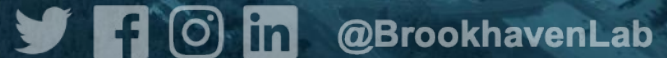




SDCC Stakeholders feedback - Belle II

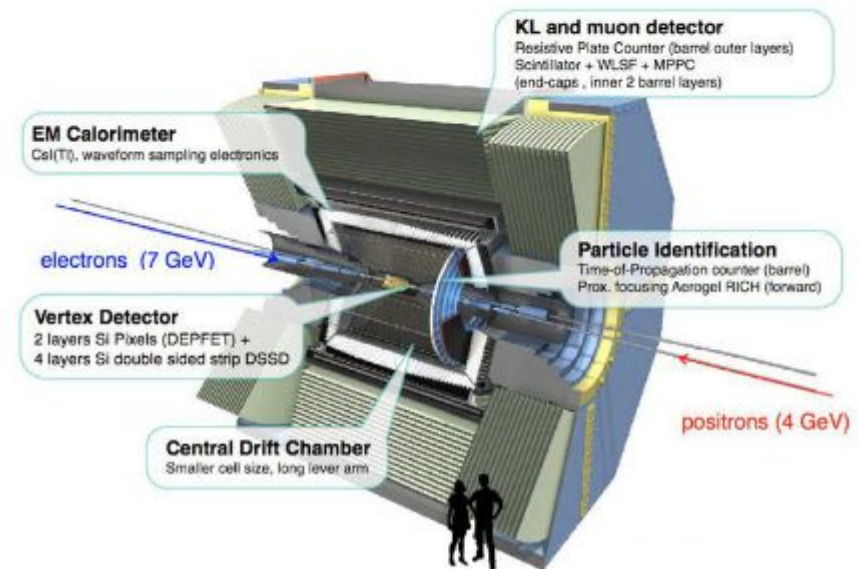
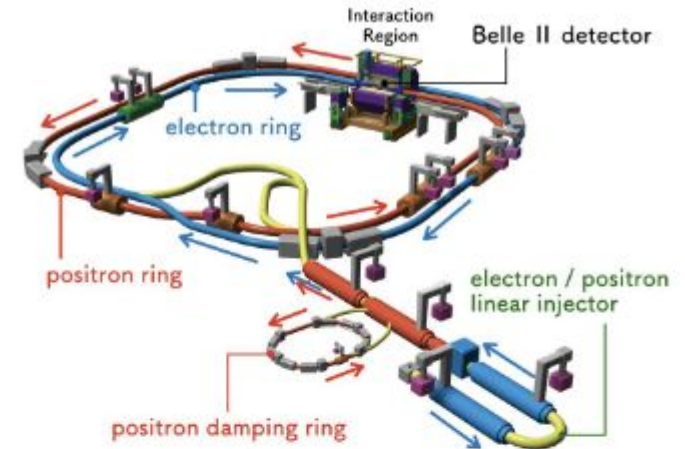
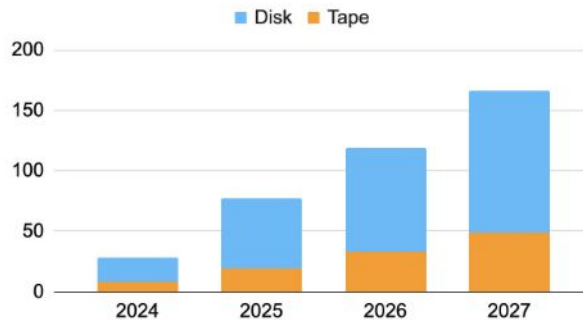
Cédric Serfon

September 6, 2024



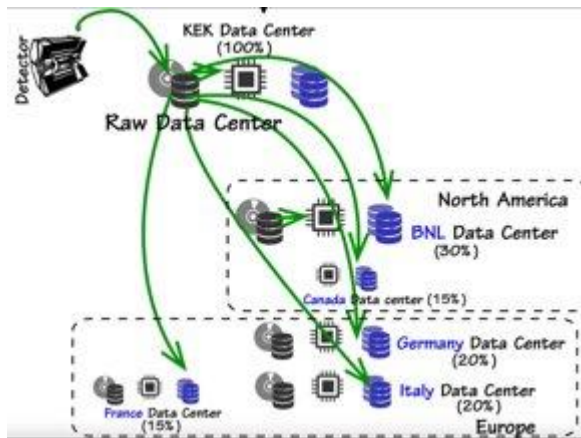
Introduction

- Asymmetric lepton collider.
- Upgrade from previous Belle experiment.
 - 50 ab^{-1} at the end of the experiment (x50 than the previous B factories)
 - Estimated size of the dataset collected by the experiment is $O(10) \text{ PB/year}$.



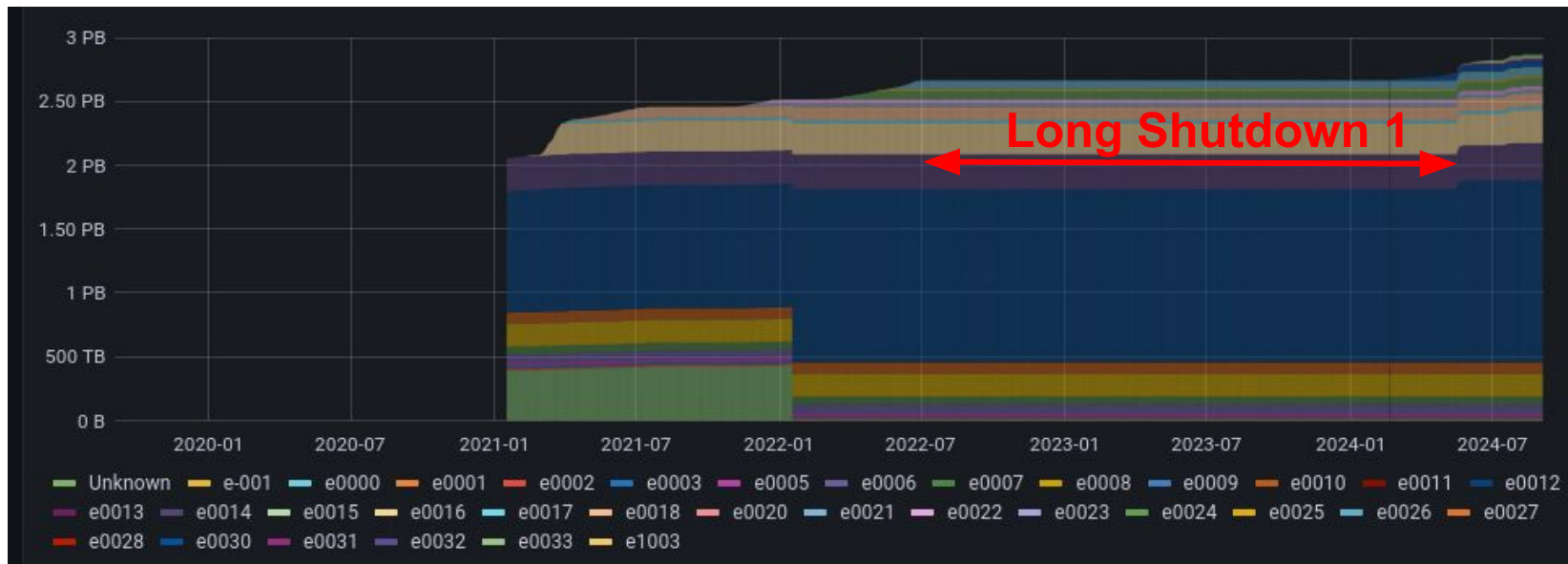
Belle II computing

- Belle II uses a grid model similar to ATLAS :
 - KEK is the Tier0 that store the primary copy of RAW data
 - A few big centers are Tier1s (aka RAW Data Centers) :
 - Store the second copy of RAW Data
 - Provide the resources for the prompt processing and reprocessing as well as MC and user analysis
 - Tier2s provide resources for MC production and analysis



TAPE

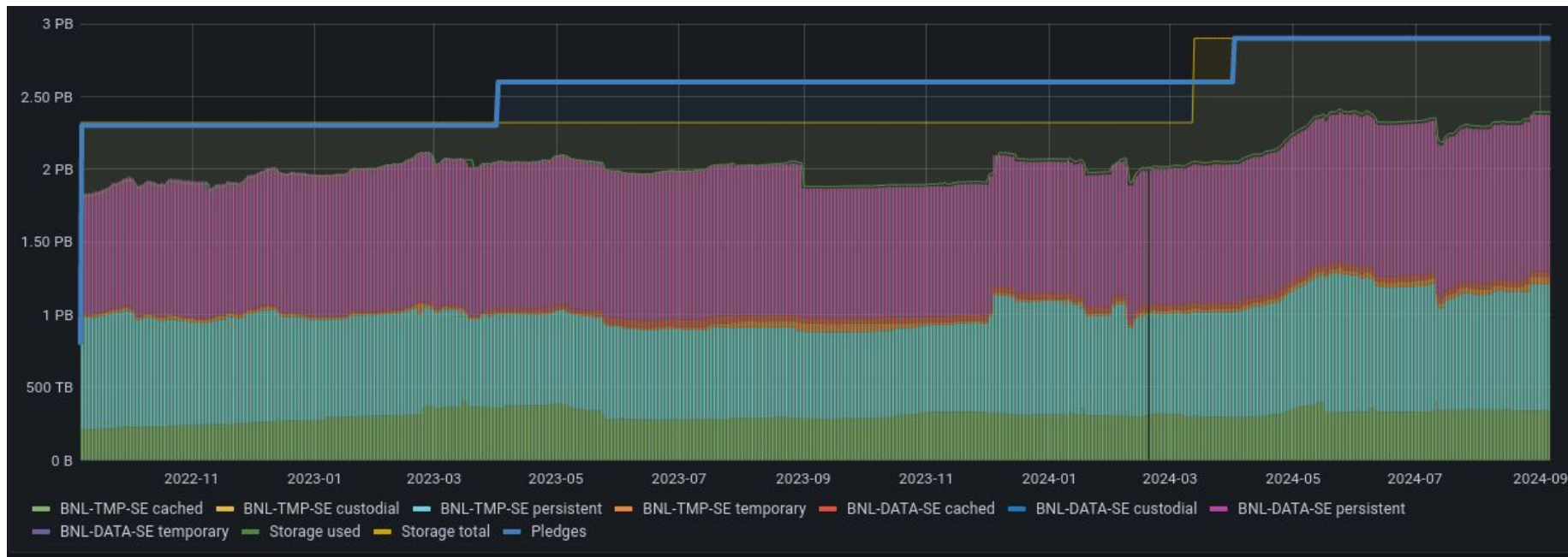
- BNL is one of the RAW Data Centers that hosts 30% of the secondary copy of RAW Data



Data stored on TAPE at BNL

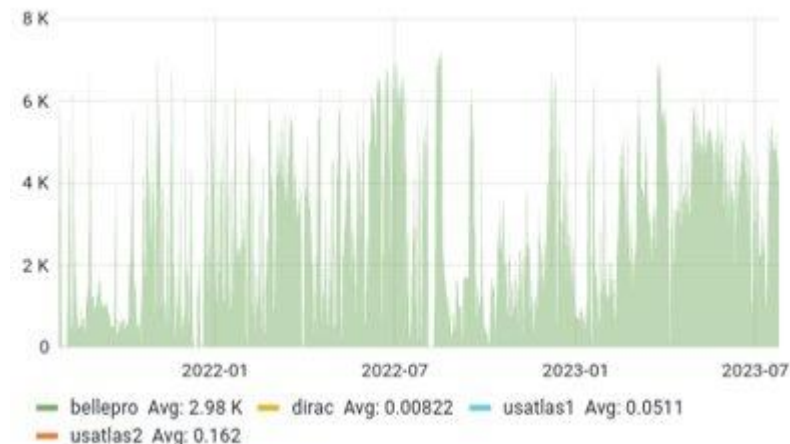
DISK

- BNL also provides disk to store data, monte Carlo and user data



Computing resources

- As a RAW Data Center, BNL provides the resources for the prompt processing
- Currently about 7200 CPU slots, i.e. ~81kHS06
- Use Belle II CPU in SDCC shared pool :
 - Full pledge available when Belle II needs it, but can be used for other DoE science when not used by Belle II
 - On the other hand, Belle II spikes can be accommodated by using extra CPU from the pool when needed (e.g. raw data processing)



Pledges

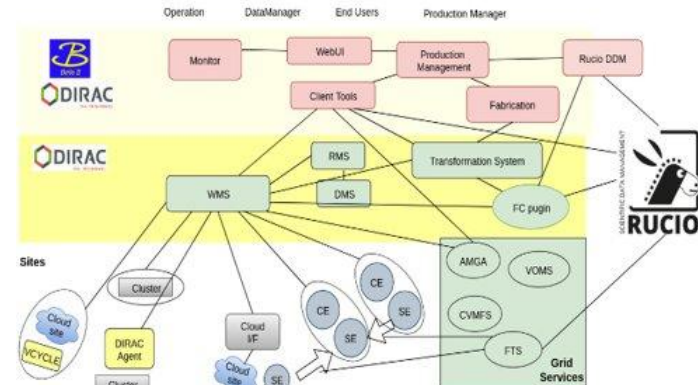
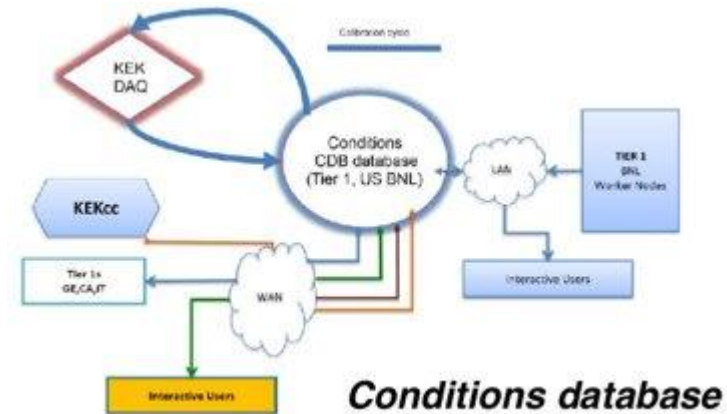
Resources in USA								Current (new) pledge		
Year	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027
Total tape (PB)	0,07	0,09	2,36	2,55	2,72	2,76	3,05	3,43	3,90	4,72
Total Disk (PB)	0,86	1,06	1,73	1,83	1,98	2,14	2,66	2,48	3,43	4,20
Total CPU crosscheck (kHEPSpec)	34,71	23,86	31,10	30,08	30,04	40,96	64,94	33,44	75,24	84,82
from 240501_ResourceEstimate_USA20240524 (previous estimate provided to SDCC)								proposed (old) pledge 20240501		
Total tape (PB)	0,07	0,09	2,36	2,55	2,72	2,76	3,05	3,43	3,90	4,72
Total Disk (PB)	0,86	1,06	1,73	1,83	1,98	2,14	2,55	2,28	2,93	3,51
Total CPU crosscheck (kHEPSpec)	34,71	23,86	31,10	30,08	30,04	40,96	62,58	31,34	64,80	71,55

- Decrease in CPU pledges explained by different factors (overestimated pledges in the past, change of frequency for reprocessing after 2024, etc.)

Belle II services operated by SDCC

- BNL (SDCC) is responsible for operating 2 critical services :

- The Condition Database :
 - Host the condition data used by every jobs
- The Data Management service :
 - Based on Rucio
 - Responsible for the management of data over their lifecycle
- After a decision in Feb 2024, Rucio will also become the metadata service of Belle II in the coming months



Other services

- FTS server (shared with ATLAS) for file transfers
- ELK stack+Grafana used for monitoring (e.g. : [Transfers Dashboard](#) or [Accounting Dashobard](#))
- DIRAC (Workload and Workflow Management Service) certification nodes and CI/CD service
- Analysis facility (JupyterHub) being evaluated

Conclusion

- SDCC has been delivering many critical services for the whole Belle II collaboration over the last years
- Increase of computing needs in the next years stay modest wrt ATLAS/HL-LHC
- Synergy with other SDCC stakeholders (ATLAS, RHIC, etc.).
- Very good interactions between SDCC and non-SDCC members
- Belle II has greatly benefitted from the relatively small, but very effective, U.S. C&S team of SDCC and NPPS members