2nd DPHEP Collaboration Workshop

Belle + Belle II

and other experiments at KEK

March 14, 2017 @ CERN

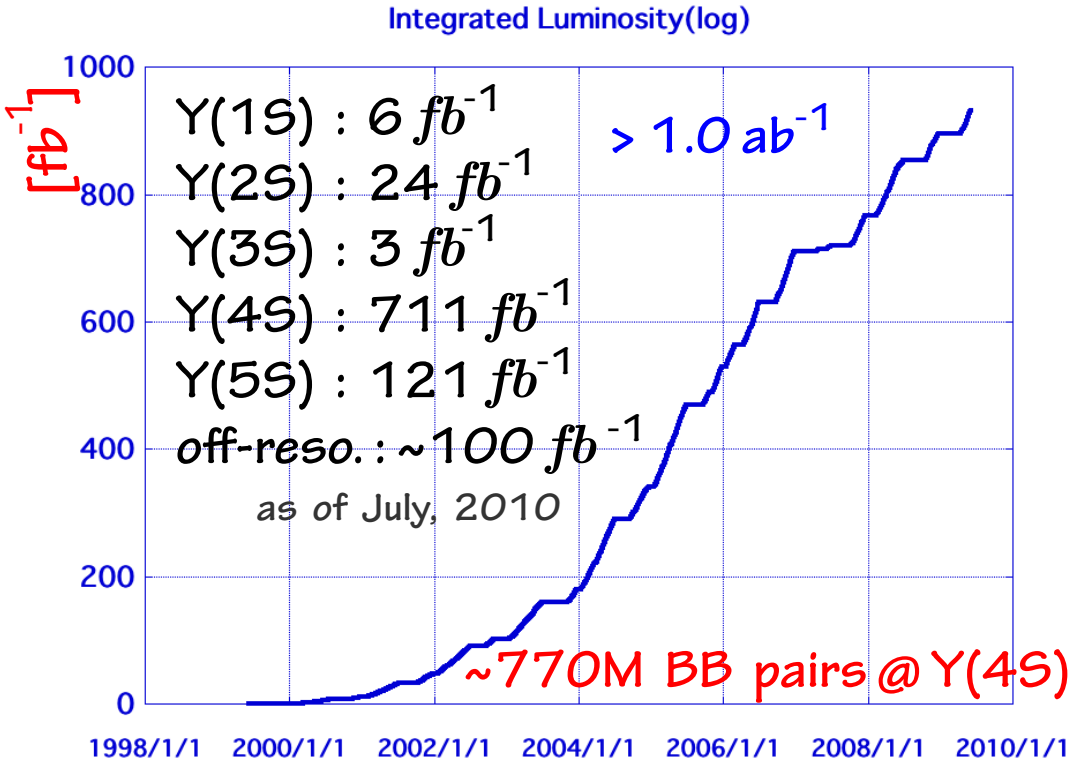
takanori.hara@kek.jp

Belle I Data

Belle : started in 1999, data-taking completed in 2010

still keep analysing the Belle data (>~ 20 papers / year)

in parallel with the construction of Belle II detector and computing



Size of storage for Data

RAW	raw data	~1000 TB
DST	prescaled data (1-1/400)	320 TB
mDST	reconstructed info.	150 TB

Format : panther (Belle's own bank system)

Size of storage for MC

mDST	recon. info. + MC true	800 TB

10 streams for bb + 6 streams for udsc

we have two versions of mDST

one : w/ old-tracking (=conformal finder)

the other : w/ new-tracking (=+Hough finder)

mDST stored on both Disk and Tape

raw data + DST stored on Tape

Belle I's strategy

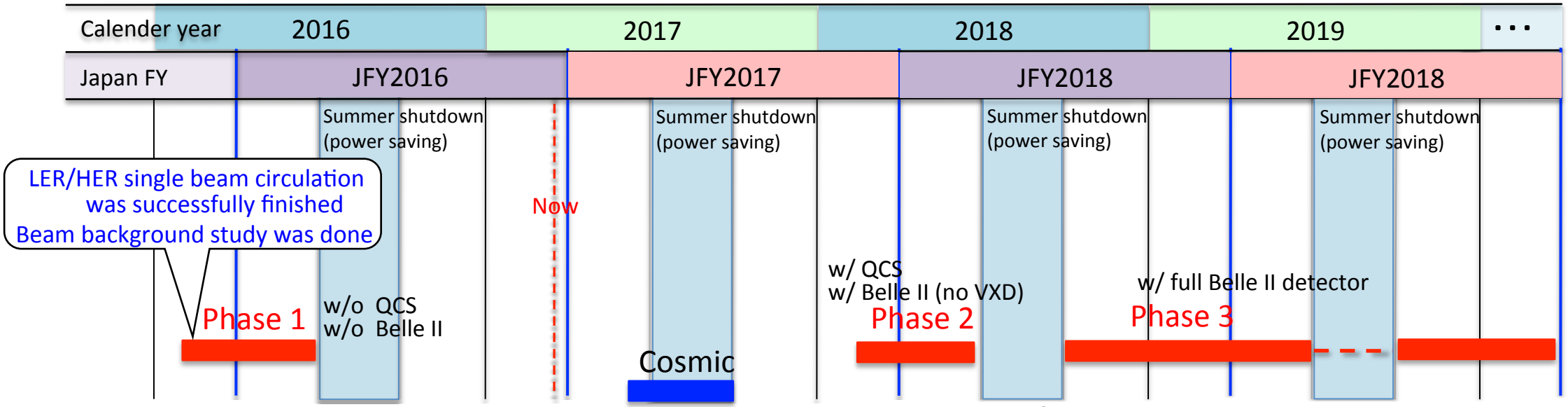
Preserve all RAW and mDST, at least.

Keep the current computing environment AMAP
(Library, Database, data format)

Belle I data will be used by the Belle community
until the time the statistics of Belle II > Belle I data set

	Belle		Belle II
Accelerator	KEKB		SuperKEKB
Beam Energy (GeV)	3.5 x 8 ($\gamma = 0.425$)		4 x 7 ($\gamma = 0.28$)
CM energy	....., Y(4S), ...		....., Y(4S), ...
Luminosity ($cm^{-2} s^{-1}$)	2.1×10^{34}	x 40 →	8×10^{35}
Total data (ab^{-1})	1	x 50 →	50
	raw data : ~1PB mDST data/MC : 0.15/0.8 PB (for one version)	Higher intensity	raw data : ~50PB (another raw data copy outside KEK)
Computing	one big center @ KEK (non-grid)	→	world-wide distriuted computing

Belle II Experiment : Time line



SINET4 → SINET5

KEKCC replacement

In 2017

Dress rehearsal is planned
(concurrent running of different type process)

→ Raw data processing

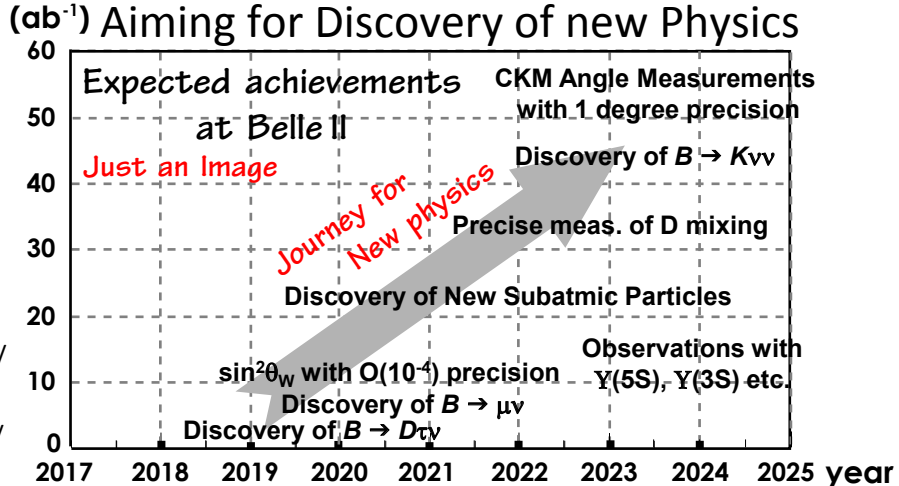
→ MC production

→ User analysis

https://indico.cern.ch/event/505613/contributions/2227937/

Fast Calibration is necessary

https://indico.cern.ch/event/505613/contributions/2227271/



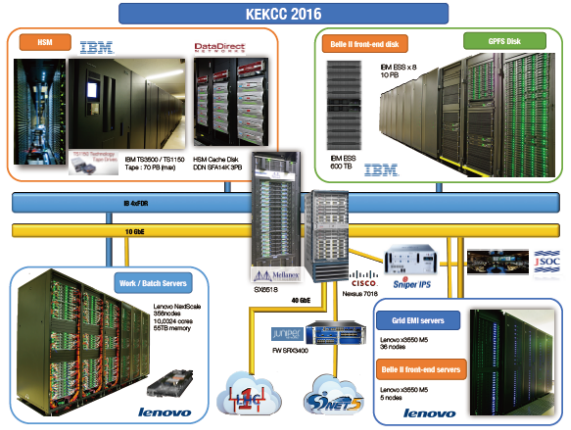
History of Belle Computer

Year (contract) / specification	1999- (4years)	2001- (5years)	2006- (3years)	2009- (3years)	2012/4- (~4.5years)	2016/9- (4years, planned)
CPU [SI2k]	~100 (WS)	~1200 (WS+PC)	~42500 (PC)	~115200 (PC)	~3500 cores ~40kHS06	~10,000 cores ~240kHS06
Disk [TB]	4	9	1,000	1,500	7,000	13,000
Tape [TB]	160	620	3,500	3,500	16,000 (at max)	70,000 (at max)
	(Belle dedicated)	(Belle dedicated)	(Belle dedicated)	(Belle dedicated)	(Belle + other KEK projects)	(Belle + other KEK projects)

KEK Computing system has to be replaced every ~4 years

Each time,.....

Data and software have to be migrated to the new system
Experiment dedicated core services system (e.g. DB)
has to be established in the new system from scratch



Belle I Data / Analysis Preservation

Data preservation

- old KEKCC → new KEKCC So far this works. But what happens in the next replacement?

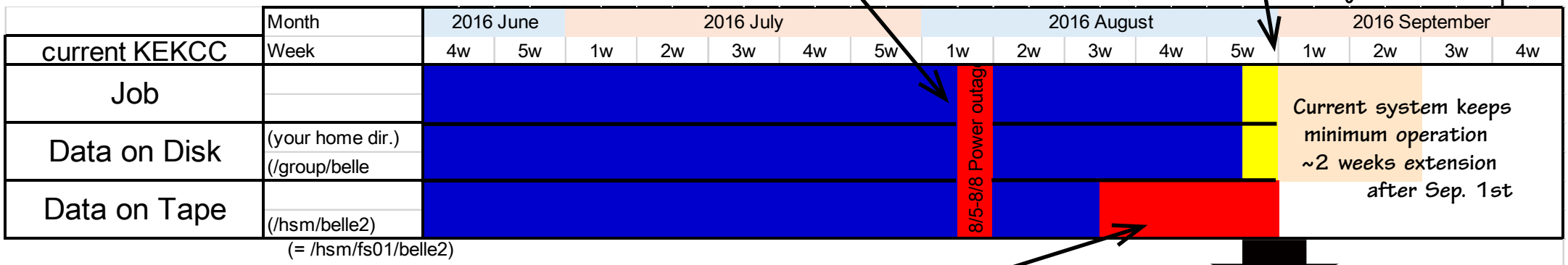
Analysis preservation

- Keep using Belle software, but OS SL5 (old KEKCC) → SL6 (new KEKCC)
  - + straighforward method to keep the analysis environment
  - + some software does not function under SL6
- Belle software (old KEKCC) → Belle II software (new KEKCC)
  - + new analysis tools can be applied even to Belle I data
  - + good exercise of Belle II analysis software workflow (before real data comes)
  - + younger researchers are not familiar with the Belle analysis software

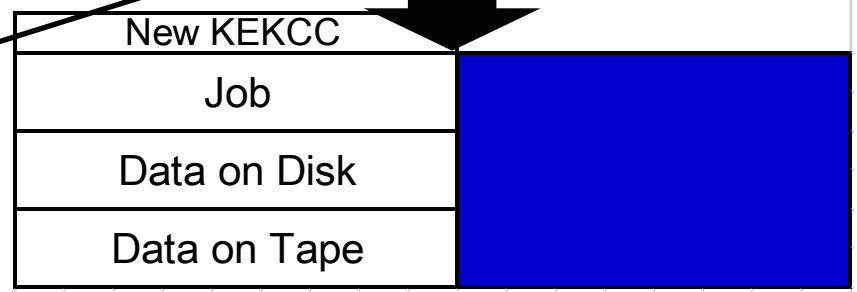
Migration to new KEKCC

8/5-8/8 : electric-power outage

User can use disk as usual until 8/29
Files updated after 8/29 have to be copied by users in September



Tape media will be moved from current KEKCC to new one during 8/15 - 8/18
(= Data on Tape)
Because the same company (IBM) won the procurement



Data on disk with rsync

migration of Belle II dedicated services

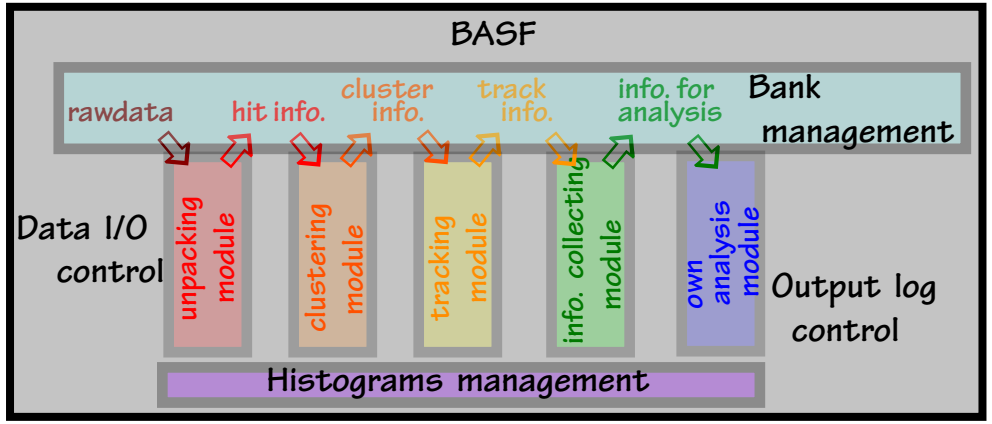
Analysis framework in Belle / Belle II

Make the Belle data readable within the Belle II software framework

Belle

"BASF" (= Belle AnalysiS Framework) is a software framework.

- . each task (e.g. tracking) is implemented as a module
- . data is exchanged through Belle Bank System
- . works under the Multi-CPU System



Panther

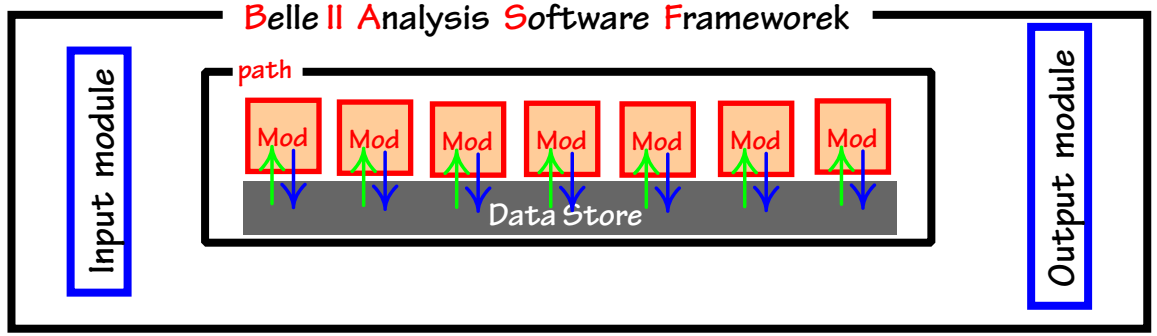
Belle own made bank system

- . Offline software : almost frozen
  software will not be updated
- . DST/MC production : not planned
  mDST is enough for analysis
- . Retention of knowledge :
  not well documented
  the number of people who knows Belle software is decreasing

Belle II

even data Input/Output is handled by module

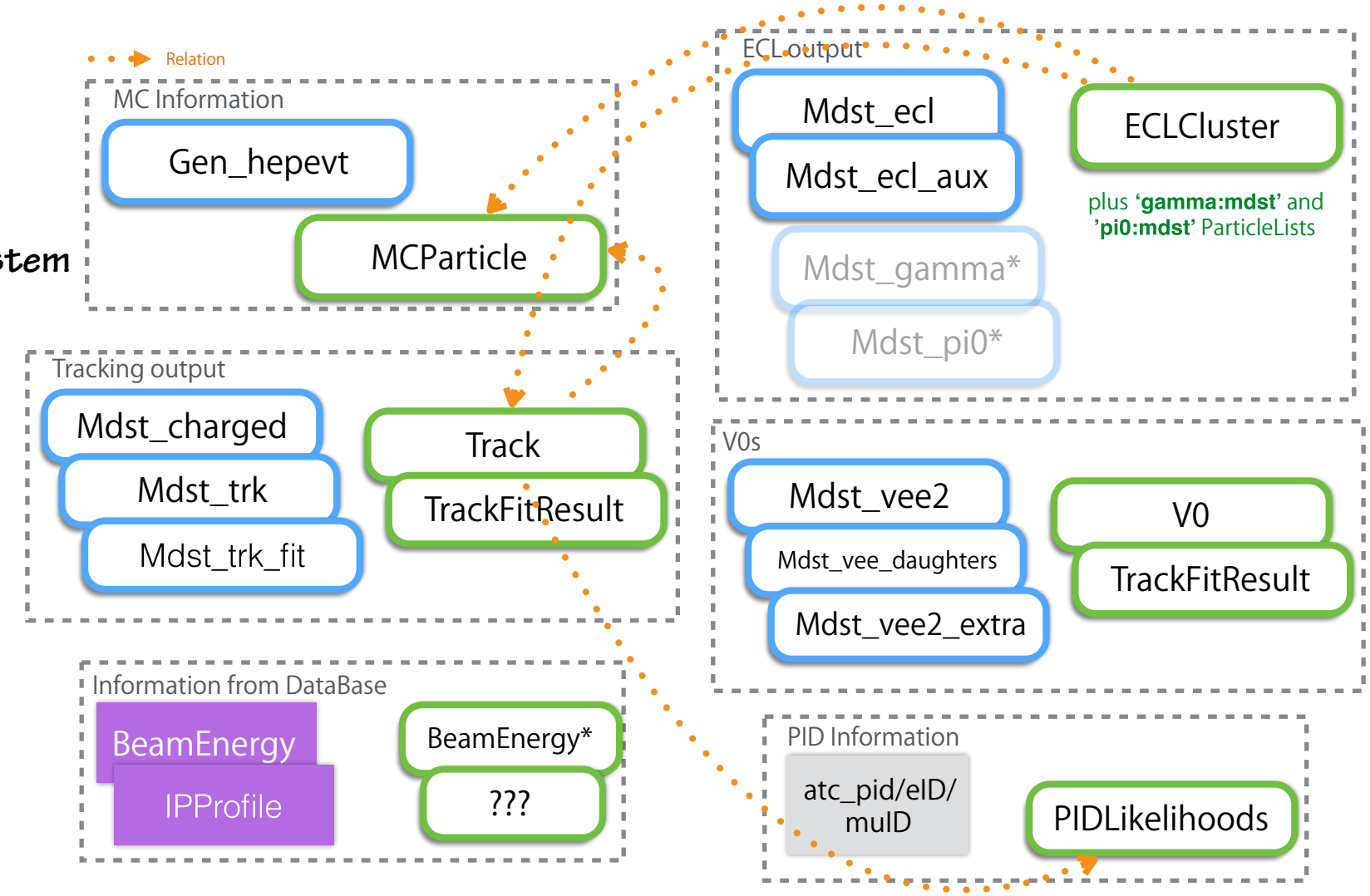
Root



Convert Belle I contents to Belle II format

Belle I

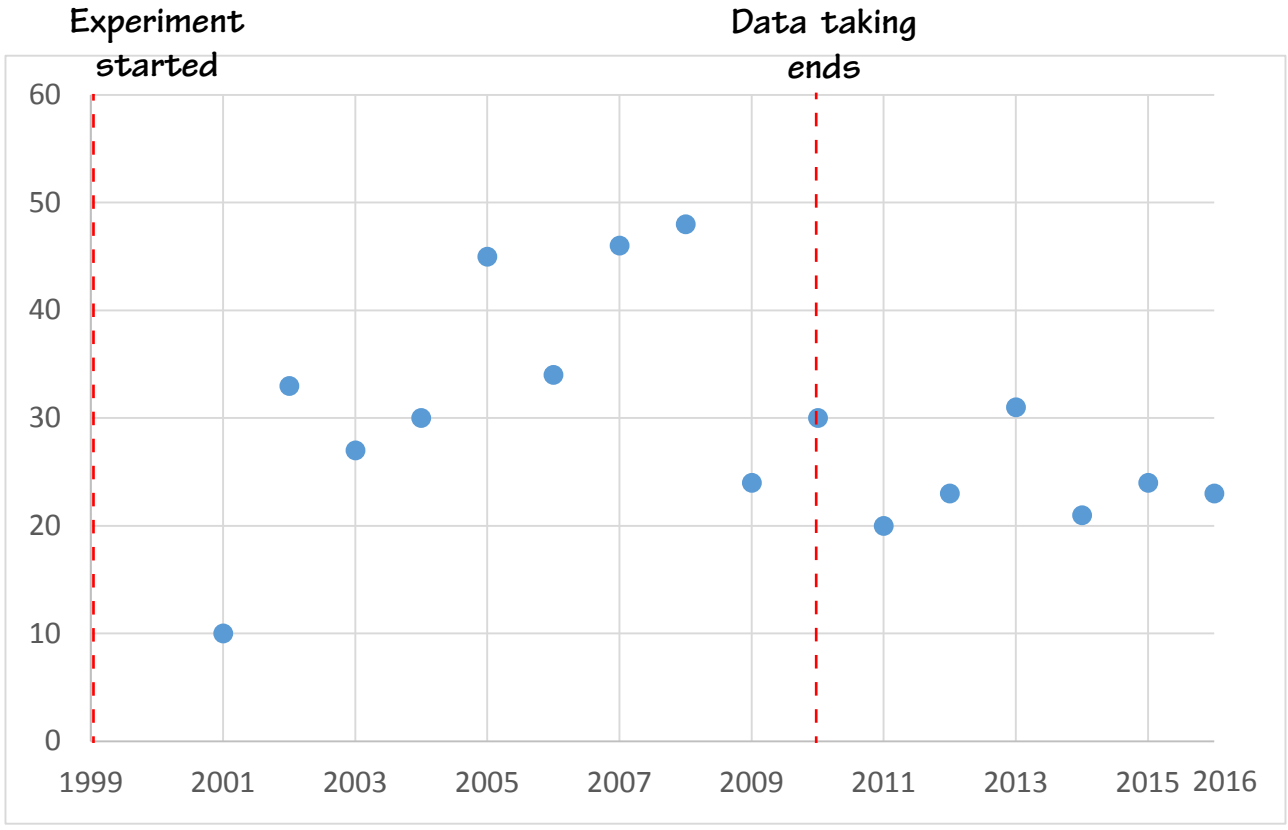
own made bank system "panther"



Belle II

ROOT

Number of publication



Even after the data taking was finished, analysis activities are vital.

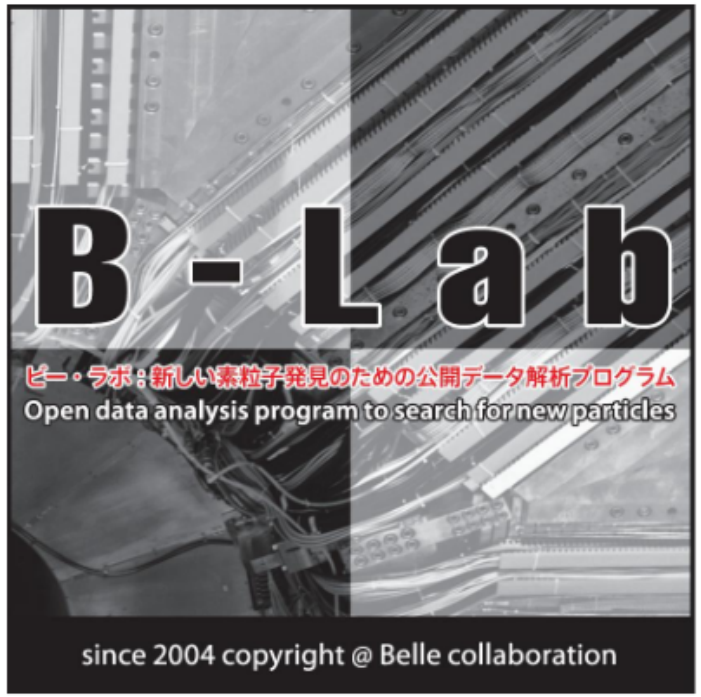
more than ~20 new results are constantly published

B-Lab (open data to public users)

Search for new particles with Belle I data !

http://belle.kek.jp/b-lab/

in Japanese...



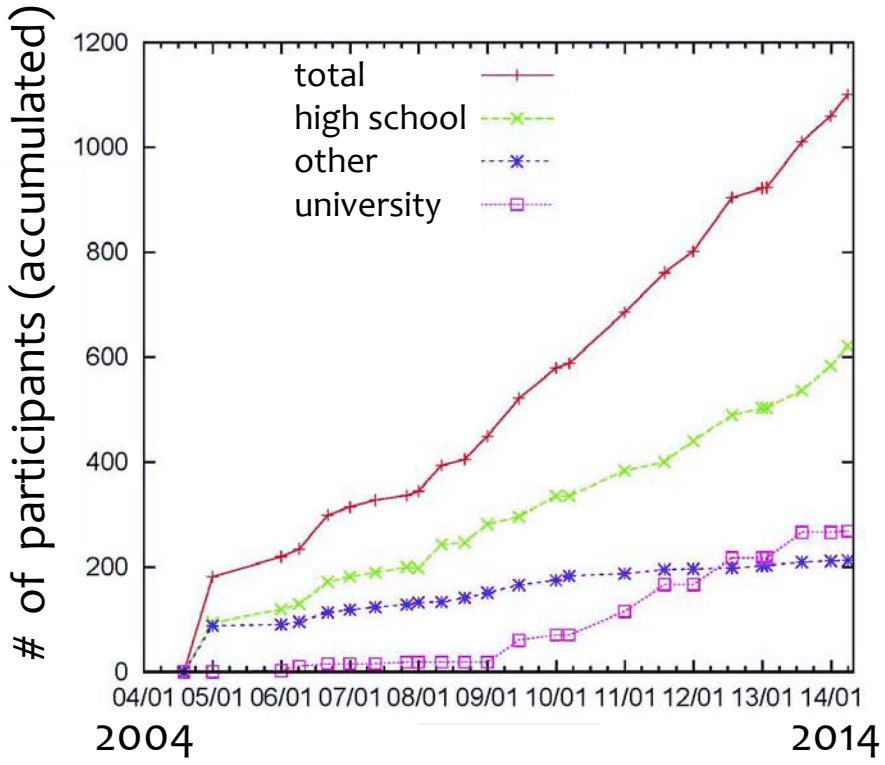
ホームページの最新更新日：2015/06/03

010514

▲Topへ戻る
■ルールの説明と参加の仕方
■解説書、データのダウンロード
■解析システムのインストールの仕方 (Windows PC)
■解析システムのインストールの仕方(Macintosh)
■新粒子の探し方
■例題とサンプルプログラム
■データ解析の模擬体験 (Windows PC)
■データ解析の模擬体験 (Macintosh)
■現在までの参加高校と参加者数
■新粒子探索結果報告
■実習授業と新聞報道
■良くある質問と答え
■B-lab プログラムの更新の記録
■B-Labの古いホームページ
■リンク集
■より高度の探索を行いたい場合

small portion of Belle I data is opened to public users (e.g. students at high school, university, etc.)
analysis tools and manuals are well prepared.

B-Lab (open data to public users)



so far, many particles are re-discovered.

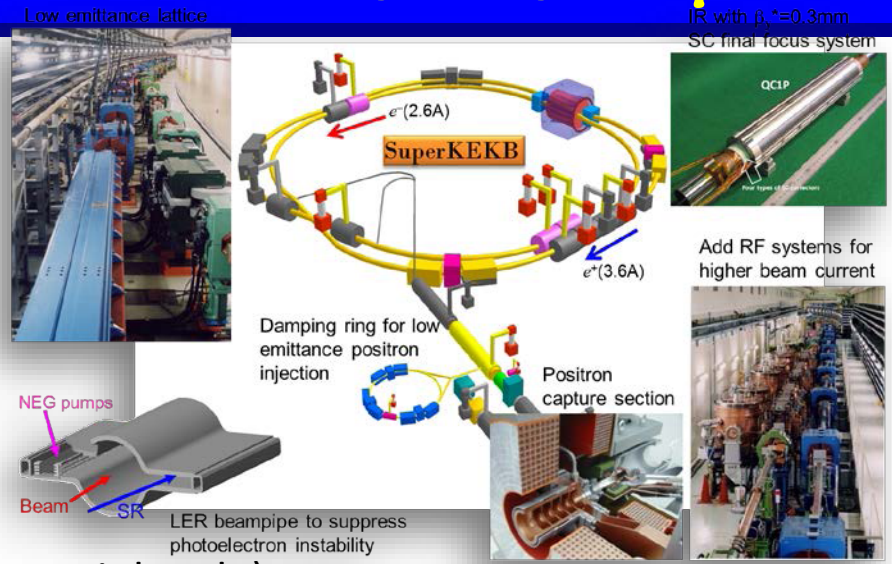
η(548MeV)
ω (782MeV)
K*(892MeV)
Λ (1115MeV)
Ds (1968MeV)
Λc (2286MeV)

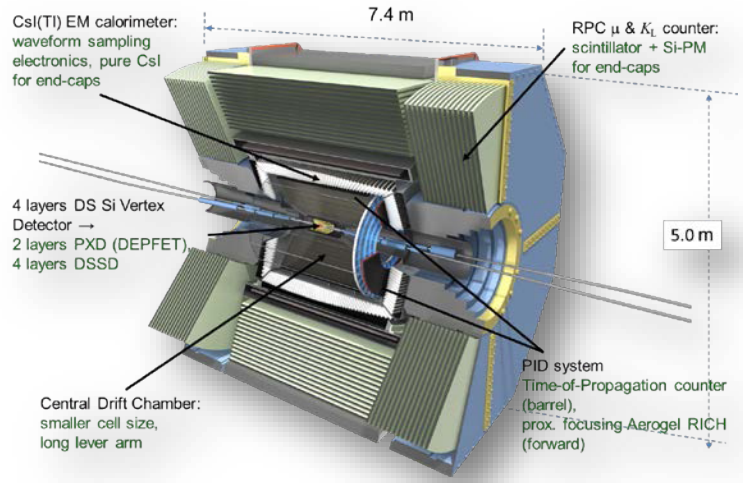
$\pi 0 \rightarrow e+e-\gamma$

$e+e- \rightarrow \rho 0 e+e- \rightarrow \pi+\pi-e+e-$

On going experiments at KEK

SuperKEK
B-Factory

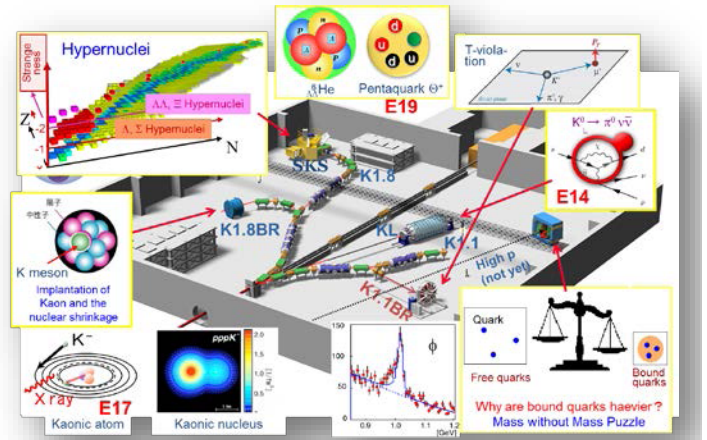




Belle II

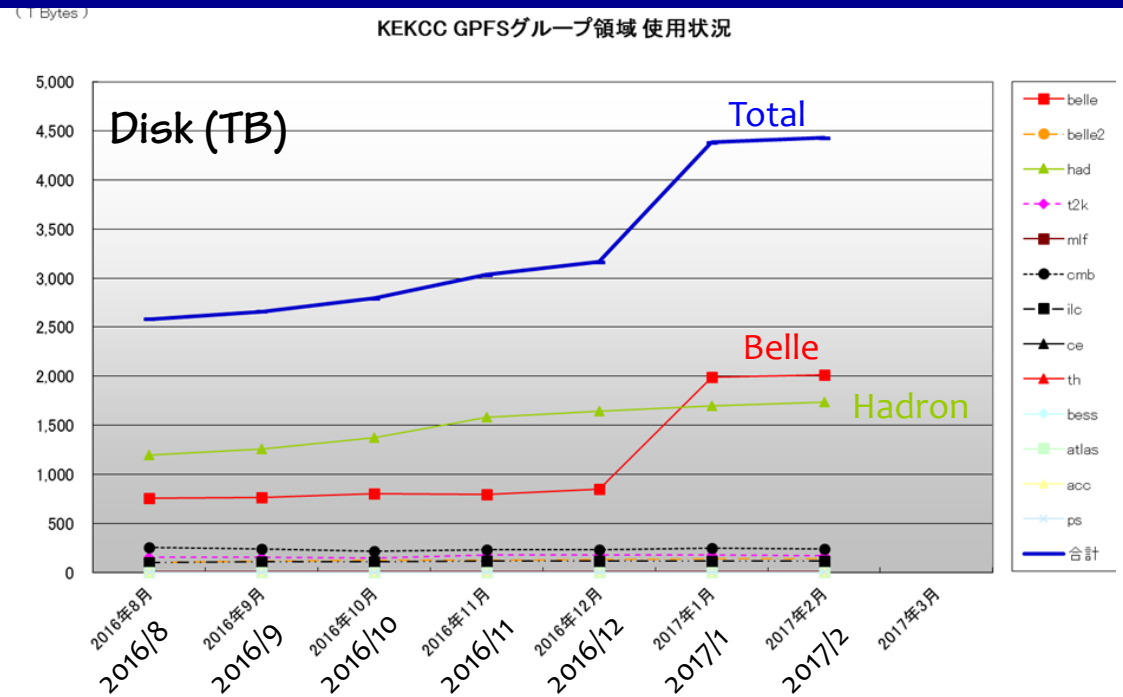
T2K (Tokai to Kamiokande)

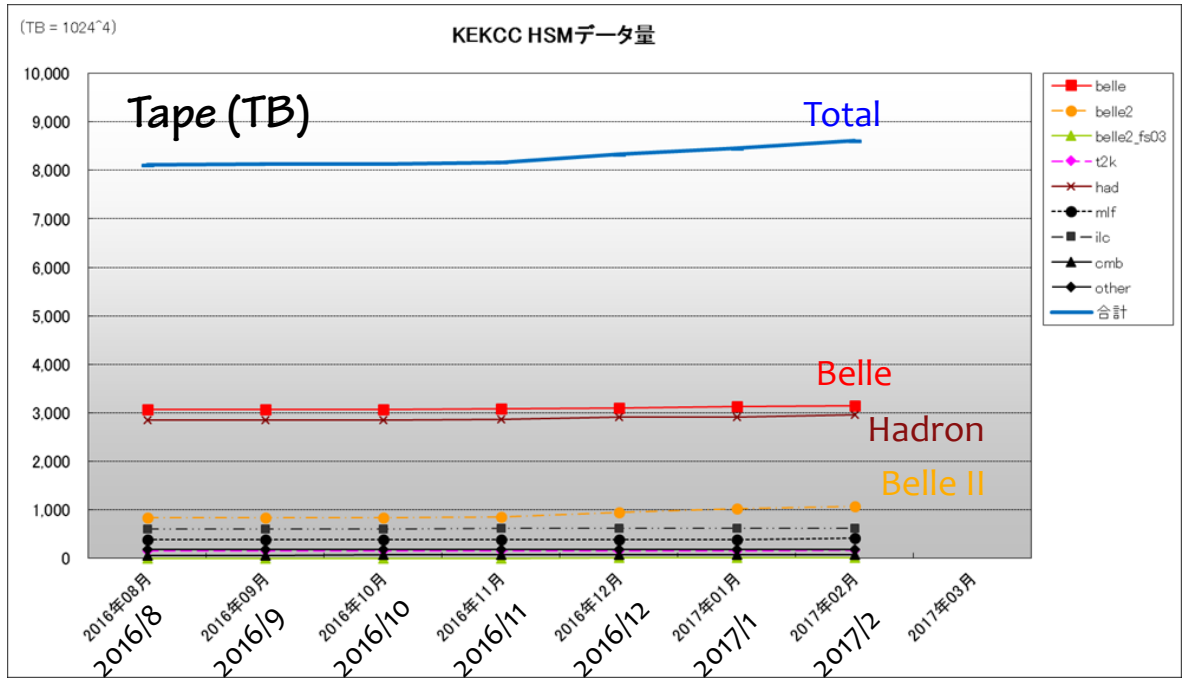
J-PARC



Nuclear/Hadron experiments

Total amount of Data on KEKCC





KOTO experiment (part of "Hadron)

RAW data : ~2PB, product DATA (for analysis) : ~0.5PB, MC : 0.13PB

(for 3 years operation)

1PB/yr RAW data is expected from now on

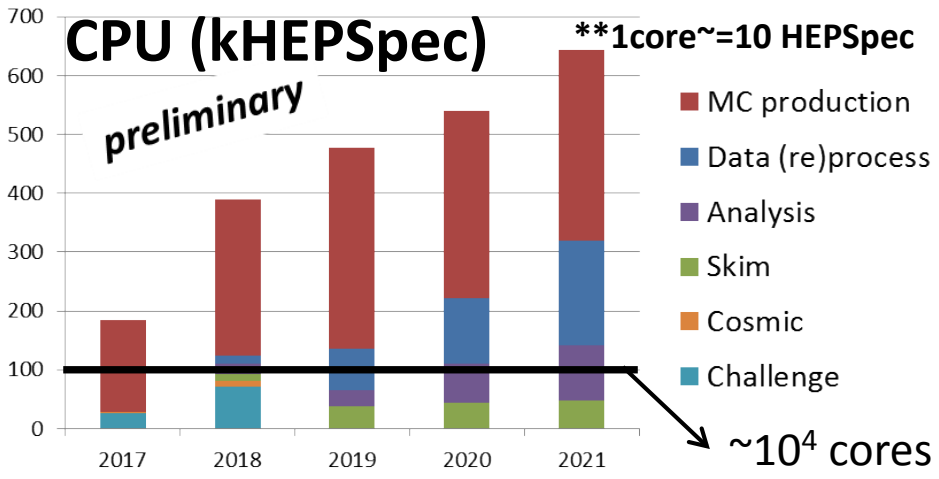
T2K (neutrino experiment, part of "Hadron)

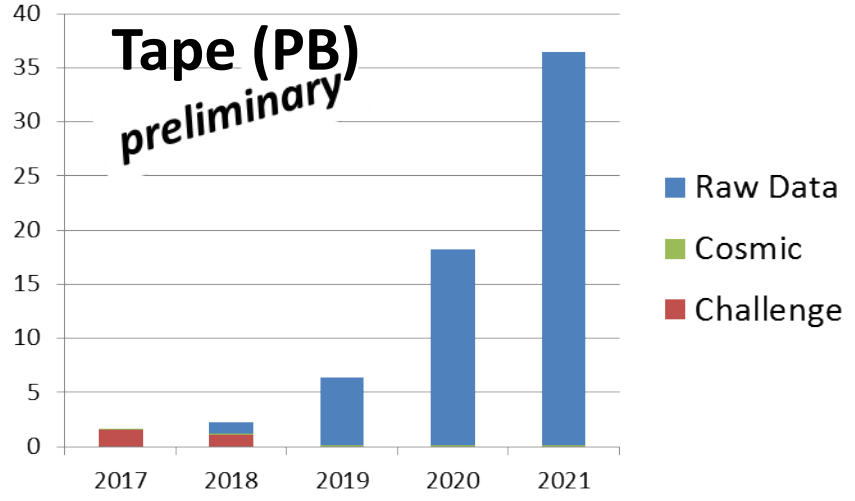
RAW data + MC : 0.05 PB, product DATA (for analysis) : 0.02PB

(for 5 years operation)

another 0.1PB RAW data + MC are stored on Data GRID @ UK + Canada

Expected Computing resoureces for Belle II

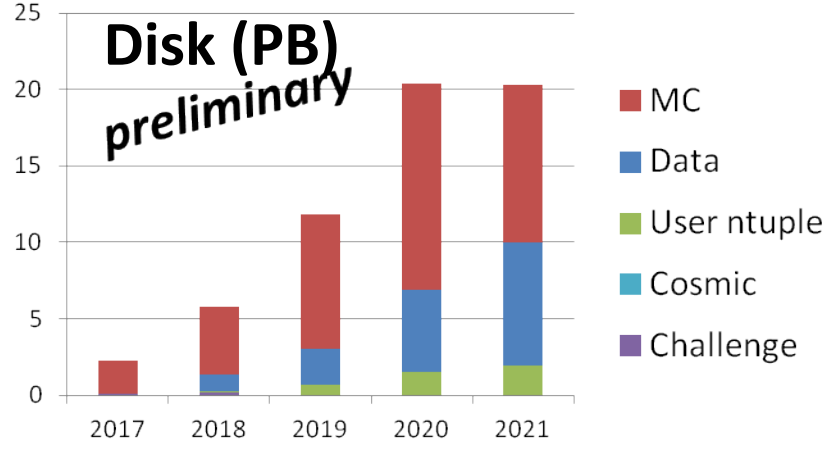




- Estimation until 2021 (~20 ab^{-1}).

- At the end of data taking (50 ab^{-1}), more than
  - 100000 core CPU
  - 100 PB storage

  are expected to be needed to store and analyze data in a timely manner.



in Summer 2020 ....

We will have a replacement of KEKCC in summer 2020, again...

Data migration is a potential concern...