B-factory Programme Advisory Committee
Full report for
Focused Review Meeting

8 and 9 November 2021 Remote Meeting

A. Andreazza* (Milano), P. Collins* (CERN), G. Corti (CERN),
M. Demarteau (ORNL), R. Forty (CERN), B. Gavela (Madrid),
S. Gori (UCSC), W. Hulsbergen* (NIKHEF), M. Ishino (Tokyo),
V. Luth (SLAC), P. McBride* (FNAL) P. Mato* (CERN),
F. Meijers* (CERN), N. Neufeld (CERN), K. Oide* (KEK),
B. Ratcliff* (SLAC), A. Petrov* (Wayne State University),
M. Sullivan (SLAC), H. Tajima (Nagoya), M. Titov* (Saclay)
and chaired by T. Nakada (EPFL)

* Expert member.

30 December 2021

1 Short summary

This focused review meeting of the B-factory Programme Advisory Committee (BPAC) took place, again remotely, with the following questions posed by the director of the Institute of Particle and Nuclear Studies:

- Long Shutdown 1 (LS1) schedule
  Is the proposed schedule for LS1 appropriate? Point out if there are any particular concerns in the preparation and execution of the detector replacement work.
- Yearly Plan
  Are the operation and study plans for machine and detector in coming months reasonable to achieve maximal integrated luminosity?
- Selected items
  - Progress in ageing study of CDC
  - Development plan of entire L1 system towards $\times 10$ luminosity
  - Improvement plan of data production and processing, especially on the recalibration
  - Solidifying software validation process

– Plan of fail-safe data management in the distributed computation

This section addresses those questions and the following sections describe more general findings by the committee in detail.

LS1 was originally planned to start in July 2022 to replace the beam pipe section around the interaction point, the Pixel Vertex Detector (PXD) and a part of the photon detectors (conventional MCP-PMTs) for the barrel particle identification system (TOP) with an ageing resistant type (life-extended ALD) MCP-PMTs. During the meeting, the committee was informed that LS1 would have to be postponed due to the delay in the production of the replacement beam pipe and uncertainties in securing foreign detector experts since the COVID restrictions could still be in place. The two scenarios presented were:

1. to start LS1 in January 2023, i.e. with a half year delay and

2. to start in July 2023 with a delay of one year.

Reported progress on the preparation shows that required hardware components will be ready for for the start of shutdown in January 2023, although the schedule is tight. The expected integrated luminosity for this scenario varies between $\sim$610 fb^{-1} and $\sim$820 fb^{-1}, depending on whether extrapolating conservatively the SuperKEKB performance or taking aimed machine improvements into account. The second scenario with one year delay is thus motivated by safely securing a data sample before the start of LS1 that would allow physics analyses beyond those by Belle, which collected a data set of $\sim$800 fb^{-1} for the $\Upsilon(4S)$ and off-resonance. The importance of this ambition is well appreciated by the committee. On the other hand, difficulties in the hardware work for LS1 due to diminishing technical expertise caused by the retirement of key persons is already noticeable. The longer the delay of LS1, the more critical this situation becomes. For the one year delay scenario, the work for replacing MCP-PMTs is planned during the regular 2022 summer shutdown starting in July with a two month extension, rather than waiting for LS1, in order to safeguard against ageing effects. The recently observed in situ behaviour of TOP MCP-PMTs indicates that efficiency drops are developing even for the ALD type MCP-PMTs, which have been observed as being more ageing resistant than the conventional ones in the accelerated laboratory test. Before replacing any MCP-PMTs, understanding the cause of the observed efficiency loss for the various MCP-PMT types, and whether it is related to ageing of the photocathodes, is essential. Starting the replacement work in July 2022 may be too early to prepare an updated plan for the MCP-PMT replacement taking this new information into account. An additional incident reported was a cold leak in the final focusing superconducting magnet (QCSR) and an in situ repair might not be possible. In this case, the repair would have to wait for LS1. Although this leak currently does not impact the SuperKEKB operation, this might change in the coming months. Considering those facts, **the committee recommends that the Belle II collaboration adopt the half year delay for LS1 as the baseline.** All the LS1 preparation work should be planned and executed according to this baseline. Successful completion of LS1 is crucial for the long term goal of the Belle II experiment. The committee also supports to assess the situation and to update the plan, if necessary, in May 2022. A detailed report during the next BPAC meeting on the progress for the LS1 preparation and SuperKEKB luminosity performance will be highly appreciated. It is still the beginning of data taking and the collaboration should make the utmost efforts to preserve technical expertise.

For the detector and machine planning before the start of LS1, collecting high quality data as much as possible is clearly important. However, maximising the integrated luminosities for this period should not be the highest priority. The stated goal of Belle II is to produce world leading results based on very large and unique data samples. This will require further improvements in the detector operation and data taking, as well as background mitigation and data processing. In parallel, further detailed machine studies and repairs and replacements of ageing components of the accelerator complex are needed to obtain stable machine operation before LS1. Beam currents should be increased while the backgrounds must be kept under control. Therefore, **the committee strongly supports devoting significant effort to machine studies during the ongoing run period.** Shortage of personnel in the machine group to perform the machine operation and studies in parallel remains a serious concern, although an international task force was established for addressing the luminosity issues. Engaging Belle II physicists to temporarily work on machine issues might significantly enhance the progress.

The committee considers the presented plan for the Central Drift Chamber (CDC) ageing studies adequate but sufficient human resources should be provided for the timely execution of the programme. On the other hand, the performance of the CDC is worrisome. After the problem of the dark current in the chamber had been mitigated by introducing H_2O in the gas mixture, CDC has been developing gain losses. Although there is some evidence that this is caused by H_2O, **further in-depth investigations must carried out to fully understand the situation before any drastic decision should be made**.

The presented plan for the future activities of the Level 1 trigger is well worked out and covers all the aspects needed for operating at luminosities well above 10^{35} cm^{-2}s^{-1}, i.e. $\sim$10 times more than now. On the other hand, **the committee is concerned about the lack of human resources to ensure smooth operation and further implementation of the system**. Efforts to simplify the system and to develop commonalities are highly recommended and should reduce the maintenance effort. New resources could become available by attracting new groups for the trigger upgrade to extend the physics scope such as searches for long living particles. However, a careful balance is needed between maintaining the running system and developing upgrades.

Three other important issues are related to the offline activities. A framework based on the distributed computing for the production of simulated data and processing of raw data is in place and working well. Production centres are all functioning well. A major outstanding issue is the slow production of the run-dependent simulation samples, which are crucial for the precision of physics analyses. In addition, delays in recalibration of the raw data are occurring due to the need of its validation by subdetector experts. **Careful coordination with the subdetector groups is required for the planning to ensure speedy recalibration.** A backlog in data processing affects not only physics analysis activities but also the computing resource requirement and should be avoided. The status of the Belle II software is good. **Careful validation of upgrades to external software, such as GEANT4, PYTHIA and EVTGEN, is essential** before being included in a new major release for production. The committee suggests introducing software changes incrementally in shorter intervals, combined with regression tests and potential intermediate software development releases. The committee strongly supports the collaboration's efforts to further introduce and improve automation to avoid human errors and eventually reduce the need for staff. Although automation has already reduced the needs for human resources, **lack of redundancy in key personnel is a concern**.

2 Machine and background

2.1 Status

The accelerator team has accomplished a new peak luminosity of 3.12×10^{34} cm^{-2}s^{-1}, a new world record. The specific luminosity for the operation with $\beta_y^* = 1$ mm has improved. The vacuum pressures in the rings continue to improve allowing for more beam current. There have been some unexpected thyratron firings in the Low Energy Ring (LER) kicker system that may have caused some damage to a collimator. A cold leak to air has developed in the QCSR cryostat. It is currently manageable but worrisome. Improvements in the injection have been and are being made in several parts of the Linac. It should be noted that the Linac and transport lines are servicing the two B-factory rings as well as the two Photon-factory rings. This greatly increases the need for a steady reliable linac system. The committee congratulates the Linac team in accomplishing improvements in positron production. There are still outstanding issues in the injection that need further study and/or improvement: in particular, the emittance blowup in the Beam Transport line (BT) for the High Energy Ring (HER) and the need to increase the positron yield.

The background group continues to make progress in the understanding of the various background sources and their effects on the Belle II experiment. Four out of five storage ring background components now agree with the simulation results. Background rates in each detector subsystem, the PXD, SVD, CDC, and TOP all show good agreement with those from the simulation, while the endcap particle identification system and the K_L^0-Muon detector system (KLM) background rates still differ from the simulation results. The Electromagnetic Calorimeter system (ECL) has not been studied yet. Great progress has been made in further understanding of the injection backgrounds as the simulation results can now match the injection backgrounds seen by the CDC. Backgrounds from local neutron sources have been measured and improved shielding is being designed and built.

2.2 Concerns

- Although the current injection rate is adequate, improvements will be needed to achieve higher stored beam currents and shorter ring lifetime.
- Due to instabilities in injection, where their causes are still not understood, the collimator settings cannot be optimised.
- The positron yield is below the design value.
- The HER injection rate also is low and there is still the emittance increase in the BT to the ring.
- The root cause of the QCSR cryostat cold leak has not been identified.
- The LER thyratron kickers are unexpectedly firing and need to be replaced.
- The Transverse Mode Coupling Instability (TMCI) threshold is lower than originally calculated.
- The root cause of the LER catastrophic beam losses that occur somewhere in the storage ring has not been understood.

2.3 Recommendations

- Further studies on the two-bunch injection for the HER is encouraged, most likely requiring dedicated machine time.
- The committee supports the effort to study the use of the PF-AR transport line as a possible better tunnel for injecting beam into the HER.
- The committee encourages expanding the effort to improve the modelling of the current BT in order to understand where the emittance blowup is occurring; e.g. how much a lattice error needs to be increased in order to induce the observed emittance blowup.
- The positron yield needs to be improved in order to keep up with the expected short lifetime of the LER as beam currents increase.
- A close watch on the leak in the QCSR is very important. When trying to fix the leak, bolts should not be overtightened in order to avoid getting too close to the yield point of the material.
- It is important to control the unexpected thyratron firings as this makes it dangerous to increase the LER stored current. Perhaps one can wire the thyratron trigger such that if any single thyratron fires then it will force all thyratrons to fire.

- The committee also encourages further studies of collimator settings with the aim of improving injection quality and beam lifetime. This will also reduce the risk of damage to both the detector and to the accelerator. These studies could include the option of adding more adjustable and also fixed collimators that are much less expensive and use less space. They need to be closely associated with thresholds for transverse mode coupling instabilities.
- The catastrophic beam losses remain an issue and the committee highly encourages the background team's efforts to further narrow down the possible location of the source.
- In summary, the committee is convinced that addressing the previous three recommendations, especially increasing the LER current and stability, will be critical to make major progress towards significant improvements of the machine performance.

3 Selected hardware issues

While the performance of the Belle II detectors has generally been stable and efficient since the previous BPAC meeting, two issues concerning the detector hardware were raised: first, the gradual reduction in the gain of the CDC signals, and second, the reduction in the signal efficiency observed for the MCP-PMTs of the TOP.

3.1 Central Drift Chamber operation and ageing study

3.1.1 Status

The gradual loss in gain seen in the CDC that was reported at the previous BPAC meeting corresponds to a gain reduction of about 20% over the period from the end of 2020 until May 2021. There is a concern that this is related to ageing of the chamber due to deposits on the wires, but the integrated charge of less than 0.1 C/cm is an order of magnitude below the level at which ageing effects may show up. The gain is also affected by the level of water (H_2O) added to the gas mixture.

Water has been added to the gas mixture since 2019, after an increase in leakage currents was observed. After its introduction, the leakage currents stabilised. In studies made with cosmic rays, the gain shows a clear dependence on the amount of water added. However, during the period that the gain drop was observed, the monitored level of water content of the CDC gas was measured to be constant at 2000 ppm. But there is some concern that the monitoring may not reflect accurately the actual level of water content. The amount of water has now been reduced to 1500 ppm, as a response aiming to recover the gain loss.

In addition to understanding the best choice of operating parameters for the CDC, a dedicated ageing study has been resumed using a test chamber. Improved quality control is being imposed for this chamber construction compared to the previous study made in 2019, aiming to reproduce the conditions in the actual CDC, and irradiate the test chamber with a 30 MBq ^{90}Sr source to accumulate 1 C/cm in a few months.

3.1.2 Concerns

- Correlation between parameter settings for the CDC operation and its behaviour is not well understood.
- Water content of the CDC gas is not well known inside the chamber.
- If the water content is reduced too far, the increase of leakage currents seen in 2019 may return.
- With the current level of effort, results from the study using the revived ageing set-up in preparation will be available only in the timescale of about a year.

3.1.3 Recommendations

- There is a need to gain better understanding of the effect of the parameters, such as the actual level of water content in the gas mixture, on the operation of the CDC.
- Reducing the amount of water added to the gas mixture to recover from the gain loss is reasonable up to a certain extent, but may be dangerous if taken too far, with the risk of the increase in leakage currents returning. Such changes should be performed cautiously, with careful monitoring of relevant parameters.
- The ageing study should be pursued with high priority, with increased effort to arrive at a timely decision for the future operation of the CDC based on its conclusions.

3.2 Photon detectors for the barrel particle identification system

3.2.1 Status

A recent analysis of the number of hits per track in dimuon events recorded by the TOP MCP-PMTs has indicated an unexpected decrease in the number of detected hits in recent runs, although with large fluctuations due to changing of the run configuration. The integrated charge at the MCP-PMT photocathodes is of order 0.1 C/cm^2 for the conventional tubes, and typically about twice that for the ALD and lifetime-extended ALD tubes. This is still well below the expected lifetime at an exposure of 1 C/cm^2 for the conventional MCP-PMTs, as determined from test-bench measurements. The effect is also seen for some ALD coated MCP-PMTs, whereas their lifetime was expected to be substantially enhanced, by around an order of magnitude. The loss of hits since March 2021 averaged over azimuthal sectors (slots) varies between 2% and 8%, with some indication that it depends on the production batch of the MCP-PMTs. It is not yet clear if the effect corresponds to a real loss of quantum efficiency (QE) of the photocathode, or just a loss of gain which could be recovered by adjusting the operating voltage. In principle this should be determined using the laser calibration system, but there is some concern that the effect may be rate dependent, complicating the comparison of the calibration and physics data.

3.2.2 Concerns

- The loss of signal efficiency currently seen in the TOP MCP-PMTs is unexpectedly higher than what is expected from the accumulated charge, in particular for a subset of the tubes.
- It is not yet clear that the observed QE loss, if real, is directly related to output charge, as measured in the bench tests. For example, it might be that some other in situ or environmental factor (like neutron dose, or elevated tube temperature) is playing a major role, either directly or by mediating more rapid deterioration with output charge.
- If it is confirmed that the QE is deteriorating, then this will have a significant impact on the planned replacement of MCP-PMTs during LS1. Since some of the ALD coated type are also suffering degradation, the original plan of just replacing the conventional MCP-PMTs with lifetime-extended ALD coated MCP-PMTs may need to be reconsidered.

3.2.3 Recommendations

- The cause of the efficiency loss in the TOP MCP-PMTs must be studied and understood with high priority; if the loss can be recovered the calibration procedure will need to be adapted.
- If the efficiency loss cannot be recovered, then reliable extrapolations of the long-term impact of the effect should be developed for the scheduling of the MCP-PMT replacement campaign, as well as considering possible changes to the operating conditions for the MCP-PMTs in the ongoing run to minimise the impact.
- The gain calibration and adjustment procedure should be made more robust, if it is indeed found to depend on the hit rate of the events concerned.

3.3 Other detector issues

3.3.1 Status

The committee greatly appreciated the informative Belle II overview report, including the summary of concerns.

The DAQ transition from COPPER to the PCIe40 based read-out for TOP and KLM worked fine for KLM but caused significant problems and down-time in the TOP[1]. The

[1]Detailed information was received after the BPAC meeting.

roll back to the COPPER based read-out was successful, and the problem was traced later to a software bug in a control process, which was exposed only when a large number of channels were monitored. A fix was found afterwards and successfully verified.

3.3.2 Concerns

- Changeover from COPPER to PCIe40 in the DAQ system was not as transparent as anticipated.

3.3.3 Recommendations

- It is helpful to continue to include the “Summary of Concerns” for Belle II in future overviews.
- Changes to the readout system should be always tested at full-scale and with the final software stack in place, before the start of physics data taking.

4 Long Shutdown 1

4.1 Status

Due to an observed gold delamination of the new beam pipe for the installation of the vertex detector (VXD) with the upgraded pixel vertex detector, PXD2, the collaboration is forced to delay the start of Long Shutdown 1 (LS1). Two scenarios are being considered. Scenario one calls for a half-year delay of the start of LS1, with LS1 starting on January 2023 with a duration of nearly ten months. The LS1 work would encompass the replacement of the MCP-PMTs for the TOP detector and the extraction of the VXD with current pixel detector, PXD1, and installation with the PXD2. Checking and possible repair of the QCSR leak would occur during the 2022 summer shutdown. Scenario two has a year delay of the start of LS1, with LS1 starting in July of 2023 with a duration also of 10 months. Under this scenario the work on the MCP-PMTs would be done during the regular 2022 summer shutdown extended by two months. As already conveyed in the short summary of the meeting, the committee recommends the first scenario, that is, a half year delay of LS1. This mitigates the potential loss of expertise and allows the integration of the MCP-PMT replacement in the same shutdown.

To mitigate the effects of synchrotron radiation on the PXD, a new beam pipe was being produced for the PXD2 that has a shorter straight section on each side, by 1 cm in the forward region and 0.5 cm in the backward region, additional gold plating on the inner pipe, and a slightly modified profile at the crotch part. Unfortunately delamination of the gold plating of the beam pipe was observed, which cannot be repaired. A new inner section had to be procured, and the order was placed in September 2021. The schedule calls for the gold sputtering on the new inner pipe to take place in May 2022. After further integration with the heavy metal part the new beam pipe should be ready by October 2022. The cause of the delamination is at the moment not fully understood. The thought is that an oxide layer had developed that prevented the gold to adhere to the titanium. Further studies, such as X-ray photoelectron spectroscopy will be carried out to investigate the condition of the oxide film layer on the inner pipe. The PXD2 half-shells are scheduled to arrive at KEK in September 2022, so the availability of the new beam pipe drives the overall schedule.

The team has a fallback plan in case there are any further issues with the new beam pipe. A spare cylinder set exists, but at the moment it is not known if it can actually be used, given the issue reported with this set.

The PXD2 schedule has seen a shift due to some new issues that emerged. The foil that is needed to isolate the backside of the module electrically from the Support Cooling Block (SCB), which is at ground potential, suffered from some cracks. This has been traced to the presence of burrs from the melted adhesive when the foils are laser-cut. Foils are now cut by knife and with punch pliers. This has resolved the issue. Also a serious misalignment was uncovered during the installation of dummy modules on a half-shell. The stiff CO_2 and N_2 pipes exert force on the SCBs and cause a misalignment, which leads to stress on the ladders. This has contributed to the crack that was observed in one module. This will be avoided by pre-bending the pipes.

Populating the final half-shells with production modules has started. One layer 1 module was mounted and the revamped electrical tests uncovered several disconnected lines. The team is to be commended for expanding the electrical tests. Continuing the population of the final half-shells has been halted until all the issues with the dummy half-shells are fully understood.

At the time of the review four grade A and two grade B ladders were ready for layer 1. Two more are needed plus spares. There are enough good modules to build another four layer 1 ladders. For layer 2, eleven grade A ladders are available; one more is needed, plus spares. There are enough modules available for four additional layer 2 ladders. The rework that some modules required has been successful but is time consuming.

The current schedule calls for the PXD2 to be completed by mid-June 2022, which gives the project about three months of schedule contingency.

The schedule and procedures for extraction and reinstallation of the VXD with the new PXD2 are being developed and documented and are being reviewed, however, without several key experts present. Several scenarios have been developed including scenarios where repairs will have to be made to the SVD. This level of contingency planning and risk mitigation is very valuable and the committee was very impressed with the proactive approach. The chicane at the forward QCS is tight for the full VXD cable plant. The situation is being mitigated by a modification of the QCS head-plate and the CDC support ring. The modification of the infiniband cables will provide more space.

The forward and backward beam background shields around the bellows is proceeding well and there are no schedule concerns.

Overall, the project is well-prepared for the activities planned for LS1 under scenario 1.

4.2 Concerns

- The schedule for the new beam pipe is very tight.

- It was not clear whether the spare beam pipe set is really usable given the known issue with the principal set.
- Given the part inventory, it appears that the collaboration has a very limited number of components, modules and switcher chips, to complete the assembly of the required ladders, with a limited number of spare parts. This holds especially true for the switcher chips.
- Loss of expertise for the pixel detector is a key concern, in particular for the installation of the pixel detector.
- All half-shell transports were done with steel ladders in L2 so far. The feasibility of safe transport of half-shells made from only silicon dummy ladders is at the moment unknown.
- Most of the disassembly operations have been documented and exercised without some key experts, which is a real concern.
- Although there is an excellent effort to document the extraction and installation procedure, to mitigate the impact of disappearing expertise and the possibility of continuing travel restrictions, the committee notes that there is no substitute for hands-on presence.

4.3 Recommendations

- Carry out studies to understand the cause of the gold delamination expeditiously and perform tests on the bench to verify the process under as many conditions as possible to ensure success of the final gold sputtering of the inner beam pipe. Engage the subject matter experts, both from KEK and the company, during all steps of the process.
- Verify that the spare beam pipe set can actually be used and that the observed issues are not a showstopper. If the set can be used, carry out all necessary steps to be able to immediately resort to the fallback plan to use this spare set if needed.
- Regarding PXD construction, give special attention to test all procedures in advance given the limited number of still available components, in particular the half shell assembly and its transportation. Using schedule contingency is preferable over taking the untested path with its associated risks.
- Practice the installation of the VXD often to minimise the risk of accidents, mimicking the real detector as closely as possible including a realistic modelling of the cable plant. It is strongly recommended to practice the beam pipe assembly and the VXD support assembly with the key experts. Effort should be given to early and advanced planning to allow for the possibility of former experts to participate in the extraction and installation process.

5 Level 1 trigger

5.1 Status

The L1-trigger must keep the accepted event rate within the limit sustainable by the Data Acquisition (currently < 14 kHz, in the future up to 30 kHz) and provide ~ 10 ns precision for the time stamp of an event, while maximising the efficiency for $B\overline{B}$-pairs. The system uses principally charged tracks in the CDC, signals from the ECL and also the information from the TOP and the KLM. A central trigger decision is calculated in the Global Decision Logic. All L1-trigger systems, except for the TOP that is required mostly for precise timing, are completed and operational.

There are convincing plans and work-programs in place for improving the selectivity while maintaining a very high efficiency by implementing hardware upgrades and deploying more sophisticated configurations. In particular, an L1-trigger configuration is needed for operating at luminosities well above $10^{35}\text{cm}^{-2}\text{s}^{-1}$, i.e ~ 10 times more than now.

5.2 Concerns

- While there is a clear work-program, with priorities, to realise the various gains in trigger performance, the human resources available for this program look quite limited for many items, and for some, albeit rated with lower priority, they are today altogether lacking. For the neural trigger, longer-term support and development depend on the success of an upcoming funding request.
- A lot of critical tasks depend on a single tenured person.

5.3 Recommendations

- The efforts should be continued to attract more and new institutes to contribute to the trigger work, not only for new features or future improvements but also for long-term continuing maintenance.
- It would be highly advantageous if a second tenured scientist could be identified, with the L1-trigger being one of the priority tasks, who could also act both as a back-up as well as share the load in the central coordination of the trigger work.
- The L1-trigger team is encouraged to pursue the studies to mitigate the performance degradation in a high luminosity and large background environment for all subsystems. The most important one seems clearly to be the upgrade of the CDC track trigger for the reduction of fake tracks.
- Efforts to simplify the system and to develop commonalities among the different hardware components are highly recommended and should reduce the maintenance effort.

6 Data calibration and quality validation

6.1 Status

There has been substantial progress in the development of the rather complex calibration processes based on information extracted from the detector components. These calibrations are of great importance for high precision analyses, specifically the production of the run-dependent simulation samples.

The prompt calibrations are performed just after the data are recorded. They now include KLM and ECL crystal timing, PXD gains, the beam energies, and the recently added ECL leakage corrections which are now based on random background data, rather than simulations, as previously. The calibrations are almost fully automated. Based on recent improvements, they now take about ten days to complete. These calibrations are supervised by detector experts who check and approve the results.

The CDC dE/dx calibration is very challenging. It requires accurate path-lengths and is performed after the tracking calibration, first with electrons and then for muon and for hadrons with mass-dependent specific treatment. These very important calibrations are still being fine-tuned and are not yet automated.

Simultaneous internal and relative global alignment of PXD, SVD and CDC (layers) is part of the automated prompt calibration. It is using tracks from collisions combined with cosmic ray tracks recorded during physics data-taking. For the proc12 reprocessing (and since Run 2021c already in prompt calibration), full alignment of PXD, SVD and CDC, including geometrical alignment of the wires using collision, off interaction point (IP) and cosmic tracks combined with vertex and IP-constrained di-muon events is performed. In addition, run-dependent alignment for PXD sensors, VXD half-shells and CDC layers per every approximately 100k of recorded di-muon events is done to mitigate the observed time evolution of the misalignment.

Recalibrations have been performed on earlier software releases *proc9* to *12*. Raw and post-tracking calibrations on *release-06* data are being planned for *proc13*. The presently estimated time for this enormous task is four to five months, thus it should be completed by late summer of 2022. What is presently missing is the validation of the prompt and recalibration processes. The stability and the precision of the various calibrations and their impact on the event reconstruction are of great importance. Plans for a fast and broad coverage of monitoring the calibrations and the resulting data quality are being developed. An assessment and visualisation of the time dependence of the performance for the broad spectrum of calibrations would be very valuable. The implementation using the Mirabelle system is expected to be ready by the summer 2022.

6.2 Concerns

- In the meeting there was little information given on the results obtained with the calibration procedures, such as their precision and stability over time, and the impact and corrections for malfunctions in the automated procedure. It is somewhat surprising that the internal alignment of the subdetectors is part of the prompt and more frequent calibrations, given that individual sensors should be firmly mounted. This is in contrast to the global alignment of the PXD and the SVD relative to the drift chamber which might be sensitive to diurnal changes unless the temperature control is perfect. For instance, changes of 1°C could result in significant relative movements and may require alignments every few hours.

- It was reported that the global alignment of the tracking detectors depends on cosmic rays recorded with zero magnetic fields on a few occasions per year. This is impractical as other global movements are calibrated at much smaller time-scale. It is perhaps also inappropriate as the position of the detectors may depend on the magnetic field.

6.3 Recommendations

- The validation of the calibration processes is of high importance and should be given full priority, both in terms of the measurements as well as the complex software tools. The development of the calibration algorithms is covered at the level of the subdetectors. While the current support by subdetector experts appears to be adequate, additional support may be needed to oversee the development, maintenance, and control of the workflow, thereby avoiding delays.
- In order to maximise the sensitivity to variations with time and yet simultaneously reduce the sensitivity to statistical fluctuations it should be studied at which time scale and at which granularity the alignment and calibration constants vary. The calibration strategy should be adjusted accordingly. The developments in the Mirabelle system that facilitate these studies should be given high priority.
- Alternate methods for the detector alignment should be explored to eliminate the dependence on magnet-off data. For instance, in addition to muon pairs and cosmic rays, the kinematic constraints from low multiplicity decays can be used to constrain the alignment degrees of freedom corresponding to curvature bias (see e.g. arXiv:1207.4756).

7 Software validation and data processing

7.1 Status

The status of the Belle II software is generally good. All source code and scripts are made publicly available since summer 2021. New features and bug fixes are regularly deployed. A software validation process with dedicated tools is in place. A suite of test runs checks compatibility of the modifications submitted for the Belle II Analysis Software Framework (*basf2*) and flags obvious consequences resulting from the changes. Nightly tests process small data samples of simulated and real events and compare the event signatures with reference ones. Following these regular "low level" checks, validations are carried out for different types of software release. Patch releases are generally to provide bug fixes, in particular for the HLT, and prompt calibration and reconstruction. For this reason, specific tests are made on small samples of raw data to check the memory usage and processing time as well as the trigger performance on the online system before deployment. Major releases include all approved new features developed over time and are made once a year. A limited amount of new features, usually of a specific purpose, are added in minor releases. Major releases require a thorough validation of their functionalities and output results in close collaboration with the Data Production and Physics Performance Groups. Special processing of data *buckets* and production of simulation (MC) samples on the Grid, is carried out to evaluate the performance of "low level" algorithms, like tracking, as well as high level tools for PID, FEI and flavour tagging.

The analysis software also relies on external packages such as ROOT and, in case of simulation, on EvtGen, Pythia and Geant4. The update of these packages in *basf2* is usually done a few months before the freeze of a major release. Only basic technical compatibility checks with the Belle II software are carried out for them. Their physics functionality is evaluated through their effect on the validation of a major release.

The validation of *release-06* was expected to be carried out over the summer to allow an early start of *proc13* in order to provide the data set for summer 2022 analyses. Various issues were observed in high-level quantities in the first Monte Carlo validation samples, in particular a significant degradation of the flavour tagging performance. Many potential sources for the problem were considered due to the wide range of changes introduced in the release. The problem was identified in the version change of the Pythia generator. Additional fixes and improvements were included in *release-06-00-00* deployed on the HLT nodes at the beginning of October. The whole validation process took fourteen weeks. It cannot be ruled out that an unexpected effect of changes will only be found in the last phases of the validation for a major release and it is difficult then to predict how long it may take to identify the cause. This should be taken into account in the planning of the deployment of new major releases. Further automation of the validation procedure is under discussion. A strategy for early identification of the possible implications of updates of external packages should be devised.

Since 2016 Belle has used Atlassian collaborative tools for software management. The Belle II software group is looking for alternative products due to the cost increase in licences. The move to the most likely candidate, GitLab, is not expected to have a massive impact on the experiment.

The infrastructure and framework for data processing and Monte Carlo production in the Grid distributed computing environment is in place and operating smoothly. The five raw data production centres are performing well and provide sufficient storage and CPU resources. Smooth operation was also reported for user analysis jobs, which are increasingly running on the distributed computing system.

The production of data sets for analysis targeting Moriond 2022, *proc12* and *MC14* was almost completed at the time of this BPAC meeting. Production of *skims*, i.e. events selected for specific analysis, was in progress for both collected data and simulation. Reduction of the number of skims and automatic submission of skim jobs are under consideration to ensure future sustainability of such productions.

An accidental removal of the whole run-independent MC data set *MC14ri_a* after completion was reported. About a quarter of the samples were recovered at KEK and DESY and will be redeployed on the distributed Grid storage. Work to reproduce the remaining samples in a new *MC14ri_d* production campaign was in progress. The aggressive job scheduling provided a very useful stress test of the distributed computing system: a record of over $30k$ concurrent jobs was reached in October 2021. A generic MC sample ($700\,\mathrm{fb}^{-1}$) was produced in only 10 days, much faster than any previous sample. One of the reasons is the fact that no other major processing took place at the same time. The other reason is the focused effort by the data production and computing groups to minimise the time for job preparation and to continuously look after the submission and start of productions. Increasing the automation of the latter aspects would result in faster Monte Carlo productions without undue strain on people.

Experience continues to be gained on the production of run-dependent Monte Carlo samples. Preparation and processing of this type of MC production is slow due to the complexity, requiring a preparation of specific running conditions and payloads for corresponding runs and their dependence on the background input files. The number of productions to manage is also an issue. To speed up the collection of the payloads, the same AirFlow mechanism used for the calibration of the collected data has been put in place. It is foreseen to be eventually integrated into the automated workflow for the preparation itself of the payloads. Streamlining of the whole process is expected to reduce the time needed to prepare each run-dependent MC production. In addition to automation of the run-dependent MC production, work has started to implement requests for signal MC production via AirFlow. An appropriate number of experts needs to deploy such automation in a timely manner.

Prompt data processing for *2021c* data taking will use *release-06* and follow the well established procedure of *2021a* and *2021b*. For the first time, run-dependent MC for prompt data is also planned to follow as soon as the prompt calibration at BNL is done, with a target latency of ~ 1 month. The start of reprocessing *proc13* with *release-06* has been delayed and the processing will start after the full validation on collected data is completed. The critical point remains to be delays in the recalibration. Effort is foreseen in further automating the process. A proposal for the data set to be used for physics analysis targeting the summer was under discussion with the Physics and Physics Performance groups. For real data the Moriond data set[2] will be complemented with the *prompt21c* samples. Simulated samples produced with corresponding *MC14* and *MC15* have to be ready for the physics analysis. The Belle II collaboration intends to make the run-dependent MC as the default for the analysis targeting for Summer 2022. A validation plan of *MC14-rd* with comparison with data will be in place for the next months.

The committee was pleased to hear about the effort by the computing and data processing groups to continue the automation of the various operation aspects, in particular for data management. The need to do so, in particular for file deletion, is signified by the accidental removal of all files produced in the latest run-independent *MC14ri_a*

[2]reprocessed *proc12* and *prompt21ab* samples produced with *release-05*

simulation campaign. This mistake occurred during the manually triggered file deletion of cancelled jobs. Fully automating the deletion of files has high priority in their work plan. An independent verification of the list of files to be deleted has been put in place as a mitigating measure until then. Central removal of user files was planned soon after the BPAC meeting to reduce the $\sim$ 800 TB of Grid space they use. Different strategies on how to keep this under control for the future are under discussion.

7.2 Concerns

- Belle II is now in a phase where operation with stable software has to be sustained; at the same time new features and bug fixes need to be deployed in a timely manner to improve the precision of physics measurements. A careful balance of the two aspects has to be ensured.
- Allocation of human resources for data processing operation and for further automation of the operation has to be carefully balanced.
- Careful resource planning and follow up is required for the deployment of the planned automation in preparation of the run-dependent Monte Carlo production.
- Automation is reducing the need for human resources. Nevertheless redundancy in key personnel is still required, in particular to ensure knowledge transfer during unavoidable turnover of personnel.

7.3 Recommendations

- To ensure stable software operation, it is unfeasible to foresee more than one major release per year. The Belle II software group should investigate how to best exploit development releases and to introduce changes incrementally with regression tests to monitor basic performance of algorithms. Such development releases may also be used in the identification of problems observed in a subsequent major release.
- It is essential that a careful validation of external packages is carried out as soon as possible, and not wait for a major release, although a thorough validation requires the production of many new simulated samples. The committee suggests that checks of simple quantities for generators and Geant4 are carried out when changes are introduced.
- While moving to GitLab collaborative tools should not affect the software per se, it will still require some change of work habits. Users and developers should be given clear recipes on how to perform their tasks with the new tools.
- The computing group is encouraged to timely complete the remaining work for the automation. This will help in relieving the burden on key personnel.