Belle II CPU Benchmark
Randall Sobie   University of Victoria

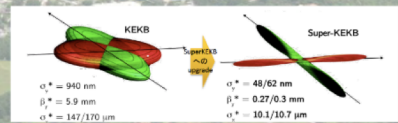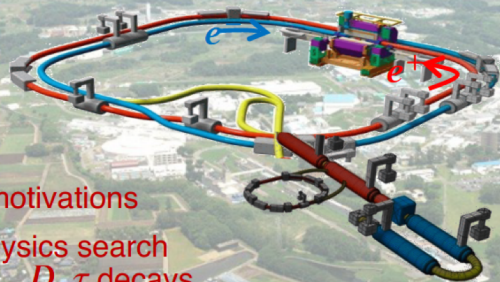Belle II is a detector designed to record electron-positron collisions at 10.58 GeV centre-of-mass energy (asymmetric, high intensity collider)

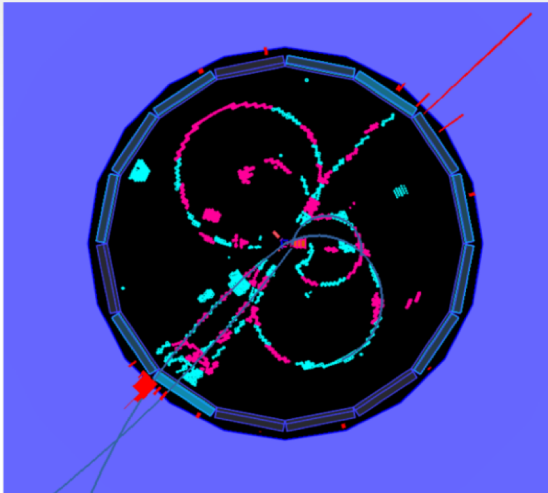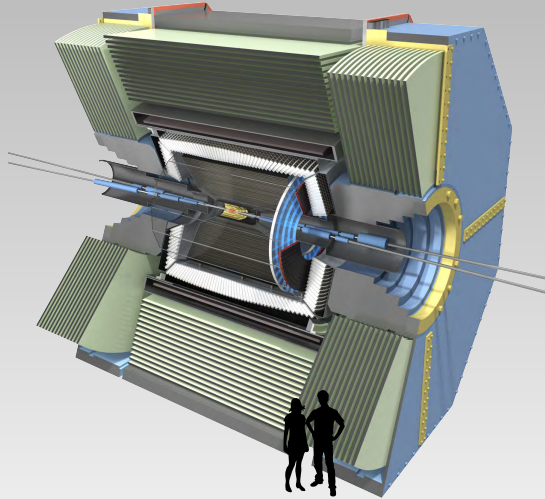KEK Laboratory in Tsukuba Japan

Early years of operation
90 fb-1 of integrated luminosity
Need to install pixel vertex detector in 2022
Goal is 50,000 fb-1 in this decade

Previous generation (Belle/KEK and BaBar/SLAC) collected less than 1,000 fb-1

Belle II – event topology

Low multiplicity events:
Di-lepton (2 tracks per event)
Tau pairs (2-10 tracks)
qqbar events (10-50 tracks)

Backgrounds from photons generate by the beams
(no pile up events)

Software has typical HEP work flow
(Gen, Sim, Reco, Analysis)

Gen is very small
Analysis is unpredictable (probably not for a benchmark)

How does Belle II use it benchmarks

Computing pledges are determined by number of PhD researchers

Best effort basis: many researchers from developing nations
(lots of "small" sites)

Many sites provide resources to the LHC experiments
Some sites report usage to the WLCG accounting system
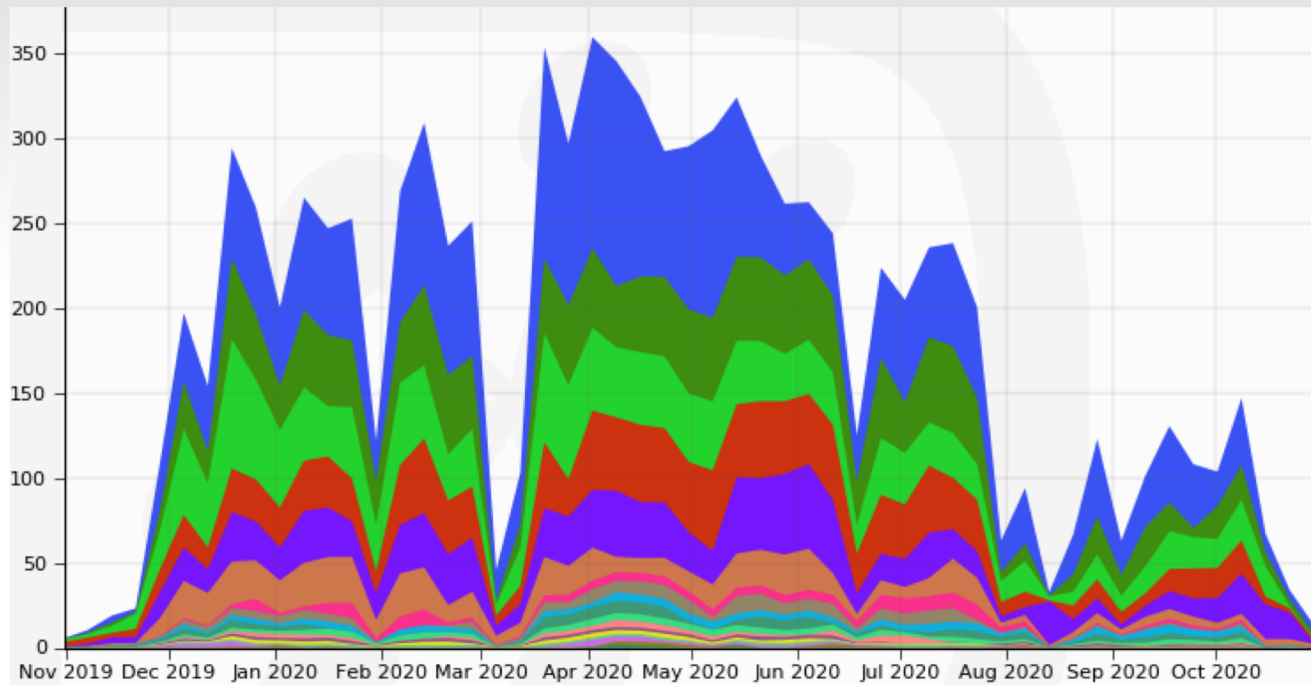
Belle II uses the DIRAC workload manager

DIRAC runs the DB12 (LHCb) benchmark at the start of each pilot job to estimate HS06

Usage plots show the DIRAC DB12 data

Measurements sensitive to CPU usage on nodes

Okay for now but could be better

Belle II benchmark

Identified a representative workload from Belle II
MC production of B0 B0-bar mesons (b-quark mesons) – busiest type of events

Event generation (Gen)   2%
detector and trigger simulation (Sim)   37%
reconstruction (Reco)   61%
Percentage is based on one run of 48 copies on a local 48-core machine
Major software releases once a year that impact the event/second result
CentOS7 - use singularity for jobs that need SL6

B2 runs single core jobs for all workloads

B2 benchmark runs N-jobs to fill N-cores
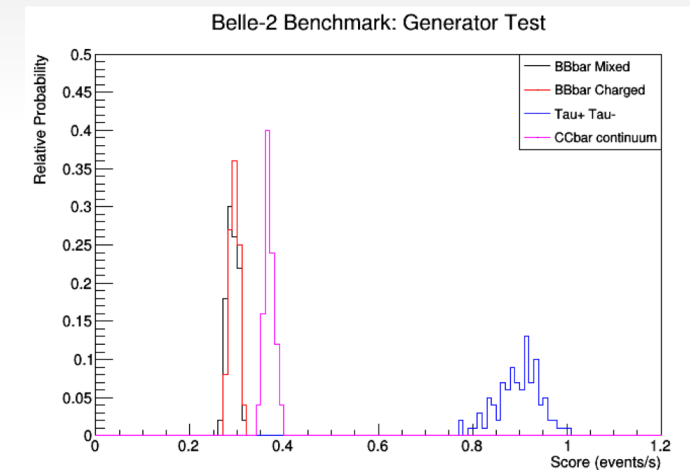No input data, CVMFS not required

Belle2-gen-sim-reco: 50 events with 1 thread takes 5-10 minutes
Not sure it makes sense to run (Gen/Sim/Reco) separately

Currently no use of GPU's, HPC's or ARM processors (maybe later?)

B2 Benchmarks on different UVIC VMs
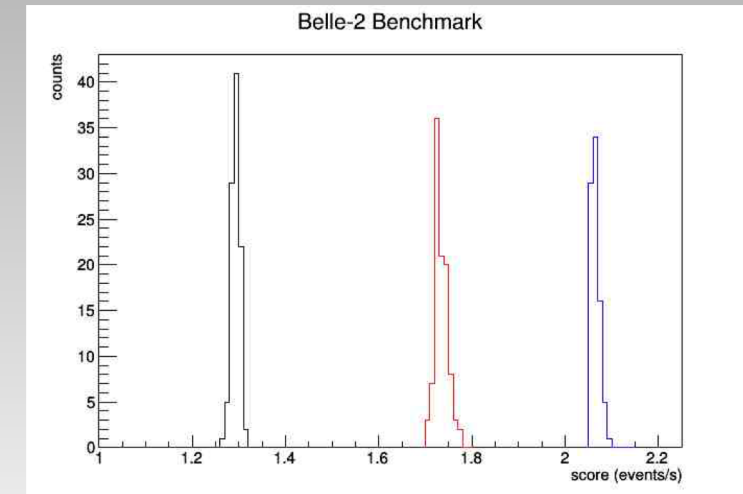

Different physics generators

Integration with Benchmark Suite

We (UVIC post-doc and me) have joined the Working Group
Regularly presenting results to B2 Collaboration

Followed the instructions of the Working Group to containerize the B2 workload
Runs multiple copies of code (one for each core)

Plan to integrate B2-Bmk into the Suite in the next release
Analyze data on more CPU-type and compare with other expts

Trying out the existing WG analysis code
Contribute to interpretation of the data
What constitutes a "good" benchmark?

Correlation of HS06 With B2 Benchmark
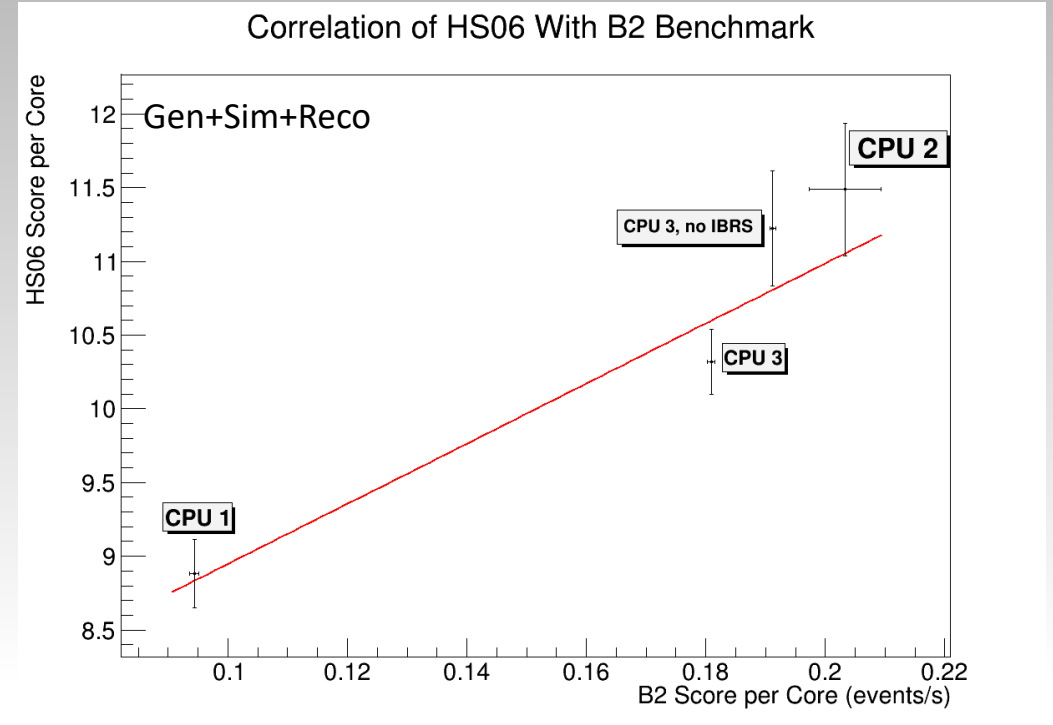
Ratio of Sim/Reco constant
But need more data