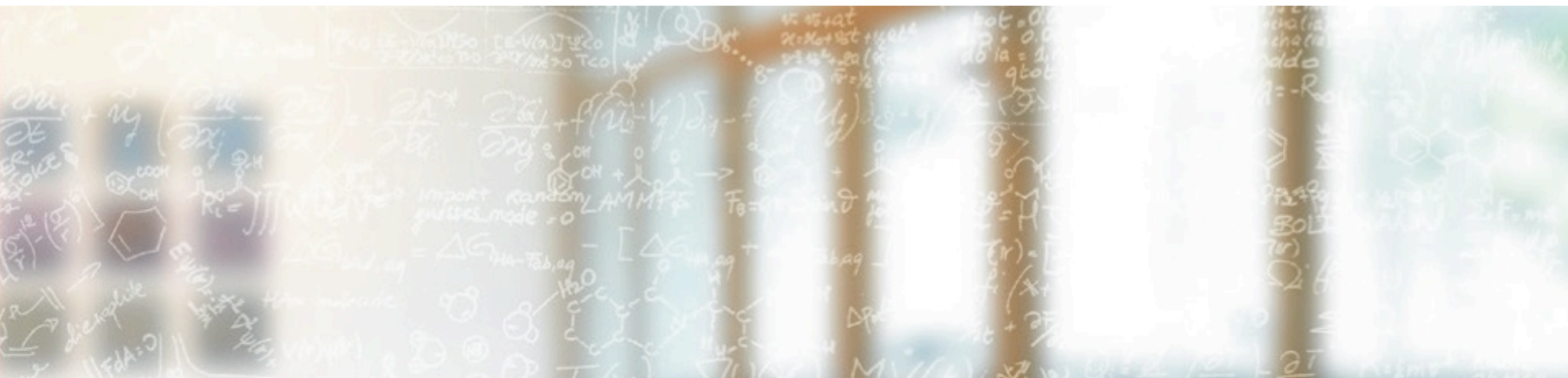




CSCS

Centro Svizzero di Calcolo Scientifico
Swiss National Supercomputing Centre

ETH zürich



CSCS-LCG2 Site report

HEPiX Spring 2017

Dino Conciatore - System Engineer - CSCS

Monday 24 April, 2017



CSCS

Centro Svizzero di Calcolo Scientifico
Swiss National Supercomputing Centre

ETH zürich

Swiss National Supercomputing Centre

Swiss National Supercomputing Centre

- Located in Lugano, Switzerland
- Unit of the Swiss Federal Institute of Technology in Zurich (ETH Zurich)
- Flagship machine: Piz Daint
 - Hybrid Cray XC50/XC40 system
 - Intel E5-2690v3 - NVIDIA Tesla P100 / Intel E5-2695v4



Lugano Lake View



ETH Zürich View

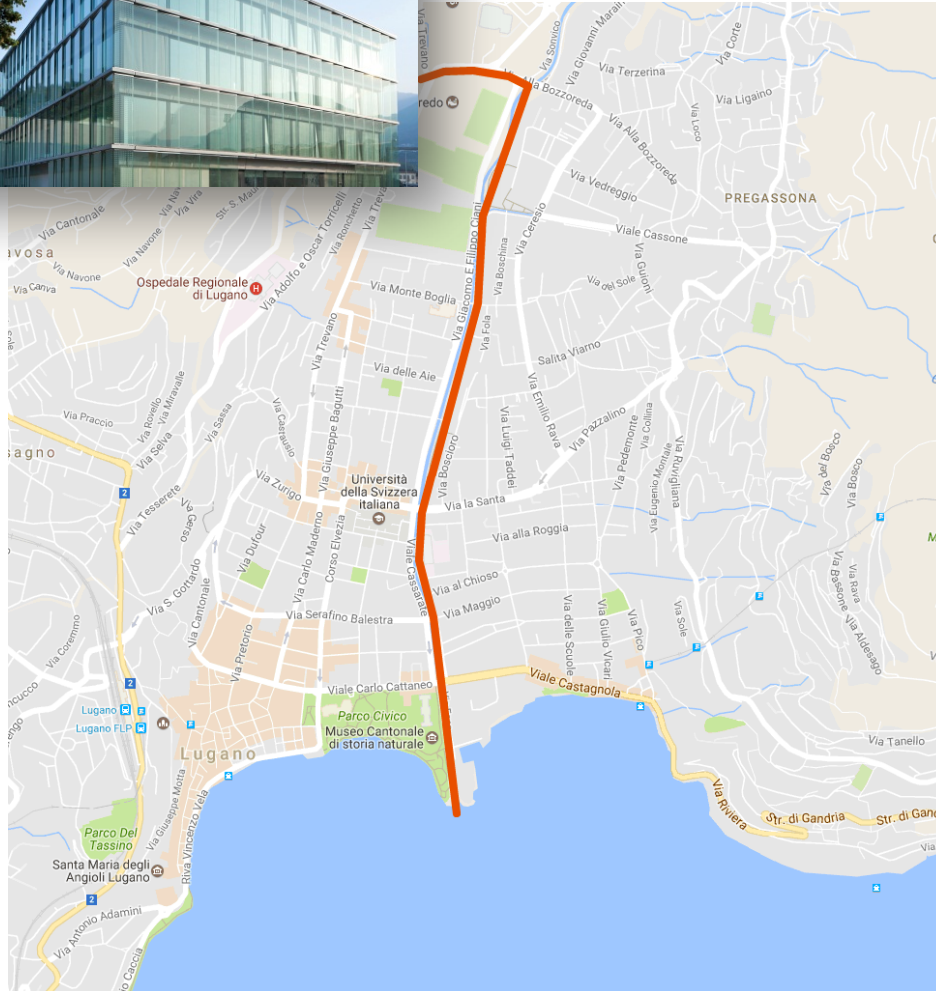
CSCS Facility - Offices and machine room



CSCS Facility – Machine room – Piz Daint



CSCS Facility – Free cooling (lake water)



- Pipeline length: 2.8 km
- High difference: 30 m
- Max. flow rate: 760 l/s
- Water taking depth: 45 m
- Water temperature: 6 °C





CSCS

Centro Svizzero di Calcolo Scientifico
Swiss National Supercomputing Centre

ETH zürich

CSCS-LCG2 Site report

Dedicated T2 Cluster

- 3 VOs
 - ATLAS – CMS - LHCb
 - Shared resources
- Compute
 - 6240 Job slots (SC and MC)
 - 3 CPU generations
 - Total ~ 70 kHS06
 - ARC middleware
 - Fully puppetized

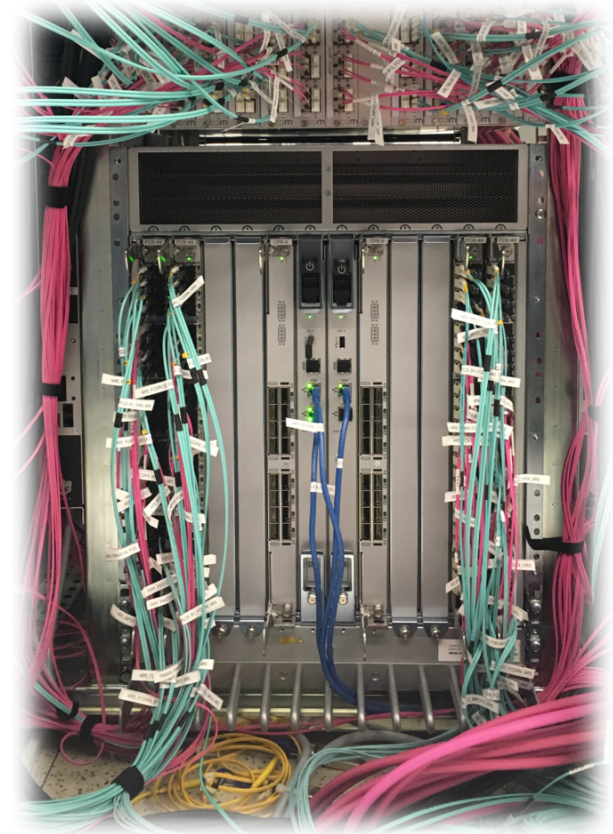
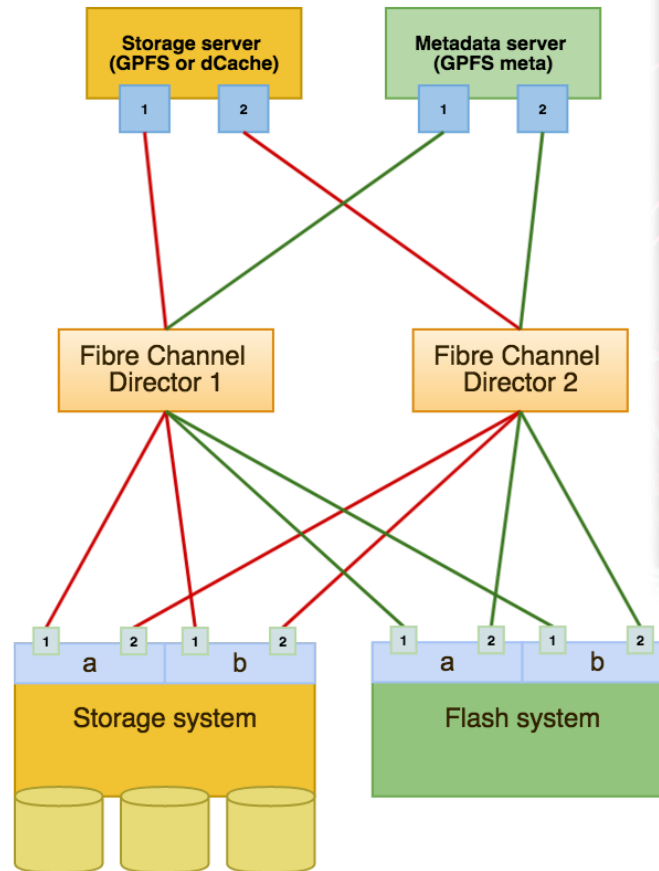


■ Network

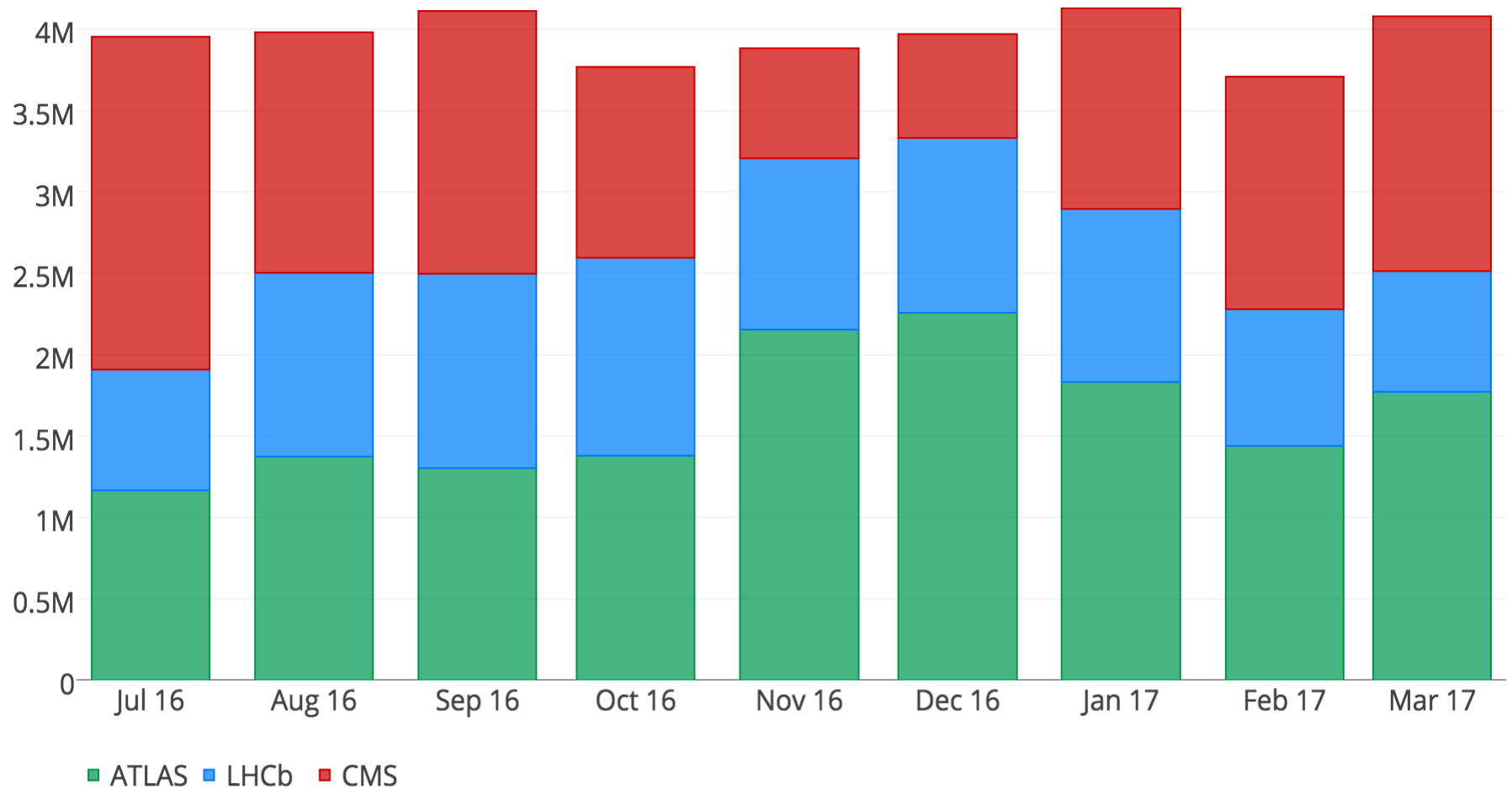
- Infiniband (EDR / FDR)
- Ethernet
- Transparent IB to ETH Bridge @ 80Gbps
- 100Gbps Internet (shared)

Dedicated T2 Cluster – SAN Storage

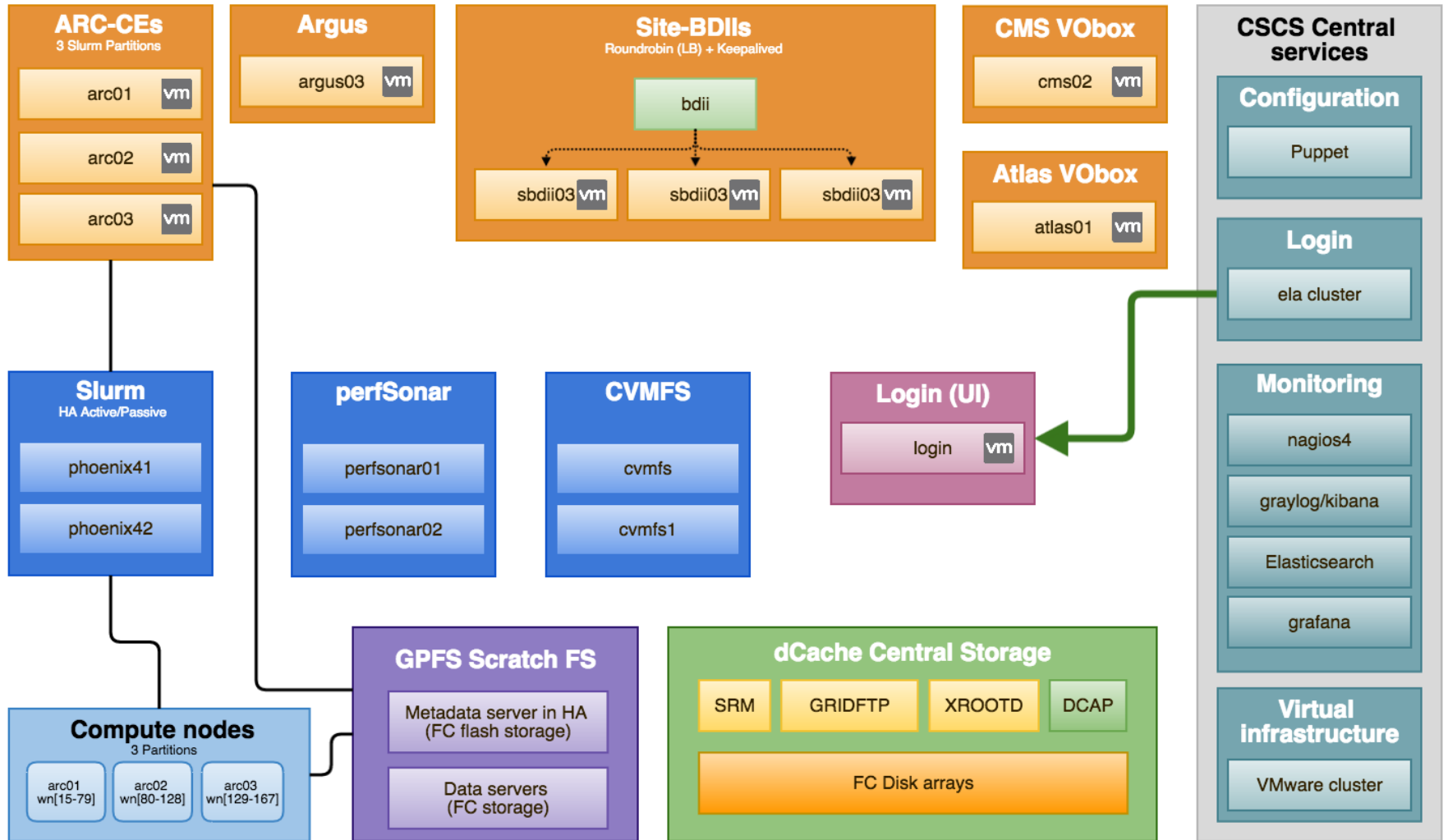
- dCache (xrootd, dcap, gftp, srm)
 - 4.2 PB in total
- /scratch – Spectrum Scale (formerly GPFS)
 - DATA
 - 641 TB available
 - META
 - Flash Systems
 - 1TB available
 - 340M inodes



Cluster usage (walltime in hours)



Services overview



Cluster Status and Job Monitoring

- Python script collecting Slurm metrics (pyslurm lib)
- Slurm elasticsearch jobcomp plugin





CSCS

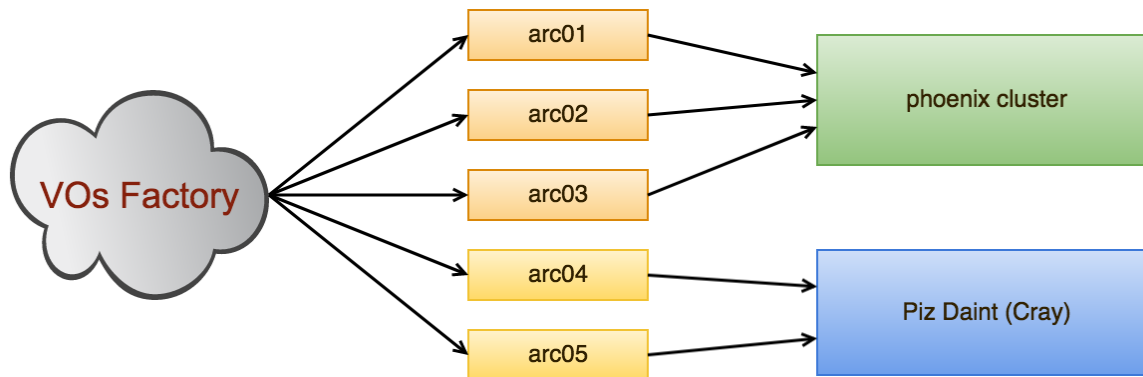
Centro Svizzero di Calcolo Scientifico
Swiss National Supercomputing Centre

ETH zürich

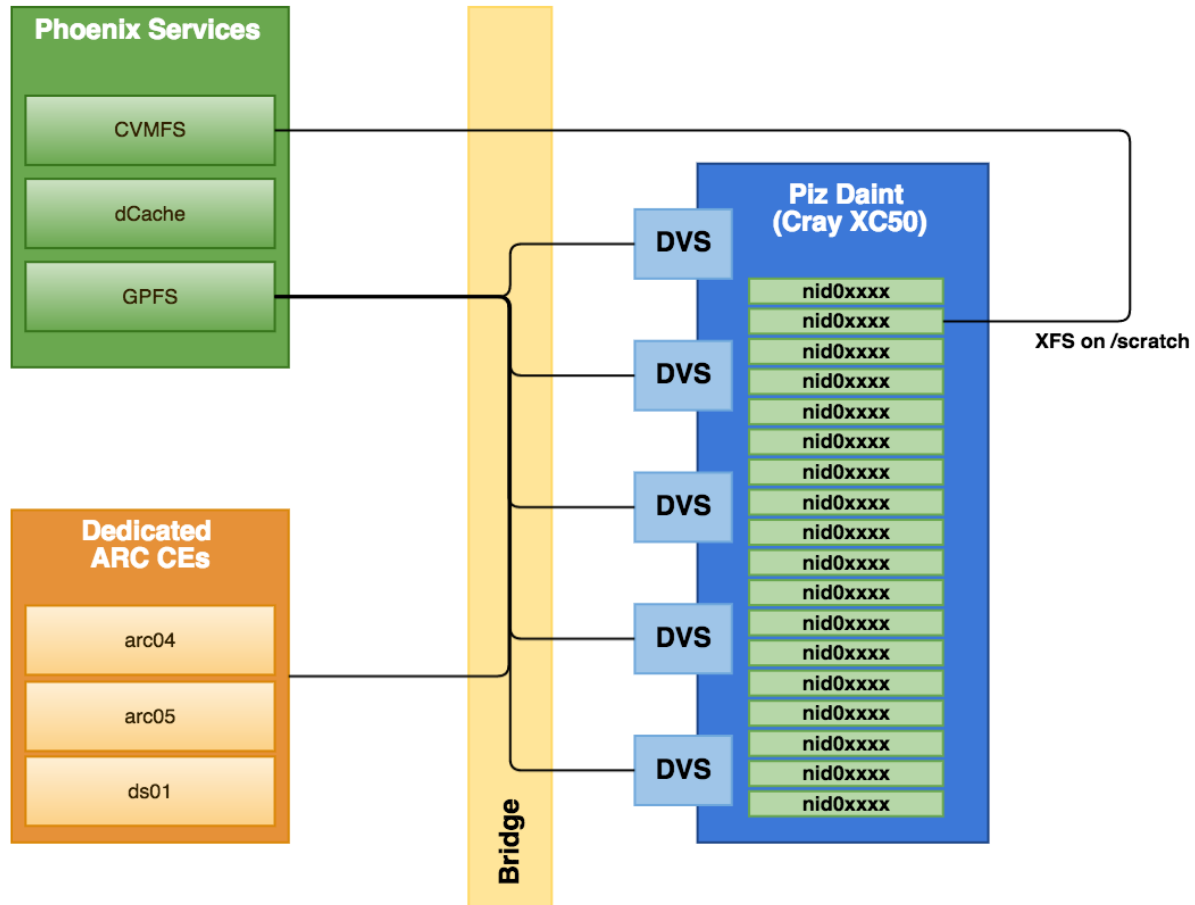
LHConCray project

LHConCray project

- Consolidation project to run LHC jobs on Piz Daint
 - Started one year ago
 - Recently started production
 - The goal is to run production jobs without changes to the workflow
- Normal workflow:
 - Jobs submitted via ARC
 - Running in containers (Shifter)
 - CVMFS Native on cray nodes



LHConCray project





CSCS

Centro Svizzero di Calcolo Scientifico
Swiss National Supercomputing Centre

ETH zürich

Any Other Business

Challenges

- Massive file creation (1M per minute, 250M in total)
 - GPFS with Flash metadata and small files in metadata (feature)
 - Quota per user/VO

- 3 VOs
 - SLURM fairshare, 10k jobs in queue, cluster still full if one VO is not running for some reason
 - Dealing with changes affecting all VOs
 - dCache configuration complexity
 - VOs are impacting each other in memory usage (swap)

Links

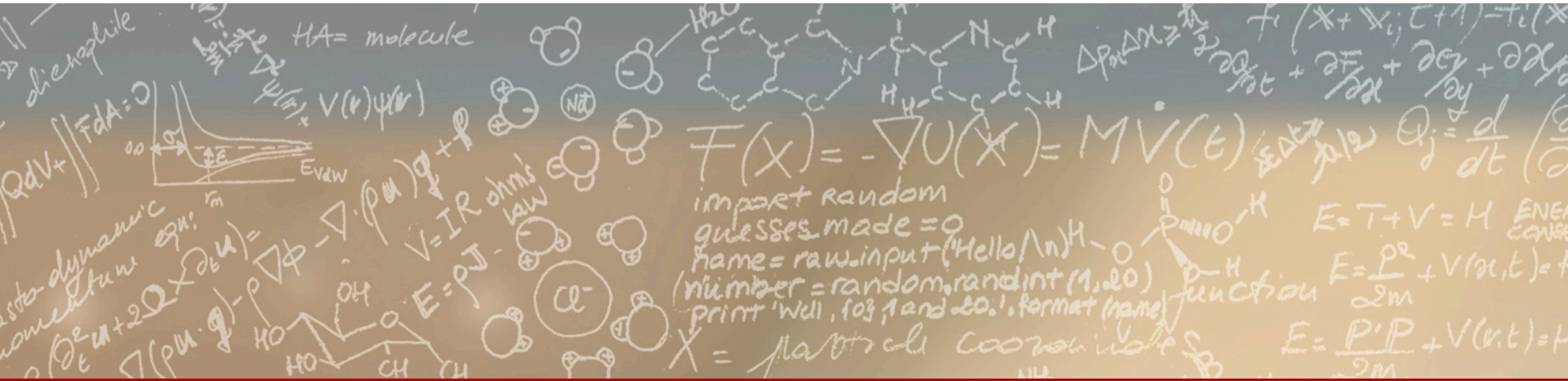
- CSCS
 - <http://www.cscs.ch/>
- Pyslurm
 - <http://www.gingergeeks.co.uk/pyslurm/index.html>
 - <https://github.com/Pyslurm/pyslurm/>
- Jobcomp (SchedMD)
 - <https://github.com/SchedMD/slurm/tree/master/src/plugins/jobcomp/elasticsearch>
- Shifter
 - <https://github.com/NERSC/shifter>



CSCS

Centro Svizzero di Calcolo Scientifico
Swiss National Supercomputing Centre

ETH zürich



Thank you for your attention.