

Computing system at Belle II experiment

Kiyoshi Hayasaka¹

*Kobayashi-Maskawa Institute for the Origin of Particles and the Universe, Nagoya University
Chikusa-ku Furo-cho, Nagoya, Japan*

E-mail: hayasaka@hep1.phys.nagoya-u.ac.jp

We make a brief introduction for computing system at Belle II experiment, that will start 2015 and is one of big projects for KMI. Also we discuss our contribution from KMI Tau-Lepton Data Analysis Laboratory: We have built a grid site in our laboratory. This is a first site from Japanese university to join Belle II computing and our site is only one operated by the scientists among Belle II computing sites over the world. We consider this is a big advantage because many problems on the site can be investigated, understood and solved by ourselves and this knowledge helps to construct the Belle II computing system since it is still being developed. We have made big contributions for CPU resources at simulated-event production tests as well as for the stable operation for the whole Belle II computing.

*KMI International Symposium 2013 on "Quest for the Origin of Particles and the Universe"
11-13 December, 2013
Nagoya University, Japan*

¹Speaker

1.Introduction

After LHC experiments, a computing system is one of important key issues to make a success at collider experiments, similarly to an accelerator and a detector. A “good” computing system can accept, handle and store high-rate signals from so many channels of the detector as well as so many users belonging to a big collaboration. One possible solution for the “good” system is a distributed computing system since recent required resources are too huge to handle one computing center. Roughly say, the distributed computing is a technique to make computing resources over the world look like one computing system where users do not need to care where their job can be executed because the system automatically finds a suitable resource.

2. SuperKEKB/Belle II and its computing system

The SuperKEKB/Belle II is a next-generation B factory.[1] It aims the 40 times higher luminosity than that of KEKB/Belle with two times stronger beam current and twenty times narrower beam size. Finally, Belle II experiment will correct 50 ab^{-1} data sample. According to the estimation, this requires 500k HEP SPEC cpu resources, 200TB hard disk and 200TB tape in total. Here 500k HEP SPEC cpu resources roughly corresponds to 50,000 cores. Thus, Belle II experiment also introduces the distributed computing system to handle these huge resources. The structure of the Belle II computing system has just two layers, where upper and lower layers mean main and local sites, respectively, as shown in Fig. 1. KEK and Pacific Northwest National Laboratory (PNNL) become main sites and others are local sites.

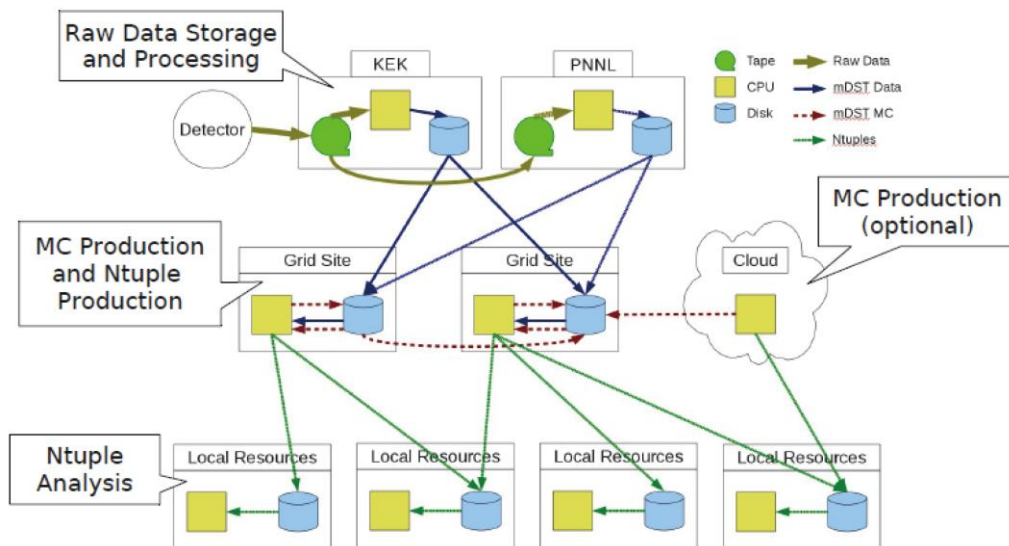


Fig. 1 Layer structure of Belle II computing system. Upper two layers are handled by the distributed computing system.

Since the institutes and universities utilize different grid systems, we need a middleware to handle them seamlessly. We employ the middleware called DIRAC (Distributed Infrastructure with Remote Agent Control) [2], that is originally developed for LHCb experiment, with our original add-on, called “Belle-DIRAC” to handle Belle II computing system. In addition, we also develop “gbasf2” (on Grid, Belle2 Analysis Software Framework). [3] Usually, we make analyses using “basf2” (Belle2 Analysis Software Framework), which is also now being developed. The same function as that of “basf2” on the distributed system is provided by “gbasf2”. By using this, usual users can utilize our distributed computing system for their analysis without any special knowledge for the computing. Now, we have many computing site in Europe and North America while only a few site in Asia, one site in Korea, two sites in Japan, i.e., KEK and KMI as shown Fig. 2.

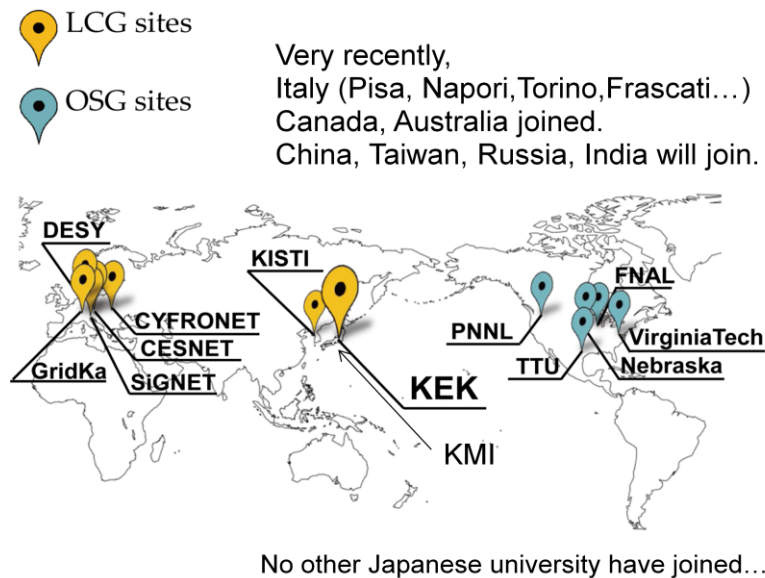


Fig. 2 Sites joining Belle II computing system.

As a world-wide test, Belle II computing group has performed “MC campaign”; By generating huge MC sample, a real-size operation has been tested, some bottle-necks have been searched for and a stability of our software has been checked. We had the first MC campaign at the end of Feb, 2013 and the second one at the end of July, 2013. In the first one, 11 sites (i.e., institute or university) joined, where total CPU resource is around 40kHEP SPEC times day, and 60M $B\bar{B}$ events were generated for three weeks. In the second one, 13 sites joined, where total CPU resource reaches about 700kHEP SPEC times day, 560M $B\bar{B}$ events were generated for 7 weeks.

3. KMI GRID computing system

Tau-Lepton Data Analysis Laboratory has newly built a GRID system, which has 10 worker nodes equipping 2 Intel Xeon E5-2690 (6cores), a storage element whose disk size is 300TB and 6 servers to host virtual machines where server to realize some basic grid service work. By operating the grid services on the virtual machines, when the machine trouble occurs, the grid service can work continuously on the virtual machine by changing the host machine from broken one to the healthy one. Here, as a grid middleware, we use EMI3 provided by European Middleware Initiative, that is taken by the LHC experiment groups and provide grid services such as CE, SE, top/site BDII, ARGUS, WMS, L&B, PROX. At the present, this system is dedicated to Belle II computing.

Our KMI computing system joins Belle II computing from 2nd MC campaign. During the 2nd MC campaign, our system has processed 45,000 jobs. After joining 2nd MC campaign, with one week operation, we have achieved the 3rd biggest contribution, that corresponds to 10% of the total amounts of CPU resources. The reason why our small resources can make such a big contribution is a stable operation. As shown in Fig. 3, KEK (dark purple) and KIT (bright purple) have made big contributions. But, their contribution width as well as that from other sites is unstable. On the other hand, our KMI contribution is very stable except 15th Aug when we had some trouble.

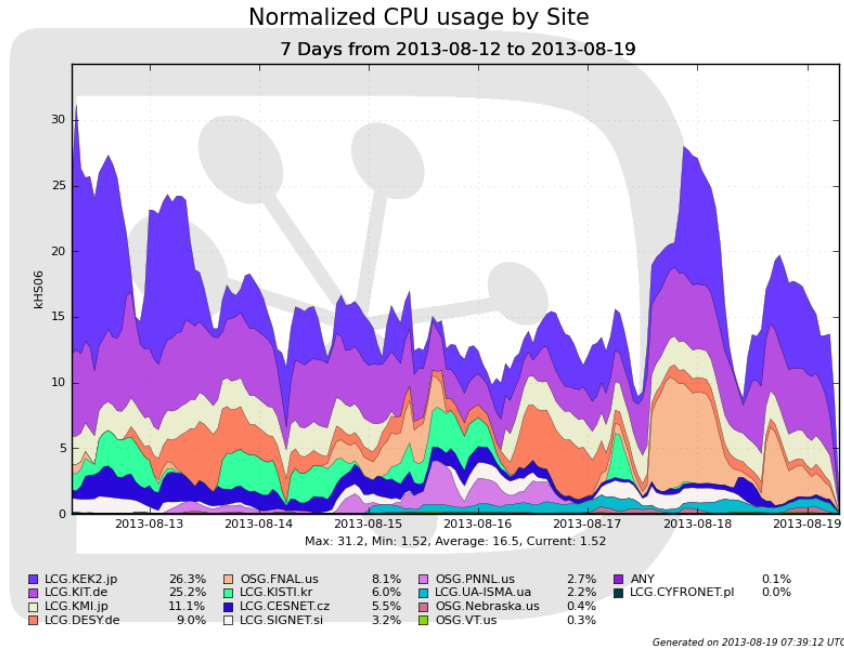


Fig. 3 Contribution of CPU resources for 2nd MC campaign from each site. The biggest one is KEK (Japan), the next one is KIT (German) and the third one is KMI.

We consider that the site construction and operation not by the computer professionals but by the scientist is important because we can investigate any problem by ourselves when the problem occurs. During 2nd MC campaign, KMI reveals the reason of several problems and solve some of them. This helps the stable operation for other sites which increases their computing time. It is also a big contribution to the collaboration since it equals to increase of the CPU resources.

One example we solved is a job failure without any clear reason which is provided by DIRAC. But, information obtained by DIRAC is very restricted. On the other hand, information inside the site, which usually the maintainers do not like to show outside because of the security reason, may have a hint to solve this problem. It is difficult to check such kind of information for the scientists even in KEK while it is possible in KMI grid site. By checking the details of inside information, consequently, it turns out that proxy is expired due to the long wait inside CE. Therefore, by making proxy's lifetime long, this problem is solved. As another example, we found the inconsistency of the number of the running jobs between DIRAC inside and the batch job system. Sometimes new jobs are not submitted due to this problem though many computing nodes have no jobs at all. In this case, we could not find any good solution. On the other hand, we have observed that this problem suddenly disappears. Therefore, as a workaround, by increasing the number of waiting jobs, we have succeeded in not making "no-job-running time" on the site.

4. Summary

Nowadays, computing system is one of important elements for high energy experiments and the distributed computing is an only solution to realize the system having huge resources. Belle II experiment also introduces the distributed computing and is now developing the system. KMI is the first site from

Japanese universities to join Belle II computing and plays an important role on the Belle II computing system. In particular, KMI is only one site which can provide the detailed information of various troubles as well as the most stable site. This is achieved by building and operating the site by scientists.

References

- [1] T.Abe et al, KEK Report 2010-1, arXiv:1011.0352
- [2] S K Paterson and A Tsaregorodtsev, J. Phys.: Conf. Ser. **119** (2008) 062040
- [3] A. Moll, J. Phys.: Conf. Ser. **331** (2011) 032024