

ATLAS job submission system for Salomon HPC based on ARC-CE

M. Svatoš, J. Chudoba, P. Vokáč

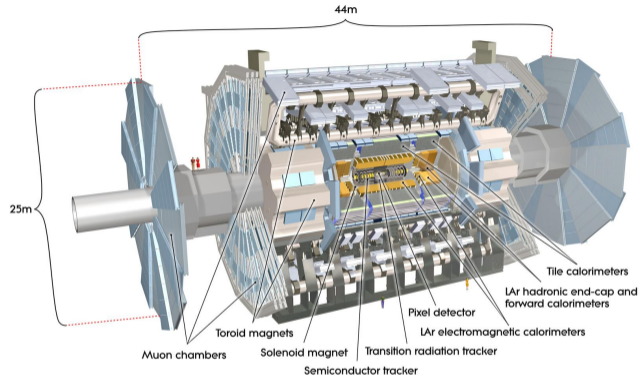
On behalf of the ATLAS Collaboration

HPCSE 2019

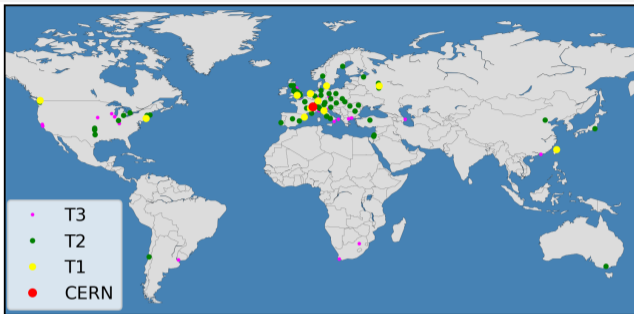
20-23.5.2019

The ATLAS Experiment

- located at the Large Hadron Collider (LHC) at CERN near Geneva
- the detector is cylindric, 44m long, 25m in diameter, weighting 7,000 tonnes
- the collaboration comprises about 3000 scientific authors from 183 institutions, representing 38 countries
- its physics programme consists of analyzing products of elementary particles collisions
 - one of its greatest achievements was discovery of Higgs boson for which the Nobel price was awarded in 2013

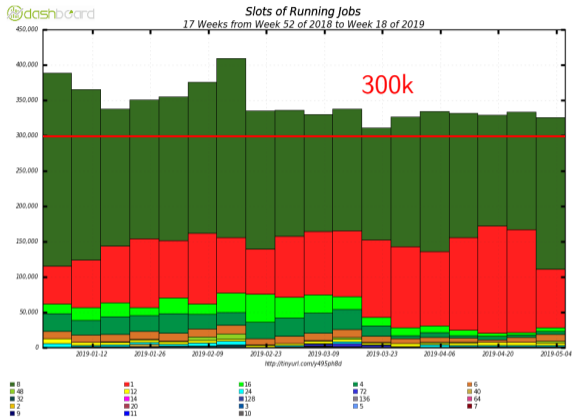


The ATLAS Distributed Computing (ADC)



- manages more than 400 PB of data
- manages more than 700 storage endpoints (on more than 150 sites located around the world).

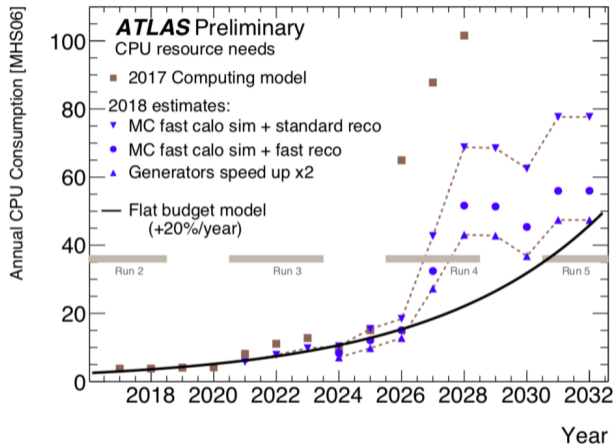
The ATLAS Distributed Computing (ADC)



- uses heterogeneous computing resources - Worldwide LHC Computing Grid (WLCG) sites, cloud resources, HPCs, volunteer computing (BOINC) resources, etc.
 - usually runs above 300k cores
- ADC uses Salomon HPC opportunistically (only filling gaps between projects with allocation)

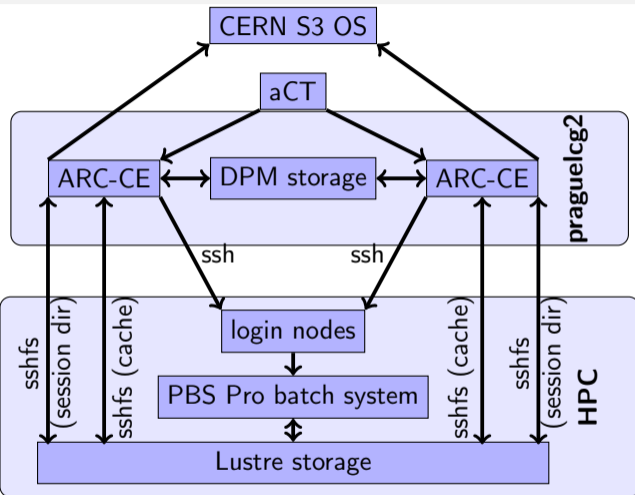


The ATLAS Distributed Computing (ADC)



- future upgrade of the LHC will cause increase in amount of data
- if the 2017 computing model is used, required computing resources would need to increase by one order of magnitude

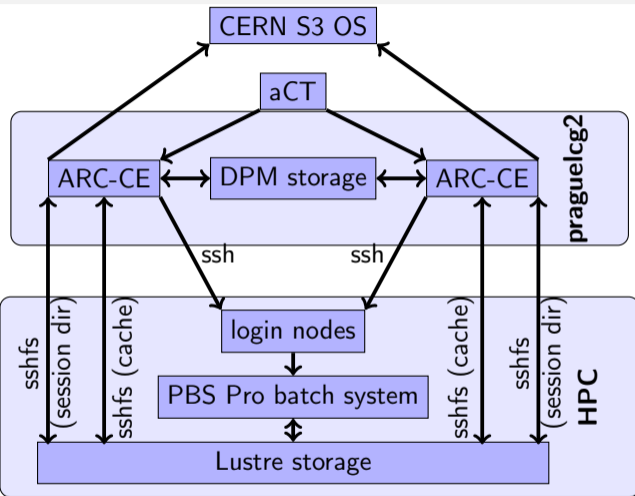
Submission system overview



Customized job submission workflow:

- the ARC Control Tower (aCT) submits job description into one of the ARC-CE machines located at the computing center of the Institute of Physics of the Czech Academy of Sciences
- the ARC-CE translates the job description into a PBS script
- the ARC-CE puts necessary scripts into the session directory which is shared with scratch on Salomon via sshfs

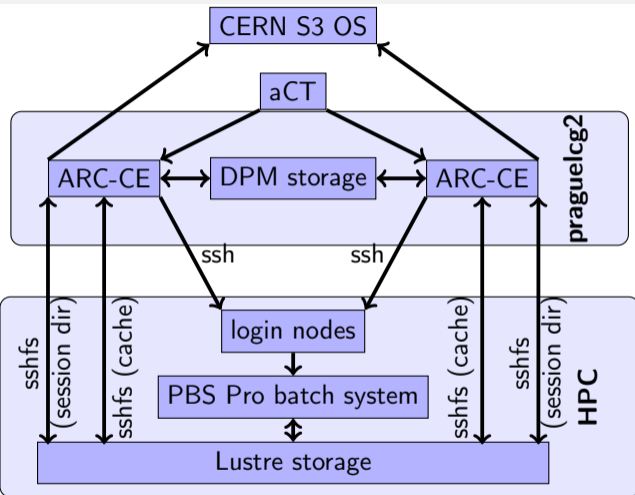
Submission system overview



Customized job submission workflow:

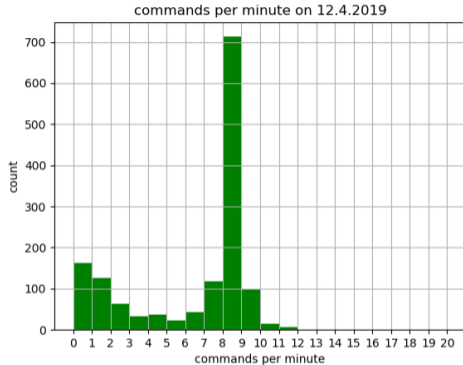
- the ARC-CE gets input files - either it links them to the session dir from a cache dir (also on the scratch of Salomon) or copies them there from local DPM storage
- the ARC-CE submits a job to the PBS via ssh connection to login node
- running job uses software stored on scratch
 - the number of jobs limit on Salomon is 100 jobs per user; each ARC-CE submits as a different user

Submission system overview



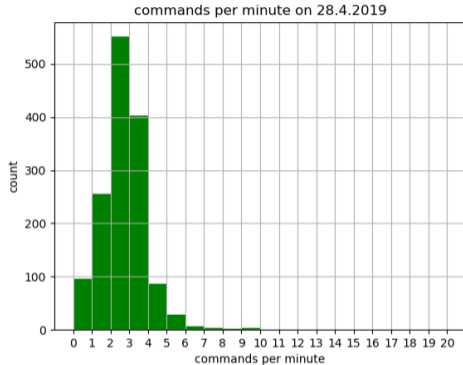
Customized job submission workflow:

- when the job finishes, the output and logs are located in the session dir and is accessed by the ARC-CE via sshfs
- the job output and log is copied to S3 Object Store in CERN in Geneva
- a copy of the log is stored in the local DPM storage



- the ARC-CE interact intensely with the PBS - apart from job submission and deletion, it performs rounds of checking via `qstat`, `pbsnodes`, and `qmgr`, because it does not have access to PBS logs
- there is a limit on number of PBS requests to protect the system and exceeding this limit puts user into a blacklist which can lead to failed submissions
- the ARC-CE with default setting usually performs 8 PBS commands per minute

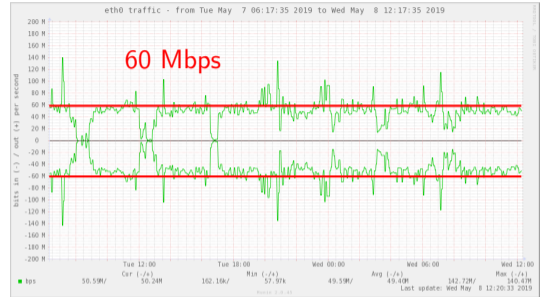
Histogram of frequency of PBS command execution - default

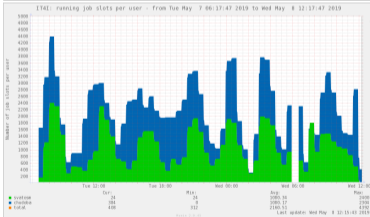
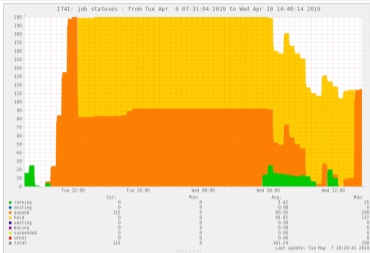


- the ARC-CE commands were modified to slow down the round of checking - after that, the number of PBS commands is usually below 5 per minute

Histogram of frequency of PBS command execution - modified

- the bottleneck in the system seems to be the sshfs connection of session directories
- the ARC-CE are run as Virtual Machines
 - the hardware has 10 Gbps networking
 - traffic reaches plateau around 60 Mbps
- ongoing testing of sshfs parameters
 - no compression - no significant speed up
 - faster encryption algorithm (aes128-ctr)
 - no significant speed up
 - caching - being tested
- the problems is probably in number of files within shared area - running job can reach more than thousand files, i.e the sshfs has to handle up to $\mathcal{O}(100k)$ files

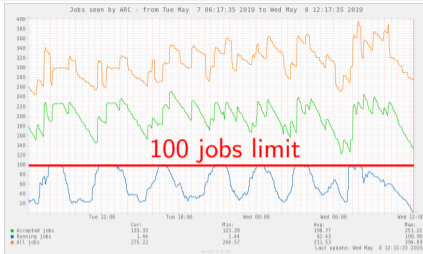




monitoring of jobs in the batch system

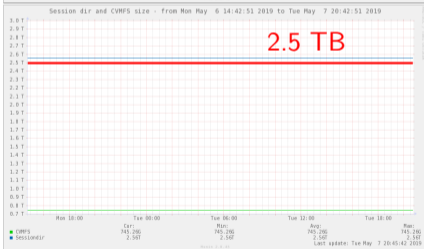
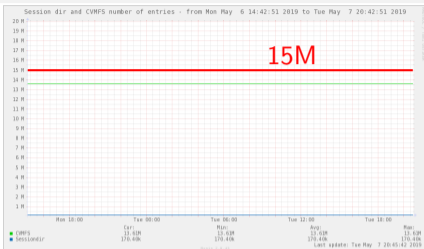
- using munin framework
- low load on PBS server - one qstat per 30 minutes
- monitoring of job states
 - allows discovery of problems in the batch system, e.g. there were jobs in held state (scheduler problem)
- monitoring of jobs per user
 - if one of the ARC-CE machines would have problem





monitoring of jobs in the ARC-CE

- using munin framework
- queries the ARC-CE every 5 minutes
- shows how long it takes to refill the machine after old jobs are finished



monitoring of used space

- using munin framework
- via `it4i-disk-usage` command executed once a day
- number of entries and disk size in `/scratch/work` and `/scratch/temp`
- the software (CVMFS) is in `/scratch/work` - about 14M entries
- session and cache dirs are in `/scratch/temp` - most is taken by the cache

cross-check with IT4I monitoring

- <https://extranet.it4i.cz/dash/salomon>
 - if no jobs are running, to see if the machine is full
- <https://extranet.it4i.cz/motd/all>
 - the Message Of The Day - to check informations about downtimes, problems



- jobs of the ATLAS experiment are successfully running on the Salomon HPC
- number of PBS requests performed by the ARC-CE was decreased to protect the HPC and to avoid job submission failures
- tuning of sshfs parameters is ongoing - changes in compression and encryption had no significant effect, caching is being investigated
- various aspects of the system are monitored and frequently checked

