Chapter 1
Distributed Computing Model for Belle II Experiment

T. Hara, S. Ahn, M. Bračko, K. Cho, Z. Drasal, T. Fifield, R. Grzymkowski, B. Y. Han, S. Hwang, Y. Iida, R. Itoh, G. Iwai, H. Jang, N. Katayama, Y. Kawai, C. Kiesling, B. K. Kim, J. H. Kim, T. Kuhr, D. J. Kong, S. Lee, A. Moll, H. Nakazawa, S. Nishida, H. Palka, K. Prothmann, M. Röhrken, T. Sasaki, M. E. Sevior, M. Sitarz, S. Stanič, Y. Watase, H. Yoon, J. Yu, M. Zdybal
(the Belle II Computing Group)

Abstract In order to search for the new physics beyond the Standard Model, an upgrade of the Belle, so-called the Belle II experiment is being planned. We intend to commence this experiment in 2013 with the target instantaneous luminosity $8 \times 10^{35}/\mathrm{cm}^2/\mathrm{s}$. One of the important points to make the experiment successful is the computing model, including the data handling system, the date distribution and the large-scale simulation/reconstruction. We have started to draft the possible distributed computing model. The basic concept of the Belle II computing model is presented.

1.1 Experimental environment

In order to explain the absence of antimatter in our Universe the violation of *CP* symmetry is believed to be essential. The *CP* violation in the Standard Model (SM), which is a quantum field theory of quarks, leptons, and gauge bosons based on the $SU(2) \times U(1)$ local gauge symmetry, stems from an irreducible complex phase in the weak interaction 3×3 quark-mixing (CKM) matrix [1]. Since 1999 the Belle experiment [2, 3] has started taking data at the energy-asymmetric e^+e^- collider KEKB with the primary aim of studying the *CP* violation in the decays of neutral *B* mesons. Up to now (as of July, 2009), we have logged $\mathscr{L} = \sim 950\mathrm{fb}^{-1}$ of data and have provided many epoch-making results, for instance, a large *CP* asymmetry in neutral *B* meson system [4], discovery of new particles [5, 6, 7, 8, 9], and the direct *CP* violation in charged *B* meson decays [10]. Though the several results show possible deviations from the SM, which may indicate the new physics, the statistical uncertainties in many measurements remain still large. To search for a new source of *CP* violation beyond the SM that may revolutionize the understanding of the origin

Takanori Hara
High Energy Accelerator Research Organization, 1-1 Oho, Tsukuba, Ibaraki, Japan, e-mail: takanori.hara@kek.jp

of the matter-dominated Universe, we need to make more precise measurements. For this purpose, SuperKEKB, a major upgrade of the current KEKB aiming at the instantaneous luminosity of $8 \times 10^{35}/\mathrm{cm}^2/\mathrm{s}$, is being proposed. Figure 1.1 shows the prospect of the integrated luminosity as a function of a year. After a 3-year shutdown that is needed to upgrade KEKB and the Belle detector, we will start Belle II experiment [11] from 2013. The final goal of the integrated luminosity is 50ab $^{-1}$ which is roughly 50 times as large as the amount of data accumulated by KEKB. Such large amount of data allows us to examine various Flavor Changing Neutral Current (FCNC) processes, such as the radiative decay $b \to s\gamma$, the semileptonic decay $b \to sl^+l^-$, and the hadronic decays $b \to sq\bar{q}$ and $b \to dq\bar{q}$, with unprecedented precision. In addition, new transitions that are currently out of reach will also become accessible.

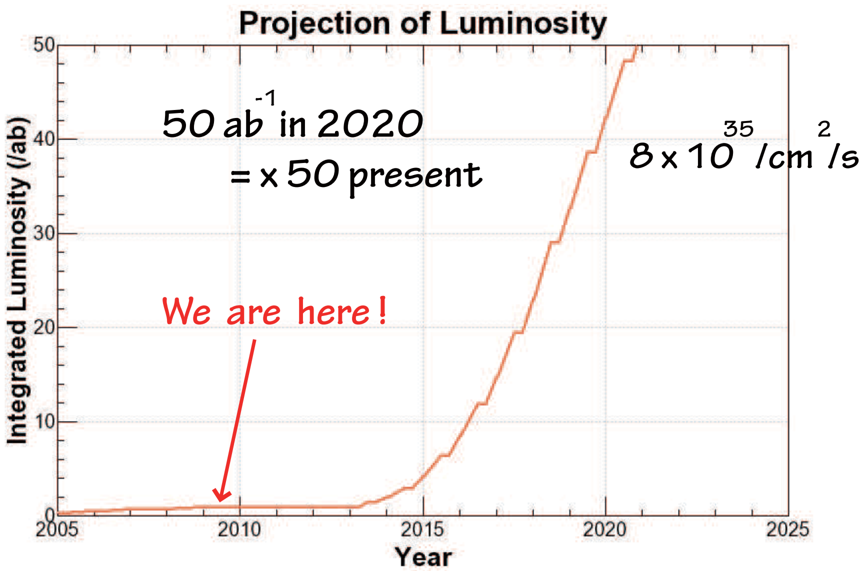

Fig. 1.1 The prospect of the integrated luminosity of the SuperKEKB.

1.2 Current Belle Computing System

Figure 1.2 shows a concept of the computing model for the current Belle experiment. As can be seen, we adopted a centralized computing system wherein almost every service is performed at the computers in KEK. Under this situation, we have to process beam data without any delay for the online data acquistion and reprocess it whenever we update reconstruction software as well as calibration constants.

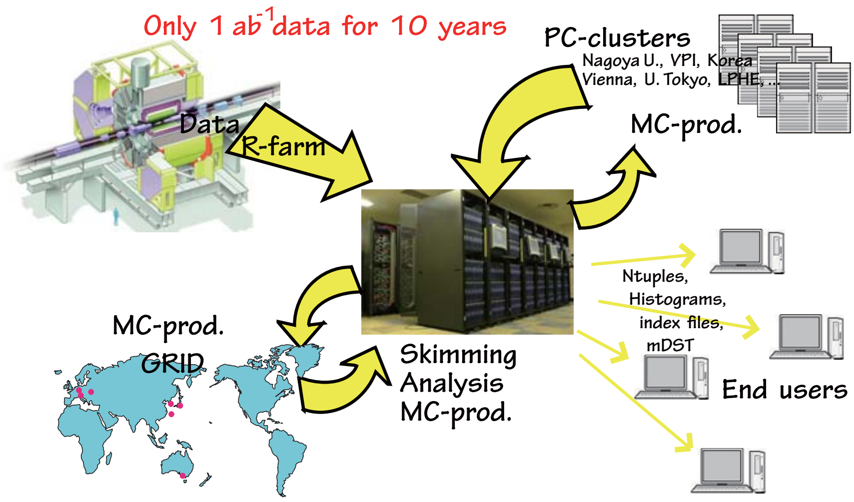

Fig. 1.2 The concept of the computing system for the current Belle experiment.

At the beginning of the experiment, the integrated luminosity per day was several 100pb^{-1}. As time advanced, thanks to the excellent operation of KEKB, the integrated luminosity per day has been approaching 1500fb^{-1} (as of July 2009). To cope with these increasing demands for the data processing and analyses, the computing resources in KEK have been updated several times as shown in Table 1.1

Table 1.1 History of computing resources for Belle experiment

	1999-	2001-	2006-	2009-
CPU [SI2k]	∼100	∼1200	∼ 42,500	∼ 115,200
Disk [TB]	4	9	1,000	1,500
Tape [TB]	160	620	3,500	3,500

For the MC data, we require an amount of data, at least 3 times larger than beam data to evaluate systematic uncertainties due to detector geometrical acceptance, reconstruction procedures and so on. Basically the KEK computers contribute to the MC production as well as the data processing. To augment this resource, we asked some institutes in the Belle collaboration to take part in the MC production with PC farms located in each site. Furthermore, since 2007 we have also started the Belle Virtual Organization (VO) to use the grid resources for the MC production.

For the current Belle experiment, this centralized computing system has worked smoothly.

1.3 Belle II Computing System

As mentioned in 1.1, the Belle II computing system has to handle an amount of data eventually corresponding to 50 times the current Belle level. However, we assume a two-step upgrade for the computing system, dictated by the period of rental agreement for the computing resources. The first step aims to satisfy the requirements for the computing resources at around 2015 and the next step for 2020. Even for the first step, this is a challenging attempt in terms of the CPU power and the storage capacity.

To establish the Belle II computing design, we need to check the expected event rate and data size and also have to consider how to handle these data.

1.3.1 Expected Event Rate

The expected event rate can be roughly calculated from the instantaneous luminosity and the event production cross section. The SuperKEKB accelerator will be operated at energy of the $\Upsilon(4S)$ resonance ($= 10.58\mathrm{GeV}/c^2$) and the $b\bar{b}$ production cross section at this energy is roughly equal to 1nb. Though the target instantaneous luminosity is $8 \times 10^{35}/\mathrm{cm}^2/\mathrm{s}$, the SuperKEKB accelerator will be operated with a lower luminosity at the early stage of the experiment and then gradually increase the luminosity.

In this estimation, we assume that the instantaneous luminosity at around 2015 to be $2 \times 10^{35}/\mathrm{cm}^2/\mathrm{s}$ and the running-time of the accelerator to be roughly two thirds of a year. The rate for $b\bar{b}$ event is calculated to be ~ 200 Hz and the number of $b\bar{b}$ event is to be 4×10^9 events per year. The current Belle has accumulated the integrated luminosity $\sim 950\mathrm{fb}^{-1}$ for 10 years operation, and this corresponds to roughly 1×10^9 $b\bar{b}$ events. That is, a 4 times amount of the entire Belle data will be accumulated in a year even in the early stage of the experiment.

Table 1.2 Cross section for each physics event

	cross section (nb)
$b\bar{b}$	$\sim$1
$\tau^+\tau^-$	$\sim$1
$q\bar{q}$ (continuum)	$\sim$3

However, $b\bar{b}$ events are not the only event in the Belle II experiment. We need to take into account other $q\bar{q}$ events as well as $\tau^+\tau^-$ pairs. Table 1.2 shows the cross section for each event type. Only for these physics events, we need to handle the order of the 10^{10} events per year.

Furthermore, it is very hard to reduce background events due to lost beam particles, beam-gas interaction and synchrotron radiation. Because the beam background estimation is currently being done, we do not quote its value at this moment. Assuming the purity of the $b\bar{b}$ events in the current Belle experiment, the number of events could easily increase by the beam background effect by a factor of three or more.

1.3.2 Expected Event Data size

The Belle II detector will be a large-solid-angle magnetic spectrometer that consists of a pixel-type vertex detector (PXD), a four-layer silicon strips tracker (SVD) and a central drift chamber (CDC) as tracking devices, a time-of-propagation (TOP) counter and aerogel rich (A-RICH) counters for particle identification in the barrel and the forward endcap respectively, thallium doped or pure CsI crystals as an electromagnetic calorimeter (ECL), and a K_L^0/μ identifier (KLM).

The data from each detector is collected by data acquisition system (DAQ) and passed to an event builder. The present estimation of the event size for each detector is summarized in Table 1.3. Because the detailed design of each detector has not been fixed yet, the values are derived from these for the current Belle experiment. The total data size excluding the PXD detector is expected to be $40-70$kB. For PXD, the event data size will likely exceeds $1,000$kB, which DAQ may not handle. Therefore the PXD group is making effort to reduce the data size. In this estimation, the total event size is assumed first to be $\sim$ 50kB without PXD detector.

Table 1.3 Breakdown of event data size

detector	event size (kB)
PXD	$\sim$ 1000
SVD	8
CDC	18-36
TOP	1-6
A-RICH	4-7
ECL	12
KLM	2-4
Trigger	unknown
Total	40-70 (w/o PXD)

From 1.3.1, the raw data size for the early stage of the experiment is calculated to be $\sim$ 6PB per year including the beam background effect, which is empirically estimated from the present Belle. In the data processing, we also make skimmed datasets by applying loose event selection cuts, so-called physics skims, so that the users can analyze the data without processing the whole set by themselves. The same skim process is applied to MC events, too. The event size for a skimmed event

for both data and MC events is estimated to be 30kB. Including the raw data and skimmed data, totally $\sim$ 15 PB of storage capacity is required per year.

Here, the raw data size for PXD is not taken into account in this estimation. Furthermore, the instantaneous luminosity will increase gradually up to $8 \times 10^{35}/\mathrm{cm}^2/\mathrm{s}$ by 2020. To ensure the Belle II computing, serious efforts to reduce the data size are anticipated.

1.3.3 Possible Computing Model for Belle II

As discussed in 1.3.1, we will have an amount of raw data of the order of 10^{10} events. We need to process these data without any delay of the data acquisition as well as the production of MC events corresponding to at least 3 times of beam data. Moreover, we have to provide CPU power for physics analyses. If we still keep using the centralized computing system for Belle II, we must face serious problems, for example, network traffic surge at KEK, heavily-loaded servers, and so forth. To avoid these problem, we will adopt a distributed computing design based on the Grid.

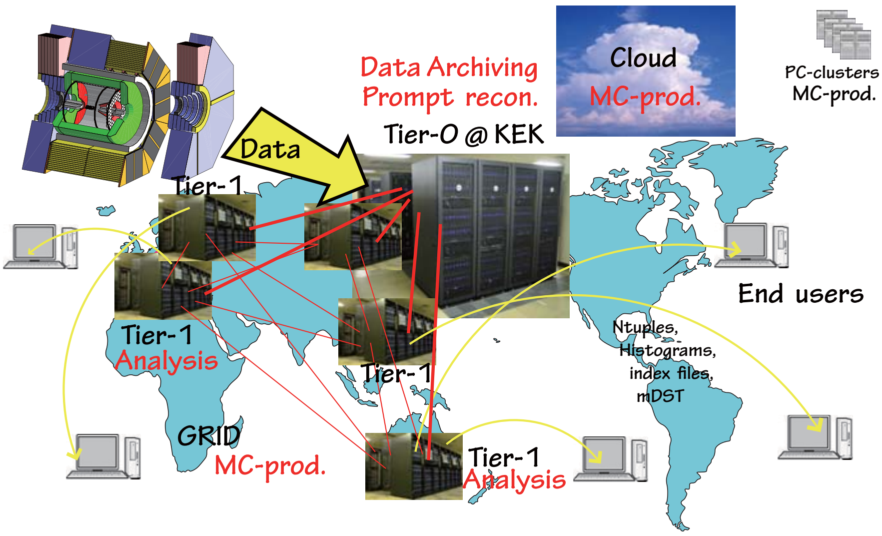

Fig. 1.3 The concept of the computing system for the Belle II experiment.

In this design, the main computer facility, KEK, and the remote computer facilities in the world will play different roles. KEK is a core site for data processing of beam data and the raw/skimmed data archive. KEK also contributes a fraction of the MC production, but the main part will be done at the remote facilities. The produced MC data is skimmed at each site and then skimmed datasets are copies to

KEK. Every analysis user can access to the nearest computer facility to analyze the datasets through the data handling system explained later. The concept of the Belle II distributed computing model is depicted in Fig. 1.3.

In parallel with the Grid-based computing system, we are considering to incorporate commercial Cloud Computing facilities, such as the Amazon Elastic Computing Cloud (EC2), for the MC production. So far, the computing design has been done so as to meet the requirements for the computing resources at the peak demand. However, it is a great cost to own and maintain the resources not used in off-peak time, and it might be possible that the total cost can be reduced, when commercial cloud computing is partially incorporated. Therefore, we have already started to test MC production on Amazon EC2 to estimate benefits.

1.4 Data Handling System for Belle II

The data handling system is very essential to enable users to effectively and smoothly analyze datasets. On the other hand, the available computing resources for Belle II are limited and the schedule for preparation of the computing system is rather tight. Therefore, it is necessary to design a technologically-feasible data handling system.

The basic policies are, using concepts of metadata and projects already adopted in D0/CDF experiments, keeping the system as simple as possible, and using grid (gLite) services as much as possible. According to these policies, the analysis sequence and data flow are proposed as follows (also shown in Fig. 1.4);

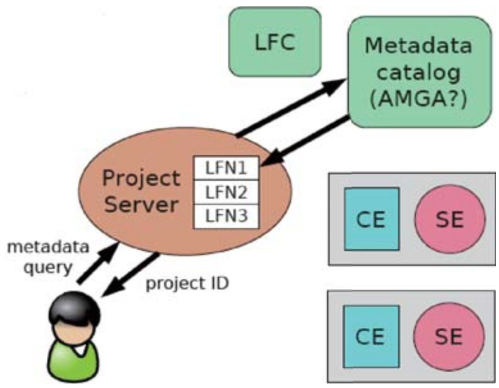

Fig. 1.4 Proposal for the data handling system. CE : Computing Element, SE : Storage Element, LFG : LCG File Catalogue, AMGA : ARDA Metadata Grid Application and Project server for arranging the delivery of the set of files.

1. User starts a project using a metadata query
2. Project server accepts this query and translates it to a list of files with consulting the metadata catalogue server. Then a unique project ID is generated by the Project server and sent back to User. Project server also sets up a database to store the project ID and the list of files.

3. User obtains a project ID from the Project server and submits a grid job.
4. Project server obtains a request for files from the grid jobs based on its project and job ID. And finally it returns a physical file name to the job. Then the job starts running.
5. When the job finishes, it notifies the Project server. Then User can get the output at the local storage element.

This is the basic concept for the data handling system. In order to realize this, it is required to make a test bed and examine the feasibility of this system as soon as possible. For the metadata catalogue, ARDA Metadata Grid Application, AMGA, is employed and the KISTI group has begun adapting it to our usage. Furthermore, a group in University of Melbourne has made a prototype of the project server and confirmed that it works up to the item 2 in the above list.

Though the feasibility check is still going on, we have been advancing steadily toward the establishment of the Belle II distributed computing system.

1.5 Summary

To search for the new physics beyond the SM that may provide us a hint for the matter-dominated Universe, SuperKEKB, aiming at the instantaneous luminosity of $8 \times 10^{35}/\mathrm{cm}^2/\mathrm{s}$ and the tolal integrated luminosity $50\mathrm{ab}^{-1}$ by 2020, is being planned. We will start the Belle II experiment with the SuperKEKB accelerator from 2013, and the expected amount of data will be comparable to LHC experiments even in the early stage of this experiment. Furthermore, we have only 3 years to prepare the computing system. This means that it is necessary to design a technologically-feasible computing system including the data handling scheme as soon as possible. For the computing system, we propose the grid-based distributed computing system incorporating commercial cloud computing. We are also considering integrating the concept of the project server which has been applied in D0/CDF experiments as a data handling system. We are now testing the idea of the incorporation of commercial cloud into the data handling system. We hope that the distributed computing design will be fixed soon after doing the more detailed tests.

Acknowledgements We acknowledge technical support from Institut für Experimentelle Kernphysik in Universität Karlsruhe, Institute of Particle and Nuclear Studies and Computing Research Center in High Energy Accelerator Research Organization (KEK), the NSDC and HEP teams in Korea Institute of Science and Technology Information, Korea University, H. Niewodniczanski Institute of Nuclear Physics in Krakow, J. Stefan Institute and University of Nova Gorica in Slovenia, University of Melbourne, Max-Planck-Institut für Physik in München, National Central University in Taiwan and Charles University in Prague.

References

1. M. Kobayashi and T. Maskawa, *Prog. Theor. Phys.* **4**9, 652 (1973)
2. KEKB B Factory Design Report No.95-1, (1995).
3. A. Abashian *et al.* (BELLE Collaboration), *Nucl. Instrum. Meth.* **A** 479, 117 (2003).
4. K. Abe, *et al.*(Belle Collaboration), Phys. Rev. D. **6**6, 071102(R) (2002).
5. S. K. Choi *et al.* (BELLE Collaboration), *Phys. Rev. Lett.* **9**1, 262001 (2003).
6. B. Aubert *et al.* (BaBar Collaboration), *Phys. Rev. Lett.* **9**5, 142001 (2005).
7. K. Abe *et al.* (BELLE Collaboration), *Phys. Rev. Lett.* **9**8, 082001 (2007).
8. S. K. Choi *et al.* (BELLE Collaboration), *Phys. Rev. Lett.* **9**4, 182001 (2005).
9. S. K. Choi *et al.* (BELLE Collaboration), *Phys. Rev. Lett.* **1**00, 142001 (2008)
10. S. W. Lin *et al.* (BELLE Collaboration), *Nature* **4**52, 332 (2008)
11. K. Abe *et al.* : http://superb.kek.jp/documents/loi/img/LoI_physics.pdf,
http://superb.kek.jp/documents/loi/img/LoI_detector.pdf,
http://superb.kek.jp/documents/loi/img/LoI_accelerator.pdf