Belle II Distributed Data Management System

MALACHI SCHRAM
ON BEHALF OF THE BELLE II DDMS
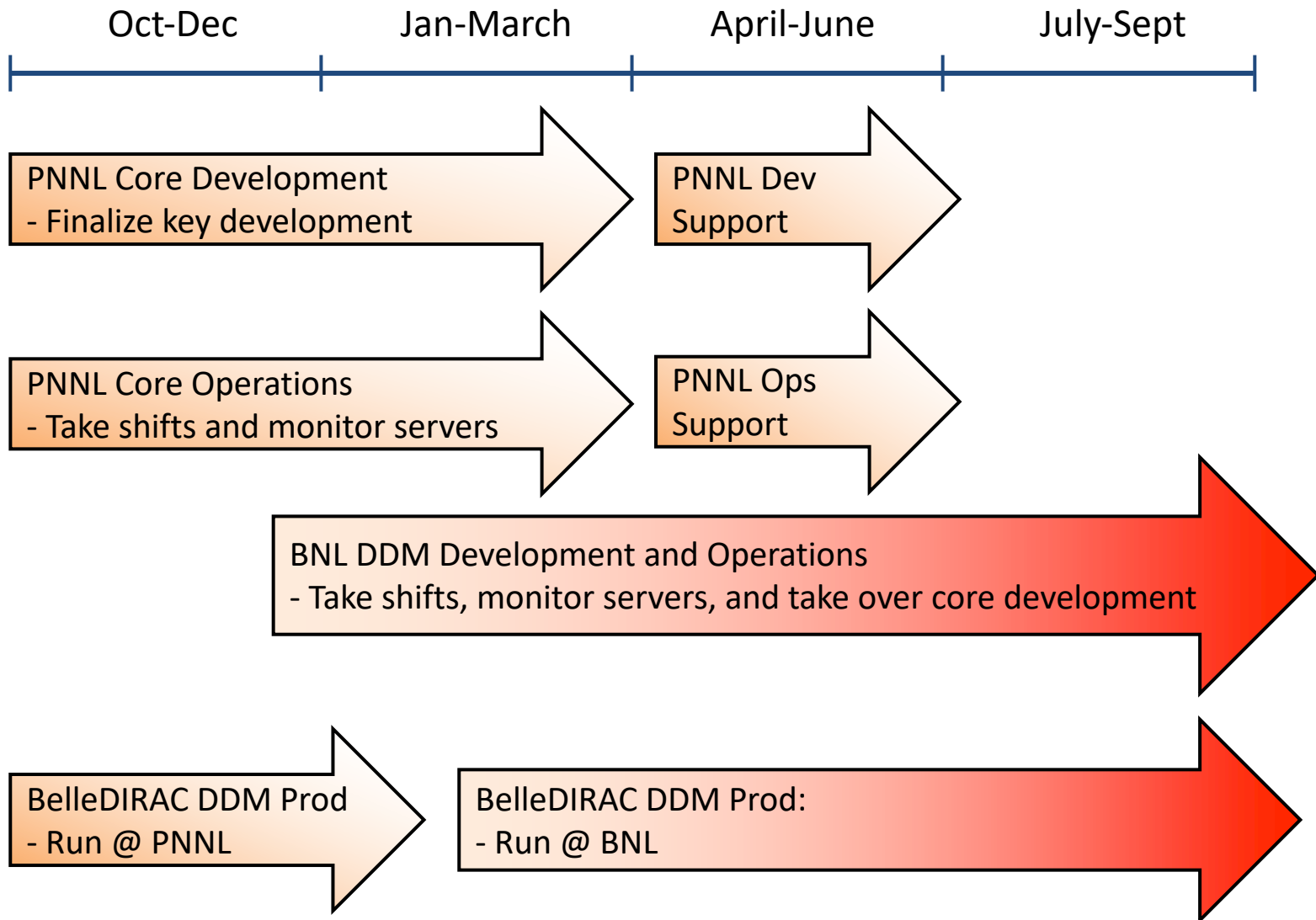
U.S. contributions to Belle II DDMS

▶ **<u>DOE-SC HEP decided to move the U.S Belle II computing/software effort from PNNL to BNL</u>**

▶ **<u>Significant delays in development</u>**
- ■ Focused on site transition (CPU, storage, services, etc.)
- ■ Premature departure of key members

▶ **<u>Personnel</u>**
- ■ PNNL:
  - ● Malachi Schram: contribution ends at the end of June
  - ● Kevin Fox: contribution ends at the end of June
  - ● Antonio Ledesma: contribution ends at the end of June
  - ● Vikas Bansal: resigned (last contribution was in November 2017)

- ■ BNL:
  - ● Alex Undrus (PAS)
  - ● Siarhei Padolski (PAS)
  - ● Hironori Ito (RACF)

Timeline for U.S. FY2018

| Oct-Dec | Jan-March | April-June | July-Sept |

PNNL Core Development
- Finalize key development

PNNL Dev Support

PNNL Core Operations
- Take shifts and monitor servers

PNNL Ops Support

BNL DDM Development and Operations
- Take shifts, monitor servers, and take over core development

BelleDIRAC DDM Prod
- Run @ PNNL

BelleDIRAC DDM Prod:
- Run @ BNL

DDMS development & operation short term (CY2017)

▶ **Development with DDM v0.2**

- ■ Minor feature and bug fixes within the existing DDMS framework
- ■ Initial refactoring of the deletion component
  - ● Currently the deletion component is the limiting factor for the DDMS
- ■ Update the raw data transfer agent
- ■ Scaling tests for new components
  - ● *FanoutAgent* and Deletion components

▶ **Operations**

- ■ Expert shifts for MC production
- ■ Migrate Belle II samples from PNNL to BNL
- ■ RAW data transfer workflow between KEK to BNL
- ■ Prepare for the full dress rehearsal

DDMS development & operation mid-term (CY2018 until June)

► **Development with DDM v0.2**

- Minor feature and bug fixes within the existing DDMS framework
- Develop the ReplicationPolicy components (service/agent/db)
- Develop the ReplicaAndPopularity components (service/agent/db)
- Multithread agents (focus on limiting factors)
- Generalize DDM to handle different SE plugins
- Update/re-implement end-user command tools to use DDMS
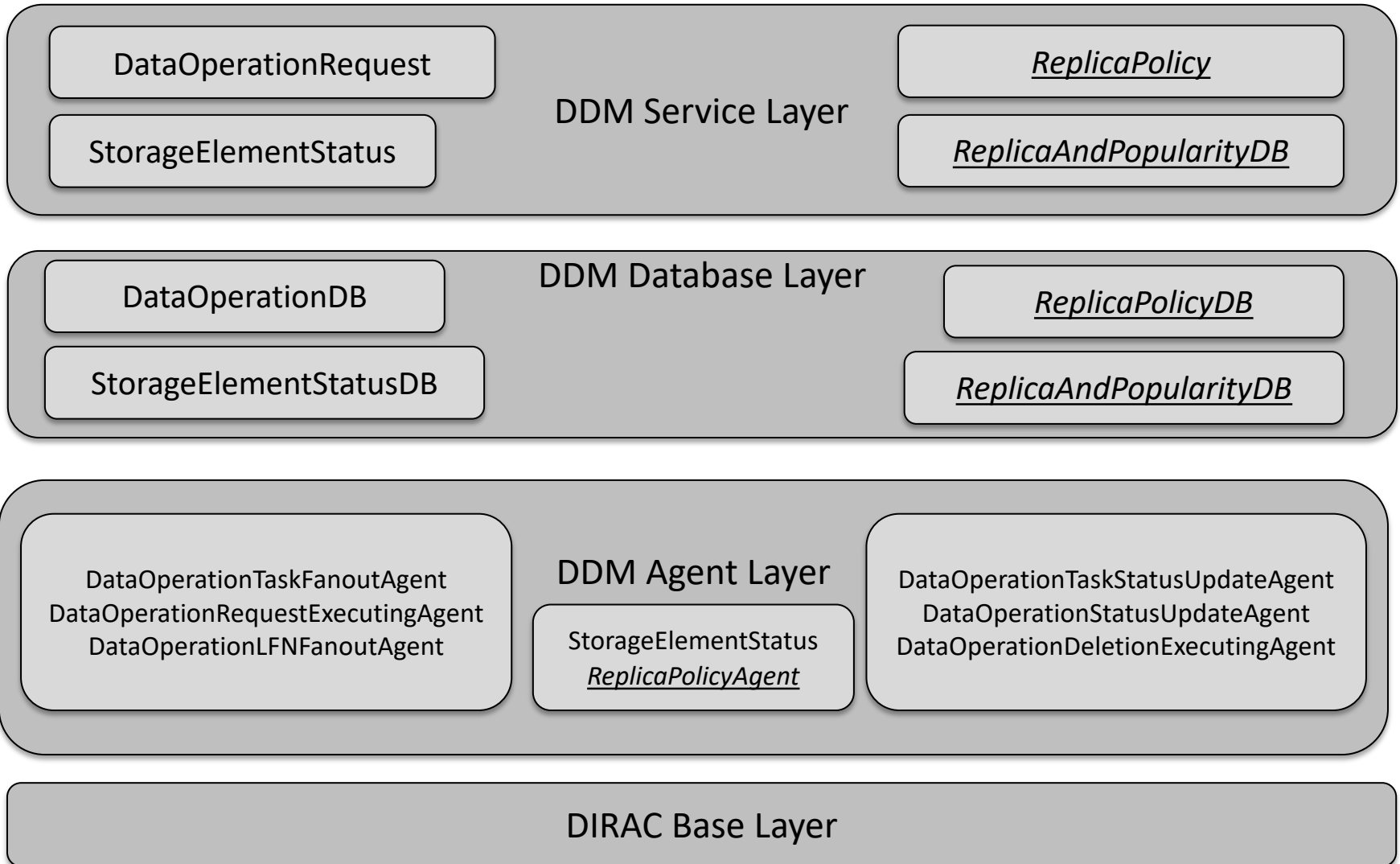- Recursive datablock transfers

► **Development for DDM v0.3**

- Make the deletion component it's own separate subsystem (like RMS)
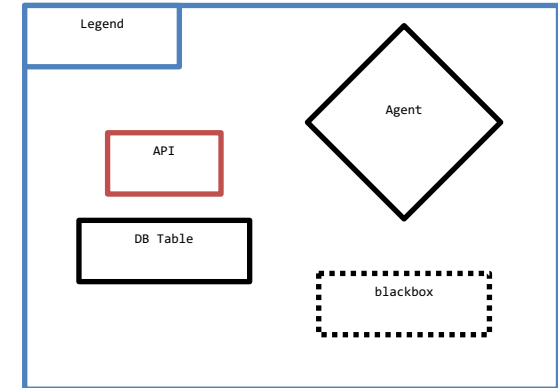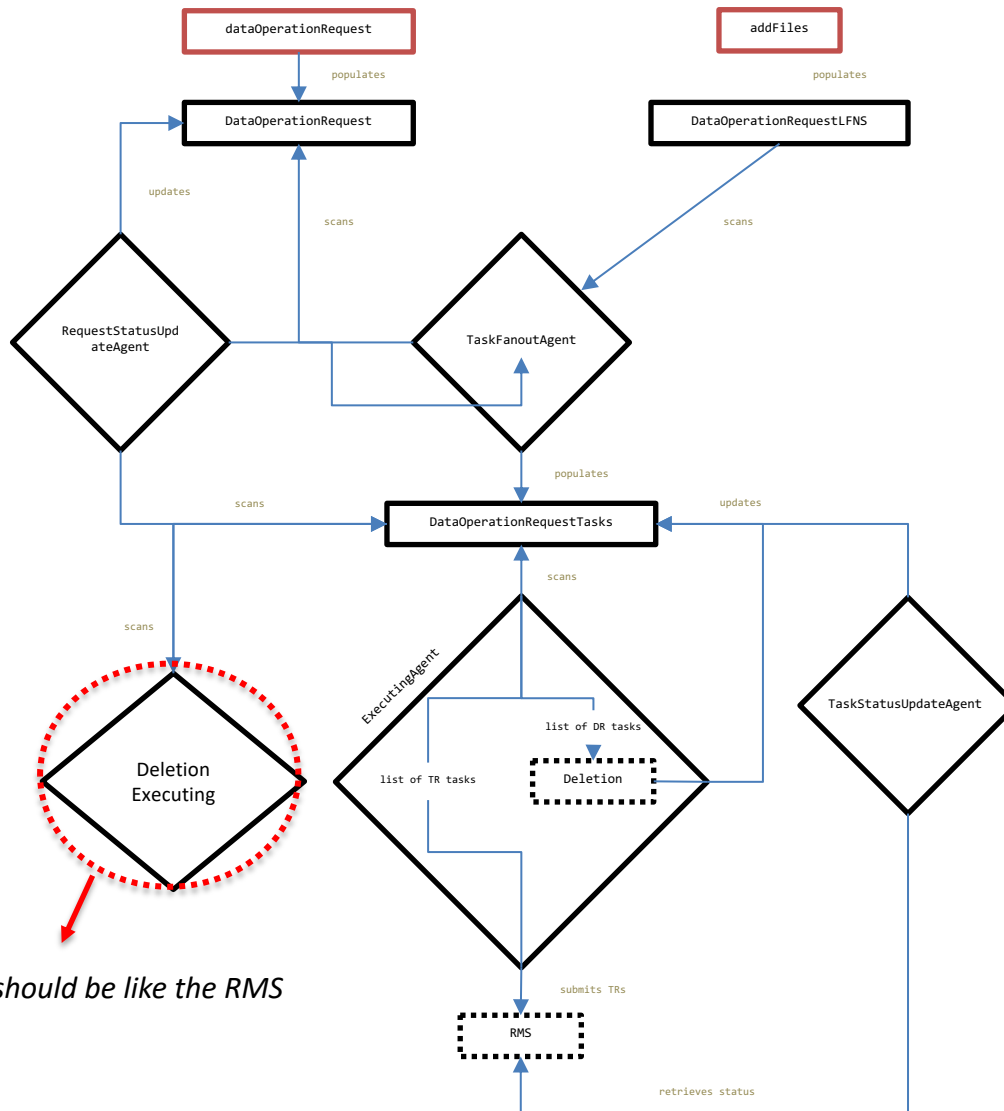- Integrate the rebased DIRAC FTS system into the DDMS

► **Operations (CY2018)**

- Expert shifts
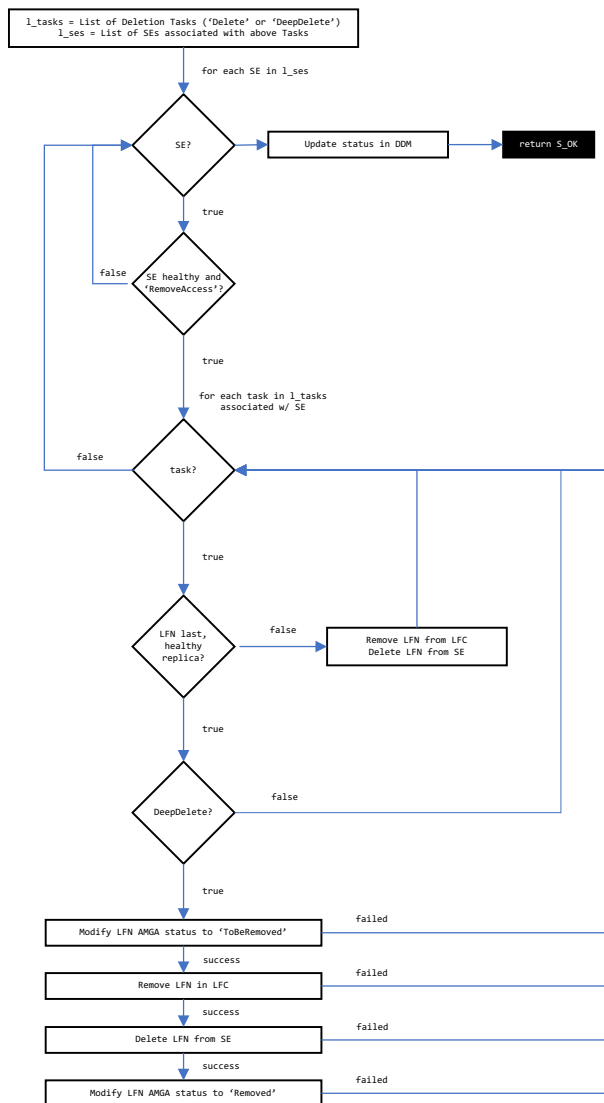- Prepare for early data taking

DDMS (version 0.2)

DataOperationRequest

ReplicaPolicy

DDM Service Layer

StorageElementStatus

ReplicaAndPopularityDB

DDM Database Layer

DataOperationDB

ReplicaPolicyDB

StorageElementStatusDB

ReplicaAndPopularityDB

DataOperationTaskFanoutAgent
DataOperationRequestExecutingAgent
DataOperationLFNFanoutAgent

DDM Agent Layer

StorageElementStatus
ReplicaPolicyAgent

DataOperationTaskStatusUpdateAgent
DataOperationStatusUpdateAgent
DataOperationDeletionExecutingAgent

DIRAC Base Layer

Data operation workflow
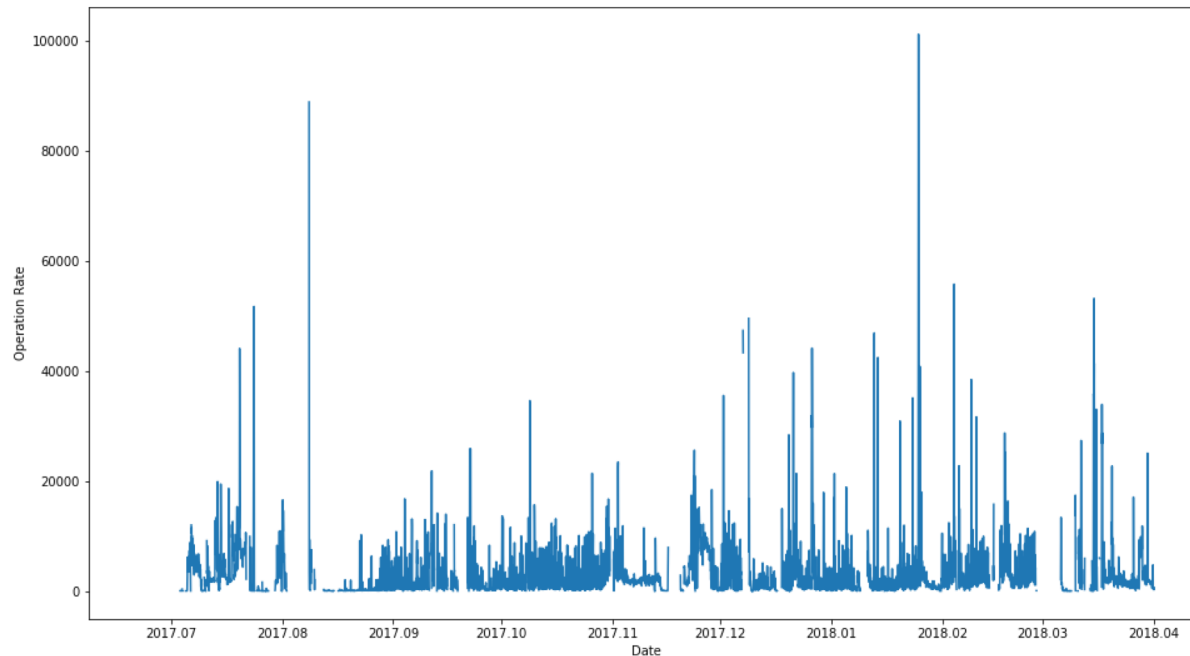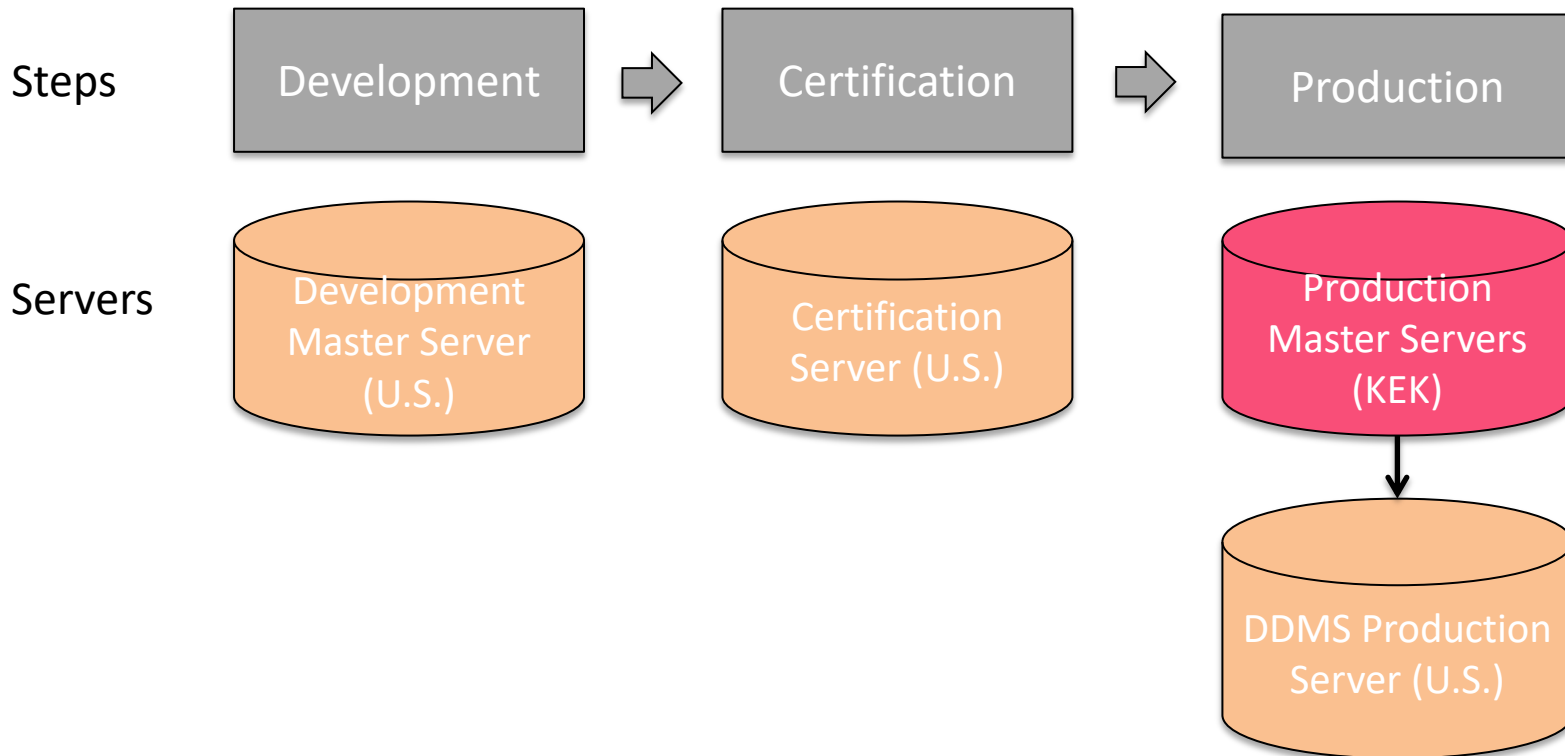
Deletion sub-system workflow

DDMS v0.2 by the numbers

▶ **New DDM v0.2 was put into production at PNNL from July 2017 to April 2018**

▶ **Total data operations**

  ■ 19.8M finished data operations

  ■ Maximum deletion operation rate in an hour was over 12k

  ■ Maximum replication operation rate in an hour was over 50k

  ■ Failure rate is 0.57%

Software release cycle

► DDM and the rest of the BelleDIRAC effort uses the following development release cycle

Steps

Development → Certification → Production

Servers

Development Master Server (U.S.)

Certification Server (U.S.)

Production Master Servers (KEK)

DDMS Production Server (U.S.)

Summary

▶ Ongoing work:
  - ◼ Finalize deletion and policy replication sub-systems
  - ◼ Enhanced monitoring
  - ◼ General bug fixes and updates

▶ Using the new FTS rebase would still be very nice ☺

▶ Plans for DDMS development at BNL starting this Fall:
  - ◼ Monitoring enhancements
    - ● Operational issues analysis
    - ● Improved data flow control
    - ● Alarm systems
  - ◼ Replica and popularity subsystems
    - ● Make optimum number of replicas at locations close to users that need them
    - ● Predictive usage intensity tools (save on SE space)
  - ◼ End-user tools improvements