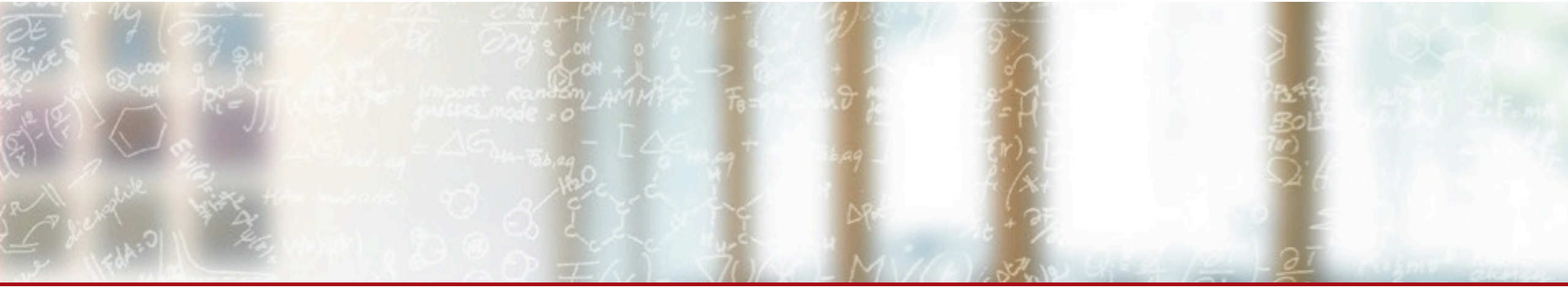
CSCS
Centro Svizzero di Calcolo Scientifico
Swiss National Supercomputing Centre

ETH zürich



CHIPP - CSCS F2F meeting

CSCS, Lugano
January 21st, 2019

CSCS
Centro Svizzero di Calcolo Scientifico
Swiss National Supercomputing Centre

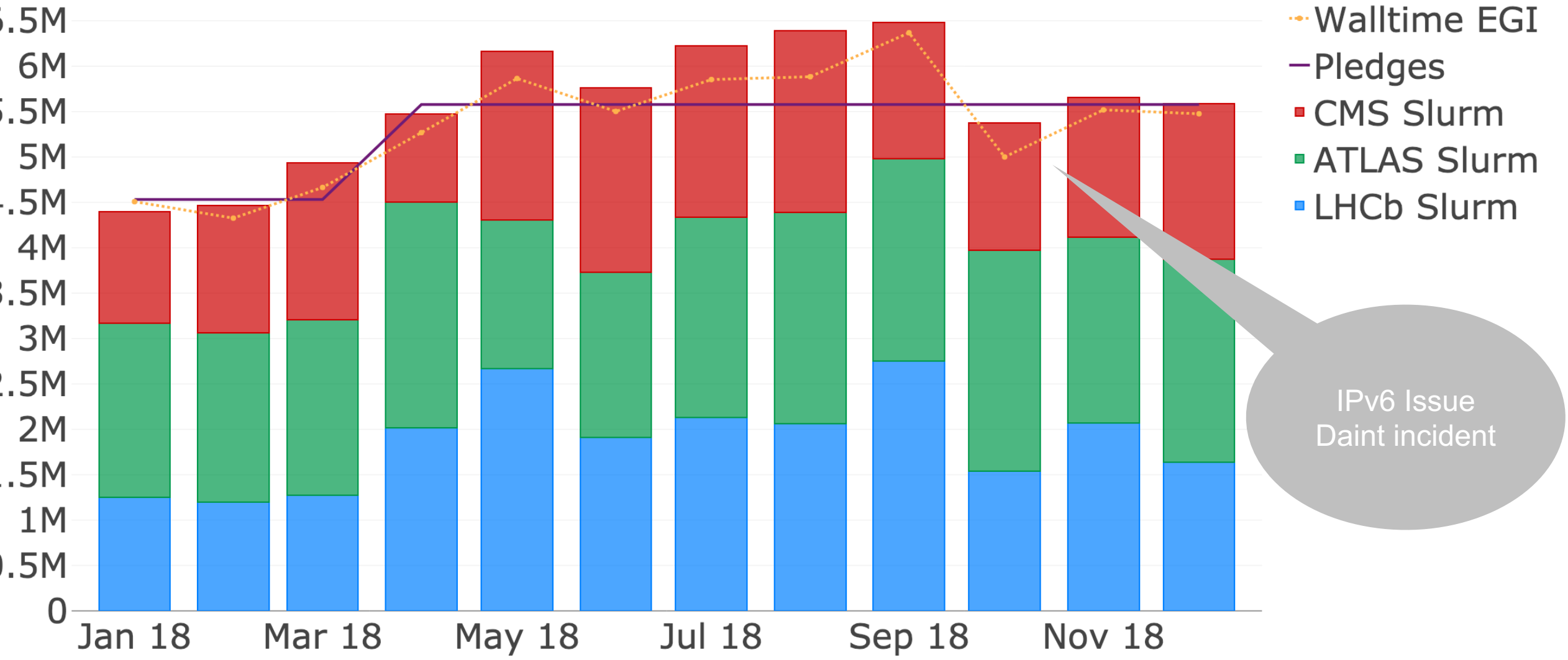
ETH zürich

Tier 2 status and plans
CSCS

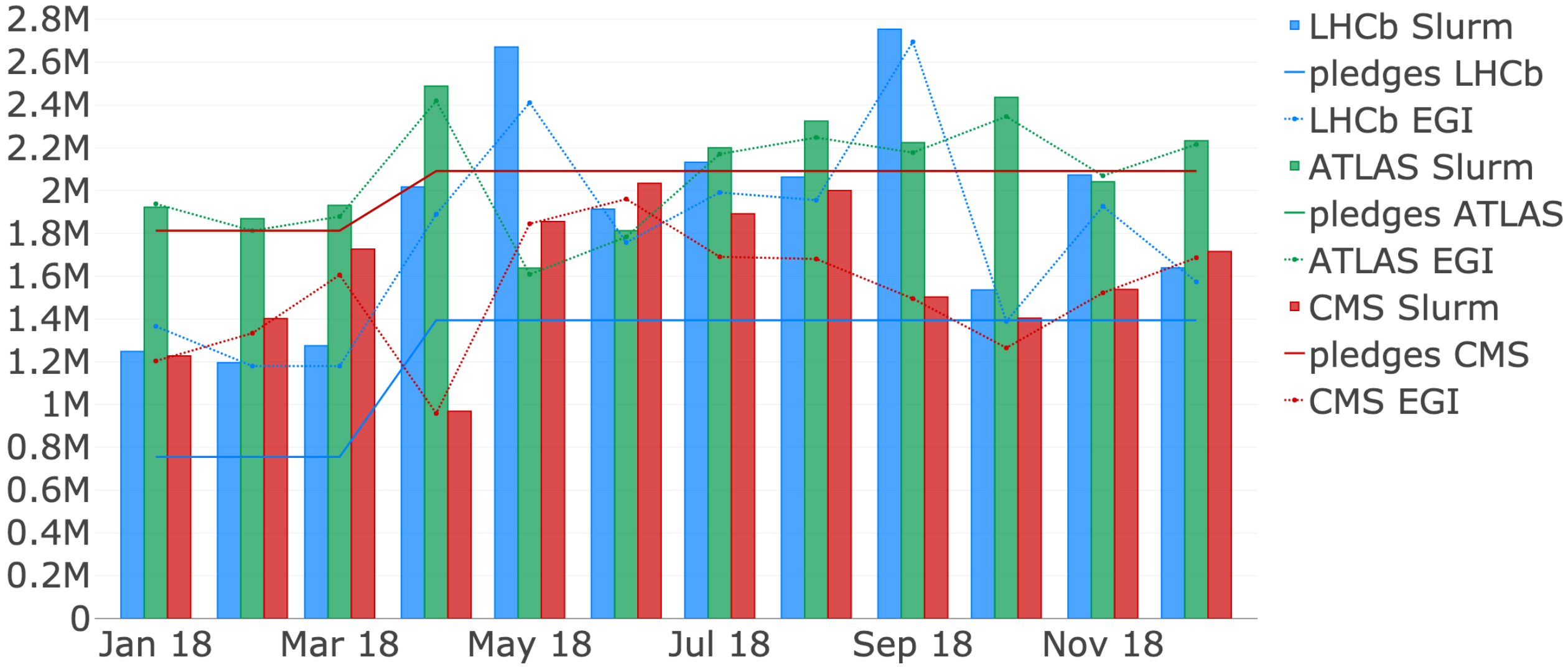
1. **Statistics**
   a. Availability/Reliability
   b. CPU usage
   c. Storage usage
2. **Operations**
   a. Updates
   b. Main Issues
3. **Plans**
   a. Resources Overview

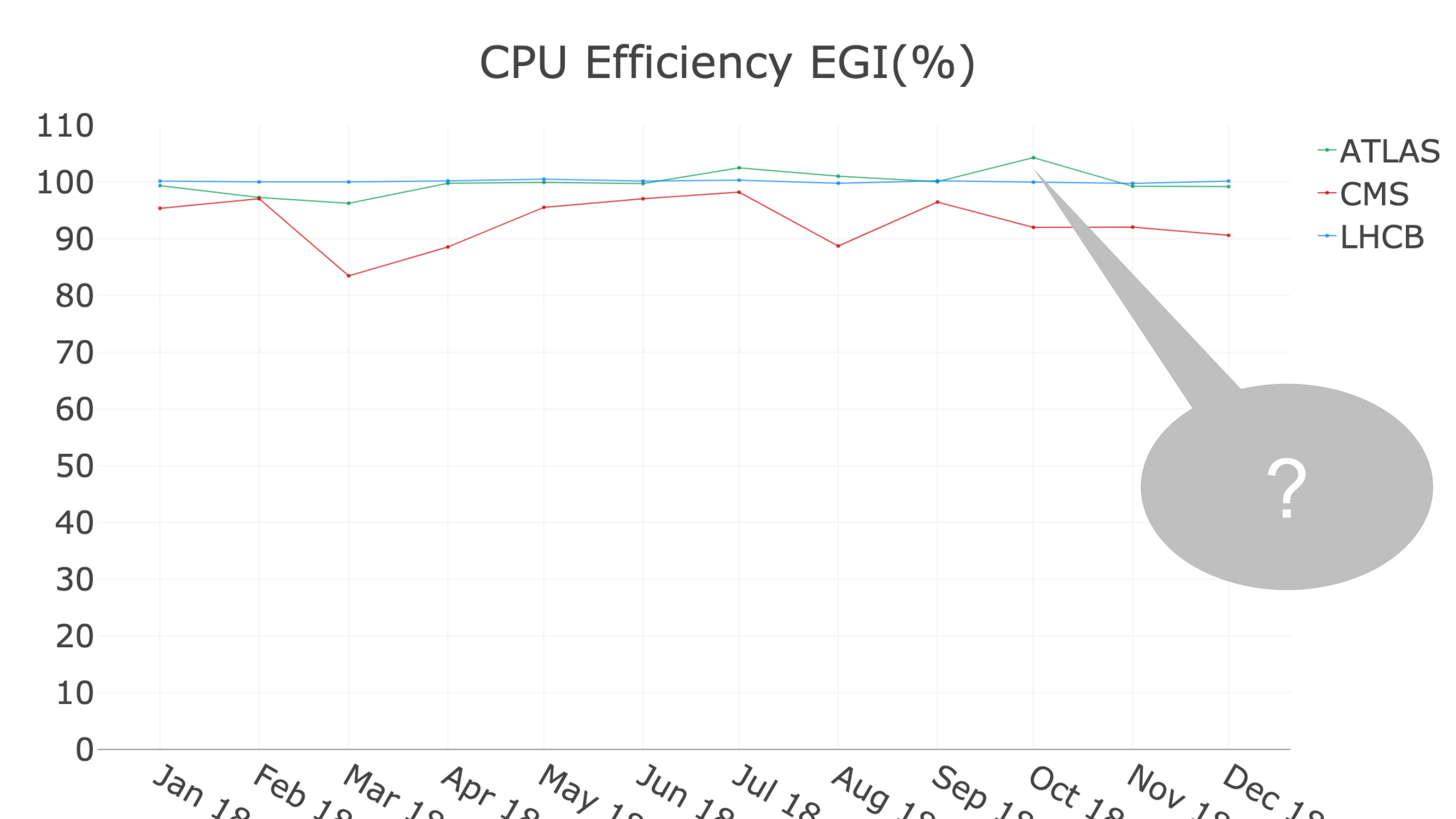
Statistics

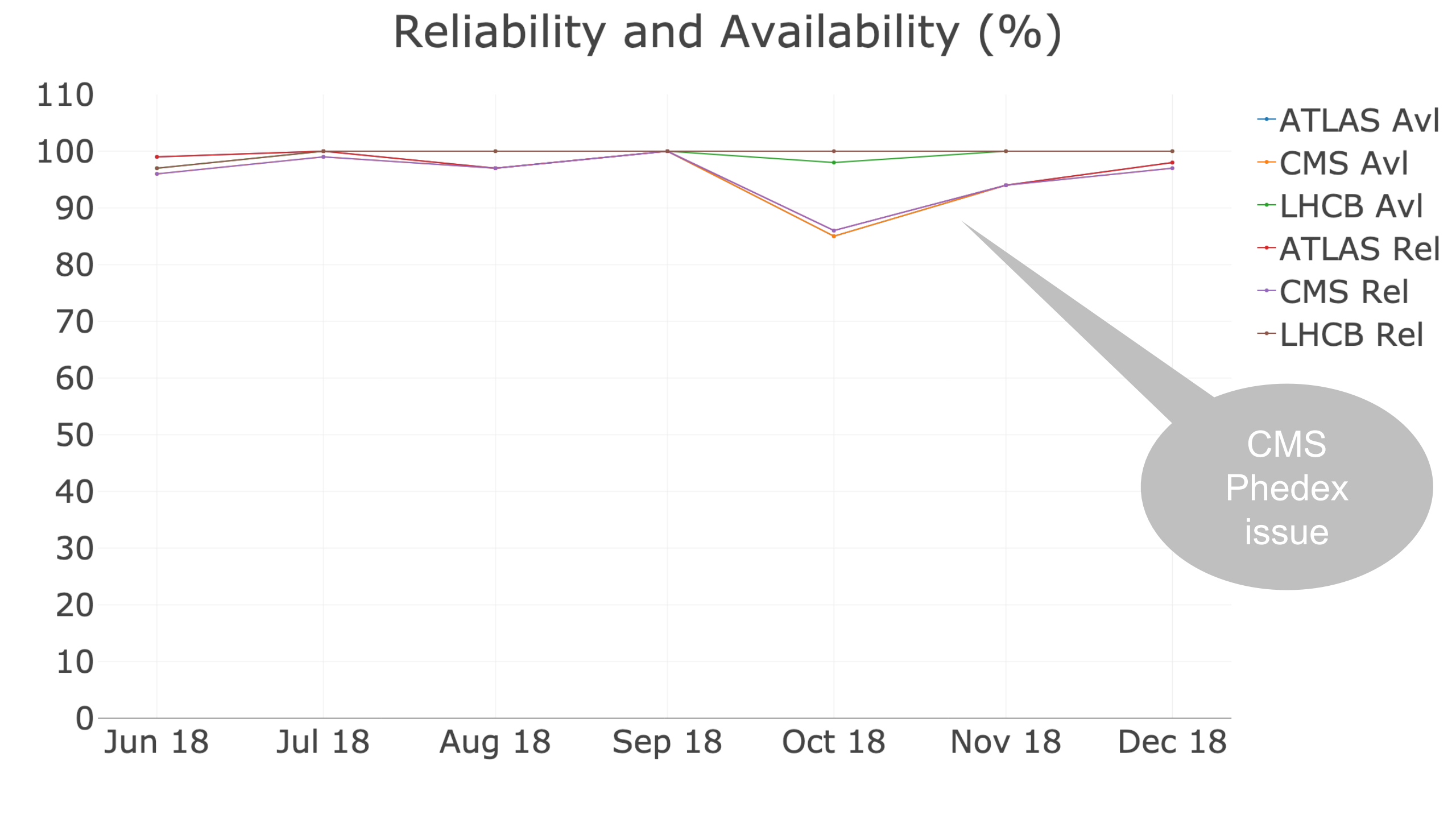
Accounting in CPU hours (Piz Daint, Phoenix)



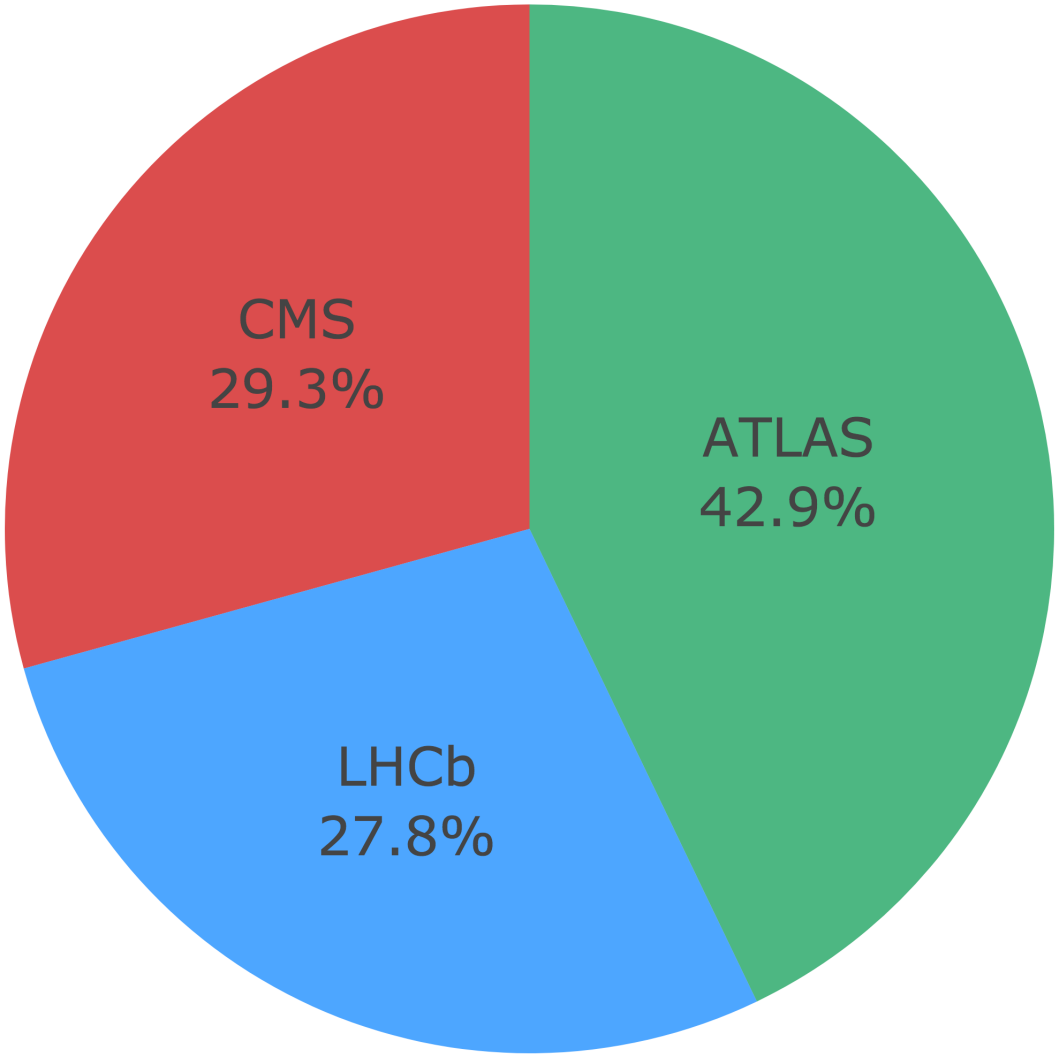
Accounting in CPU hours (Piz Daint, Phoenix)




CPU Efficiency EGI(%)
110
100
90
80
70
60
50
40
30
20
10
0
ATLAS
CMS
LHCB
?
Jan 18
Feb 18
Mar 18
Apr 18
May 18
Jun 18
Jul 18
Aug 18
Sep 18
Oct 18
Nov 18
Dec 18


Reliability and Availability (%)
110
100
90
80
70
60
50
40
30
20
10
0
Jun 18
Jul 18
Aug 18
Sep 18
Oct 18
Nov 18
Dec 18
ATLAS Avl
CMS Avl
LHCB Avl
ATLAS Rel
CMS Rel
LHCB Rel
CMS Phedex issue

Statistics – Usage per VO 2018

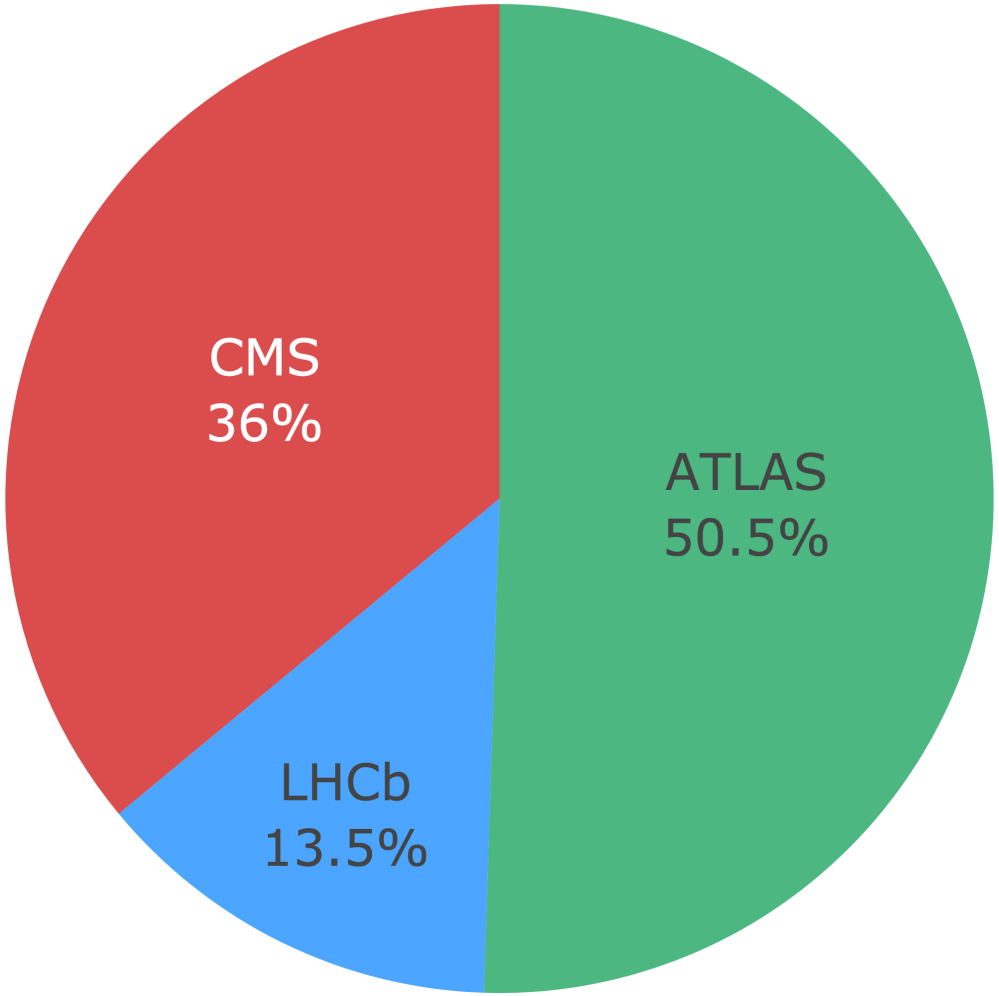


Statistics – Storage usage 2018

ATLAS
Total: 2’736.07 TB
Used: 1’535.54 TB **PL19: 1'720TB**
Free: 1’100.53 TB

CMS
Total: 1’950.29 TB
Used: 1’733.66 TB **PL19: 1'720TB (!)**
Free: 216.65 TB

LHCb
Total: 734.20 TB
Used: 521.51 TB **PL19: 860TB**
Free: 212.69 TB

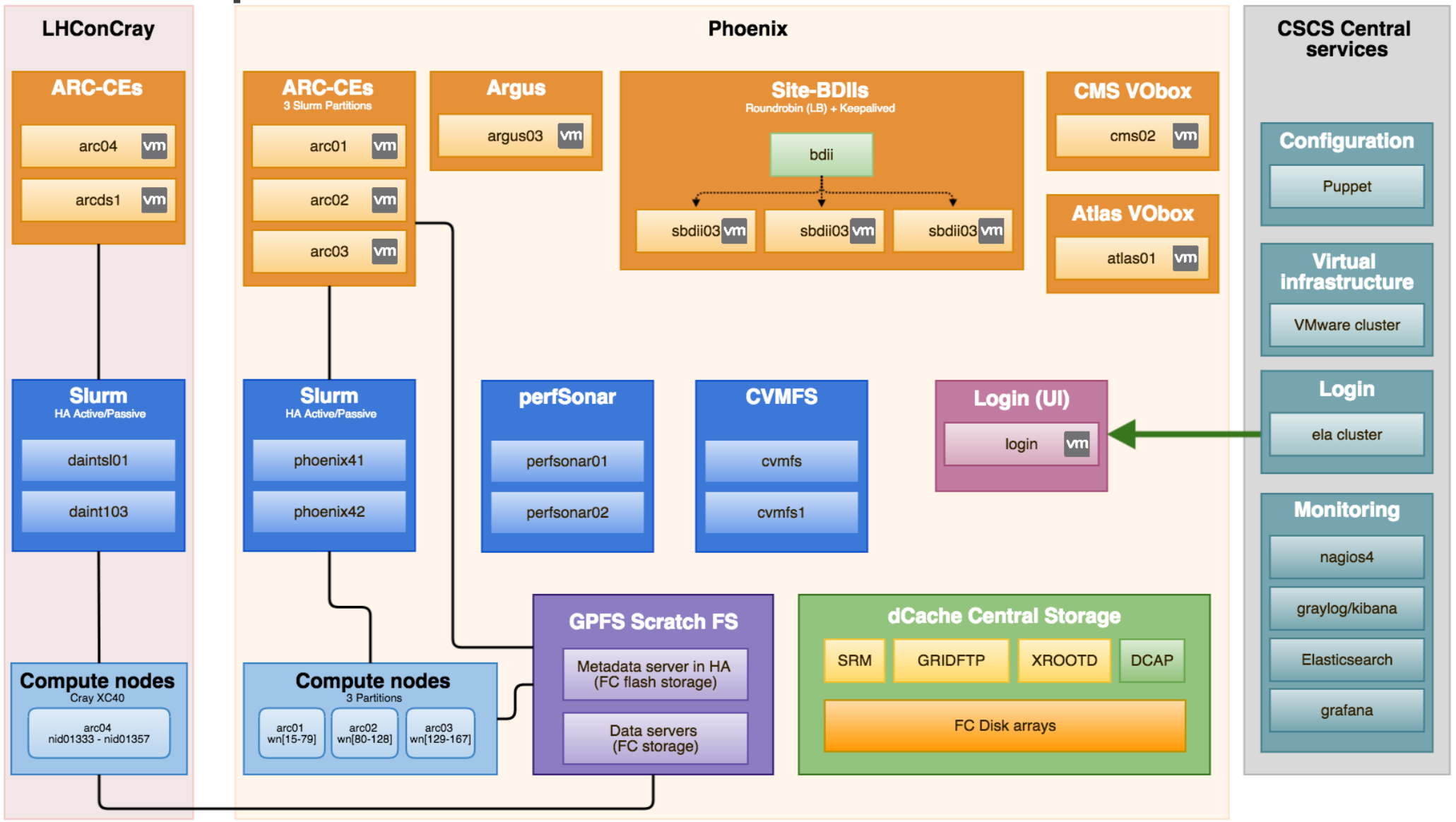


T2 Space allocation %

Is it a problem if T0 space will stay until April ‘19?

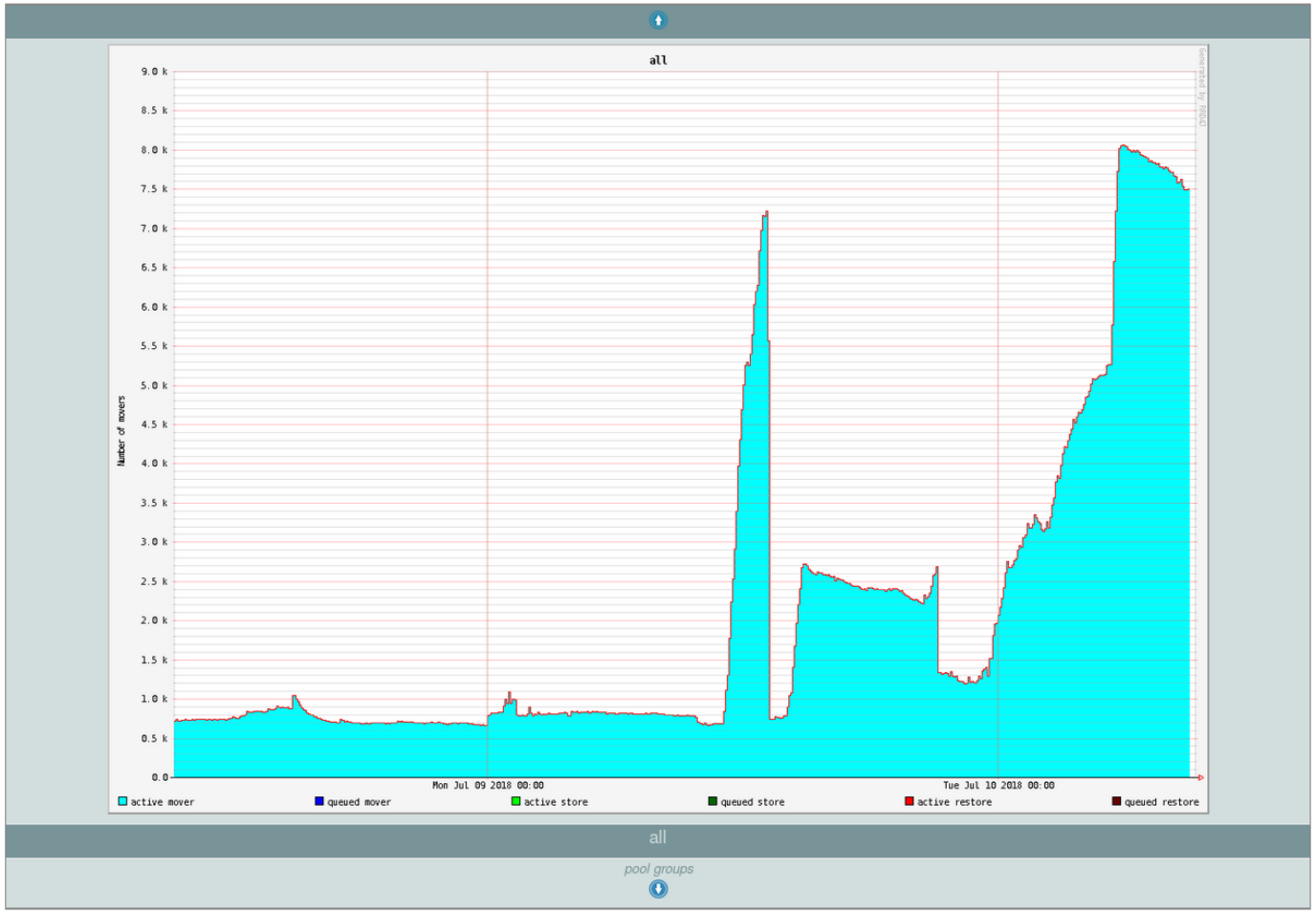
Operations

Operations – Updates: CHiPP Services Overview



dCache in 2018

- IPv6
  - All nodes moved to Ipv6
  - All nodes moved to Ethernet 25G
- Upgrade to 3.2



Ticket report 2018

Total of 80 tickets

- **ATLAS: 15 tickets**
- **CMS: 57 tickets**
  - 23 Phedex related (CMS02)
  - 7 Proxy related (CMS02)
- **LHCb: 8 tickets**

Ticket report 2017

Total of 56 tickets

- **ATLAS: 10 tickets**
- **CMS: 32 tickets**
  - 5 Phedex related (CMS02)
  - 2 Proxy related (CMS02)
- **LHCb: 14 tickets**

LHCb Issue

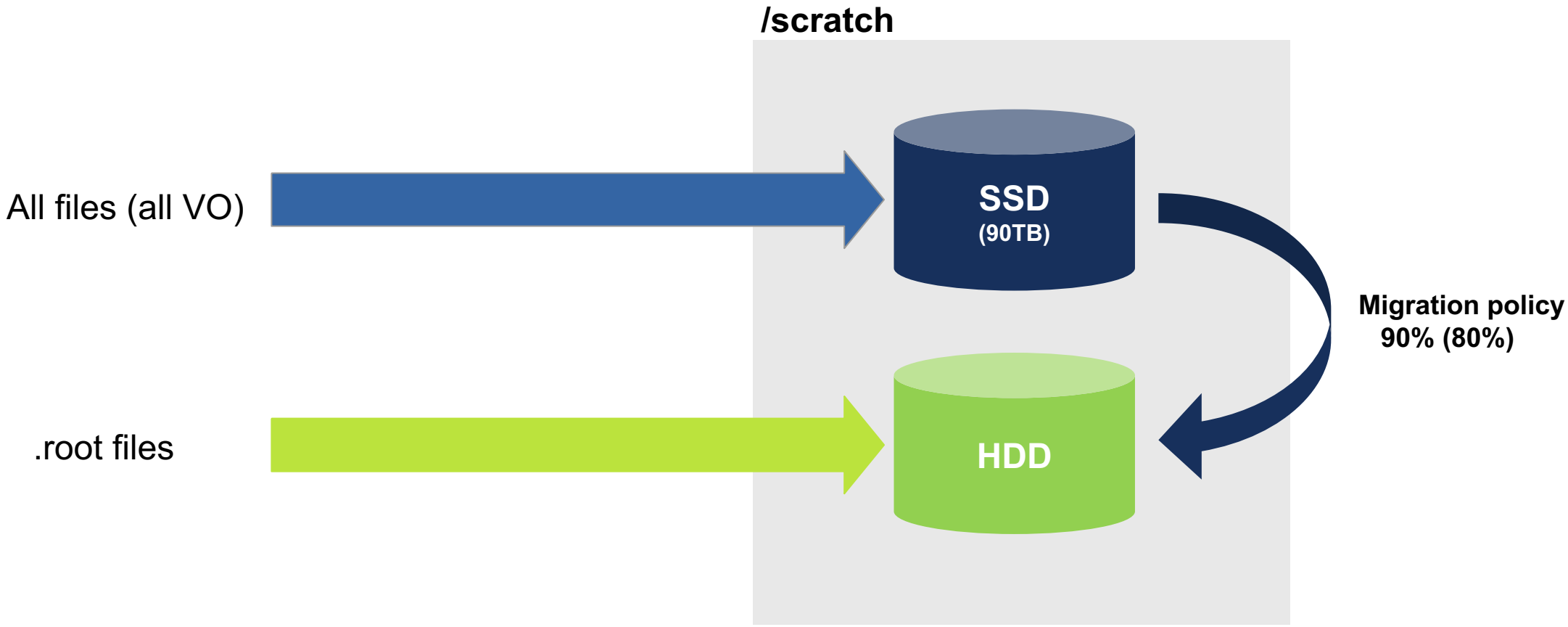
- We have had a few tickets regarding LHCb jobs stalling on Piz Daint
- There were two main reasons for this:
  - Incorrect publication of `{machine,job}features` scripts (CSCS). This are not used on Phoenix, yet they don't seem to affect LHCb jobs there.
  - Incorrect definition of the queue (LHCb)

- In both cases extensive debugging took place and took many, many hours of digging thru the LHCb job environment to get these two figured out

- Having a better understanding of the LHCb workflow would help to minimize downtime and debugging efforts

- This can be extensible to all VOs...

Scratch Filesystem



Scratch Filesystem – Performance

Other

- LHC on Cray advertising
  - HEPiX (Marco, Gianfranco)
- PerfSonar on IPv6
  - Old PerfSonar (40G) updated with dual stack Ipv4 and IPv6
  - New server (100G) will be ready soon
- VO Boxes
  - More VO Boxes for CMS (Do we need them?)
  - IPv6 ? (CMS02 is already on Ipv6)
- DVS Issue:
  - Bad Uplink caused DVS performance to be really bad and fail/drop connections in a non-consistent way. Very tricky to detect
- Argus host at CSCS (argus03): It does not seem to be used since early 2018. Is the service needed any more?

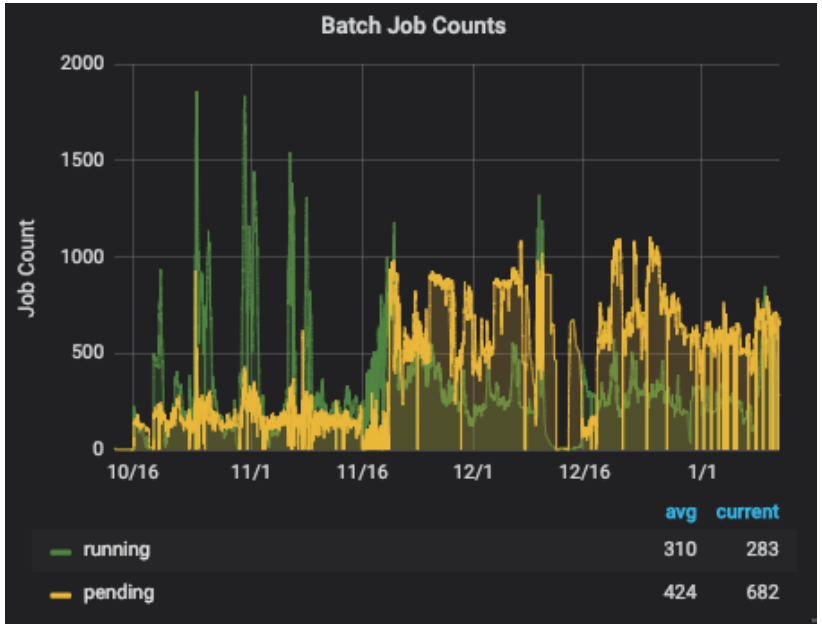
Fair Share

- A significant effort was invested, findings presented at special meeting
- Changes Since That Time…
  - Wallclock limit set to 2 days
  - Least Loaded Node (LLN) set to NO

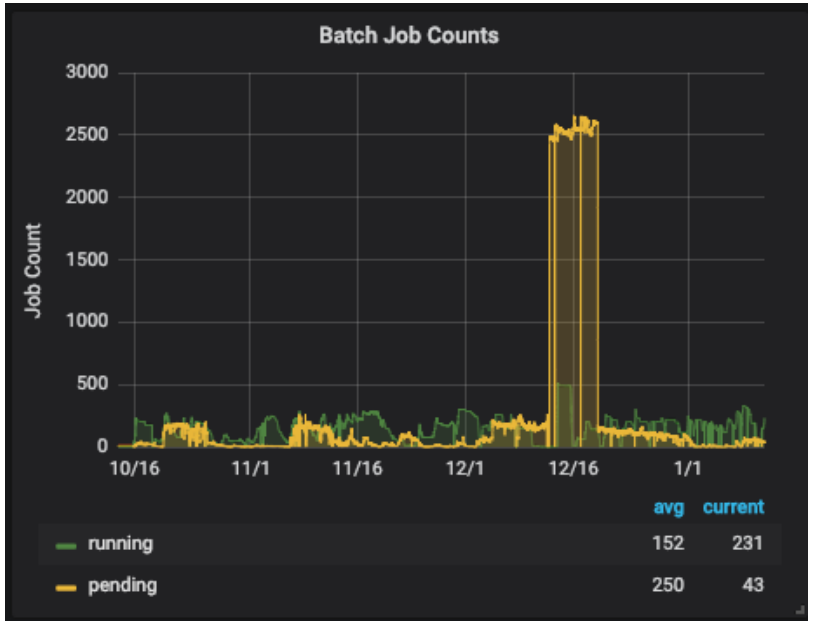
Batch Job Counts

Indication of how much work is being sent to Piz Daint.

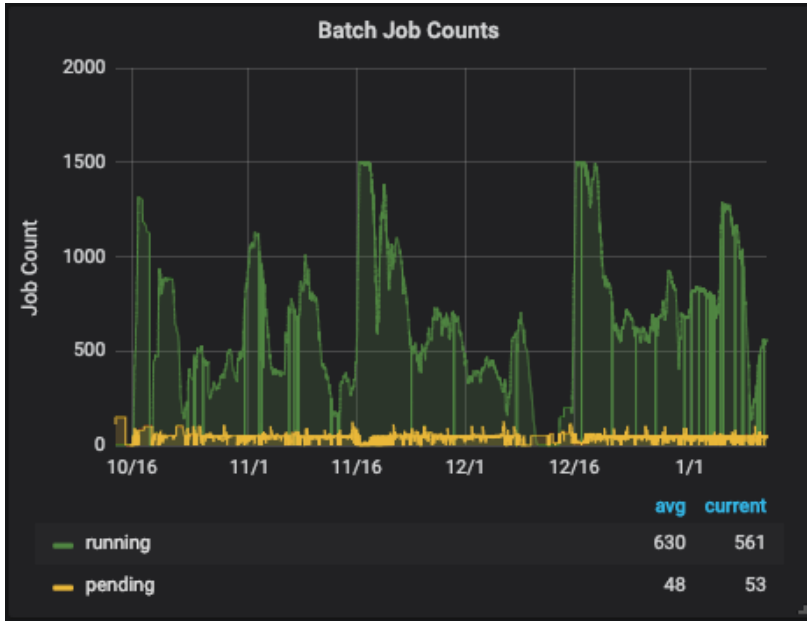
Atlas



CMS



LHCb

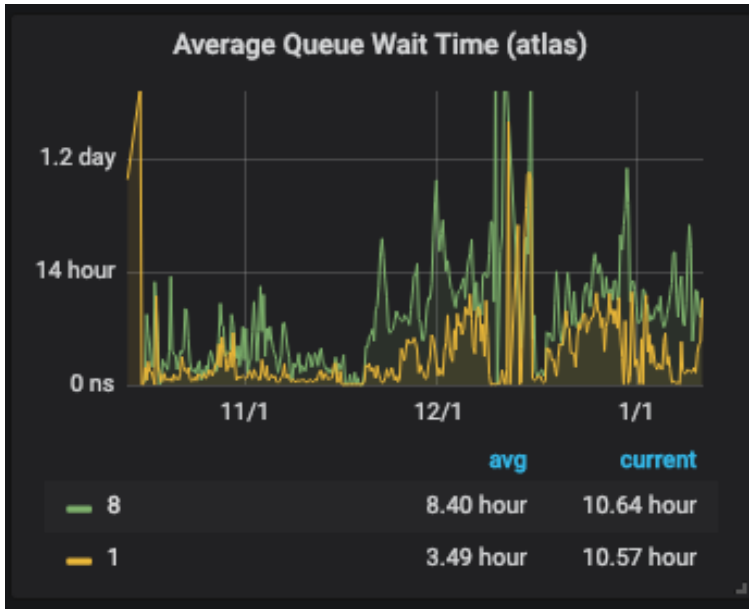


Job Submission Increased

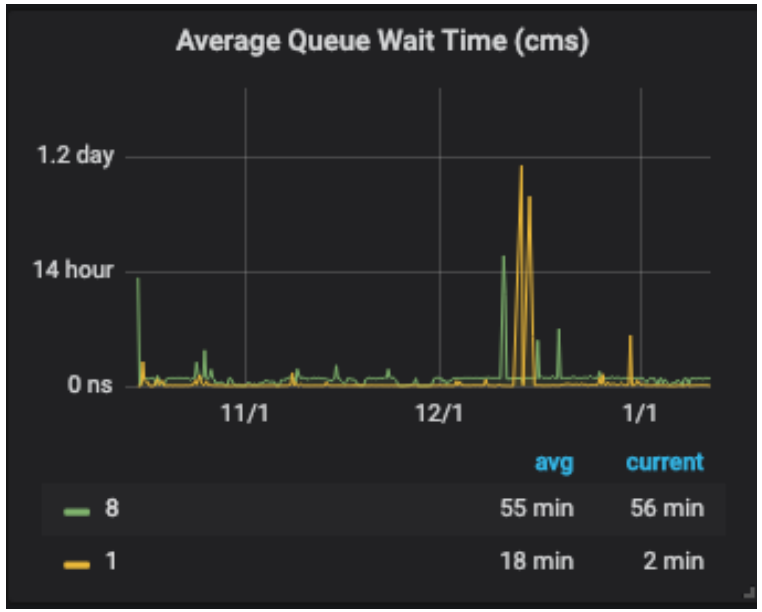
Piz Daint Only

Average Queue Wait Time

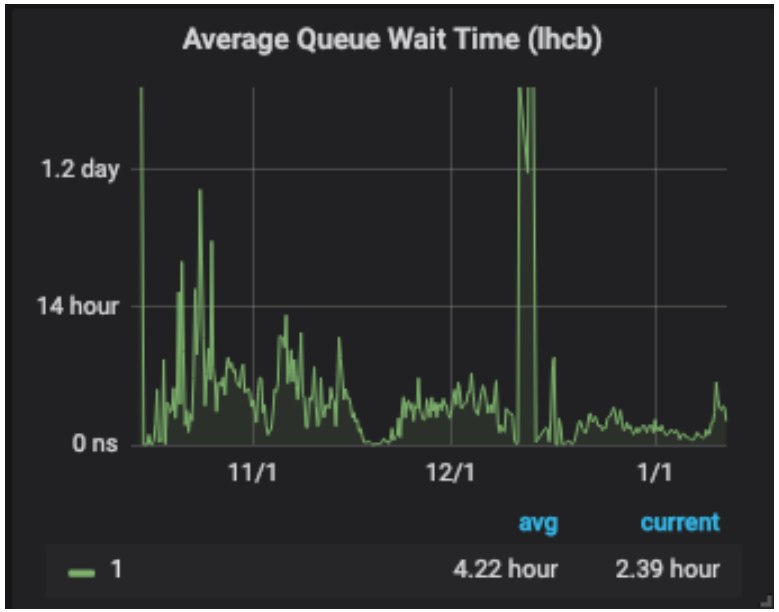
Atlas



CMS



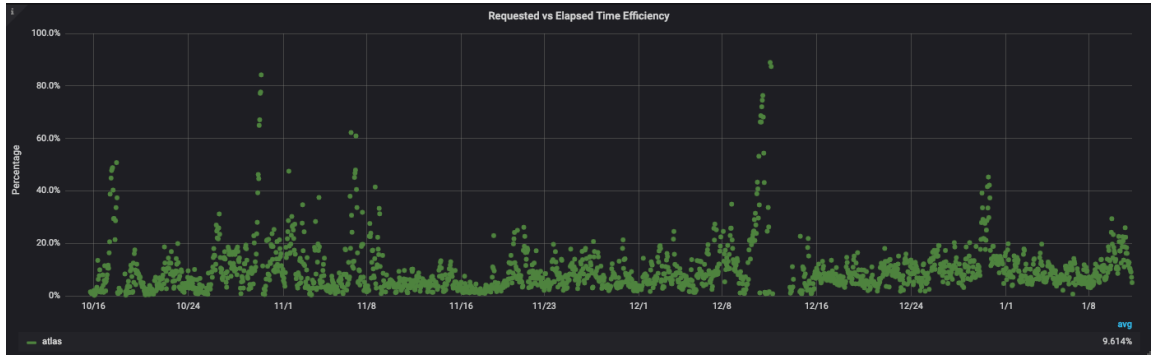
LHCb



Queue Wait time increased with increased job submission rate

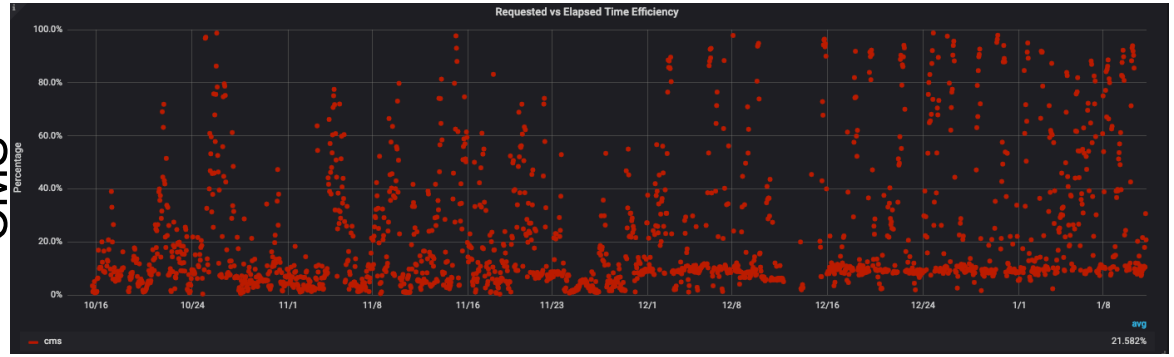
Requested vs Elapsed Efficiency

Atlas



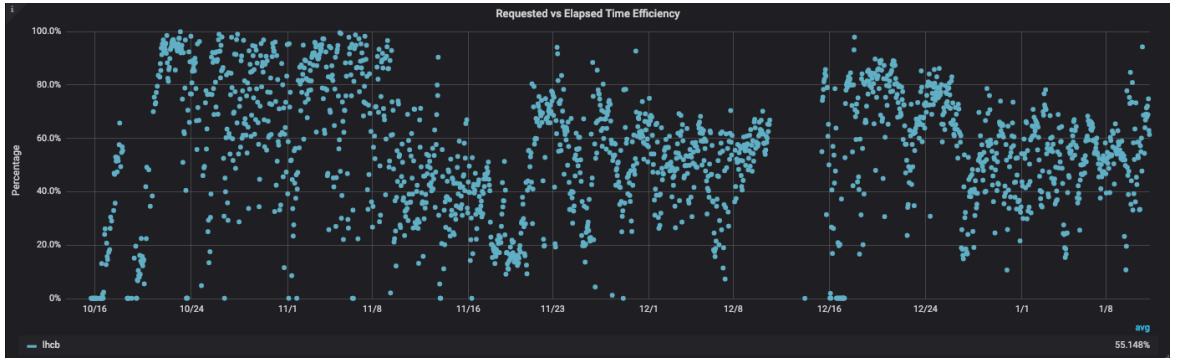
Ratio of requested to Used time is still low

CMS



Each point is 1 job. The higher the point on the graph, the more the requested and elapsed times are equal. Higher is better.

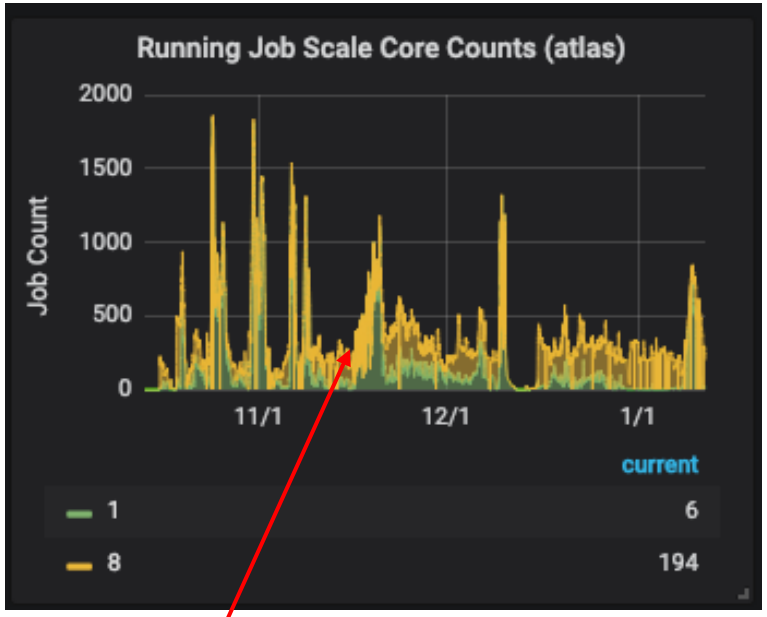
LHCb



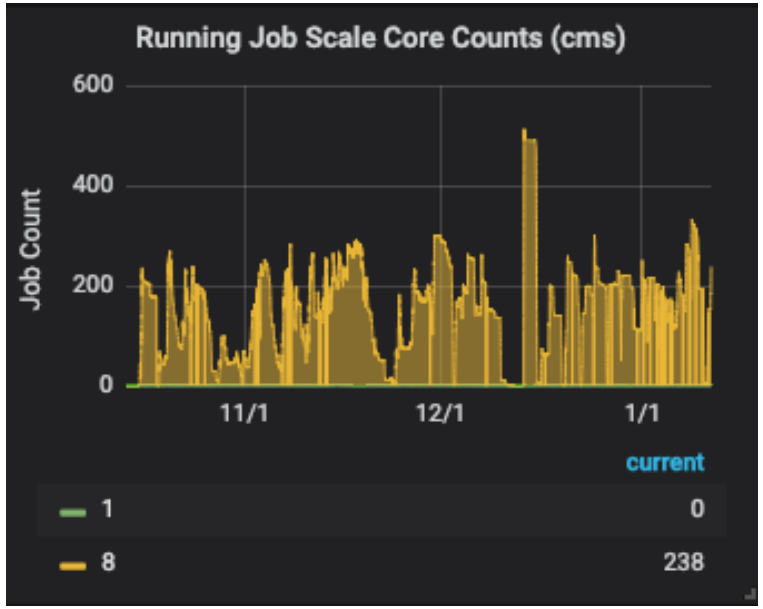
Piz Daint Only

Running Core Counts

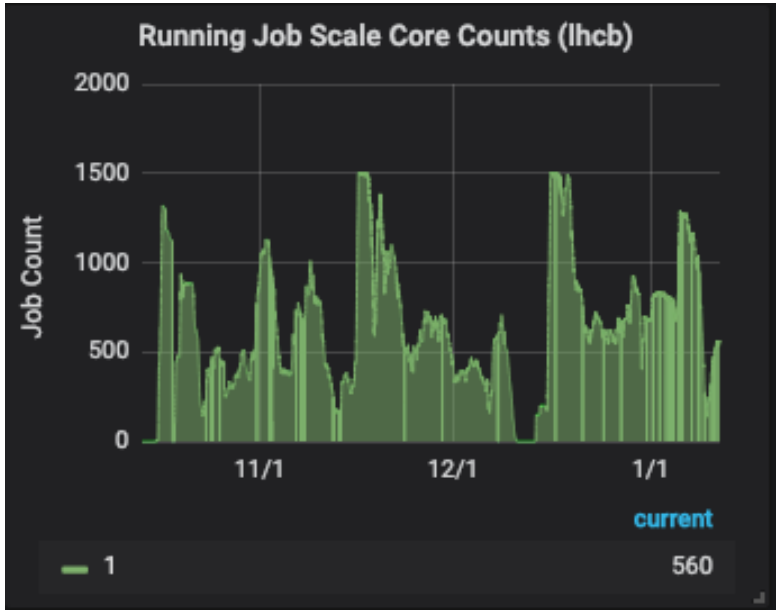
Atlas



CMS



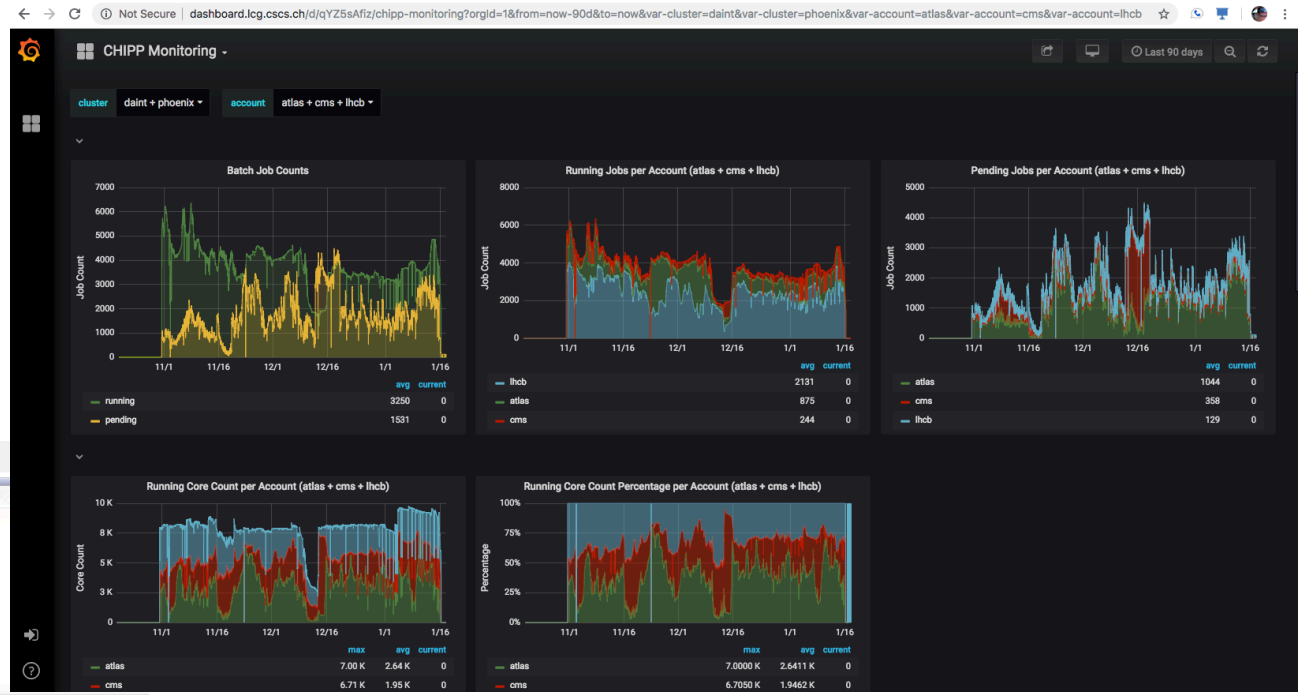
LHCb

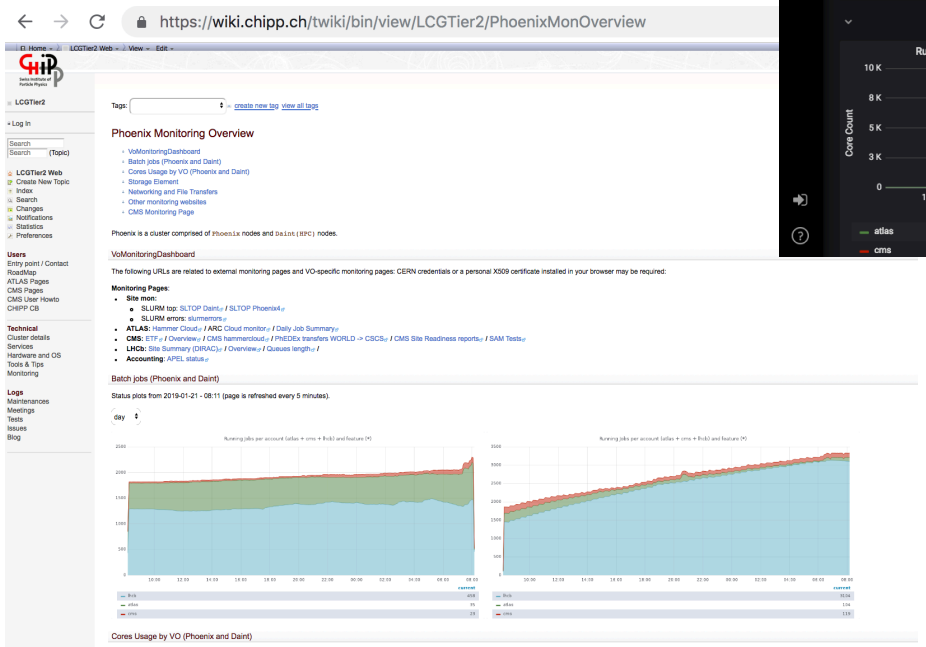


Atlas Backlog Increased

VO Rep one-on-one followups

- Dashboard released
- CHiPP Wiki Monitoring Page: cleaned up and updated



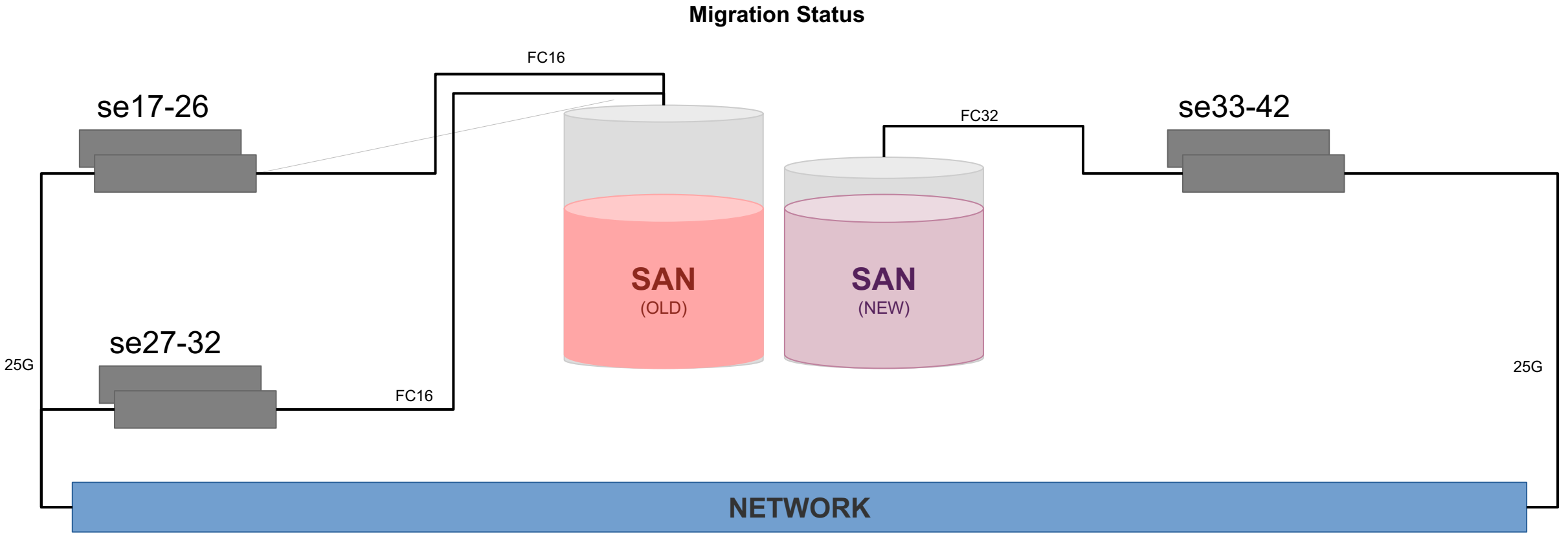


Plans

dCache 2019 migration

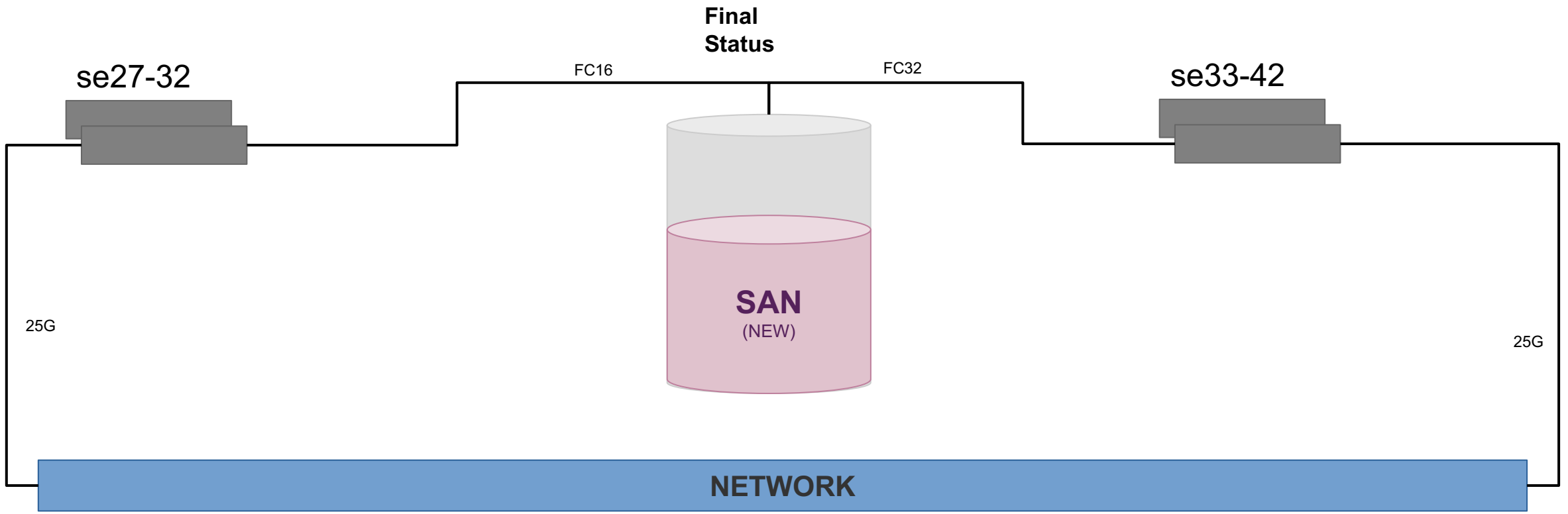
Step 1: Deploy se33-42 on the new SAN and migrate data from se17-32 (starting from se27-32)

*Free space could be limited for some days



dCache 2019 migration

Step 2: Deploy se27-32 on the new SAN and restore the full space
Step 3: Complete migration and decommission se17-26



Phoenix Compute Shutdown Schedule

	Daint Nodes		ARC01 Nodes		ARC02 Nodes		ARC03 Nodes	
Date	Change	Total	Change	Total	Change	Total	Change	Total
Feb. 04	+30	94	-15	50	-10	30	-10	30
Feb. 11	+30	124	-20	30	-10	20	-10	20
Feb. 18	+24	148	-20	10	-10	10	-10	10
Feb. 25	+24	172	-10	0	-10	0	-10	0

Future

Operations:

1. Reinstall ARC for Daint
2. Sarus
3. ARC HA
4. Services reinstall on new Puppet

Daint update:

- CLE update (April ‘19, October ‘19)

Center-wide outage

- May ’19
- Sept ’19

Events:

- HEPIX (San Diego March ‘19 – Amsterdam Oct ‘19)
- dCache (Madrid May ‘19)
- ARC events?

Resource Planning

Compute 2019

- 127 kHS06 Installed
- 110 kHS06 pledged
- Split TBD ?

Storage 2019

- 4600 TB installed
- Split TBD ?

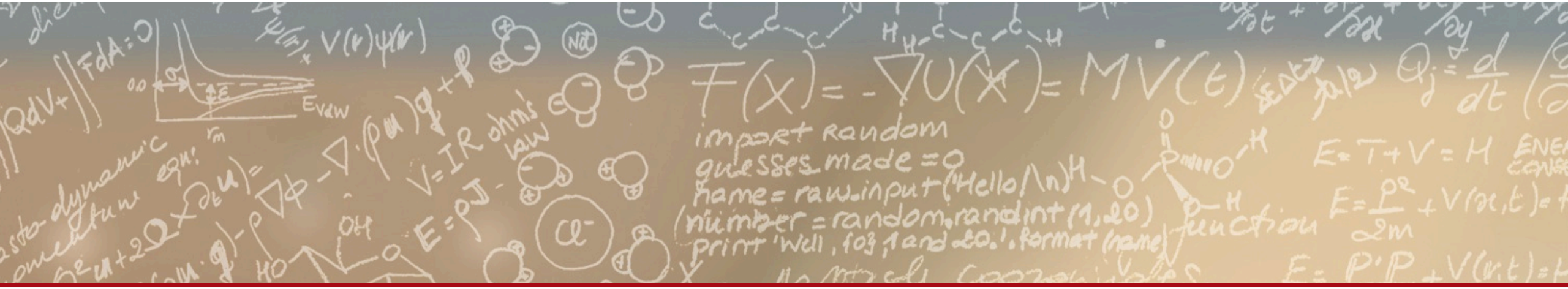
	Jan-16	Jan-17	Jan-18	Jan-19	Jan-20		
	(DONE ALREADY)	(DONE ALREADY)					
	LOC Pilot	LHC on CRAY (+27%)	LHC on CRAY (+61.29%)	LHC on CRAY (Phoenix + 104.83%)	LHC on CRAY (Phoenix + 160.14%)		
	Phoenix	Phoenix (No growth)	Phoenix (No growth)				
Phoenix Compute HS06	69000	66930	64929	0	0	Decay over the time	
Phoenix Storage TB	3070	2715	2106	1057	0	Decay over the time	
Phoenix FTE (CHIPP)	1.50	1.00	1.00				
Phoenix FTE (ETHZ) - not paid	1.00	1.00	1.00	1.00	1.00		
Shared Compute HS06		20500	46116	141016	179090		
Shared Storage TB		1300	1910	3560	5310		
Shared FTE (CHIPP)		0.50	0.50	1.50	1.50		
Shared Nodes		25	57	172	219		
Shared Cores		1600	3648	11008	14016		
Total FTE (CHIPP)		1.50	1.50	1.50	1.50		
						CHF 4 years	CHF / year
Shared Compute CHF		65700	149795	452016	575532	1243044	310761
Shared Storage CHF		80600	118420	176588	263368	638976	159744
Shared Service/MW/Netw CHF		40000	40000	40000	40000	160000	40000
Phoenix maintenance CHF		50000	30000	150000		230000	57500
Phoenix + Shared FTE CHF		210000	210000	210000	210000	840000	210000
TOTAL COSTS CHF		446300	548215	1028604	1088900	3112020	778,005
Deviation		333700	231785	-248604	-308900	Current (Jan-17) funds:	780,000
TOTAL Compute HS06	69000	87430	111036	141016	179090		
TOTAL Storage TB	3070	4015	4016	4617	5310		
kHS/TB ratio	22	22	28	31	34		

27.0%	Annual Growth (Compute resources)
15.0%	Annual Growth (Storage resources)
820	HS06/node
64	Cores/node
2628	CHF/node/year (without electricity, cooling nor manpower, 12[illegible] GB RAM)
62	CHF/TB/year 2017-2018 (without electricity, cooling nor manpower)
50	CHF/TB/year 2019-2020 (without electricity, cooling nor manpower), tentative
140000	CHF/FTE

COMPUTE	Effective Capacity	78,687	99,932	126,914	161,181
	Pledges	78,000	96,000	110,000	124,000
STORAGE	Effective Capacity	4,015	4,015	4,617	5,310
	Pledges	4,000	4,000	4,300	4,300

PHOENIX	*Compute effective Capacity*	*79.091*	*97.27*	*110.909*	*127.273*
	Storage Capacity	*3500*	*400*	*4300*	*4800*

cscs



Thank you for your attention.