Distributed computing infrastructure for the SPD experiment

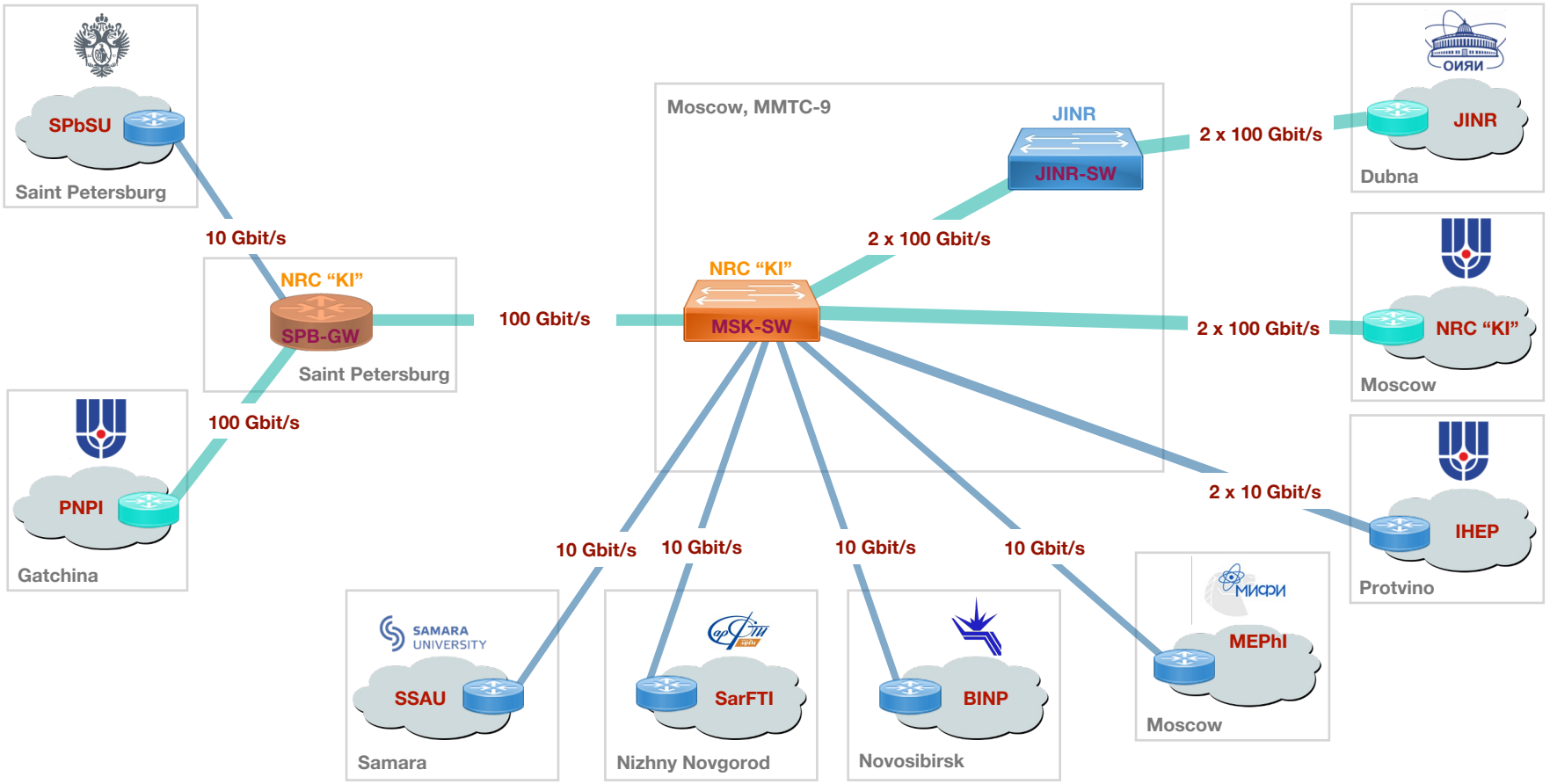
Andrey Kiryanov, NRC KI – PNPI

GRID’2025, 7-11 July 2025

Introduction

- SPD relies heavily on a distributed offline computing, which is currently in active development
- Collaboration participants can provide computing and storage resources with standard interfaces
- SPD provides central facilities for security, job and data management, software distribution, monitoring
- Russian Scientific Backbone provides broadband network infrastructure for the most major scientific organizations in Russia
  - Peering with Russian and international networks

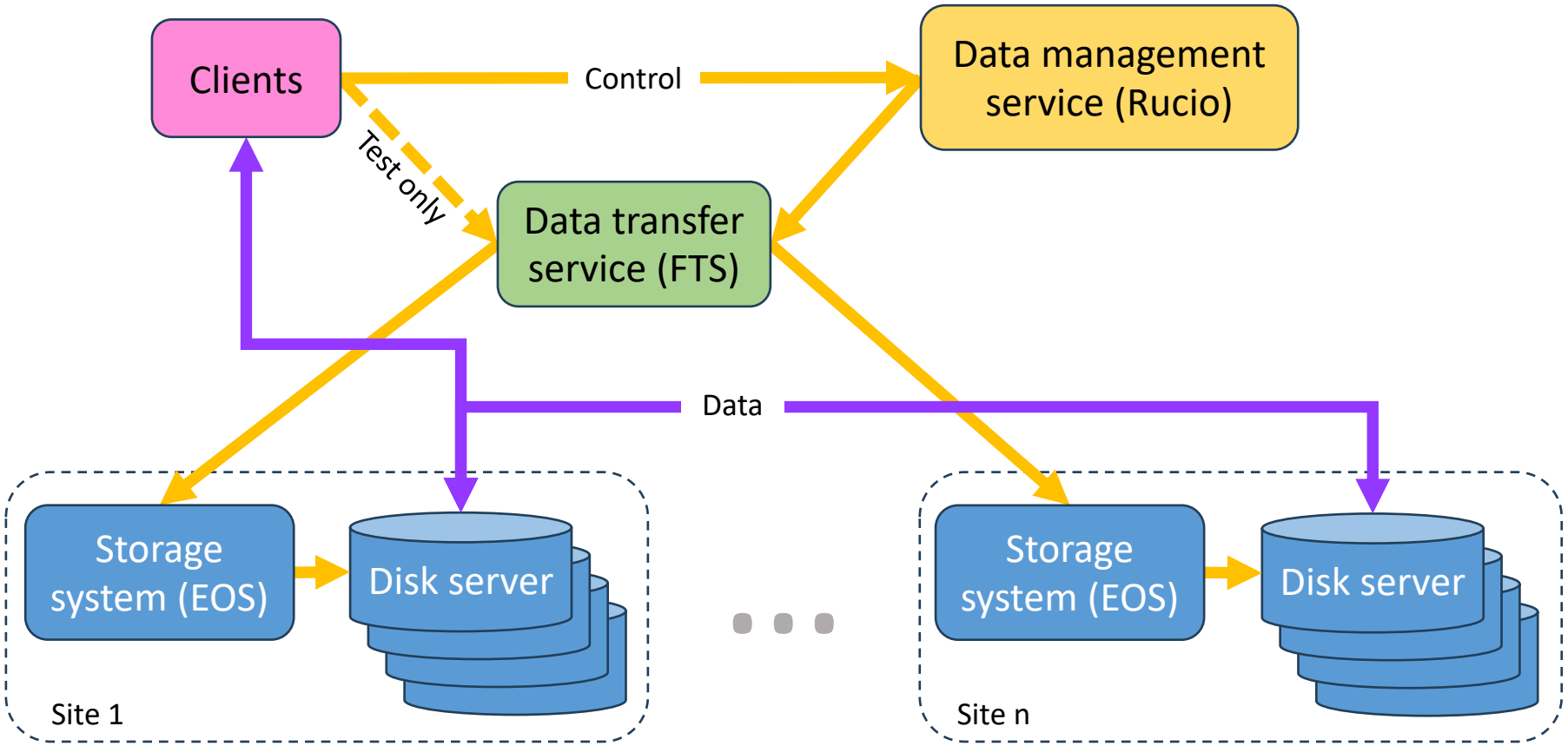
Russian Scientific Backbone



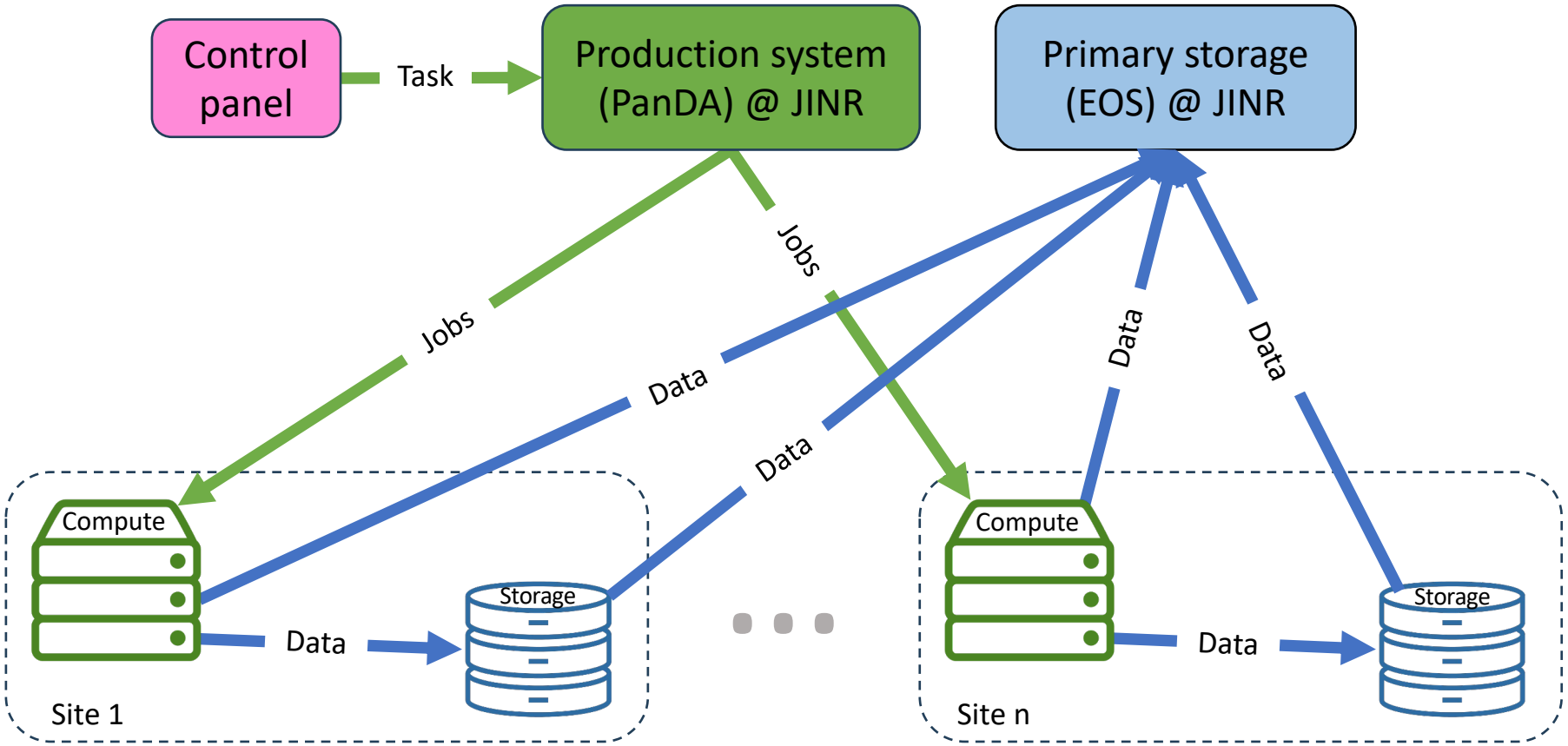
Computing building blocks

- Core of the production system deployed at JINR
  - IAM – the root of trust
  - CRIC – resource registry
  - CVMFS Stratum 0 – software repository
  - PanDA/Harvester – job management
  - Rucio/FTS – data management
  - Monitoring for all of the above
- Sites are expected to deploy a limited set of ”extra” software
  - ARC CE (or HTCondor if you really need to) – job submission interface
  - Slurm (or Condor) – local batch system
    - PBS, LSF, SGE and clones are NOT recommended
  - NFS (or Lustre, CephFS, etc.) – local shared filesystem
  - CVMFS – software distribution
  - Local monitoring (Zabbix, etc.)

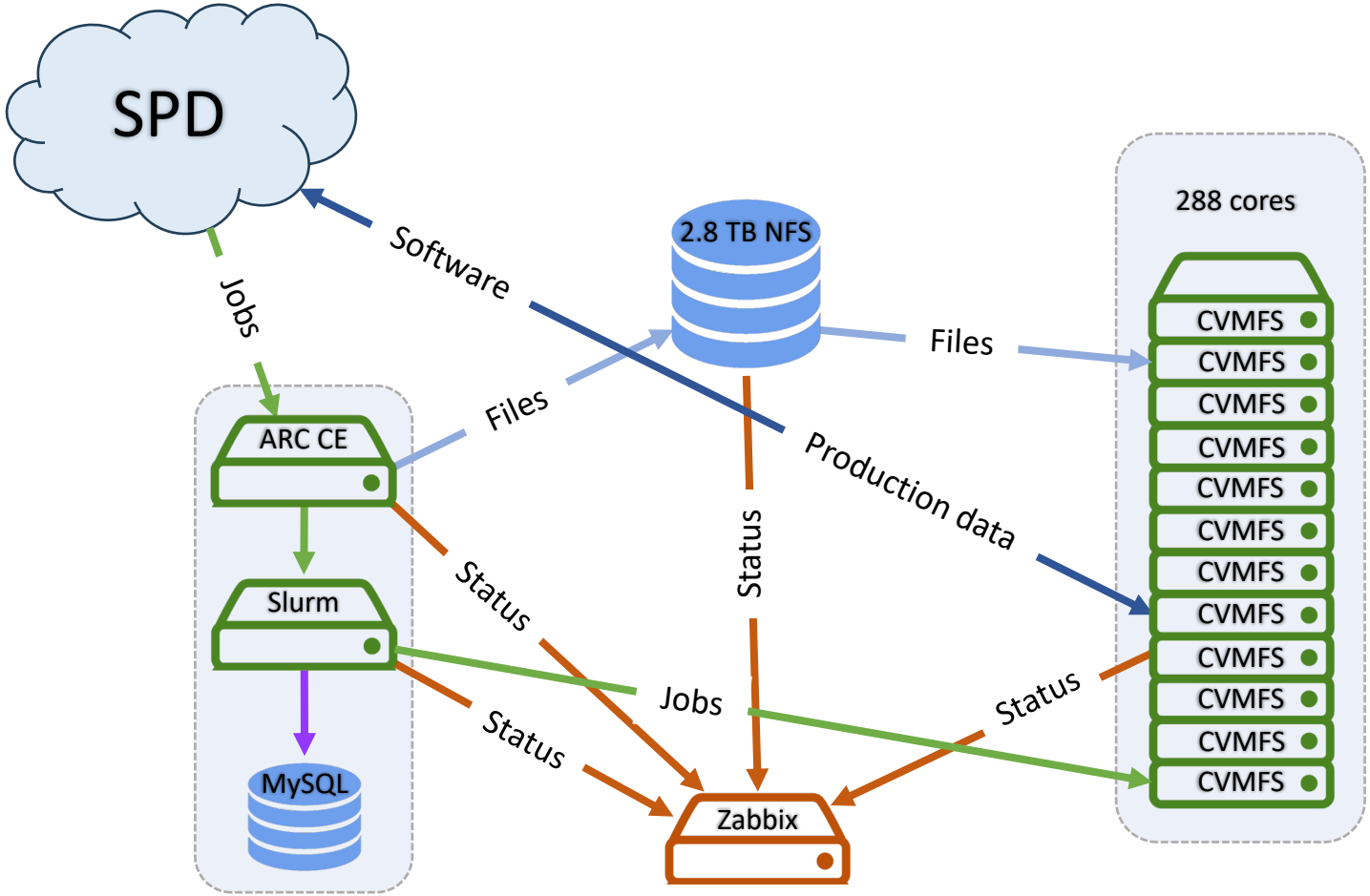
Data management in SPD



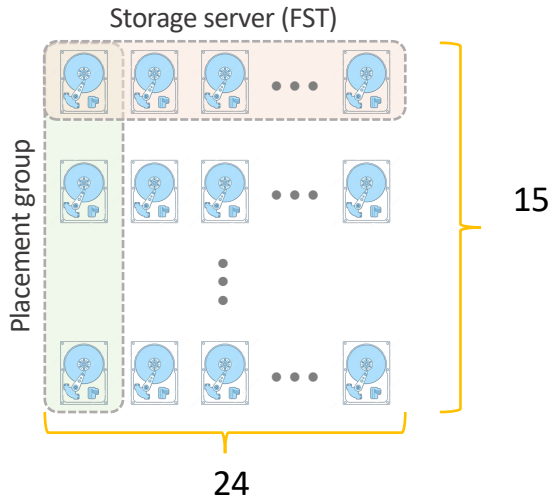
Distributed computing in SPD



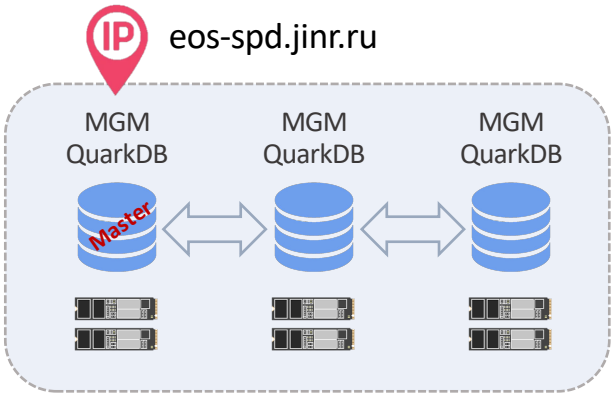
Computing deployment at PNPI



EOS deployment at JINR



7.2 PB of raw disk storage
24 placement groups
15 stripes per group
11+4 QRAIN layout (27% overhead)
5.3 PB of usable space

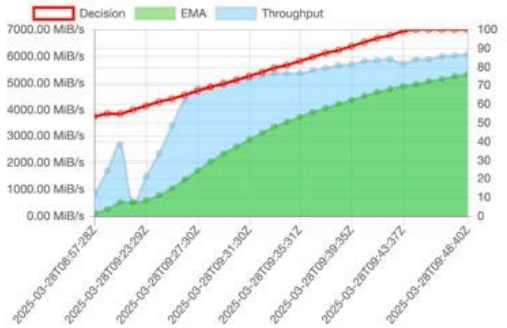


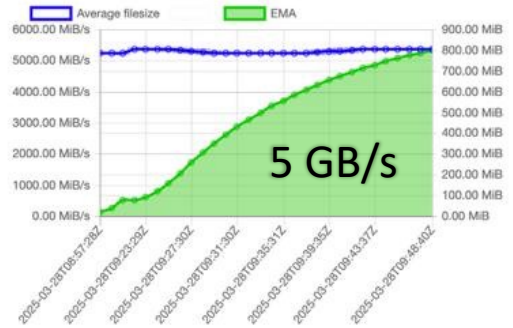
HA MGM/QuarkDB on 3 nodes
1.5 TB of namespace
Floating IP alias

Data transfer rates

Details for https://juno-se-dr01.jinr.ru → https://eos-spd.jinr.ru

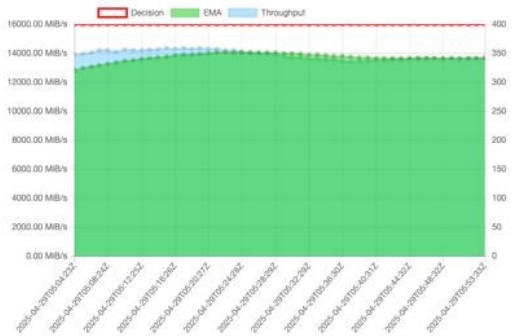
Storage	Current active transfers	Max. active transfers
https://juno-se-dr01.jinr.ru	0	200
https://eos-spd.jinr.ru	0	200

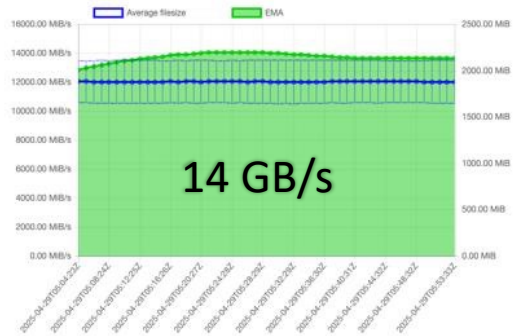




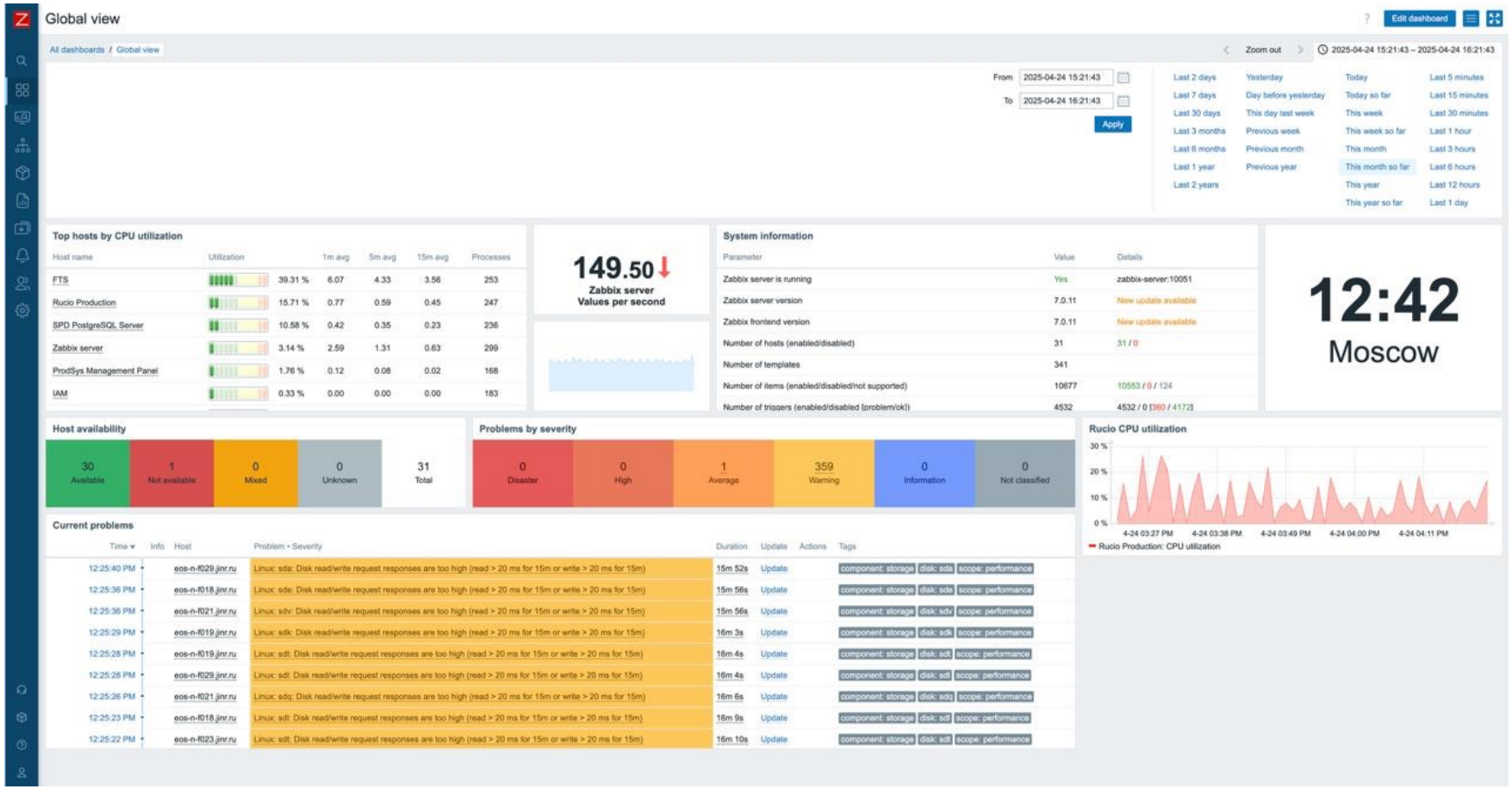
Details for https://eos.jinr.ru → https://cta-eos.jinr-t1.ru

Storage	Current active transfers	Max. active transfers
https://eos.jinr.ru	389	800
https://cta-eos.jinr-t1.ru	389	1000





Zabbix monitoring at JINR



What is currently deployed

- JINR
  - Production system services (prod and devel instances)
  - Computing (2200 cores)
  - Storage (7.2 PB raw with 27% redundancy = 5.3 PB)
  - Monitoring
- PNPI
  - Computing (288 cores)
  - Storage (190 TB redundant)
  - Monitoring
- SSAU
  - Computing (256 cores)
  - Storage is on the way (240 TB raw with 17% redundancy = 200 TB)
- MEPhI
  - Ongoing negotiations

What is yet missing

- Monitoring coverage for core services
  - We need to monitor the health of services, not just hardware
  - Visualization (dashboards, etc.)
- Periodic infrastructure tests
  - Network monitoring
  - Job submission tests
  - Worker node health tests
  - Data transfer tests
- Full transition to tokens
  - EOS-side configuration
  - ARC CE-side configuration
  - User manual

Thank you!