Fermilab

Sep 17, 2026

XRootD Redirectors at Fermilab: Architecture and Operations

Lorena Lobato Pardavila
(in collaboration with Yujun Wu)

U.S. DEPARTMENT of ENERGY
Fermi National Accelerator Laboratory is managed by FermiForward for the U.S. Department of Energy Office of Science

FERMI NATIONAL ACCELERATOR LABORATORY

What I will cover today

Agenda

Introduction

Why XRootD Redirectors Matter?

Services Overview

XRootD Redirector Services at Fermilab

Operations & Monitoring

How we operate and monitor the services?

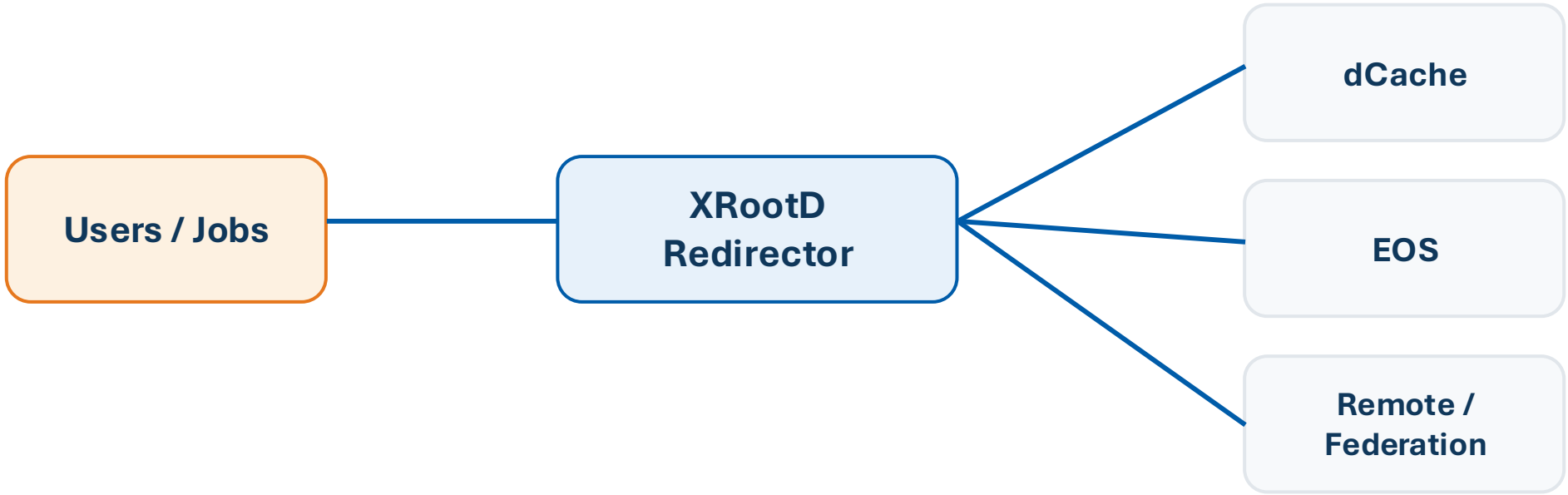
What's next?

XRootD as part of Storage Evolution

Takeaways

Why XRootD redirectors matter?

Redirectors hide data location from the client



The value is simple: one access point, many data sources

Key distinction

Redirector = control path: locate the file and redirect the client. Data then flows directly between client and data server.

Fermilab runs three redirector roles across two federation models

CMS AAA (Any Data, Anytime, Anywhere) : 1 regional + 3 sites

- Global → Regional → Site hierarchy
- Regional data access for CMS
- FNAL site redirectors front local storage
- No caching at the redirector

StashRedirector :Fermilab origin access

- Part of the OSG Data Federation
- Exposes Fermilab-hosted data through OSDF, while remote sites operate the caches.
- Acts as an origin/redirector
- Does not operate a local cache

What changes operationally

- CMS services run on hosts
- StashRedirector runs in OKD
- Different monitoring and failure domains

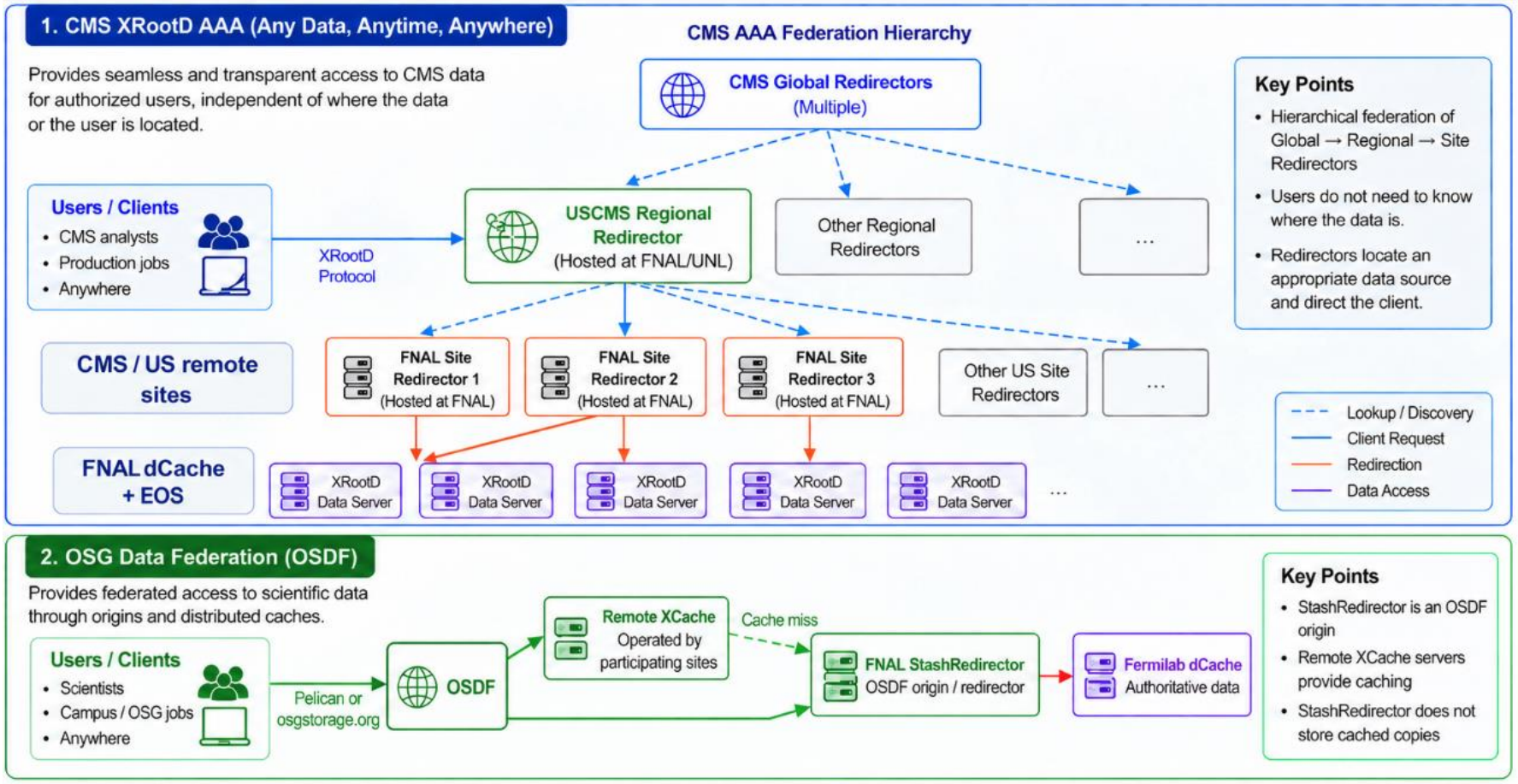
Same technology. Different communities, workflows, and operational needs

At Fermilab, redirectors sit between clients and storage services

Key Questions

Where is the data?

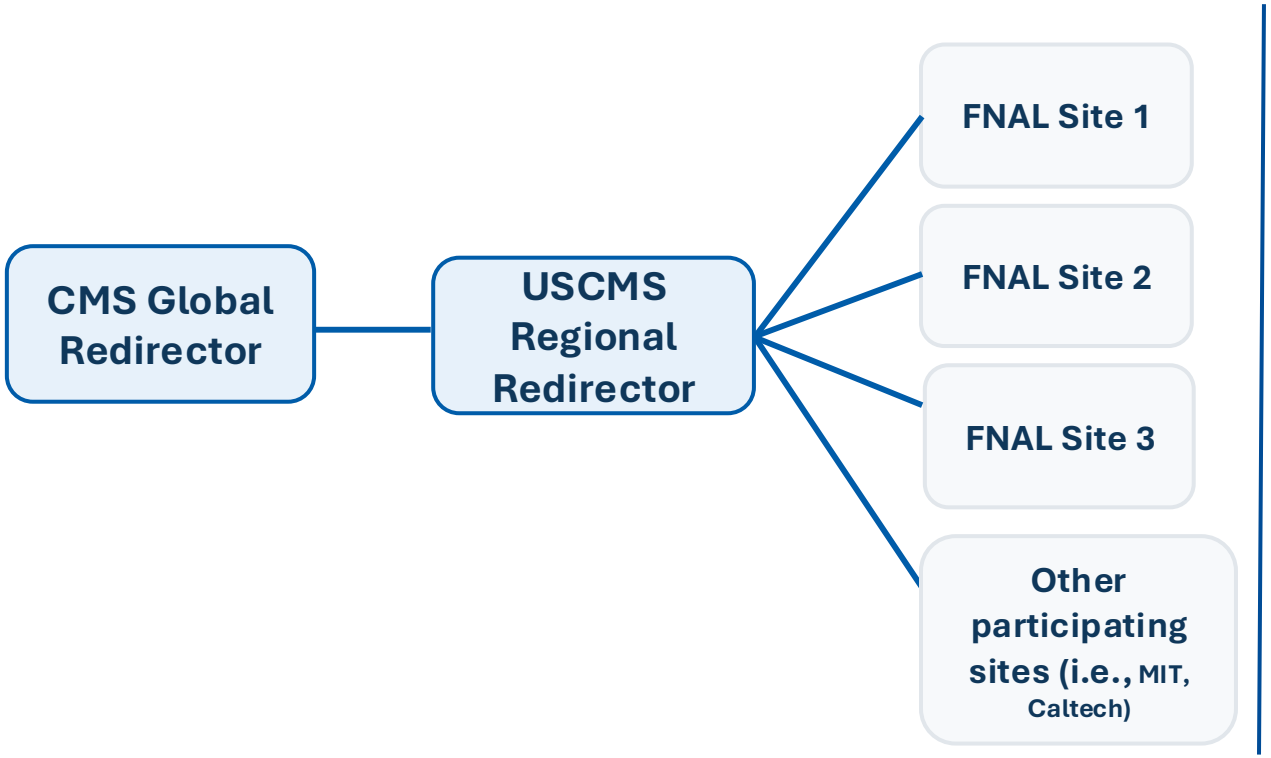
Is there a nearby cached copy?

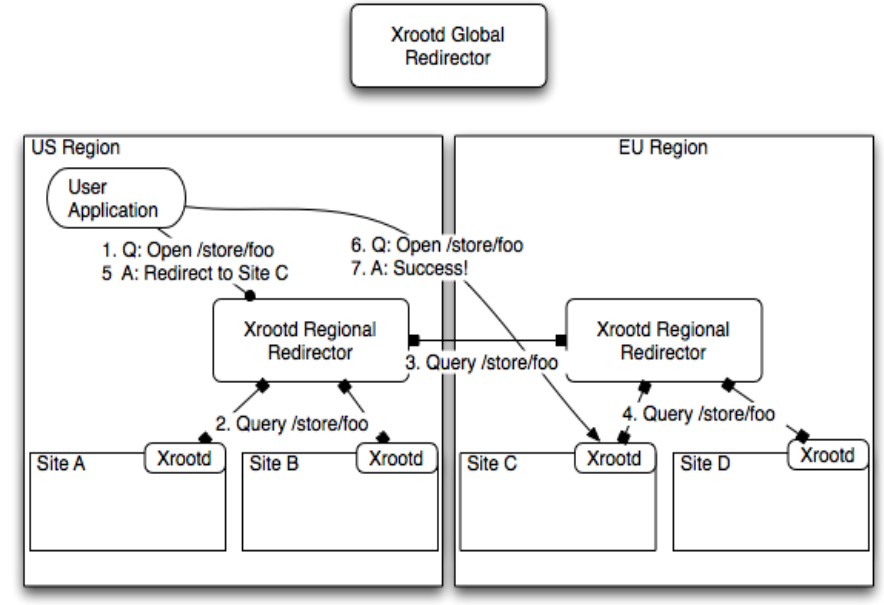


Operating principle Locate the data and, in OSDF, reuse a nearby cached copy when available

CMS AAA uses a global → regional → site hierarchy

Fermilab participates at both the regional and site levels



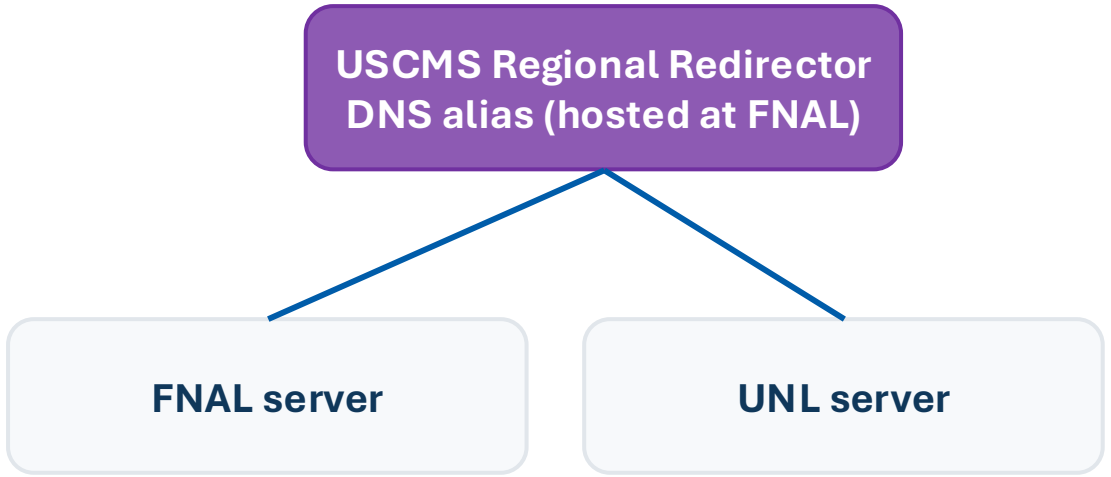


Reference:Illustration of the CMS XRootD AAA architecture for cross-region redirector (*Courtesy of CMS Data Operations)*

- Clients contact a regional redirector (or the global redirector), and AAA handles the rest.
- Fermilab hosts one of the two USCMS regional redirector servers; the other is at the University of Nebraska–Lincoln (UNL).
- 17 USCMS site redirectors are connected to both regional redirectors. 3 XRootD site redirectors are hosted at FNAL.

The USCMS regional redirector is redundant at two levels

Two servers at FNAL + UNL provide geographic redundancy



Why this matters operationally

A regional redirector failure should not become a regional data-access outage.

Fermilab regional host

- AMD EPYC 7543
- 96 GB RAM
- AlmaLinux 9.8
- XRootD 5.9.1-1.3

The Redirector is not a single server. It is a DNS alias backed by two servers, using round-robin DNS to distribute requests across the two servers and provides geographic redundancy

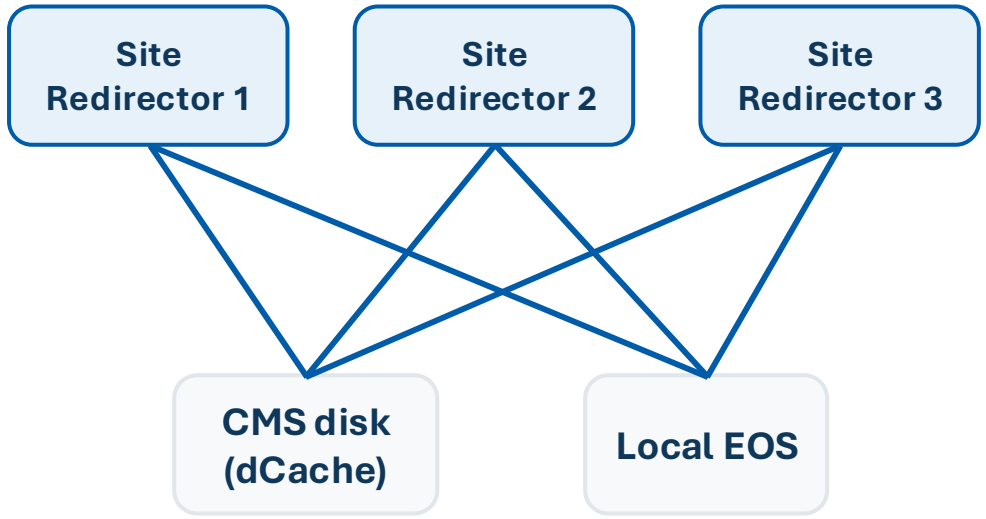
FERMI NATIONAL ACCELERATOR LABORATORY

FNAL site redirectors unify access to dCache and EOS

Three site redirectors front Fermilab storage for CMS and local workflows

User Experience

The storage backend can change without changing the client's access model (XRootD interface)



Fermilab site redirector hosts (3)

- AMD EPYC 7543
- 96 GB RAM
- AlmaLinux 9.8
- XRootD 5.9.1-1.3

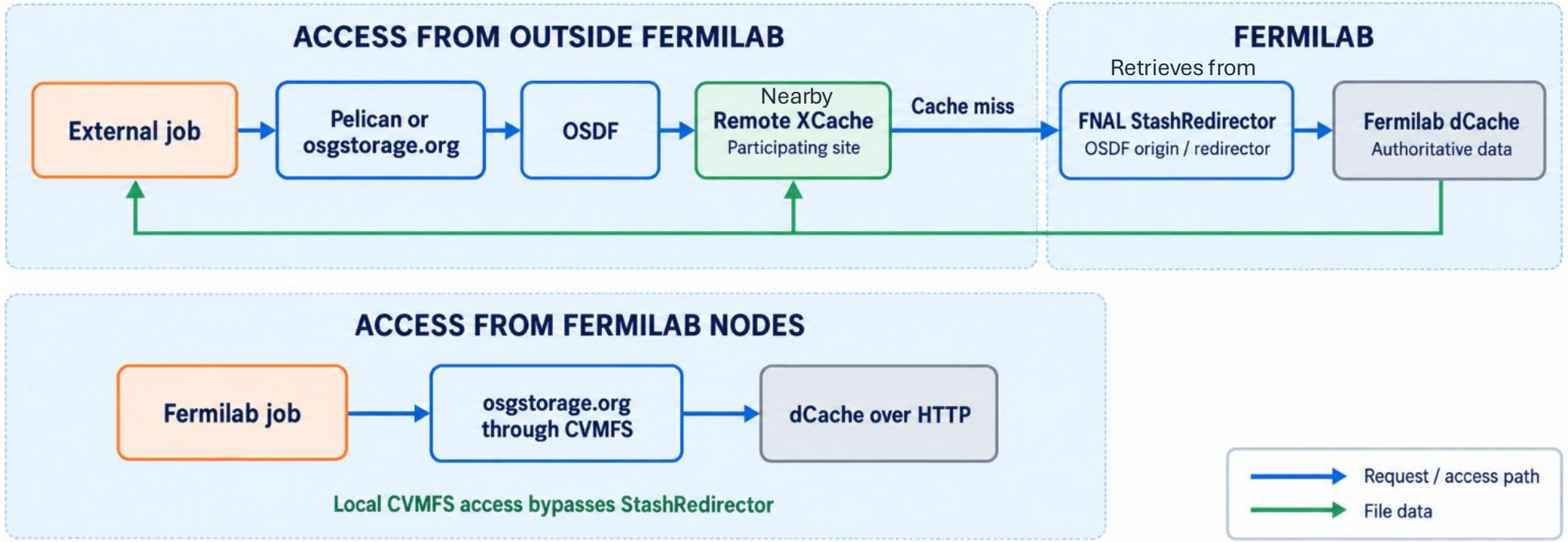
The Redirector is not a single server. It is a DNS alias backed by 3 servers, using round-robin DNS to provide load balancing and high availability.

Important Distinction

The site redirector handles local discovery, while the CMS framework controls the fallback.

StashRedirector connects the OSG Data Federation to Fermilab storage

External OSDF access uses the federation. Our local CVMFS access bypasses it



StashRedirector is containerized and managed as an OKD service

The Fermilab redirector runs as a single containerized service

Resources

- 1 pod
- 4 CPU
- 2 GB RAM

Processes managed by supervisord with automatic restart.

Configuration

GitLab
Deployment configuration
Documentation

Image supply

Harbor
Private container registry

Endpoints

stashredirector.fnal.gov
osdforigin.fnal.gov (alias)

Monitoring has to answer two questions: “Is it up?” and “Can users use it?”

Availability alone is not enough for a federated data-access service

CMS redirectors

CHECK_MK

- Host/service health
- Process availability

CMS SAM

- Site redirector tests
- Listening / read / authenticated write
- Federation/subscription health

StashRedirector

OSG monitoring / Check_MK

- Service health
- osdforigin visibility

Logs + supervisorctl

- XRootD status
- CMSD status
- Process failures

Operational question

If a user cannot access a file, what layer has failed?

1. **Access service:** CMS redirector or StashRedirector origin
2. **Platform or network:** Host, OKD, DNS, connectivity
3. **Storage backend:** dCache or EOS
4. **Federation or authentication** CMS AAA or OSDF registration and authorization

Monitoring should narrow the search.

An alert is only the first step in the investigation. Good monitoring should help us narrow down which layer needs attention.

The hardest incidents are often below the redirector

A healthy redirector does not guarantee a healthy data-access path

A user-visible XRootD error can originate at any layer.

Observed lesson

Slow client connections can create heavy redirector load.

More common root cause

Many reported access problems have been associated with underlying storage.

Operational response

Correlate logs, monitoring, connectivity, and storage health before restarting anything.

We always begin with the same question: "Which layer failed?" But the tools we use next depend on the deployment.

The deployment model determines the operational playbook

Host-based CMS services and containerized StashRedirector have different failure domains

CMS redirectors — host based

Monitor
- Checkmk + CMS SAM

Investigate
- XRootD logs
- Host / connectivity
- dCache / EOS health

Operate
- Configuration management with Puppet
- systemctl for service control

Resilience
- Regional redundancy at FNAL + UNL

StashRedirector — containerized

Monitor
- OSG / Check_MK

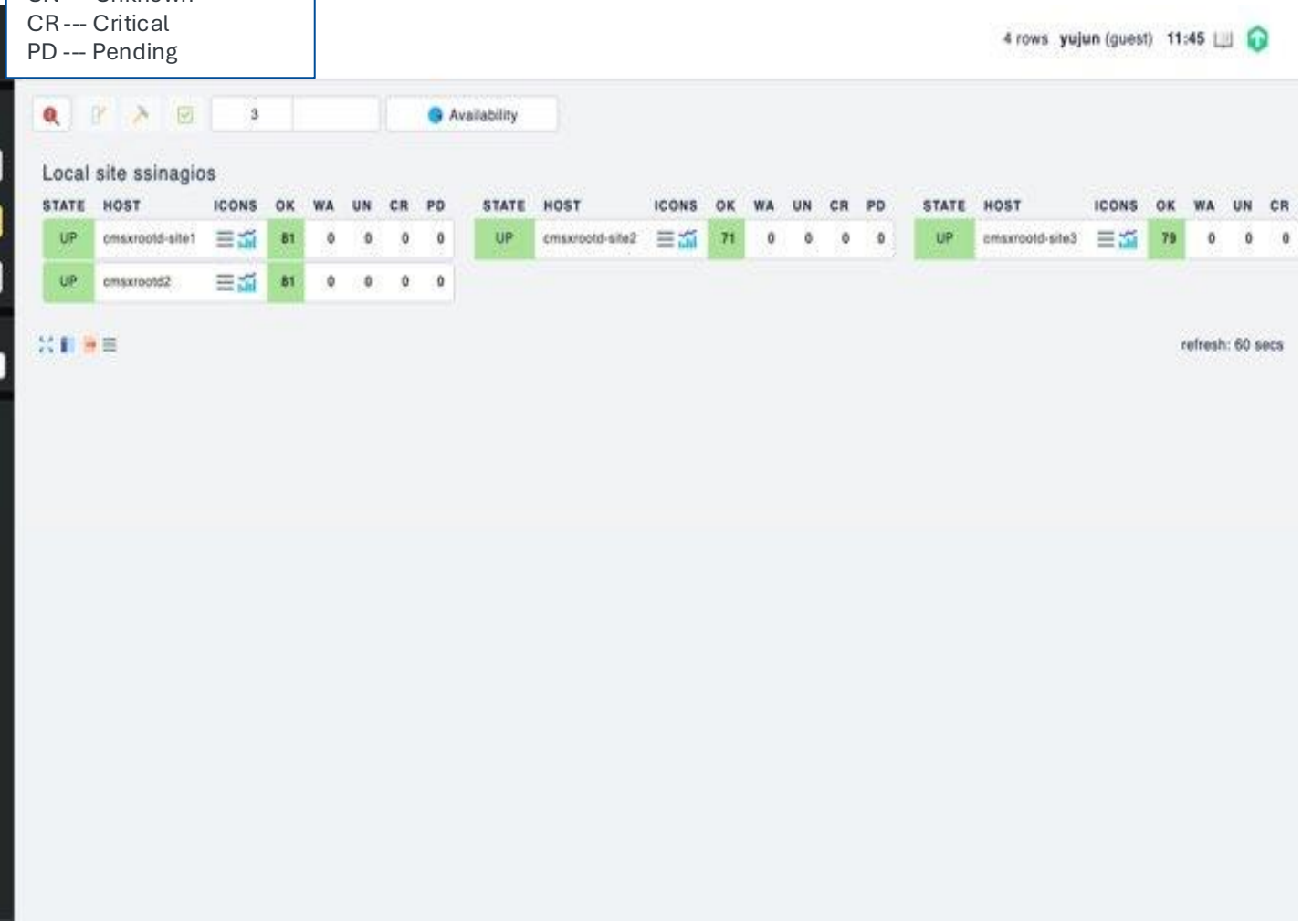
Investigate
- XRootD + CMSD logs
- supervisorctl
- Pod / OKD / network / dCache backend

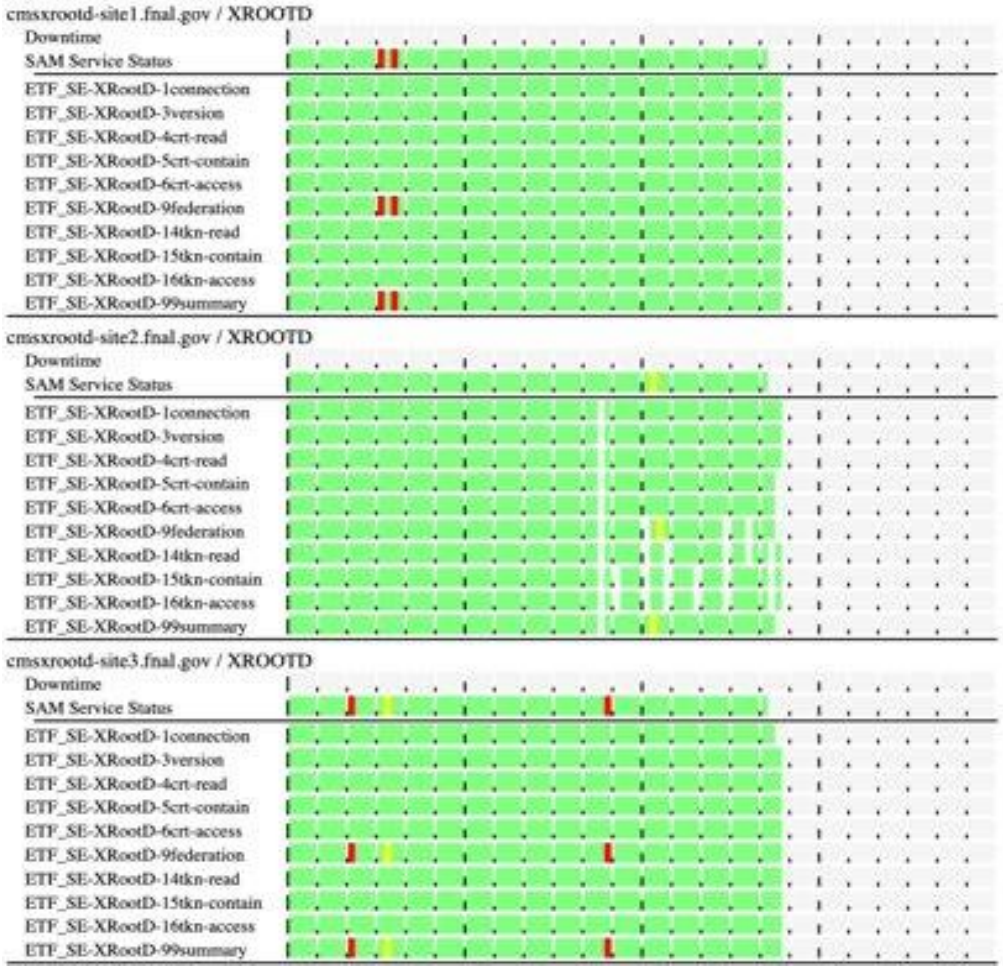
Operate
- GitLab + OKD deployment
- supervisord autorestart

Dependency
- Federation connectivity + FNAL storage

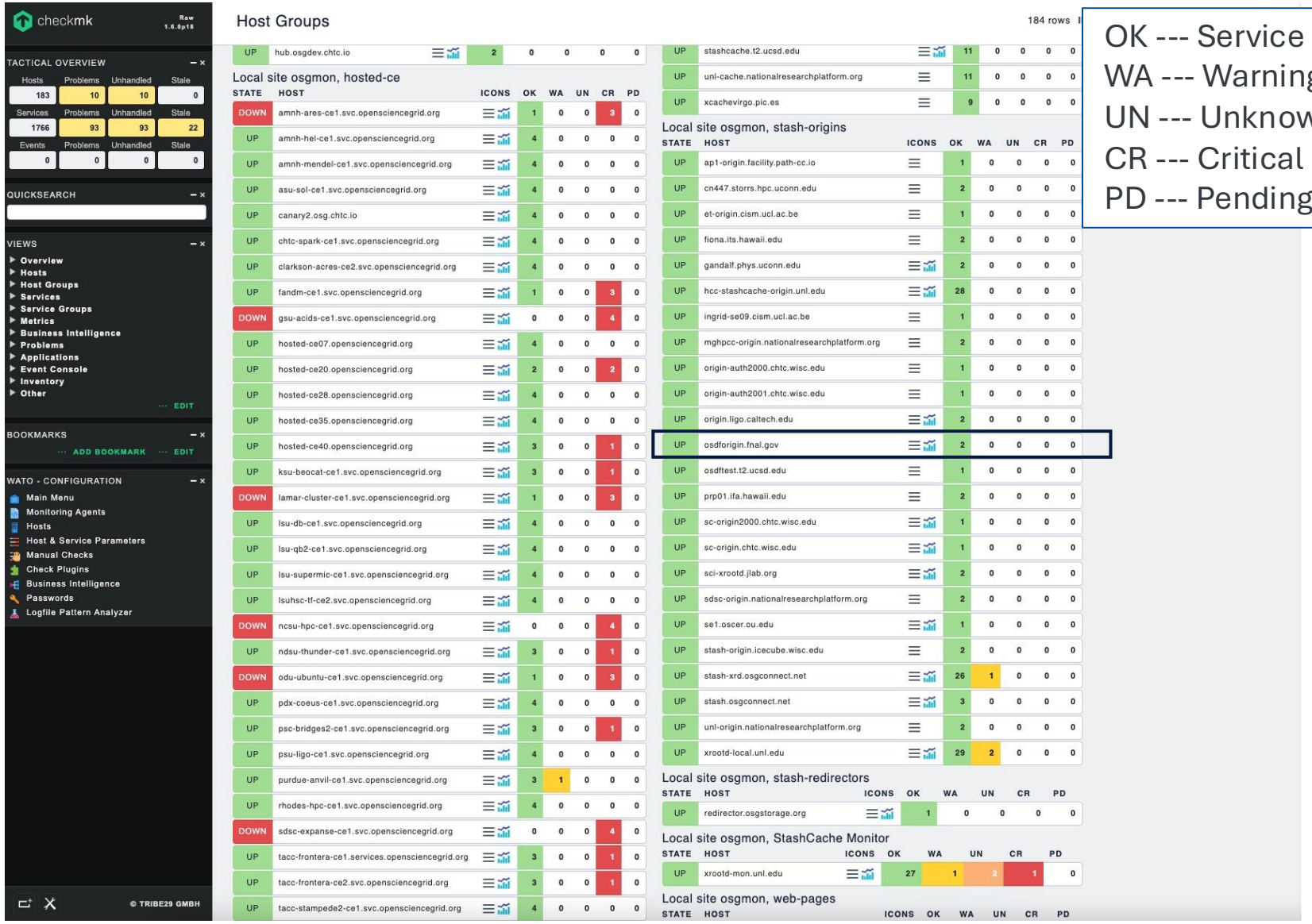
CMS XROOTD AAA MONITORING

OK --- Service is healthy
WA --- Warning
UN --- Unknown
CR --- Critical
PD --- Pending





STASHREDIRECTOR MONITORING



OK --- Service is healthy
WA --- Warning
UN --- Unknown
CR --- Critical
PD --- Pending

XRootD is part of Fermilab's storage evolution, not just today's operations

The same protocol is being evaluated across future storage architecture

Why it matters

Fermilab's Disk Storage Evolution Project is evaluating future HEP storage architectures for the HL-LHC era.

CMS and DUNE workflows rely on the XRootD protocol for data access.

Proxy test

An XRootD proxy configuration for EOS behind a firewall. Initial tests showed performance comparable to direct access.

CephFS integration

An XRootD server on CephFS enables XRootD protocol access to a distributed filesystem, opening another architecture option.

The goal is simple: the storage behind the service may change, but users can continue using the same access method.

3 takeaways from operating XRootD redirectors at Fermilab

Architecture, operations, and observability are inseparable

Different Roles

- CMS AAA uses hierarchical redirection
- StashRedirector connects the OSG federation to Fermilab storage . Remote XCache servers provide caching closer to jobs
- Same protocol, different operational behavior.

Availability is end-to-end

- A redirector can be healthy while the user still cannot access data.
- Monitor the complete path to the redirector, network, federation, and storage.

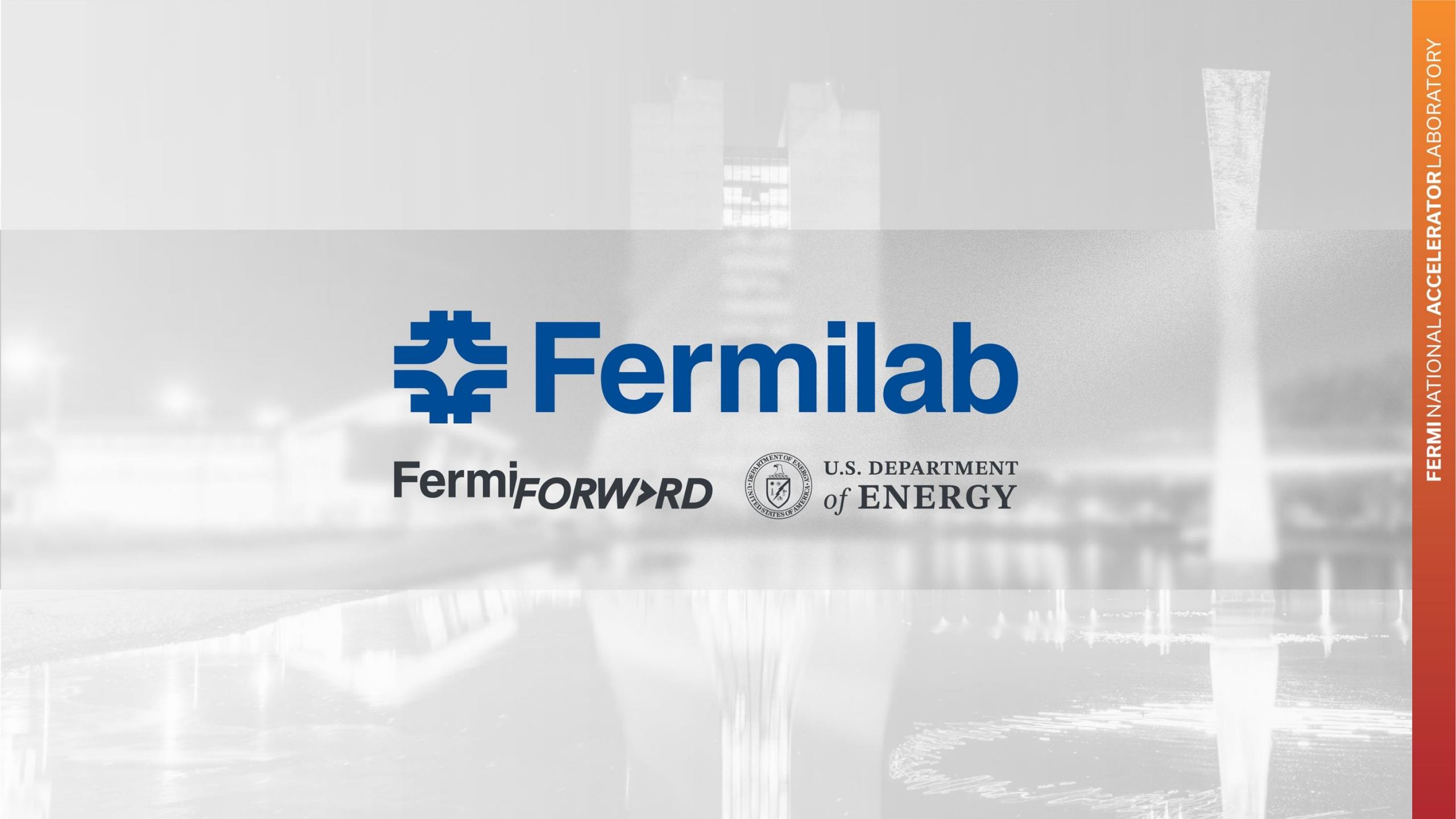
Keep evolving the operations

Automation, observability, and clear failure isolation become more important as storage architectures change.

Reliable distributed data access depends on the complete path

Thank you!

Q&A


Fermilab
FermiFORWARD
U.S. DEPARTMENT of ENERGY
FERMI NATIONAL ACCELERATOR LABORATORY