Operating the Belle II Collaborative Services and Tools

Andreas Gellrich[1], *Daniel* Knittel[1] , and *Thomas* Kuhr[2]

[1]Deutsches Elektronen-Synchrotron, DESY, Hamburg, Germany
[2]Ludwig-Maximillians Universität, LMU, Munich, Germany

Abstract. In 2016 the Belle II collaboration reviewed and consolidated their collaborative services and tools. Services and tools were migrated into the existing IT infrastructure at DESY. Missing components were added, workflows adapted, and a membership management system was developed. Since the Belle II collaboration was already alive and working, the collaborative work could not be completely started from scratch. The introduction of new technologies - in particular up-to-date security mechanisms - required users to get used to new systems, workflows, and look-and-feels. In this contribution we review the migration process and thoroughly discussed experiences which we gained while supporting the daily work of the Belle II collaboration.

1 Introduction

The Belle II experiment at the SuperKEKB e+ e- collider at KEK in Tsukuba, Japan is the successor of the Belle experiment. Belle II started data taking in 2019 and plans to run until the end of this decade, aiming on collecting an integrated luminosity of 50 ab^{-1}. As of July 2020 the Belle II collaboration consists of more than 1000 scientists from ~100 institutions from 23 countries in 4 continents covering 19 time zones.

In the first part of this paper we describe the current status of Belle II collaborative services. In the latter part we summarise experiences gained during the migration process from the initial services and tools hosted at KEK, as well as experiences from almost four years of operating the new services.

2 Collaborative Services and Tools

2.1 Motivation

Collaborative services and tools are needed to efficiently communicate within a global virtual community such as Belle II. They facilitate the sharing and exchange of relevant information among the members as well as presenting selected information to the public mainly through web services.

In 2016 the Belle II collaboration decided to consolidate their collaborative services and tools to meet the requirements for at least a decade of experimental work in high energy particle physics (HEP). As a result of a thorough review all services and tools were migrated into the existing IT infrastructure at DESY which already supports various experiments and groups such as the European XFEL, CFEL, and other HEP experiments at DESY. Most of the services and tools were subject to modifications and adaptions to meet new demands in technology and security. Missing services and tools were added or developed. Details on the concept and the migration process were already presented in a poster session during CHEP 2016 in San Francisco and are available in the proceedings [1].

2.2 Requirements

The following aspects and requirements were identified and considered essential:

Table 1. Requirements to the collaborative services and tools.

Requirement	Comment
Security	The General Data Protection Regulation (GDPR) of the European Union must be fulfilled. Data and services must be protected against unauthorized access. Authentication and authorization for all users according to law. Centrally controlled updates and maintenance. Secured web services use officially issued certificates. Servers run in a dedicated protected subnet.
Access	Access is restricted to authenticated users (via account/password). User accounts are managed centrally. A membership management system is the source for all user associated services such as mailing lists and author lists.
Legacy	Already existing information must be kept.
Reliability, Stability, Availability	The experiment must be operational and stable at any time. Integrity of data and information is crucial.
Sustainability, Maintainability	Minimise the use of self-developed and self-maintained tools. The collaboration will exist for >10 years. Services embedded into a existing IT infrastructure which is used by other groups too (*synergy*). Usage of state-of-the art technologies.
Flexibility	New requirements demand new features or new services and tools. Self-maintained or self-developed services are run on virtual hosts.

2.3 IT-Infrastructure

The DESY IT service infrastructure is operated to support the scientific computing performed by the various groups at DESY. Among them are WLCG HEP groups, ILC, the European XFEL, and CFEL. By way of the *Atlassian* [3] tool suite a portfolio of collaborative services and tools is provided. Additional services, in particular small-scale self-maintained or even self-developed tools, are operated in Linux-based virtual hosts in a dedicated *XEN*-cluster [4].

Security at DESY

Data and services must be protected against unauthorized access at all times and the General Data Protection Regulation (GDPR) of the European Union must be fulfilled. Users and accounts are managed with an Identity and Access Management (IAM) system which requires individual credentials (account/password) to access services at DESY. The DESY password policy asks for a renewal every 180 days.

All server hosts are regularly scanned for security holes such as rootkits and kernel vulnerabilities. Open network ports are regularly tested and monitored.

Network

Services – in particular web services - which allow access form outside DESY are operated in dedicated subnets with strict port protection by a firewall - a so-called de-militarized zone (DMZ). Web servers are equipped with server certificates and only support secure web access (https).

Systems

Services which are maintained and operated by users, such as Belle II users, run in virtual machines in *XEN*-clusters. For Belle II Linux hosts are offered with centrally-installed, configured, and maintained operating systems (Ubuntu, Scientific Linux, CentOS), running in an DMZ.

Data

Essential files are stored on distributed file systems. In addition *TSM* backup can be configured.

IT services

- Dedicated domain belle2.org and the website www.belle2.org
- DESY user registry (user accounts)
- *Sympa* (mailing lists service)
- *Indico* (agenda service)
- *ZMS2* (website service)
- *NetApp*-based distributed file system

- *TSM* backups

Atlassian tool suite

- *Confluence* (wiki)
- *JIRA* (issue tracker)
- *Stash* (git web UI)
- *Bamboo* (build system)

Additional services

For services and tools which are not part of the DESY portfolio a set up virtual hosts are made available in a *XEN*-cluster [4]. All hosts are operated in a DMZ and are centrally maintained and updated. *TSM* backup for crucial data is available. Web services (Apache2) is available with DFN (*Deutsches Forschungs Netz Verein*) certificates and access to the registry (user accounts) via LDAP. The actual service is then operated and maintained by a Belle II responsible:

- *Invenio* (documents)
- electronic logbook (*eLog*)
- shift tool
- chat service (*Rocket.Chat*)
- questions (*AskBot*)
- monitoring services (*Apache2*)
- software building and validation (*Bamboo*, *BuildBot*)

2.4 Membership Management System

Due to the lack of existing software products a membership management system was developed for Belle II at DESY in 2017. It is in production since January 2018.
The Belle II Membership Management System (B2MMS) plays a central role for the organization and management of the collaboration:

- It keeps track of all users who are connected to Belle II.
- It is the source for all Belle II member related issues such as:
  - mailing lists
  - author lists
  - voting
  - shift lists
  - Maintenance & Operation (M&O) costs

The membership management system implements the following paradigms:

- It manages all persons who have some involvement with Belle II; this includes non-Belle II collaboration members such as externals, summer students etc. who need access to the Belle II collaborative services and tools.
- In order to access the system, DESY login credentials are needed. It acts as an interface to the DESY registry to provide credentials (account/password).
- Person data are in general not public. Only selected information is presented to the Belle II community.
- The institutional representatives (**IR**) in the institutional board (**IB**) play a crucial role. They take care of the users of their institution.
- *Authentication* of persons are done via Grid user certificates or via and identity check by the responsible IR.
- Users can see and edit their personal data. The change of some data, e.g. membership status, requires approval of an authority.
- Since the membership category determines the role in Belle II collaboration which in turn effects the M&O costs; it requires the approval of the IR.
- The registration of new persons is handled by approval workflows depending on the membership category.
- The workflows are set up in accordance to the Belle II by-laws.
- All data including the status can be changed at any time.
- Users must register themselves (by Grid certificate or on invitation) because it is mandatory to accept the DESY usage rules. Without such a permission by the users DESY is not allowed to store personal data.
- It is based on the DESY Identity and Access Management (IAM) (*Apex*)

3 Experiences

The migration to and introduction of a new set of collaborative tools affected both existing, and new users. Existing collaborators had to relinquish systems and workflows they had become accustomed to, and were asked to adopt new ones. In particular colleagues who worked with the predecessor experiment Belle, which was mainly operated at KEK, had to abandon acquired habits. We identified the following topics that will be further discussed below:

- Major changes in access policies and user management.
- Re-implementation of already used services to new technology.
- Introduction of new services.
- System aspects of the IT infrastructure.

3.1 Migration Process

The migration process was started in 2016 immediately in spring after the decision for a change of the hosting infrastructure and the initial set-up of the services at DESY. A hard deadline was set by the KEK computer facility's summer shutdown in August 2016. The goal was to finish the migration and operate fully-functional collaborative services and tools at DESY before August 2016.

Since the Belle II collaboration was already alive and active in 2016 the migration had to be carried out while users were working. Details are discussed in the proceedings of CHEP2016.

3.2 Access policies and users management

Security

All Belle II users have to register with the Belle II Membership Management System to obtain personal credentials (account/password). The access to the collaborative services and tools always requires users to present those credentials at login. The old-fashioned best practice to provide anonymous access or the share accounts and passwords is forbidden.

Membership Management

Since the Belle II collaboration was already active, most users had to go through a registration process to be able to access the new services and tools. Already existing accounts of Belle II members were adopted. Since the beginning of 2018 the membership management system is the exclusive source for all member-based services. The author lists of all published papers of Belle II were deduced from the system. The Belle II member institutions use the system to manage their Belle II users from summer students to emerita. As of mid April 2020 the system contains 1337 user records of which 1044 are listed as collaboration members according to the by-laws. At the time of writing, the last 10 weeks saw ~5000 distinct sessions to search the database, and 300,000 modifications were performed.

User Support

User support is a not linearly-scaling field and person-power and effort tend to be underestimated. Therefore emphasis was put on thoroughly documenting and maintaining at a central location (confluence):

- the procedure to request an account,
- contact addresses for support,
- lists of available services and tools.

Prime user issues are:

- not reading documentation (to the end),
- forgotten account names and passwords,
- using personal email contacts rather than support mailing lists for a helpdesk,

- ignoring multiple requests to change password (every 180 days by DESY policy),
- leaving unmaintained and out-dated information in the wiki.

3.3 Migration of service and tools

Since the collaboration was alive and working before the consolidation of the collaborative services and tools a start from scratch was not possible. A significant amount of information such as a web site, wiki documents and attachments, agendas (*Indico*), issue tracker data, and software had already been generated. On first sight most of the information was considered as valuable and therefore indispensable. All of this information had to be migrated into the new services. Many of them are using different technologies. Moreover, the work of the collaboration went on during the migration process which lasted roughly 3 months.

Table 2. Migrated services

Original	New	Comment
twiki	Confluence	~4000 pages and attachments were copied manually - the by far most time consuming task. The majority of pages has never been access though.
redmine	JIRA	It was decided to start from scratch.
Indico (KEK)	*Indico* (DESY)	Security concerns with non-Belle II data in the KEK *Indico* inhibited export/import of old data. The Belle II groups started new events in the new instance.

3.4 New services and tools

For services which are not in the DESY IT-infrastructure portfolio, virtual hosts for small-scale stand-alone web servers are offered, e.g. detector performance monitors, run, publication, and author databases. The services are installed and maintained by Belle II members whereas system aspects are taken care of by DESY IT. Since 2016 roughly a dozen hosts and services were set up. The concept of sharing system aspects and the service between DESY IT and Belle II has proven to be appropriate to meet the needs of the collaboration.

3.5 System aspects

The utilization of collaborative services and tools of the IT-infrastructure at DESY allowed an immediate start of the migration process in 2016. Most of the Belle II collaborative services and tools are operated in the newly introduced domain belle2.org which is hosted at DESY. In order to protect data and systems, services which need to be accessed from outside the DESY intranet are operated in a special encapsulated part of the DESY network

(DMZ). Only selected ports are opened on the firewall. All web services are equipped with server certificates and allow for secure access. The server hosts are regularly updated. Crucial data are stored on distributed file systems and have backups on *TSM*.

3.5 Lessons learned

As might be expected from an innovative science community new implementations are readily adopted. While the new collaborative services and tools were set up, new ideas and requests came along. From the infrastructure provider's point of view those changes are not always easy to implement if they don't fit into the initial concept. A notable example being that of instant messaging (chat). There are currently two chat instances in use by the collaboration. We also have some evidence of the adoption of new tools at the expense of cleaning up existing ones. In particular more than half of the wiki pages which were migrated have yet to be accessed but remain on the system.

4 Summary and Conclusions

The Belle II collaboration took the unique opportunity to revise and re-think the Belle II collaborative services and tools just before the physics program started. Although the time scale was tight, the migration into a production-grade IT infrastructure at DESY was successfully carried through in 2016. This included the adaption of recent security standards and the introduction of state-of-the-art technologies. A crucial new component, the membership management system, was developed. It allows to efficiently manage the work of more than 1000 Belle II users, it serves as an interface to create DESY accounts, and it is the source for mailing and author lists. Although security is considered as mandatory ingredient by all participants, the handling of individual accounts with passwords makes the work of scientists potentially clumsier. Confirmed habits such as anonymous accounts or shared passwords had to be given up finally. Even more, users must manage their accounts including timely renewing and remembering passwords.

The Belle II management was deeply involved in the planning of the set-up and the migration process. The progress was reported in respective meetings. The Belle II collaboration was kept up-to-date in the general collaboration meetings. Comments and wishes were discussed and included in the developments. The new services and tools were received very positively and are regularly mentioned to be useful and very valuable.

References

[1] N. Braun et al, *Migrating the Belle II collaborative services and Tools*, *Proceedings of CHEP2016*, J. Phys.: Conf. Ser. **898** 102014 (2017)

[2] Belle II website: www.belle2.org

[3] *Atlassian* tool suite: www.atlassian.com

[4] XEN: xenproject.org