The data management of heterogeneous resources in Belle II

MALACHI SCHRAM

On the behalf of the Belle II Distributed Computing Group

CHEP2018: Sofia, Bulgaria
Belle II Computing

- Belle II is an international experimental collaboration of over 800 members at 108 institutions in 25 countries. PNNL is one of the main contributors.
- Belle II detects the particles resulting from collisions of electrons and positrons accelerated to high energy.
- Expected data rates from the Belle II experiment are high:
  - At peak luminosity we expect 11 PetaBytes per year, 100 PetaBytes raw data volume by 2024, total data volume 200 PB
  - Processed data samples will be distributed worldwide
DIRAC provides interoperability across a number of heterogeneous resource providers:

- **Computing interfaces:**
  - Clusters: SLURM, HTCondor, Torque, etc.
  - Clouds: AWS, CERN, Private, etc.
  - HPC: NERSC, Constance, etc.

- **Storage interfaces:**
  - HTTPS, SRM, XROOTD
  - Data transfers: GLOBUS, FTS, etc.

- **Integrated File Metadata**
BelleDIRAC is a DIRAC extension that provides dedicated Belle II services

- Production Management System
- Fabrication System
- Distributed Data Management System
- Specialized user client tools
- Dedicated monitoring tools
Belle II Distributed Data Management System

- Provides a single point to make any data operation request.
- US leads key contributions to the design, deployment, and operations of the computing effort, including DIRAC extensions for the Distributed Data Management system.
- In the past nine months, the latest version of the DDM achieved:
  - 19.8 M data operations
  - >12 k/hr deletion operations
  - >50 k/hr replication operations
All production data operations are submitted to the Belle II Data Operation service.

Fundamental data operations:
- Replicate
- Delete

Compounded data operations:
- Move: replicate A→B then delete A
- Migrate: replicate A,B,…N→ X then delete A,B,…N
The replication policy dynamically replicates and/or removes files based on Belle II data placement policies.

A policy is defined by:
- Base Path: Corresponds to a group of datasets
- Data Type: Data or Monte Carlo
- Data Level: RAW or processed level
- Number of replicas

Based on defined policy the proper Storage Elements are used for replication.

To be deployed this summer
Belle II Distributed Computing Summary

Since January 2018:
- Average 17.3k concurrent jobs
- Average 18.0k successful transfers per hour
- Current max 24.9k/hr concurrent job
- Current max 31.6k/hr successful transfers (including non-DDM transfers)
Thank you