FZU CC Site Report

Jiří Chudoba
with help from: L. Míča, A. Mikula, M. Svatoš, P. Šesták, J. Uhlířová, P. Vokáč (all FZU)
J. Dvořáček for CEICO (FZU)
D. Adamová (NPI Řež)

FZU Fyzikální ústav Akademie věd České republiky

Hepix Fall 2023, 17.10.2023

FZU Computing Center

- Distributed resources with all central services and most of hardware at FZU
- **Several clusters under one HTCondor LRMS**: 10000 (-9%) jobslots (5000 physical cores)
  - 1 cluster (from 2016) switched off in June 2023 (15 kHS06, 1440 jobslots) and moved to MFF
  - 138 kHS06 (-7 %), avg: 13.86 (+3 %) HS06/core
  - 4 types of hardware, oldest from 2017 (Intel Xeon E5-2650 v4)
  - 1240 jobslots (-21 %) at MFF UK, 16 kHS06 (-16 %)
    - very old servers from 2012 switched off

Local Resources Fully Used (?)

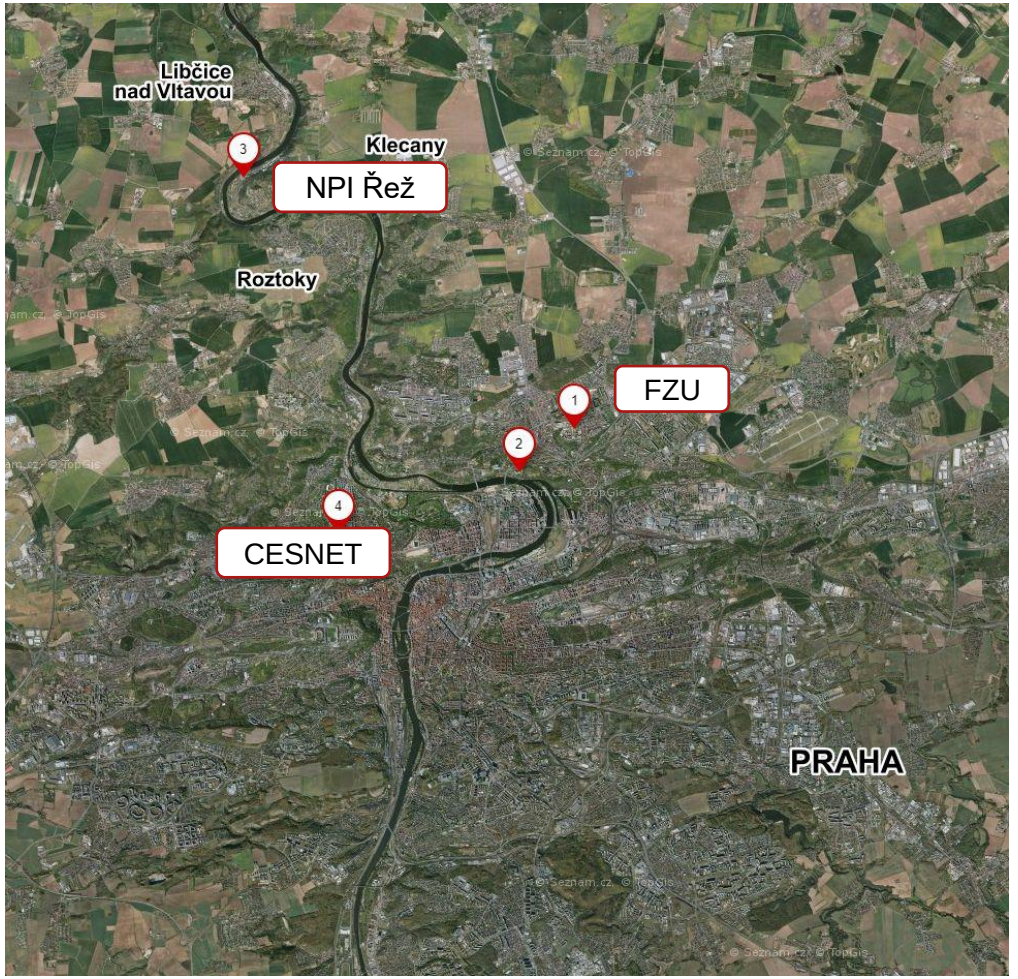
Name	Last *	Min	Mean	Max
Total Cores	10000	1320	10960	13772
Total w/t Offline	9936	1288	10850	13704
Total w/t Offline and Draining	9936	1288	10850	13704
ATLAS multicore	6720	216	4683	10832
ALICE	240	0	2826	8016
Auger User	1634	0	1564	6433

only subset of VOs supported on remote nodes

- NFS overload
- ALICE storage

Storage

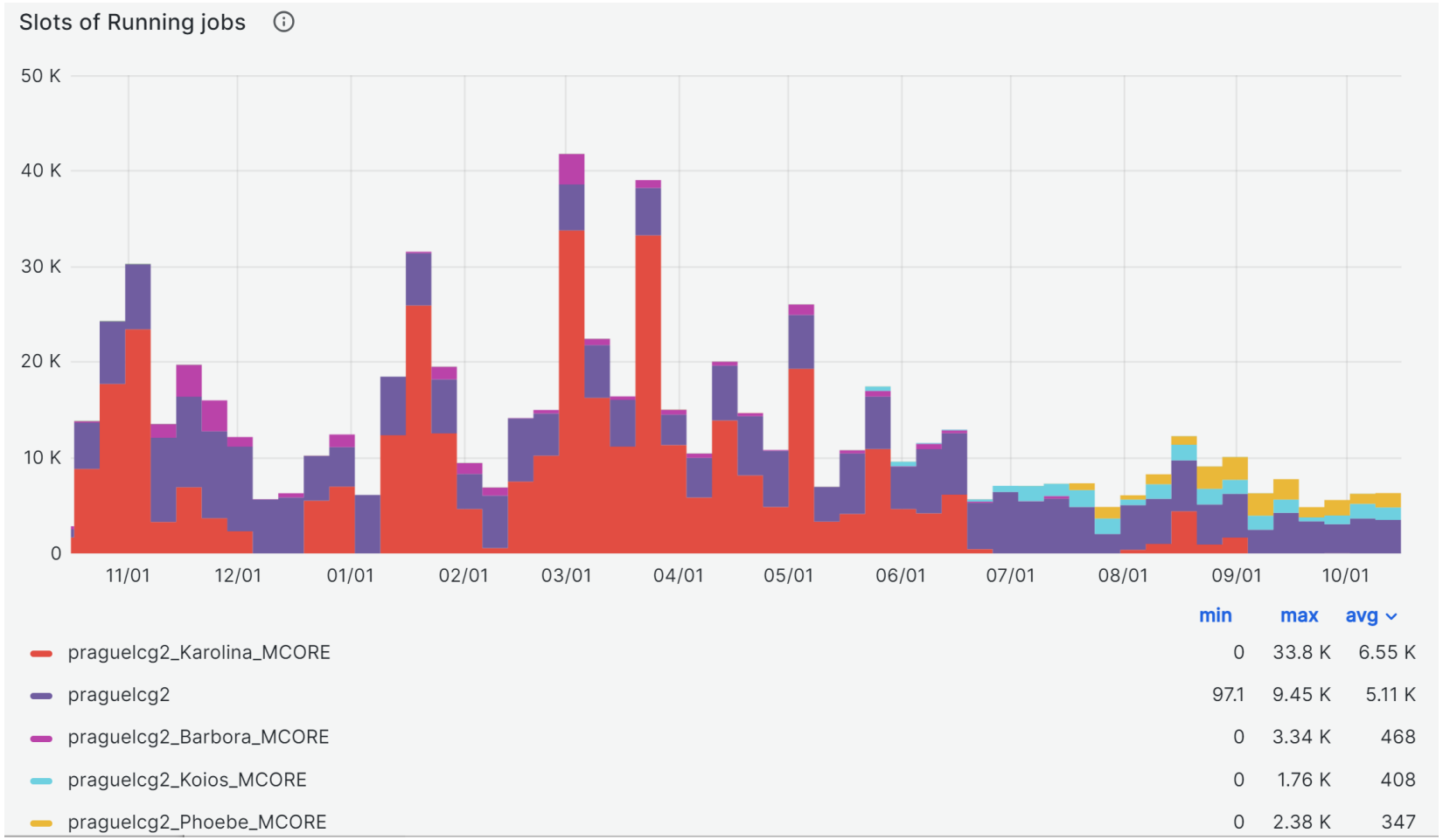
- **7 PB dCache** (-10%)
- **3.14 PB xrootd** (only ALICE), 4 servers at NPI Řež
  - Cooling upgrade during summer 2023, ALICE storage not available for several weeks
  - Recent failure at CESNET site (power cut) caused unavailability of the storage for ALICE
  - 100 GBps connection to NPI in preparation



More CPUs – IT4I

- National Supercomputing Center IT4Innovation in Ostrava
  - **200 000** node hours in the project OPEN-27-57
    - **153 % used** (100 % with a full priority, 50 % reduced, the rest in preemption queue)
  - new project since September 2023: OPEN-29-6
    - multiyear
    - **2x (150 000 + 50 000)** node hours (Karolina + Barbora, 2023 + 2024)
    - test access to **LUMI**

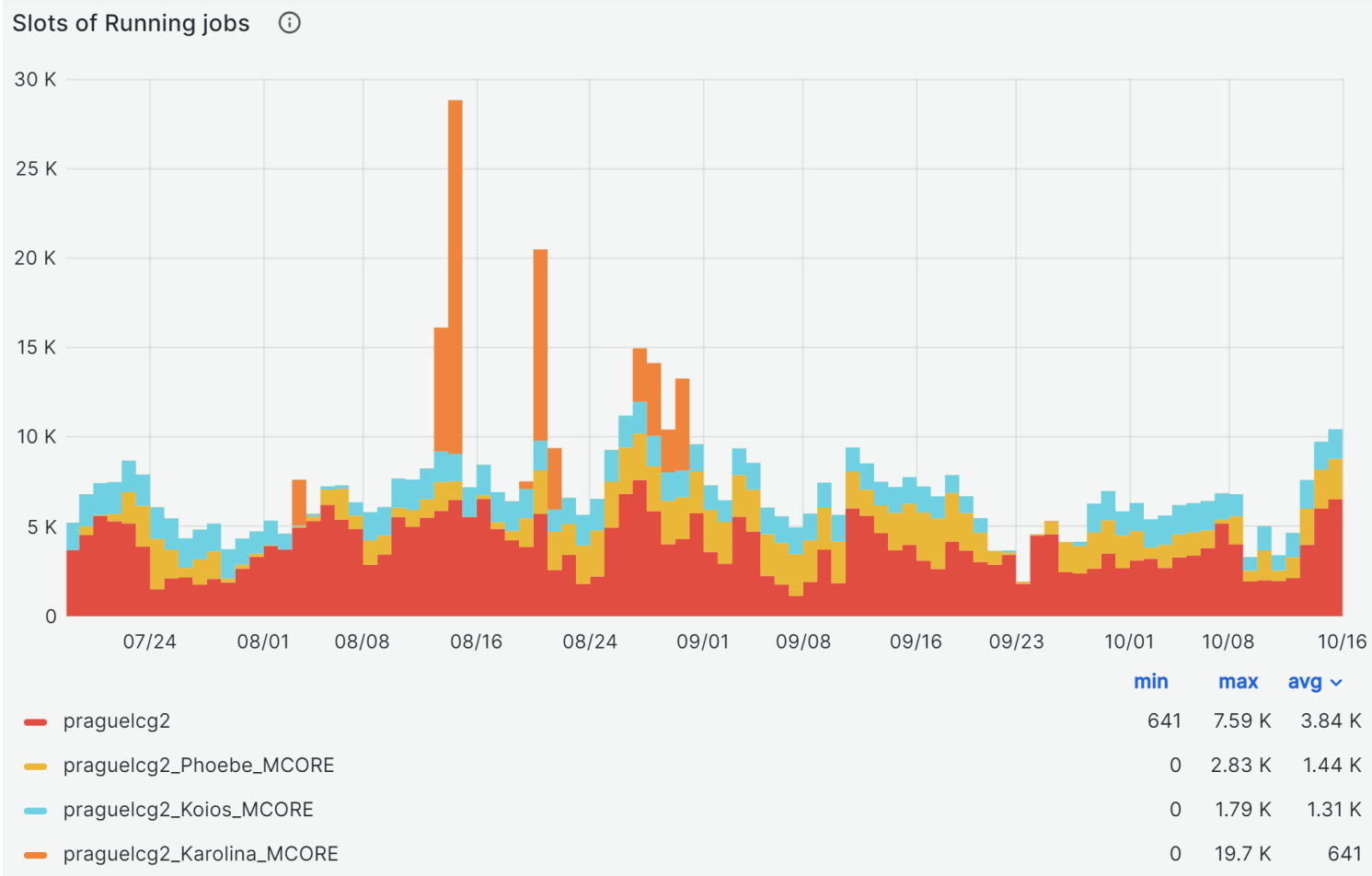
More CPUs – IT4I



More CPUs

- 2 local clusters using Slurm dedicated to cosmology group
  - legacy - “Koios” – 2017, 30 servers á 2 x Skylake CPU (Xeon Gold 6130), 0.3TB RAM, 100Gbit IB, 1 GBE ethernet
  - new - “Phoebe” – 2022, 22 servers á 2 x Epyc 3rd gen (EPYC 7543), 0.5TB RAM, 100Gbit IB, 10 GBE ethernet
- Now used also by ATLAS via dedicated CEs
  - preemtion: SUSPEND - suspending preempted job into node memory using SIGSTOP standard linux signal

More CPUs


Slots of Running jobs
30 K
25 K
20 K
15 K
10 K
5 K
0
07/24
08/01
08/08
08/16
08/24
09/01
09/08
09/16
09/23
10/01
10/08
10/16
min
max
avg
praguelcg2 641 7.59 K 3.84 K
praguelcg2_Phoebe_MCORE 0 2.83 K 1.44 K
praguelcg2_Koios_MCORE 0 1.79 K 1.31 K
praguelcg2_Karolina_MCORE 0 19.7 K 641

HTCondor updates

slide provided by Petr Vokáč

- New version of HTCondor-CE no longer supports GSI
- Recent versions allows SSL connection with client certificate
  - identity mapping doesn't support VOMS roles during authentication
  - experiments with one (few) grid users (pilot factories) can be mapped by certificate subject
- SSL client and server configuration described in GDB talk
  - side server easy to configure (still support VOMS classAds, e.g. for job routing)
  - require configuration changes on client side (pilot factories)
    - **unable to move without coordination with experiments that still rely on GSI**
    - most recent HTCondor client necessary and configuration changes to support SSL
    - prevent us to move our infrastructure to supported HTCondor 23
- FZU HTCondor-CE 6.0 (HTCondor 10.9.0) testbed
  - works fine with SSL
  - took 1/2 hour to configure and test

Grid storage

slide provided by Petr Vokáč

Migrated from DPM to dCache in May 2022 (details)
- smooth operation, we should have moved years ago…

Production storage already upgraded to the new golden dCache 9.2
- incompatibilities section for **all** branches (8.2 → 9.0 → 9.1 → 9.2)
  - `cleaner` service replaced by `cleaner-disk` (and `cleaner-hsm`) in 9.1
- communication issues with mixed 8.2 and 9.2 dCache services
  - upgraded all services at same time to 9.2

Access with OIDC tokens (WLCG JWT flavour)
- available already in 8.2 (support in 7.2 not sufficient for WLCG)
- tricky to configure correctly
  - issuer vs. group vs. scope based authorization (capability access preferred by WLCG)
- tokens configured for ATLAS experiment on our production storage
  - transfers with tokens @ WLCG Data Challenges 2024

Summary

- Small decrease of dedicated local CPUs over-compensated by non dedicated resources
- OS: WN on Centos7, services mostly on AL9
  - planned update of AliceBox for next week
- New project for WLCG dedicated hardware submitted
  - for 2024 – 2026
  - 400 Gbps router

We greatly appreciate network, computing and storage resources provided by e-INFRA.cz project

This presentation is supported by MEYS project LM2023061 (FERMILAB-CZ)