Data Management in the Belle-II experiment

Cédric Serfon
_Brookhaven National Laboratory_
for the Belle II DDM team
Belle II is a B-physics experiment located at KEK (Japan)

Start of data taking last year
Belle II Distributed Computing model

- The Belle II computing model is based of a geographically distributed environment which aim at accomplishing several tasks:
  - RAW data processing and reprocessing
  - Monte Carlo Production
  - Physics analysis
  - Data Storage and Data Archiving
Disk Storage estimation

Storage resource estimation including disk for RAW Data. Storage for MC production and analysis, and storage for miniDST and uDST data will be shared among the different countries according to the PhD count.
RAW Data distribution

The second copy of RAW Data is currently stored at BNL. From 2021 the second copy of RAW will be distributed in different countries: USA, Italy, Germany, France and Canada.

<table>
<thead>
<tr>
<th>SITE</th>
<th>2019-2020</th>
<th>2021-2024</th>
</tr>
</thead>
<tbody>
<tr>
<td>BNL - USA</td>
<td>100%</td>
<td>30%</td>
</tr>
<tr>
<td>CNAF - Italy</td>
<td>0%</td>
<td>20%</td>
</tr>
<tr>
<td>DESY - Germany</td>
<td>0%</td>
<td>10%</td>
</tr>
<tr>
<td>KIT - Germany</td>
<td>0%</td>
<td>10%</td>
</tr>
<tr>
<td>IN2P3CC - France</td>
<td>0%</td>
<td>15%</td>
</tr>
<tr>
<td>UVIC - Canada</td>
<td>0%</td>
<td>15%</td>
</tr>
</tbody>
</table>
Raw Data Cycle
Current Infrastructure

- 15.4 PB Disk space currently managed
- More than 20 Storage Elements distributed all over the world
- Current infrastructure can work without the support of SRM
  - gsiftp and HTTP/WebDAV are election protocols.
  - Some gsiftp only end-point already registered
- Accounting: Mix of SRM based accounting and JSON accounting
Services used for the Data Management

- Belle II uses DIRAC for the Workload and Data Management:
  - Extension called BelleDirac was done to fit Belle II’s needs (more details in the next slide)
- 2 external services are used:
  - File catalog based on the LCG File Catalog (LFC): One LFC instance at KEK
  - FTS for file transfers: 1 FTS instance at KEK and 1 at BNL
**DIRAC Data Management**

- Current Distributed Data Management (DDM) is part of this BelleDirac:
  - Original design by PNNL group respecting Dirac paradigms, good for Belle II customisation but all development effort must come from Belle II
  - Looking ahead we saw lots of development work, why not use Rucio instead
Moving to Rucio

- Work ongoing to move DDM to Rucio
- Lots of new features developed to fit Belle II’s need:
  - Change the current DDM API to use Rucio: i.e. the API methods names do not change but Rucio is used behind. This allows the other services interacting with DDM not to change anything.
  - Rucio File Catalog plugin in BelleDirac (will eventually be merged in Vanilla Dirac)
  - Chained subscriptions (for RAW data export)
  - New lightweight daemon in Rucio to submit to external services (InfluxDB, ActiveMQ, ElasticSearch)
  - New dashboards for transfers/deletion monitoring as well as accounting
Moving to Rucio

● Multiple tests conducted:
  ○ To validate the new features of Rucio and Rucio based DDM
  ○ Scaling tests
  ○ Migration tests of LFC content into Rucio

● All the tests were successful so far. Plan for the final migration to Rucio during winter shutdown (probably mid-January) under discussion

![Load on the migration machine during the LFC test import (80M files)](image)
Conclusion

● The Belle II experiment will have a significant increase of the data produced and stored in the coming years
● To accommodate this increase, changes in the computing model are being done (e.g. moving from 1 to 6 RAW Data Centers)
● An other big change is the migration to Rucio where we expect to leverage the experience of bigger collaboration like ATLAS
● Some of the development done like the integration a Rucio File Catalog plugin into Dirac can potentially be used by other communities
Acknowledgements

- Thanks a lot to the whole Belle II DC team in particular to: J. De Stefano, T. Hara, Y. Kato, H. Ito, P. Laycock, R. Mashinistov, H. Miyake, S. Pardi, I. Ueda, M. Villanueva
- Special thanks to S. Pardi for providing some of the materials of this talk