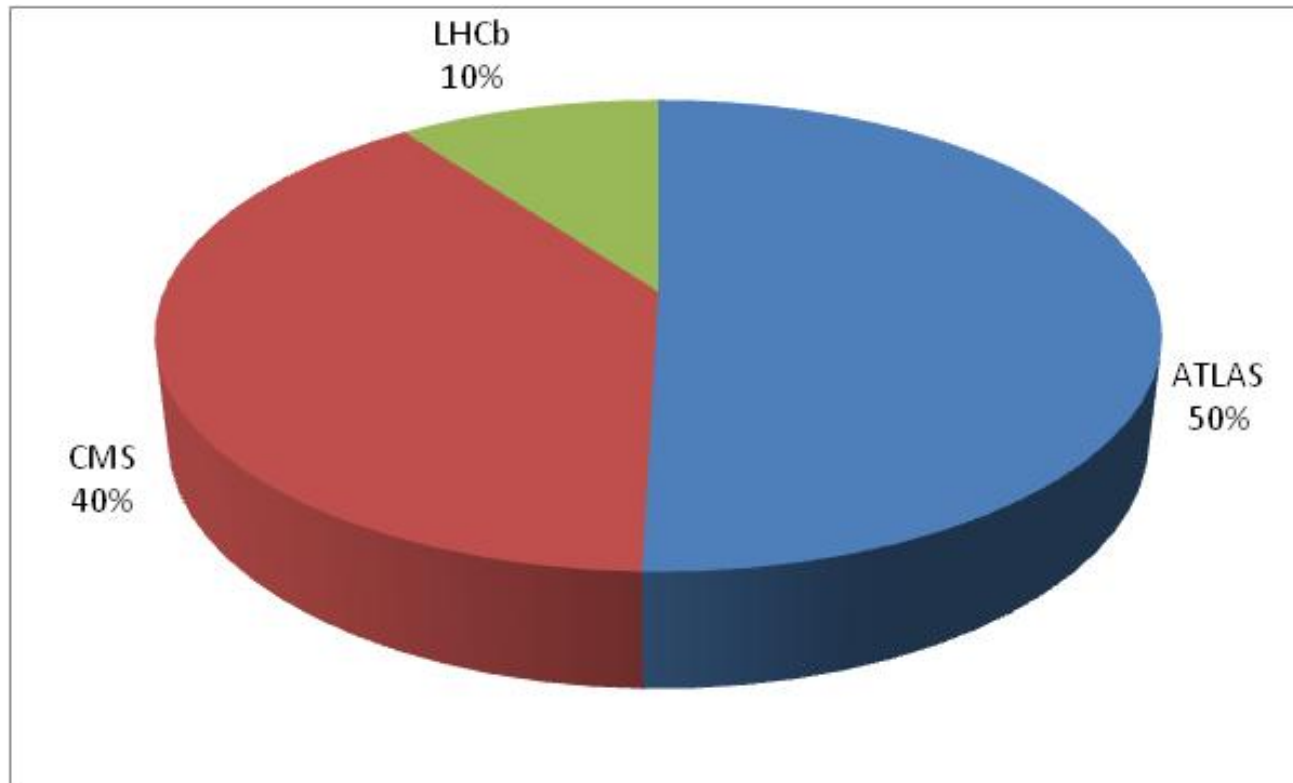## REBUS: Capacity and Pledge Comparison

Topology	Pledges ▾	**Capacities** ▾

Capacities > Capacity and Pledge Comparison

Country ▲	Federation ⇕	Physical CPU ⇕	Logical CPU ⇕	HEPSPEC06 ⇕	CPU Pledge ⇕	Total Online Storage (GB) ⇕	Disk Pledge ⇕
Federation							
Slovenia	SiGNET	875	2,866	27,985	12,000	37,687	900,000
Spain	ATLAS Federation	1,135	4,851	53,348	13,300	6,288,975	2,350,000
Spain	CMS Federation	827	3,617	34,308	15,750	2,648,009	1,300,000
Spain	LHCb Federation	468	1,917	13,394	2,800	2,509	1,000
Sweden	SNIC Tier-2	771	771	3,400	7,870	1,983,958	920,000
Switzerland	CHIPP	98	1,411	14,154	17,670	1,148,867	1,226,000
Taipei	Taiwan Analysis Facility Federation	285	1,119	11,326	5,320	0	600,000
Turkey	Turkish Tier-2 Federation	333	656	4,361	9,800	1,801,958	900,000

Our pledged resources for 2012 come late (20% CPU missing)…
However: Switzerland is just about average … (compare Slovenia, Portugal….)

- Details see: http://wlcg-rebus.cern.ch/apps/capacities/pledge_comparison/
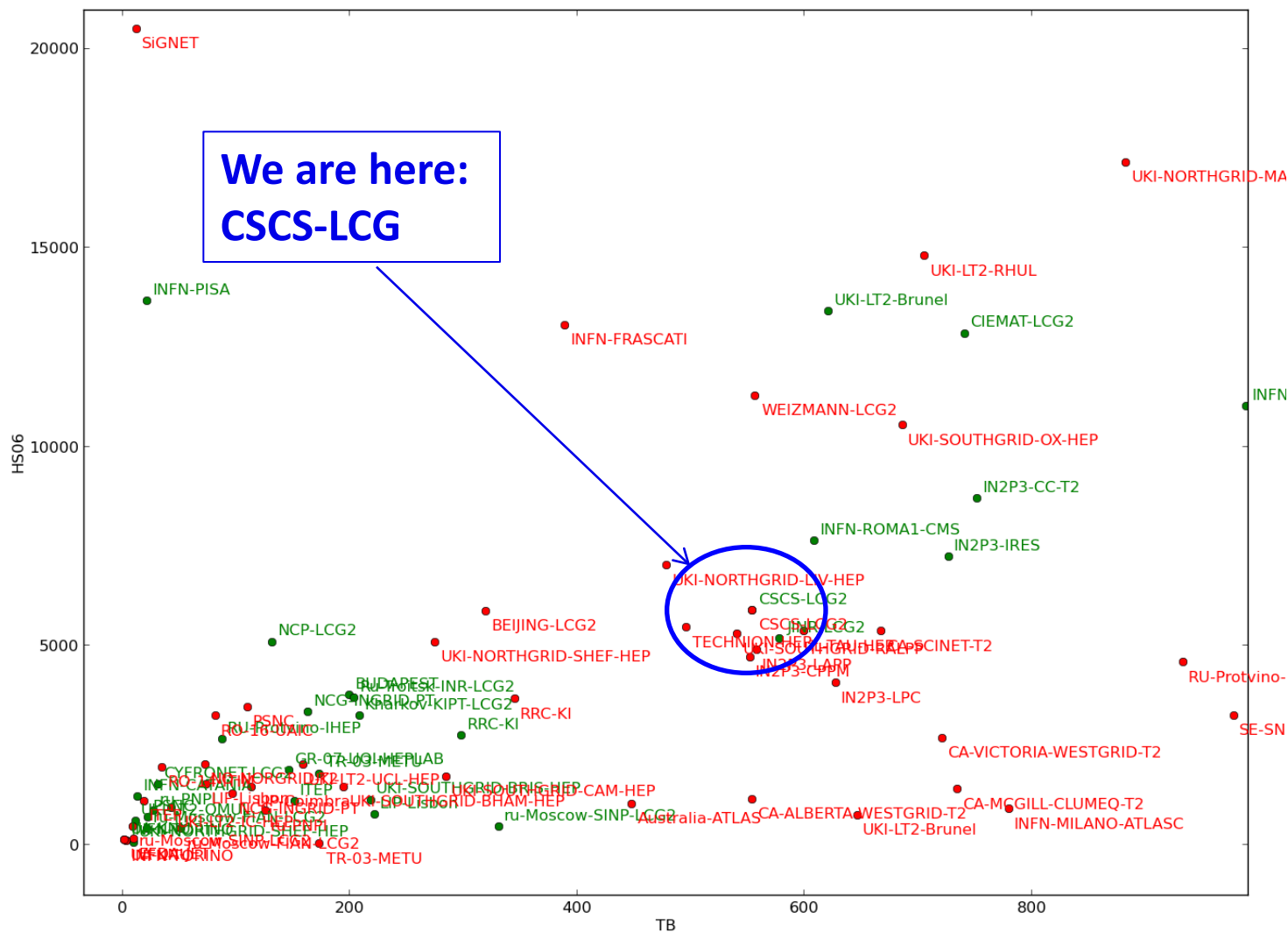
- Swiss Tier-2 is positioned about average wrt other tier-2s

Planning for 2013 and beyond

- Planning and operational issues are discussed within the CHIPP computing board (CCB) and also directly with the tier-2 representatives from CSCS.

- We do a yearly planning of hardware upgrades and replacements at Tier-2 for the next phase.

- Input are the experiments "Computing requirements documents" – and reality of available resources.

- **Investments in 2012**: used for moving and
  - To replace old CPU and old storage (Thumpers + Thors)
  - Catch up in fulfilling 2012 pledges in CPU in Jan/Feb. 2013

- **Request investments from FLARE/SNF for 2013 :**
  - Request in Sep. 2012 for funds in 2013
    - HW resources to be spent in late 2013 to meet the 2014 "pledges" (in time for pledge date) ~ **610 kCHF**

- **Planned investments (FLARE/SNF) for 2013 - 2016** :
  - Request for replacements and additions :
    - hardware for 550 – 600 kCHF / year
  - ONE-TWO FTE persons to support Tier-2 operation at CSCS
    - recurring power costs, if ETH/CSCS does not continue MoU asis
    - additional costs for EGI contributions (level 30-50 kCHF/a)

Note: all numbers are today's; will change with time, and possibly $ vs CHF ??

➢ Public procurement procedure tightened on Federal level due to recent happenings in the departments

➢ "Freihaendige Vergabe" increasingly difficult .

➢ Consequences:
- We need to do an public WTO-tendering process again
- CPU: Our order for 2012 purchase of CPU upgrade is delayed;

- We join with CSCS a "common, general tendering process", separately for CPU/compute and for storage

Tier-2: Long term Planning issues

Include running experience to optimize the cluster usage

- **CPU or storage : which should be invested in ?**
  - Ratio of CPU/Storage is different in ATLAS (6), CMS (13) and LHCb (4200)!
  - storage is more expensive
- **How to split** resource sizes between ATLAS, CMS and LHCb?
  Up to now we used a model of 2:2:1 according to personell.
  Question: is this still true ? Present user numbers at tier3 ~70:50:40.
- **Layout/configuration of nodes?**
  - memory/core requirements is growing and uncertain
  - ATLAS >3 GB/core, CMS (2-3 GB), LHCb (<=2 GB)
- **Usage of bandwidth to and dimensioning of scratch ?**
  - Requirements different in ATLAS (>5 GB/s), CMS and LHCb (<=1 GB/s)
- Number of FTEs needs to follow increase in size+complexity
  - 2 FTEs at CSCS; + about 0.8 FTE from institutes

- Need to re-negotiate MoU CHIPP <-> CSCS./ ETH now !

- Solution / optimization need to be found within CHIPP/CCB !

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Summary of Swiss Tiers Efforts (Q3/12)

Site (~#users)	Nr cores	CPU (HS06)	Storage (TB)	Comments
Swiss Tier-2 √	**1472**	**~ 17538**	**1114**	
ATLAS BE (15)√ GE (~60) √	472 352	3294 4980	252 460	BE: +1536 HT-cores+250 TB Implementation under way GE: 10 Gbps line to CERN. Next upgrade is under way.
CMS ETHZ, √ PSI, UZH (~50)	392	~5887	509	GRID SE + UI :direct GRID access.
LHCb EPFL (30) √ UZH (12) √	480 48+96sh	~3600 835+847sh	143 88	EPFL; UZH:MC production; shared
Total Tier-3		**~19000 HS06**	**1450**	1 kSI2k ~ 200-250 HS06

- Tier-3 capacities : similar in size as the Tier-2

Distribution of old hardware (2012)

Was agreed to re-distribute the Tier-2 hardware to the Swiss tier-3s for continuous usage (without warranty); from move, we gave:

CPU : To ATLAS Tier3 (Bern)

- SunBlades full Rack (96 nodes)
  + 2 eth switches (with cables) + 16A power cables

In case you want to know more details …

**Infiniband-SWITCH : To ATLAS Tier3 (Bern)**

- M9 IB Switch, with power cables, fibre cables (special connector) to the SunBlades, and fibre cables with QSFP for normal nodes.

Storage : To ATLAS Tier3 (Bern+Geneva) and to CMS tier3 (at PSI)

- 22 Thumpers to UniBe (10 old, 10 new, 2 broken)
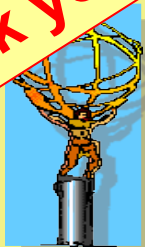- 2 older + 2 newer thumpers to CMS (PSI, CERN)
- 2 newer thumpers to UniGe

Summary ..

- Cluster runs well and reliably !
- We contribute(d) to physics exploitation !
- We deliver (nearly) as pledged to WLCG !

- 2013 will be again different
- NO additional data – LS1
- Need to optimize the cluster setup+operation…..
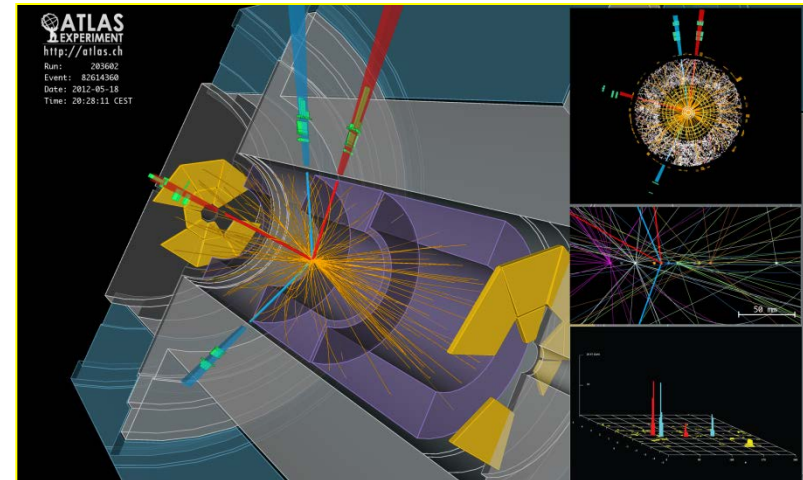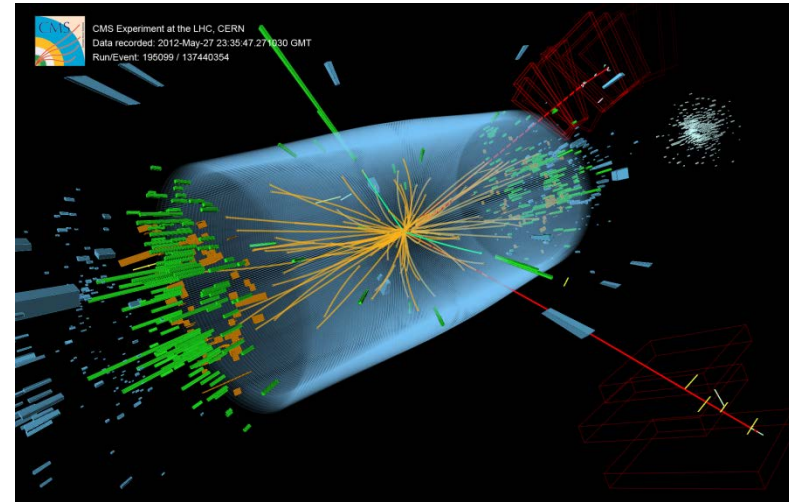
Thank you …

Coordinates the Tier-2 activities
representatives of all institutions and experiments
included also the tier-3 expert

S.Gadomski, P.Jacobucci (UNI Ge)

S.Haug, H.P.Beck, G.Sciacca (UNI Bern)

C.Grab (ETHZ) chair CCB

D.Feichtinger (PSI) vice-chair CCB

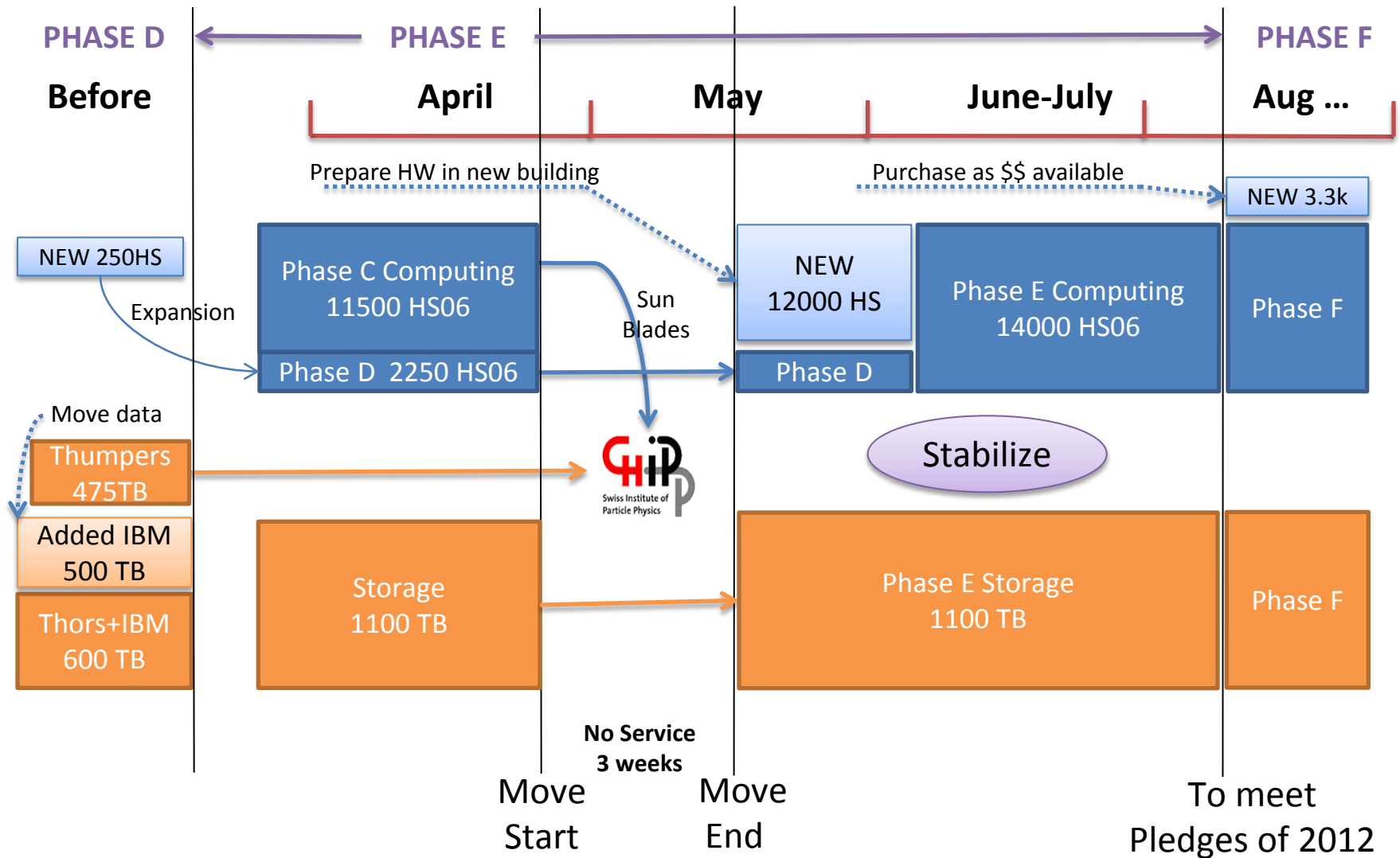D.Meister (ETHZ), F.Martinelli

R.Bernet (UNIZH)

S.Tourneur (EPFL)

P.Fernandez, M.De Lorenzi (CSCS)

Swiss NGI and EGI

- **SWING took over the Swiss NGI international representation in** 2012 from SWITCH (which changed strategy + organisation)

- Actual resources needed to operate a national NGI are provided by:
    - Operational (national+international) tasks done by users [T2, T3, SMSCG].
    - Project management by SWITCH; coordination+representation by SWING
    - Funding by SBF, SWING, CHIPP.

- **Support for international Grid operation (EGI)**
    - In 2012: EGI costs were 270 kCHF; covered by
        - SBF with 210 kCHF (one-time only)
        - European membership fee for 2012 was paid by CHIPP (60kCHF)
    - **For 2013**: costs budgeted to 140 kCHF, planned to be covered by
        - SWING with 110 kCHF and
        - CHIPP with 30kCHF   ← **expected total from CHIPP in 2012**

        (cover this globally from CHIPP LHC computing budget)

- 2014 ff: expect new project contributions by SUK and SER

In case you want to know more details …

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Phase E: move to Lugano done (opt)

Phase E completed in Aug 12

- **Fully Water-cooled facility**
  - Thermally isolated island, 2x cool, 1x warm corridor.
  - Two water circuits warm lake water from 6 to 29 Celsius degrees.
  - Higher inlet temperature, more air pressure, same cooling capacity
- **New Infiniband QDR production network**
  - Fat Tree topology: 2 core + 5 leaf switches.
  - 192 Gbit/s (24 GB/s) aggregate bandwidth (worst case).
  - 32 Gbit/s (4 GB/s) from each node (upgr. to 64 Gbit/s if needed).
- **Transparent IB/Ethernet bridge**
  - Voltaire 4036E in HA mode (active/passive).
  - 20 Gbit/s to Ethernet hosts (virtual machines, CSCS network).
  - Fully transparent: no need for additional software layers.
  - Internet connection upgraded to 10 Gbit/s (upgr. to 20 Gbit/s).
- **36 new compute nodes from Dalco**
  - 2x Intel SandyBridge E5-2670 @ 2.6 GHz. 410 Watts max.
  - 32 job slots, 2 GB RAM / slot.
  - 333 HS06 performance (50% faster than AMD).

Swiss Tier-2/3: Network (opt)

- **Network traffic**:
  - Still: present bandwidths are sufficient …
  - Internally of PHOENIX 2x10 Gb/s
  - Internally at CSCS 10 Gbps (at Lugano)
  - externally now to SWITCH 16 Gbps.

- **routing within Switzerland via SWITCH :**
  two redundant lines >10Gbps to CERN and Europe

  - If needed, we can upgrade within months …!

Still not a real issue …

Swiss Network (opt)

SWITCHlan Backbone 2012

Legend:
- Darkfibers
- SWITCHlan backbone node
- SWITCHlan node with external peerings
- Project — Lightpaths
- Provider — Internet transit
- IX — Internet exchange
- Network — Research and education network

Served well by SWITCH …

SWITCH
Serving Swiss Universities

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Swiss Institute of Particle Physics