

ARC-CE+HYPERQUEUE BASED SUBMISSION SYSTEM OF ATLAS



JOBS FOR KAROLINA HPC

M. Svatoš, J. Chudoba, P. Vokáč



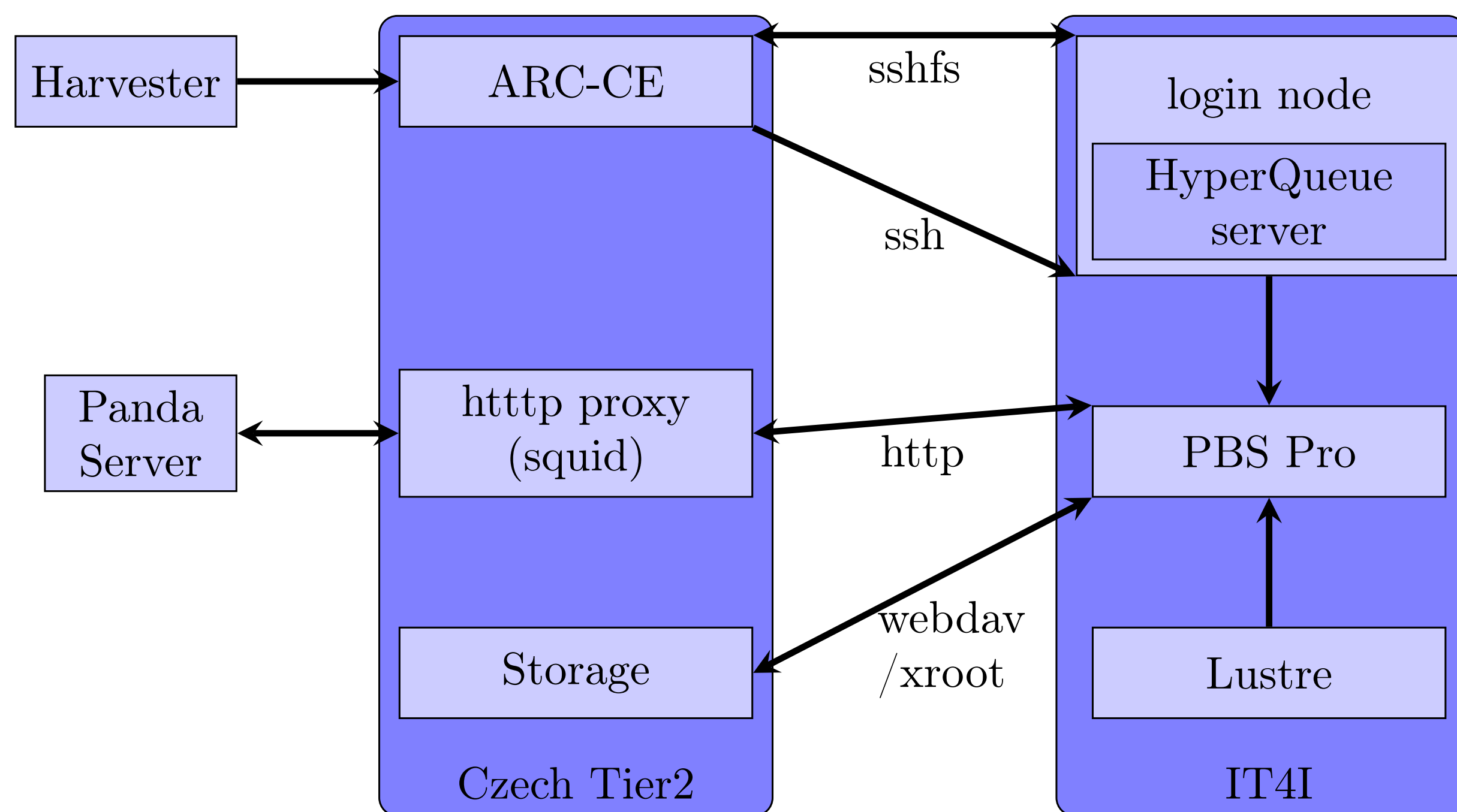
6th Users Conference of IT4Innovations

3.-4.11.2022

Introduction

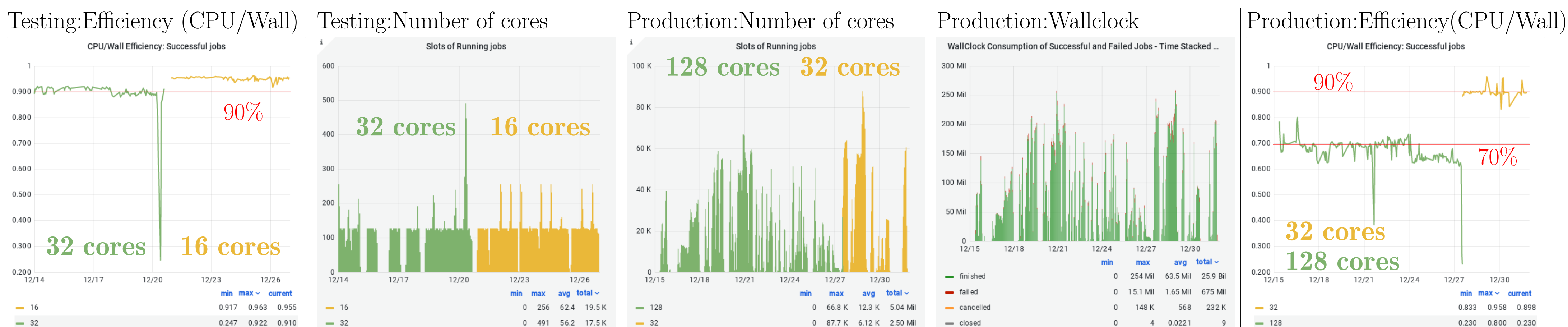
For several years, the distributed computing of the ATLAS experiment at the LHC (ADC) has been granted opportunistic use of computing resources of the Czech national HPC centre, IT4Innovations. Within the ADC, resources of IT4Innovations are attached to the Czech Tier2. Implementation of HyperQueue (a tool designed to simplify execution of large workflows on HPC clusters, developed by IT4Innovations) into the previously used ARC-CE based submission system allows tweaking of submitted jobs and hopefully helps with handling of pre-emption at IT4Innovations.

Submission system



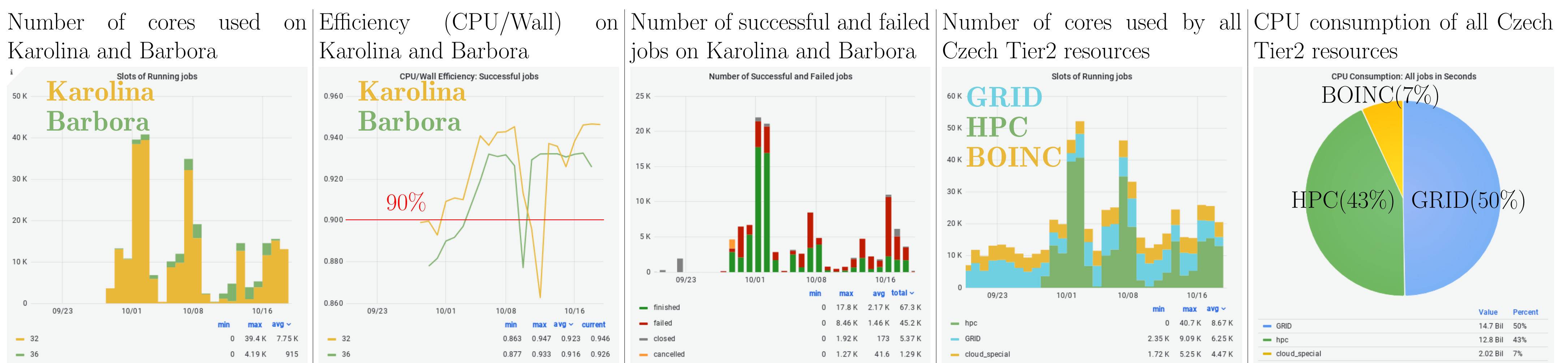
- the ARC-CE receives a pilot job, translates the job description into a script that can be run in the batch system, puts necessary files into a folder shared with the HPC via sshfs, and submits the job via ssh connection to the HyperQueue server running on a login node
- the HyperQueue server buffers the jobs and when there are enough of them, it submits jobs into the PBS Pro
- when the PBS job starts, HyperQueue jobs start in it (in sufficient numbers to fill the worker node - if available)
- in each HyperQueue job, pilot starts
- pilot contacts panda server through http proxy (Czech Tier2 squid) to receive payload job (as http is one of few open ports)
- when it receives a payload job, it gets input file from the Czech Tier2 storage via xroot or webdav (in a rucio container)
- then it starts the calculation (in software container)
- when the payload job finishes, it sends outputs to the Czech Tier2 storage via xroot or webdav (in a rucio container)
- when this is finished, pilot will request another payload job (if it can expect that the batch queue setting would allow it to finish)

2021: Migration to HyperQueue



• In December 2021, the HyperQueue implementation was tested on Karolina using 16 and 32 cores and then put into production a few days before the year-end

2022: Living with pre-emption



- In 2022, new project which can use only pre-emptive queues started in September.
- Development to tune the submission system continues.
- Contribution of Barbora, which is not using the HyperQueue yet, is not significant in comparison with Karolina.
- Even with pre-emption, the IT4I provides significant resources to Czech Tier2 with high efficiency.

Acknowledgement

This work and computing resources at FZU were co-financed by projects CERN-C (CZ.02.1.01/0.0/0.0/16013/0001404) and CERN-CD (CZ.02.1.01/0.0/0.0/18_046/0016013) from EU funds and MŠMT and projects CERN-CZ (LM2018104) and LTT17018.