Belle II Computing and Networking Requirements

MALACHI SCHRAM
Pacific Northwest National Laboratory
LHCOPN-LHCONE Meeting University of Michigan (US)
Overview

► **Computing & Networking Requirements**

► **Belle II Networking Workshops:**
  - Pacific Network and Computing Requirements Workshop hosted by PNNL October 2012
  - European Networking Workshop hosted in Vienna – October 2013
  - General Networking Workshop during SC14 – November 2014

► **Data Challenges and VC setups:**
  - General
  - ANA-100
  - KEK-PNNL

► **Plan Moving Forward**
Belle II Collaboration

ATLAS, 38 countries, 177 institutes, ~3000 members
CMS: 42 countries, 182 institutes, 4500 members
ALICE: 36 countries, 131 institutes, 1200 members
LHCb: 16 countries, 67 institutes, 1060 members

23 countries/regions
97 institutes
577 colleagues

as of June 30, 2014

Asia: ~45%
N. America: ~15%
Europe: ~40%

Japan: 137
Korea: 34
Taiwan: 22
India: 20
China: 15
Australia: 18

Germany: 83
Italy: 59
Russia: 37
Slovenia: 14
Austria: 14
Poland: 11
Belle II Luminosity and Data Rate

SuperKEKB Commissioning starts in 2015

Physics run starts in 2017

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Event size</th>
<th>Rate @ Storage [kB]</th>
<th>Rate @ Storage [event/sec]</th>
<th>Rate @ Storage [MB/sec]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Belle II</td>
<td>300</td>
<td>6,000</td>
<td>1,800</td>
<td></td>
</tr>
<tr>
<td>ALICE (Pb-Pb)</td>
<td>50,000</td>
<td>100</td>
<td>4,000</td>
<td></td>
</tr>
<tr>
<td>ALICE (p-p)</td>
<td>2,000</td>
<td>100</td>
<td>200</td>
<td></td>
</tr>
<tr>
<td>ATLAS</td>
<td>1,500</td>
<td>600</td>
<td>700</td>
<td></td>
</tr>
<tr>
<td>CMS</td>
<td>1,500</td>
<td>150</td>
<td>225 (&lt;1000)</td>
<td></td>
</tr>
<tr>
<td>LHCb</td>
<td>55</td>
<td>4,500</td>
<td>250</td>
<td></td>
</tr>
</tbody>
</table>

(LHC experiments: as seen in 2011/2012 runs)
The Belle II Computing model will manage in a geographically distributed environment the following main tasks:

- RAW data processing.
- Monte Carlo Production
- Physics analysis
- Data Storage and Data Archiving

On going activities

- Resource Estimation
- Define strategy for analysis and data distribution
- Evaluating technologies
The Belle II Computing Sites are categorized as follow:

- **Raw data Center**: Stores the RAW Data and data processing and/or data reprocessing.

- **Regional Data Center**: Large data center that stores mDST and participates at the Monte Carlo production.

- **MC Production site**: Data Center that produces and stores Monte Carlo simulations, that included:
  - Grid Site
  - Cloud Site
  - Computing Cluster Site
Data Volume Estimation

Storage estimation based on:
- RAW data
- mDST after data taking
- mDST during data reprocessing
- mDST-Monte Carlo related the data
- mDST-Monte Carlo related data reprocessing

The current parameters for data estimation are
- Event Size x RAW data: 300Kb
- Event Size x mDST: 40Kb

<table>
<thead>
<tr>
<th></th>
<th>Year1</th>
<th>Year2</th>
<th>Year3</th>
<th>Year4</th>
<th>Year5</th>
<th>Year6</th>
<th>Year7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Event/year</td>
<td>1.2</td>
<td>3.1</td>
<td>29.6</td>
<td>70.7</td>
<td>87.8</td>
<td>89.3</td>
<td>93.5</td>
</tr>
<tr>
<td>Integrated</td>
<td>1.2</td>
<td>4.3</td>
<td>33.9</td>
<td>104.6</td>
<td>192.4</td>
<td>281.7</td>
<td>375.2</td>
</tr>
</tbody>
</table>

September 15-16, 2014 LHCOPN-LHCONEN Meeting University of Michigan (US)
Belle II Computing Resource Estimation

**Yearly integrated luminosity**

**CPU (kHEPSpec)**

**Disk space (PB)**

Tape (PB) for raw data

- 2° COPY
- KEK
- Challenge

**September 15-16, 2014**

LHCOPN-LHCONE Meeting University of Michigan (US)
Belle II Raw Data Distribution

Scenario 1
(copy from KEK)

KEK Data Center (100%)

North America
PNNL Data Center (30%)
Canada Data Center (10%)

Germany Data Center (20%)
Italy Data Center (20%)

Asia
India Data Center (10%)
Korea Data Center (10%)

Scenario 2
(2step copy, KEK → PNNL → Europe)

KEK Data Center (100%)

North America
PNNL Data Center (70% → 30%)
Canada Data Center (10%)

Germany Data Center (20%)
Italy Data Center (20%)

Asia
India Data Center (10%)
Korea Data Center (10%)
Belle II Raw Data Network Requirements

**Scenario 1**

**RAW Data - Outbound**

- KEK
- PNNL

- RAW Data – Inbound

- India
- Canada
- Italy
- Germany
- Korea
- PNNL

**Scenario 2**

**RAW Data - Outbound**

- KEK
- PNNL

- RAW Data – Inbound

- India
- Canada
- Italy
- Germany
- Korea
- PNNL

*September 15-16, 2014  LHCOPN-LHCOME Meeting University of Michigan (US) 10*
Belle II Total Data Network Requirements
Scenario 2

Total Data - Inbound
- Viet Nam
- Ukraine
- Turkey
- Thailand*
- Taiwan
- Spain*
- Saudi Arabia
- Russia
- Poland
- Mexico
- Malesia
- India
- Czech Rep.
- China
- Austria
- Canada
- Australia
- Slovenia
- Italy
- Germany
- Korea
- PNNL
- KEK

Total Data – Outbound
- Viet Nam
- Ukraine
- Turkey
- Thailand*
- Taiwan
- Spain*
- Saudi Arabia
- Russia
- Poland
- Mexico
- Malesia
- India
- Czech Rep.
- China
- Austria
- Canada
- Australia
- Slovenia
- Italy
- Germany
- Korea
- PNNL
- KEK
The purpose of this workshop was to begin preparation for addressing the wide-area networking requirements for science in general and of the Belle II experiment in particular.

Report can be found at: http://www.es.net/assets/pubs_presos/Belle-II-Experiment-Network-Requirements-Workshop-v18-final.pdf

Various goals were defined, for example preliminary data challenge goals are:

<table>
<thead>
<tr>
<th>Date</th>
<th>Summer 2013</th>
<th>Summer 2014</th>
<th>Summer 2015</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rate</td>
<td>100MB/sec</td>
<td>400MB/sec</td>
<td>1000MB/sec</td>
</tr>
<tr>
<td>Duration</td>
<td>24 hours</td>
<td>48 hours</td>
<td>72 hours</td>
</tr>
</tbody>
</table>
The purpose of this workshop was to begin preparation for addressing the wide-area networking requirements for science in general and of the Belle II experiment in Europe.

Report is ongoing, potential milestones:

<table>
<thead>
<tr>
<th>Date</th>
<th>Winter 2013</th>
<th>Summer 2014</th>
<th>Summer 2015</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rate</td>
<td>100MB/sec</td>
<td>200MB/sec</td>
<td>400MB/sec</td>
</tr>
<tr>
<td>Duration</td>
<td>24 hours</td>
<td>48 hours</td>
<td>72 hours</td>
</tr>
</tbody>
</table>
First Belle II Data Challenge
KEK to PNNL & PNNL to GridKa

- Transfer rate from KEK to PNNL during 48hr stability test meet summer 2014 goal
- Transfer rate from PNNL to GridKa during 24 hr stability test was >100MBps (only 8hrs shown in figure).
Deploy new FTS3 server at PNNL

Evaluating FTS2 vs. FTS3

KEK to PNNL throughput for FTS3 was initially half that of FTS2

Fine tuning FTS3 is ongoing
FTS3 Throughputs

GridKa (Germany)  U of Victoria (Canada)

DESY (Germany)  INFN Torino (Italy)

SiGNET (Slovenia)  U of Melbourne (Australia)

KEK to PNNL
Tuning FTS3 allows us to reach ~8.8 Gbps transfer rates from KEK to PNNL.

Working on cloud implementation to provide load balanced FTS3 services.
Belle II ANA-100 Setup

Goal
- Test the ANA-100 Trans-Atlantic link
- Test/tune/profile the performance of current Belle II data transfer tools

Dates:
- June 2nd to 6th use dedicated VLAN
- June 10th to 20th extended use dedicated VLAN

Network Setup:
- Network providers (Geant, ESnet, GARR, DFN, etc.) setup the VLAN
- Local network providers and sites coordinated final configurations
- Sites must configure hardware interface to match destinations

Testing Tools:
- **Traceroute** was used to confirm the routing to each DTN
- **Iperf** was used to do initial network transfer rate tests
- **gridftp** and/or **srm-copy** was used to test site
- **FTS3 server** at GridKa was used to schedule data transfers

September 15-16, 2014

LHCOPN-LHCONE Meeting University of Michigan (US)
First few days we conducted tests using *iperf* for true network testing

- Required several parallel transfers to reach network saturation
- Reached ~9.6GBps (>2x the Tier-1 EU site requirements)

Transitioned to FTS3 server
- Reached network saturation but rates fell very quickly
- Large amount of drop packets
- Satisfies the incoming network requirements for Tier1 EU sites up to calendar year 6
Challenges encountered:
- The main issue was the configuration of the local network apparatus.
- Having all the servers at each site using/checking the proper network route.
- Hardware limitation (router, storage, etc.).
- Not having dedicated setups (shared with ATLAS, etc.).

Modification to sites to accommodate the increased rates:
- Modification of TCP windows was performed at PNNL and Italy.
- Routing hardware interface.
- Configure/tune network interrupts for multicore.
- Modification of the FTS3 optimization & global-timeout.
The KEK-PNNL VC and endpoints are configured.

Iperf tests were performed yielding a 1-3Gbps throughput.

Additional testing required.

Ideally, we would like to setup a gridftp server on KEK PC and start FTS3 transfers.
LHCONE is for LHC experiments
- Canadian and European sites are already part of LHCONE
- KEK and PNNL, key sites, are not part of LHCONE

Belle II thoughts and consideration:
- Belle II would like to have a closed network similar to LHCONE
- Configure LHCONE-like VRF layer?
  - Complicates configurations and operations for sites that are already part of LHCONE
- Can Belle II join LHCONE?
  - Is it difficult to expand LHCONE to non-LHC experiments?
  - Belle II traffic would be shared/compete with LHC experiment?
  - Easier to coordinate sites under one umbrella

Your comments/suggestions are invaluable to find the best solution!