

# Computing and Software Services II Database Services

Dmitry Arkhipkin

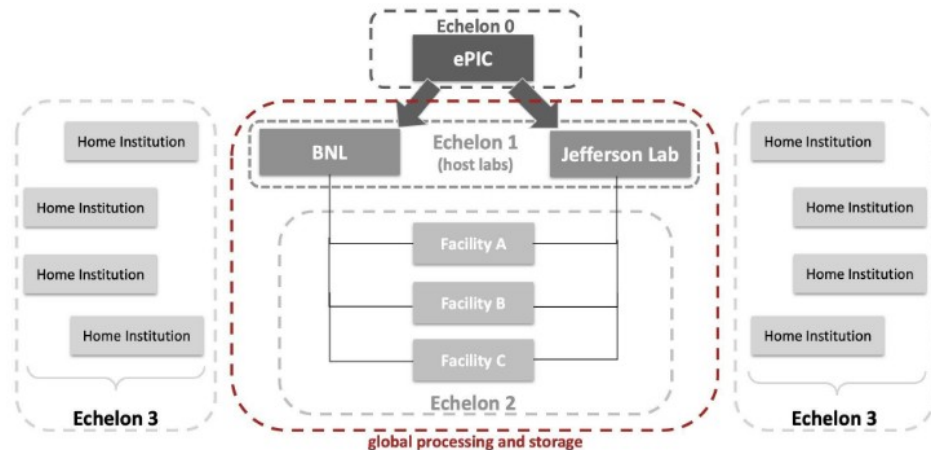
2025-12-05



@BrookhavenLab

# Overview

“...Provisioning and operating standard infrastructure solutions consistent with supported lab infrastructures and community best practices...”



Use Case	Echelon 0	Echelon 1
Streaming Data Storage and Monitoring	✓	✓
Alignment and Calibration		✓
Prompt Reconstruction		✓
First Full Reconstruction		✓
Reprocessing		✓
Simulation		✓
Physics Analysis		✓
AI Modeling and Digital Twin		✓

Echelon 0: ePIC experiment, DAQ system, **“ONLINE” DATABASES**, write-dominated

Echelon 1: two host labs, two primary ePIC computing facilities, **“OFFLINE” DATABASES**, read-dominated

Echelon 1s uniquely perform the low-latency streaming workflows consuming the data stream from Echelon 0

- Archiving, monitoring, prompt reconstruction, rapid diagnostics

# Database Services & Clients Overview



## “Expert”-level Databases:

- Detector Conditions DB
- DAQ Control DB
- Run Control DB
- Online QA (fast) DB
- “AI for the EIC” Dbs
- ...

Hardware nodes

## “User”-level Databases:

- Detector Calibrations DB
- Detector Geometry DB
- Offline QA (detailed) DB
- File Catalog DB
- Workload Management DB
- “AI for the EIC” Dbs
- ...

Virtual nodes (OpenShift?)

self-hosted DB Replicas  
...Or...  
access to Echelon 1 DBs  
via HTTP API

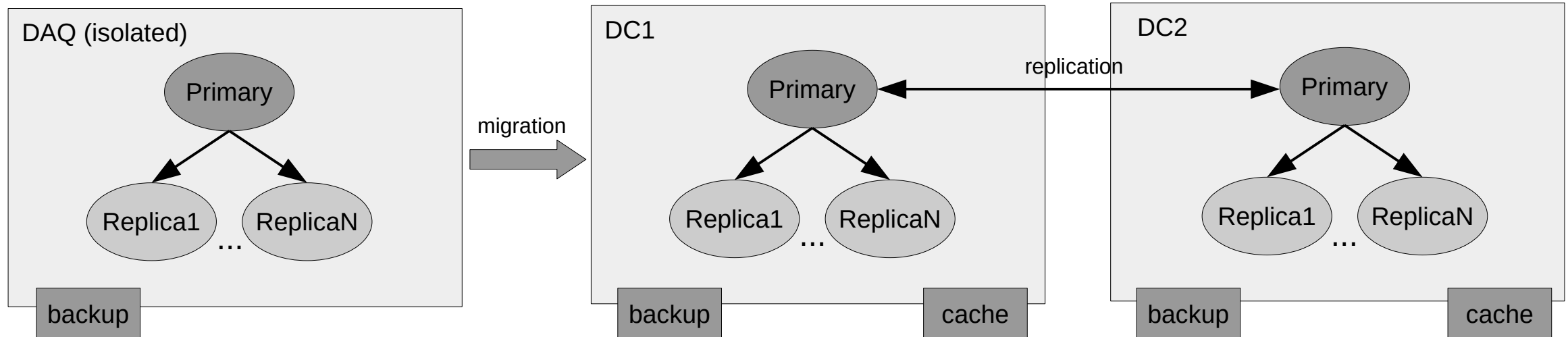


# Challenges

- **high-throughput, high-availability, disaster recovery / automatic failover, large size**
  - total data size: multi-TB online, <1TB offline – per year, +archives
  - clients: 1k online, 150k...500k offline
  - numerous solutions available, mostly based on replication / clustering with horizontal scaling, varying level of automation and recovery
- **DAQ enclave (isolated)**
  - “online” to “offline” data migration across network domain borders
- **two-zone multi DataCenter setup** (BNL, JLAB)
  - multi-master setup with HA, failover
  - symmetric infrastructure and support
  - local backups, db-aware monitoring

Database high-availability solutions must perform the following tasks:

- Elect a primary node as the leader.
- Direct all write operations to the primary node.
- Replicate all changes on the primary to all active replica/secondary nodes.
- Monitor the status of the primary node and identify any failures.
- In the event of a primary failure, promote one of the replicas to become the new primary.



# DB Planning FY2026

A detailed plan covering the following key points is required:

- **Automated on-demand database provisioning for ePIC / EIC**
  - evaluate currently available technologies for clustering, HA, and automatic provisioning
    - focus on multi-DC multi-client deployment
    - minimize the costs of handling two distinct enclaves, two different deployment setups
    - plan for known and emerging services both Online and Offline
  - plan for the unified and automated db disaster recovery, backups, db-specific monitoring
  - enable centralized access control and security audit, including hot-patching of vulnerabilities
- **Estimation of the storage and network capacity**
  - reliable multi-TB data storage for two distinct domains of operation
  - network throughput used by data replication streams and migrations
- **Multi Data Center coordination**
  - evaluate the possibility of having a fully symmetric database setup at BNL and Jlab
    - asymmetric setup based on asynchronous streaming to read-only replicas is already tested
  - develop procedures to coordinate the DB management efforts and maintenance across two Labs
    - Including monitoring and timely response to incidents 24/7