COMPUTING FOR ALICE IN THE CZECH REPUBLIC
2022-2024

D. Adamova (NPI ASCR)
J. Chudoba, A. Mikula, L. Míča, M. Svatoš , P. Šesták , J. Uhlířová , P. Vokáč (all FZU ASCR)**, J. Hampl** (NPI ASCR)

Seoul, 18.04.2024

HEP Computing in Prague: WLCG site praguelcg2
(a.k.a. the farm GOLIAS)

- A national computing center for processing data from various HEP experiments
- Distributed resources with all central services and most of hardware at the central site: FZU ASCR
- Basic infrastructure already in 2002.
- *One of the 13 sites participating in the first LCG campaign in 2003*

- *April 2008, WLCG MoU signed* by Czech Republic (ALICE+ATLAS)
- *Certified as a Tier2 center of LHC Computing Grid (praguelcg2)*
- Collaboration with various other Grid projects

- *Very good network connectivity* provided by CESNET / e-INFRA CZ.
- Multiple dedicated 10 – 100 Gb/s connections to collaborating institutions, 100 Gb/s connection to LHCONE.

- Provides computing services for ATLAS + ALICE, Auger, NOVA, Fermilab, Astrophysics …

Czech republic Tier-2 center - geographical layout

- **Distributed resources**
- **Several clusters under one HTCondor LRMS**: 10000
- jobslots (5000 physical cores)
  - 4 types of hardware, oldest from 2017 (Intel Xeon E5-2650 v4)
  - 1 cluster (from 2016) switched off in June 2023 (15 kHS06, 1440 jobslots) and moved to MFF UK
  - Plus 1240 jobslots at MFF UK, 16 kHS06
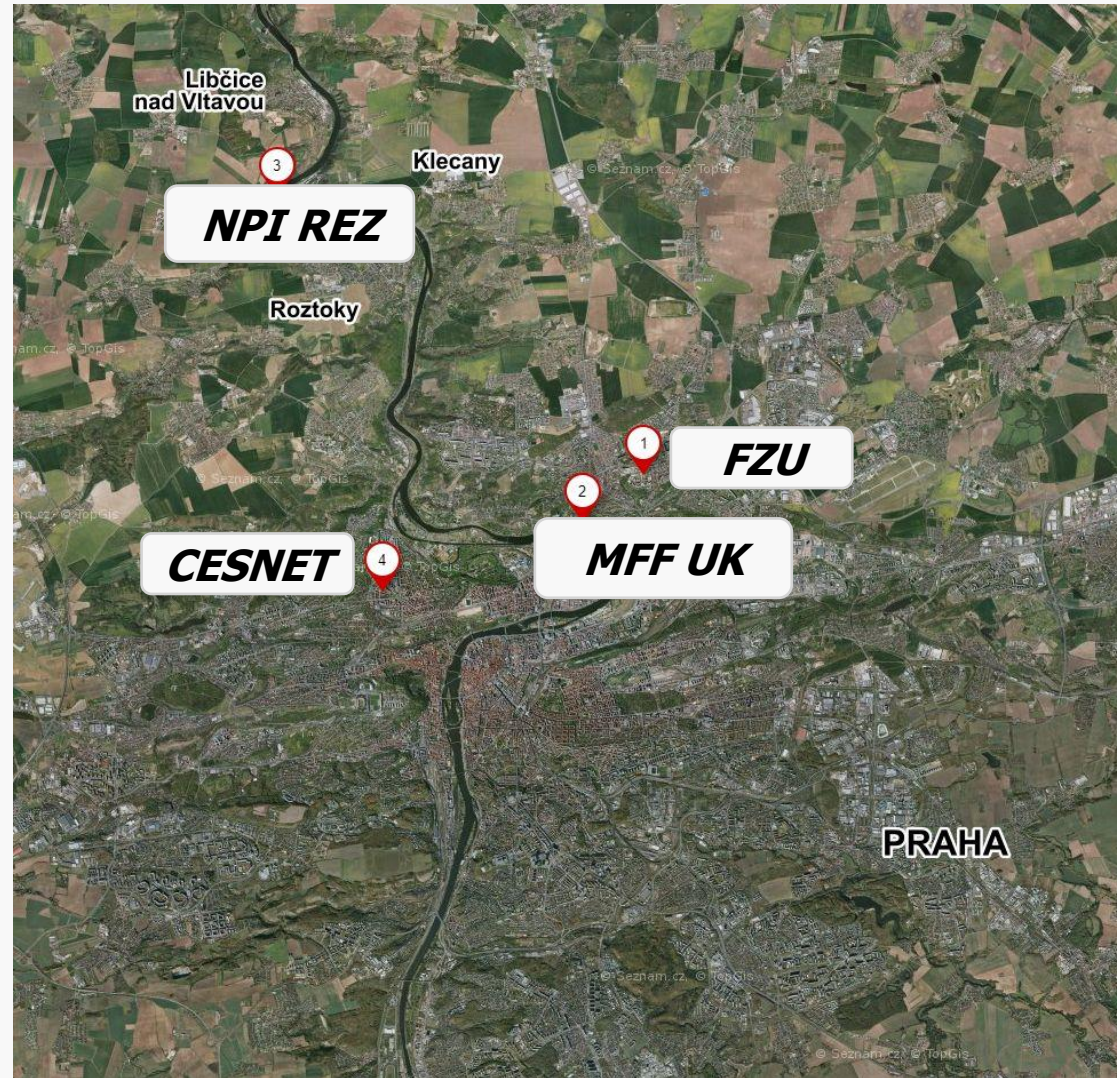
Extended geographical layout

Other participating institutions – in Prague
Charles University (UK), Faculty of Mathematics and Physics
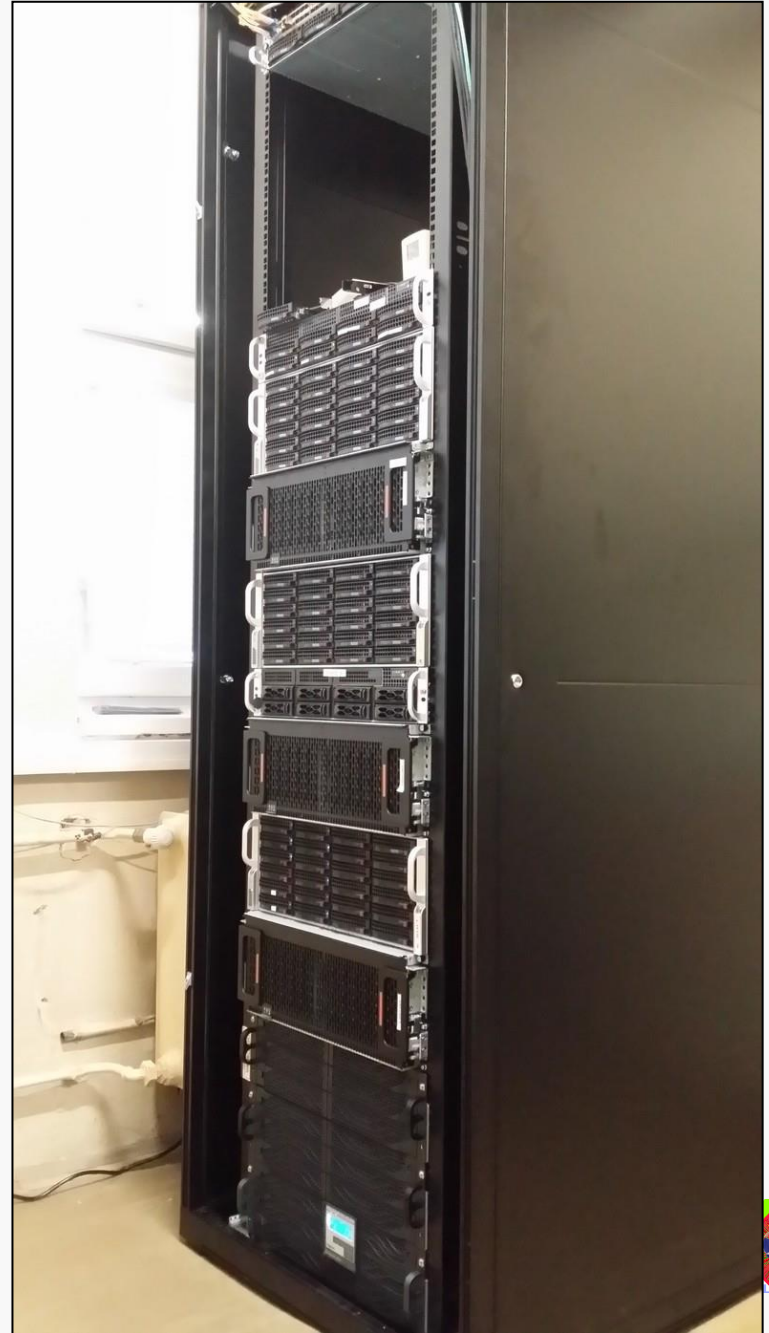Czech Technical University (CTU)
CESNET MetaCentrum
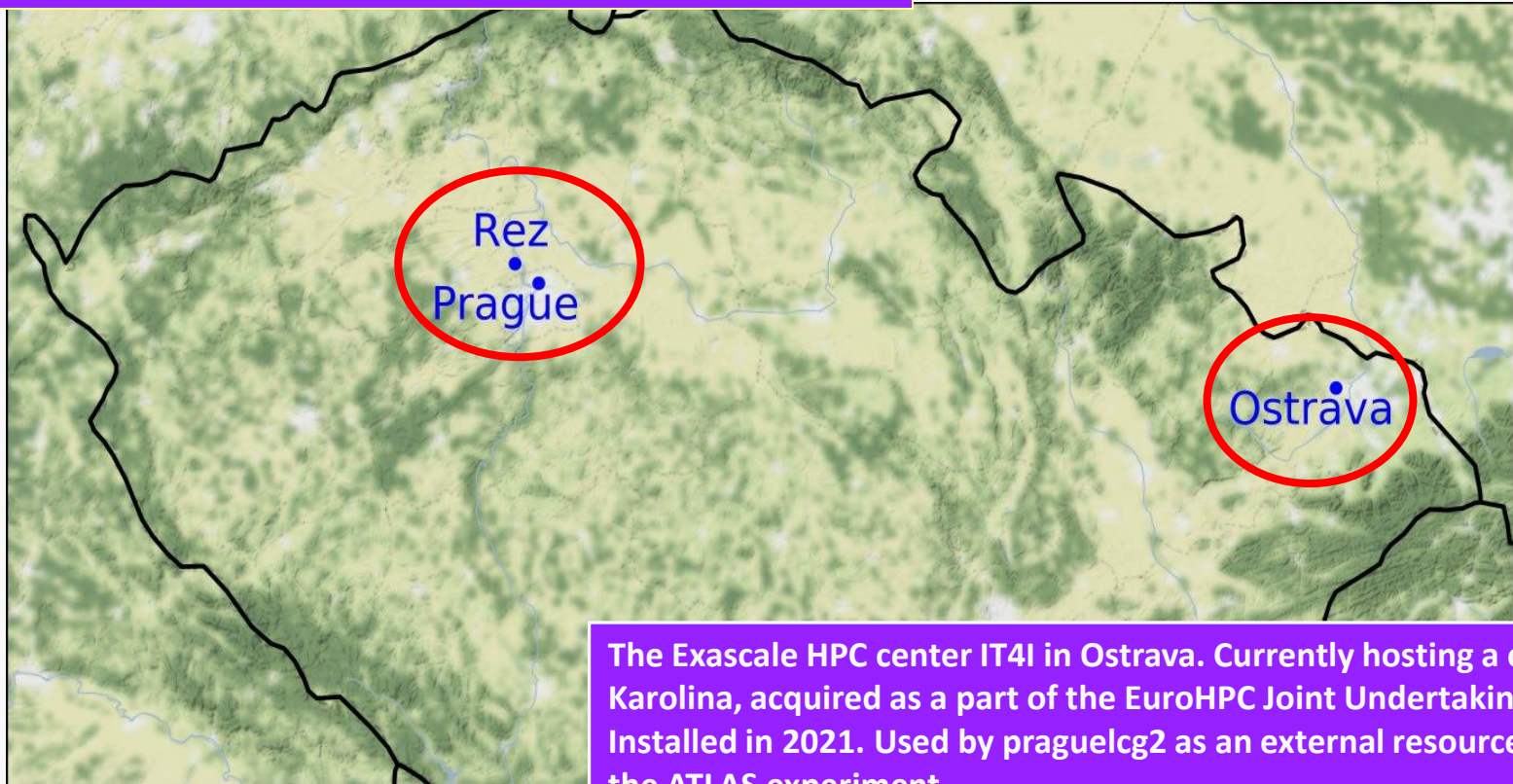
The Exascale HPC center IT4I in Ostrava. Currently hosting a cluster Karolina, acquired as a part of the EuroHPC Joint Undertaking. Installed in 2021. Used by praguelcg2 as an external resource for the ATLAS experiment.
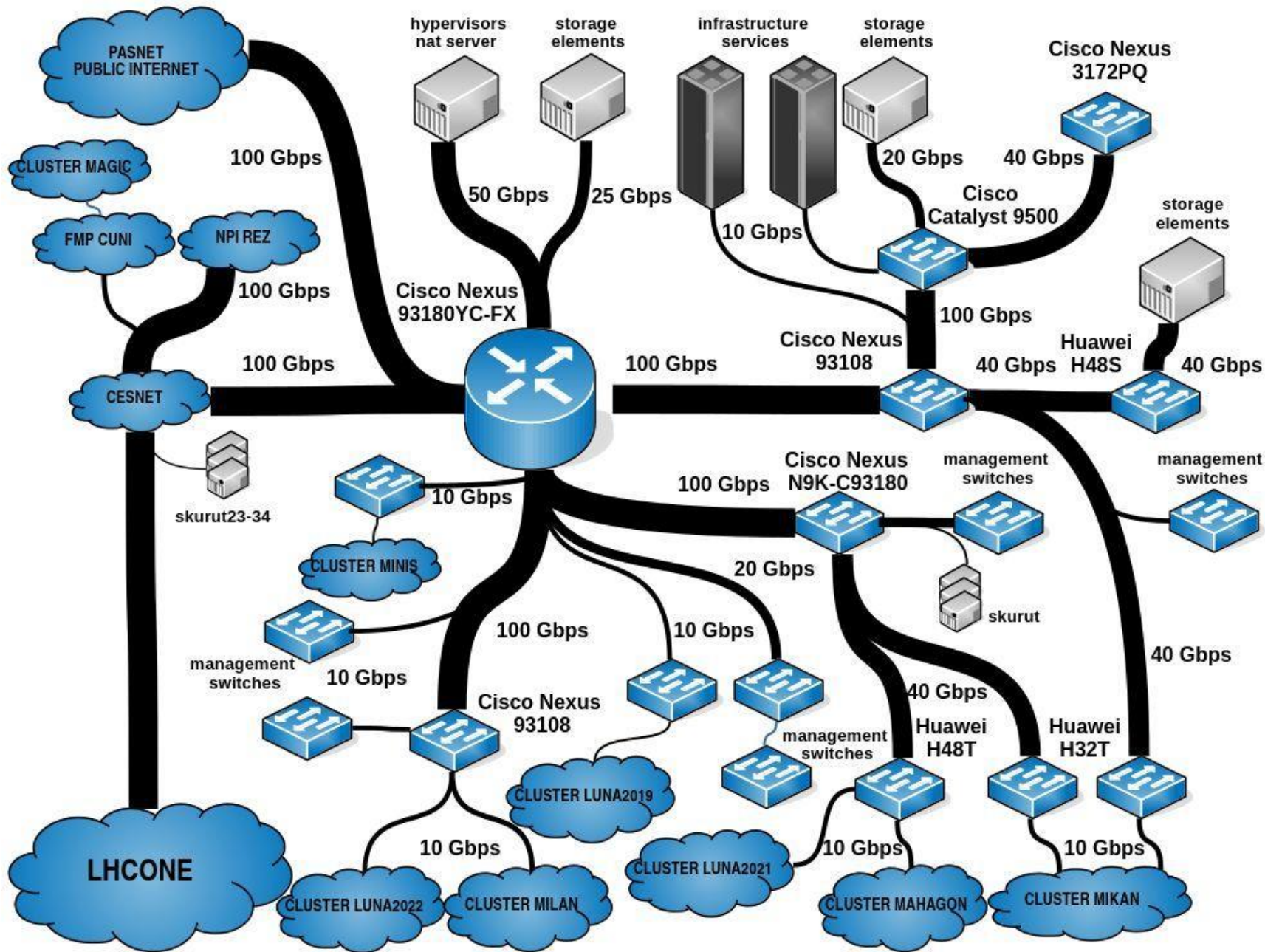IT4I is also involved in the LUMI consortium, thus providing access to the LUMI supercomputer (2022).

Current resources

- Dedicated server room 62 m^2 (2015)
- 1 batch system (HTCondor)
- 2 main WLCG VOs: ALICE, ATLAS
- ~ 10900 cores (~149kHS)
-
- ~ 10.3 PB in total on disk grid storage on site and at NPI (dCache, XRootD)
- ~ 1PB on NFS servers
- OS: WN on Centos7, services mostly on AL9
- planned: update of AliceBox

- Monitoring: Grafana and icinga
- Configuration management by Puppet
- Provisioning and SW management by Foreman

- ALICE XRootD storage at NPI – 3.33 PB of disk space
- Dedicated optical connection between NPI and FZU: throughput 100 Gb/s.

PASNET PUBLIC INTERNET

hypervisors nat server

storage elements

infrastructure services

storage elements

Cisco Nexus 3172PQ

CLUSTER MAGIC

100 Gbps

50 Gbps 25 Gbps

20 Gbps 40 Gbps

FMP CUNI **NPI REZ**

Cisco Catalyst 9500

storage elements

100 Gbps

10 Gbps

Cisco Nexus 93180YC-FX

Huawei H48S

100 Gbps

100 Gbps 100 Gbps

Cisco Nexus 93108

40 Gbps 40 Gbps

CESNET

skurut23-34

10 Gbps

Cisco Nexus N9K-C93180

management switches

management switches

CLUSTER MINIS

100 Gbps

20 Gbps

skurut

management switches

100 Gbps 10 Gbps

40 Gbps

management switches 10 Gbps

Cisco Nexus 93108

10 Gbps

management switches

40 Gbps **Huawei H48T**

Huawei H32T

LHCONE

CLUSTER LUNA2019

10 Gbps

10 Gbps 10 Gbps

CLUSTER LUNA2022 **CLUSTER MILAN**

CLUSTER LUNA2021

CLUSTER MAHAGON **CLUSTER MIKAN**

Recent upgrades, new hardware

- *Recent : 2022 - 2023*

- 2022 NPI : new server rack cabinets installed + new UPS units with bigger capacity

- 2023 NPI : new air conditioning of the NPI server room

- In 2023 applied for ASCR investment for disk server, capacity ~ 1.32 PB. The approval process took 1 year, currently the purchase finished, installation and tests ongoing.

- November 2023 : increase of capacity of the link NPI – FZU to 100 Gb/s.

Current perspectives for 2023 - 2026

- Min. of Education investments project OPJAK (2023-2026)

- In it for ALICE: Disk storage: 0.9 PB & CPU: 12.5 kHS23

- Additional support needed for sufficient disk storage capacity by 2026

- CPU remains to be largely dependent on opportunistic resources

New server rack cabinets at NPI

# Delivery of Czech computing resources during 2019 - 2024

All pledges fulfilled, only 1 exception

|	REQUESTED	DELIVERED
2019		
CPU/year (HS06 hours/HS06)	152,928,000 / 17700	227,570,650 / 26339
Disk (PB)	1.6	2.0
2020		
CPU/year (HS06 hours/ HS06)	157,248,000 / 18200	294,191,948 / 3405
Disk (PB)	2.04	1.9
2021		
CPU/year (HS06 hours/ HS06)	154,656,000 / 17900	321,702,339 / 37234
Disk (PB)	1.09	3.09
2022		
CPU/year (HS06 hours/ HS06)	186,192,000 / 21550	419,203,708 / 48519
Disk (PB)	2.2	2.9
2023		
CPU/year (HS23 hours/ HS23)	175,680,000 / 24400	342,700,809 / 39664
Disk (PB)	2.5	2.9 (till March) / 2.79 (since March)
2024		
CPU (HS23)	24300	29487
Disk (PB)	2.6	2.7 (3.33 raw)
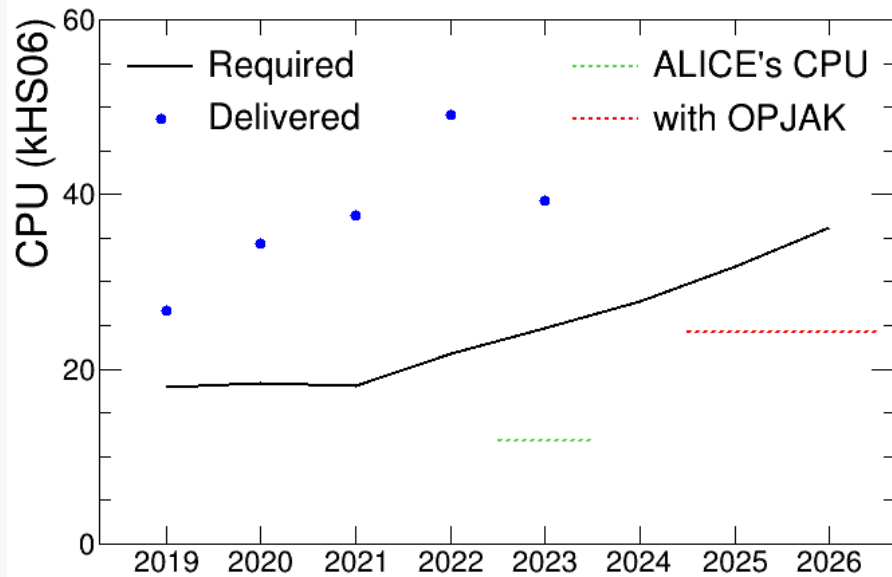
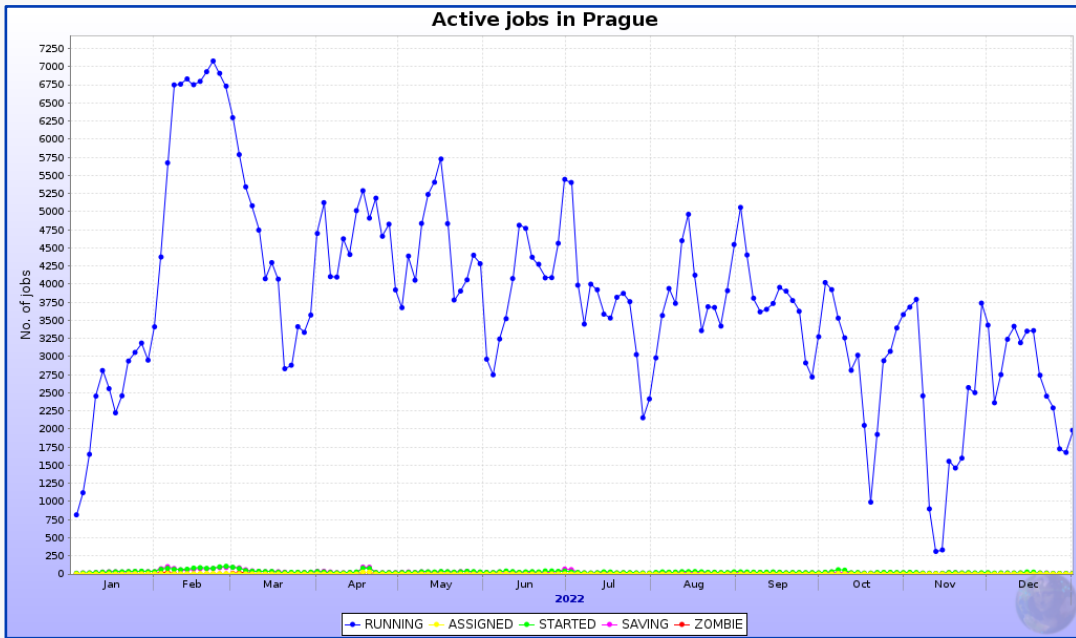*Based on earlier estimates of pledges for 2025 and 2026*
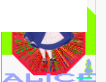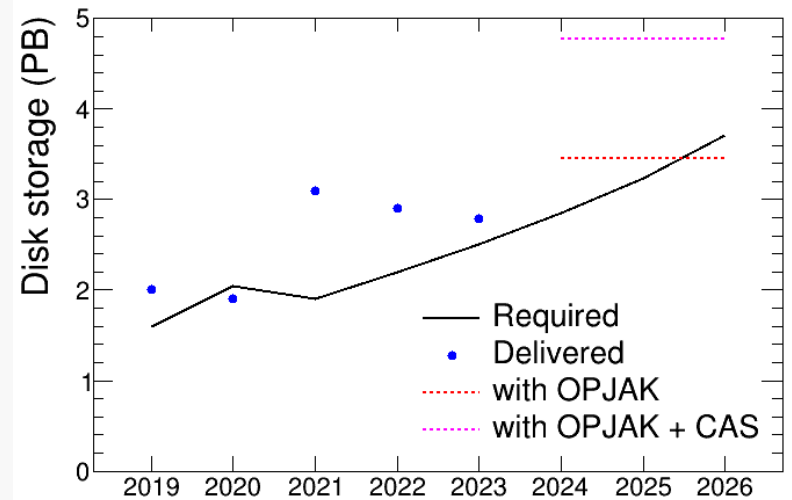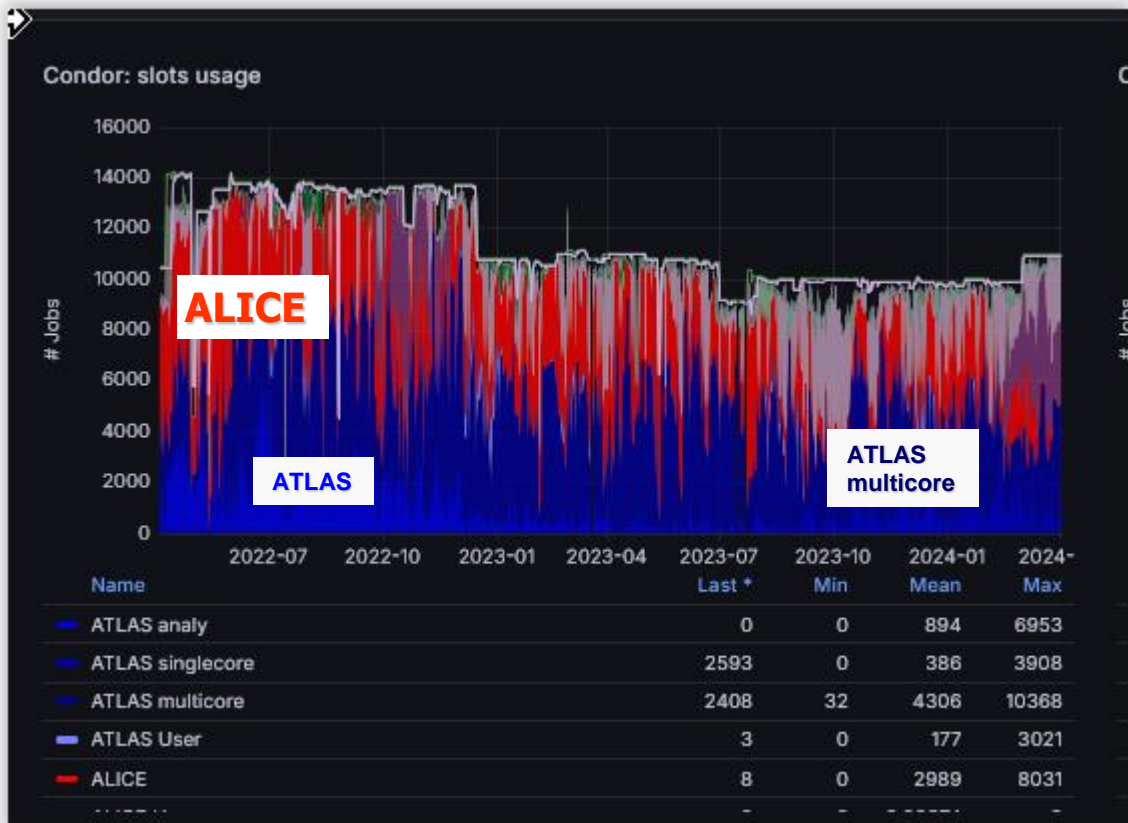
Running jobs profile during last 2 years, by the local monitoring at the central site in Prague. *Main CPU consumers are ALICE and ATLAS.*
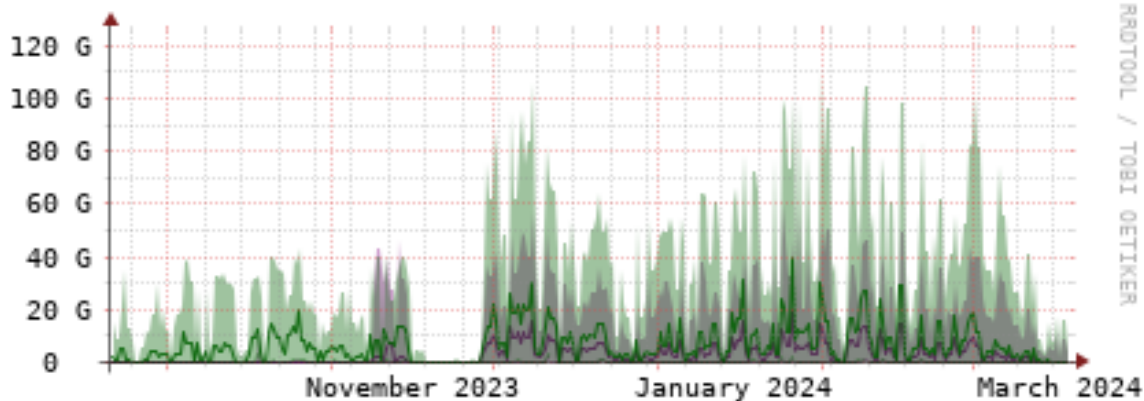
Network traffic between the main site of praguelcg2 and NPI storage cluster during last 6 months. *In November, the link was elevated to 100 Gb/s.*

Active jobs in Prague

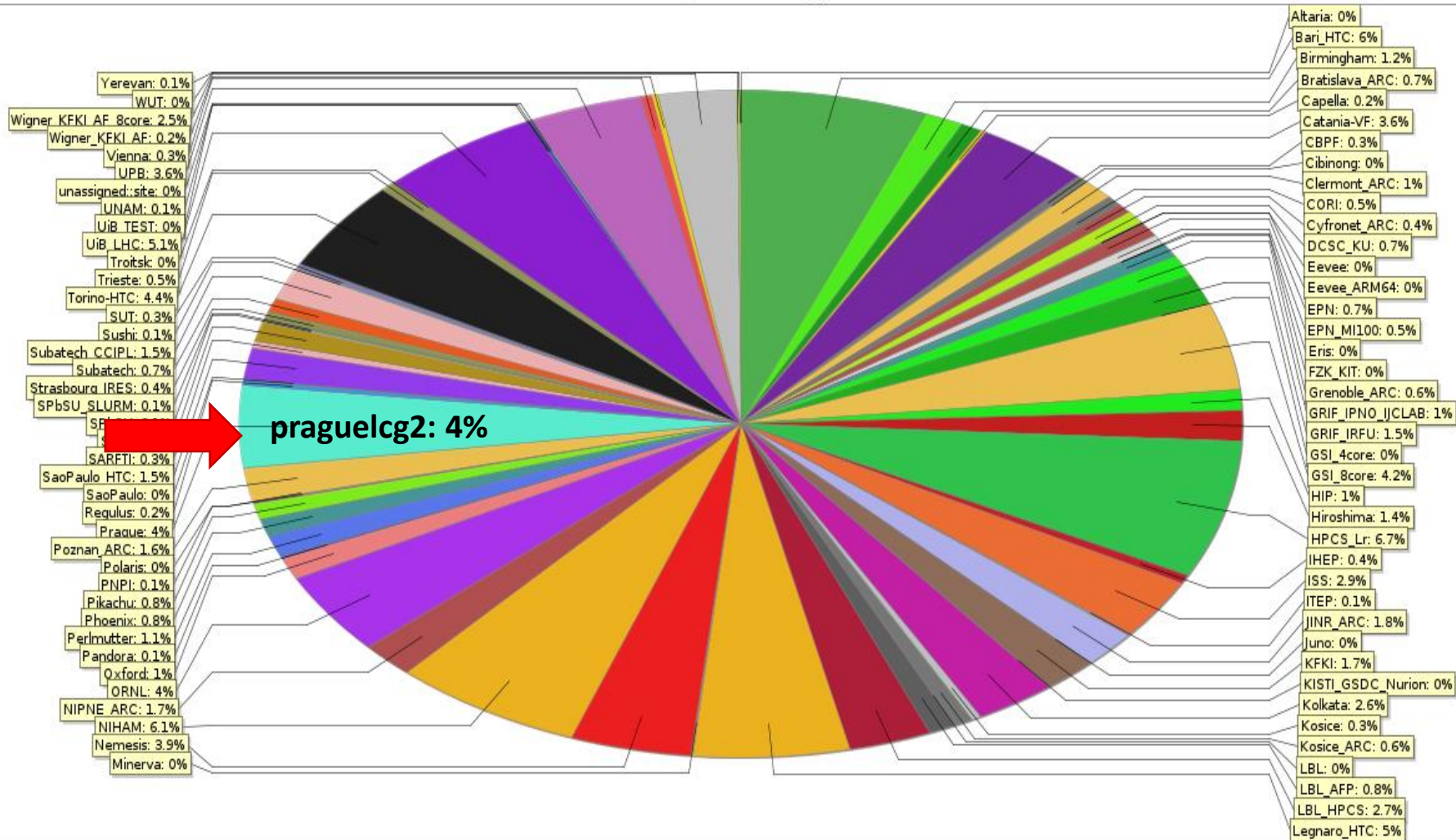Running jobs profile in MonaLisa in 2022.
Avg. 3756, Max. 8913

Active jobs in Prague
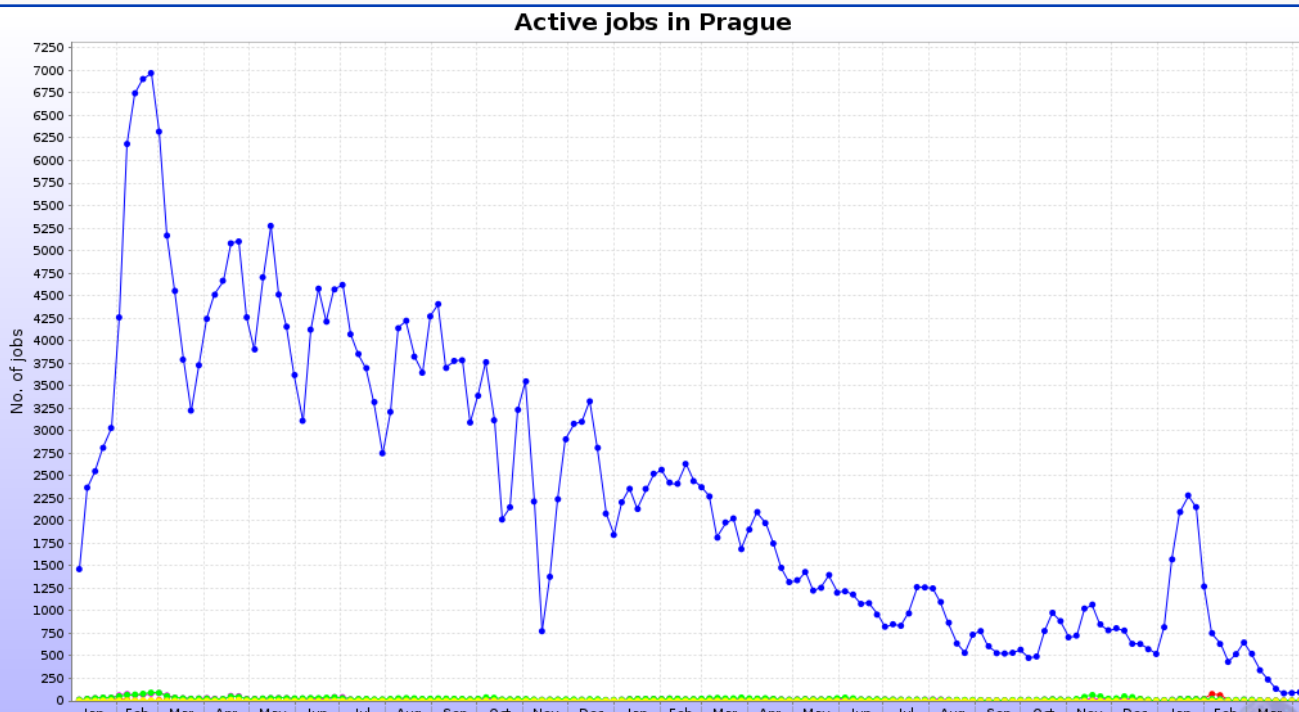
Running jobs profile in MonaLisa,
01/2022-03/2024.
Avg. 2325, Max. 8913

For 01/2023-03/2024
Avg.1215, Max. 7670

Average running jobs

Altaria: 0%
Bari_HTC: 6%
Birmingham: 1.2%
Bratislava_ARC: 0.7%
Capella: 0.2%
Catania-VF: 3.6%
CBPF: 0.3%
Cibinong: 0%
Clermont_ARC: 1%
CORI: 0.5%
Cyfronet_ARC: 0.4%
DCSC_KU: 0.7%
Eevee: 0%
Eevee_ARM64: 0%
EPN: 0.7%
EPN_MI100: 0.5%
Eris: 0%
FZK_KIT: 0%
Grenoble_ARC: 0.6%
GRIF_IPNO_IJCLAB: 1%
GRIF_IRFU: 1.5%
GSI_4core: 0%
GSI_8core: 4.2%
HIP: 1%
Hiroshima: 1.4%
HPCS_Lr: 6.7%
IHEP: 0.4%
ISS: 2.9%
ITEP: 0.1%
JINR_ARC: 1.8%
Juno: 0%
KFKI: 1.7%
KISTI_GSDC_Nurion: 0%
Kolkata: 2.6%
Kosice: 0.3%
Kosice_ARC: 0.6%
LBL: 0%
LBL_AFP: 0.8%
LBL_HPCS: 2.7%
Legnaro_HTC: 5%

Yerevan: 0.1%
WUT: 0%
Wigner_KFKI_AF_8core: 2.5%
Wigner_KFKI_AF: 0.2%
Vienna: 0.3%
UPB: 3.6%
unassigned::site: 0%
UNAM: 0.1%
UiB_TEST: 0%
UiB_LHC: 5.1%
Troitsk: 0%
Trieste: 0.5%
Torino-HTC: 4.4%
SUT: 0.3%
Sushi: 0.1%
Subatech_CCIPL: 1.5%
Subatech: 0.7%
Strasbourg_IRES: 0.4%
SPbSU_SLURM: 0.1%
SPbSU:
SARFTI: 0.3%
SaoPaulo_HTC: 1.5%
SaoPaulo: 0%
Regulus: 0.2%
Prague: 4%
Poznan_ARC: 1.6%
Polaris: 0%
PNPI: 0.1%
Pikachu: 0.8%
Phoenix: 0.8%
Perlmutter: 1.1%
Pandora: 0.1%
Oxford: 1%
ORNL: 4%
NIPNE_ARC: 1.7%
NIHAM: 6.1%
Nemesis: 3.9%
Minerva: 0%

praguelcg2: 4%

Czech republic share in 01/2023 – 03/2024 was 4% (In 2022 it was 7%)

More CPUs for praguelcg2 – IT4I

National Supercomputing Center

IT4Innovation in Ostrava

- **200 000** node hours in the project OPEN-27-57 (end 2023).

  - **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)
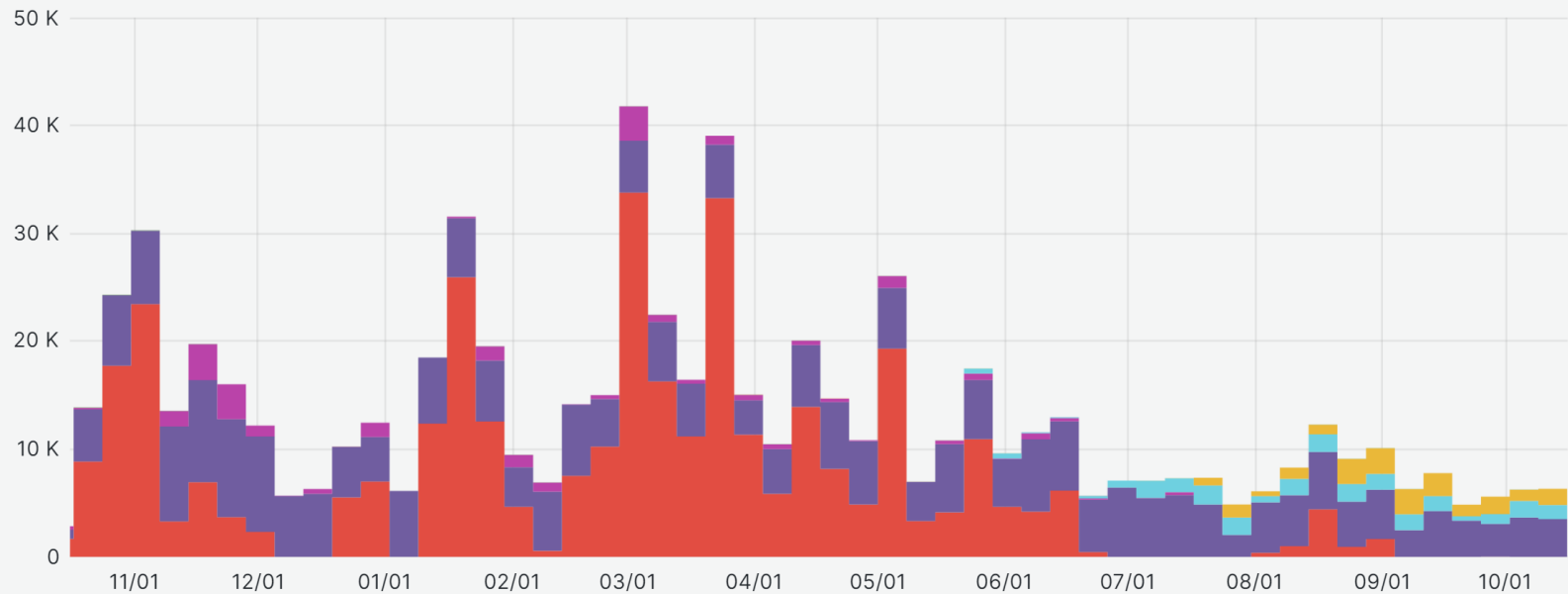
  - test access to **LUMI**

LUMI: Supercomputer in Finland (2022). Joint venture EuroHPC and the LUMI consortium. Access of the Czech republic thanks to the involvement of the IT4I in the LUMI consortium.

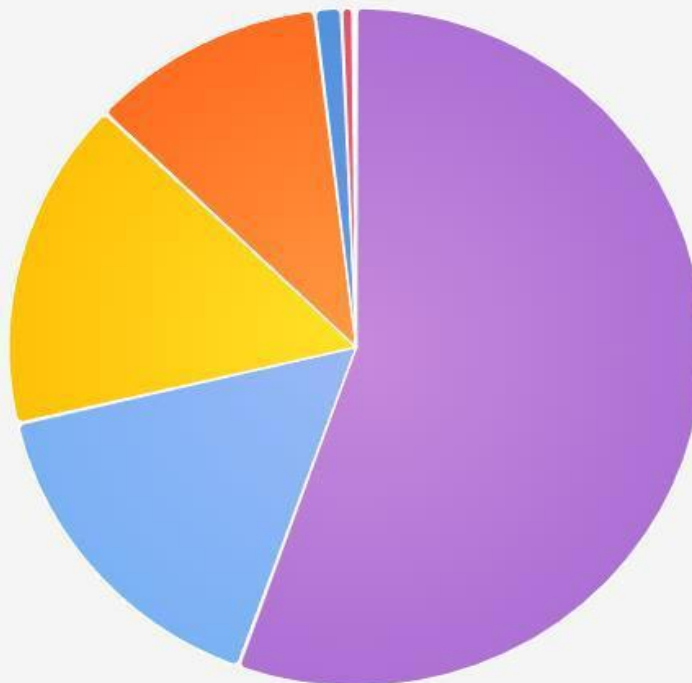# IT4I: ATLAS running jobs slots. 11/2022 – 10/2023

Slots of Running jobs ⓘ

	min	max	avg ⌄
🔴 praguelcg2_Karolina_MCORE	0	33.8 K	6.55 K
🟣 praguelcg2	97.1	9.45 K	5.11 K
🟣 praguelcg2_Barbora_MCORE	0	3.34 K	468
🔵 praguelcg2_Koios_MCORE	0	1.76 K	408
🟡 praguelcg2_Phoebe_MCORE	0	2.38 K	347

# ATLAS jobs wall time distribution. 11/2022 – 10/2023

Wall clock time. All jobs (HS23 seconds)

ATLAS: almost 50% of computing resources delivered from external resource.
Good for ALICE: more opportunistic resources at the main site.

	Value	Percent
praguelcg2	459 Bil	56%
praguelcg2_Koios_MCORE	131 Bil	16%
praguelcg2_Phoebe_MCORE	130 Bil	16%
praguelcg2_Karolina_MCORE	90.0 Bil	11%
praguelcg2_BOINC	9.64 Bil	1%
praguelcg2_Barbora_MCORE	3.73 Bil	0%
praguelcg2_fzu_TEST	570 Mil	0%
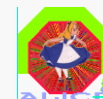
Summary

- Decommissioned older CPU servers compensated by non-dedicated resources

- New project for WLCG dedicated hardware accepted (2023-2026).

- Czech Tier-2 center is successfully delivering CPU and storage resources to the Grid and local users with minimal outage.

- Receiving only 8-core jobs for ALICE limits unfortunately the extensity of appropriate opportunistic resources.

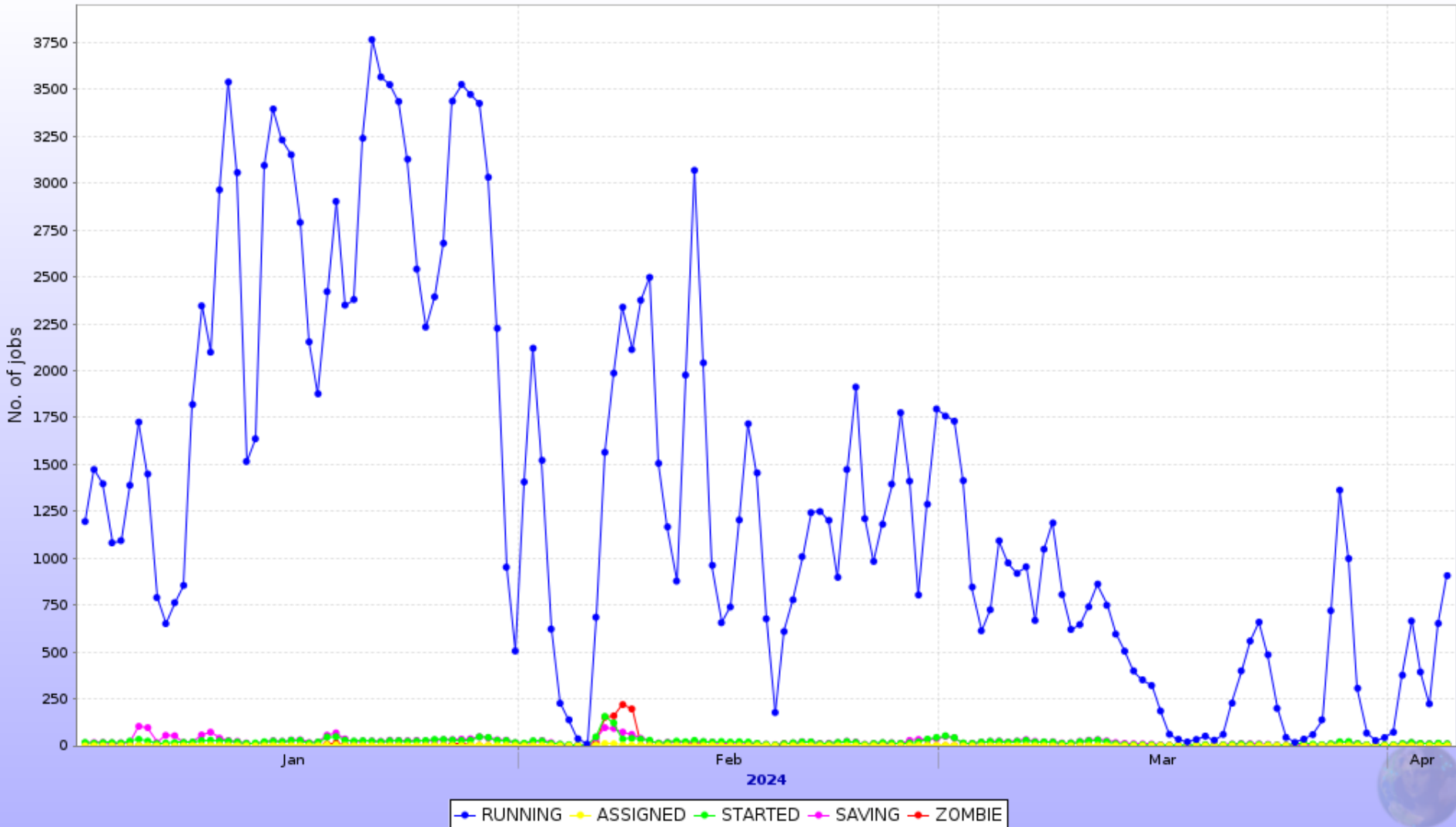- We are prepared to make sure we deliver resources for ALICE as requested in the coming years.

BACKUP

Running jobs Legnaro_HTC 01.2024 – 03.2024

Active jobs in Legnaro_HTC

Use of the resources of IT4I, current status, by ATLAS

- 2023: Standard project with highest allocation this year
- 25.6M corehours
- allocation almost used up already within 2 months
- equivalent to 66000 FZU cores used for 2 months

IT4I cores used by ATLAS jobs

	min	max	avg ⌄
▬ praguelcg2_Karolina_MCORE	0	22.5 K	2.16 K
▬ praguelcg2_Salomon_MCORE	0	2.51 K	369