# Federated storage initiatives at NRC "Kurchatov Institute"

Andrey Kiryanov

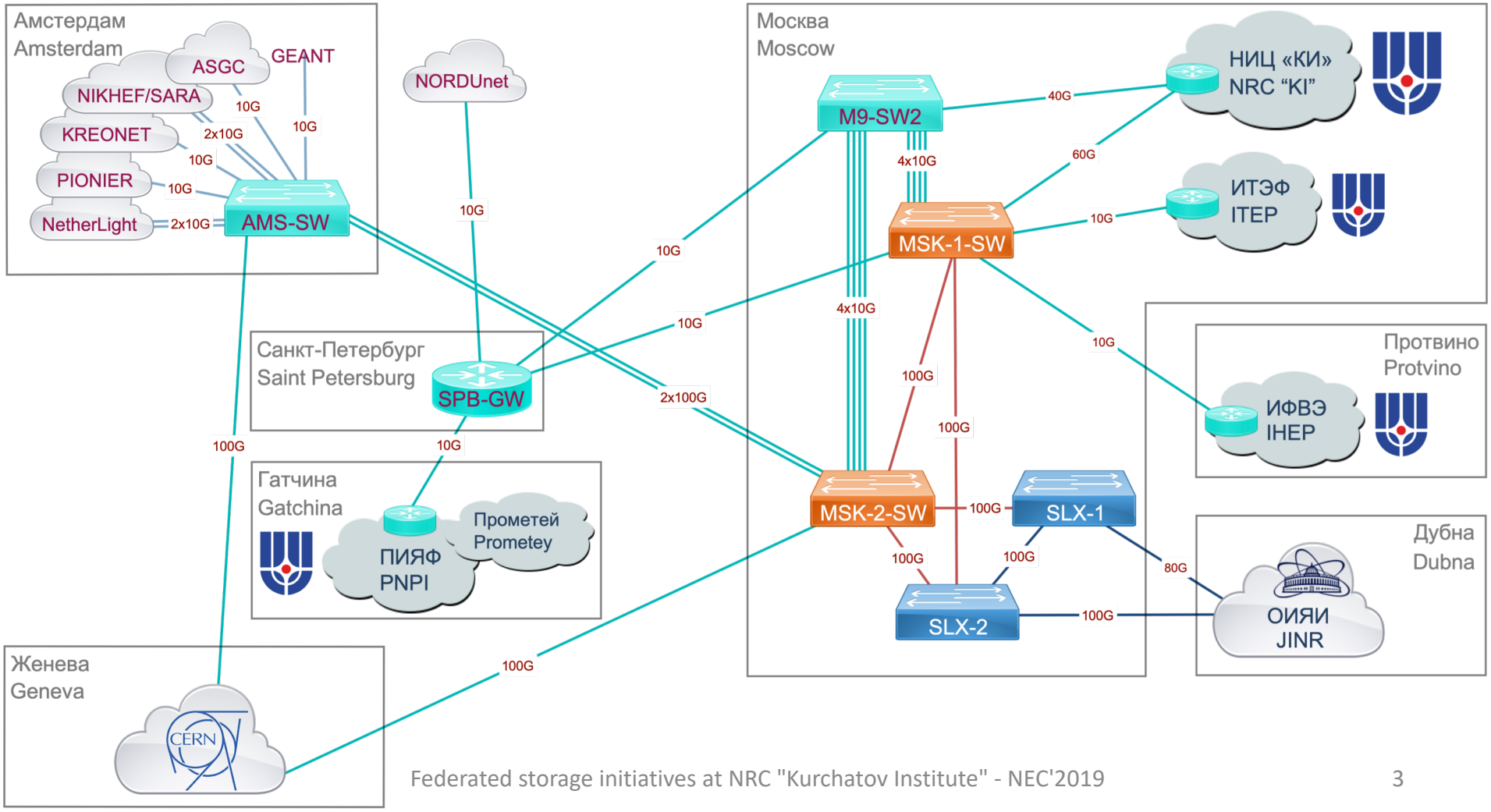
# Rationale

NRC "Kurchatov Institute" is a geographically distributed research center

- 7 research institutes, most of which have their own compute facilities
- Moscow district
  - KI
  - IHEP
  - ITEP
  - IREA
  - GOSNIIGENETIKA
- Saint Petersburg district
  - PNPI
  - Prometey

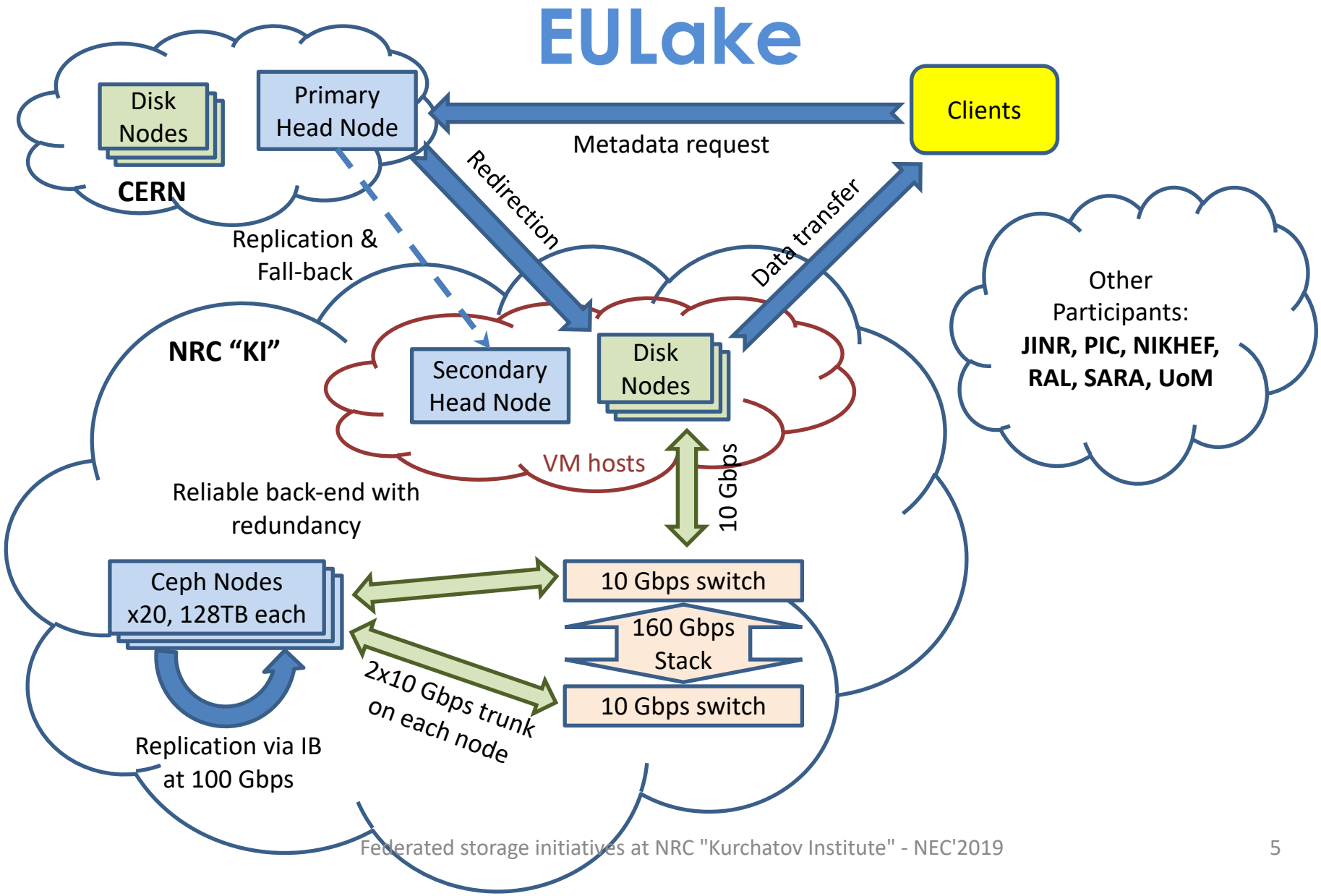
# Network backbone

## for international Megascience-class projects



# Participation in R&Ds

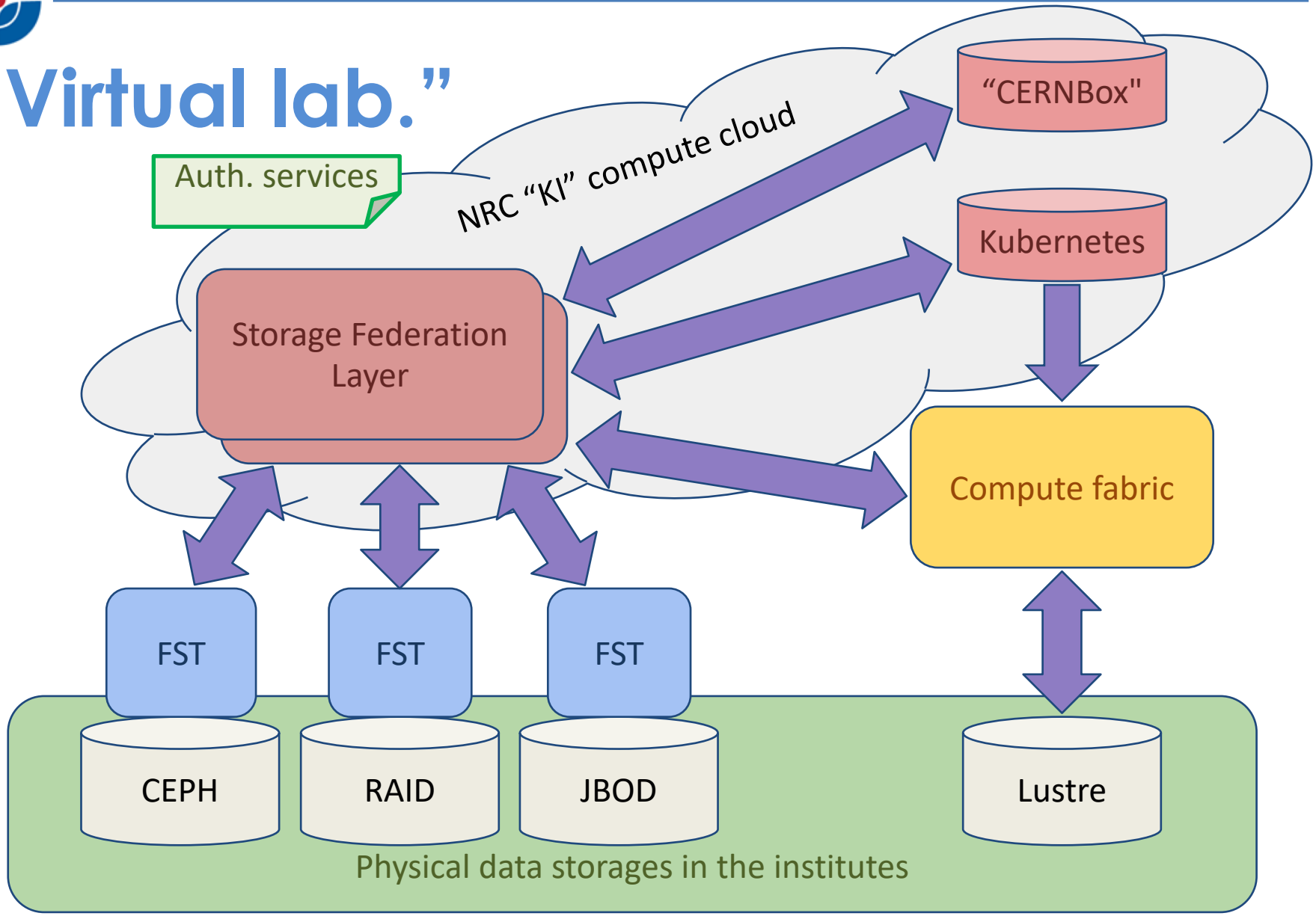
- 2015 – Russian Federated Data Storage
  - EOS + dCache
- 2018 – EU DataLake (EULake)
  - EOS
- 2019 – Russian DataLake
  - EOS + dCache + xCache + …


EULake
Disk Nodes
Primary Head Node
CERN
Clients
Metadata request
Redirection
Data transfer
Replication & Fall-back
Other Participants:
JINR, PIC, NIKHEF, RAL, SARA, UoM
NRC "KI"
Secondary Head Node
Disk Nodes
VM hosts
10 Gbps
Reliable back-end with redundancy
Ceph Nodes x20, 128TB each
10 Gbps switch
160 Gbps Stack
10 Gbps switch
2x10 Gbps trunk on each node
Replication via IB at 100 Gbps

# Ultimate goals

- Unification of access to supercomputer resources of NRC “KI”
- Unification of data storage models for scientific experiments at NRC “KI”
- Unification of user authentication and authorization

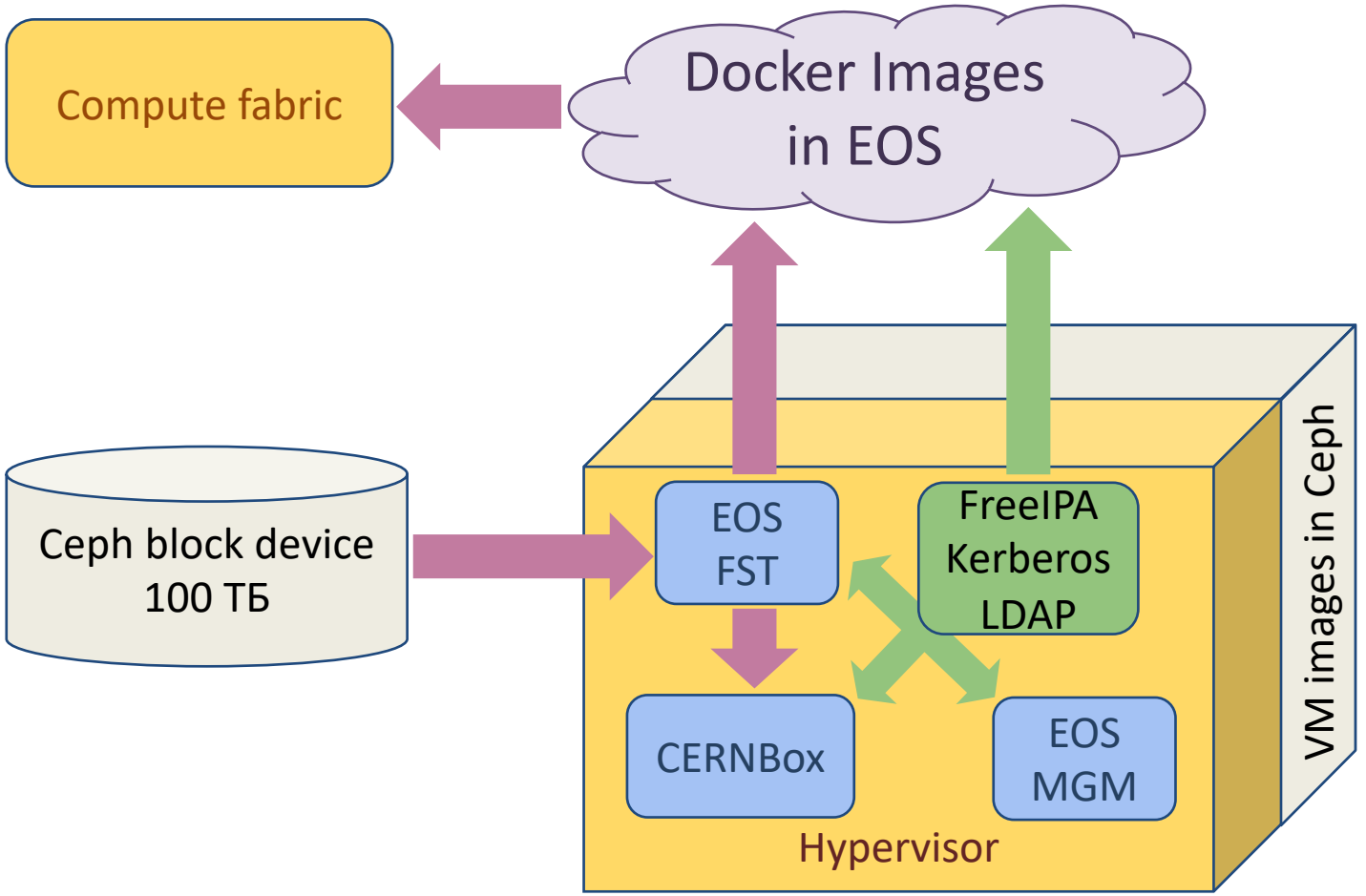
# "Virtual lab."


Auth. services
NRC "KI" compute cloud
"CERNBox"
Kubernetes
Storage Federation Layer
Compute fabric
FST
FST
FST
CEPH
RAID
JBOD
Lustre
Physical data storages in the institutes

# Current status

- Pilot project between KI and PNPI
- Unified registration service
- Two replicated FreeIPA instances
  - Kerberos auth.
- EOS with MGM at PNPI and two FSTs at PNPI (Ceph) and KI (physical storage)
- CERNbox on top of EOS
- Kubernetes orchestrator
- Apache Guacamole for desktop access

# Current status



# Things TBD

- SSO across all web services
  - Moving away from classic login/password
  - OTP?
- Expansion of the pilot onto other NRC ”KI” institutes
- Validation by real experiment’s use cases
- Cybersecurity
  - Moving scientific data unencrypted is OK as long as there’s auth. and integrity verification
  - Moving user’s data across site’s boundaries needs strong encryption
  - Active intrusion prevention for supercomputer resources


This work was supported by the NRC "Kurchatov Institute" (order No. 1571)