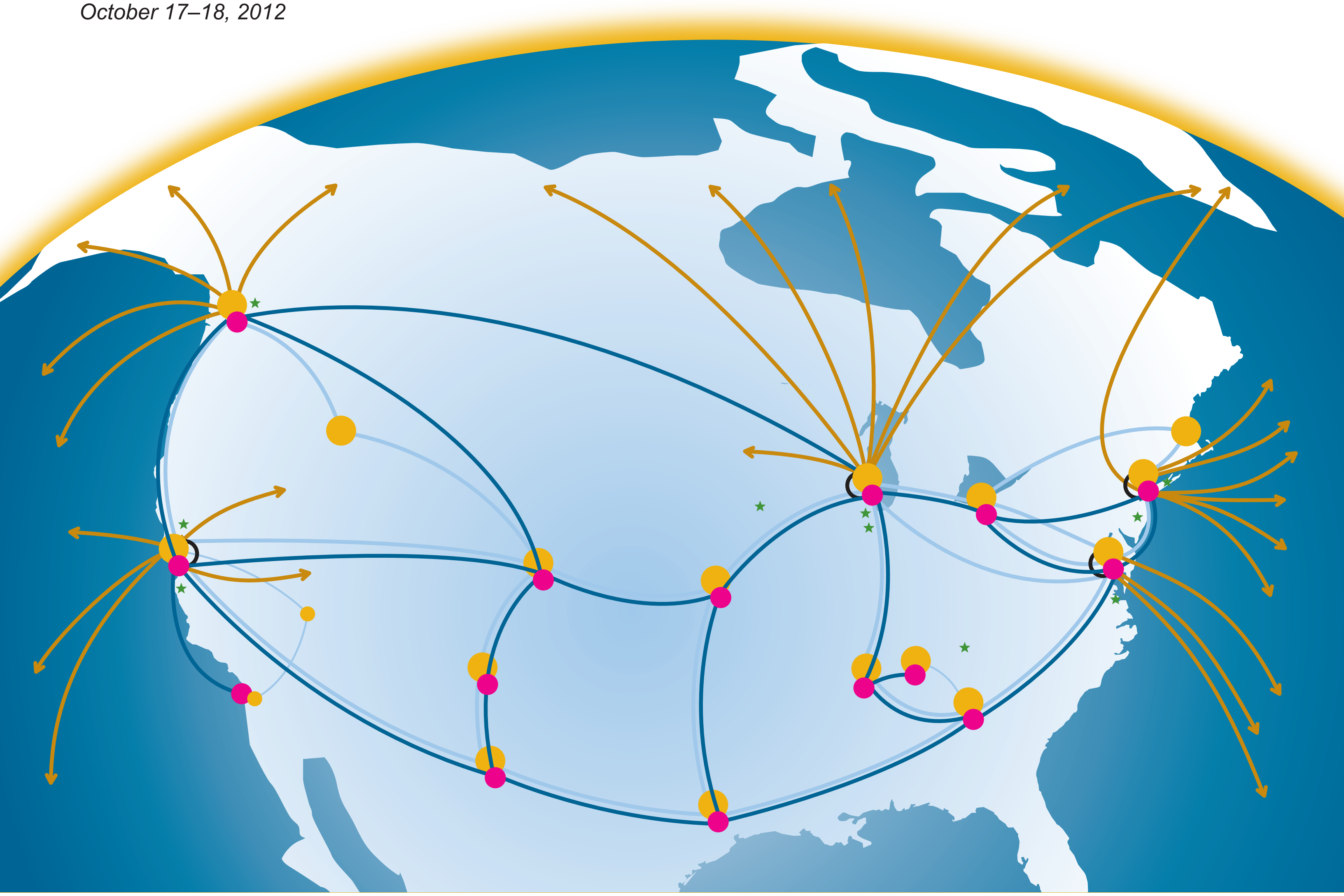
Belle-II Experiment Network Requirements

Belle-II Experiment Collaboration
Energy Sciences Network

October 17–18, 2012



PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01830

Printed in the United States of America

Available to DOE and DOE contractors from the
Office of Scientific and Technical Information,
P.O. Box 62, Oak Ridge, TN 37831-0062;
ph: (865) 576-8401
fax: (865) 576-5728
email: reports@adonis.osti.gov

Available to the public from the National Technical Information Service
5301 Shawnee Rd., Alexandria, VA 22312
ph: (800) 553-NTIS (6847)
email: orders@ntis.gov <http://www.ntis.gov/about/form.aspx>
Online ordering: http://www.ntis.gov

This document was printed on recycled paper.

(8/2010)

Belle II Experiment Network Requirements Workshop

Final Report

Pacific Northwest National Laboratory
October 17, 18, 2012

Participants and Contributors

David Asner, PNNL
Greg Bell, ESnet
Tim Carlson, PNNL
David Cowley, PNNL
Eli Dart, ESnet
Brock Erwin, PNNL
Romulus Godang, University of S. Alabama
Takanori Hara, KEK
Jerry Johnson, PNNL
Ron Johnson, UW
Bill Johnston, ESnet
Kerstin Kleese-van Dam, PNNL
Toshiaki Kaneko, KEK
Yoshihiro Kubota, NII
Thomas Kuhr, KIT
John McCoy, PNNL
Hideki Miyake, KEK
Inder Monga, ESnet
Motonori Nakamura, NII
Leo Piilonen, Virginia Tech
Ruth Pordes, FNAL
Douglas Ray, PNNL
Richard Russell, PNNL
Malachi Schram, PNNL
Jim Schroeder, PNNL
Martin Sevior, University of Melbourne
Surya Singh, PNNL
Soh Suzuki, KEK
Takashi Sasaki, KEK
Jim Williams, Indiana University

Editors

David Asner, PNNL — david.asner@pnnl.gov
Eli Dart, ESnet — dart@es.net
Takanori Hara, KEK — takanori.hara@kek.jp

Table of Contents

1 Executive Summary

The Belle experiment, part of a broad-based search for new physics, is a collaboration of ~400 physicists from 55 institutions across four continents. The Belle detector is located at the KEKB accelerator in Tsukuba, Japan. The Belle detector was operated at the asymmetric electron-positron collider KEKB from 1999-2010. The detector accumulated more than 1 ab^{-1} of integrated luminosity, corresponding to more than 2 PB of data near 10 GeV center-of-mass energy. Recently, KEK has initiated a $400 million accelerator upgrade to be called SuperKEKB, designed to produce instantaneous and integrated luminosity two orders of magnitude greater than KEKB. The new international collaboration at SuperKEKB is called Belle II. The first data from Belle II/SuperKEKB is expected in 2015.

In October 2012, senior members of the Belle-II collaboration gathered at PNNL to discuss the computing and neworking requirements of the Belle-II experiment with ESnet staff and other computing and networking experts. The day-and-a-half-long workshop characterized the instruments and facilities used in the experiment, the process of science for Belle-II, and the computing and networking equipment and configuration requirements to realize the full scientific potential of the collaboration's work.

The requirements identified at the Belle II Experiment Requirements workshop are summarized in the Findings section, and are described in more detail in this report.

KEK invited Belle II organizations to attend a follow-up meeting hosted by PNNL during SC12 in Salt Lake City on November 13, 2012. The notes from this meeting are in Appendix C.

2 Findings

To accommodate Belle II's anticipated data rates, network upgrades to allow 100 Gbps data rates will be necessary at both KEK and PNNL.

KEK is currently connected to SINET at 20 Gbps (two 10-Gbps links) at Tsukuba. SINET will be transitioning to a new network infrastructure – SINET5 – in 2016. This will be close to the time when the Belle II experiment begins production operation. It will be important for KEK, SINET, ESnet, and PNNL to collaborate closely so that SINET is aware of the needs of the Belle II experiment in time to incorporate them into the plans for SINET5 and ensure a smooth transition.

As part of the SINET5 network to be deployed in 2016, SINET is expected to provide a 100Gbps trans-Pacific connection between Japan and Seattle. Depending on the available capacity between Japan and the United States before SINET5 is deployed, there may be a shortage of Japan-US bandwidth in 2015.

PNNL has an optical transport system that provides connectivity between PNNL and ESnet at two locations – Seattle and Boise. The current system does not have 100 Gbps capability. It is likely that upgrades to the PNNL optical transport system will be necessary to support the Belle II experiment.

In order to provide resiliency at 100 Gbps speeds, ESnet will consider adding 100 Gbps capability at its Boise location. ESnet currently has 100 Gbps capability at Seattle, the location of the primary connection between ESnet and PNNL. However, 100 Gbps connectivity for the backup connection between ESnet and PNNL at Boise should be considered.

Data and service challenge exercises have been used successfully by other experiments to verify and harden the networking, computing, and software infrastructures used in the conduct of the science. The Belle-II experiment plans several such challenges. These efforts will require coordination between the scientific, computing, and networking organizations that support the Belle-II experiment.

It is expected that PNNL will serve data to European as well as U.S. institutions. This will have implications for trans-Atlantic capacity network capacity.

Additionally, the replacement of the KEKCC computing infrastructure in 2015 may place additional demands on the trans-Pacific network infrastructure because of the desire to replicate data to the PNNL data center, both to maintain analysis continuity and to mitigate the risk of data loss during the upgrade. The KEKCC computing infrastructure is expected to be replaced again in 2018.

Undersea network cables outages are days to weeks longer than is typical for terrestrial cables. In light of this, the Belle-II experiment should consider diverse paths for network connectivity between KEK and PNNL.

3 Action Items

In light of the high data rates that will be sustained by the Belle II experiment, engineering resiliency and service continuity for Belle II's experiment data flows may pose significant technical or financial challenges. The Belle II collaboration and the relevant network infrastructure providers should develop a plan for network service resiliency in both the terrestrial and undersea portions of the network.

ESnet, KEK, PNNL, and SINET will work together to bring up perfSONAR tests on the network paths that will support the data transfers from KEK to PNNL.

ESnet, KEK, PNNL, and SINET will work together on data and service challenges to ensure that the experiment infrastructure is ready for production when data taking begins.

4 Background

In August 2011, the US Department of Energy Office of Science (SC) Office of High Energy Physics (HEP) issued a Mission Need Statement for Next Generation B Factory Detector Systems that calls out the upgrade of the current Belle detector at KEK as the preferred option for making precision measurements of the properties of B and D mesons and tau leptons to indirectly probe for scientific discoveries at and beyond the energy range of the Large Hadron Collider. In September 2012, the US Belle II Project managed by PNNL achieved the CD-1 milestone. The US has key roles in the Belle II detector upgrade that are required to match the ongoing upgrade in accelerator capabilities.

The Belle II experiment is part of a broad-based search for new physics. The Large Hadron Collider (LHC), which plans to resume operation in 2015 with high luminosity at a center-of-mass energy of 13 TeV, is designed to search for new physics at the energy frontier. Its high center-of-mass energy may allow it to produce heavy, as-yet-undiscovered particles such as supersymmetric partners of quarks and leptons, or new particles linked to extra dimensions. The SuperKEKB/Belle II facility searches for new physics by colliding very high intensity particle beams, i.e., by precisely measuring and comparing with theory a number of observables that are difficult or infeasible to measure at the energy frontier. In the past, measurements of processes involving internal loops have provided access to high-mass scales before accelerators were available to probe these scales directly. To continue this paradigm-shifting pursuit of flavor physics, about two orders of magnitude more data is now needed.

Following the earthquake and tsunami of March 12, 2011 off Japan, the electrical power available to KEK was dramatically reduced, and most of Belle's computing—centralized at KEK—was offline. The KEK Director General requested emergency computing assistance from PNNL. The DOE-supported Belle computing at PNNL came online in July 2011. Approximately 100 Belle users now have accounts at PNNL and access to both Belle data transferred over the network from KEK and Monte Carlo (MC) samples generated at PNNL. In November 2012, a Belle Virtual Organization (VO) utilizing Open Science Grid (OSG) middleware was established at PNNL with funding support from both DOE and the US-Japan Collaboration (nichi-bei). This VO will evolve into a Belle II "grid site," satisfying the US obligation to the planned distributed computing system of Belle II.

The Belle II computing system has to be able to handle an amount of data eventually corresponding to ~50 PB per year (the anticipated production at SuperKEKB design luminosity). In order to achieve Belle II's physics goals within a timely manner, the network must be equipped to facilitate the processing of the raw data without any delay to the experiment's data acquisition, in addition to the production of Monte Carlo samples corresponding to several times the beam data. To achieve this, Belle II has adopted a distributed computing model based on the grid. A key component of this model is the establishment of a remote data center at PNNL, where the raw data can be reprocessed in parallel with KEK within a Belle II distributed computing framework.

5 The Belle II Experiment

The Belle experiment at the KEKB accelerator [1] in Tsukuba, Japan, is a collaboration of ~400 physicists from 55 institutions across four continents. The Belle detector [2] recorded electron-positron interactions near 10 GeV center-of-mass energy over the period 1999-2010 and has published ~350 papers from these data to date. KEK has initiated an accelerator upgrade, SuperKEKB [3], designed to have instantaneous and integrated luminosity two orders of magnitude greater than those of KEKB. The new international collaboration at SuperKEKB is called Belle II [3].

In November 2009, the Belle groups at the Universities of Cincinnati, Hawaii, and Virginia Tech submitted a White Paper to the DOE/SC-HEP for US participation in the upgraded Belle II experiment. The DOE/SC-HEP subsequently requested that a formal proposal be submitted and subsequently presented at an "Intensity Frontier" Review held August 10-11, 2010. As a result of this review, the DOE/SC decided to support US participation in Belle II, and requested that another proposal be prepared that would better align US responsibilities for Belle II detector construction with funding profiles proposed by the DOE/SC. Since that review, several additional US institutions have joined Belle and Belle II, including PNNL, Carnegie Mellon University, Indiana University, Kennesaw State University, Luther College, University of Pittsburgh, University of South Alabama, and Wayne State University. In August 2011, the US Department of Energy Office of Science, Office of High Energy Physics approved a Mission Need Statement for a Next Generation B-Factory Detector Systems. In that Mission Need Statement, the upgrade of the current Belle detector at KEK was called out as the preferred option for making precision measurements of the properties of B and D mesons and tau leptons to indirectly probe for scientific discoveries at and beyond the energy range of the Large Hadron Collider. In September 2012, the US Belle II Project managed by PNNL achieved the CD-1 milestone. The international Belle II Collaboration now includes more that 400 scientists from 21 countries – Japan, US, Australia, Austria, Canada, China, Czech, Germany, India, Korea, Malaysia, Poland, Russia, Saudi Arabia, Slovenia, Spain, Taiwan, Thailand, Turkey, Ukraine, Vietnam.

5.1 Physics Case

Research in flavor physics is an essential component of the future US program in particle physics. The design luminosity of SuperKEKB is around 1×10^{36} $cm^{-2}s^{-1}$[1], approximately fifty times higher than what has been achieved at the KEKB accelerator. This will allow a data sample with an integrated luminosity of 50 ab^{-1} to be accumulated. This can be compared to the 1 ab^{-1} data sample obtained over a decade of Belle running with KEKB. Belle, together with BaBar, established the existence of large charge-parity (CP) violation (i.e., matter- antimatter asymmetry) in the b quark system, in agreement with the expectation of Kobayashi and Maskawa (KM). In contrast to the kaon system (strange quarks), the observed CP violation effects for b quarks are of order one rather than 10^{-3}. This critical experimental contribution of the B factories was explicitly recognized in the 2008 Nobel Prize in Physics citation.

The B factory results show most of standard model CP violation can be explained by the single irreducible complex KM phase in the weak interaction coupling. Nevertheless, the possibility of contributions from new physics that are O(10%) of the size of the standard model contribution are not ruled out. Moreover, the matter-antimatter asymmetry of the universe cannot be explained by the KM phase alone. The standard model KM explanation of the baryon asymmetry of the universe falls short by ten orders of magnitude. This demonstrates that there must be new sources of CP violation and new heavy particles that remain to be discovered.

The Belle II experiment is part of a broad-based search for new physics. The Large Hadron Collider (LHC), which plans to resume operation in 2015 with high luminosity at a center-of-mass energy of 13 TeV, is designed to search for new physics at the energy frontier, i.e., its high center-of-mass energy may allow it to produce heavy, as-yet-undiscovered particles such as supersymmetric partners of quarks and leptons,or new particles linked to extra dimensions. The SuperKEKB/Belle II facility searches for new physics using the very high luminosity of the Intensity frontier, to compare theoretical calculations and precision measurements of branching fractions, angular distributions, CP asymmetries, forward-backward asymmetries, and a host of other observables that are difficult or infeasible to measure at the LHC.

At the LHC and in the future at the International Linear Collider (ILC), the particles that are responsible for the electroweak force in the Standard Model will be studied. These include the W, Z^0 and the recently discovered Higgs boson. In the upcoming round of neutrino and muon experiments, which will feature improved measurements of neutrino mixing angles and a determination of whether future neutrino CP violation experiments are feasible, the properties of leptons will be explored. However, these experiments will not explore the new physics possibilities of the flavor- and heavy-quark–sector that will be studied at SuperKEKB / Belle II.

In the past, before accelerators were available to directly probe high-mass scales, measurements of processes involving internal loops provided researchers access to these scales. For example, the suppression of $K_L \rightarrow \mu^+\mu^-$ decays allowed theorists to infer the existence of the charm quark; the charm quark mass was subsequently estimated from the observed rate of neutral Kaon meson oscillations. The unexpected observation of CP violation in K^0 meson decays was used to predict the existence of a third generation of quarks. The unexpected discovery of large neutral B meson oscillations indicated that the top quark was very heavy, contrary to the theoretical prejudice at the time (and contrary to where experiments were looking). These processes, as well as the violation of CP symmetry, are quantum mechanical phenomena sensitive to very high energy scales and have revolutionized the thinking about extensions of the Standard Model. To continue this paradigm-shifting pursuit of flavor physics, about two orders of magnitude more data is needed. Such a data set would tell us whether the CP violation effects observed in B decays are consistent with the Standard Model. Searching for flavor-changing neutral-currents (FCNC) with such a data set would probe a mass range of 1-100 TeV, which is mostly beyond the reach of direct searches at the LHC. If supersymmetry is discovered during LHC operation at 13 TeV, a Super B Factory could help determine how the supersymmetry was broken [2].

6 Computing Environment

6.1 Background

The SuperKEKB accelerator is designed to deliver an instantaneous luminosity of 8 x 10^{35} $cm^{-2}s^{-1}$ in 2022; thus, the Belle II experiment will collect a data sample of 50 ab^{-1}, eventually corresponding to a few hundred petabytes in total data. The Belle II computing system is required to process this ever-increasing data sample without any delay to the experiment data acquisition, and to produce Monte Carlo events and physics analysis. The computing resources required for these purposes increase faster than the projected performance of CPUs and storage devices. Under these circumstances, it cannot be expected that one centralized computing center such as KEK will be able to provide all computing resources for the whole Belle II collaboration. Meanwhile, the Belle II collaboration is expanding to more countries throughout the world, and it would be reasonable for all collaborators to contribute the Belle II computing system. In fact, many of the institutes and universities that belong to the Belle II collaboration have already started operating grid sites for the LHC experiments. After evaluating these factors, the Belle-II collaboration adopted a distributed computing model based on the grid as a baseline design.

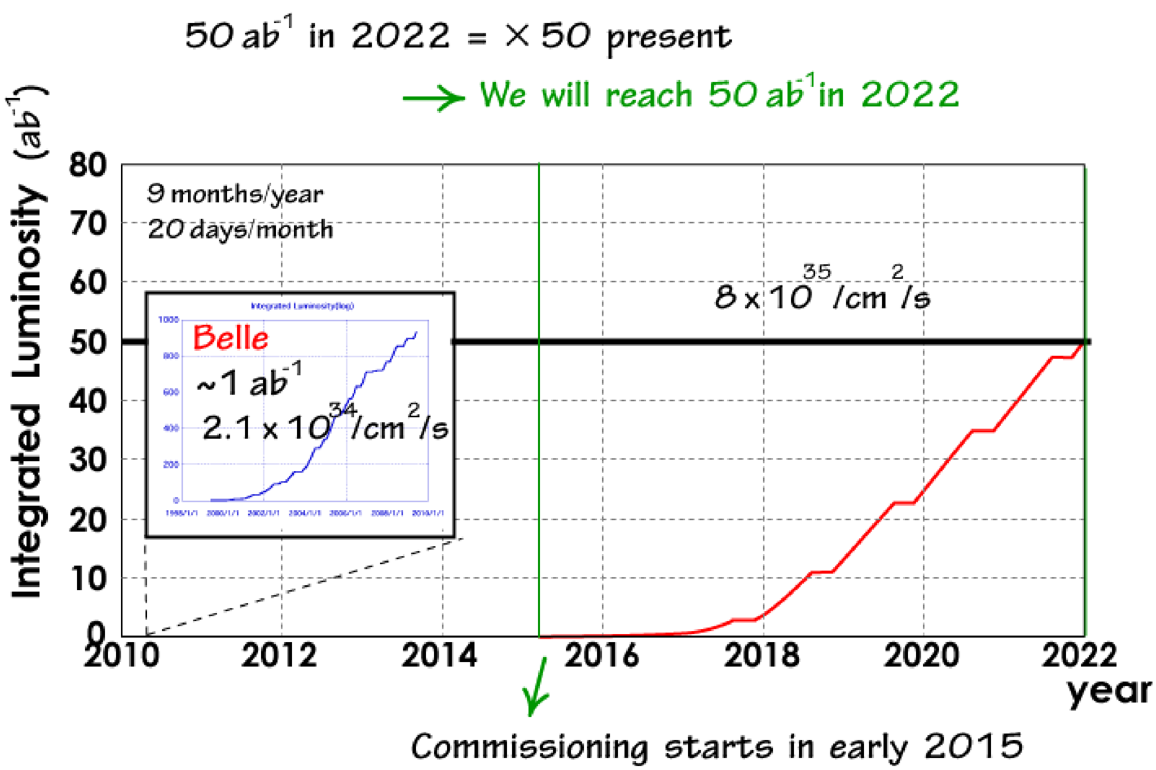


Figure 6.1: Integrated luminosity expectations profile for Belle II

6.2 *Belle II Distributed Computing Design*

The Belle II computing system has to accomplish several tasks, e.g., raw data processing, MC sample production, physics analysis processes, and archiving the data resulting from each process. The Collaboration decided to process and store raw data at KEK. The resulting output (called “mDST”), is roughly one tenth of the size of the raw data and will be distributed to the other grid sites. This reduces unnecessary network traffic from KEK to each grid site. In other words, we will not distribute the raw data to each Tier1-level grid site. This is a simpler model than the current LHC experiments utilize. However, the Collaboration also decided to have one duplicate of the raw data at PNNL. One reason for this is that the Belle experiment faced a temporary suspension of all of analysis activities after the earthquake in 2011 because all of the processed data were stored at KEK only. Another reason is the important role of the reprocessing center outside KEK. In terms of computing, KEK’s first priority is the data acquisition and processing of raw data. In parallel with this, at the early stage of the experiment (i.e., until we understand the detector performance well), the software and the detector constants must be updated often; consequently the raw data has to be reprocessed frequently. This reprocessing can be performed only at a place where the raw data is stored. If KEK were the only place to have the raw data, it would have to be done at KEK, making KEK’s workload much heavier. On the other hand, if we have another place to store raw data, the reprocessing can be performed there, reducing the load on KEK. Finally, this arrangement will make the physics results available faster. Therefore, another data center, such as PNNL, is very important for the Belle II experiment from the perspective not only of a backup copy of the raw data but also of flexibility in raw data reprocessing. However, the data transfer from KEK to PNNL will require a one order of magnitude higher network bandwidth between KEK/Japan and PNNL/US.

In the experience of the Belle Collaboration, MC samples at least six times larger than the beam data are required to produce precision measurements. As the MC production does not need large input files, this task can be distributed easily to the grid sites. In order to reduce peak demand, optional processing by cloud computing is planned. The output format of the MC events is the same as that of the mDST from the beam data. The MC samples will be placed on the disks on the grid site where it was produced, and at least one replica will be distributed to other grid sites.

As with the Belle II computing design, we expect users to perform analysis processes on the mDST files on the grid and to transfer the resulting lighter output (Ntuple—see Figure 6.2) to the local resource. The local resources will ideally be grid-enabled, but we explicitly include non-grid resources, like private clusters at institutes.

Figure 6.2 shows the concept of the Belle II computing model.

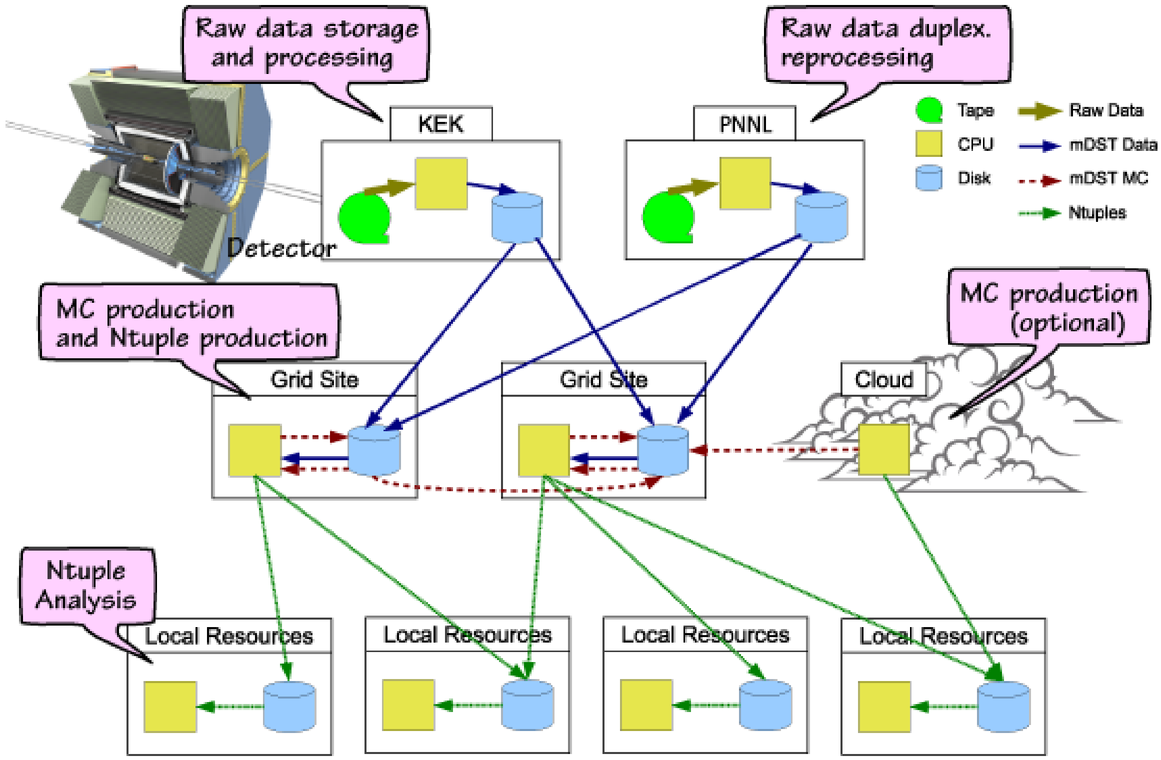


Figure 6.2: Concept of Belle II computing model

Based on the Belle II computing design and the luminosity expectations for the SuperKEKB accelerator, we can estimate the total required computing resources as a function of the calendar year as shown in Table 6.1.

Year	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
Tape [PB]	2.8	2.8	2.8	2.8	19.24	54.43	103.55	153.89	204.64	255.39
Disk [PB]	4.00	4.00	5.00	8.00	27.98	79.17	115.68	153.10	190.82	228.55
CPU [kHepSPEC]	45.00	45.00	50.00	55.00	328.31	568.98	567.54	609.45	643.14	672.60

Table 6-1: Total required Belle II computing resources

KEK is expected to perform the raw data processing and storage. In addition, the tasks of MC production and user analysis are allocated in proportion to the number of Belle II PhD physicists at that site (25% is assumed for KEK in this estimation). Table 6.2 shows the expected computing resources at KEK. The resulting mDST files are transferred to each grid site to create at least one complete copy across the Belle II computing system, and the raw data will also be transferred to PNNL.

KEK resources:

Year	2013	2014	2015	2016	2017	2018	2019	2020	2021
Tape [PB]	2.80	2.80	2.80	2.80	9.62	27.22	51.77	76.94	102.3
Disk [PB]	3.00	3.00	3.00	3.00	4.57	12.94	22.44	32.17	41.98
CPU [kHepSPEC]	40.00	40.00	40.00	40.00	93.60	168.01	184.65	196.20	204.98
WAN [Gbit/s]	0.50	1.00	2.50	4.00	9.71	18.83	24.25	24.86	25.06

Table 6-2: Required Belle II computing resources at KEK

KEK has 2.8PB tape and 3.0PB disk storage space and 40kHepSPEC CPU power now. KEK will keep using the current computing system until 2015 summer, of which resources (The tape/disk storage and CPU) has satisfied the expected requirement.

PNNL, as a raw data storage center, plays an important role in data reprocessing. We assume that the reprocessing will be repeated most frequently in the first year of the data collection, then the number of reprocesses is expected to decrease as the reconstruction software matures. Finally, after four years of operation, the Collaboration must stop reprocessing activities, except in the case that a more sophisticated reconstruction algorithm is invented. On the other hand, the amount of beam data will increase as the instantaneous luminosity increases. PNNL will mainly handle the reprocessing in the early stage of the experiment and evolve into a data storage role in the latter stage. PNNL will store the latest version of the mDST and the second-latest version of mDST. As with the reprocessing of the raw data, the corresponding MC samples will also be produced in proportion to the number of PhD physicists in each grid site (15% for PNNL). Another role of PNNL will be to distribute the reprocessed mDST to the Belle II grid sites. Table 6.3 shows the required computing resources for PNNL.

PNNL resources:

Year	2013	2014	2015	2016	2017	2018	2019	2020	2021
Tape [PB]	0.00	0.00	0.00	0.00	9.62	27.22	51.77	76.94	102.3
Disk [PB]	1.00	1.00	2.00	5.00	12.00	17.00	22.00	27.00	32.00
CPU [kHepSPEC]	5.00	5.00	10.00	15.00	59.11	95.81	76.58	82.65	87.63
WAN [Gbit/s]	0.50	1.00	2.50	4.00	8.65	15.75	18.82	19.29	19.44

Table 6-3: Required Belle II resources at PNNL

As stated above, each regional grid center is expected to perform MC production and user analysis in proportion to the number of PhD physicists assigned to that site, but this will not be the case with raw data storage. As a case study, the required computing resources for GridKa in Germany, where a 14% of the MC sample is produced, is summarized in Table 6.4. GridKa could become a major Belle II data center in Europe and is expected to have a full copy of the mDST for beam data and the aforementioned proportionate amount of the mDST for MC.

Regional Center resources:

Year	2013	2014	2015	2016	2017	2018	2019	2020	2021
Disk [PB]	0.00	0.00	0.03	0.27	3.22	9.12	15.16	21.35	27.59
CPU [kHepSPEC]	0.00	0.00	0.68	4.34	41.40	71.94	72.21	77.93	82.63
WAN [Gbit/s]	0.00	0.00	0.04	0.27	2.65	4.55	3.86	3.96	3.99

Table 6-4: Required Belle II computing resources at 14% Regional Grid Center

The regional center will have storage space and CPU for the Belle II experiment even before taking the beam data, e.g., 2013 and 2014 for the Monte Carlo event production campaign. But this table does not include the currently available computing resources at regional centers.

6.3 Belle II Distributed Computing Software

The Belle-II experiment has adopted the grid computing model to enable the processing of the very large volume of experimental data and Monte Carlo samples that the Collaboration must analyze. In order to realize this, we also need to access to different types of computing resources, such as gLite middleware in Europe/Japan, OSG middleware in US, cloud computing and local resources. DIRAC, which was originally developed by LHCb and is now an independent project, can provide this environment. First of all, DIRAC backend allows us to process jobs on the heterogeneous computing systems listed above, once the backend interface corresponding to each system is prepared. Another feature of DIRAC is the pilot job, which provides more reliable job scheduling. Furthermore, the DIRAC monitoring system is useful to check the status of jobs and resources. For metadata service software, we are examining AMGA (ARDA Metadata Grid Application), which provides us efficient and scalable metadata searching. These two servers are the core of the Belle II distributed computing system and now operated at KEK. However, in terms of redundancy, it is preferable to have alternate servers at another site where the raw data reprocessing and the large-scale MC production are planned, e.g., PNNL.

File-level metadata catalogues (AMGA, AMI etc.) are used to improve the identification of suitable event samples. While tremendously helpful, these methods are increasingly seen as too coarse to provide an efficient analytical environment for large-scale data. The emerging concept of event-metadata-based selection and access is promising; however, current systems do not provide the necessary scalability and functionality to handle the timely analysis of extreme-scale data. Therefore, it is important that new research be performed to improve the existing event-metadata-based technology.

As previously stated, there are several tools that are used to deploy and monitor each component of the grid infrastructure. However, this monitoring information appears not to be collected presently. Developing new services/agents that collect the monitored information to optimize the overall performance of the grid virtual organization would decrease the time needed for new discoveries as well as the load on resources. Examples of improvements include network redirection/throttling, CPU and storage-load balancing, and others. The Belle II experiment provides an ideal environment in which to test new ideas, which can in turn be applied to other scientific endeavors.

To perform raw data processing, MC sample production, and user physics analysis, we developed the common software framework named “basf2.” It has a software bus architecture, and a large-scale application such as MC sample production is realized by plugging in a set of a building block unit “modules”—for example the event generator, the full detector simulator, and reconstruction tools—into this framework. The modules are controlled by a “path” in which the order of the modules is defined in the steering file written in Python. basf2 also supports parallel processing. For the grid environment, we developed a user interface, “gbasf2,” in which the steering file is the same as basf2 jobs with some supplemental information such as the grid project name.

7 Data Analysis Workflow

7.1 Background

After data collection at KEK, the raw data will be copied from online disk storage to offline tape storage. Each sub-detector expert starts to prepare the precise calibration and alignment constants online, based on the information stored in Data Quality Monitor. Once these constants are ready, the raw data will be processed to produce the data for the physics analysis, called “mDST,” which includes information about the reconstructed track and particle identification as well as metadata for the run. This process is defined as the DST production and is accomplished within the basf2 software framework. The procedure to prepare the precise constants within a proper time interval (less than a week) is still under discussion. The data flow after the online system is depicted in Figure 7.1.

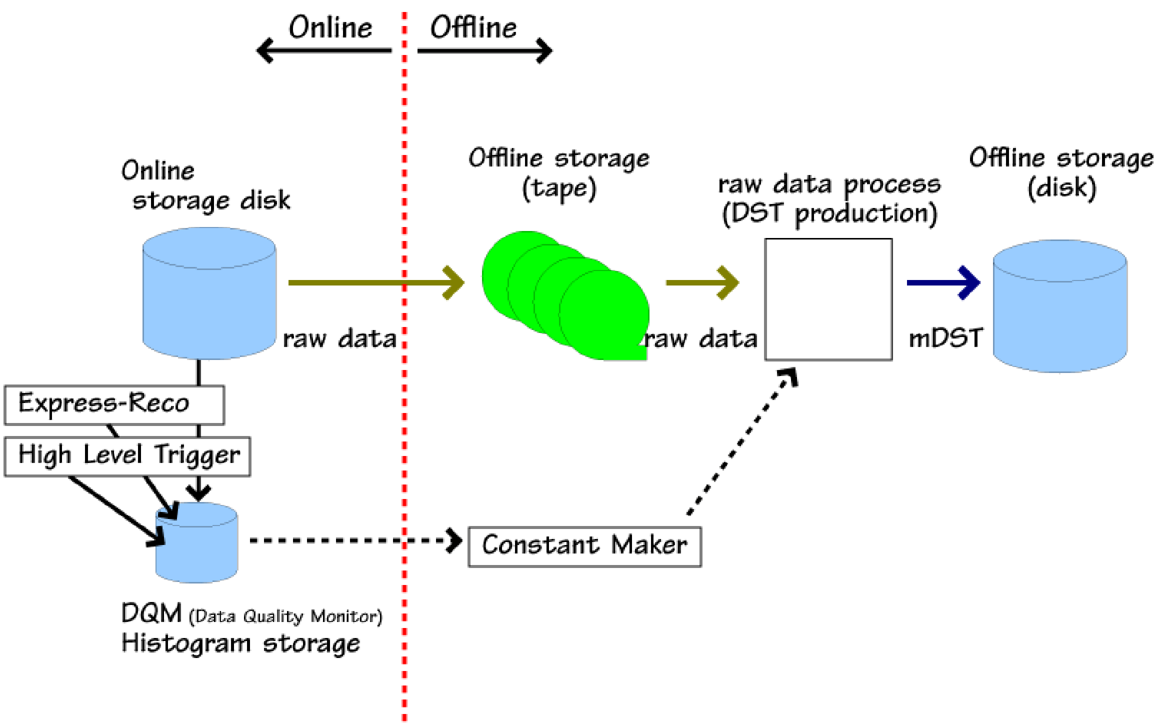


Figure 7.1: Belle II data flow

To understand the signal efficiency and background components, we need to produce the MC data, taking into account the beam profile condition and the beam background level as realistically as possible.

The results of the raw data processing are stored on disk at KEK and PNNL. In addition, each of the analysis sites will produce Monte Carlo data in proportion to the number of physicists at the analysis site as a percentage of the total collaboration. The Monte Carlo data is expected to be six times the experimental data in volume.

7.2 Instruments and Facilities

7.2.1 KEK Belle II Computing

The KEK computing system for the Belle II experiment is located at the KEK Computing Research Center. Previously, there were two computing systems at KEK, one for the Belle experiment and the other for the remaining on-going experiments and projects. However, these two have been unified into a single system since April 2012. This new system, known as “KEK Central Computing System (KEKCC),” has roughly 4,000 cores of calculation servers, 7 PB disk space and a 16-PB capacity tape storage system; it will continue operating until summer, 2015. The grid system is a part of the KEKCC, with 3,000 of 4,000 cores available for grid jobs, which corresponds to 44 kHEP-SPEC[1]. Though the total resources of KEKCC meet the requirements of the Belle II experiment until 2015, as shown in Table 6.2, they have to be shared with other experiments—for example J-PARC, ILC, Experimental Cosmology, Accelerator, Theory group and on-going analysis jobs for Belle. Figure 7.2 illustrates the configuration of the KEKCC system with the network connection.

[1] **HEP-SPEC** is the High Energy Physics-wide benchmark for measuring CPU performance, developed in 2009

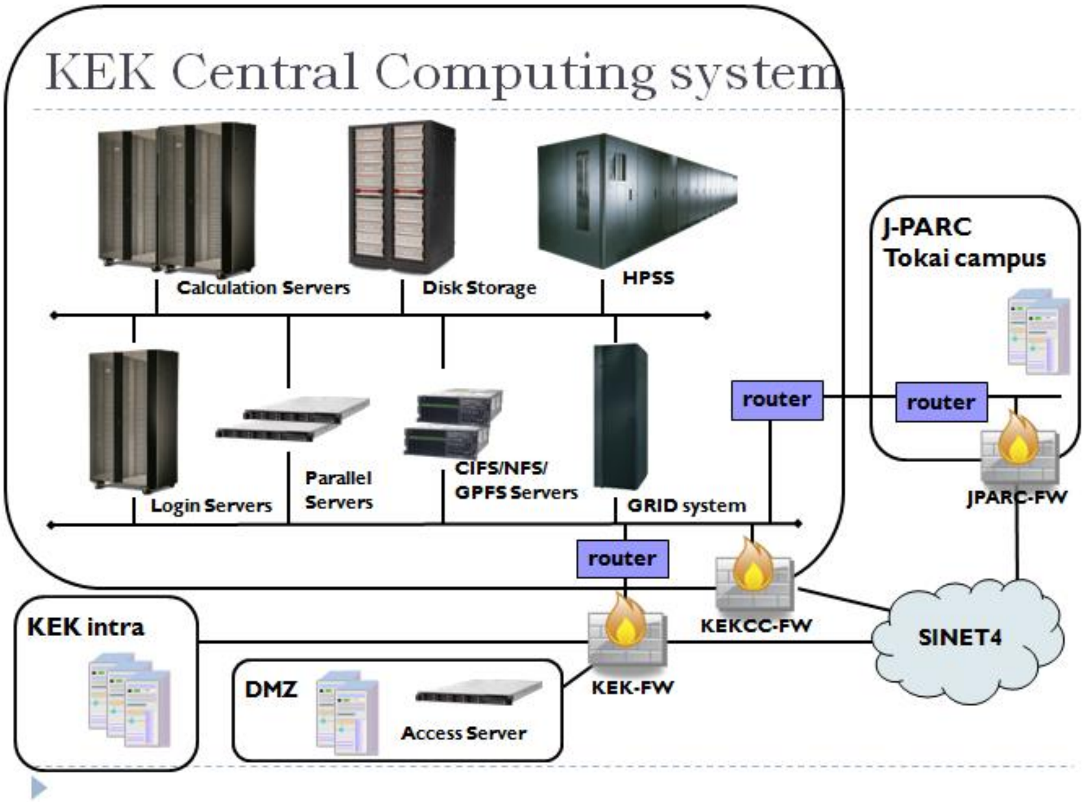


Figure 7.2: Configuration of KEK Central Computing

7.2.2 PNNL Belle II Computing

Computing infrastructure for Belle II is located on PNNL's main campus in Richland, Washington USA. Compute systems and data storage will be located in the Computer Science Facility and/or the William R. Wiley Environmental Molecular Science Laboratory (EMSL) buildings. Belle II computing will be supported initially by a mixture of both dedicated and shared resources belonging to PNNL's Institutional Computing (PIC), and EMSL. Procurement of new, dedicated resources for compute and data storage are planned as the collaboration matures and funding for Belle II at PNNL grows.

At present, Belle II uses computer time and stores data on PNNL's shared computer cluster, Olympus. The collaboration also utilizes space in the EMSL High Performance Storage System (HPSS) archive and operates its own cluster of 1024 cores used for grid computing. Operational support is provided by EMSL staff in partnership with PNNL's Belle II physicists.

PNNL provides a data transfer node with 10 Gbs connectivity to both the Internet and the multiple-petabyte Lustre file system; this node provides shared data access to Olympus and to the Belle II grid nodes.

PNNL plans to meet growth in Belle II computational and data storage needs by procuring additional servers and storage annually to maximize resources per dollar as much as possible. PNNL intends to leverage this strategy to allow all Belle II data to be kept on disk. If needed, PNNL will fall back on tape storage as a cost-cutting contingency.

7.3 Process of Science

Under the Belle II computing design, every possible grid site is expected to have some number of output files resulting from raw data processing/reprocessing and MC events in proportion to the number of PhD physicists assigned to that grid site. Further, GridKa in Germany, KISTI in Korea and PNNL in the US will have full copies of the mDST of the beam data. The Belle II grid sites are listed in Table 7.1, each with its status as of summer 2012. Most of the sites are Tier 1/2 centers of the LHC experiments and now they also support the Belle Virtual Organization (Belle VO). GridKa and DESY in Germany, KISTI in South Korea, SiGNET in Slovenia, CesNet in Czech Republic, Cyfronet in Poland and KEK employ gLite middleware, and PNNL began operations with OSG middleware in 2012. Furthermore, we are working with academic cloud projects in Australia and Poland towards using DIRAC for Belle II distributed computing.

As the Belle II collaboration is still growing, the proportion of the MC sample production and storage for the mDST files assigned to each center will vary year by year. We are also in discussions with other grid sites, such as Russia and China, that have not yet officially agreed to support the resources for Belle II distributed computing.

Country	Sites	Belle VO	Comment
Australia	Tier2/3	Supported	Cloud system planned
Austria	Tier2		
China	Tier2		DIRAC server
Czech Republic	Tier2	Supported	
Germany	Tier1/2	Supported	
India	Tier2		New data center planned
Japan	KEK	Supported	
Korea	Tier2	Supported	
Poland	Tier2/3	Supported	Cloud system developed
Russia	Tier2		
Slovenia	Tier2	Supported	
Taiwan	Tier1/2		
USA	OSG	Supported	Site at PNNL is set up

Table 7-1: Status of Belle II grid sites

The main datasets from the raw data processing are the B-pair, tau-pair and continuum event samples. The High Level Physics Trigger will be applied to reduce the beam background contribution as much as possible. The cross section of the tau-pair event is comparable to that of the B-pair events. By contrast, the amount of the continuum events is three times larger than the B-pair event sample.

Assuming the running time of SuperKEKB accelerator to be two-thirds of a year with scheduled instantaneous luminosity and an mDST event size of 40 kB, the total expected data size of these physics events (B-pair, tau-pair, and continuum) are calculated as shown in Table 7.2.

mDST:

Calendar Year	2013	2014	2015	2016	2017	2018	2019	2020	2021
Size per year [PB]	0.00	0.00	0.01	0.06	0.78	1.56	2.18	2.24	2.26
Total size [PB]	0.00	0.00	0.01	0.07	0.85	2.42	4.60	6.84	9.09

Table 7-2: Belle II data mDST size

As noted earlier, the performance of precision measurements will require MC samples corresponding to at least six times the beam data to perform the precision measurement. The original MC samples are kept at the grid site where they are produced, and one copy will be distributed to the other grid sites. Table 7.3 shows the expected data volume of the MC samples.

MC data:

Calendar Year	2013	2014	2015	2016	2017	2018	2019	2020	2021
Size per year [PB]	0.00	0.00	0.05	0.38	4.69	9.39	13.10	13.42	13.53
Total size [PB]	0.00	0.00	0.05	0.44	5.13	14.52	27.61	41.04	54.57

Table 7-3: Belle II Monte Carlo mDST size

Physics run data will not be obtained until 2016. However, to ensure a smooth start to Belle II's distributed computing operations once the physics run starts, it will be critical to conduct substantial data challenges and MC sample production dry runs using the actual data analysis codes and workflow tools over the next two years. This verification is essential to understanding the potential bottlenecks in the computing design, including the network and software infrastructure. Since the raw data replication workflow and the data analysis workflow will run concurrently with the experiment, it is expected that at least the 2014 and 2015 data challenges will run concurrently with those for the raw data replication workflow. A notional schedule for the data analysis workflow data challenges is described in Table 7-4.

Date	Summer 2013	Summer 2014	Summer 2015	Production
Number of Sites	25%	70%	100%	100%
Data Rate, %	25%	70%	100%	100%
Data Rate, Actual	100 MB/sec	400 MB/sec	1000 MB/sec	1000 MB/sec
Duration	24 hours	48 hours	72 hours	24 hours/day

Table 7-4: Schedule for the data analysis workflow data challenges

7.4 Instruments and Facilities, Next 2-5 Years

The SuperKEKB accelerator is scheduled to begin commissioning in 2014, and the physics run will start in 2016. During this period, as stated above, we will iterate large–scale, extensive MC sample production tests and data challenges for the data analysis workflow and the raw data replication workflow.

Once the physics run starts, the Collaboration's work priorities will be 1) raw data processing, 2) calibration/alignment constant evaluation, and 3) MC event run production and 4) user analysis on the new KEKCC.

The required computing resources for KEK during this period are summarized in Table 6.2. However, the current KEKCC system will be replaced entirely with the new system in the summer, 2015, just before the first physics run. We have considered carefully how to transfer the data stored in the current system to the new system. If PNNL can have the entire data stored before the KEKCC is replaced, the replacement of the KEK computing system will go more smoothly.

7.5 Process of Science, Next 2-5 Years

Over the next two to five years, before the physics run begins, the cosmic and/or beam test data must be conducted with the actual Belle II detector. Although the raw data will not be in perfect format, it will be adequate to check the data analysis and raw data replication workflows at least partially. In parallel, KEK should carry out realistic and repeated runs of raw data processing with MC events, and concurrently, PNNL should conduct similar tests of raw data reprocessing.

Once the physics run starts, the data analysis workflow and the raw data replication workflow from KEK to PNNL will run concurrently. Furthermore, while the software is tuned up with the real beam data, the data may need to be reprocessed frequently, especially in the first couple of years of the experiment. After that, reprocessing will be necessary less frequently. Table 7.5 shows the reprocessing plan schedule.

Schedule:

Calendar Year	2013	2014	2015	2016	2017	2018	2019	2020	2021
Reprocessings per year	0	0	4	3	2	1	0	0	0

Table 7-5: Belle II raw data reprocessing plan

Tables 7.6 and 7.7 show the required additional storage capacity, CPU power and the network bandwidth for the reprocessing of the beam data and the MC samples, respectively.

Data Size:

Calendar Year	2013	2014	2015	2016	2017	2018	2019	2020	2021
Total size version -1 [PB]	0.00	0.00	0.01	0.07	0.85	2.42	2.42	2.42	2.42
Required @ PNNL [kHepSPEC]	0.00	0.00	0.36	2.10	15.21	19.52	0.00	0.00	0.00
WAN per copy [GBit/s]	0.00	0.00	0.01	0.09	0.68	0.97	0.00	0.00	0.00

Table 7-6: Resources required for Belle II raw data reprocessing

MC Size:

Calendar Year	2013	2014	2015	2016	2017	2018	2019	2020	2021
Total size version -1 [PB]	0.00	0.00	0.05	0.44	5.13	14.52	14.52	14.52	14.52
Required [kHepSPEC]	0.00	0.00	2.82	16.58	120.1	154.2	0.00	0.00	0.00
WAN per copy [GBit/s]	0.00	0.00	0.09	0.52	4.09	5.79	0.00	0.00	0.00

Table 7-7: Resources required for Belle II Monte Carlo reprocessing

As raw data processing with real beam data commences with the new computing system, unexpected problems may arise with the new system; for example, wrong initial settings and unknown hardware malfunctions. These problems will have to be troubleshot one by one to finally stabilize the system.

7.6 Instruments and Facilities, 5+ Years

As the experiment increases in capability and reaches full luminosity, the data rates will continue to increase. The required resources for the data analysis workflow for raw data processing and MC sample production during this period are summarized in Table 6.1 and 6.2, respectively.

However, again, another replacement of the entire KEK computing system is scheduled in summer 2018. Unexpected hardware malfunctions and software glitches in that new system could once again cause problems. Furthermore, because the KEK computing system might have roughly 30 PB of total data by the summer of 2018, data migration could present serious issues to overcome. It may be necessary to migrate legacy data back from PNNL to KEK at the same time that new raw data is moving from KEK to PNNL.

7.7 Process of Science, 5+ Years

After the first five years, no reprocessing is expected to be needed except in the case that a more sophisticated reconstruction algorithm is invented.

8 Raw Data Replication

8.1 Background

Assuming the expected instantaneous luminosity and the raw data event size of 300 kB, the data rate at KEK is estimated to be 1.8 GB/s. Assuming the maximum file size of the raw data to be 4 GB, a raw data file will be generated roughly every two seconds, ultimately amounting to more than ten thousand files created in a typical physics run. Once the raw data file is closed on the online storage disk, the Data Acquisition System (DAQ) will return an acknowledgement to the offline computing system, and then the data transfer to the offline tape storage can start. During this procedure, the file metadata is extracted and registered in AMGA so that it can be accessed from the DST production expert via grid jobs (this scheme is still under discussion). After the completion of the data transfer, the raw data will be processed on the grid system, and the resultant mDST file is stored on the offline disk storage at KEK. Because the DAQ network should be separated from the Internet, we will have a special network path between the online storage disk and the offline computing system.

The raw data and metadata are replicated from the offline tape storage at KEK to disk at PNNL. It can be then processed in parallel to the raw data processing at KEK and/or reprocessed later with the updated detector calibration constants. The data files that result from the raw data processing will be kept on disk for distribution to scientists for analysis.

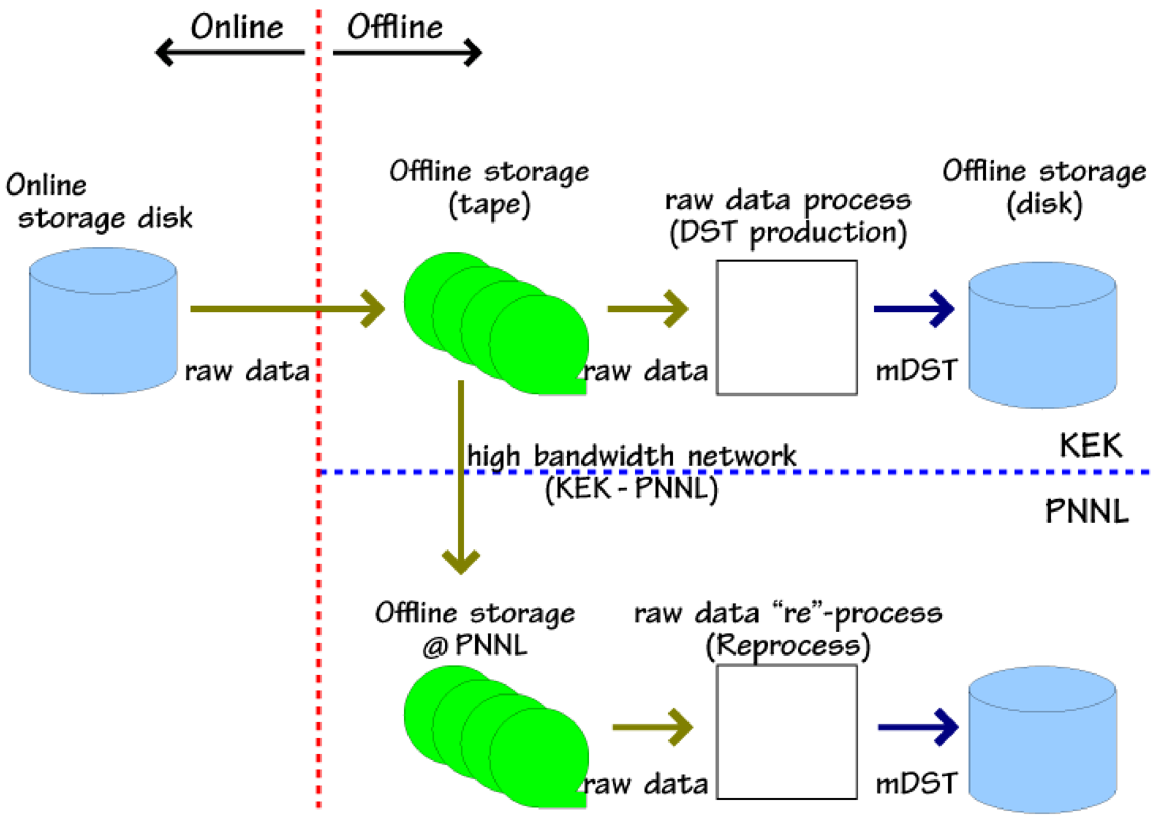


Figure 8.1: Belle II data flow

8.2 Instruments and Facilities

Over the next two years, the infrastructure for replicating the raw data will be developed, deployed, and tested. This will require the development of network configuration, data transfer node configuration, security policy development, and workflow integration.

Several aspects of these tasks were discussed at the Belle II Experiment Requirements workshop. One aspect is whether to use a standard routed network service or a virtual circuit service for data replication. The consensus of the group was to explore a virtual circuit service because of the additional capabilities of traffic isolation and traffic engineering that a virtual circuit service provides—these were seen as advantages over a best-effort routed service.

In addition, perfSONAR resources will need to be deployed so that the infrastructure for data replication can be tested, performance problems can be diagnosed, and so forth. KEK, PNNL, and ESnet currently have perfSONAR servers deployed.

8.3 Process of Science

Over the next two years, the data replication workflow must be developed and tested.

The Belle II Experiment Requirements workshop attendees discussed the use of data challenges, wherein the data replication workflow is run for a period of time with simulated data. Each data challenge would have a performance target, with each successive challenge having a higher performance target until the final challenge, which would run at the peak performance level expected for the first year or two of production physics runs on the Belle II experiment.

The workshop reached consensus that the first data challenge would be held by the summer of 2013. A table containing the notional goals first challenge and two additional data challenges is below:

Date	Summer 2013	Summer 2014	Summer 2015	Production
Rate	100MB/sec	400MB/sec	1000MB/sec	1000MB/sec
Duration	24 hours	48 hours	72 hours	24 hours/day

Table 8-1: Goals of the Belle II data challenges

It is likely that some portion of the data transfer nodes, storage, and network equipment that will be used when the experiment begins production operation will be purchased sometime in 2015. The data challenge in the summer of 2015 should be conducted using the equipment that will be used in production operation of the experiment.

Since the raw data replication and data analysis workflows will run concurrently when the experiment is running in production, it is expected that at least the 2014 and 2015 data challenges will run concurrently with the data challenges for the data analysis workflow.

8.4 Instruments and Facilities, next 2-5 years

Production operation of the Belle II experiment is scheduled to begin in 2016. Once production operation begins, the raw data replication workflow is expected to run for 2/3 of the year, increasing in data volume as the capabilities of the detector increase. The expected data production volume and data rate of the raw data replication workflow for the years 2016, 2017, and 2018 is contained in the following table.

Year	2016	2017	2018
Tape [PB]	0.82	9.62	27.22
Disk [PB]	0.39	4.57	12.94
WAN [Gbit/s]	0.84	9.71	18.83

Table 8-2: Expected Belle II raw data production and replication volume

8.5 Process of Science, 2-5 years

As the performance requirements for the raw data replication workflow increase, there will be a need for development, test, and measurement of additional systems and software capabilities. It is expected that these activities will be conducted during accelerator downtime. This will require coordination between the operational groups responsible for the different parts of the infrastructure, including KEK, SINET, ESnet, and PNNL.

8.6 Instruments and Facilities, 5+ years

As the experiment increases in capability and reaches full luminosity, the data rates will continue to increase. Table 5.2 and 5.3 contains the expected data rates and volumes including the raw data workflow for KEK and PNNL, respectively.

8.7 Process of Science, 5+ years

Very little process change from the 2-5 year case is expected.

8.8 Network Resiliency and Undersea Cables

Undersea cables typically have longer repair times than terrestrial circuits due to the challenges of a marine environment. The time to repair undersea cable faults is typically several days to a few weeks, where the time to repair terrestrial circuits is typically less than 24 hours, and usually less than two days.

8.9 Network Infrastructure Upgrades

In order to provide resiliency at 100 Gbps speeds, ESnet will consider adding 100 Gbps capability at its Boise location. ESnet has 100 Gbps capability at Seattle now, and Seattle is the location of the primary connection between ESnet and PNNL. However, 100 Gbps connectivity for the backup connection between ESnet and PNNL at Boise should be considered.

PNNL has an optical transport system that provides connectivity between PNNL and ESnet at two locations – Seattle and Boise. The current system does not have 100 Gbps capability. It is likely that upgrades to the PNNL optical transport system will be necessary to support the Belle II experiment.

KEK is currently connected to SINET at 20 Gbps (two 10 Gbps links) at Tsukuba. It is likely that this connectivity will have to be increased in order to support the Belle II experiment.

SINET will be transitioning to a new network infrastructure – SINET5 – in 2016. This will be close to the time when the Belle II experiment begins production operation. It will be important for KEK, SINET, ESnet, and PNNL to collaborate closely so that SINET is aware of the needs of the Belle II experiment in time to incorporate those needs into the plans for SINET5 and ensure a smooth transition.

8.10 Summary Table

Feature	Key Science Drivers		Anticipated Network Requirements	
Time Frame	**Science Instruments and Facilities**	**Process of Science**	**Local Area Network Bandwidth and Services**	**Wide Area Network Bandwidth and Services**
Near-term (0-2 years)	• Development of the Belle II raw data replication workflow system and infrastructure	• Development, test, verification, and commissioning • Periodic data challenges to ensure data replication workflow is ready	• 2 PB data to be copied from online disk to offline disk to test workflow • 2.5 GBit/sec bandwidth to computational analysis for testing prompt reconstruction workflows	• Virtual circuit configuration for raw data replication workflow • 100 MB/sec for 24 hours in first data challenge • 400 MB/sec for 48 hours in second data challenge • 1000 MB/sec for 72 hours in third data challenge • Periodic test flows for debugging and performance analysis of workflow systems
2-5 years	• First few years of physics using Belle II • 300 kB event size • Raw data files of 4 GB, more than 10,000 files from each run • Increasing data production as experiment is refined	• Replication of raw data from KEK to PNNL • Processing of raw data into mDSTs at PNNL and KEK • Data challenges for increased replication rates during experiment shutdown periods	• 5 PB data to be copied from online disk to offline disk • 10 GBit/sec LAN bandwidth to computational analysis for prompt reconstruction	• 80 MB/sec from KEK to PNNL for raw data replication in first year • Growth to 1500 MB/sec for raw data replication by 2018
5+ years	• Progression to full luminosity at Belle II	• No change	• 110 PB data to be copied from online disk to offline disk • 25 GBit/sec LAN bandwidth to computational analysis for prompt reconstruction	• 1500 MB/sec from KEK to PNNL for raw data replication in 2018 • Growth to 1900 MB/sec for raw data replication by 2022

9 Glossary

ab^{-1}	an inverse attobarn (10^{42} cm^{-2}) is a unit of integrated luminosity
AMGA	ARDA Metadata Grid Application - the EGI gLite service that allows metadata handling on the grid.
AMI	Amazon Machine Image
ARDA	A Realization of Distributed Analysis developed for LHC experiments
Babar	B-physics experiment at SLAC
Belle	B-physics experiment at KEK
CP	Charge-parity
CesNet	is the operator of academic network of the Czech Republic
Cyfronet	academic computing center in Poland
DESY	Deutsches Elektronen-Synchrotron (German Electron Synchrotron) is a national accelerator laboratory in Germany
DIRAC	Distributed Infrastructure with Remote Agent Control
DOE(/SC)	U.S. Department of Energy (/Office of Science)
DST	Data Summary Tapes – files of reconstructed events
ESnet	Energy Sciences network
EMSL	William R. Wiley Environmental Molecular Science Laboratory (at PNNL)
GB/sec or Gbps	Gigabytes per second – a measure of network bandwidth or data throughput
Grid	Global computing network
HEP	High Energy Physics
KEK	*Kō Enerugī Kasokuki Kenkyū Kikō* – Japanese high energy physics research laboratory with campuses in Tsukuba and Tokai
KEKB	Accelerator for B physics at KEK
KEKCC	KEK Central Computing
KISTI	Korea Institute of Science and Technology Information
KM	Kobayashi and Masakawa quark mixing matrix
ILC	International Linear Collider
LHC	Large Hadron Collider
MB/sec or Mbps	Megabytes per second – a measure of network bandwidth or data throughput
MC	Monte Carlo
mDST	Mini-DST
OSG	Open Science Grid
PB/sec or Pbps	Petabytes per second – a measure of network bandwidth or data throughput
perfSONAR	PERFormance Service Oriented Network monitoring

	ARchitecture is a tool for end-to-end monitoring and troubleshooting of multi-domain network performance
PNNL	Pacific Northwest National Laboratory
SiGNET	Slovenian Grid NETwork
SINET	Science Information Network – Japanese academic computer network
SuperKEKB	Upgrade to KEKB
TB/sec or Tbps	Terabits per second – a measure of network bandwidth or data throughput
VO	Virtual Organization

10 Appendix A – Belle II Data Networking Workshop Agenda

OCT 17, 2012 (DAY 1)

TIME	DESCRIPTION	PRESENTER	LOCATION
9:00-9:30 a.m.	Welcome	Douglas Ray, ALD,FCSD	Darwin 1007
9:30-10:00 a.m.	Workshop Kickoff	David Asner	Darwin 1007
	Overview of Collaboration Context		
	Go around the room, introductions, logistics		
	Belle Collaboration		
	Participating funding agencies, countries, institutions		
10:00-10:30 a.m.	Introduction of KEK Computing Research Center	Toshiaki Kaneko	Darwin 1007
10:30-10:45 a.m.	Break		
	Wide Area Networks – Architecture and Capabilities		
10:45-11:15 p.m.	Report on ESnet	Greg Bell	Darwin 1007
11:15-11:45 a.m.	Report on SINET	Professor Nakamura	Darwin 1007
11:45-12:15 p.m.	Report on IRNC, TransPAC3	Jim Williams	Darwin 1007
12:15-1:15 p.m.	Lunch		
	The network as an enabler for distributed large-scale science		
1:15-2:25 p.m.	Lessons from the LHC	Bill Johnston	Darwin 1007
2:25-2:55 p.m.	Science DMZ and perfSONAR	Eli Dart	Darwin 1007
2:55-3:15 p.m.	PerfSONAR	Soh Suzuki	Darwin 1007
3:15-3:30 p.m.	Break		

3:30-5:00 p.m.	Belle II Case Study Discussion	David Asner and Eli Dart	
	Case Study 1: Experiment to Tier0 data flow		Darwin 1007
	Case Study 2: Tier0 to analysis sites data flow		Pasteur 2019
5:00 p.m.	Summary and Wrap-up of Day One		Darwin 1007
	Agenda checkup for day two and make changes as necessary		

ADJOURN

6:30 p.m.	No Host Dinner via own transportation		Anthony's at Columbia Point, 550 Columbia Point Drive, Richland

OCT 18, 2012 (DAY 2)

TIME	DESCRIPTION	PRESENTER	LOCATION
9:00-9:15 a.m.	Introduction to Day 2	David Asner, Eli Dart	Darwin 1007
	Recap of previous day		
	Brief round robin to raise thoughts		
9:15-9:45 a.m.	KEK Grid Status	Takashi Sasaki	Darwin 1007
	Belle II Grid and Computing Environment		
9:45-10:15 a.m.	Belle II Computing Design	Martin Sevior	Darwin 1007
10:15-10:30 a.m.	Break		
10:30-11:00 a.m.	Belle II Data Model	Hideki Miyake	Darwin 1007
11:00-11:30 a.m.	Status of Belle II Grid Environment and Performance	Takanori Hara	Darwin 1007
11:30-12:00 p.m.	Grid Technologies	Ruth Pordes	Darwin 1007
12:00-1:00 p.m.	Lunch		
1:00-1:45 p.m.	Case Study 3 – Computing Environment	David Asner, Eli Dart	Darwin 1007

1:45-3:00 p.m.	Belle II Experiment Organizational Structures Group Discussion Networking Group Performance Measurement Group (close to networking group) Systems Group Software / Tools / Grid Group	David Asner, Eli Dart	Darwin 1007
3:00-3:15 p.m.	Break		
3:15-5:00 p.m.	Open Discussion Near-term Activities (perfSONAR deployments, etc.) Report Outline Report Content		Darwin 1007

ADJOURN

11 Appendix B – Notes from KEK SC12 Meeting

1. KEK invited Belle II Consortium members to attend a meeting at SC12. The organizations that attended were: Cyfronet, Poland; DESY, Germany; FZU Czech Republic; KEK Japan; PNNL, US

3. It was pointed out that the Belle II organization is flat as compared to the LHC hierarchy. Current plans have both KEK and PNNL hosting Belle II's raw data.

4. There is an assumption that most Belle II partners will obtain data from PNNL as network connectivity from the partner country to the US may be easier, or less costly, to implement than to Japan. In other words, PNNL will be a cache for all of the Belle II data.

5. NII in Japan will decide on an appropriate network path(s)to PNNL and other Belle II sites based on cost

6. The FZU representative commented that they may have an issue duplicating the LHCONE infrastructure for Belle II.

7. There was general concern that Belle II Physics leadership needs to put together a schedule for data challenges. The data challenges would:

    a. Allow perfsonar everywhere

    b. Ensure that TCP/IP parameters are correct

    c. Ensure that the partner institutions are ready to participate in science when data become available.

    d. Accomplish a test showing all infrastructure is in place during 2013

    e. Accomplish a test demonstrating network bandwidth of say 30% of the anticipated maximum network bandwidth by 2014.

8. The DESY representative recommended that an "abstraction" model for data be set up

9. Monte Carlo results will stay at the partner site where they were created.

10. There appears to be a need for a common I/O framework for Belle II.

11. There was general concern that Belle II needs to organize a networking group that can provide input/recommendations to the physics group.

12. It was agreed that Sasaki-san would communicate key questions to the Physics leadership of Belle II at the upcoming conference. David Asner to help formulate the questions.

13. Participants agreed the meeting at SC12 was beneficial and there was consensus to hold a similar meeting at SC13 in Denver. PNNL offered its assistance to KEK in setting up such a meeting.

12 Acknowledgements

This work would not have been possible without the contributions and participation of those who provided information and attended the **Belle II Experiment Network Requirements** workshop.

Lower left cover image courtesy Rey.Hori / Takada Office / KEK

Lower right cover image courtesy Scott Butner / PNNL

PNNL is operated by Battelle for the U.S. Department of Energy under contract DE-AC05-76RL01830

ESnet is funded by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research (ASCR). Vince Dattoria is the ESnet Program Manager.

ESnet is operated by Lawrence Berkeley National Laboratory, which is operated by the University of California for the U.S. Department of Energy under contract DE-AC02-05CH11231.

This work was supported by the Directors of the Office of Science, Office of Advanced Scientific Computing Research, Facilities Division, and the Office of High Energy Physics.

This is KEK Proceedings 2013-1, LBNL report LBNL-6268E, and PNNL report PNNL-22373