Belle II Report

1. Experiment status
2. Site issues

April 3, 2019
@ GDB in Taiwan
Takanori HARA (KEK IPNS)
Belle II Report

1. Experiment status
2. Site issues
Motivation from Physics

**Accelerator**
- KEKB
  - Beam Energy (GeV): $3.5 \times 8$ ($\gamma = 0.425$)
  - CM energy: $\ldots, Y(4S), \ldots$
  - Luminosity ($cm^2 s^{-1}$): $2.1 \times 10^{34}$
  - Total data ($ab^{-1}$): $1$

**SuperKEKB**
- CM energy: $\ldots, Y(4S), \ldots$
- Luminosity ($cm^2 s^{-1}$): $8 \times 10^{35}$
- Total data ($ab^{-1}$): $50$

**Computing**
- one big center @ KEK (non-grid)

**Expected achievements at Belle II**
- Just an Image
- Discovery of $B \gamma D \tau \nu$
- Discovery of New Subatomic Particles
- Precise meas. of D mixing
- Observation with $\Upsilon(5S), \Upsilon(3S)$ etc.
- CKM Angle Measurements with 1 degree precision

**Motivation from Physics**
- Need large-scale computing resources and effective operation

**Physics Achievements at Belle**
- Serendipity
  - First tetra-quark $X(3872)$
  - Direct CPV in $B \rightarrow \pi \pi$ decay
  - $B \rightarrow Kll$ decay
  - $B \rightarrow \tau \nu$ decay
- Bottomed tetra-quark $Z_b$
- CPV in neutral $\pi$ system
- CPV in $B \rightarrow \pi \pi$
- $b \rightarrow d\gamma$ transition
- $D^0 \bar{D}^0$ mixing

**Expected achievements at Belle II**
- Discovery of $B \rightarrow K \nu\nu$
- Discovery of $B \rightarrow D \nu$
- Precise meas. of $D$ mixing
- Observation with $Y(5S), Y(3S)$ etc.
**Belle II Experiment: Global Schedule**

- **2016**
  - Phase 1: Global Cosmic run
  - w/o QCS, w/o Belle II
  - SINET4 → SINET5
    - Japan-US: 100Gbps
    - Japan-EU: 20Gbps
  - LER/HER single beam circulation was successfully finished
  - Beam background study was done
- **2017**
  - Summer shutdown (power saving)
  - New KEKCC operation started
  - Start using LHCONE
- **2018**
  - Power saving after 17 July, 2018
  - Phase 2 (MR): ARICH + Fwd ECL installation was completed
- **2019**
  - Phase 3: First official real data w/ CDC+TOP+ECL+KLM
  - FWD/BWD QCS was connected with VXD system
  - VXD installation
  - Summer shutdown (power saving)
  - W/ full Belle II
  - Plan to start on March 11, 2019

**Important Dates:**
- **2018**
  - Mid. Feb.–mid Jul. 2018: Phase 2 (MR) operation
  - Summer shutdown after 17 July, 2018

**Calendars:**
- JFY2016
- JFY2017
- JFY2018
- JFY2019

**Networks:**
- SINET4 → SINET5
- Japan-US: 100Gbps
- Japan-EU: 20Gbps

**Specific Events:**
- CDC installation was completed
- LER/HER single beam circulation was successfully finished
- Beam background study was done
Phase 3 beam run has started on 3/11!

SuperKEKB Phase 3 (Belle II Physics Run) Starts

Belle II and SuperKEKB are poised to become the world’s first Super B factory facility. Belle II aims to accumulate 50 times more data than its predecessor, Belle, and to seek out new physics hidden in subatomic particles that could shed light on mysteries of the early universe.

On March 11th, 2019, Phase 3 operation of the SuperKEKB project began, marking a major milestone in the development of Japan’s leading particle phase will be the physics run of the project, in which the Belle II experiment is taking data with a fully instrumented detector.

The KEKB accelerator, operated from 1999 to 2010, currently holds the world record for the most粲夸克-反夸克对产生


https://physicsworld.com/a/major-revamp-complete-at-superkekb-particle-physics-facility/
The first “B - anti-B like” event in the Belle II Phase3 physics run

Peak luminosity is still $\sim 10^{33} \text{ cm}^{-2} \text{s}^{-1}$ (1/800 of the target luminosity)

Electron-Positron collision has restarted on 3/25!

Day time : accelerator machine tuning
Night time : physics run
**RAW data flow**

Distributed Computing

Detector

2TB/day @ current lumi.
(tested up to 90 TB/day)

Belle II Detector

DAQ unit

DAQ unit

DAQ unit

Online disk

Online Database

Offline Database

Offline Database (Master)

BNL

(re)process

mDST

copy to BNL

KEK

/ghi/fs03

process

mDST

KEK

/ghi/fs01

permanent storage

SROOT data transfer

SROOT data transfer

SROOT → ROOT conversion, merge

Frontend server:

temporary storage

Core services for distributed computing

.....: DIRAC

.....: AMGA

.....: FTS

Database

Offline

Online

SROOT data transfer

RAW data registration

2TB/day @ current lumi.
(tested up to 90 TB/day)

KEKCC

Database

Offline

Online
**Distributed Computing**

- **Detector**
  - Belle II Detector
  - Online disk

- **Data**
  - 2TB/day @ current lumi.
  - (tested up to 90 TB/day)
  - **Frontend server:**
    - Temporary storage
    - SROOT -> ROOT conversion, merge
  - RAW(SROOT) -> RAW(ROOT)

- **Offline Database**
  - KEK
  - BNL
  - KEK /ghi/fs01
  - BNL /ghi/fs03

- **Core services for distributed computing**
  - DIRAC
  - AMGA
  - FTS

- **KEKCC**

**RAW data flow**

- **Online disk**
  - RAW(SROOT)
  - RAW(ROOT)

- **Database**
  - Offline (Master)
  - KEK
  - BNL
  - KEK /ghi/fs01
  - BNL /ghi/fs03
Data copy snapshot

From Jan. 17 to Feb. 9
193k SROOT files (=374TB)

High Energy News Vol.33 No.3 2014
http://www.jahep.org/hepnews/2014/14-3-2-BelleIIDAQ.pdf
**RAW data flow**

- **Belle II Detector**
- **Online disk**
- **Frontend server:**
  - temporary storage
  - SROOT → ROOT conversion, merge
- **Online Database**
- **Offline Database**
- **Database**
  - Offline (Master)
  - Online
- **KEKCC**
- **Distributed Computing**
  - Core services for distributed computing
    - DIRAC
    - AMGA
    - FTS
- **mDST**
- **2TB/day @ current lumi.**
  - (tested up to 90 TB/day)
- **RAW data transfer**
- **SROOT data transfer**
- **KEK**
  - /ghi/fs03
  - permanent storage
- **BNL**
  - (re)process
  - copy to BNL
- **mDST**
- **KEK**
  - /ghi/fs01
  - permanent storage
- **RAW data registration**
- **SROTOR**
- **Online disk**
- **Frontend server**
- **SROTOR**
- **ROOT conversion, merge**
- **mDST**
- **copy to BNL**

---

**Additional Information:**

- 2TB/day at current luminosity
- Tested up to 90 TB/day
- Database synchronization
- KEK and BNL locations
RAW data registration / replication

M.H. Villanueva (CINVESTAV), I. Ueda (KEK)

Raw Data Registration

Aim
- Upload files onto the Grid storage at KEK
- Register files into the file catalog and the metadata catalog

Initial Step
✔ Prepared a command-line tool by adapting the existing "end-user tool" for Raw data use case
  - To read the metadata embedded in the Raw (root) file
  - To receive extra metadata via a command-line option
  - To compose the logical file path accordingly
✔ Uploaded + Registered the GCR data with the tool

Second Step
✔ Built a system to do the work automatically
  - with a database with attributes and status for each file
  - with metrics in case of failures
✔ Uploaded + Registered the GCR data to Grid without a list from the FE system
✔ Initial small set of phase-2 data has been treated with a proper list from the FE system

cf. UNIC
The Online code (DataMover) is responsible
- To register raw data to BNL (metadata)
- To request upload from Online to Grid storage
The DIRAC RequestIngestAgent
- Executes the upload
The LHCBDIRAC RAWintegrityAgent
- verifies checksum of the uploaded files
- registers files to the File Catalog

https://atlas.web.cern.ch/Atlas/GROUPS/LHCB/DIRAC/Online GET

Exp3 data files (~31,000 files) were registered
Exp5 data files (~7,500 files RunType "cosmic" only) were registered

Raw Data Replication

Aim
- Make two replicas on the Grid (KEK and BNL for the first three years)
- Using the Belle II Distributed Data Management (DDM) system used in production

Initial Step
✔ Prepared a script to make requests to DDM specifying the destination (BNL)
✔ Treated the GCR data (up until late May) with the tool

Second Step
✔ Establish a system to automate requests and follow-up of their completion
✔ Replicated the GCR data (since the end of May) to BNL
- The phase-2 data yet to be treated

Third Step
- Extend the DDM and utilize the new feature, policy-based distribution
  - To be able to treat multiple Raw Data Centers
  - Choose a destination for each data block following the share and according to the status
- and adapt the Raw Data Registration system accordingly

RAW data files are being copied from KEK to BNL
RAW data registration / replication
Belle II Distributed Computing Structure

**Human**

- Software interface
  - Interware extension
  - Analysis user interface

**Production Manager**
- Client Tools
- Production Management System
- Fabrication System

**Data Manager**
- Web Portal
- Distributed Data Management System
- Monitoring

**End Users**
- Web Portal

**BelleDIRAC**
- v4r6p0

**Interware**
- v6r20p26

**Cyberinfrastructure**
- GRID services for Belle II

**Platform**
- GRID Middleware
- OS
- Hardware
- Network

**Sites**
- Cloud site
- Vcycle
- DIRAC slave
- Cloud I/F
- Cluster
- SE
- CE
- LFC
- AMGA
- CVMFS
- RMS
- DMS
- FTS

**Cloud**
- Site
- Infrastructure

CE: grid computing element
SE: grid storage element

Belle II computing sites

GRID sites
KEK, BNL, DESY, GridKA, KISTI, CNAF, many European sites
~30 sites: ~75%

Cloud sites
Univ. of Victoria, Univ. of Melbourne
several sites: ~15%

Computer cluster sites
Many Universities in Japan, Korean, Indian, China, Russia, Mexico,
~25 sites: ~10%

Normalized CPU usage by Site
46 Weeks from Week 52 of 2017 to Week 46 of 2018

Max: 339, Average: 183, Current: 61.6

~25 sites: ~10%
15
Status: Operation evolution

Continuous and stable operation

Light I/O MC production
Heavy I/O MC production
Skim production
RAW data process

Evolution of the Production Management / Fabrication system

- v4r4p25 (July 6)
  - CPU Time calculation: bug fix and improvement

- v4r5 (Aug 7)
  - Switched to new basf2 tools (b2*)
  - Enable to add user files to output sandbox
  - Allow multiple output job

- v4r5p1 (Aug 9)
  - Validate all files by b2file-check
  - Fix for flat ntuple output

- v4r5p2 (Sep 12)
  - Check basf2 and gbasf2 releases if allowed on GRID

Switched from “Proto”-Production system to “full” Production system in Sep. 2018

H. Miyake (KEK)
Produced MC/data files are collected in “Primary SEs”
- 10 primary SEs (Asia:3, US:1, Europe: 6) among ~30 SEs.
- Replica creation/clearning/BG distribution still in the manual basis
- Working on the automation.
  - Improvement of BelleDIRAC DDM
  - Implementation of Rucio into DIRAC
• ~2 JIRA tickets/day are submitted.
• 90% of expert slots are occupied by 5 members in 2018:
  - Difficult to continue 10 years operation.
  - Missing great opportunity for students to learn DC world.
• Development of training course.
• Automated JIRA/ggus submission under development
Local Grid
Set of user analysis clients (gbasf2) are developed
- Submission of job to grid/Check status
- Download output files

Still several hundreds of jobs in average.
- Approximately half of Belle II collaborators are registered on DIRAC
- Approximately half of them have experience to submit jobs.

Expected to increase with the start of data taking.
Belle II Report

1. Experiment status
2. Site issues
BNL increased the Storage capacity

BNL joined
Oct. 2017

PNNL decommission was completed

Total storage for Belle II

10 PB
wrong operation in some European sites

BNL increased the Storage capacity

PNNL decommissioned

KEK2-TMP-SE 22.8% NTUCC-TMP-SE 3.3% CESNET-TMP-SE 1.4% Frascati-TMP-SE 0.2%
PNNL-TMP-SE 12.9% CNAF-TMP-SE 3.1% TAU-TMP-SE 0.4% NTL-TMP-SE 0.1%
BNL-TMP-SE 12.9% KEK-RAN-SE 3.1% ULAKBIM-TMP-SE 0.3% Roma2-TMP-SE 0.0%
DESY-TMP-SE 7.1% KEK-TMP-SE 3.1% BNL-TAPE-TMP-SE 0.3% Torino-TMP-SE 0.0%
CERN-TMP-SE 6.3% KEK-TMPE-SE 2.0% IPPM-TMP-SE 0.3% Pisa-TMP-SE 0.0%
KIT-TMP-SE 5.5% Adelaide-TMP-SE 1.9% Melbourne-TMP-SE 0.2% CINEVESTAV-TMP-SE 0.0%
Napoli-TMP-SE 5.1% UFRJ-TMP-SE 1.8% CYFRONET-TMP-SE 0.2% IPHC-TMP-SE 0.0%
Karlsruhe-TMP-SE 3.9% VSI-TMP-SE 1.7% HEPHY-TMP-SE 0.2%
OSG stopped the Certificate Authority service at the end of May 2018

26 countries/region
113 institutes
856 colleagues
(as of Oct. 23, 2018)
OSG stopped the CA service

OSG stopped the Certificate Authority service at the end of May 2018

KEK GRID CA provides the certificate for US members as a temporal solution

BNL started providing the certificate (CILogon Silver CA 1) for US members

26 countries/region
113 institutes
856 colleagues
(as of Oct. 23, 2018)
Belle II Computing model

RAW data + produced mDST
end of year 3

KEK Data Center
BNL Data Center
Raw Data Center
Regional Data Center

GRID sites
Cloud sites
Computer cluster sites
HPC sites

Local resource

Storage for original + copy
Storage for copy
Temporary storage

CPU
Disk
Tape

Raw data
mdst Data
mdst MC
dashed
inputs for
udst
Ntuple

Detector

MC production and
Physics analysis skim

RAW data
mdst Data
mdst MC

user analysis
(Ntuple level)

MC production site

Asia

Europe site B

Europe site A
Belle II Computing model

RAW data + produced mDST
start year 4

Detector

Raw Data Center

KEK Data Center
(100%)

Regional Data Center

MC production site

GRID sites

Cloud sites

Computer cluster sites

HPC sites

North America

BNL Data Center
(30%)

Canada Data center (15%)

Europe

Germany Data Center
(20%)

Italy Data Center
(20%)

Local resource

MC production and
Physics analysis skim

User analysis
(Ntuple level)
In 2018
Japan-EU : 2 x 20 Gbps
Japan-US : 100 Gbps

From 2019 February
Japan-EU : 100 Gbps
Japan-US : 100 Gbps

100Gbps : around the North Hemisphere

Network between Japan and EU was improved

20 Gbps LINK via LONDON

Latency Drop between KEK (JP) perfsonar and KIT (DE) perfsonar from 91ms to 82ms one way

NEW 100Gbps LINK VIA AMSTERDAM
“Run-dependent” MC production on GRID

- Huge beam BG compared with Belle.
  - Possible efficiency/resolution degradation
- Essential to implement properly
  - Largely depend on the accelerator condition.
    Need run dependent BG simulation.
- BG files are prepared beforehand, and “overlaid” in simulated event.
  They are distributed to SEs or shared places.
- Even in early phase3 only, total amount is several TB to assure randomness.
  - Difficult to put in local cluster sites.
  - Put part of BG files depending on the CPU resources.
- How to distribute run dependent beam BG is under discussion.