Improvements in utilisation of the Czech national HPC center

M. Svatoš, J. Chudoba, P. Vokáč

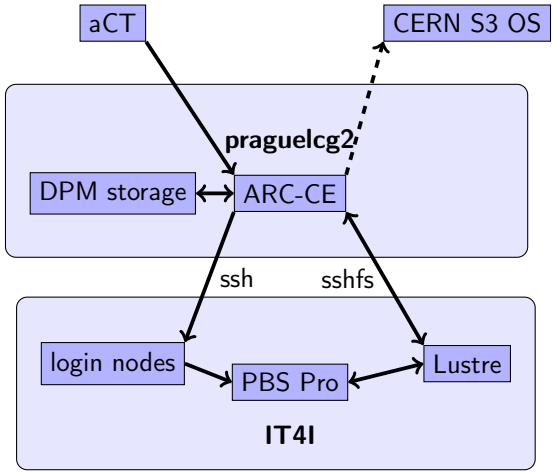
On behalf of the ATLAS Collaboration

CHEP 2019, 4-8.11.2019

Introduction

- the ATLAS experiment is opportunistically using computing resources of the Salomon HPC located at the Czech National HPC Center IT4Innovations (IT4I) in Ostrava
- Salomon HPC:
  - put into production in 2015 (ranked 39th in Top500 at June 2015)
  - currently ranked 282nd in Top500 (as of June 2019)
  - it consists of 1008 WNs
  - WNs:
    - 24 cores of Intel Xeon E5 CPUs
    - 128 GB of RAM
    - Infiniband (56 Gbps)

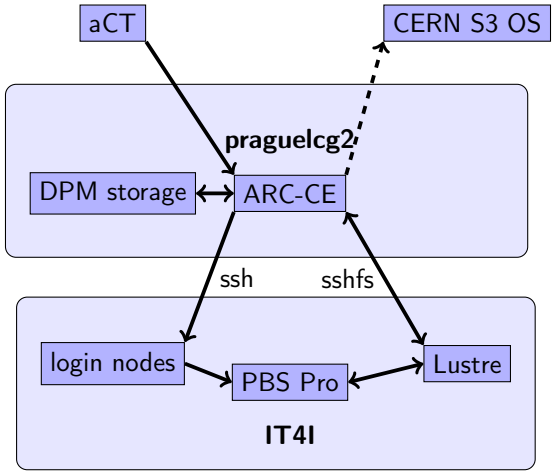
Settings



Short submission system overview:

1. the ARC Control Tower (aCT) submits job description into one of the ARC-CE machines installed at Czech Tier2 site (praguelcg2)
2. the ARC-CE translates the job description into a PBS script
3. the ARC-CE puts input files (from the local DPM storage or cache on the Lustre) necessary scripts and into the session directory (shared with Lustre storage on Salomon via sshfs)
4. the ARC-CE submits a job to the PBS via ssh connection to login node

Settings



Short submission system overview:

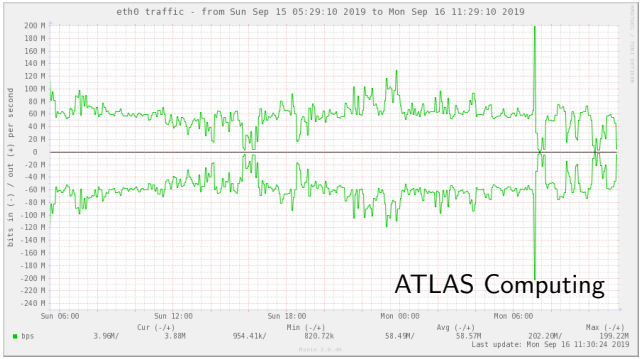
5. the output and logs of finished job are located in the session directory
6. output:
   - running job uses software stored on scratch (located on the Lustre)
   - Standard jobs: the job output and log are copied to the local DPM storage
   - Event Service jobs: the job output and log are copied to S3 Object Store in CERN (a copy of the log is stored in the local DPM storage)

details in `https://doi.org/10.1051/epjconf/201921403005`

Pre-Emption

- In October 2018, the management of the Salomon HPC decided to change conditions of opportunistic usage.
  - jobs using the Salomon HPC opportunistically could be pre-empted
  - the decision was not put in action so far
- ATLAS has a system which handles such jobs called Event Service.
  - The submission system ran in that mode for several months while receiving gradual updates and tweaks.
  - During this time, the submission system demonstrated its preparedness for pre-emption.
  - Later, the submission system was switched back to standard jobs as those are more efficient.

Sshfs



- sshfs seems to be the bottleneck of the submission system
  - the speed reaches a plateau around 60 Mbps
  - in throughput test, the machine reached 500 Mbps
- probable reason is huge number of small files in shared area (1k+ files per one Event Service job)

Sshfs

Several parameters of the sshfs were tested in attempt to improve throughput:

- no compression
- faster encryption (aes128-ctr)
- caching (tested parameters: kernel_cache, noauto_cache, cache_timeout=300, cache=no, no_readahead)

Testing:

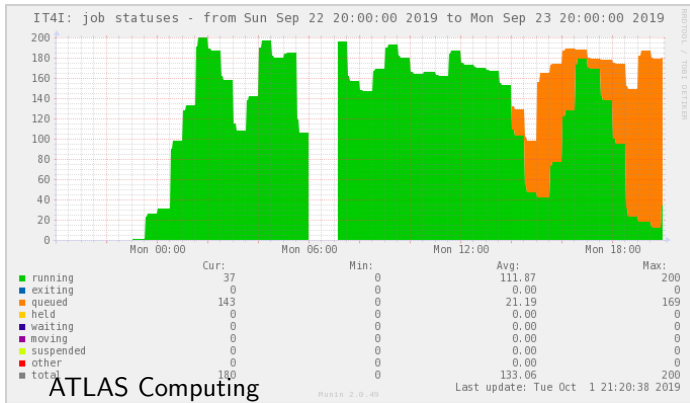
- two identical ARC-CE machines (jobs spread moreless evenly)
- change one sshfs parameter on one of them

Results:

- no visible effect on throughput plateau
- cache=no and no_readahead decreased volume of data being transferred

Sshfs

- 4 ARC-CE machines are submitting jobs to the Salomon HPC
  - each machine has limit of 50 jobs (allowed limit is 100 single-node jobs in the batch system per user with two user accounts are used to submit jobs)
- each machine has 2 sshfs connections:
  - session directory
  - cache of input files

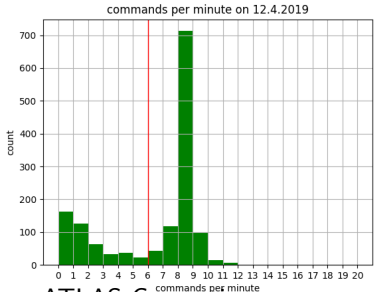


ATLAS Computing

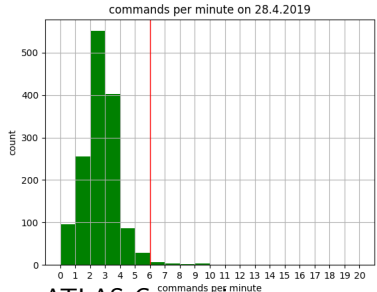
- with this setup, we are able to fill the HPC relatively quickly and keep it full

Number of PBS requests

before modification

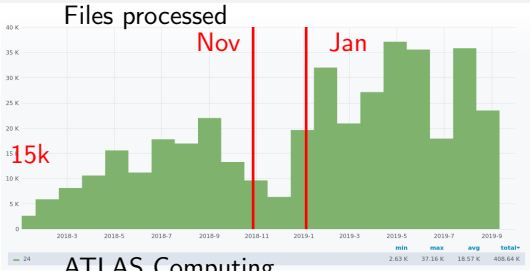


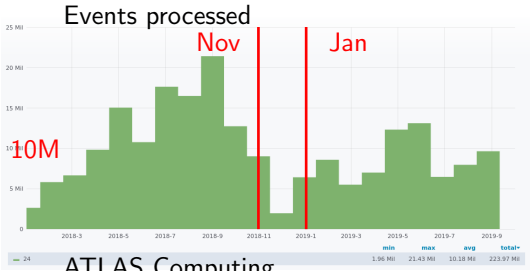
after modification



- the HPC has a limit on number of PBS requests, exceeding this limit can cause job submission to fail
- the ARC-CE in default setting exceeds the limit
  - the frequency cannot be tuned in ARC 5 (fixed in ARC 6.1)
  - the ARC-CE commands (`qstat`, `pbsnodes`, and `qmgr`) were modified to slow down the interaction with the HPC
- failing submission became rare after the modification

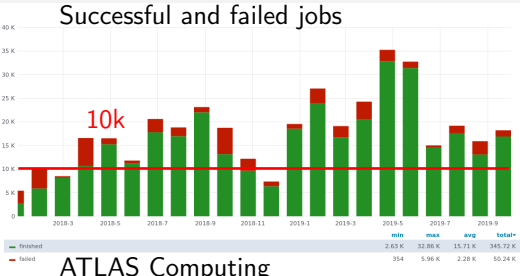
Performance





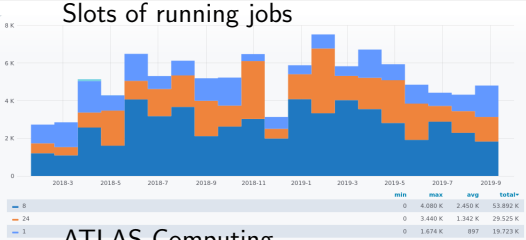
- In January 2019, submission limit increased from 100 (i.e. 2.4k cores) jobs to 200 (i.e. 4.8k cores)
  - number of processed events decreased because the Event Service jobs (running from end of November 2018 to beginning of September 2019) process less events per job

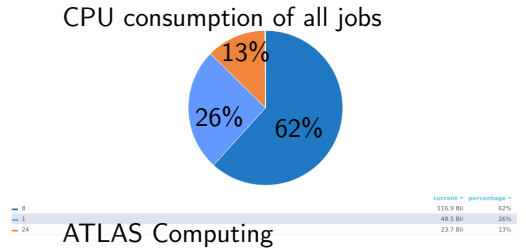
Performance



- number of failed jobs is reasonable (13%)
- error distribution is similar to normal grid queue (software errors, storage problems, problems of central services, etc.)

Performance





- praguelcg2 site statistics:
  The Salomon HPC provides significant amount of unpledged resources to ATLAS (about 13% of CPU time of the site).

Summary and Conclusions

- improvements:
  - pre-emption readiness
  - tuning of sshfs
  - addition of more ARC-CE machines
  - tuning of PBS command frequency
- with those improvements, the submission system
  - can fill provided opportunistic resources relatively quickly and keep them full
  - has reasonably low failure rate
- the Salomon HPC provides significant amount of unpledged resources to ATLAS

Acknowledgement

Computing resources at FZU are co-financed by projects Research infrastructure CERN (CERN-CZ) and OP RDE CERN Computing (CZ.02.1.01/0.0/0.0/16013/0001404) from EU funds and MŠMT. This presentation is supported by project LTT17018.
We gratefully acknowledge National Supercomputing Center IT4Innovations (IT4I) for computing time provided by the Salomon supercomputer.

EUROPEAN UNION
European Structural and Investment Funds
Operational Programme Research,
Development and Education

MINISTRY OF EDUCATION,
YOUTH AND SPORTS