1 Nov. 2023

Takanori HARA (KEK IPNS / SOKENDAI)

Belle II computing

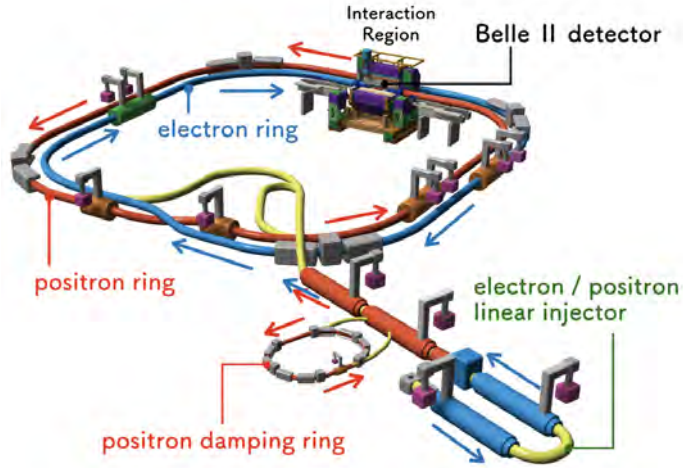
Slides are collected and based on recent presentations done by
I Ueda (KEK) , Cedric Serfon (BNL), Silvio Pardi (INFN / Napoli),
Hiroaki Ono (the Nippon Dental University)

The 7th Asian Tier Center Forum (ATCF7) at

Introduction

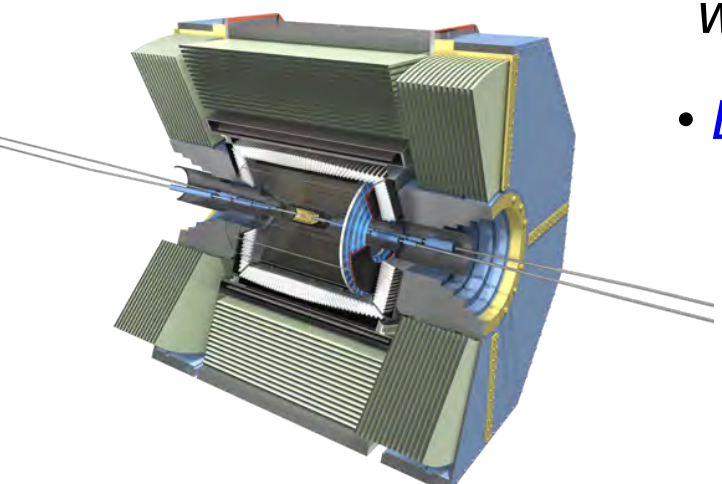
Belle II is a B-factory based at KEK (Tsukuba Japan)
to search for new physics beyond the Standard Model

target integrated luminosity : 50^{-1} ab = 10^{11} B meson sample
(corresponding to ~10 PB RAW data per year @ designed instanteneous luminosity)

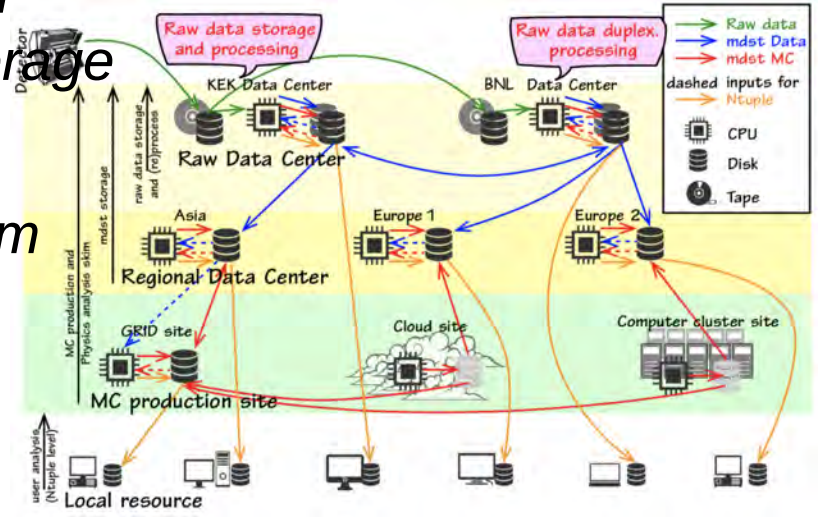


- *"SuperKEKB" accelerator : Asymmetric energy e^+e^- collider*

- *"Belle II" detector : a general purpose particle detector with almost full solid angle coverage*



- *Belle II computing :*
  a distributed computing system based on grid services with heterogeneous resources

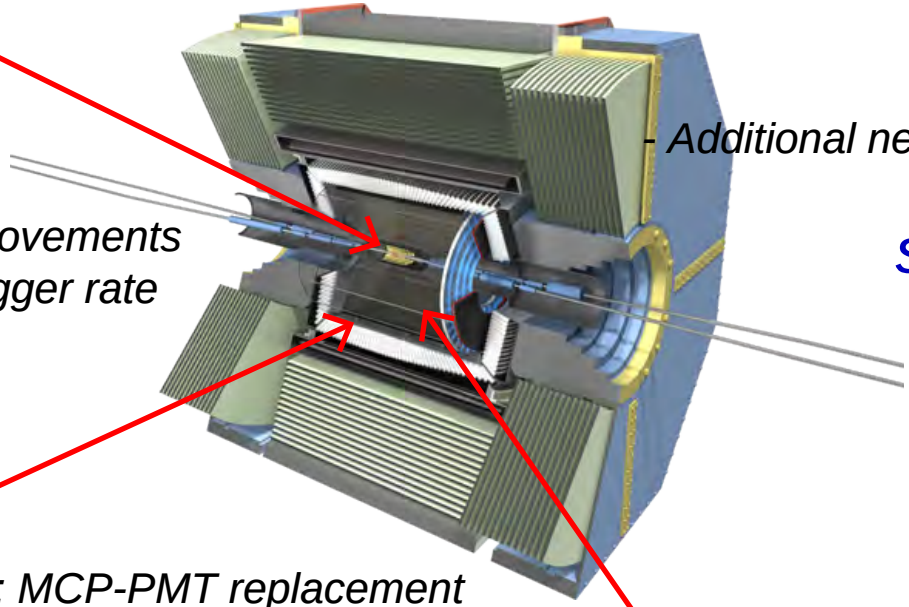


Belle II status : now in Long-shutdown

Since July 2022,

Belle II detector in a long maintenance

- *Vertexing (PXD)*
  Installation of new "full" Pixel-type vertex detector



- *Additional neutron shields*

- *DAQ : improvements for higher trigger rate*

- *PID (TOP) : MCP-PMT replacement*

- *Tracking (CDC) : improvement to mitigate the gain drop due to beam background increase*

What Belle II has achieved by July 2022

-reached world record instantaneous luminosity:
$4.7 \times 10^{34}\ cm^{-2}\ s^{-1}$,

collected up to $15\ fb^{-1}$ per *week*

-recorded luminosity at Belle II: $424\ fb^{-1}$

(Belle $988\ fb^{-1}$, BaBar $513\ fb^{-1}$)

same as the SuperKEKB accelerator ...

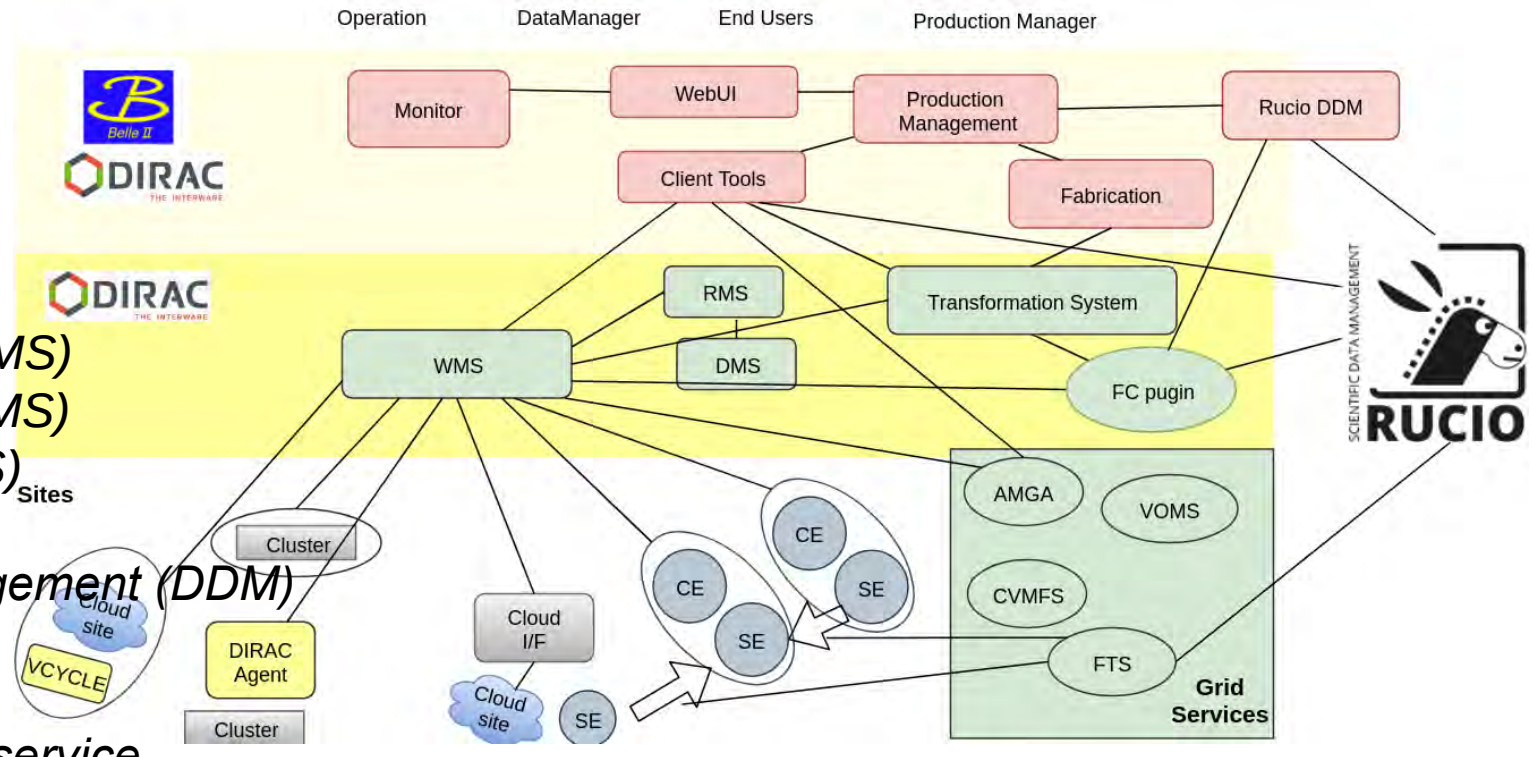
- *IR radiationn shield modification for BG reduction*
- *Non-linear collimater for impedance and BG reduction*
- *New beam pipes with wider aperture at HER (electron beam) injection point for improvement of injection efficiency*

Plan to restart the operation in Jan 2024

Belle II Collaboration

Computing Architecture

Belle II uses DIRAC with a specific extension called BelleDIRAC



- *BelleDIRAC as*
  - *WorkLoad Management (WMS)*
  - *Production Management (PMS)*
  - *Request Management (RMS)*
  - *Data Management (DMS)*
- *Rucio as Distributed Data Management (DDM)*
- *Rucio as File Catalog (RFC)*
- *AMGA as metadata Catalog*
- *CVMFS as software distribution service*
- *FTS for fle transfers.*
- *VOMS as Virtual Organizations and attribute-based authorization*

Distributed computing infrastructure at Belle II

Surveyed In mid April 2023

18 countries / region are contributing

- 55 sites providing pledged and opportunistic resources
- 29 Storages
- 5 Tape systems

TYPE	Resource provided	c.f. Pledged for 2023 JFY
CPU Pledge	451.6 kHS06/kHS23 + Opportunistic CPU	404 kHS06
DISK	16.8 PB*	18.9 PB
TAPE	11.9 PB	8.5 PB

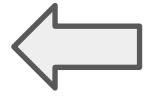
31 kjobslots pledged and 33 kJobslots opportunistic

NEW CHALLENGES FOR SITES

- Token Based Authentication
- End-of-life of storage technologies (DPM, gsiftp, srm)
- Update the Operative system (RHEL9/Almalinux9)
- Network Operation (Link update, Jumbo Frame)

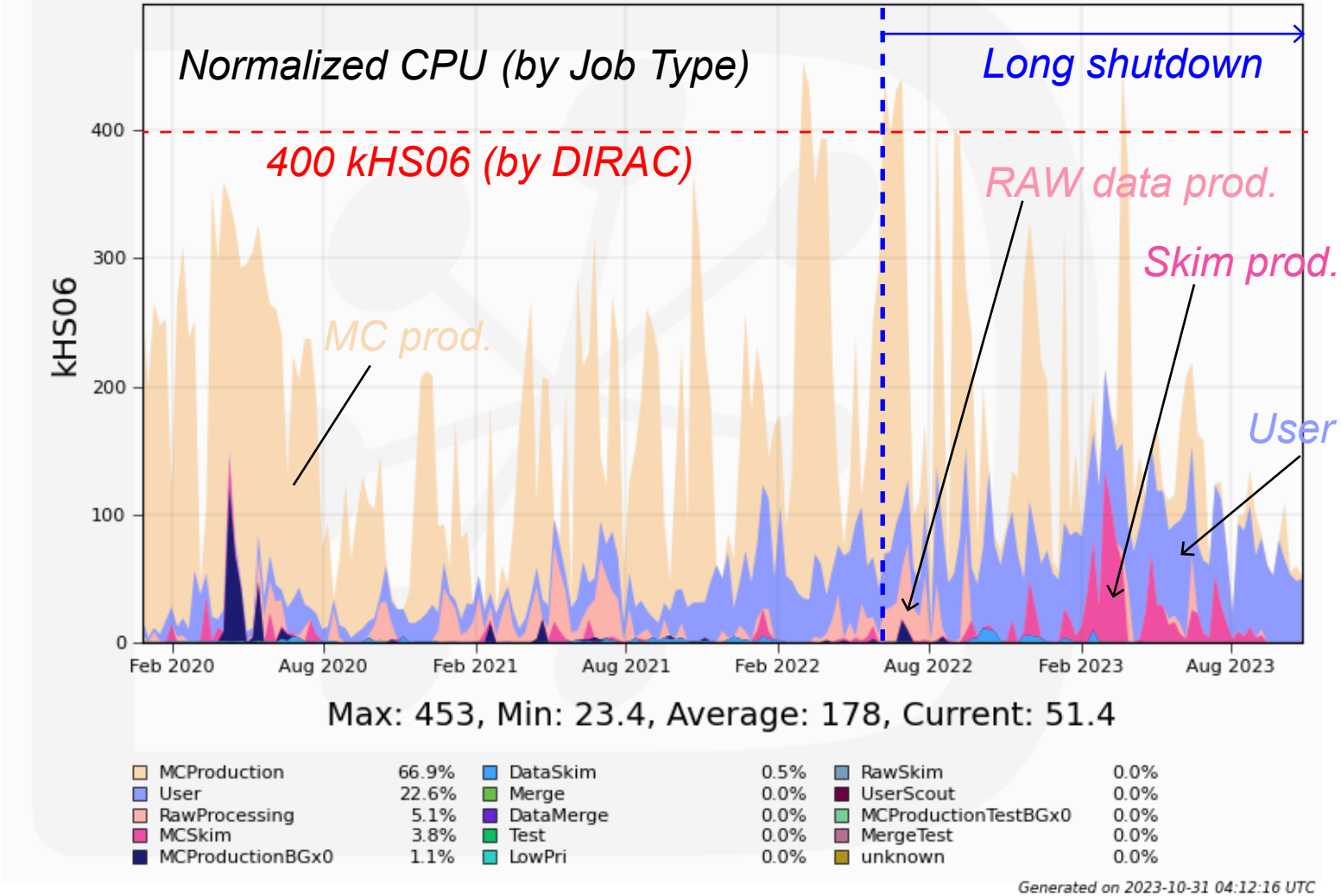
*Additional storage under implementation in some of the sites

TYPE	Resource provided
CPU	36,7 kHS06/HS23
DISK	550 TB

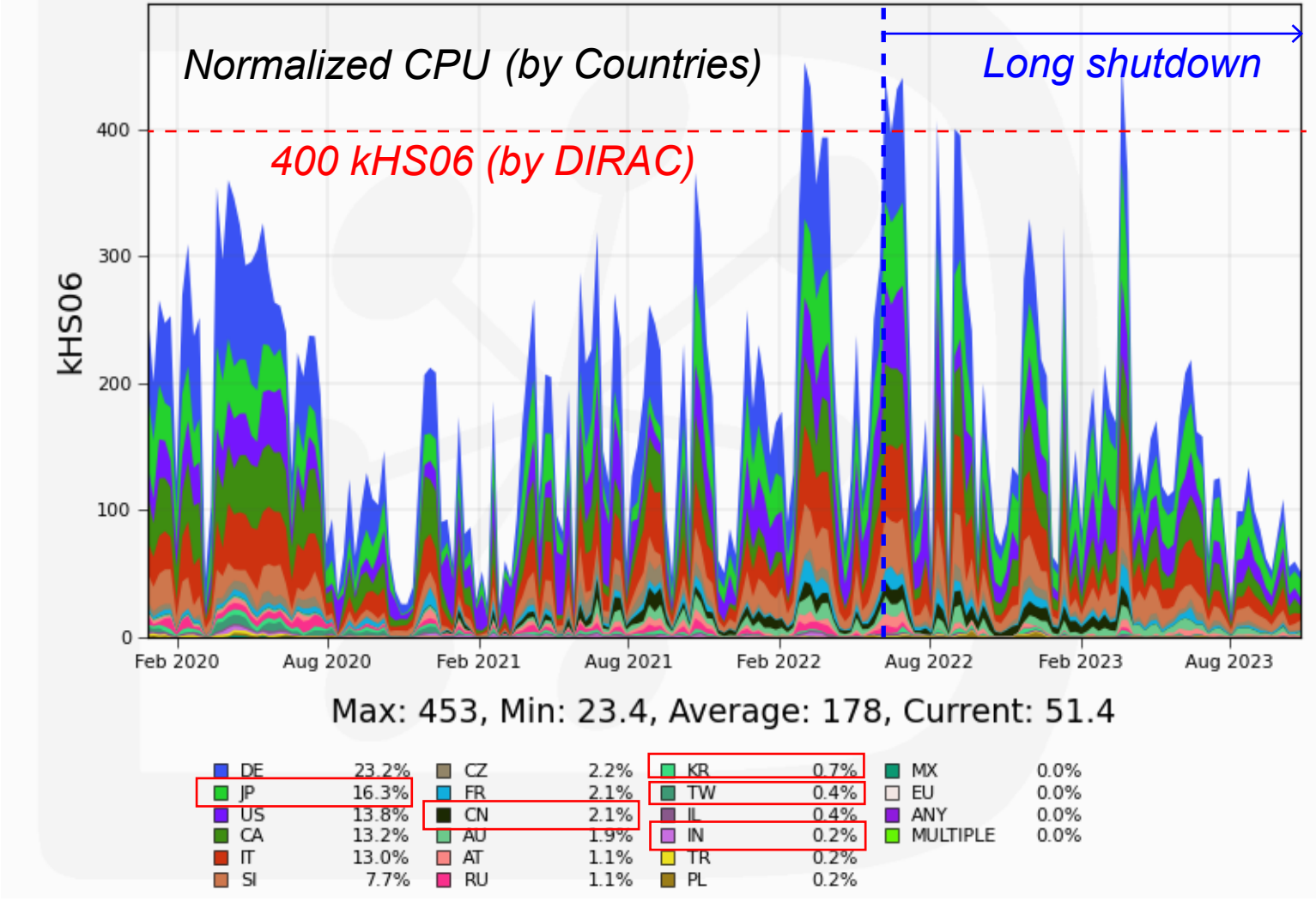


Resource for calibration

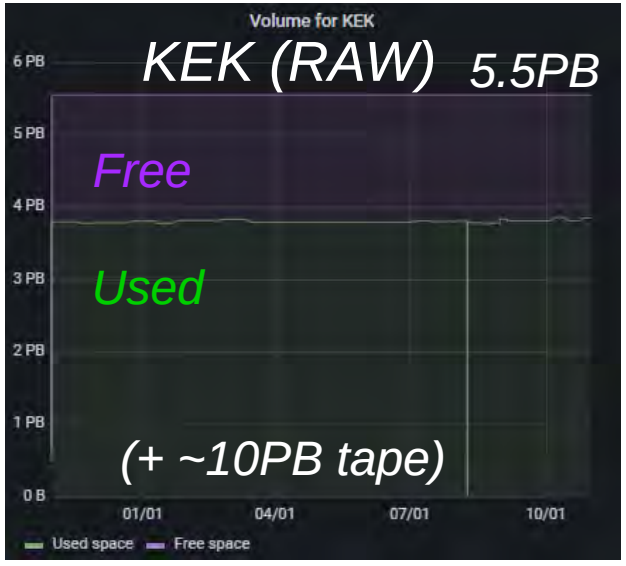
Compute Element

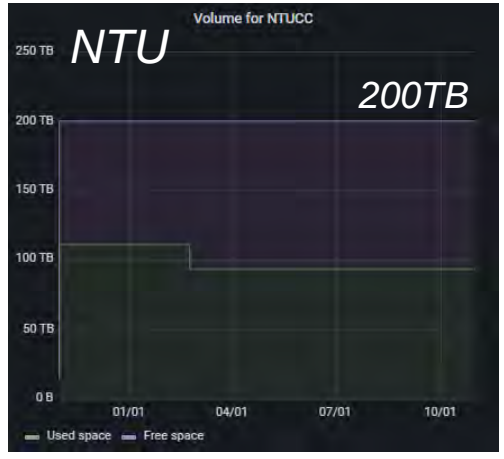


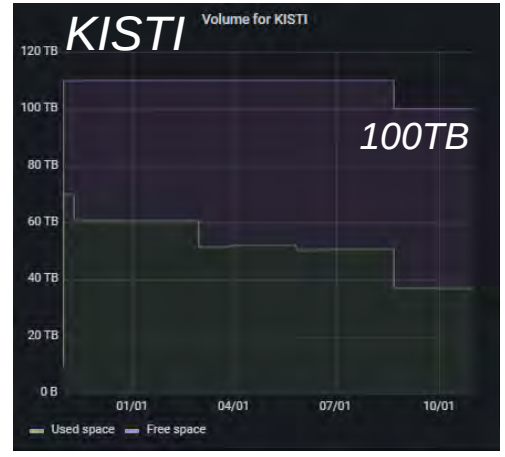
Compute Element

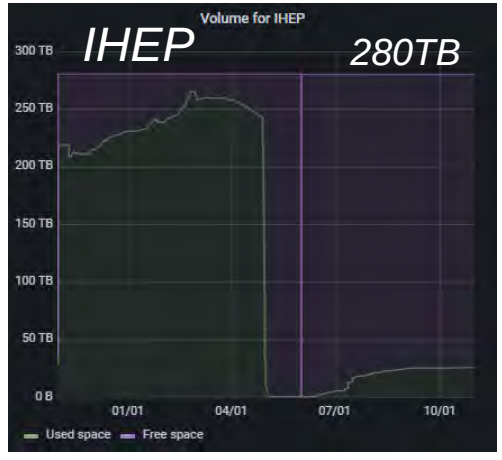


Storage Elements in Asia









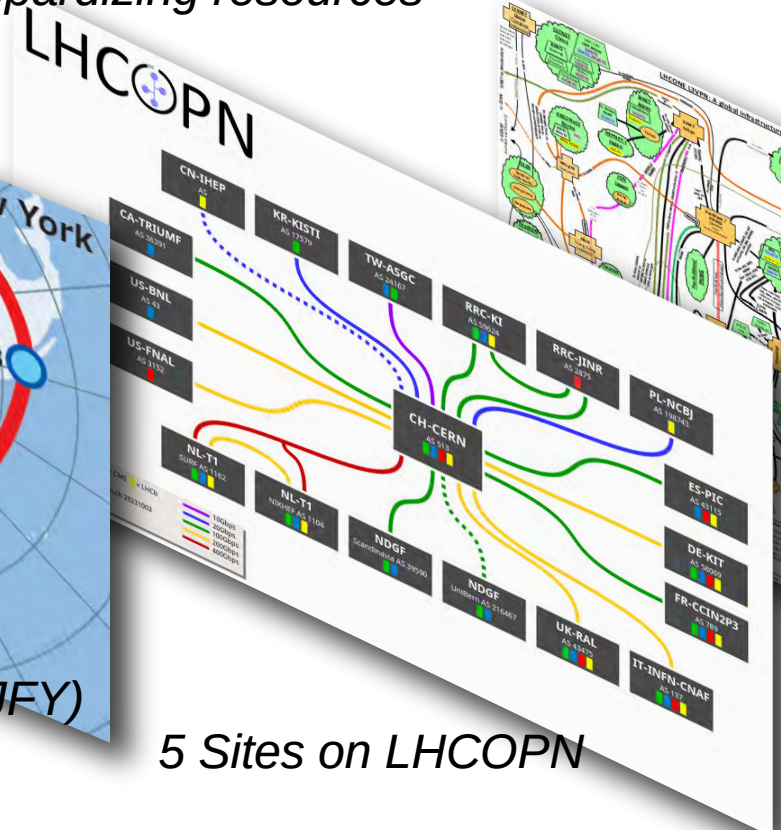
c.f. DISK (pledged in Belle II total) : 18.9PB

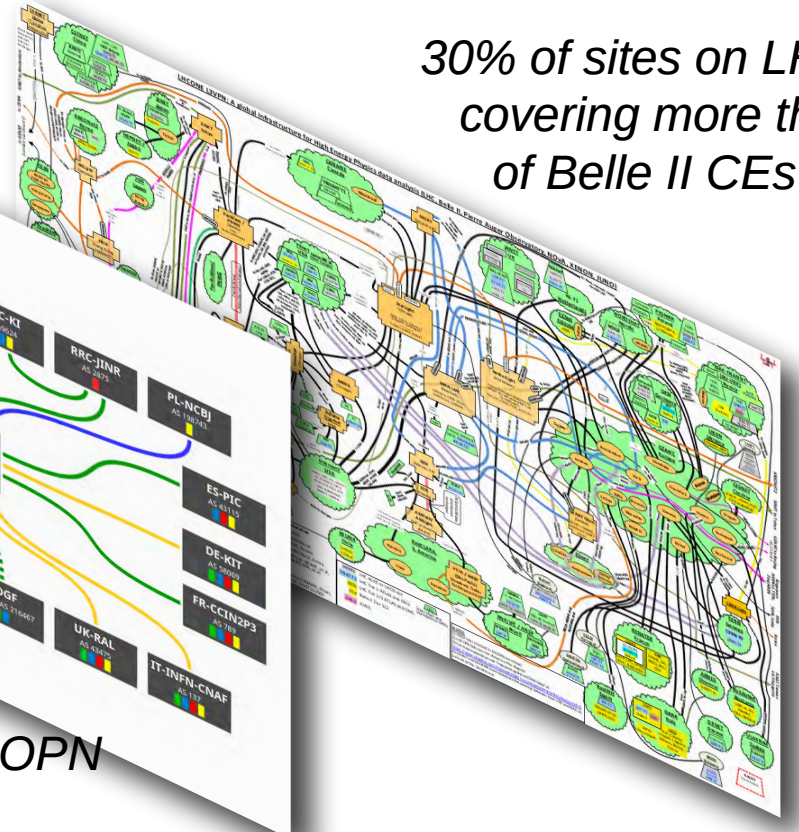
Network environment for Belle II

100G Global Ring runned by SINET6

LHCOPN optical infrastructure that can be used without jeopardizing resources

LHCONE L3 VPN Connecting all the major Data Centers





30% of sites on LHCONE covering more that 80% of Belle II CEs and SEs

Japan to
NY,LA : 100Gbps x 2
Amsterdam : 100Gbps x 2 (in 2024 JFY)

5 Sites on LHCOPN

Chages since the last ATCF6

@ ATCF6

Rucio integration (Jan 2021) → smooth operation

+ automatic deletion of unnecessary files

+ nice monitors

+ quota setting per user

RAW data centers (BNL, KIT, CNAF, DESY, CC-IN2P3, UVic)

Since then...

Belle II is still in the long shutdown...

No big changes and updates in the Belle II computing ...

Adaptation to python3

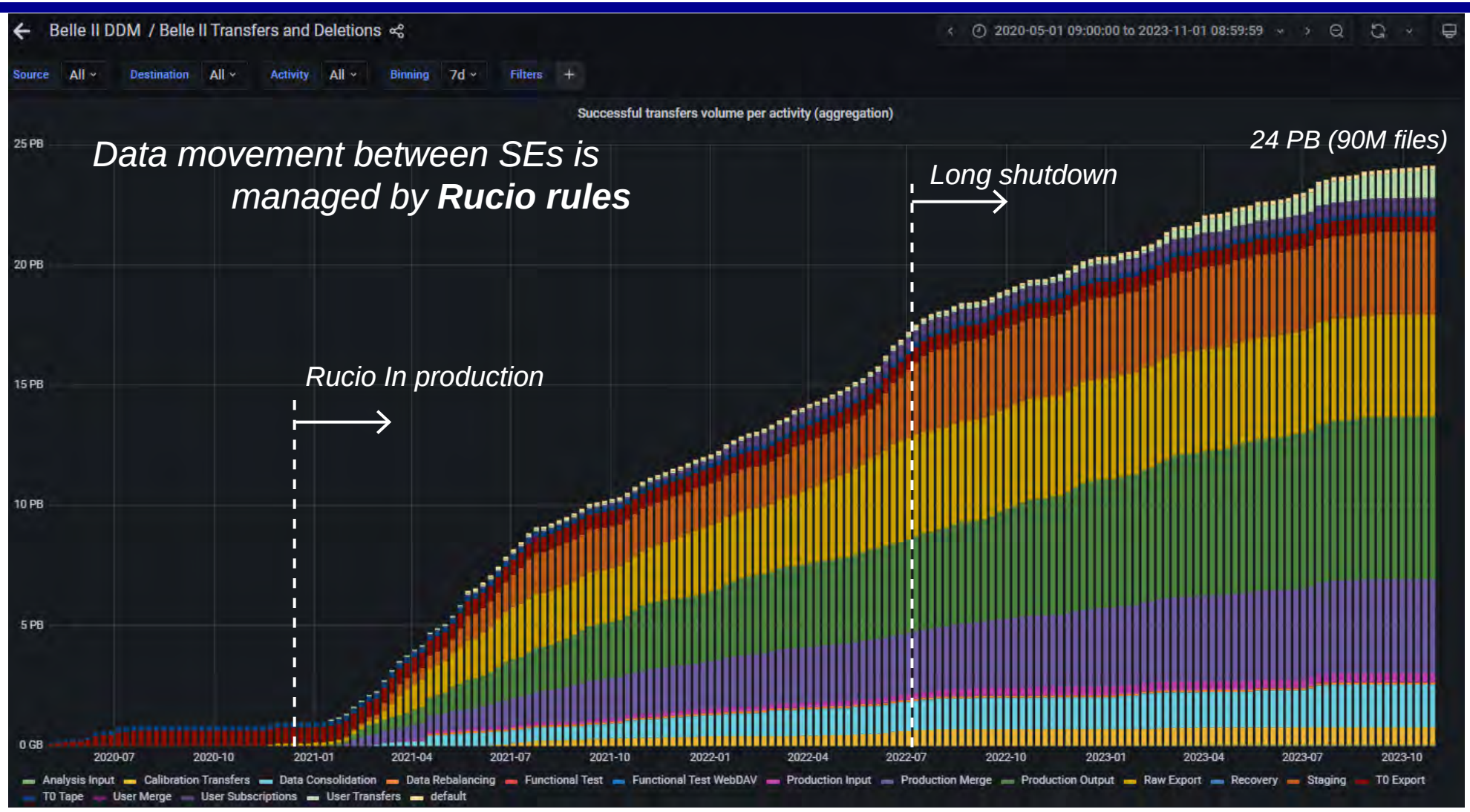
Most of the servers are running with python3

Clients (gbasf2) installation on cvmfs currently wih python3 + python2

VOMS and X.509 proxy → IAM and JWT-base authentication

Preparation work on-going....

Data Transfer Performance

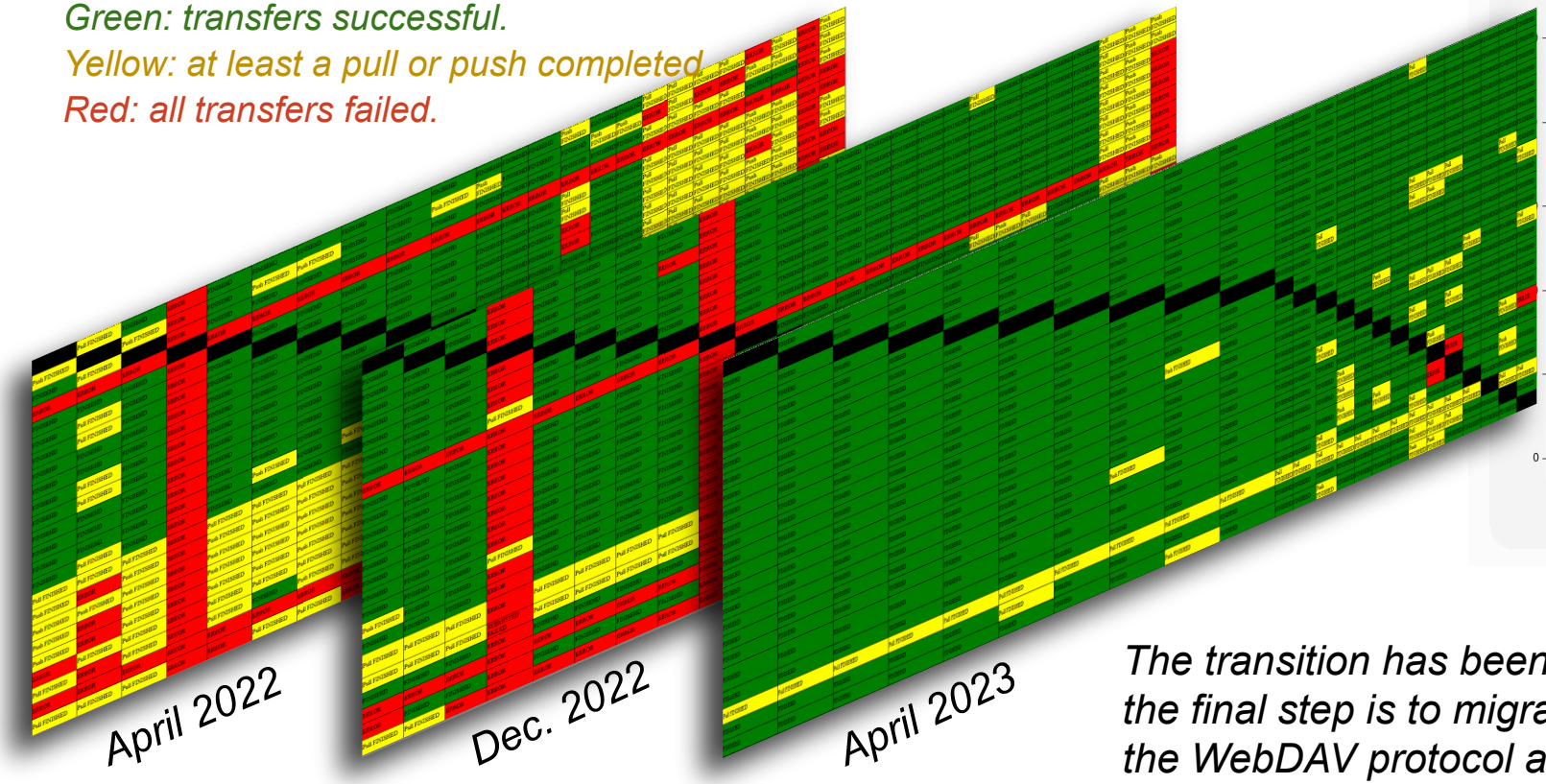


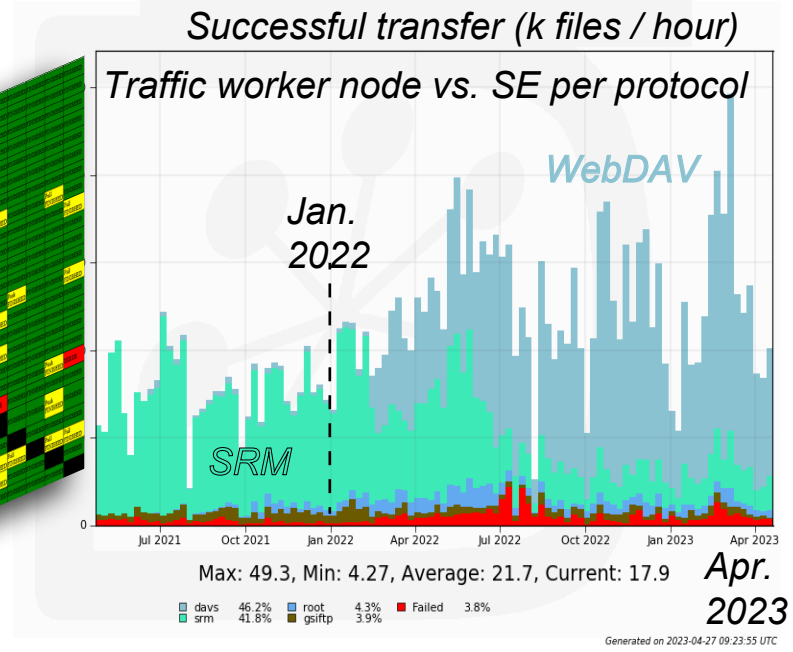
Data access and transfer with WebDAV

Belle II experiment adopted WebDAV as the main protocol for data access and third-party protocol
The migration process (started in 2022) required a large effort to ensure a smooth transition while keeping the infrastructure operational.

Third-party-copy monitor :

Green: transfers successful.
Yellow: at least a pull or push completed.
Red: all transfers failed.

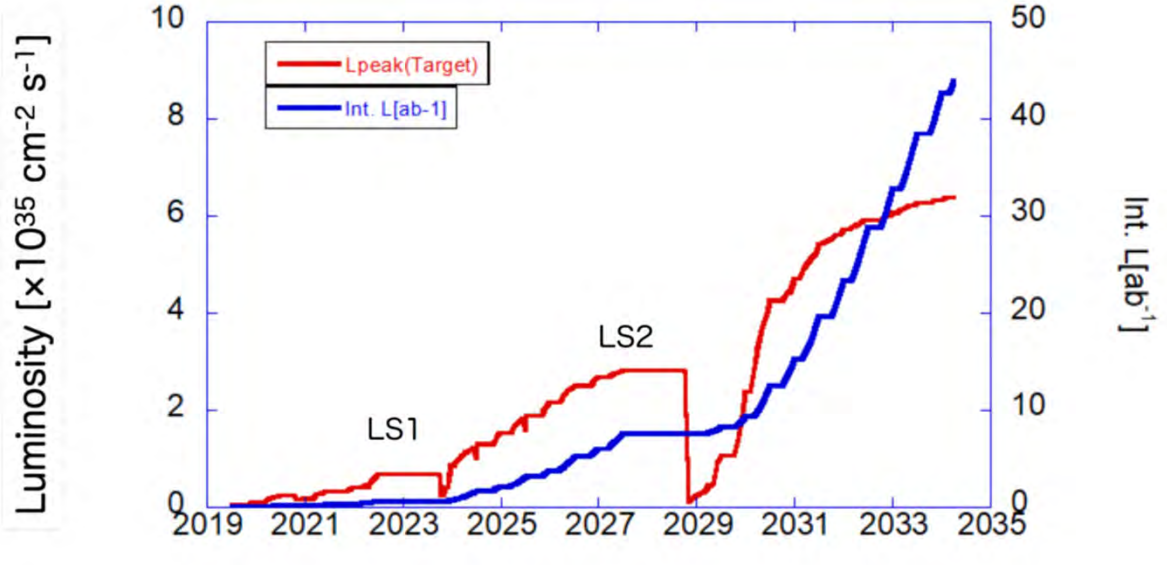




The transition has been completed for disk storages, the final step is to migrate the TAPE system to the WebDAV protocol and adopt REST API for tape staging.

Experiment plan

Experiment plan



Ultimate goal: reach 50 ab^{-1} by operating at the design luminosity of 6 x 10^{35} cm^{-2} s^{-1}

What Belle II has achieved in Run1

-reached world record instantaneous luminosity: 4.7 x 10^{34} cm^{-2} s^{-1},

collected up to 15 fb^{-1} per week

-recorded luminosity at Belle II: 424 fb^{-1}

(Belle 988 fb^{-1}, BaBar 513 fb^{-1})

completed

-LS1 in 2022-23 for the full pixel vertex detector (PXD) installation & partial replacement of MCP-PMTs in TOP

-Run2 (2024-2028) : target luminosity 2 x 10^{35}cm^{-2}s^{-1} integrated luminosiry 5-10 ab^{-1}

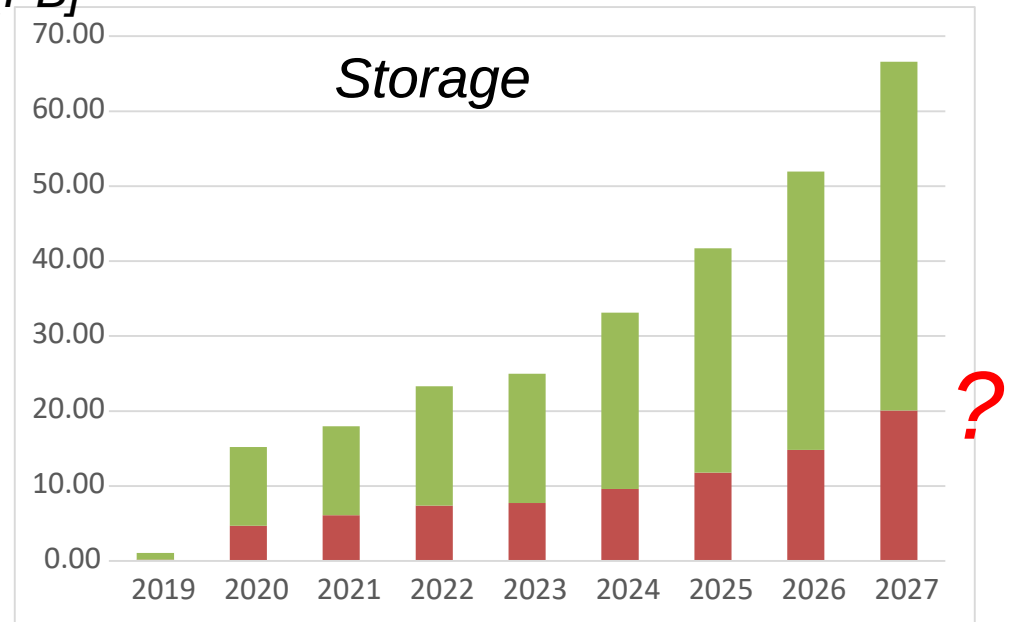
-options for an interaction region upgrade (LS2) ≳ 2028 under study → https://arxiv.org/abs/2203.11349

Beyond: discussions of physics and detector options with an upgraded accelerator to reach an even larger data sample of ~250/ab

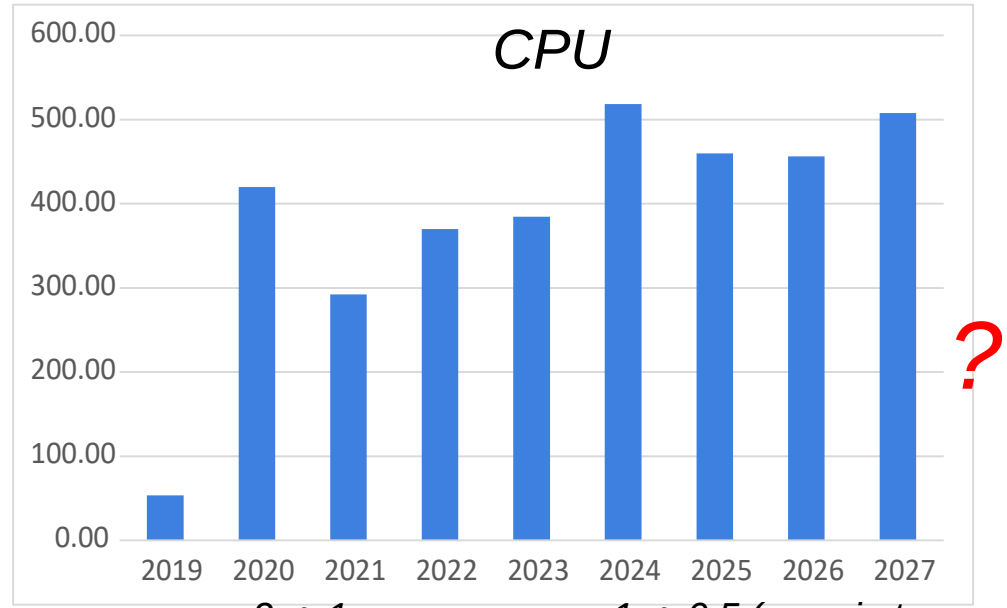
Computing Resource Estimation

Belle II Computing Steering Group esitmates the resource requirement for the next and succeeding three years (2024, 2025-2027)

[PB]



[kHS06]



2 → 1 *1 → 0.5 (once in two years)*

Number of MC production campaign with newer software will be reduced

{+ software will be more stable
+ data will be too much to simulate in short period

but we may not be able to keep using the same MC samples for more than two years....

Resource estimation heavily relies on LS2 schedule...

Distibuted Computing

external factors

Adaptation to python3

major effots to adapt the code to python3-compatible,
and to prepare the installaton of gbasf2 and DIRAC servers with python3

VOMS and X.509 proxy → IAM and JWT-base authentication

IAM : Identity and Access Management
JWT : JSON Web Token

Preparation work on-going, another upgrade of base DIRAC to 8.0 requires certain efforts

Belle II internal development items

Production system

Tuning and fixing after enabling production output without run boundary

New features in certification process

***Staging daemon** : Will be able to automatically stage data from tape based on request*
from the production system. Major improvement for reprocessing

***Metadata in Rucio** : Having metadata in Rucio will provide better overview of our data*
and should be able to improve user analysis and data production
→ more robust metadata catalog system in future

***Data popularity** : Client/pilot will report file access to Rucio allowing*
to measure number of accesses and time of data blocks
→ can introduce popularity-based data replication

Still in development

Should reduce some latency observed when Production System contacts Rucio
New agent to manage campaign lifecycle (campaign archiving g reduction g deletion)
→ will remove load on people managing the disk space

Summary

- *Belle II experiment*
  - *Now under LS1 (long shutdown 1) to replace PXD/TOP.*
  - *Ready for the new data taking (will be resumed from Jan 2024)*
- *Distributed computing*
  - *No big changed since the ATCF6*
    - *+ smooth DIRAC / Rucio operation*
- *Experiment plan*
  - *LS2 is under study and “the beyond” is under discussion*
    - *→ depending on these, compute/storage estimation will be changed ...*