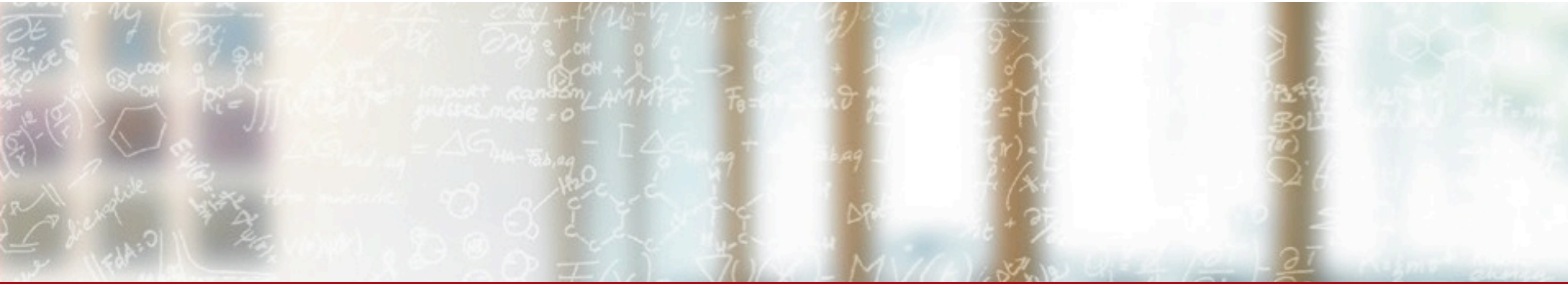




CSCS

Centro Svizzero di Calcolo Scientifico
Swiss National Supercomputing Centre

ETHzürich



LHConCRAY

Acceptance Tests 2017 – Run4 System Report

Miguel Gila, CSCS

August 03, 2017

Table of Contents

1. Changes since Run2/3
2. DataWarp
3. Current configuration
4. System statistics
5. Next steps



CSCS

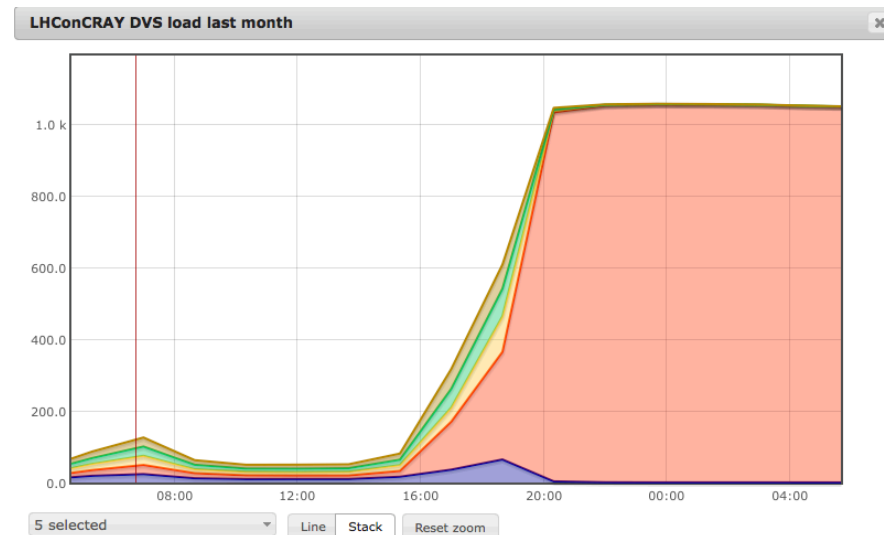
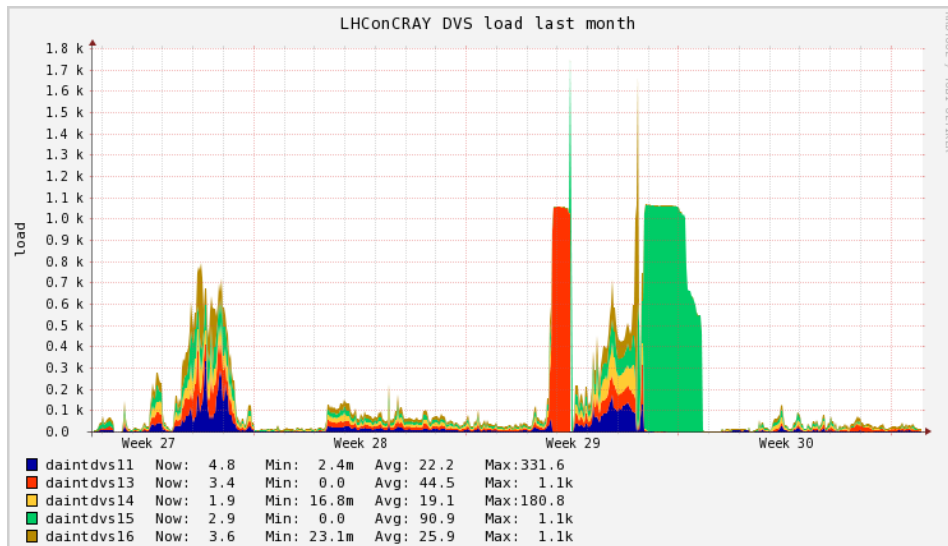
Centro Svizzero di Calcolo Scientifico
Swiss National Supercomputing Centre

ETH zürich

Changes since Run2/3

Problems with DVS

- DVS load
 - At times load on DVS would shift from being almost balanced to target a single node
 - Cray identified this to be a GPFS problem (being unresponsive).
- DVS hung and XFS filesystems on files crashing
 - When a DVS server is unavailable, CNs use another one after a non-configurable timeout
 - However, when an XFS filesystem is mounted off the failed DVS server, caching and open descriptors often corrupt the xfs filesystem
 - This is, at times can go unnoticed and cvmfs continue serving files with incorrect checksums
- These two problems have affected us in the evaluated period at least 2 times

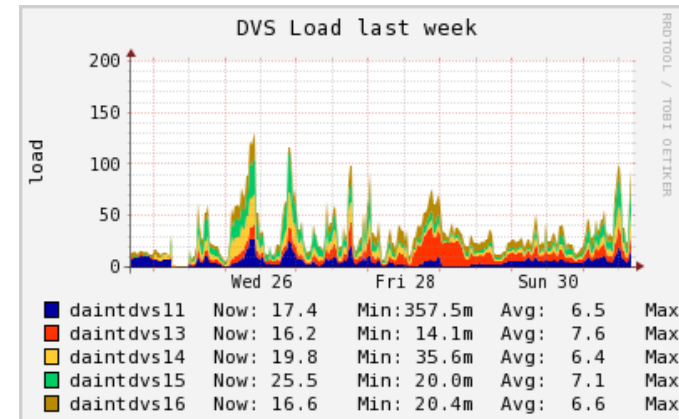
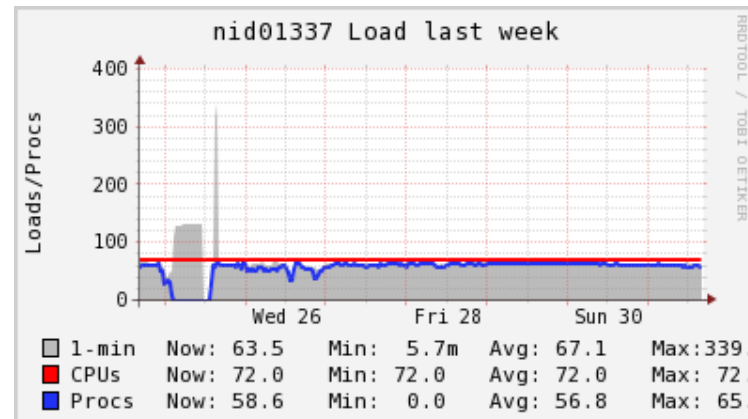


Changes due to DVS issues 1/2 (CVMFS)

- CVMFS was originally configured to use a XFS filesystem on top of DVS as local cache
- However, CVMFS developers created two new features that allow us to store data directly onto DVS projected filesystems (no more XFS!):
 - Workspace: since DVS does not support `flock()`, with the workspace setting it is now possible to set all locks relative to the cache local to the node (on ramdisk)
 - Tiered cache with in-ram storage: with enough RAM, it's now possible to instruct cvmfs to store its cache on memory, without the need for local storage. This can dramatically increase performance
 - In order to reduce repetitive queries outside, all relevant cvmfs repos have been preloaded on GPFS thanks to a new, fast service provided by CERN for HPC sites. This gets updated several times a day. Note that cvmfs is smart enough to query outside if the file is not on the preloaded cache
- These two changes were applied on all 25 CNs on **July 19**

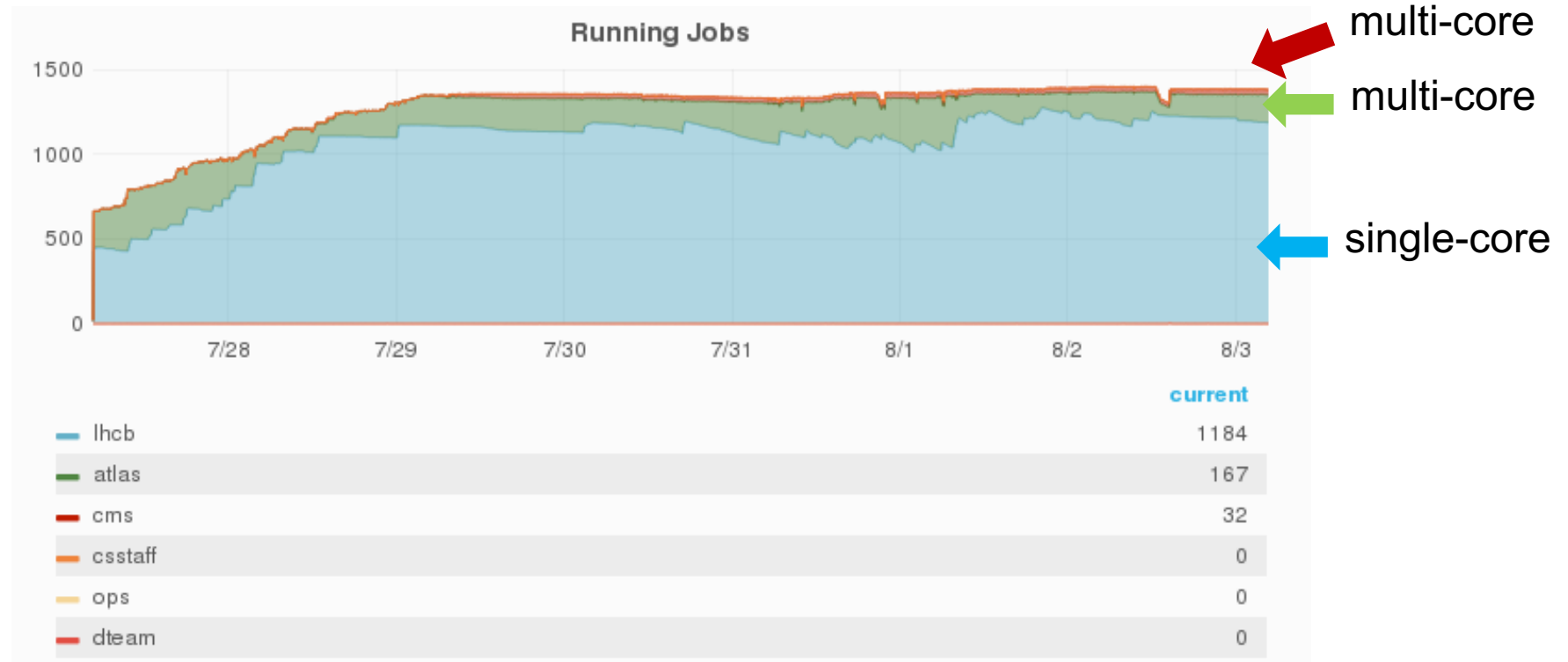
Changes due to DVS issues 2/2 (DVS itself)

- As it was configured earlier, DVS client would calculate the hash of a file (or a block) and pick a DVS server of the list of available nodes.
- As reported, at times IO would shift and target just a single DVS node. It could be a group of jobs doing a lot of metadata operations on blocks with the same hash, or on similar paths.
- In order to reduce the chances of this happening, we forced the this hash to be calculated per node, so each CN will target one DVS no matter what file/block it is trying to use.
- As a result, we can see load peaks on CNs, but not on DVS servers.
- This was introduced to all CNs on **July 25**.



CMS started running production jobs

- On July 28 CMS started running almost production jobs (empty payloads)
- On July 31 CMS started running real production jobs
 - About 30 jobs are in the system at all times



Public monitoring

- (still) Working on a proper dashboard
- For now a Twiki page with public charts is enabled on <https://wiki.chipp.ch/twiki/bin/view/LCGTier2/LHConCRAYMonitoring>

- Links to VO monitoring pages
- Selectable time window
- Notably missing per-VO CPU allocation statistics.

- Instant view (SLTOP) available on http://ganglia.lcg.cscs.ch/ganglia/sstop_lhconcray.html

LHConCRAY MonOverview

- External links
- Jobs
- System totals (Grafana)
 - Utilization
 - CPU statistics
 - Queue status
 - VO totals (Grafana)
 - Running jobs per VO
 - Pending jobs per VO
 - Node performance (Ganglia)
 - Backers
 - Compute nodes
 - How this page is generated

External links

- SLTOP: http://ganglia.lcg.cscs.ch/ganglia/sstop_lhconcray.html/
- Common ETP view: https://efl-11.com.ch/eflcheck_mk/index.py?view_name=searchhost&host_regenar04.lcg.cscs.ch&filter_inifirst#/ (not working)
- ATLAS:
 - Last 12h of ParDA jobs in CSCS-LCG2-HPC: https://bigpanda.com.ch/job/compilinggate-CSCS-LCG2-HPC&hours=12&display_min=100/
 - Last 12h of ParDA jobs in ANALY_CSCS-LCG2-HPC: https://bigpanda.com.ch/job/compilinggate-ANALY_CSCS-LCG2-HPC&hours=12&display_min=100/
 - ParDA CSCS-LCG2-HPC: <https://bigpanda.com.ch/job/CSCS-LCG2-HPC/>
 - ParDA CSCS-LCG2-HPC_MCOORE: https://bigpanda.com.ch/job/CSCS-LCG2-HPC_MCOORE/
 - ParDA ANALY_CSCS-HPC: https://bigpanda.com.ch/job/ANALY_CSCS-LCG2-HPC/
- HammerCloud full site last 7 days: <http://hammercloud.com.ch/ho/2016/09/07/hammercloud-view?site=CSCS-LCG2&start=Time&end=Time&templateType=Golden>
- ETP: https://efl-11.com.ch/eflcheck_mk/index.py?start_url=%2Feflcheck_mk%2Fview.py%3Fview_name%3Dhost%2Bhost%3Dcar04.lcg.cscs.ch%2Bstate%3D/

CMS:

- Gridin configuration: http://gridin-3b.grid.in.edu/factory/moonrakerFactoryEntryStatusNow.htm?Entry=CMSHTPC_T2_CH_CSCS_0n04
- HammerCloud (requires CMS account): <http://hammercloud.com.ch/ho/2016/09/07/>
- ETP: https://efl-11.com.ch/eflcheck_mk/index.py?start_url=%2Feflcheck_mk%2Fview.py%3Fview_name%3Dhost%2Bhost%3Dcar04.lcg.cscs.ch%2Bstate%3D/
- Site Status Board: http://statb-ssb.com.ch/eflcheck_mk/index.py?view=View-Stat-Readiness&search_0=CSCS/
- Site Readiness: http://statb-ssb.com.ch/eflcheck_mk/index.py?view=View-Stat-Readiness&search_0=T2_CH_CSCS/

LHC:

- ETP: https://efl-11.com.ch/eflcheck_mk/index.py?start_url=%2Feflcheck_mk%2Fview.py%3Fview_name%3Dhost%2Bhost%3Dcar04.lcg.cscs.ch%2Bstate%3D/

Plots

Note, the value for current may be incorrect. Ignore it.

week 2

System totals (Grafana)

Utilization

Utilization (CHIPP nodes)

min	max	avg	current
49%	89%	80%	89%
89%	89%	89%	89%
89%	89%	89%	89%
89%	89%	89%	89%
49%	89%	80%	89%
89%	89%	89%	89%
89%	89%	89%	89%

SLTOP

Version: 0.8.10/daint
Page last updated at Thu Aug 3 11:49:04 CEST 2017

1488 jobs running/147 jobs pending || CPUs 1600/1800 (88.88%)

Job summary

Partition	Running	Pending
wlog	1488	146
Total:	1488	147

VO

	Running	Pending
atlas	107	91
cms	16	4
lhcb	1365	50
ops	0	1
dteam	0	0

Resource allocation per partition

PARTITION	AVAIL	TIMELIMIT	CRUS(A/T/O/T)	NODELIST
wlog	5-45-05-00	up	1600/200/0/1800	nid(01333-01357)

Resource allocation per DR

Running jobs:

- 16 dn-/DC-ch/DC-cern/OD-computers/CN-ompiplot02/vocms080.cern.ch
- 107 dn-/DC-ch/DC-cern/OD-Organic Units/OD-Users/CN-atlas1/CN-55105/CN-Robot: ATLAS_ACT 1
- 1365 dn-/DC-ch/DC-cern/OD-Organic Units/OD-Users/CN-romanov/CN-427293/CN-Vladimir Romanovskiy

Pending jobs:

- 1 dn-/DC-ch/DC-cern/OD-computers/CN-ompiplot02/vocms080.cern.ch
- 91 dn-/DC-ch/DC-cern/OD-Organic Units/OD-Users/CN-ellect1/CN-55105/CN-Robot: ATLAS_ACT 1
- 4 dn-/DC-ch/DC-cern/OD-Organic Units/OD-Users/CN-lbdiaz/CN-574609/CN-Robot: LHCB_Direct Service Account
- 45 dn-/DC-ch/DC-cern/OD-Organic Units/OD-Users/CN-romanov/CN-427293/CN-Vladimir Romanovskiy
- 3 dn-/DC-ch/DC-cern/OD-Organic Units/OD-Users/CN-sciaba/CN-43078/CN-Andreas Sciaba
- 1 dn-/DC-BU/DC-BOI/C-BR/O-Robots/O-SRCE/CN-Robot:argo-egifcro-ngr

Node status

Node	MEM	CPU	A/T	Load	State
nid01333	0%	64/72	(56.77)	MIXED	
nid01334	0%	64/72	(60.35)	MIXED	
nid01335	0%	64/72	(63.35)	MIXED	
nid01336	0%	64/72	(64.02)	MIXED	
nid01337	0%	64/72	(64.23)	MIXED	
nid01338	0%	64/72	(64.25)	MIXED	
nid01339	0%	64/72	(61.13)	MIXED	
nid01340	0%	64/72	(64.33)	MIXED	
nid01341	0%	64/72	(64.31)	MIXED	
nid01342	0%	64/72	(64.34)	MIXED	
nid01343	0%	64/72	(59.71)	MIXED	
nid01344	0%	64/72	(60.26)	MIXED	
nid01345	0%	64/72	(61.79)	MIXED	
nid01346	0%	64/72	(60.40)	MIXED	
nid01347	0%	64/72	(51.51)	MIXED	
nid01348	0%	64/72	(64.47)	MIXED	
nid01349	0%	64/72	(64.24)	MIXED	
nid01350	0%	64/72	(64.00)	MIXED	
nid01351	0%	64/72	(64.29)	MIXED	
nid01352	0%	64/72	(62.28)	MIXED	
nid01353	0%	64/72	(64.07)	MIXED	
nid01354	0%	64/72	(53.95)	MIXED	
nid01355	0%	64/72	(56.38)	MIXED	
nid01356	0%	64/72	(60.34)	MIXED	
nid01357	0%	64/72	(63.13)	MIXED	

Running jobs:



CSCS

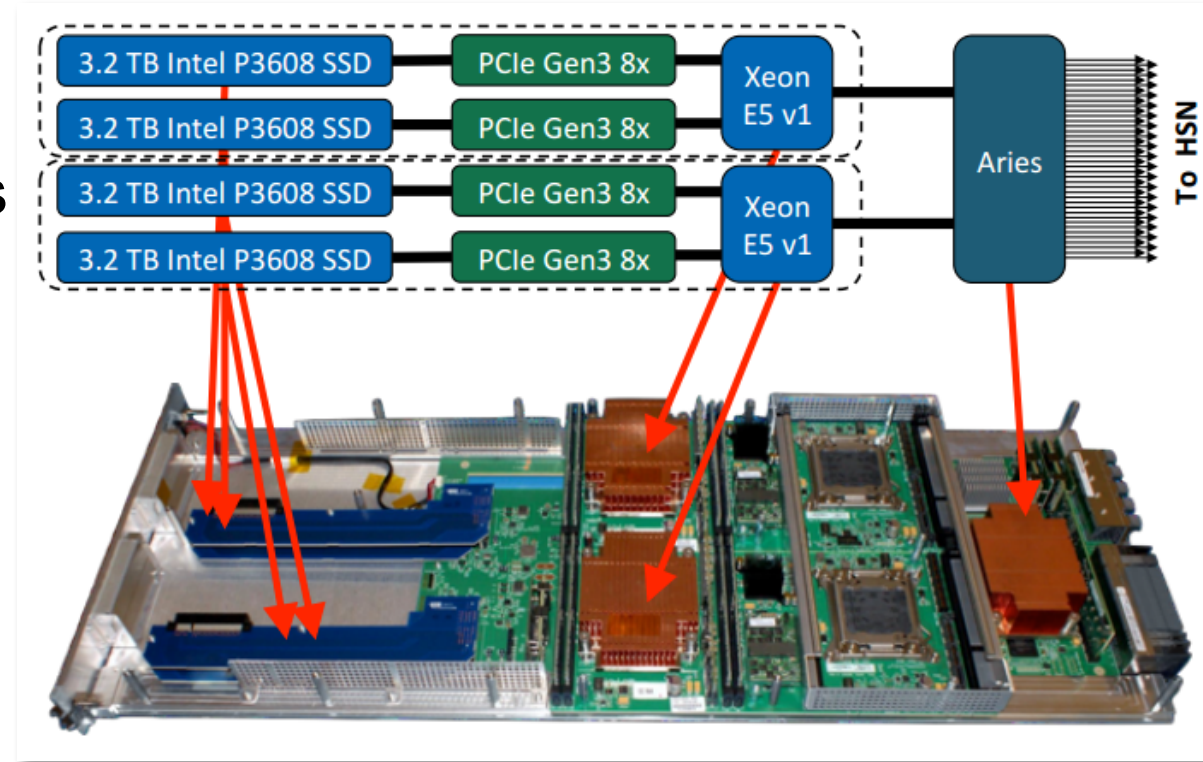
Centro Svizzero di Calcolo Scientifico
Swiss National Supercomputing Centre

ETH zürich

DataWarp

Concepts

- DataWarp is a Cray product that exposes very fast SSD-based storage to CNs
- Three main use-cases
 - Swap space
 - Scratch filesystem
 - ~~Cache (Lustre, not implemented)~~
- DWS activations can be done
 - On-demand per-job (1job/node)
 - Static, set up by an admin
- In our context, static DWS activations can be potentially used for
 - Swap
 - CVMFS
 - Scratch



Test bed

- 4 DataWarp nodes dedicated for LHConCRAY
 - 5.82TiB x 4 = 23.28TiB
- Static DWS activations were created on June 30
 - SWAP: 64 GB per node, 1.56TiB total
 - CVMFS: /dws/cvmfs (shared across all nodes), 6.25TiB total
 - SCRATCH: /var/opt/cray/dws/lhc_scratch (shared across all nodes), 15TiB total

Test bed - Swap

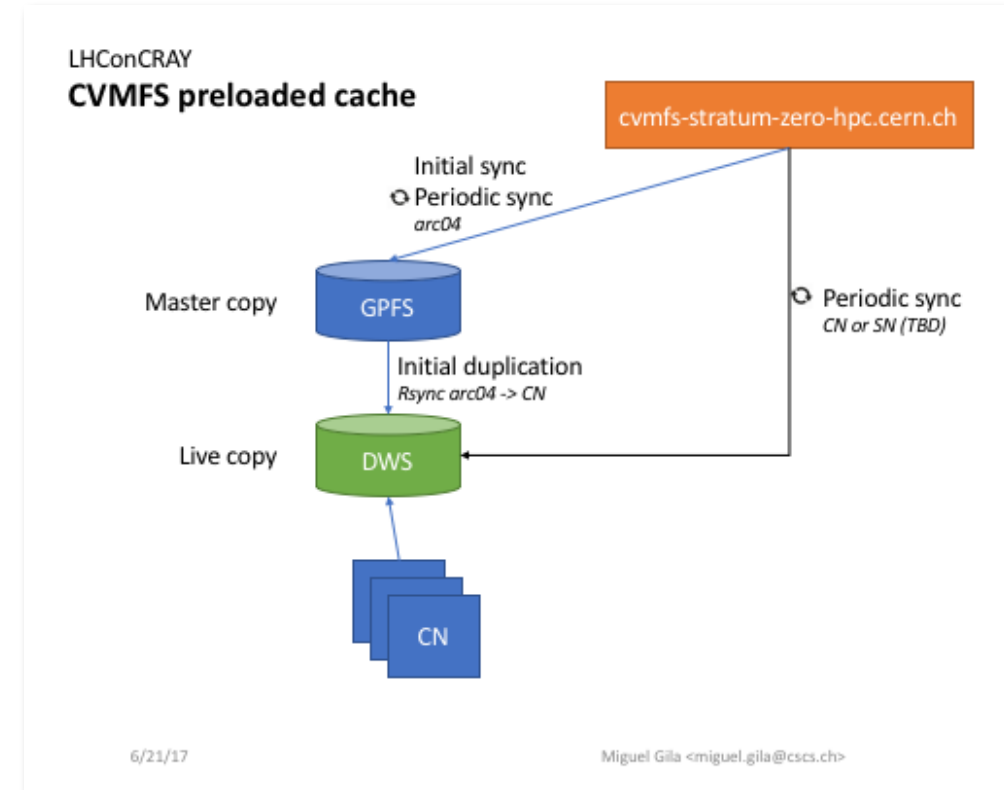
- Swap was activated on **Jun 19**
- One iSCSI device per node with 64 GB each
 - Reduced to 16GiB at the moment, more on this later
- No issues so far!

```
[544547.040890] iscsid[45170]: iscsid: Login failed to authenticate with target iqn.1988-11.com.cray.datawarp:sn80
[544547.076287] scsi 16:0:0:0: Direct-Access      LIO-ORG  FILEIO          4.0  PQ: 0 ANSI: 5
[544547.089719] sd 16:0:0:0: [sdb] 134217728 512-byte logical blocks: (68.7 GB/64.0 GiB)
[544547.102387] sd 16:0:0:0: [sdb] Write Protect is off
[544547.108990] sd 16:0:0:0: [sdb] Mode Sense: 43 00 00 08
[544547.109164] sd 16:0:0:0: [sdb] Write cache: disabled, read cache: enabled, doesn't support DPO or FUA
[544547.124060]  sdb: unknown partition table
[544547.129938] sd 16:0:0:0: [sdb] Attached SCSI disk
```

```
|nid01334:~ # free -h
|          total        used         free       shared    buffers     cached
|Mem:      126G         105G          20G         465M       1.0G        31G
| -/+ buffers/cache:      72G          53G
|Swap:      15G           708K          15G
```

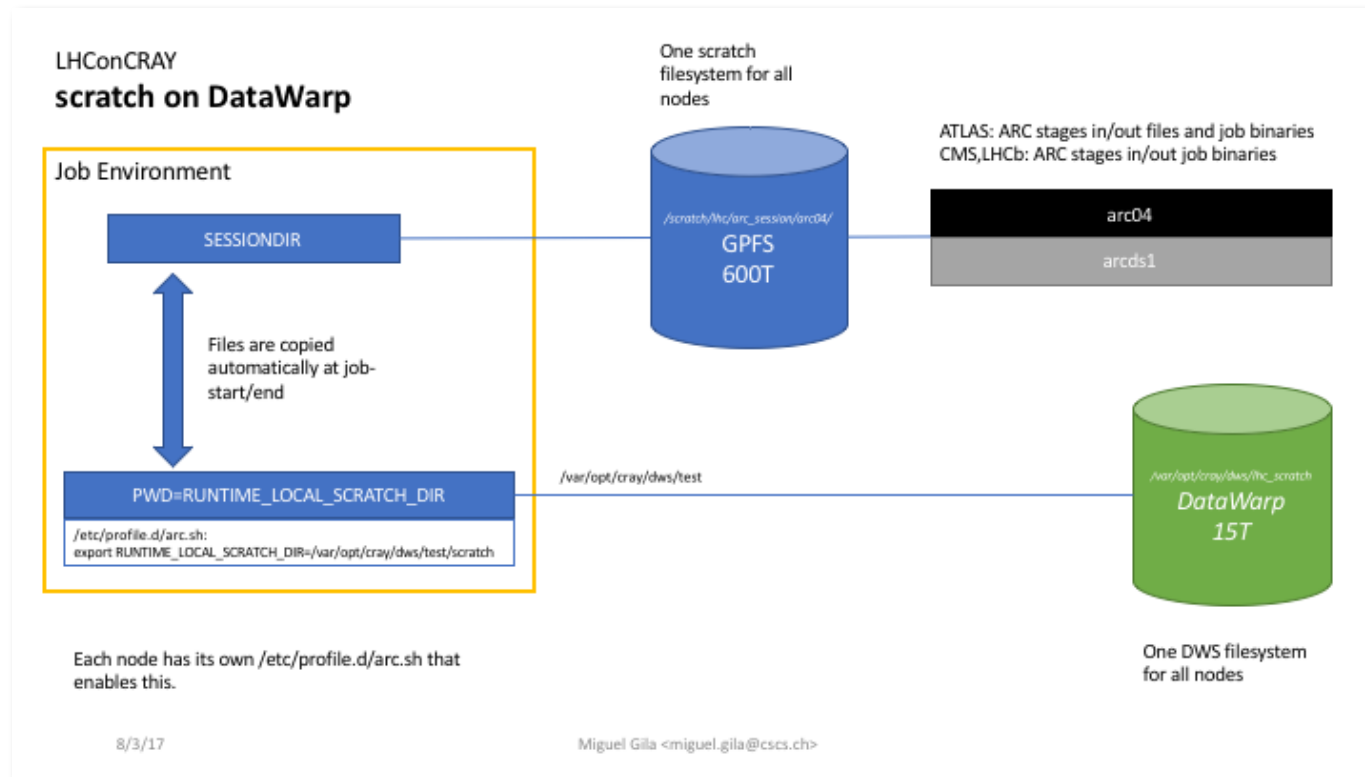
Test bed - CVMFS

- Tested on a single node (nid01357)
- Single mountpoint shared across all nodes for maximum performance and reduced data duplication
- Leveraging CVMFS tiered cache structure
- Early results showed good performance
- But after deeper testing we detected transparent **data corruption**



Test bed - Scratch

- Tested on a single node (nid01357)
- Single mountpoint shared across all nodes. Easy to setup, limited space distributed
- Using ARC's `RUNTIME_LOCAL_SCRATCH_DIR` variable that copies sessiondir to DWS
- Early results showed good performance
- **Data corruption** was detected with ATLAS jobs (2 files with same filename=data corruption)



A possible solution

- Cray detected the corruption problem and issued a Field Notice and a Patch:
- FN #6178a - DataWarp potential data corruption in CLE
- Corruption does not affect swap (they're iSCSI devices)
- Patch PS56 released and installed on Piz Daint
- Images rebuilt on 31.07.2017
- DWS server nodes for LHConCRAY rebooted with new image
- Currently **having problems activating swap** using these nodes
 - Added 16GB of SWAP from the rest of the DWS server nodes as a temporary measure
- Testing continues



CSCS

Centro Svizzero di Calcolo Scientifico
Swiss National Supercomputing Centre

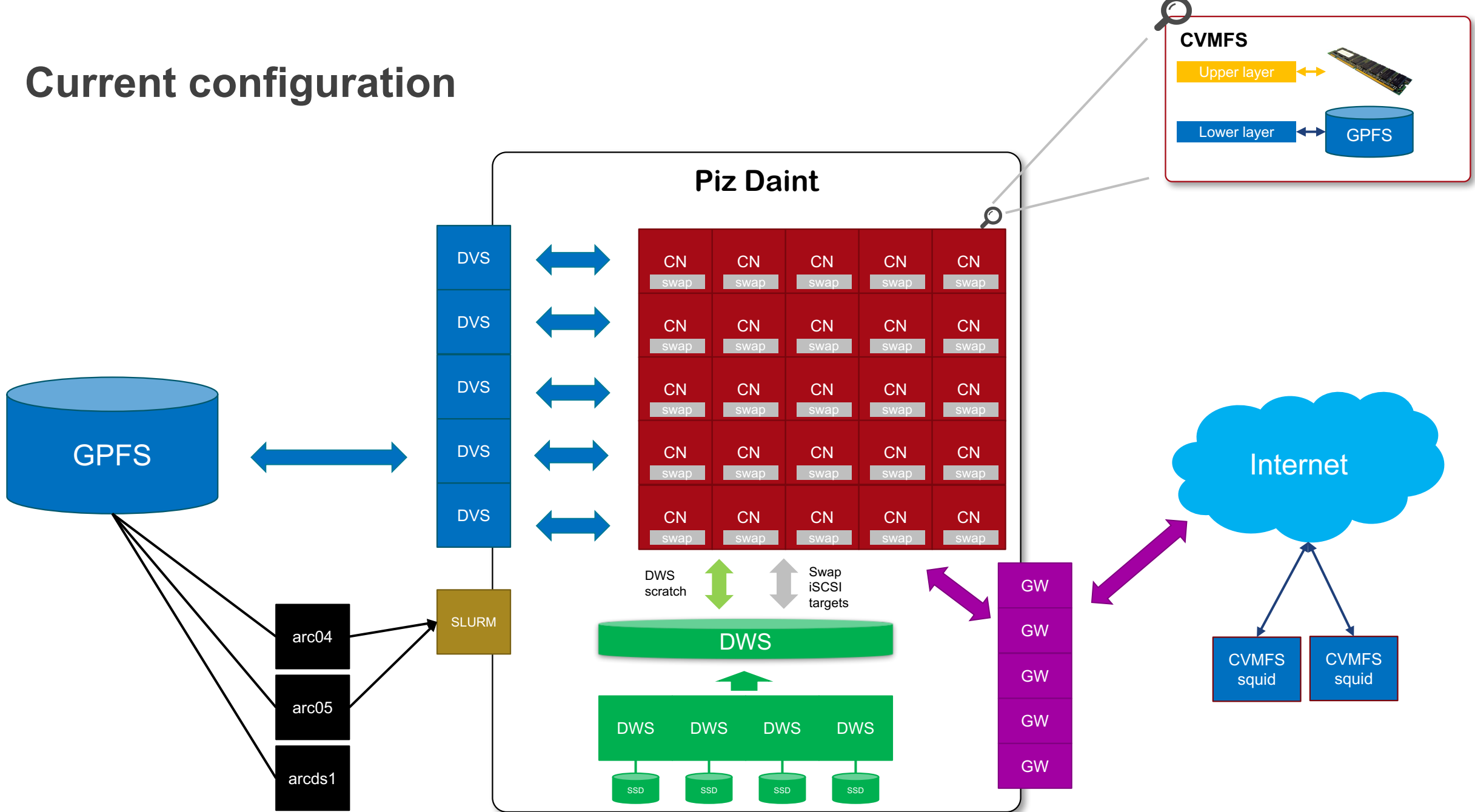
ETH zürich

Current configuration

Current configuration

- 25 compute nodes (72core, 128GB RAM, diskless)
- 1 production ARC + 1 data staging ARC + 1 test ARC (internal)
- Dedicated GPFS filesystem shared with Phoenix
- 5 DVS nodes exposing GPFS to CNs
- CVMFS running natively on CNs using tiered cache and workspaces
 - Upper layer: 6 GB in-RAM shared to all experiments
 - Lower layer: preloaded cache on GPFS, mount on CNs RO with caching enabled
- Memory limits NOT really enforced
 - Hard limit of 6000MB/core to catch rogue jobs
- **Swap** on DataWarp **enabled** (64GB 16GB per node)
- ARC caching not enabled (ATLAS)
 - Each job has a copy of the files, even if they're the same on multiple jobs.

Current configuration





CSCS

Centro Svizzero di Calcolo Scientifico
Swiss National Supercomputing Centre

ETH zürich

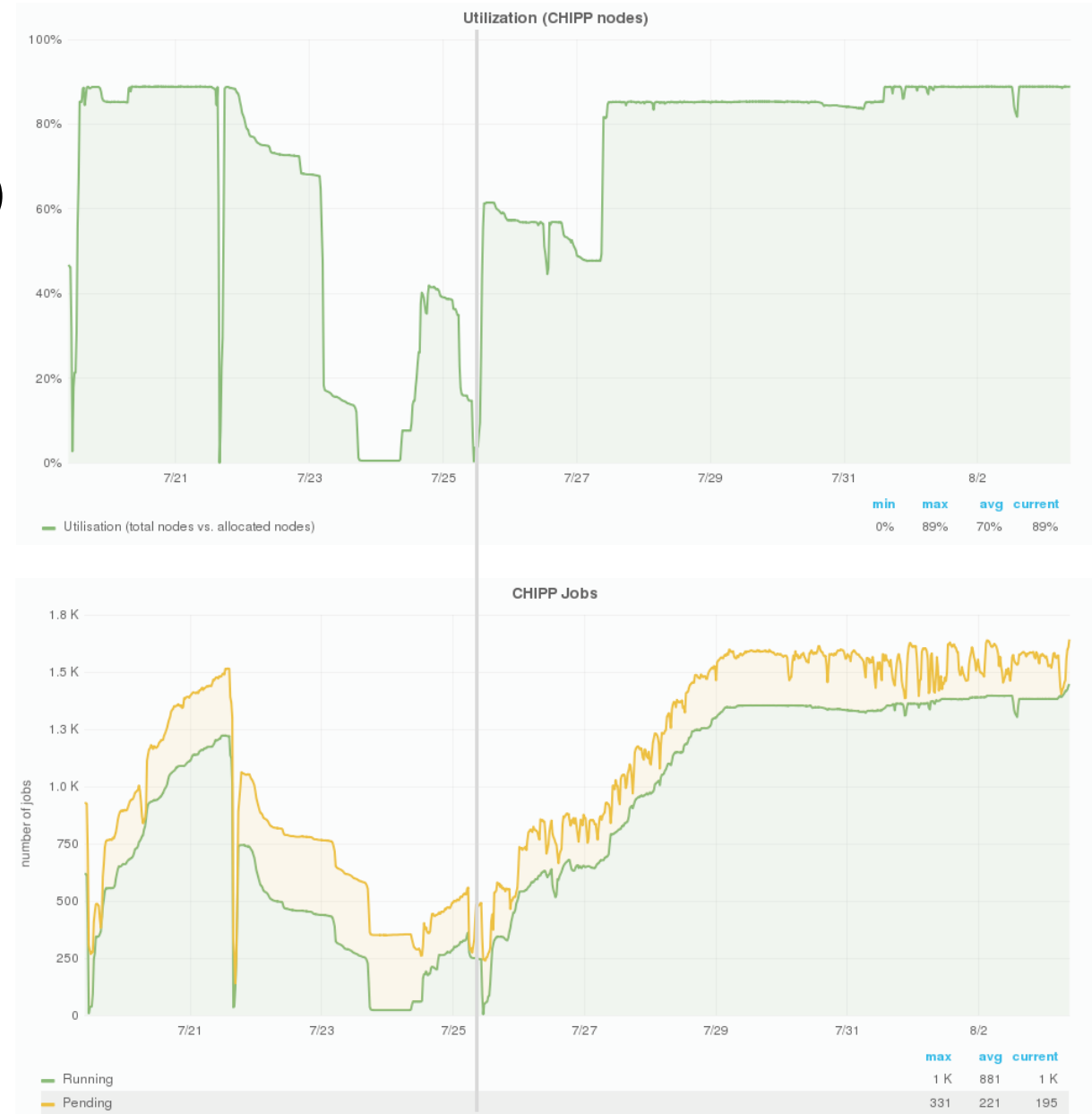
System statistics

System utilization (since Jul 19)

- Core allocation up to ~90% (1600/1800) with 64core/node (out of 72).

- Queues filled with pending jobs:

VO	Running	Pending
atlas	197	115
cms	30	3
lhcb	1170	53
ops	0	0
dteam	0	0



System utilization (since Jul 19)

- Load

- Number of procs in line with load
- Some load peaks due to IO

- CPU utilization

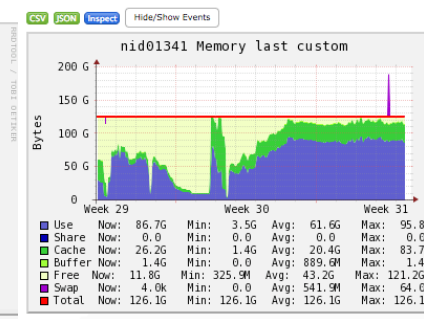
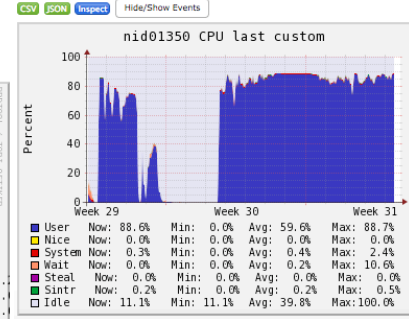
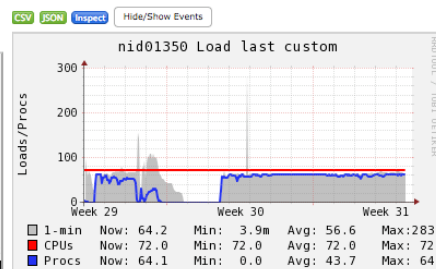
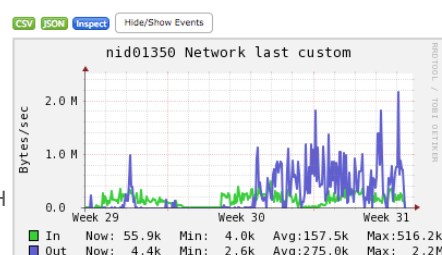
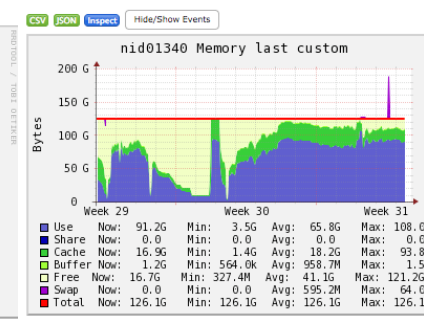
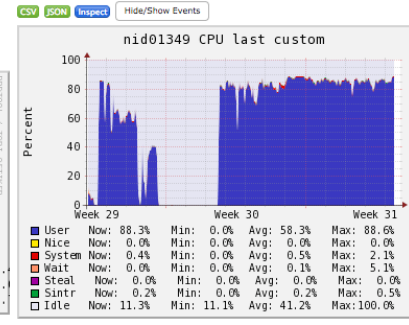
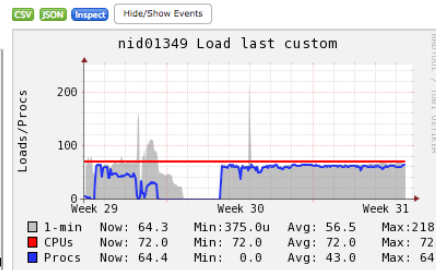
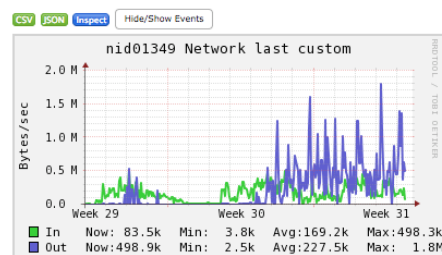
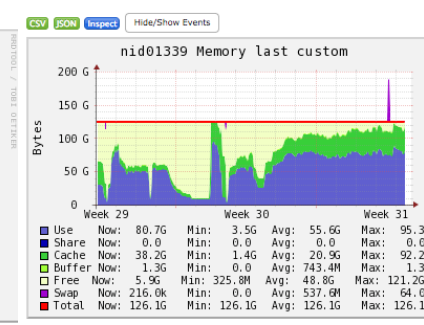
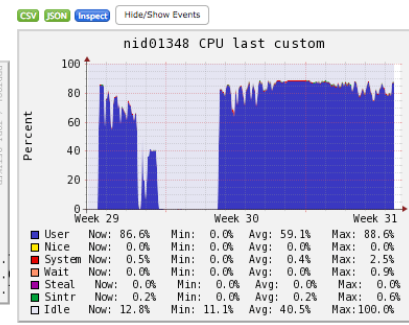
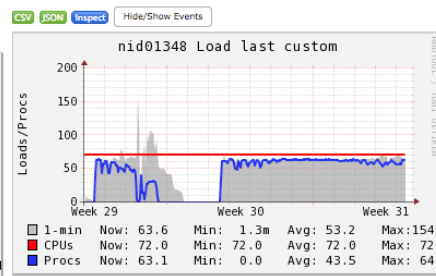
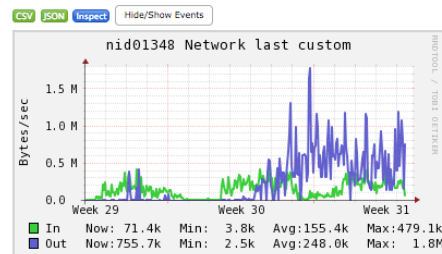
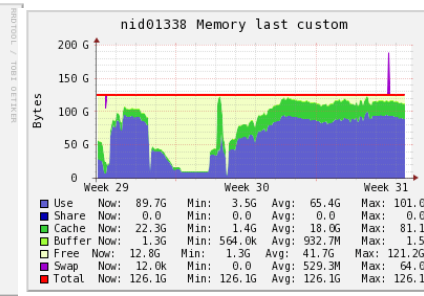
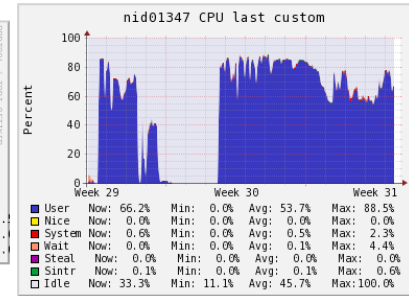
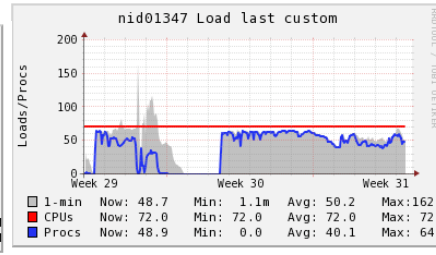
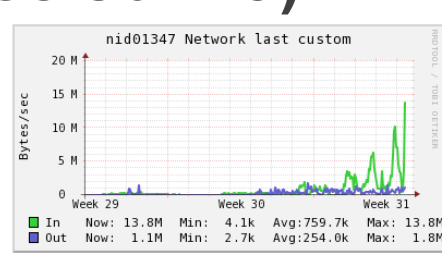
- Almost flat on ~85%
- IO wait negligible

- Memory utilization

- About 16GB in cache
- About 8-16GB free on average

- Network

- No significant activity



Next steps

Next steps

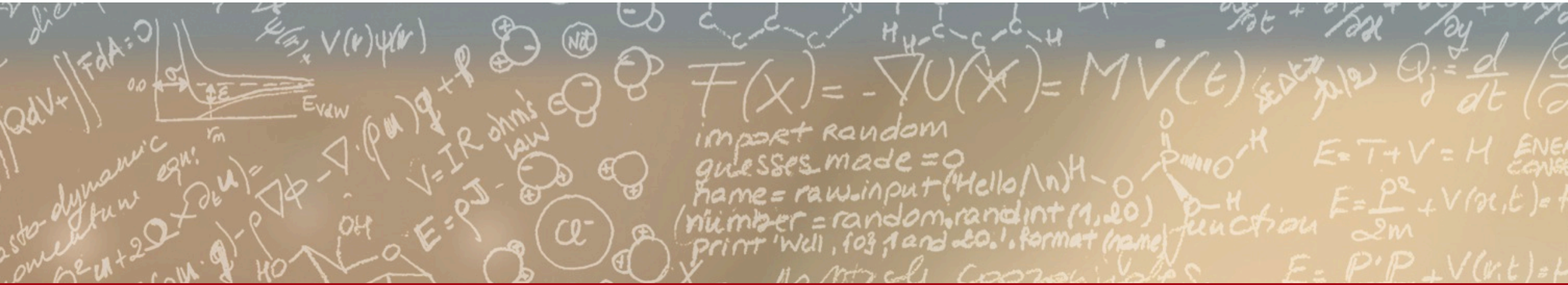
- Memory utilization suggests that there might be room to squeeze in a few cores and allocate ~66 or 68 cores instead of 64 cores
 - This could be positive (more CPU used!)
 - Or negative if nodes start swapping
 - Any comments?
- DataWarp evaluation ongoing
 - The intention is to test small (1-2 nodes) and evaluate from there
 - Any comments or ideas?
- Extended maintenance around Sep. 27 2017
 - Length to be confirmed
 - Upgrade to CLE 6.0.UP04 being tested



CSCS

Centro Svizzero di Calcolo Scientifico
Swiss National Supercomputing Centre

ETH zürich



Thank you for your attention.