

Migrating the Belle II Collaborative Services and Tools

N. Braun², D. Dossett³, M. Dramburg¹, O. Frost¹, A. Gellrich¹, J. Grygier², T. Hauth², D. Jahnke-Zumbusch¹, D. Knittel¹, T. Kuhr⁴, S. Levonian¹, H.-G. Moser⁵, L. Li⁶, N. Nakao⁷, M. Prim², P. van der Reest¹, F. Schwenssen¹, P. Urquijo³, B. Vennemann¹

¹DESY, ²KIT, ³U Melbourne, ⁴LMU, ⁵MPI Munich, ⁶USTC, ⁷KEK

Andreas.Gellrich@desy.de

Abstract. The Belle II collaboration decided in 2016 to migrate its collaborative services and tools into the existing IT infrastructure at DESY. The goal was to reduce the maintenance effort for solutions operated by Belle II members as well as to deploy state-of-art technologies. In addition, some new services and tools were or will be introduced. Planning and migration work was carried out by small teams consisting of experts from Belle II and the involved IT divisions. The migration was successfully accomplished before the KEK computer centre replacement in August 2016.

1. Introduction

Collaborative services and tools are essential for any (HEP) experiment. They help to integrate global virtual communities by allowing to share and exchange relevant information among members, preferably by way of web-based services. Typical examples are public and internal web pages, wikis, mailing lists, issue tracking, meeting organization, document and authorship management, electronic logbooks, as well as build services and code repositories. Moreover, a membership management system is desired.

In order to reduce the maintenance effort for solutions operated by Belle II colleagues and to achieve services at a higher level of stability and reliability, the Belle II collaboration [1] decided to migrate the current set of services into the existing IT infrastructure at DESY [3]. After the approval by the Belle II bodies, the migration process started in early spring 2016. It had to be finished before the KEK [2] computer centre replacement in August 2016.

In the contribution to CHEP2016 we described all aspects of the services and migration processes. Furthermore, we shared experiences, revealed details, and tried to give useful hints for similar approaches.

2. Conceptual Considerations

At the time the Belle II collaboration was founded some collaborative services were inherited from the former Belle experiment at KEK and many of them were newly set up by Belle II members. At the beginning of 2016 it became clear that in view of the replacement of the KEK computer centre a major

consolidation of those services and tools was desired. The following system requirements were defined. The services and tools should be:

- secure,
- highly available, stable, reliable,
- with good performance,
- sustainable,
- based on state-of-the-art technologies and products,
- long-term supported.

Furthermore, the collaborative services and tools should be integrated into an existing IT-infrastructure of a well-established and reliably operated computer centre. Although this would require more migration work and the need for getting used to new products, this concept was favoured over copying the existing services and tools to a cluster of virtual machines located outside KEK. This approach allowed to change a number of paradigms and to introduce new ones:

- No self-registration to services should be possible (if applicable).
- Every Belle II user should have a personal account.
- No shared general accounts (incl. passwords) should be distributed in the collaboration.
- One set of credentials (account / password) per user for all services should be sufficient.
- Services should run in the top-level domain 'belle2.org'.

3. Implementation of Services and Tools at DESY

The DESY IT infrastructure is built to be reliable and highly available. It has been providing sustainable services for many scientific groups such as CFEL [4] and the European XFEL [5] as well as for other High Energy Particle Physics (HEP) experiments. DESY is very conscious of data privacy and enforces strict policies to safeguard data. And DESY is bound to the strict privacy policy legislature in Germany and Europe. Therefore, the DESY IT infrastructure includes fail-over mechanisms, back-up and archiving options and regularly considers security for all services.

Where applicable, user authorization using LDAP against a central DESY authorization service is implemented. Hence all services and tools are accessible for authorized users or groups only.

User registration

DESY allows access to services and tools with individual user credentials (account, password) only. The DESY registry provides such user accounts. The registration procedure contains two steps:

Authentication: Users must authenticate by either using a Grid user certificate or by filling out and signing a paper-form, including a signature of the institute representative.

Authorization: A group responsible must decide whether the user is allowed to gain access.

Website

Web-site services hosted at the DESY are based on the open source content management system ZMS (Simplified Content Modelling) [6] which offers a modern responsive design. Web-site services usually include public pages and an internal part which is exclusively accessible with group credentials. The content can be managed by editors from anywhere.

Collaborative Services

DESY operates an *Atlassian* [7] tool suite which offers professional collaborative services:

Confluence is an enterprise wiki and team collaboration software. Issue tracking is provided by *Jira*. A migration tool for the project management web application *Redmine* [8] is available. Code management and browsing is realized by *Stash*, which is using a GIT [9] implementation. Build services are available via *Bamboo*.

Mailing lists

For mailing lists the open source mailing list manager *Sympa* [10] is used. Access controlled via registry groups. For the list management a separate registration is required.

Document management

Document management is based on the *Invenio Digital Library Framework* [11].

Agenda services

DESY operates multiple *Indico* [18] instances.

Backup

DESY operates a TSM [12] infrastructure to which all relevant services are connected.

Virtual machines

DESY operates a virtualization infrastructure using *XEN* [13]. Virtual machines (VM) can be provided for a number of operating systems used in HEP, e.g. Scientific Linux 6 (sl6), CentOS 7 (el7), Ubuntu 14.4 and 16.4 (ubu144, ubu164). The VMs can be equipped with shared files system as well as backup and archiving services.

4. The Migration

In order to make optimal use of the existing services and support structures at DESY, technology and implementation switches were necessary for some services. See table below. Hence not all services could simply be copied but had to be adapted or even restructured. In preparation for the final migration, small teams of experts from Belle II and the involved IT divisions went through the details of all services to identify problems, potentially develop alternative solutions, and finally carry out the migration. For some new services task forces were founded to develop concepts and prepare and test implementations.

Domain

The Belle II services should be easily visible in the internet. Therefore DESY purchased the domain belle2.org. If applicable, Belle II service now run under the domain belle2.org.

User registration

In order to fulfill the security requirements of DESY as well as of the Belle II collaboration, it is necessary to equip all Belle II members with credentials (account, password) to access the services. For the time being this is based on the DESY registry and provides DESY user accounts for the Belle II members. A final Belle II specific membership management system is under construction (see section 6. New Services).

User authentication is done by either using a Grid user certificate or by filling and signing a paper-form including a signature of the institute representative. For the authorization step the membership of the user and in case of the paper-form of the representative is checked against the Belle II member list and the VOMS server for the VO 'belle'. The registration procedure creates an e-mail address per user such as ['firstname.lastname@belle2.org'](mailto:firstname.lastname@belle2.org) which forwards to the physical e-mail address (users'

preferred mailbox). Moreover, user accounts are associated with groups for a fine-grained access control to services.

As of now 535 of the 743 listed Belle II members have already obtained credentials.

Website

The new Belle II web-site belle2.org was set-up from scratch on the DESY ZMS content management system. Pages were rebuilt using a modern state-of-the-art design. The contents was transferred by a Belle II responsible from the old web pages.

Wiki pages

Most of the information gathered by the Belle II collaboration is kept in a wiki. The migration of the wiki had to be carried out with the main objective to preserve all information.

The *Atlassian* wiki *Confluence* replaces the former *Twiki* [14] of Belle II. The *Twiki* was roughly structured in so-called webs which contain pages of the various Belle II groups such as 'Computing', 'Detector', 'Physics', 'Operations', 'Organization'. Since properly working automatic migration tools from *Twiki* to *Confluence* were not available, around 4000 pages with thousands of attachments had to be copied by hand. Early ideas to successively follow the tree of the *Twiki* turned out to be not successful due to many orphaned *Twiki* pages. Given the limited time, a major clean-up and restructuring of the *Twiki* by Belle II was not feasible. Hence a list of *Twiki* pages was generated and around 500 so-called relevant pages were identified and copied including their attachments. The remaining pages, e.g. shift log pages, were added later-on. For each web a responsible checked the completeness of the relevant pages and tagged them.

The *Twiki* migration to *Confluence* was by far the most resource-consuming task of the entire project. Roughly 5 FTE month were spent for this mostly manual work.

Issue tracking

The *Atlassian* issue tracking service *Jira* offers a plug-in to directly migrate the content of the project management web application *Redmine* which had been used by Belle II. Hence the migration was rather straight forward.

Code repository and browsing

The *Atlassian* tool *Stash* is based on *GIT*. After intensive discussions, Belle II decided to migrate their distributed version control system from *SVN* [15] to *GIT*.

Although tools exist to convert a *SVN* to a *GIT* repository they could not be applied blindly because the feature of subdirectory checkouts that was used extensively before is not available in *GIT*. Therefore, a careful restructuring of the code into separate repositories was carried out and coordinated with the users. Issues such as user mapping and the correct migration of branches and tags had to be solved. The total size of the repository turned out to be another issue and was partially addressed by removing large files from the history. Several established procedures and workflows, including for example commit hooks and release procedures, could not be mapped one-to-one from *SVN* to *GIT* and had to be adjusted. Overall the code repository was the conceptually most difficult part of the migration.

Software build services

Belle II used and uses *Buildbot* [16], an open-source framework for automating software build, test, and release processes. A number of virtual machines were set up at DESY using *XEN* as master / slaves for the operating systems Scientific Linux 6 (sl6), CentOS 7 (el7), Ubuntu 14.4 and 16.4 (ubu144, ubu164).

DESY offers the central build service of the *Atlassian* tool suite *Bamboo*, for which a dedicated build agent was set up which is being used heavily.

Mailing lists

For the Belle II mailing lists a separate server instance lists.belle2.org was set up based on the open source mailing list manager *Sympa*. Since *Sympa* had been in use at KEK, the migration could easily be carried out. In addition, lists of the second mailing list service *FML* [17] hosted at KEK had to be handled. In the course of the migration, a clean-up of unused mailing lists as well as some renaming of lists were done. The global collaboration members mailing list was migrated last. It is still based on a dump of the official Belle II members list of KEK since the final membership management service is not yet in place (see section 6. New Services).

Document management

The Belle II document management system was and is based on the *Invenio Digital Library Framework*. Since no technology switch was needed, the migration was straight forward and was one of the first migrated services. Access control is based on the user credentials. The service is located at docs.belle2.org.

5. Further Services

Agenda management

As in many other scientific laboratories, *Indico* [18], the open source tool for event organization, archival and collaboration, was and is used by Belle II. As of today, the migration from the KEK based system to a separate instance indico.belle2.org hosted at DESY is not finished. Since the KEK system also contains entities of non-Belle II events, a clean-up and pre-selection is necessary by experts at KEK. Work is on-going and is expected to be finished in spring 2017.

6. New Services

There are two services which were not centrally deployed for Belle II before:

Electronic logbook

So far, different groups in Belle II deployed different electronic logbook tools. A task force is currently evaluating tools for a common Belle II set-up. A virtual machine was set up which could host the electronic logbook service under elog.belle2.org.

Membership management system

So far, the Belle II member list was based on a dump from the KEK registry. Therefore, the membership in the Belle II collaboration requires a KEK account and is based on an association with Belle II. Further management of roles etc. in Belle II is handled elsewhere; partly in text files.

In order to coordinate the Belle II collaboration and to manage roles and permissions of its members, a suitable member management system is needed. A task force was founded which collects use cases and investigates possible solutions to be implemented by DESY IT.

The system shall provide all services and tools to introduce, change, and remove persons and institutes of the Belle II collaboration. It will contain portals for users and delegates and it will interface to the registry at DESY. It deploys all data security aspects in particular of personal data.

For security reasons the relational database backing the system is not directly connected to the web portals and resides in a DMZ behind the DESY firewall.

The new registration procedure will change order compared to the former approach. The Belle II membership is independent of accounts e.g. at KEK or DESY and is explicitly handled by the new

membership management system. Further applications for accounts could then be based on the Belle II membership since authentication and authorization steps are already carried out.

The membership management service should at least cover the following tasks:

- Membership life cycles.
- Requested membership.
- Active membership.
- Expiration of membership.
- Institutes joining/leaving Belle II.
- Responsibilities of members and their roles.
- Special management of students and associated non-Belle II users.

The membership management service may also be used for:

- Authorship lists.
- Voting procedures.
- Reporting to boards and member institutes.
- Information needed for internal structures and processes, e.g. policies.

7. Summary and Conclusions

In order to successfully conduct a scientific experiment such as Belle II at KEK, services and tools must be made available to share and exchange relevant information among members, preferably by way of web services. Therefore, stable, reliable, and secure collaborative services and tools are essential for the work of an international collaboration.

The Belle II collaboration decided in spring 2016 to consolidate and migrate their collaborative services and tools into the existing IT infrastructure of DESY with the goal reduce the maintenance effort for solutions operated by Belle II members. All mission critical services and tools were successfully migrated before the replacement of the KEK computer centre until end of July 2016.

Taking into account the tight time-frame with a fixed date of the KEK computer centre replacement on August 5th, 2016, the preparation and test periods before the actual migration had to be reasonably short. The assignment of small teams for the various sub-projects turned out to be a well working approach. It was clear from the start though that the difficulties and problems would differ considerably amongst the services and tools. As pointed out, the migration of the wiki was by far most resource consuming since a technology switch from *Twiki* to the *Atlassian* wiki *Confluence* had to be carried out. In retrospect it would have been more efficient to thoroughly clean up the original *Twiki* rather than copying everything. Such a clean-up could have involved a rigorous revision of the structure and a detailed planning of the migration steps. Of course, both would have needed time.

The Belle II experiment is the successor of the triumphant Belle experiment at KEK after the upgrade of the electron-positron storage ring SuperKEKB. Belle II incorporates many new institutes and universities from all over the world and became a truly international collaboration with more than 700 members. The Belle II collaboration does not only span over 20 time-zones from Melbourne via Asia, Europe, and America to Hawaii, but also incorporates many different cultures. All this had to be taken into account when setting up an IT system for the benefit of the entire collaboration.

The Belle II members realized in the course of migration services that the IT experts at DESY could and can only provide the service infrastructure and help out with hints and examples. The actual work concerning the contents had to be and must be carried out by the collaboration members themselves. It must be stressed though that the Belle II members have done a great job in filling their

new collaborative services and tools with contents while preparing the experiment for the start of the data taking in 2018.

An overview of the services and tools with their former and current technologies after the migration is shown in the appendix. As described before, one service has not yet been fully migrated and two services are still under construction.

Appendix

Service / Tool	Former	Current	Link	Status
web-site	<i>Apache</i>	<i>ZMS</i>	www.belle2.org	done
wiki	<i>Twiki</i>	<i>Confluence</i>	confluence.desy.de	done
issue tracking	<i>Redmine</i>	<i>Jira</i>	agira.desy.de	done
version control	SVN	<i>Stash / GIT</i>	stash.desy.de	done
build system	<i>Buildbot</i>	<i>Bamboo / Buildbot</i>	bamboo.desy.de https://b2-master.belle2.org/buildbot/	done
mailing lists	<i>Sympa / FML</i>	<i>Sympa</i>	lists.belle2.org	done
document service	<i>Invenio</i>	<i>Invenio</i>	docs.belle2.org	done
agenda service	<i>Indico</i>	<i>Indico</i>	indico.belle2.org	ongoing
electronic logbook	various, incl. <i>elog</i>	<i>eLog</i>	elog.belle2.org	ongoing
membership management	KEK registry / plain file	<i>IAM</i>		ongoing

References

- [1] <https://belle2.org/>
- [2] <http://www.kek.jp/>
- [3] <http://www.desy.de/>
- [4] <https://www.cfel.de/>
- [5] <http://www.xfel.eu/>
- [6] <https://www.zms-publishing.com/>
- [7] <https://www.atlassian.com/>
- [8] <http://www.redmine.org/>
- [9] <https://git-scm.com/>
- [10] <https://www.sympa.org/>
- [11] <http://invenio-software.org/>
- [12] <http://www.ibm.com/>
- [13] <https://www.xenproject.org/>
- [14] <http://www.twiki.org/>
- [15] <https://subversion.apache.org/>
- [16] <http://buildbot.net/>
- [17] <http://fml.org>
- [18] <http://indico-software.org>