

B-factory Programme Advisory Committee

Full summary report

for the annual meeting on 10–12 February 2020 at KEK

A. Andreazza* (Milano), D. Cassel (Cornell), P. Collins* (CERN),
G. Corti (CERN), M. Demarteau (ORNL), R. Forty (CERN),
B. Gavela (Madrid), S. Gori (UCSC), M. Ishino (Tokyo),
Z. Ligeti (LBL), H. Lacker (Berlin), V. Luth (SLAC),
P. McBride*& (FNAL) P. Mato* (CERN), F. Meijers* (CERN),
N. Neufeld& (CERN), K. Oide* (KEK), B. Ratcliff* (SLAC),
M. Sullivan (SLAC), H. Tajima (Nagoya), M. Titov* (Saclay)
and chaired by T. Nakada (EPFL)

* Expert members, & Remotely participated in part

11 May 2020

Preamble

Since the February BPAC meeting, and at the time of the submission of this report, the world finds itself in the midst of a fight against the COVID-19 pandemic. In the report it is recommended to increase on-site presence in several areas. But this is understandably not achievable under the current circumstances and needs to be replaced by intensified remote work. The committee very much appreciates the efforts of KEK and the Belle II collaboration to continue operating the accelerator and data taking under these unprecedented circumstances and wants to express its full and sustained support for this great physics program.

1 Short Summary

The annual review meeting of the Belle Programme Advisory Committee (BPAC) took place at KEK from 10th to 12th of February 2020, where presentations including the physics results of the Belle experiment, the machine status of the SuperKEKB, as well as the status of the Belle II experiment, were given. This section provides the most important findings and recommendations of the committee. A full report follows.

The committee is highly impressed by the productivity still maintained by the Belle physics effort, which has produced 25 refereed journal papers in 2019. For fully exploiting the potential of the Belle data, it is important that computing resources at KEK for the Belle data analysis remain adequate until the amount of Belle II data collected becomes comparable to that of Belle. The Belle physics analysis clearly benefitted from the development of the BASF2 analysis tool by the Belle II physics group. The committee

would like to see further collaborative efforts by the two physics groups, particularly in the area of dark-photons search where Belle II could have a significant impact at this early stage.

The SuperKEKB machine was able to achieve luminosities above $10^{34} \text{ cm}^{-2}\text{s}^{-1}$ with less than half the beam current of the KEKB accelerator. This is a clear success of colliding bunches with very small spot sizes achieved by the low-emittance beams. It demonstrates the potential of the SuperKEKB to further increase luminosities. However, achieving the design goal of $\sim 8 \times 10^{35} \text{ cm}^{-2}\text{s}^{-1}$ with a tolerable level of background for the experiment will still involve substantial work and new efforts, such as testing the crab-waist scheme in the next run, would be required. The committee encourages the SuperKEKB and Belle II teams to work out a plan for the future work.

There has been a continuous effort to reduce machine background and this effort will remain crucial. While a steady increase of luminosities toward the design goal is important, the current background level is already at the limit for safe operation of the Belle II detector. Therefore, efforts such as identifying the sources of backgrounds through simulation studies and improving the beam collimation and absorption of the synchrotron radiation must continue vigorously. The background condition for the injection is a worry, since the continuous injection is essential for achieving large integrated luminosities. Therefore, reducing the injection background is clearly one of the highest priorities. To facilitate this effort, the committee encourages the current close collaboration between the Belle II background study group and the SuperKEKB machine team to be extended to include the team operating the injection system.

The committee is pleased to note the progress in the commissioning work of the Belle II detector, as well as in computing and software. Although the experiment is not yet operating at full capability and further work is still needed to completely understand the detector performance, the Belle II experiment is now ready to record substantial data samples for physics analysis under stable conditions. Therefore, the committee supports the scenario where the installation of the new Pixel Detector (PXD) with two layers will happen in early 2022. This may allow the Belle II experiment to collect $\sim 1 \text{ ab}^{-1}$ of data and thereby surpass the sensitivity of the Belle data and move towards the world leading position in some analyses, before the installation of the new PXD. Since there are still some uncertainties in the evolution of the SuperKEKB luminosities and in the stability of the Belle II detector, the production of the new PXD should continue with full-speed such that it could be installed in summer 2021, in order to maintain schedule flexibility.

The recording of these large data samples requires stable operation of the detector. Therefore, the detector performance and machine conditions must be carefully monitored and further improvement in simplifying the monitoring tasks is encouraged. For collecting a large amount of data, it is equally important to have an efficient trigger and data acquisition (DAQ). The committee urges that more work should be invested in these critical areas, specifically all of the trigger components should be implemented and causes for the DAQ downtime should be carefully analysed and reduced as much as possible.

Finally, the committee would like to stress that it has now become very important

for the Belle II experiment to have sufficient running time. The KEK management is strongly urged to make all possible effort for securing resources necessary for the yearly machine operation in the coming years to be substantially more than 5.4 to 6.5 months indicated during the meeting.

2 Physics

2.1 Physics of Belle

2.1.1 Status and plan

The Belle collaboration has remained very active in 2019, nine years after the end of Belle data taking. The collaboration now consists of 450 scientists from 103 institutions in 22 countries, including three new added after the previous BPAC meeting. Utilising the world's largest data sets at the Υ resonances $1S$, $2S$, $4S$, and $5S$, the Belle collaboration maintains a very strong program of physics analyses. The 2019 highlights of Belle results include a semileptonic-tag $B \rightarrow D^{(*)}l\nu$ analysis and $R(D^{(*)})$ measurements, new measurements of CP violation, and spectroscopy.

In the past year, 25 new papers were published, increasing the total number of Belle publications to 537. This vigorous publication activity has a large impact in terms of citations and entries in the Particle Data Group reviews. This is also reflected in strong presence at conferences. Together with LHCb, Belle is co-leading the Heavy Flavour Averaging Group (HFLAV). Belle physics analyses will continue until the data set is superseded by Belle II. An important role is played by the analysis of Belle data using the Belle II analysis framework, b2bii. The b2bii framework was used to validate on Belle data the Full Event Interpretation developed by Belle II and to estimate the improvement of flavour tagging in Belle II.

Hadron spectroscopy can be explored at Belle in numerous processes, in continuum production from e^+e^- at various centre of mass energies and in decays of bottomonium resonances, B mesons and charmed hadrons. In the last year, Belle observed a new resonance decaying to $\Upsilon(nS)\pi^+\pi^-$ ($n = 1, 2, 3$), with mass (10.753 ± 0.006) GeV, in the analysis of e^+e^- collisions above the $\Upsilon(4S)$. The collaboration also observed a new charmonium-like state with mass (4.626 ± 0.006) GeV decaying into $D_s^+ D_{s1}^-$ (2536) in the analysis of ISR events. Searches for charmonium-like states were also performed in B decays. Belle continued to study a variety of non-leptonic B decays, including the first evidence for the baryonic decays $B \rightarrow \bar{\Lambda}_c^- \Xi_c^0(2790)$ and $B^0 \rightarrow p\bar{p}\pi^0$.

Belle also obtained an interesting new result for the time-dependent CP asymmetry in $B^0 \rightarrow K_S K_S K_S$ with the full data set, finding S_f to be consistent with $-\sin 2\phi_1$ and A_f consistent with 0, as expected from the Standard Model. A new measurement of ϕ_3 was performed in the decay $B^\pm \rightarrow D^0(K_S\pi^+\pi^-\pi^0)K^\pm$, where $K_S\pi^+\pi^-\pi^0$ is another final state accessible in both D^0 and \bar{D}^0 decay, with results consistent with the world average.

Important semileptonic B decay measurements were also performed, primarily focused on resolving the hints of lepton flavour universality violation and the tensions between determinations of $|V_{cb}|$ from inclusive and exclusive semileptonic B decays. With

$B \rightarrow D^{(*)} \ell \bar{\nu}$ decays, new $R(D^{(*)}) = B_\tau / \langle B_{\mu+e} \rangle$ measurements were made, where B_τ is the semileptonic branching fraction with τ and $\langle B_{\mu+e} \rangle$ the average of the semileptonic branching fractions with e and μ . The new Belle result is more consistent with the SM than earlier analyses. However, this measurement is also consistent with the hint of an anomaly seen by the LHCb collaboration in this decay mode. Belle also measured the τ polarisation and the longitudinal polarisation fraction of the D^* in $B \rightarrow D^{(*)} \tau \bar{\nu}$ decays, in agreement with SM predictions. New measurements for so-called R_K and R_{K^*} , the ratios of the branching fractions for $B \rightarrow K^{(*)} \mu$ over that for $\rightarrow K^{(*)} e$, were also obtained, with results compatible with the SM and uncertainties larger than those at LHCb. A new improved analysis of the purely leptonic decay $B^+ \rightarrow \mu^+ \nu$ was also performed, utilising the inclusive tagging approach to reconstruct the second B meson produced in the $\Upsilon(4S)$ decay. The measurement of the muon spectrum has also been used to search for a massive sterile neutrino with a mass below 1.5 GeV.

The Belle collaboration plans to continue an ambitious analysis and publication record during 2020. This includes the publication of the $R(D^{(*)})$, $R_{K^{(*)}}$ results, new hadron spectroscopy results, new searches for lepton flavour violating B meson decays, as well as new dark sector searches, with the dark sector particle produced from a B meson decay.

2.1.2 Concern

No particular concern.

2.1.3 Recommendations

- The BPAC congratulates the Belle collaboration on continuing to successfully produce important physics results. The BPAC encourages continued analysis of Belle data with the b2bii framework exploiting fully the Belle data.
- The committee encourages a smooth transition of the analysis effort from Belle to Belle II. The collaborations may consider merging all analysis groups (particularly in the area of dark sector searches where Belle II could have a significant impact at this early stage) and transferring the Belle roles in HFLAV to Belle II.

2.2 Physics of Belle II

2.2.1 Status and plan

The Belle II collaboration has completed its first physics paper that will appear in Physical Review Letters. This paper focuses on a new dark sector search for an invisibly decaying Z' boson produced in association with $\mu^+ \mu^-$ or $\mu^\pm e^\mp$ and is based on 276 pb⁻¹ of Phase 2 data. This search has been interpreted in terms of a dark sector model based on gauging the $L_\mu - L_\tau$ symmetry. The bound is not yet competitive with other experimental bounds on the Z' associated with this symmetry, but it will quickly become competitive when more data is collected and analysed. The analysis could be updated in the near future using either 10 fb⁻¹ or 100 fb⁻¹ of Phase 3 data. An additional

dark sector search based on Phase 2 data is close to completion. This is the search for axion-like particles decaying into photons, $e^+e^- \rightarrow a\gamma$ and $e^+e^- \rightarrow a \rightarrow \gamma\gamma$.

In addition to the Phase 2 data set, the Belle II detector has collected an additional $\sim 10.6 \text{ fb}^{-1}$ of Phase 3 data, of which 0.8 fb^{-1} is off-resonance, with a peak luminosity of $1.138 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$. This data has enabled the collaboration to rediscover several B meson decays, such as $B \rightarrow DK$, $B \rightarrow K\pi$ and $B \rightarrow K^*\gamma$. Several tens of events have been recorded for each of these decay modes. The $B \rightarrow DK$ analysis is a good candidate for the first Belle and Belle II combined analysis. The analysis of $\sim 2.66 \text{ fb}^{-1}$ data has also enabled the collaboration to perform the first studies of neutral B meson oscillation, using the decay mode $B \rightarrow D^{*-}\ell^+\nu$, $D^{*-} \rightarrow D^0\pi^-$. The fraction of un-mixed events has been determined as a function of the B decay time and this is the first milestone to performing time dependent CP violation analyses.

A large set of dark sector analyses have been also started. This includes the important search for an invisible dark photon (mono-photon signature). An analysis based on only 20 fb^{-1} data could give a reach on the dark photon coupling to standard model particles that is ~ 3 times better than the bound obtained by the BaBar experiment. The KLM veto is crucial for this analysis. The collaboration has also started looking for visible signatures of dark sector models, such as (i) prompt multi-muon signatures arising in dark Higgs models in which a dark Higgs, h' , is produced in association with a dark photon, A' ($e^+e^- \rightarrow h'A'$, $A' \rightarrow \mu^+\mu^-$, $h' \rightarrow$ invisible); (ii) displaced tracks signatures arising in models for inelastic dark matter ($e^+e^- \rightarrow \chi_2\chi_1$, $\chi_2 \rightarrow \chi_1 h^+ h^-$, where h is either a lepton or a charged hadron). The search (i) was investigated in the past only by the KLOE experiment but neither by Belle nor BaBar.

The collaboration has ambitious near term goals, in view of the Moriond conference and of the Summer 2020 conferences. In particular, for Moriond, several analyses are expected to be performed using $\sim 9 \text{ fb}^{-1}$ of Phase 3 data. This includes additional analyses for the B lifetime and oscillation using several $B \rightarrow D^{(*)}\pi/\rho$ decay modes (flavour tagging will be very important for these analyses); semileptonic B decays and corresponding measurement of $|V_{ub}|$; and first measurements of the τ mass. A large array of additional analyses is also in the pipeline and is planned for the Summer conferences and for the end of 2020. New results are expected from several working groups. Several core analyses are expected to become competitive with the corresponding LHCb analyses by the end of the present run. This includes tests of lepton flavour universality through the measurement of $R(D^*)$, the exclusive measurement of $|V_{ub}|$ and the measurement of the CKM phases $\sin(2\phi_1)$. The collaboration will also perform measurements that have not been done by Belle or BaBar, or that have been done but not with the full luminosity. On the longer term, two significant milestones with large impacts and visibility that would be remarkable to achieve before replacing the PXD would be the search for an invisible dark photon with 200 fb^{-1} ($e^+e^- \rightarrow \gamma +$ missing) and the test of lepton flavour universality violation in B decays with 1 ab^{-1} .

The collaboration is considering to have some short $\mathcal{O}(10 \text{ fb}^{-1})$ run above the $\Upsilon(4S)$ and, in particular, on the $\Upsilon_b(10750)$ or $\Upsilon(6S)$, to enrich its physics program. An additional goal could be the measurement of the D meson lifetime that has not been performed in the last ~ 20 years.

2.2.2 Future upgrade

The Belle II Upgrade working group was formed in October 2018. It has developed a well organised structure of studies to assess the feasibility and physics impact of a “high-luminosity” run and of polarised beams. The working group has already organised two workshops in 2019 and is open to the participation of people outside of Belle II for technology development.

The goal will be to investigate the feasibility for the machine and physics impacts of a run at $4 \times 10^{36} \text{ cm}^{-2}\text{s}^{-1}$ luminosity, that is 5 times larger than the Belle II design luminosity and a run with polarised beams (“chiral” Belle). The working group has a very ambitious plan ahead. In particular, the group that studies the possibility of realising a polarised e^- beam at SuperKEKB aims at finalising a white paper by August of 2020. The working group has also started identifying the most interesting channels that could benefit from a higher luminosity. Examples are $B \rightarrow \mu\nu$, $B \rightarrow \tau\nu$, $B \rightarrow \rho\gamma$, $B \rightarrow K^{(*)}\nu\nu$.

2.2.3 Concerns

- Progress in physics analysis is now crucially depending on the rapid collection of data.
- Full understanding of the detector performance needed for precision measurements has not yet been achieved.

2.2.4 Recommendations

- The committee encourages the collaboration and KEK to make every effort to achieve the goal of collecting $100 - 150 \text{ fb}^{-1}$ data by the 2020 Summer shutdown. This would allow obtaining several world leading physics results and some of the core physics results might start to be comparable to those of Belle.
- The committee encourages systematically developing and pursuing a program for low multiplicity searches, that could lead to world leading sensitivity to several light dark sector scenarios, with the data already available in 2020.
- The committee encourages finalising some dark sector searches possibly by the summer. A particularly interesting one could be the invisible dark photon search (mono-photon + invisible).
- The run plan presented to BPAC foresees only 5.4 or 6.5 months of SuperKEKB running in FY2020, a worse prospect than the one considered last year. KEK management and the collaboration are encouraged to find the resources for longer data taking in FY2020, as well as in the following years, enabling the stated luminosity and physics goals.
- The committee encourages KEK and the collaboration to do whatever it takes to collect at least a 1 ab^{-1} data set usable for physics analyses, prior to the long

shutdown to replace the PXD. This is needed for competitiveness with LHCb and to achieve high impact and world leading results soon, which will be necessary to maintain the excitement about Belle II and crucial to attract and retain top talent.

- We encourage the Belle II Upgrade Working Group to continue to explore the physics case and technical feasibility of maximising the scientific return on the investment in SuperKEKB and Belle II.

3 SuperKEKB machine

3.1 Status

The committee heard several interesting presentations on the status of the SuperKEKB machine. The peak luminosity (without detector data taking) is now $1.88 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ very close to the KEKB previous record of $2.1 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$. With the detector taking data, the luminosity is now over $10^{34} \text{ cm}^{-2}\text{s}^{-1}$. The beam current is about 500mA for LER and 360mA for HER during data taking. This indicates there is room for more luminosity, since the beam currents can be increased as the backgrounds improve. The nano-beam scheme has been shown to work as the luminosity of the machine increased when the β_y^* was lowered well below the length of the bunch. While the beam lifetimes and the beam backgrounds are slowly improving, there is still significant beam pipe “scrubbing” from synchrotron radiation especially in the LER which has the new beam pipe. The HER uses the original HER beam pipe from KEKB and hence the scrubbing advances significantly faster in the HER. Continuous injection into both rings has been implemented and the detector has instituted a blanking time around the injected bunch in order to maintain data-taking during injection. In order to reduce the data loss due to the blanking time, 2-bunch injection is being introduced. While this procedure is generally working, some of the injected pulses from the LINAC cause significantly higher backgrounds in the detector and beam aborts. The committee was informed that a dedicated time during initial commissioning has been set aside to concentrate on this issue. The committee has also learnt that the background team has uncovered a first glimpse of the expected signal related to the luminosities. This signal is expected to become dominant in the final machine operation with the design luminosity.

3.2 Concerns

- There are a subset of injected bunches that cause very high backgrounds in the detector and these particular bunches also take a long time to damp down and merge with the stored bunch.
- The specific luminosity of the colliding bunches goes down as the bunch current is increased. This implies that the bunch transverse size increases as the bunch current increases, causing a reduction in luminosity and an increase in the detector background.
- The LER collimator system appears to be insufficient.

3.3 Recommendations

- The committee strongly encourages close collaboration with the accelerator team to study the ‘bad’ injection bunches. The ability of the detector to identify exactly which injected bunch causes high backgrounds should greatly help in the effort to improve this issue.
- The committee encourages further efforts to make the beam abort as fast as possible. This includes placing “bad beam” sensors near the abort system as well as decreasing further the response time of the current beam monitors or deploying other abort signals.
- The plan to implement the crab-waist in the next run, first in the LER then possibly in the HER later, seems very attractive to suppress the beam blowup both for the core and the tail of the beam. Simulations indicate positive effects, at least for the $\beta_y^* = 1$ mm configuration, in producing a higher luminosity and a smaller beam tail from the beam-beam interactions. The committee encourages the use of the crab waist scheme as the main strategy for an intermediate luminosity step up to around 10^{35} cm⁻²s⁻¹.
- Establish a clear collimator strategy, such as in the LHC, by defining the role of each collimator. The D06 vertical collimators may work as the primary collimators if the phases between them are chosen properly. Then it is not clear if the D02 collimator still works as a secondary collimator. Perhaps the collimator optimisation code can shed light on this issue.
- The committee also encourages further studies of collimators using a low-Z material for the tip. This type of collimator could perhaps also be connected to the fast beam-abort signal.
- The committee was very pleased to hear that a background signal that looks to be luminosity related has been found and encourages further effort to dig this signal out of the general background information. This is the first time the background team has uncovered a signal that appears to be correlated to the luminosity. If the signal is coming from radiative Bhabha scatterings, as is expected, then they should try to benchmark this signal against expected background levels in various detector sub-systems using the current performance level of the accelerator.

4 Belle II detector

The committee acknowledges the excellent work done by Belle II for the improvement of data-collection efficiency and data quality, and urges the collaboration investing further efforts in this direction. Data quality of all sub-systems has to be carefully monitored in real time and all of the trigger components should be implemented. Causes for the DAQ downtime should be carefully analysed and reduced as much as possible. All of those efforts should be complemented by the efficient offline processing of data.

4.1 Vertex detector

4.1.1 Pixel detector (PXD)

4.1.1.1 Status

The PXD group presented their operational experience for the 2019c run. The data taking efficiency has been much improved and is currently at 98.4%. Inefficient clear gates have been observed in a few modules after beam losses, which contribute to the overall inefficiency. At the last meeting it was reported that the PXD suffered temporary damage due to the same incidents. A very successful set of tests is ongoing at the Mainz Microtron to understand and mitigate the effects of beam losses. The issues with the switcher chip could be completely reproduced under a similar condition seen in the experiment. The circuit board for the fast emergency power-down shows unexpected cross-currents that are being studied. A second round of tests is being planned for April in Mainz to validate the mitigation plan.

Gated-mode operation has been verified to work successfully on the majority of modules. A few modules behave differently, which will be investigated. The ROI selection was tested a couple of times and the final test, that ran for more than two days with only the ROI filtering enabled, showed no reduction of effective PXD efficiency. A data reduction factor between 3 and 9 was obtained.

The background from synchrotron radiation was studied after the machine optics switched to a β_x^* of 60 mm for the HER. The background is considerably larger than predicted by Monte Carlo and two distinct sources of background have been identified: one source that only occurs during injection, and another source from the stored beams. Dedicated studies to understand these sources are being planned. There are indications that this configuration of machine optics doubles the PXD background. Mitigation plans, such as adding a thin gold layer on the outer surface of the inner layer of the new spare beryllium beam pipe that could be installed together with the new PXD detector, are being studied.

The work to replace the current PXD with a new detector, PXD2021 with a full complement of ladders, is making steady progress. Enough ASICs are available to build the new pixel detector. Two batches of 12 wafers each are in production for the sensors. The first batch is expected to be finished by the end of February; finishing of the second batch has been put on hold at the stage prior to the metal deposition. The plan calls for 12 ladders, plus 6 spares, for the outer layer (L2), and 8 ladders, plus 4 spares, for the inner layer (L1). For L2, there is currently the material for 12 ladders, plus 3 spares, in hand. Three more spare ladders are being built. For L1, there are enough modules and ladders in hand for 6 L1 ladders, that is, two short for a full layer 1, with no spares. Six more L1 ladders are in preparation, for 8 full ladders and four spares. Module production and testing is ongoing with no unexpected surprises.

Following a prior recommendation, the PXD group has carried out an analysis evaluating the impact on the schedule of potential yield issues in L1 module and ladder assembly that may require the processing of the PXD9-21 sensor batch. This study was carried out and the group is confident that there will be enough ladders available for the PXD2021 detector. Furthermore, finishing the PXD9-21 batch would take about one

year. An equivalent study was carried out to assess the impact of moving the flip-chip bonding from HLL, which is the baseline, to IZM. The conclusion of this study was that moving the flip-chip process would be feasible but incur a one month delay.

A very preliminary simulation study to assess the impact of the second layer of the PXD on the performance for the decay time dependent CP violation measurements was carried out. The study started with an estimate of the background for the 2020 to 2022 running period based on current information. For these studies the PXD background was scaled and additional occupancy from synchrotron radiation was not included. The decay $B^0 \rightarrow J/\psi(\mu\mu)K_S^0$ was then simulated under different background conditions and for the one- or two-layer PXD configuration. The reconstruction efficiency and Δt resolution was determined for each configuration. Noting that the assumptions on the background evolution have very large uncertainties, preliminary results indicate that noticeable improvements can be expected for the time-dependent analyses with a fully equipped PXD installed in the summer of 2021.

The current schedule calls for the module assembly to be completed by May, 2020, and ladder assembly to be completed by July 2020. Half-shells would be ready for shipment to KEK in August, 2020. This delivery schedule has ten weeks of schedule contingency. The PXD team is also engaged in a concerted collaboration-wide effort to find more personnel for the PXD group. This effort, to guarantee that expertise and sufficient staff are available to complete the PXD2021 construction and installation, and have enough staff for the long-term operation of the detector, is highly appreciated.

The committee commends the sustained efforts of the collaboration and remains impressed by the dedication of the team. The path to a full PXD detector is in clear focus with a reasonable time profile and schedule contingency.

4.1.1.2 Concerns

- The schedule for the new PXD detector remains a vulnerability for the project.

4.1.1.3 Recommendations

- Although the schedule is in clear focus, the committee would just like to remind the team of the expression: “Man soll das Fell des Bären nicht verteilen, bevor er erlegt ist” or “don’t count your chickens before they hatch.” The team is encouraged to remain as vigilant as ever in their final stretch to complete the detector.

4.1.2 Strip vertex detector (SVD)

4.1.2.1 Status

Since the last BPAC review in October 2019, the SVD participated in the 2019c run. Operation went smoothly, following the model of having an on-site and a remote shifter. According to the online monitoring, hit efficiency has been above 99.5% both in the U ($R\Phi$) and V (z) measuring strips. With the current background level, occupancy was below 0.3% and did not affect the detector operation. The total collected dose on L3 during Phase 3 has been estimated to be 20 krad. It caused a limited increase of leakage

current and a 20% increase in L3 noise. Also the CO₂ cooling system operated reliably during data taking. An increased leakage current correlated with continuous injection has been observed on one L6 ladder. It is now operated at a shifted voltage, and tentative explanations, which are under active and thorough investigation, have been provided.

The SVD group is looking carefully at the data quality and performing detailed data/simulation comparisons. The dE/dx studies show the energy deposited by charged particles follows the expected pattern for 320 μm detector thickness and incident angle. The dE/dx distributions also match between data and simulation. As far as the resolution is concerned, measured values qualitatively follow the expected shape for $R\Phi$ and z measurements as function of layer and incident angle, but they show a performance systematically slightly worse than the simulation prediction. This seems correlated to the different cluster size in data and simulation.

Attention has also been devoted to the study of possible solutions to the increasing occupancy at future luminosities. Occupancy above 2.5% will hit data rate limitations, causing hit inefficiency, and above 3% tracking performance will be affected. These limits will be exceeded at the nominal luminosity. The data rate limitation can be addressed by reducing the digitisation to three-fold sampling and by exploiting the DAQ upgrade to increase the readout bandwidth. The tracking limitation can be overcome by rejecting out of time hits. A time resolution of 2.5 ns is observed with Bhabha events. Spurious hit rejection can tolerate a 10 ns T_0 jitter, therefore this rejection filter should work with ECL timing, and eventually with CDC-timing, after the performance of events with CDC-only timing is understood. Test runs in 3-samples readout mode are foreseen in 2020 to verify this hypothesis.

The committee commends the SVD group members for their effort in operating and understanding the detector and its excellent performance.

4.1.2.2 Concerns

- The discrepancies found in data/simulation comparison shows that the cluster size, and therefore the occupancy, are higher than the simulation predictions. That reflects in a worse than expected position resolution. These effects may limit the overall tracking resolution, especially if the PXD encounters any performance limitation before its replacement.
- Even the updated background simulation indicates that the SVD will exceed its safe occupancy target when running at full luminosity.
- The installation of the complete PXD will require a partial dismounting of the SVD. This is a delicate operation that needs to be planned carefully, since no SVD replacement is foreseen. The committee appreciates the effort put into the documentation of the assembly procedures. Despite the fact that many people involved in the assembly are permanent staff, missing any part of the documentation and training now may cause difficulties at the time of the PXD replacement.

4.1.2.3 Recommendations

- Efficiency, resolution and timing should be checked on the sensor that showed an increase in leakage current. It will be interesting to know if this is a peculiar case or if this behaviour will extend to other ladders.
- The SVD group is encouraged to closely investigate the reason of the different resolution and cluster size between data and simulation. Inserting missing effects in the simulation, like the Ramo potential, in order to better reproduce the cluster size, may provide an indication whether charge sharing could be the only source of the worse resolution. In parallel an analysis of the alignment constants and residues, especially in the overlap region, should provide an alternative estimation of the point resolution and more insights about systematic effects and residual misalignments.
- The committee values the proposal to test the 3-sample readout mode in the coming 2020a run and also invites the SVD group to evaluate the performance for the fraction of events having CDC-only timing.
- The ongoing effort on the documentation of the SVD deinstallation, reinstallation, and dismounting should be continued. It should be possible, for the persons that will take part in this operation, to exercise the relevant procedures, therefore it should be understood if the available parts and mock-ups are suitable for such a training. Documentation and training should also include the replacement of a ladder with a spare. This operation is not expected to happen, but it may be necessary if serious damage on some ladders will happen during the future Belle II running or as an accident during the PXD replacement.

4.2 Central drift chamber (CDC)

4.2.1 Status

The committee was pleased to learn that the CDC performed well with high efficiency during the autumn 2019c run. With 2000 ppm of water added to the gas mixture, layer 54 turned off, and reduced voltage (-25 V) in layers 53 and 55, the CDC base current during collisions was near or below $50 \mu\text{A}$, dominated by machine background. Due to these actions, there were no persistent growing CDC dark currents during the entire run, a particularly gratifying outcome of those decisions. Furthermore, there was no substantial degradation of CDC performance due to the addition of water and the potentially lower efficiencies in layers 53–55. Background improvements have substantially decreased CDC currents, compared to the spring 2019a run, and thus the rate of charge accumulation in the detector. Up to now, the total accumulated charge is ~ 5 mC/cm for outer and 10–20 mC/cm for inner CDC layers.

Moreover, background, hardware, and software upgrades have improved some key performance measures. Ratios of CDC tracking efficiencies per layer, divided into earlier and later Belle II running periods, are all consistent with 1.0 within statistical uncertain-

ties. In several layers, some “islands” corresponding to 1 FEE board, have efficiencies about 10% lower, most probably due to an electronics issue.

An ASIC cross-talk filter was introduced, which was optimised to retain actual hits, while rejecting most extraneous cross talk hits. This improvement substantially increased track finding efficiencies in the forward and backward regions beyond the central region. Short track logic in the trigger algorithm that was implemented in the 2019c run further improved trigger efficiencies in the same regions. Also, a 3D neural net was added to the CDC trigger in the 2019c run. It reduced the trigger rate by about 30%, with at most about a 1% reduction in hadron event efficiency.

Detector resolutions were measured with $e^+e^- \rightarrow \mu^+\mu^-$ events. Resolutions measured in data and Monte Carlo simulations are generally comparable. The mean value of $\sigma(p_t^*)$ improved substantially, while the mean value of $\sigma(d_0)$ shifted somewhat.

For the immediate future, a number of efforts to improve CDC performance and utility are underway.

One of the major sources of DAQ downtime is current spikes during injection, which trip the CDC high voltage and thereby induce DAQ downtime. For a few days, injection current spikes were studied with analog current monitoring in layers 14–17. This test established that the CDC current surges during LER injection (but not during HER injection) and that a current surge is also observed if the kicker magnet is activated but no LER current is injected. The CDC current monitoring is too slow to provide essential information to the machine group for effective injection tuning. An online leak current monitor with sampling rate > 1 Hz would substantially improve the information required for tuning. A new fast current monitoring device was developed at DESY and a prototype will be tested on two layers in the 2020a run.

The CDC 3D tracker became available after the 2019c run. With this trigger, z resolution is worse than it was with the 3D neural net. This problem is due primarily to the algorithm to measure event timing. An upgrade is underway and is expected to be ready in a few weeks. Implementation of the 3D tracker in the online trigger is expected within the next year.

CDC timing was introduced in global decision level L1 in the 2019c run to complement ECL timing. This introduced a strange peak in the SVD timing distribution, which is under investigation.

Some important histograms were missing from the online data quality monitor due to a software bug. Updated software is being tested in cosmic ray runs.

To clarify responsibilities, the CDC group established an organisational chart with essential tasks clearly delineated and assigned to individuals and collaborating institutions. Even though new groups have joined the CDC effort and assumed responsibilities for key tasks, additional personnel are urgently required, particularly in the software effort. To enhance communication, there are regular common group meetings that include individuals from both hardware and software efforts. The committee is pleased with this significant step forward in planning.

As described in the October 2019 BPAC, an older test chamber built for another purpose was adapted for ageing studies in order to get a handle on the cause of the CDC persistent dark current. To reproduce the CDC behaviour, a “target” gold-plated

Tungsten sense wire as well as the surrounding Aluminum field wires were replaced with wire left over from the Belle II CDC construction. Furthermore, the normal CDC gas mixture and flow rate were used. A region of the target sense wire (and the surrounding field wires) was irradiated through a hole in the chamber wall with a ^{90}Sr source. Two other holes in the chamber were made in order to compare ^{55}Fe spectra in irradiated and non-irradiated spots of the sense and field wires. After about 40 mC/cm of irradiation, the sense wire gain was reduced by about 20%, but there was no indication of a persistent dark current.

The target sense wire and two of the surrounding field wires were removed and studied with SEM analysis in order to determine the presence of various elements on the wire surfaces. For the field wires, no significant difference was found between the irradiated region and the non-irradiated ones. In particular, silicon or hydrocarbon deposits that could lead to the Malter effect were not observed on the field wires. On the other hand, Si traces were found on the irradiated region of the sense wire but no Si was found at the non-irradiated reference spots. The Si in the irradiated region is thought to be the reason for the sense wire ageing.

Since this test has not identified the origin of the CDC persistent currents, plans are underway to build a new test chamber with cell sizes, materials, and wires that are as close as possible to the Belle II CDC. Careful monitoring of oxygen and water content during the new ageing studies are also envisaged. According to the current schedule, it would be possible to resume ageing tests in May 2020 with the current He:ethane gas mixture. These tests would be followed with other tests involving alternate (e.g. non-hydrocarbon) gas mixtures. Belle II management is actively attempting to involve additional personnel in this test program.

Although there has been excellent progress with CDC operation and utilisation of data, much remains to be done, including some concerns from the October 2019 BPAC meeting that are still open.

4.2.2 Concerns

- Although stable CDC operation has been achieved for the time being, there is no guarantee that the measures adopted so far will be adequate to maintain CDC performance for the rest of the Belle II program. Special precautions and development of advanced online monitoring tools are required to reveal, as early as possible, the onset of ageing effects in the CDC chamber in order to timely prevent their propagation throughout the chamber.
- Shortage of personnel involved in the CDC effort remains critical.
- Studies of CDC performance with He:CO₂ or He:CF₄:CO₂ gases, as examples of mixtures without hydrocarbons, have begun but simulation work is still needed.
- The essential elements of how a replacement of the CDC chamber should be designed or built differently are still not clear.

4.2.3 Recommendations

- The efforts to understand the reasons for the high CDC dark currents and possible mitigation strategies must continue with very high priority. Construction of the new test chamber for ageing studies is mandatory in order to try to reveal the original reason for the CDC dark currents and to advance solutions for “in-situ” remediation of the problem.
- Every effort should be made to ensure that the level of personnel engaged in CDC hardware is adequate to address concerns with long-term future central tracking in the Belle II detector.
- The Belle II management should be prepared for a major repair or even construction of a new CDC.

4.3 Particle identification

4.3.1 Barrel particle identification (TOP)

4.3.1.1 Status

The committee commends the TOP group for their continued hard work and progress on several fronts since the last meeting. They continued to make substantial progress in solving operational issues and improving data taking, and appear to be ready for 2020a running. The Monte Carlo and data results show some convergence since the last meeting, and some specific possible sources for the remaining discrepancies have been ruled out, but substantial differences between simulation and data remain, and work to understand these continues. The replacement plan for the 224 conventional MCP-PMTs, which was originally scheduled for 2021, is being revisited in the context of the PXD installation also required at Belle II, which could occur in 2022. To insure survival of the tubes until then, and also to make sure all tubes survive until 50 ab^{-1} have been attained, the PMT rate for Touschek plus beam gas must be limited to 1 MHz/PMT for all future running.

TOP continues to work to improve operations for 2020 and beyond. A number of operational improvements have been made including an integrated GUI for TOP experts, with improved monitoring of injection rates, single event upsets, MCP-PMT performance, and streamlined calibration. Care is being taken to understand and mitigate the backgrounds. More experts are being trained. There is extensive effort toward understanding the TOP trigger and improving its efficiency. The TOP trigger is not yet planned to be included in the overall trigger for the 2020a run. The fraction of uncorrected single event upsets is about 1%. The biggest operational headache is SCROD processing resets or stops. There are a number of problems of varying severity with board stacks and more broadly with the DAQ that are not yet fully resolved. In most cases, at the system level these can be reasonably accommodated during running. In more specific cases, e.g. Board-stack 6a, the link is permanently lost, and repair must await full access that will come when the MCP-PMTs are replaced in the summer of

2021 or 2022. About 4% of the front-end channels need replacement. More details were given in the talk.

The overall TOP performance continues to be studied intensively. Kaon identification efficiencies and pion fake rates are being studied using identified kaons and pions from D^* tagged D^0 s and pions from K_S^0 events, as well as di-muon samples. Some improvements in data–simulation agreement especially in the low momentum region for pion misidentification have been observed with the most recent processing (proc10). However, substantial discrepancies remain overall. Detailed studies on single photon resolution, photon yield, chromatic dispersion, start time jitter and so on have been done, and continue. The group is trying to understand limitations to performance due to the use of analytical PDFs, and how these might be improved.

The production of the new MCP-PMTs is proceeding in a timely manner at Hamamatsu, and it seems that the testing and assembly work required can be completed well in advance of the 2021 (or 2022) summer installation period. The performance of the latter good batches seems to be adequate. However, the 70 PMTs from the earlier bad batches all need to be replaced with good ones, and it is unclear to the committee how or when this will be done. The solution may require additional funds as well as more testing. To deal with PMT lifetime concerns, as simulated by the background group, the limit of the PMT hit rate due to Touschek and beam gas has now been set to be 1 MHz/PMT until 2027 when 50 ab^{-1} are expected to be attained.

4.3.1.2 Concerns

- Accepting a running condition limit of 1 MHz/PMT rate for Touschek plus beam-gas backgrounds implies that the acceptable total backgrounds will change as a function of luminosity, requiring some care in its actual implementation during operations. Tracking the total accumulated rate for each tube would seem to be important, and a very useful diagnostic.
- It is still unclear that the bad batches of new PMTs can be fully replaced.
- Intense analysis work continues to be needed to understand the TOP performance and simulation better.

4.3.1.3 Recommendations

- Understanding and reducing LER backgrounds are very high priorities for the TOP and background groups, impacting both the short-term and long-term health of the TOP system.
- Continue to closely monitor the present round of ALD MPC-PMT production and testing, while working with Hamamatsu and lab management to obtain further high quality tubes as needed.
- Understanding the discrepancies in TOP’s particle identification performance between data and simulation is another high priority for the TOP group. With

sufficient data now becoming available, it should be possible to bin the PID data in different regions of phase space, and/or physical or component space, in order to clarify the origins of the differences.

4.3.2 Endcap particle identification (ARICH)

4.3.2.1 Status

The committee congratulates the ARICH team for significant progress made since the last BPAC meeting: understanding PID performance for the data already taken is progressing and the system seems to be well prepared for the coming run periods.

An update on the 2019a dataset performance was reported. The data–simulation disagreement for PID efficiencies in particular in the outer ring of the ARICH reported at the June 2019 BPAC meeting can be reduced by an alignment of the aerogel tiles in direction and in their tilt and then of the mirrors. However, the required z -shifts in the outer ARICH ring, ranging up to the cm range, appear to be unrealistically large, at least in part, and hence seem to be partially compensating for effects not taken into account by the procedure. The PDFs in the simulation used to calculate the Likelihood were tuned to agree with data. Nonetheless, the kaon and pion efficiencies/misidentification rates obtained from reconstructed D^* and K_S^0 samples are still over/underestimated by the simulation. The observed high pion misidentification rates are likely to come from tracks with scattering and/or decays in flight, which become especially relevant when the particle track points close to the gap region of the tiles. It is then possible that the data–simulation discrepancies are caused by an imperfect alignment of the aerogel gaps. This hypothesis is supported by the fact that the overall pion misidentification rate and the data–simulation disagreement decrease with decreasing pion momentum when only tracks are studied that are not close to the tile edges and in particular when also an ECL hit associated with the track is required. Confirming this requires future alignment studies. Further insight might be obtained by studying tracks of different track-fit quality.

In preparation for the fall data-taking period (2019c), the forward endcap was extracted, with all ARICH cables disconnected in July and reconnected in September. This resulted in 21 HAPDs out of 420 not working due to electronics/DAQ problems (5%). While one part of the related front-end boards can be recovered, e.g. by firmware changes, another part can only be fixed by a future physical intervention. For about 3.6% of the HAPDs, the bias and/or high voltage cannot be applied for various reasons, which in part can be easily fixed. The overall number of unusable HAPDs for the 2020a run is estimated to be around 4%.

4.3.2.2 Concerns

- The required alignment shifts are unrealistically large in some cases.

4.3.2.3 Recommendations

- To reduce and also to understand the data–simulation disagreements in the ARICH PID performance, a systematic alignment strategy should be developed, allowing alignment parameters to change only within realistic boundaries. Such an alignment strategy may profit by using not only muon tracks but also, e. g. pion tracks categorised by track-fit quality, which may help to enlarge the sensitivity to tile gap misalignment. A careful study of the track-fit quality of the muon tracks used for the alignment may help to understand possible dependencies of the ARICH alignment on potential VXD and CDC misalignments. Separation by track charge might help identifying other sources of alignment biases.

4.4 Electromagnetic calorimeter (ECL)

4.4.1 Status

Successful operation of the ECL in Phase 3 was reported. All the counters are working properly and the DAQ system is running stably. The committee congratulates the ECL group on this excellent achievement. The pileup noise due to the beam background, which may cause degradation of the energy resolution, was a concern. The situation of the beam background in 2019c is improved with respect to the runs in spring (2019b), yet an increase of the background is observed due to the higher beam currents. Proper online histograms to monitor the hit rate as a function of time after each beam injection were introduced, and they are intensively used for machine tuning. Other useful histograms for data quality monitoring have been deployed and are useful for spotting even minor issues. Luminosity monitoring is one of the important tasks taken by the ECL. Reliable information has been provided through the Run 3 data-taking period based on the energy calibration system using test pulses.

Detailed studies of the timing and energy measurement of the ECL are progressing and careful comparisons between data and run-dependent simulations have been performed. The timing calibration has been studied using the process $e^+e^- \rightarrow e^+e^-$ for the ECL counters located in the acceptance of the CDC. Cosmic events were used for the timing calibration of the remaining ECL counters. This method will be replaced by another method using $e^+e^- \rightarrow e^+e^-\gamma$ events. The timing distribution of the data agrees with the simulation reasonably well, yet a small difference is observed. The peak position in the distribution of the photon cluster timing is off by one nanosecond. Further detailed corrections, time-walk corrections for each crystal and more realistic timing distributions of the background hits, will be introduced for achieving better agreement. The ECL team is working on the leakage correction which plays a crucial role to achieve an accuracy of better than 1%. The correction factors are obtained based on simulated events. The reconstructed invariant $\gamma\gamma$ mass peaks of the processes $\pi^0 \rightarrow \gamma\gamma$ and $\eta^0 \rightarrow \gamma\gamma$ show good agreement between data and simulation. The current inconsistency of the energy resolution between data and simulation is 0.4 MeV for π^0 and by 0.8 MeV for η^0 .

4.4.2 Concerns

- Although the situation of the beam background is getting better in the period 2019c, an increase of the noise is observed as a result of the higher beam currents.

4.4.3 Recommendations

- The ECL group together with the accelerator group should continue working on further reduction of the beam background by monitoring the situation and finding better machine parameters.

4.5 K-long Muon system (KLM)

4.5.1 Status

The report on the status of the KLM system was encouraging. Operational efficiencies have remained since the last meeting at 96.7%, 99.9% and 99.6% for the barrel scintillator, barrel RPC and endcap scintillator systems, respectively, resulting in an overall efficiency of 99.5%. The large latency issue, leading to the inability to include the KLM system into the global trigger, has been identified and was caused by the 2D hit finder. By bypassing the 2D hit finder, a trigger timing resolution of 20 ns was obtained. Full inclusion in the global trigger is planned for the spring.

So far the KLM firmware has only utilised hits on detector elements, with stable operation up to 50 kHz with occupancies of 2-3%. Full waveform sampling has been developed and is currently being tested with a target deployment in the spring. The slow control system has been updated and new functionality, such as the selective powering of individual crates, has been added. The online data monitoring has also been updated.

One source of the very low efficiencies of the barrel scintillator layers has been determined to be a software issue concerning an inconsistency between the software and the database. Although the efficiency has significantly improved, it is still much lower than the efficiency observed in the endcap. This is not understood. There are also troubling variations in efficiency in both the EKLM and BKLM. These are currently attributed to a poor understanding of the magnetic field.

The muon identification is well-advanced and steadily improving. The difference between data and Monte Carlo is currently less than 2% for momenta higher than 1 GeV, but larger for lower momenta, which is being studied. The K_L^0 identification is being studied using a boosted decision tree (BDT) method. The BDT analysis can only be done with run-dependent Monte Carlo samples that are being generated.

The organisational structure of the KLM system was reviewed and clear individual responsibilities were assigned, which should provide for an adequate coverage of all tasks.

4.5.2 Concerns

- Uncertainties in understanding KLM efficiencies still remain. The two key issues are the low efficiency for the BKLM scintillators and the variation of the efficiency in both the barrel and endcap.

- Based on current data, and hence with quite some uncertainty, a loss in efficiency of up to 25% is expected at design luminosity, which is unacceptably large.

4.5.3 Recommendations

- A very systematic study of the efficiency loss of the scintillator barrel KLM and the variation in efficiency in both the endcap and barrel is strongly recommended.
- A thorough study of the RPC efficiency at design luminosity should be undertaken. The currently anticipated efficiency loss seems unacceptably high.
- A stronger effort on the K_L identification is encouraged.

4.6 Detector Performance

4.6.1 Status

The overall efficiency of the experiment in taking good data is only about 50%, which must be improved. It includes significant losses due to detector issues such as down-time for calibration, dead time, and bad runs due to partially missing data or low detector efficiencies. The committee was glad to see that there is an active plan to address this, including ensuring expert shift coverage for each detector. Backgrounds in the detectors are high in the current mode of running, as the luminosity is increased as far as backgrounds permit. Further increases in luminosity will continue to be limited by backgrounds, therefore it is important to understand how best to handle the origin and impact of the current and future levels of machine background.

The recently established Performance Group aims to provide a stronger link between the individual sub-detector groups and the physics analysis groups by exploring in detail the reconstruction and identification of the produced particles, assessing their efficiency and resolution, and most importantly identifying various origins of performance losses, especially for measurements that involved more than one sub-detector. Comparisons between data and simulation samples are critical for the understanding of the basic detector performance and calibrations, the development and optimisation of the analysis and simulation software, and ultimately the assessment of the systematic uncertainties for signal and backgrounds in various analyses.

The new group under experienced leadership has laid out a broad program, to be supported by close collaboration of the detector and analysis subgroups. The plan covers charged particle tracking and identification, neutral particle detection and energy measurement, as well as the measurement of certain beam parameters, such as energy and spot size and position. All these studies are fundamental to the full exploitation of the expected large Belle II data samples on the way to exciting discoveries. It is hoped that the size and expertise of this group and its tight linkages to the detector and physics subgroups will grow substantially to meet the stated goals.

Recent studies of the detector performance show progress in several areas and are encouraging, though the simulation results are usually more optimistic. The results were mostly presented in terms of efficiencies as a function of momenta and angles, and

rather little on resolution or uniformity of the response in various sub-detectors. Detailed information on calibration and alignment was not presented. All this information is crucial to assess systematic uncertainties when combined with evaluations of signal selection, background characteristics, triggers, run dependence, etc. Currently the trigger performance appears not to be addressed by this working group.

4.6.1.1 Charged Particle Tracking

A variety of tag-and-probe studies of charged particle tracking efficiencies are underway. The best result is based on radiative Bhabha events with momenta above 2 GeV, resulting in good agreement with simulation at the level of 1%. Similarly encouraging is an analysis using muons from tagged tau-pairs, the data agree very well with simulation in momentum range between 0.2 and 2 GeV, based on the latest data and simulation releases. No studies of systematic uncertainties or track fake rates were presented. The analysis of low momentum pion tracks from D^* decays (< 200 MeV) is underway but due to issues with stability results were not shown yet.

In parallel, tracking efficiencies, combining CDC tracks with hits in the SVD or the single-layer PXD have been studied. The SVD track finding was checked for good CDC tracks. The resulting efficiencies for CDC and SVD track segments are close to 100% and stable over two run periods. A recent trial to match SVD-CDC tracks with hits in the PXD indicated an efficiency close to 97%.

A study of K_S^0 and Λ decays as a function of the decay distance showed sizeable differences between data and simulation for the reconstructed values of the mass and width for K_S^0 decays, and much smaller effects for Λ decays. This difference may be due to much tighter track selection criteria for the Λ (> 20 hits) analysis than for the K_S^0 (> 0 hits). The reconstructed mass for 4-body D meson decays also indicates a mass shift, by $+1.3 \pm 0.1$ MeV. This obviously needs attention, given that the detection of displaced vertices is not only a powerful test of charged particle tracking but is also critical for lifetime measurements and heavy flavour physics.

4.6.1.2 Photon Detection and Timing

Studies of the detection of photons and π^0 have been refined. A recent study of γ detection in $\mu^+\mu^-\gamma$ final states indicates a reasonable agreement (at the level of 5%) between data and simulation above 2 GeV. At lower energies, backgrounds from ISR photons, not fully simulated lead to larger differences. Studies of the reconstruction of two-photon decays of π^0 and η continue, and the most recent results indicate good agreement between simulation and data. Based on recent calibrations, comparisons of the photon cluster timing between data and simulation show agreement at the level of 1 ns for photons in hadron final states. This small difference may be attributed to imperfect timing of individual crystals and worse than expected energy resolution for simulated background events.

4.6.1.3 Lepton Identification

A number of studies of charged lepton identification are underway, based on a variety of event samples. Above 0.8 GeV, there is now good agreement between data and simulation (using recent software upgrades) with efficiencies of 94% and mis-ID at the level of 11%, based on clean $\mu^+\mu^-\gamma$ event samples. Studies of other PID methods and different event samples typically show data efficiencies of $> 95\%$ at all energies for e and μ , but below 1 GeV disagree on pion mis-ID. Unfortunately no comparison with simulation for this study was shown. In a sample of four-lepton final states, efficiencies above 95% for momenta above 1 GeV are in good agreement between data and simulation for e and μ . These analyses do not make use of the KLM timing information, which is a powerful tool to reject backgrounds. The dE/dx information from the CDC would be impacted by saturation effects and potentially also charge asymmetries.

4.6.1.4 Beam Spot Studies

Measurements of the beam spot position and dimensions are critical input to many physics studies and need to be monitored. Based on the fitted primary vertex in $e^+e^- \rightarrow \mu^+\mu^-$ events, the beam spot position and shape are derived. The measured x ($\sim 20 \mu\text{m}$) and z ($\sim 300 \mu\text{m}$) widths are dominated by the beam size, whereas with the vertical width y ($\sim 10 \mu\text{m}$) is dominated by the track measurements. The measured resolution depends critically on the angle of incidence. Simulation studies would be informative.

4.6.1.5 Hadron PID

A working group has been set up under the auspices of the Performance Group, to study the performance of hadron identification and provide feedback to the sub-detector groups. It is currently composed of a convener and five members but only 1.5 FTE, which looks low. The pion, kaon and proton identification efficiencies versus momentum are in reasonable agreement with the expectations from simulation at the few percent level, but the misidentification rates are significantly higher, e.g. by a factor of 1.5 to 2 for the pion mis-ID rate to kaons. The complex reconstruction of the TOP detector still requires significant further study. While there have been considerable improvements to the understanding of the recent ARICH data, there remain significant differences in the PID efficiencies of data with lower kaon efficiency and higher pion mis-ID compared to the simulation. Some of these differences have been attributed to the poor alignment of the aerogel tiles and mirrors, and recent realignments have significantly reduced the differences. There remains a strong momentum dependence as a function of the radial position of the incident particles.

4.6.2 Concerns

- The establishment of the Performance Group is clearly a very important part of the effort to prepare Belle II for the physics analyses. It is planned to develop a coherent approach to a detailed understanding of the performance of the whole detector and to improve the analysis and simulation software and triggers to be

used for physics analyses. Given the complexity and novelty of many detector components and associated electronics, calibrations and software, a larger group would allow more diverse and in-depth analyses of the current and potential future problems.

- Detector performance depends critically on calibrations at all levels, from single channel response and calibrations to internal alignment of sensors as well as the relative alignment of sub-detectors. At this meeting, very little was presented on the status, the various methods, sensitivity, and stability for each sub-detector. In particular, this was the case for the CDC. Misalignment was quoted more than once as the source of some efficiency problems.

4.6.3 Recommendations

- The current performance tests are largely focused on efficiency studies, based on low multiplicity, low background events, which is appropriate for initial studies. Resolution and variations of detector responses should be assessed, not just in terms of χ^2 , but with attention to residuals of the fits, number of hits per track and crystals per ECL cluster, at different energies and angles, and to the impact of backgrounds and potential sources of various observed anomalies. The presence at KEK of more members of the sub-detector groups would improve the recorded data and subsequently the physics analyses.
- The ARICH detector efficiencies point to a problem that might not be caused only by misalignment. An assessment of potential differences between the measured magnetic field of Belle and the actual Belle II field should be considered. The installation of new beam magnets near the IP might have changed the magnetic field locally, impacting the trajectories of low momentum particles in the forward directions¹. Separate samples of positive and negative tracks might also shed some light on this problem. Possibly, some data recorded at lower field or no field might also provide some insight.
- Machine backgrounds obviously affect the data analyses. Assessment of the backgrounds in the sub-detector rates and signatures and their impact on the performance of measurements and physics analyses need to be assessed now and in the future.
- The trigger does not appear to be part of the portfolio of the Performance Group, even though it is a critical part of the detector performance and impacts the sensitivity to different types of final states. A strong link to the trigger subgroup and some analysis groups should be assured.
- The close link between the sub-detector groups and the Performance Group should be strengthened. The goals, schedules and the sharing of the primary responsibil-

¹The committee was later informed that there had been a dedicated campaign to measure the combined B-field in the VXD volume and around the QCS cryostats in 2017.

ities should be formulated and appropriately adjusted with time. Equally important is the link of the Performance Group to the Physics Analyses Groups, to set benchmark criteria for optimisation of particle reconstruction and identification, methods of the assessment of systematic uncertainties, measurements of various beam parameters and determinations of luminosity, all impacted by the presence of considerable backgrounds and radiation damage. To cover these broad responsibilities, enhanced presence of group members at KEK at these critical times will be very beneficial, as it would enhance the quality of the early data and thereby foster interesting first physics results.

- The focus of the sub-detector groups should be the operational readiness of the sub-detectors, including calibrations, alignments, and analyses of basic reconstruction of particles and comparison between simulation and data. All calibration results should be integrated into a global system allowing for stability tests at any time. The committee would welcome a systematic presentation of the status of alignment and calibration studies, starting from the lower-level studies within the sub-detector groups before progressing to the combined performance and physics analyses.
- The sub-detector groups need to be strengthened to identify and address various anomalies, most importantly the CDC tracking and dE/dx , the TOP, the ARICH, and the KLM scintillator timing. More presence of sub-detector experts at KEK in the coming months would enhance the understanding in these critical areas and significantly advance the readiness for high quality physics data.
- Nevertheless, the overall impression of the committee is that the data quality is reaching a level where substantially more data taken under reasonably stable running conditions would be beneficial for physics analysis as well as calibration and also for more detailed studies and improvements of the performance.

4.7 Upgrade

4.7.1 Status

The status of the upgrade of the data acquisition system based on the PCIe40 architecture was presented. The team has been strengthened with additional personnel and liaisons with the sub-detectors have been identified. Good progress is being made on the firmware development. The mass production of the PCIe40 boards has started. The first boards are expected in April and will be tested at CERN. A staged implementation is foreseen, with full replacement of the current system by the fall of 2021.

The Upgrade Working Group, which was formed in October 2018, gave an update on their activities (see also Section 2.2.2). The current mandate of this working group is to coordinate activities in response to four considerations:

1. an evaluation of at what integrated luminosity a major detector upgrade would be required;

2. a review of technical solutions for sub-detectors that cannot withstand the expected background levels;
3. the physics impact of luminosities five times the design luminosity and how the Belle II detector would perform at those luminosity levels;
4. and an evaluation of options in response to potential forthcoming upgrades of the accelerator.

In response to the first concern, the operational limit of various sub-detectors is being studied based on the current background information. An estimate will be obtained for all detectors for the time frame when the performance has degraded so much that a replacement detector needs to be installed.

In evaluating technical solutions, most effort to date has been directed towards the study of a replacement vertex detector system based on depleted monolithic active pixel detectors. The group hopes to participate in the AIDA⁺⁺ program for the technology development. If mission need has been established for a new vertex detector, an Expression of Interest is considered for the 2021 time frame. Upgrade options for other systems are also being looked at, notably the radiation and beam abort systems.

The impact on the physics program of an increase in luminosity five times above the current design luminosity is also being studied. Many interesting physics channels, deviations from the standard model, and dark sector portals could be studied. The accessibility and sensitivity to the broader physics opportunities is to be quantified, and the requirements on detector performance should be set in relation to well-defined physics objectives. The working group recognises that significant upgrades to SuperKEKB would have to be made to achieve a five-fold increase in luminosity, which are quite challenging and for which there currently is no personnel.

The group also proposes to study the feasibility of electron polarisation in SuperKEKB. High longitudinal electron polarisation would provide enhanced sensitivity and open a unique and rich physics program. A white paper on the feasibility of polarisation in SuperKEKB is targeted for the fall of 2020.

The committee is pleased with these efforts exploring the physics case and technical feasibility of maximising the scientific return on the investment in SuperKEKB and Belle II.

4.7.2 Concerns

- A broad portfolio of studies is being proposed by the upgrade working group over a relatively long time frame with very few resources; this might affect its effectiveness.

4.7.3 Recommendations

- The upgrade of the data acquisition system to the PCIe-based configuration should be pursued with highest priority.
- The committee recommends considering the full potential of the very powerful PCIe40 platform to simplify and harmonise the readout and data acquisition.

- Most effort to date has gone towards a possible replacement of the vertex detector system. The working group is encouraged to specify the requirements for selecting the pixel detector replacement and demonstrate how the new technology will meet them.
- The working group is encouraged to narrow its scope and focus on the most promising areas that could have a substantial impact on the physics potential in the near term. The scope could be expanded when more resources become available. For each detector upgrade contemplated, the requirements on detector performance should be set in relation to well-defined physics objectives.

5 Computing and software

5.1 Online system

5.1.1 DAQ, TTD and HLT

5.1.1.1 Status

A lot of work has been done on the DAQ system. A new data flow has been introduced in the HLT based on ZeroMQ. This simplifies the overall architecture and will ease maintenance but has shown some infancy problems during the 2019c run, resulting in a deterioration in DAQ efficiency. A new monitoring system has been put in place based on the industry standard ELK stack. The pressing issues of personnel are being addressed by including junior members and the opening of M&O funded positions.

The Timing and Trigger Distribution (TTD) was working throughout, however with some issues remaining in the interface with the run-control and the slow control.

In addition to the work on the new data flow, the HLT team has increased the number of HLT units. Furthermore, it has performed some preliminary benchmarking of the HLT processing time and shown that a maximum HLT rate of 7.4 kHz was possible with currently available resources, without a Level-3 filter.

The Pixel detector Onsen RIO selection and communication with the HLT has been running passively for all the time during the 2019c data taking. It was set active for reading out the PXD data only for a limited period.

5.1.1.2 Concerns

- Judging by the level of problems associated with the new data flow system, the coverage of the tests done before was insufficient.
- There is a long standing issue with losing synchronisation on the *b2tt* links. Progress has been made by exchanging the cables for the worst cases, but there are still problematic links left, which led to some downtime.

5.1.1.3 Recommendations

- For the HLT framework, tests at maximum scale and using stress conditions should be done well in advance of the next physics runs.
- There was substantial downtime in central and sub-detector DAQ, partly due to injection. The work to improve robustness should continue.
- Powerful tools to monitor the data taking efficiency have been developed. Analysis and reporting should be done routinely as well as prioritising actions. This could possibly be done by the Run Coordination.
- The work on the monitoring system based on ELK and automation of error-handling should be continued with vigour.

5.1.2 Trigger

5.1.2.1 Status

Overall, the trigger has performed well in 2019c with stable efficiency and a global acceptance rate well within the limits of the maximum rate acceptable to the DAQ. The CDC 2-d and the newly-commissioned 3-d neural-trigger are working. The TOP, KLM and CDC 3-d tracker are still missing. Noise rejection in the CDC front-end successfully reduces the trigger rate for 2D-tracks significantly without loss in efficiency. The ECL trigger has been improved to provide a 3D Bhabha veto instead of a 2D. The KLM trigger-team has been strengthened and it is foreseen to be integrated for the 2020 data taking. There is still a significant amount of work required in the TOP-trigger and it will not be ready for 2020.

5.1.2.2 Concerns

- The TOP trigger requires still a large amount of firmware development with a risk of not being ready before 2022.
- The personnel situation has improved, but looks marginal in the mid-term for combined efforts in operation and preparing upgrades.
- Trigger problems with efficiency seem to go occasionally unnoticed by the shift-crew.

5.1.2.3 Recommendations

- It should be established in detail how much effort would be required to provide a trigger which is useful for the SVD data-reduction. If this effort seems to be too large or there is a risk of further slippage of the availability of the trigger, then the SVD team should be consulted to see if a different or modified solution can be found.

- The collaboration should continue encouraging institutes to contribute to the trigger sub-systems.
- The trigger efficiency should be monitored and automatic alarms generated to shifter console.

5.1.3 Control system

5.1.3.1 Status

The control system has made significant progress. The software deployment for the central DAQ is now based on cvmfs and a centralised continuous integration and deployment of infrastructure is in place. Restarting the full system can be done now in a few minutes. The database infrastructure has been consolidated and is now protected by a standby backup server.

Work on integrating the new PCIe40 board has started and is well under way. A schedule for integration with sub-detectors has been established.

5.1.3.2 Concerns

- Subdetector systems are not completely controlled using the central infrastructure, leading to the need of manual interventions and the risk of incoherent configurations.
- Subdetectors deviate from the agreed state-diagram for HV control, leading to lost time, because recovery actions cannot be automated.
- A review of potential hardware failures affecting the slow-control has been done. However, it seems there was no estimate done for the impact of a specific failure in terms of the luminosity loss.

5.1.3.3 Recommendations

- Subdetectors should be strongly encouraged to use centrally managed resources as much as possible.
- In collaboration with the sub-detector groups, harmonisation of the finite state machine should be pursued. This could lead to a higher run-time efficiency without compromising the safety of the detector through an increased level of automatization.
- A risk matrix should be created to establish where resources for improving the redundancy in the system are best spent to optimise the collected luminosity.

5.2 Offline software

5.2.1 General issues

5.2.1.1 Status

The committee was pleased to see that regular ways to ensure the exchange of information between the software, data production, performance and software experts in detector groups are being setup. A dedicated software and performance workshop with well defined discussion topics was scheduled to take place at the beginning of March at the time of the review; while it could not physically take place a slightly reformulated remote workshop replaced it.

A well defined release process has been established for the Belle II Software Framework, *basf2*, with a major release following the validation of its pre-release. There is a patch release with bug fixes and additional features for calibration. Alignment is then used in production and finally a light release with additional features for analysis can be deployed when necessary. The software release encompass reconstruction, simulation and analysis aspects.

The patch of release 4 was deployed in December with the 2019 Calibration Alignment Framework (CAF) updates and will be used for *2020 (a+b)* data taking, processing and analysis. A light release with updates of analysis tools was foreseen. At the same time, preparation of release 5 is in progress with the freeze of new features expected at the end of March with the major and patch releases planned respectively for end of May and June. It will include incremental improvements of the framework, simulation, reconstruction and analysis tools. This future release is intended for reprocessing of the *2020 (a+b)* in summer and will be the last deployed on the Scientific Linux CERN 6 *slc6* computing environment. The new server for Conditions Database hosted at BNL is running very reliably. An update of the interface for the use of Global Tag will be deployed in production before the start of data taking. Tools and procedures are to be improved in order to prevent most common mistakes in the management of the Global Tags. Online information is also to be transferred to the Conditions Database.

5.2.1.2 Concerns

- The committee is pleased to see that an agreement has been reached concerning the software license and copyright. The Belle II code can be made available under an open source licence and the copyright remains with the full list of contributors. This is totally adequate, however it could lead to complications in later years if the need to contact and get the agreement of all the copyright holders for any evolution in the software license would arise: the most prolific code contributors being PhD students and postdocs who may leave the HEP community.
- The group should continue its effort in motivating skilled contributors and attracting new ones. Continuity of essential tasks should be ensured also through overlapping transition periods for outgoing and incoming conveners, as much as possible. A replacement crucial for online software integration starting in mid-2020

had not been identified at the time of the BPAC review.

5.2.1.3 Recommendation

- The collaboration has developed a set of training and support tools for new collaborators. The committee encourages the software group to pursue its continuous effort in attracting, integrating and actively promoting new young members in the group.

5.2.2 Event reconstruction and calibration

5.2.2.1 Status

The committee was informed that all critical functionality is in place and work is mostly concentrating on optimisation of the reconstruction. In release 4, a cross-talk filter has been introduced in CDC tracking leading to a significant improvement in efficiency and purity, especially at low and high polar angle. Efficiency and p_T resolution have improved for tracks at large angle due to the introduction of a new SVD to CDC Combinatorial Kalman Filter (CKF). High and stable track finding efficiency was reported in both *exp8* and *exp10* with the latest processing. Although the agreement between data and simulation in hadron particle identification has improved with the latest processing *proc10/MC13a*, discrepancies are still present. Fixes for CDC calibration and ARICH alignment are being prepared in the reconstruction to address some of the discrepancies. Measured efficiency for neutral particles are in reasonable agreement between data and Monte Carlo. Photon leakage corrections are being implemented to correct cluster energies and improve γ and π^0 resolutions. In release 5, ECL track seeding will be introduced. Validation studies are being carried out to use the event T0 provided by the TOP sub-detector. that will be more precise than the current one given by the CDC. The improved T0 will then be used in calibration and reconstruction.?

The concern expressed by the committee on the lack of redundancy in expert knowledge for the CDC tracking calibration has been addressed with a second expert joining the effort. The automation of the prompt calibration using the Apache Airflow system has been deployed in production with a dedicated Calibration and Alignment Framework (CAF) based on dedicated *basf2* software components. The first fully automated calibration was in progress at the time of the review with a considerable reduction of *dead time* being observed. The goal of achieving each step within a two-days average delay was being met but with CDC calibration requiring more CPU than available from the KEKCC production queues. A strategy to speed it up was under test. Experts are still in charge of retrieving and checking calibration constants but a plan is in place to deploy the automatic production of calibration validation plots for final sign-off. Reliable operation is expected in *2020 a+b* as well as automating a few remaining manually performed calibrations. Consolidation of the whole system will be necessary with the migration of the Calibration centre from KEKCC to BNL in summer 2020 to cope with the core flat resources needed and the identification of new Airflow savvy developers and maintainers.

5.2.2.2 Concern

- The deployment of the new automatic calibration system has allowed streamlining the procedure and relieving the sub-detectors experts from routine tasks. The Apache Airflow technologies seems quite suitable to the job scheduling of the interdependent chain of tasks involved. Nevertheless even if based on Python, this requires the expert knowledge of a different product. The human resources needed for the maintenance and continuity of the system may put a drain on other existing systems.

5.2.2.3 Recommendation

- Detailed measurements of CPU performance with different levels of background show the reconstruction alone to take about 50% of the time of the combined simulation and reconstruction. While this can be partially explained by the characteristics of the experiment, it is worth to try to understand if given algorithms or functions are more CPU consuming than others and could be optimised or modified without compromising the physics results.

5.2.3 Simulation

5.2.3.1 Status

Major changes in the detector simulation are expected in the *basf2* release 5, in particular the update to a more recent Geant4 version and to a Belle II custom-made physics list. A new dedicated physics list was provided by SLAC's Geant4 experts and consists of an hadronic component of simulation optimised for performance with less accuracy at higher energy, due to the Belle II reduced requirements in this energy range. It also uses a faster standard Electromagnetic physics component offered by the Geant4 toolkit. Each change is evaluated separately using release 4 of *basf2* but a unique validation procedure is being used for all. Basic shower shape parameters for electrons, photons and charged pions as obtained from particles selected on the basis of reconstructed quantities from specific samples of $e^+e^- \rightarrow \tau^+\tau^-$ are consistent between the default and new physics list and Geant4 versions. The use of the new optimised physics list results in a 25% decrease in CPU time for the simulation. Verification of potential effects on dE/dx and particle identification of the changes are planned before their use in production in *basf2* release 5, for which validation of individualised thresholds per sub-detector is also envisaged. Updates of generators, in particular EvtGen will also take place in release 5.

The Belle II software group addressed the concerns expressed in the last review and setup a first dedicated meeting on sub-detectors software focusing on data-simulation comparisons steered by one of the simulation co-conveners. Through their software and performance liaisons, all sub-detector groups provided information for improving the agreement between data and Monte Carlo samples. Many studies have been performed using run dependent Monte Carlo and have identified some of the reasons of the discrepancies. Areas of development to introduce more realism in the simulation of

sub-detectors' response have been identified by all sub-detectors. Run dependent sub-detector response, e.g. noise and gains, are being implemented as database payloads. Updates to the geometry description of some sub-detectors are also foreseen for the next software release 5.

Longer term plans are to provide a consistent timing simulation among sub-detectors, since currently multiple clocks are used and timing effects by some detectors are modelled independently from each other. The use of timing information in random trigger data for background overlay will also be considered.

5.2.3.2 Concern

- An in depth validation procedure has been setup for monitoring the changes in the Geant4 detector simulation. While basic shower parameters and effects on higher reconstructed quantities are considered, other basic quantities such as multiple scattering and dE/dx should also be kept under control as their effects on the track quality and impact parameter resolution are non-negligible.

5.2.3.3 Recommendation

- The committee was pleased to see that a clear plan and effort has been put in place with all groups involved to better understand the discrepancies between data and Monte Carlo. Investigations and common studies with the performance group and detectors experts should continue and their outcome be accurately fed back to the simulation as well as possible.

5.2.4 Analysis

5.2.4.1 Status

Plans for short and mid-term developments for analysis tools are under discussion and were intended to be a good fraction of the topics to be addressed in the Software and Performance Workshop. In particular the possibility of setting up a dedicated framework for studies of systematics is under consideration. A fitting tools task force has been established since the last BPAC review in October 2019. A review of existing tools of interest has been carried out and feedback on the needs from the Physics Working Groups will be collected in order to recommend and support the most appropriate choices.

The use of the Conditions Database has been considerably simplified by the newly deployed Global Tag replay feature to automatically use the relevant information in the files metadata.

The Belle II collaboration is taking steps to ensure reproducibility of physics analysis, by proposing to add to the publication procedure the requirement for the review committee to verify that all software code used in an analysis is committed to a Belle II repository.

5.2.4.2 Concern

No particular concern.

5.2.4.3 Recommendation

- The committee commends the Belle II collaboration for the effort launched to review the needs of fitting tools for physics analysis. It encourages the software group to recommend and provide support for a tool-suite comprising the most relevant tools available in the HEP community.

5.3 Offline computing

5.3.1 Data Processing and Computing Operation

5.3.1.1 Status

Data preparation in 2019 nicely progressed toward automation and streamlining of the productions. The committee fully supports the priority of the Data Processing group for the coming year to provide timely and useful physics quality data in time for major conferences. Lessons learnt from 2019a+b processing with (*proc9/bucket7*) were exercised in the most recent *proc10*, in particular for ensuring the staging of the data before launching the submission of the processing jobs. Nevertheless the processing of all events took significantly longer than expected, highlighting few areas for further improvements such as monitoring of the infrastructures (e.g. cores, cvfms, automatic resubmission) and preventing human mistakes. The latest prompt *bucket8* processing was scheduled to start on the grid to process all events with parallel production of HLT skims at KEKCC soon after the review, after completion of the Airflow automated calibration. Analysis skims of good *proc10* data and *MC13a* run-independent Monte Carlo were expected to be completed soon after the BPAC review. For 2020 the goal is to simultaneously produce *mDST* and analysis skims to reduce the elapsed time needed to prepare the data for physics analysis. The new run registry was finalised by the February B2GM and will be deployed for the *2020a* data taking: it will be the one and only authoritative source of run quality information for Data Processing.

The *MC13* Monte Carlo production started in November 2019: a run-independent *MC13a* generic sample of 1 ab^{-1} had been completed and sizeable requests of up to 5 ab^{-1} by the physics group was being produced on the GRID resources. A smaller *MC13b* run-dependent production of generic samples (10 fb^{-1}) for development and performance studies was being carried out. The number of concurrent jobs at peak productions was 25000 on about 50 remote sites.

The dataset searcher has been extended to support finding Monte Carlo samples on the grid, considerably simplifying the update, maintenance and retrieval of the information. A Confluence based summary page for data readiness to be used for public results has been setup.

The whole chain of data transfer is now completely automated and ROOT formatted data is available on the the GRID within a day of completion of a run. The origin of

the discrepancy between the number of events registered online and offline is now fully understood: the online value will be taken from the HLT storage rather than the online *run record* due to latency between the issuing and the actual occurrence of the stop of a run.

5.3.1.2 Concern

- Having a unique entry point for retrieving the most up to date recommended samples for analysis is commendable. Nevertheless the long term sustainability of a web page to be edited by hand is a worry.

5.3.1.3 Recommendations

- The workflow for distributed *gbasf2* jobs by users implies manual operations to be carried out with care, e.g. submitting many jobs for a given dataset and merging of the results once downloaded locally. The committee recommends investigating whether existing tools already used elsewhere by the Belle II collaboration could be adopted to steer the workflow and to automate these repetitive actions, as this would allow to leverage on existing expertise. If leveraging and consolidating the existing expertise would not be possible, tools used by other experiments could also be considered.
- The computing group introduced an expert shift training course and called for contribution to the operation of the system. While the situation is gradually improving, there is still some shortcoming and daily operation continues to rely on computing developers. Additional support for Distributed Computing will also positively impact data processing and more personnel should be secured to ensure appropriate coverage for operation in 2020 and smooth offline data production.

5.3.2 Offline Computing System

5.3.2.1 Status

To cope with the high failure rate of users' jobs, a new functionality will be introduced with submission of the full set of users' jobs launched only after a small number of *scout jobs* terminate successfully.

Various improvements in the production system are planned, e.g. to cope with the choice of background input files for production of run-dependent Monte Carlo samples. Further automation of tasks are envisaged, in particular for the staging of raw data before processing with the BelleDIRAC Production System triggering the staging on disk.

Data Distribution over the grid with automated replica policy is in place in the current Distributed Data Management (DDM) and requires about one hour per week of human intervention by a data manager expert. The staging of RAW data from tape to the disk requires more automation and enough disk space to run a reasonable amount

of jobs concurrently. The committee expects that the adoption of Rucio would also help on the implementation of the automatic staging.

The collaboration has changed the deployment plan of Rucio from two phases to a single one following a review of the strategy by the expert that has joined the BNL DDM team. The main reason for this change is to minimise the risk of file catalog divergence, and to benefit of the full Rucio functionalities quickly and thus to relieve some of the current operation load. The committee fully supports the new deployment plan.

5.3.2.2 Concern

- It is not clear how the Data Production group will cope to submitting many smallish productions, nor if a strategy for handling backlogs due to large Monte Carlo and to skims productions draining the system is being devised.

5.3.2.3 Recommendations

- The stability of the production system against high volume productions should be closely monitored.
- The committee agrees with the new Rucio deployment plan and supports a single phase migration after all the defined milestones have been achieved, which is expected before the B2GM meeting at the end of June. Therefore a migration by early September looks very plausible.

5.3.3 Computing resources

5.3.3.1 Status

The replacement of the KEKCC resources was confirmed for summer 2020 with an increase of factors of 1.5 in CPU cores corresponding to a factor of two in HS06 units and a factor of two and 1.4 in disk and tape storage capacity, respectively. Based on the experience from the previous replacement in 2016, disruption in the operation is expected to be small.

Detailed measurements have been carried out for the latest production software release, *release-04-00-01* with different levels of background and event fractions as expected, in data and simulation. They provide input for new estimates of the experiment computing resources needed for the next Japanese fiscal year. Changes from Geant4 CPU optimisations, updated reconstruction algorithms (e.g. cluster selection, ECL seeding, electron tracking) together with high level (e.g. V0 finder) and low level software optimisation have also been taken into account. The analysis of the computing resources request and recommendations will be presented in a dedicated report.