

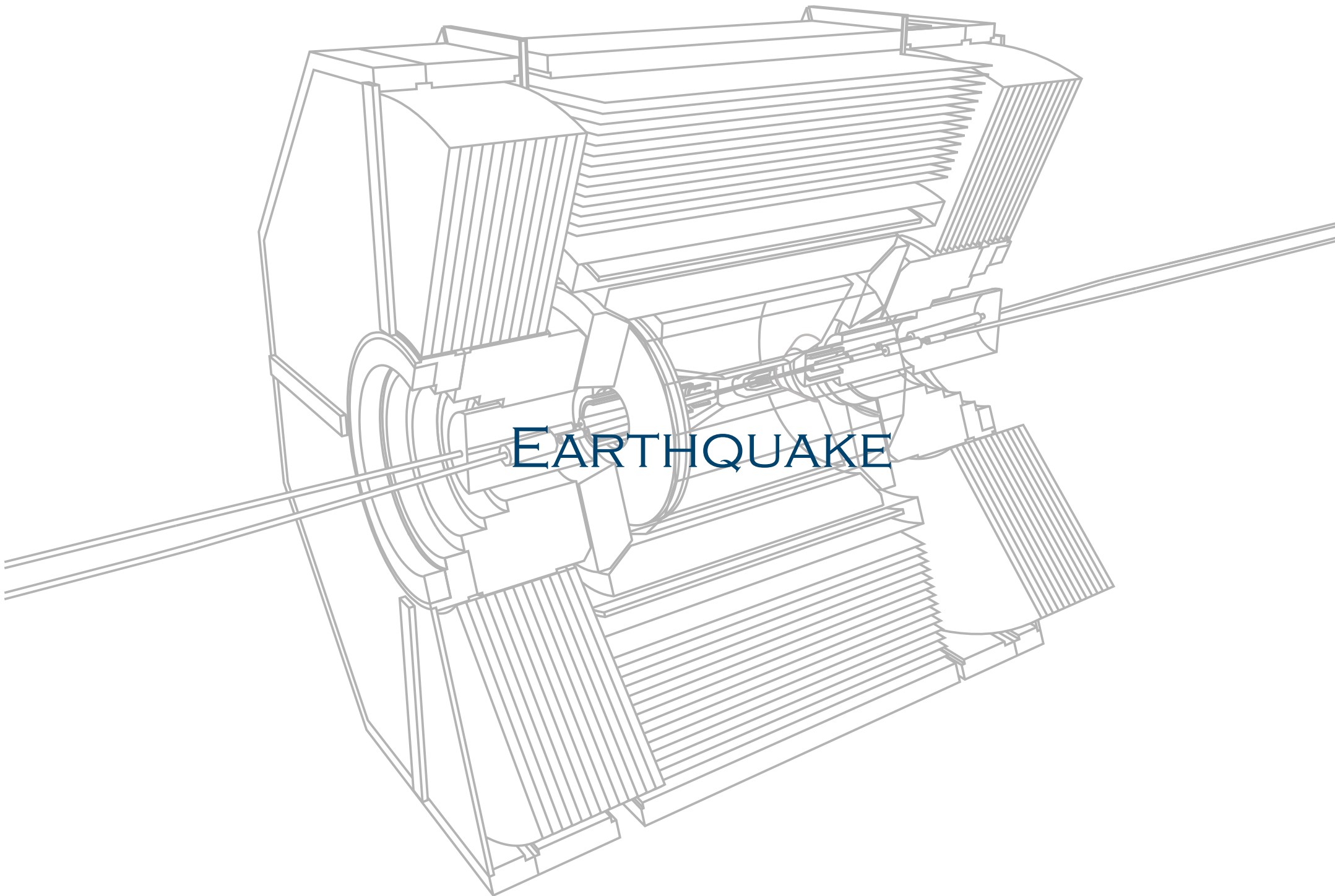
THE 5TH WORKSHOP ON DATA PRESERVATION
AND LONG TERM ANALYSIS IN HEP
16-18 MAY, 2011, FERMILAB, USA



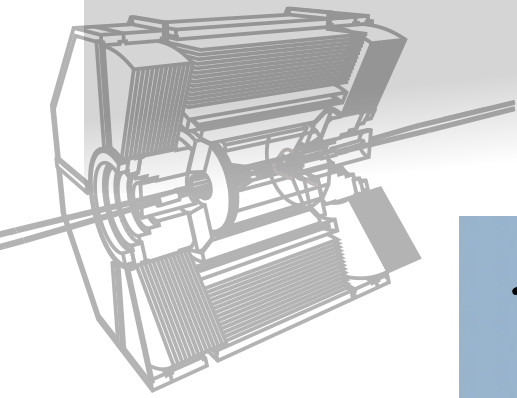
DATA PRESERVATION AT BELLE/KEK

DAVID ASNER (PNNL)

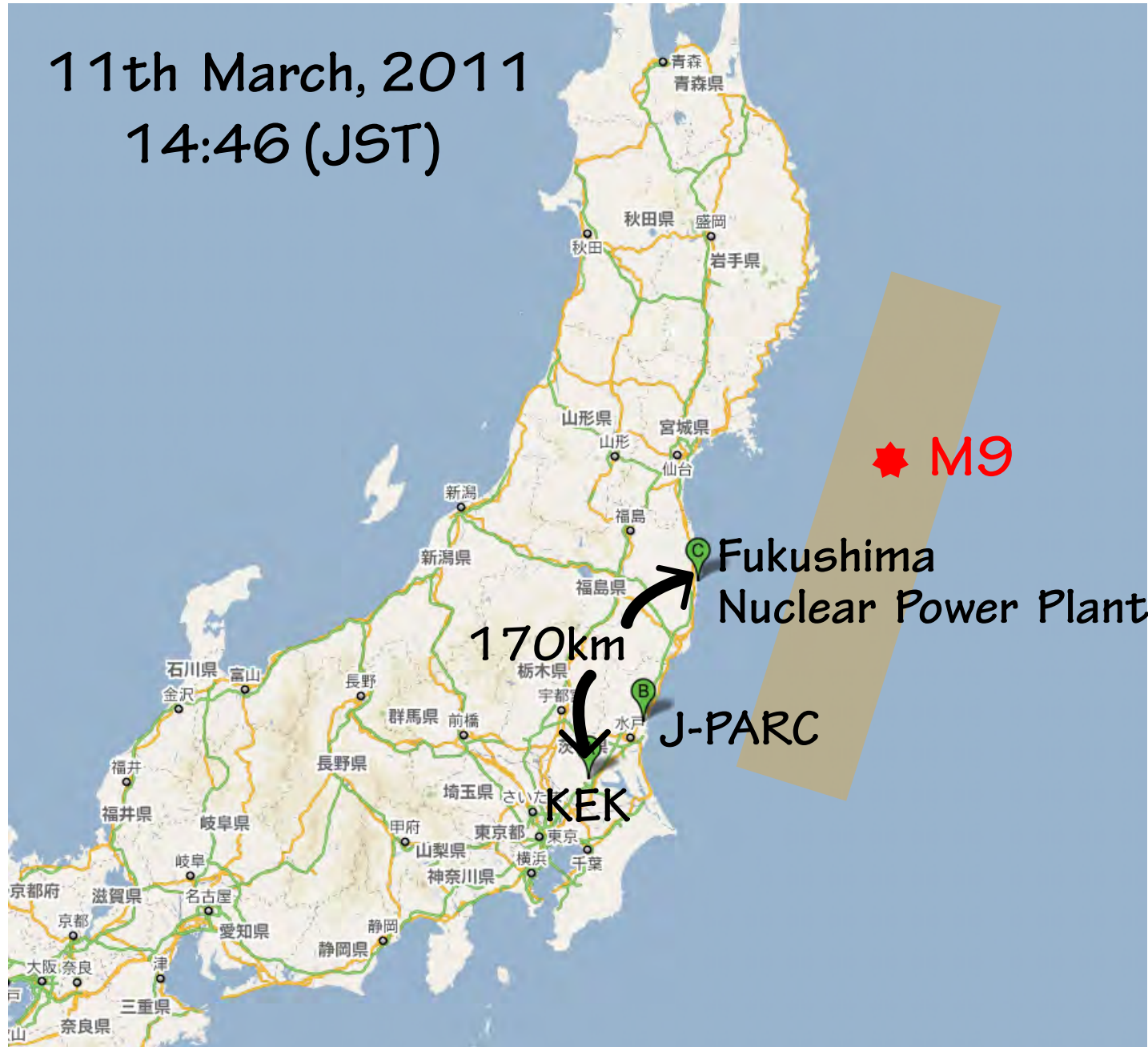
AS A PROXY OF TAKANORI HARA (KEK)



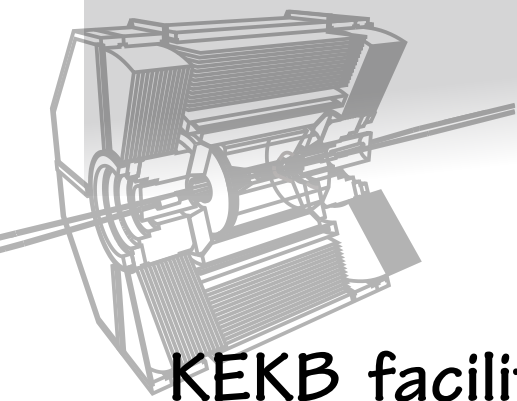
EARTHQUAKE



11th March, 2011
14:46 (JST)



EARTHQUAKE



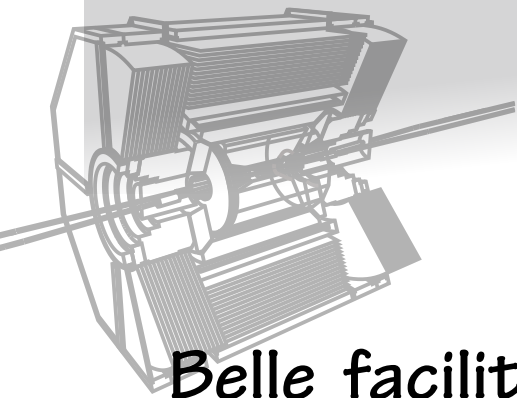
KEKB facility

- + Displacements between the expansion joints of the tunnel.
- + High Energy Ring (HER) also needs re-alignments.
- + Displacement and damages of many LINAC components
- + Klystron cooling facility

Fortunately we are in the process of disassembly ...

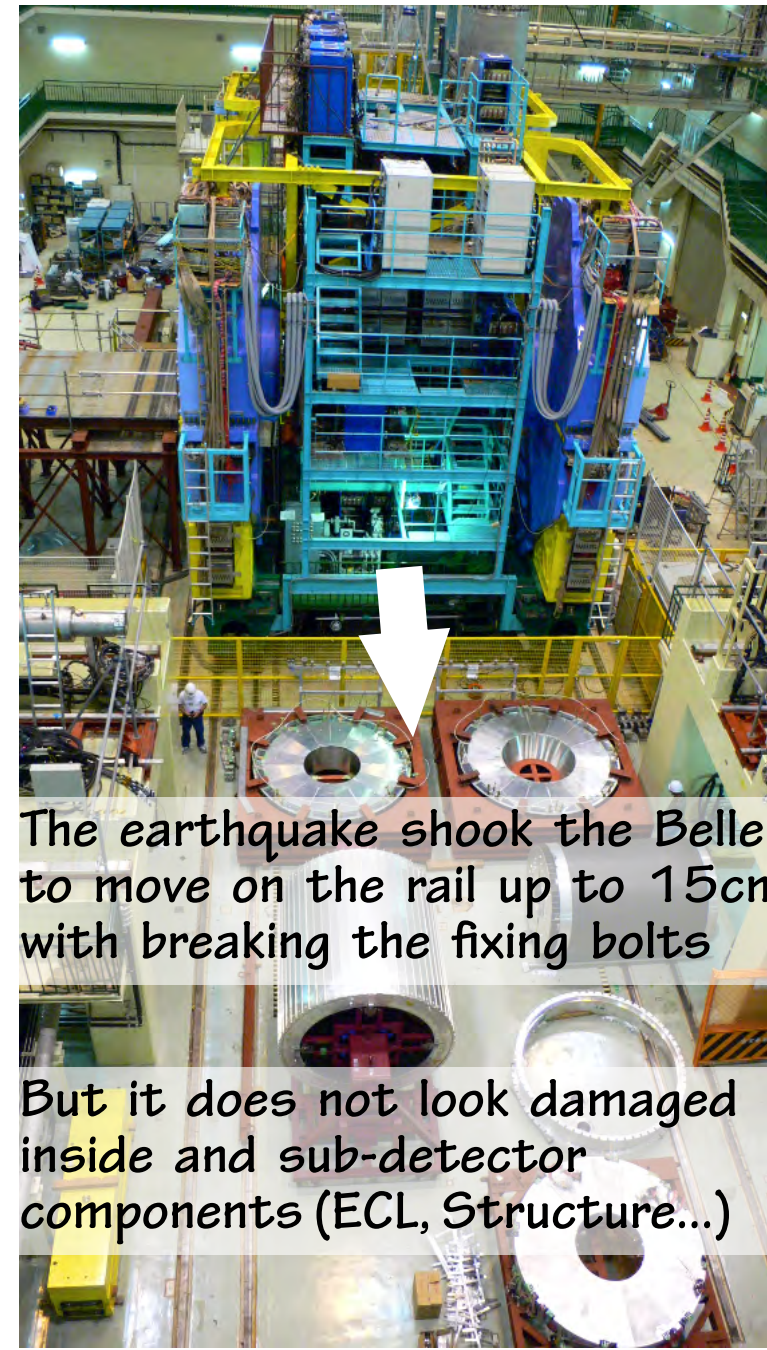
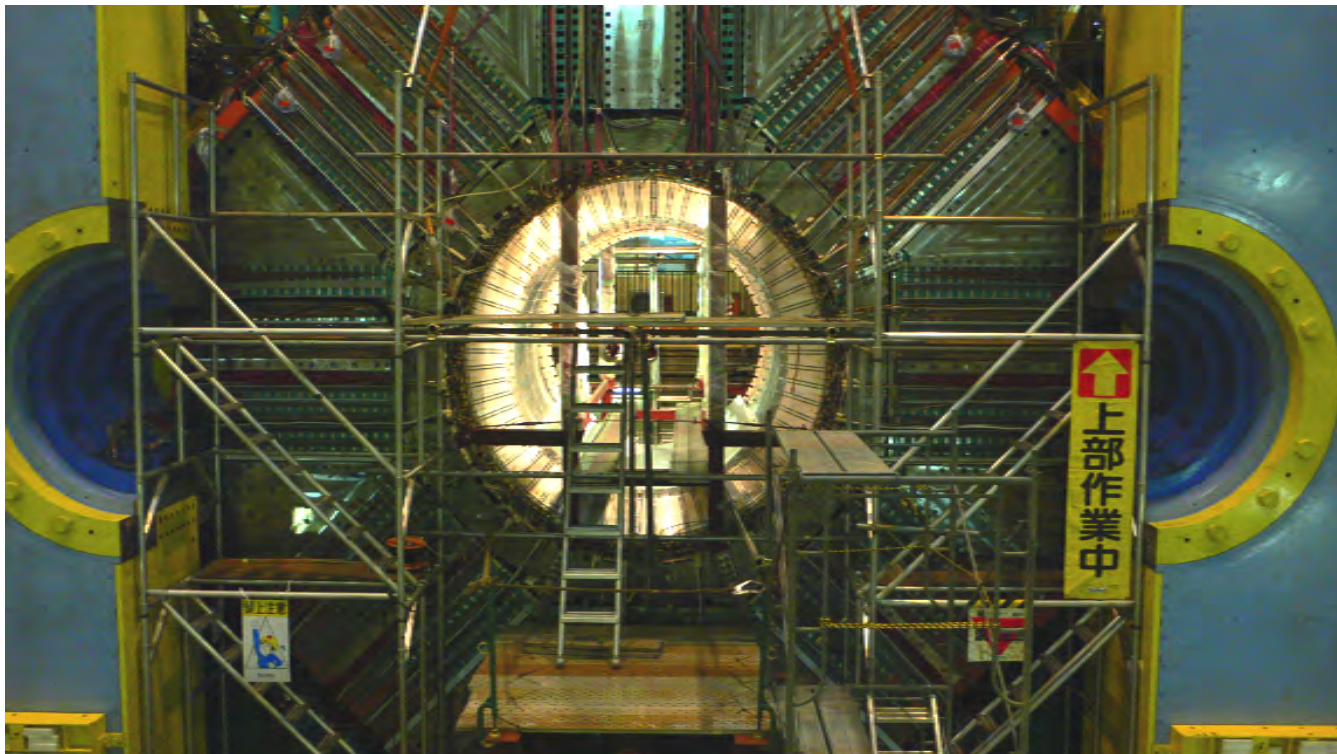


EARTHQUAKE



Belle facility

Fortunately again, Belle has been rolled out and unloaded by the end of January

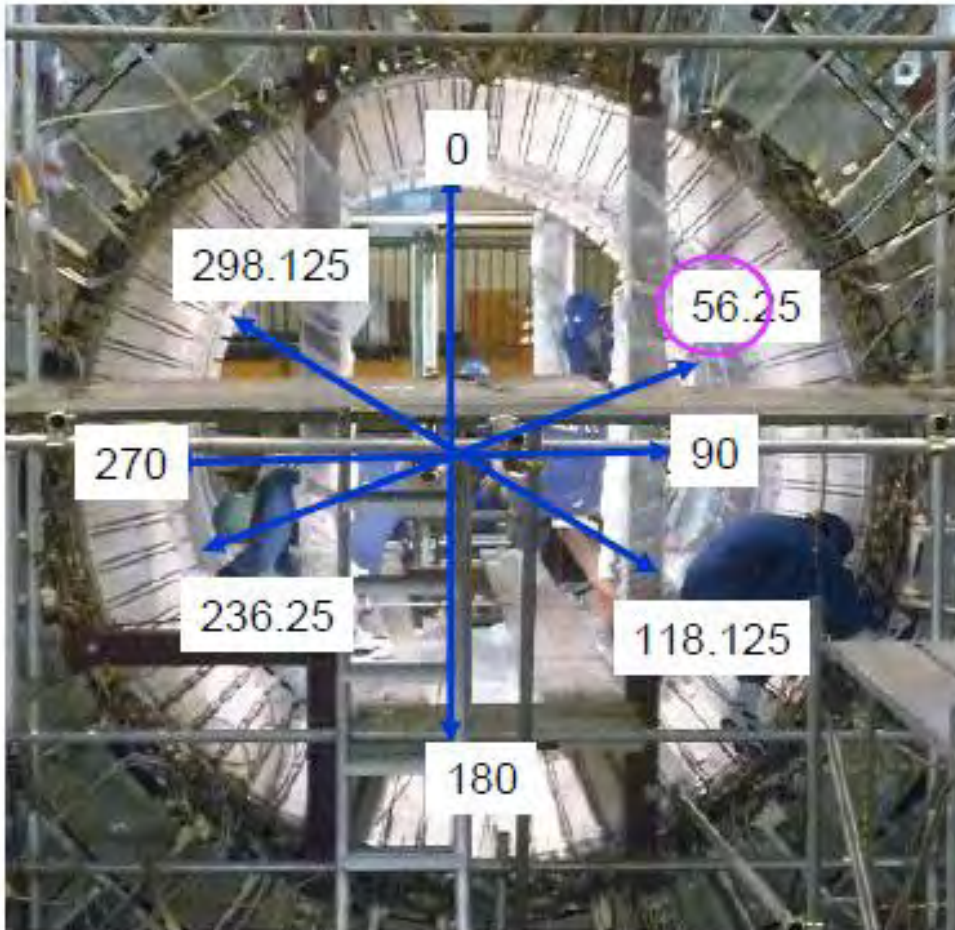


The earthquake shook the Belle to move on the rail up to 15cm with breaking the fixing bolts

But it does not look damaged inside and sub-detector components (ECL, Structure...)

ECL: small deformation ?

Backward (日光側)のECLフランジの直径比較

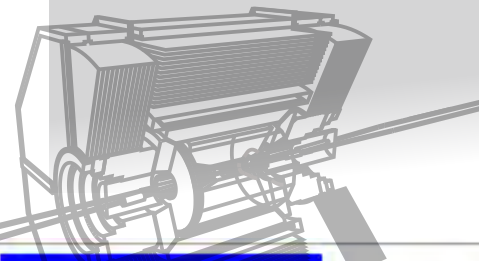


比較(設計値2360.2) reference value

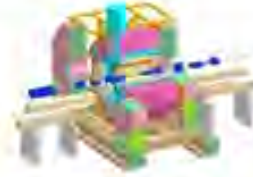
Flange(Forward)

	Jan. 24	May 10
Location	Dia.(mm)	Dia.(mm)
0-180deg.	2358.5	2358.1
56.25-236.25	2360.6	2360.8
90-270	2360.6	2361.3
118.125-298.125	2361.5	2361.3

EARTHQUAKE



**Belle
Collaboration**



Belle is an experiment at the [KEK B-factory](#). Its goal is to study the origin of CP violation.

[\[一般向 \(日本語\)\]](#) [Introduction \(English\)](#)



[Belle
members
only](#)



[KEK
\[English\]](#)



[KEKB](#)



[Belle II](#)

Belle Calendar [2011](#)

On Earthquake

As is now well known, Japan suffered a terrible earthquake and tsunami on March 11, which has caused tremendous damage, especially in the Tohoku area. Fortunately, all KEK personnel and users are safe and accounted for. The injection linac did suffer significant but manageable damage, and repairs are underway. The damage to the KEKB main rings appears to be less serious, though non-negligible. No serious damage has been reported so far at Belle.

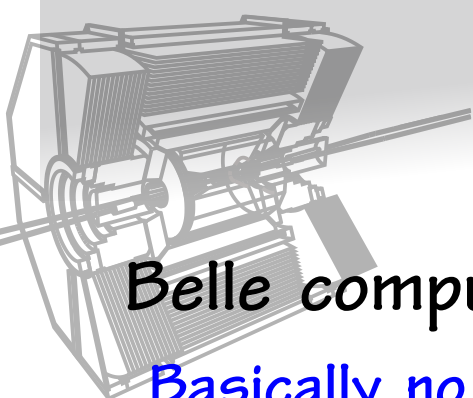
Further investigation is necessary. We would like to convey our deep appreciation to everyone for your generous expressions of concern and encouragement.

[Recent Physics Results](#)

[Conference Papers 2010](#)

We got many kind offers from the world (inside/outside of the Belle community), e.g. temporary storage space, CPU powers. We cordially appreciate that.

ROAD TO RECOVERY



Belle computer

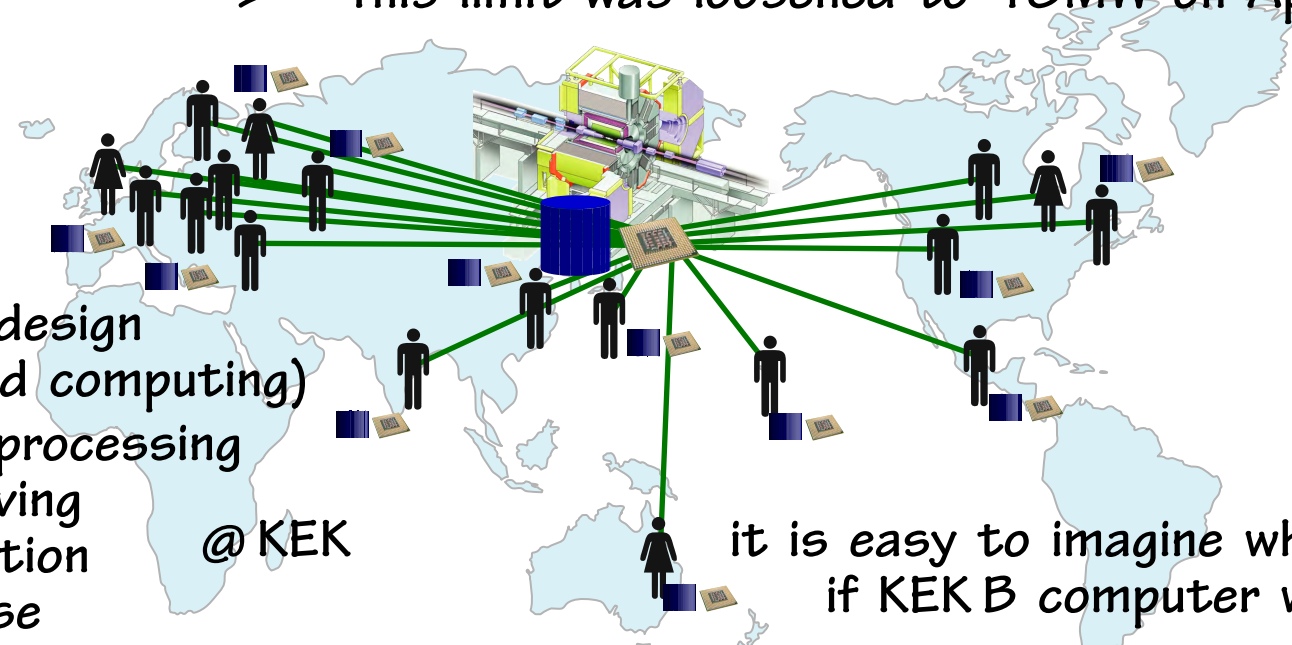
Basically no damage in the hardware of the B computer

Severe limitation of electricity usage in KEK in March and April

Serious shortage of electricity supply in the area hit by the earthquake
Scheduled electricity cut-off was performed

➔ Electricity usage in KEK was limited to 2MW
(usually 20MW without running KEKB accelerator)

➔ This limit was loosened to 18MW on April 28th



Belle computing design
(= KEK-centralized computing)
Raw data processing
Data archiving
MC production
Analysis use
@ KEK

it is easy to imagine what would happen if KEK B computer were not available...

ROAD TO RECOVERY

March 2011

sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30	31		

- Red numbers (6, 11, 13, 20, 27) indicate days with bwg.
 - Blue numbers (12, 19, 26) indicate days with bwg.
 - Black arrows indicate 'off' periods: March 13-16, 20-25, and 28-31.
 - Blue arrows indicate 'bwg' periods: March 11-12, 18-19, and 27-28.
 - A red 'X' is drawn over the cell for Friday, March 11.

bwg = login servers

100kW cfg = *bwg*+HSM+ file servers + 1/5 comp. servers

~normal = almost full operation (usually 3/5 comp. servers for Belle I)

remaining 2/5 for GRID operation

April 2011

sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
					1	2
3	4	5	6	7	8	9
10	11	12	13	14	15	16
17	18	19	20	21	22	23
24	25	26	27	28	29	30

- Red numbers (3, 10, 17, 24) indicate days with bwg.
 - Blue numbers (1, 2, 9, 16, 23, 30) indicate days with bwg.
 - Black arrows indicate 'off' periods: April 1-2, 4-7, and 11-12.
 - Blue arrows indicate 'bwg+HSM' periods: April 1-2, 3-4, 5-6, 7-8, 9-10, 15-16, 17-18, 23-24, 29-30.
 - Purple arrows indicate '100kW cfg' periods: April 10-11, 12-13, 17-18, 24-25, 26-27.
 - A red arrow indicates '~normal' operation from April 29-30.

(we expect the current level of operation even in summer season)

ADVERSITY AS AN OPPORTUNITY

Limited resources oblige us to use CPU/Storage more effectively
(as well as our brains...)

We should keep the tension for Physics analysis under adverse conditions.

→ decided to do a coordinated “Official Physics Skimming”

reduce DISK I/O, CPU load

March 2011

sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30	31		

almost could not do anything but we discussed a coordinated official skimming

April 2011

sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
					1	2
3	4	5	6	7	8	9
10	11	12	13	14	15	16
17	18	19	20	21	22	23
24	25	26	27	28	29	30

bwg+HSM call for skimming modules

deadline for skimming modules

~70 modules were prepared

start of official skimming

Y4S on-res. : finished by May 3

Y5S on-res. : finished by May 2

Y4S MC (10streams) : on-going
(will finish by May 20)

Y5S MC (6streams) : by April 28

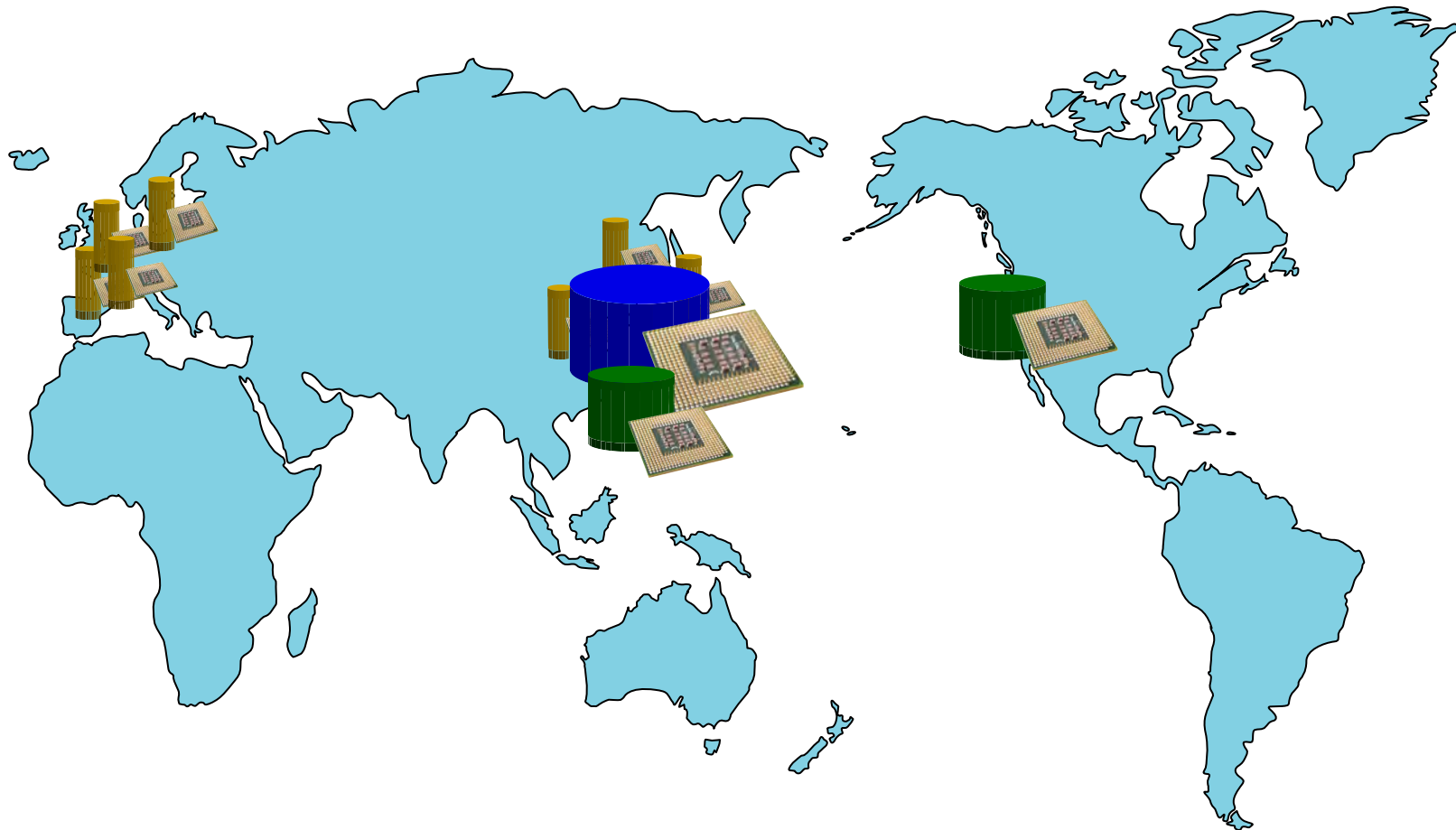
LEARN FROM THE EARTHQUAKE

Produced skimmed data will be distributed to the “local analysis centers”

Karlsruhe (Germany), KISTI (Korea), IFJ (Poland), Ljubljana (Slovenia),
MPI (Germany), Nara WU, RCNP, Tohoku U (Japan), Vienna (Austria), VPI (USA)

and two big analysis centers

Nagoya U (Japan), PNNL (USA)



LEARN FROM THE EARTHQUAKE

Produced skimmed data will be distributed to the “local analysis centers”

Karlsruhe (Germany), KISTI (Korea), IFJ (Poland), Ljubljana (Slovenia), MPI (Germany), Nara WU, RCNP, Tohoku U (Japan), Vienna (Austria), VPI (USA)

and two big analysis centers

Nagoya U (Japan), PNNL (USA)

Belle computing at KEK is now in almost normal operation, and we expect the operation will be something similar to the current level even in hot summer season in Japan.

But,

it depends on the situation of electricity power supply in Tokyo-area

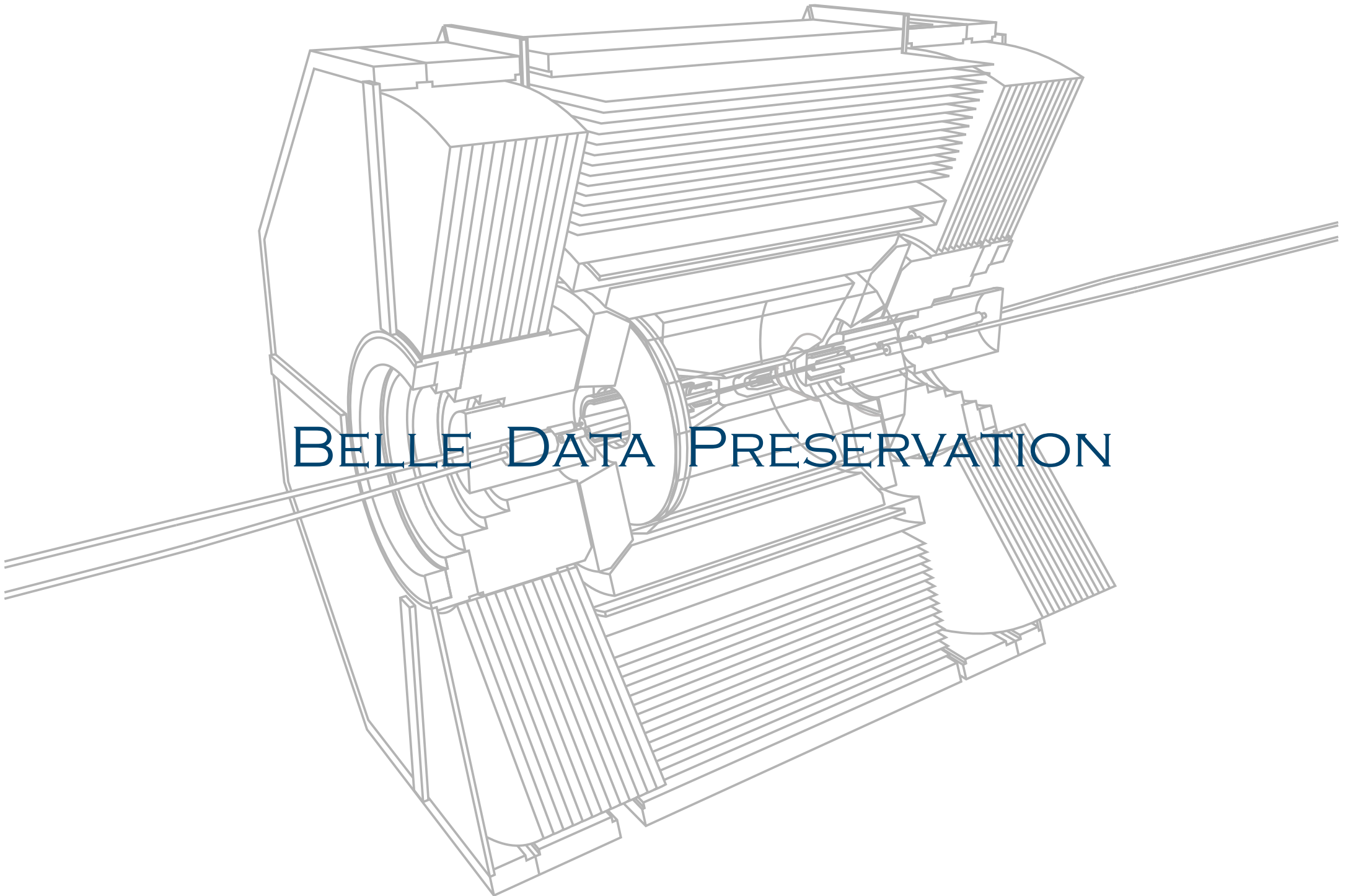
→ longer summer shutdown or minimum operation for certain period may happen in this summer

Plus,

Another long-term B-computer shutdown (~40 days) may happen, discussion is on-going for the replacement of the current system with new one. because of budget, limited available space, etc. (for schedule, see page 16)

Therefore,

To avoid a “possible” period of unavailability of Belle data, it is important to have alternative Belle data/analysis centers outside KEK and a good exercise for Belle II distributed computing

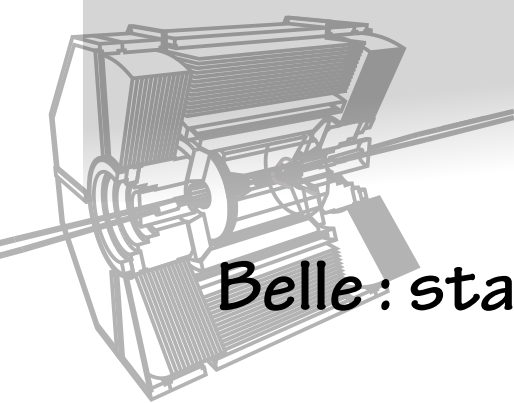


BELLE DATA PRESERVATION

BELLE DATA

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

now moved to “Intense Analysis Phase”



Size of storage for Data

RAW	raw data	>1000 TB
DST	prescaled data (1-1/400)	3-400 TB?
mDST	reconstructed info.	140 x 2 TB

Format: panther (Belle's own bank system)

Size of storage for MC

mDST	recon. info. + MC true	600 x 2 TB
------	------------------------	------------

10 streams for bb + 6 streams for udsc

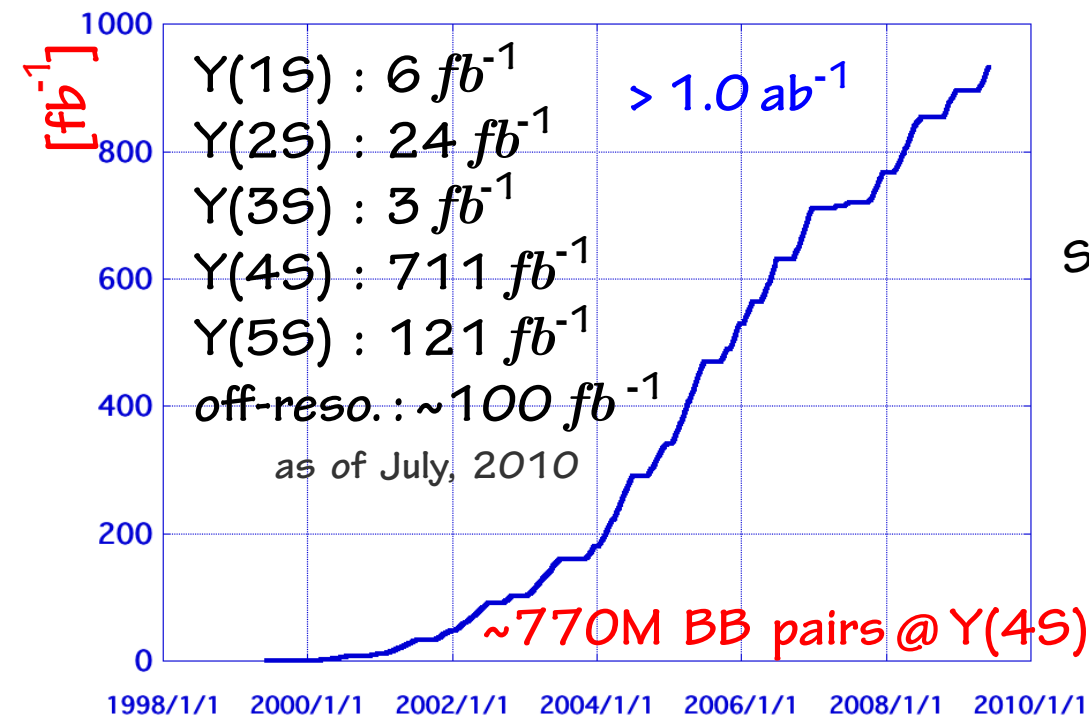
we have two versions of mDST

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

Integrated Luminosity(log)

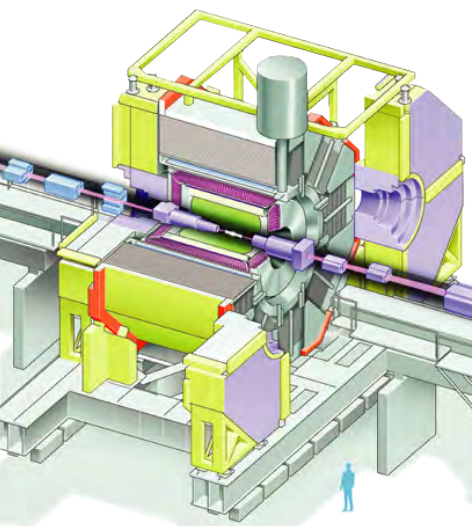


BELLE'S POLICY

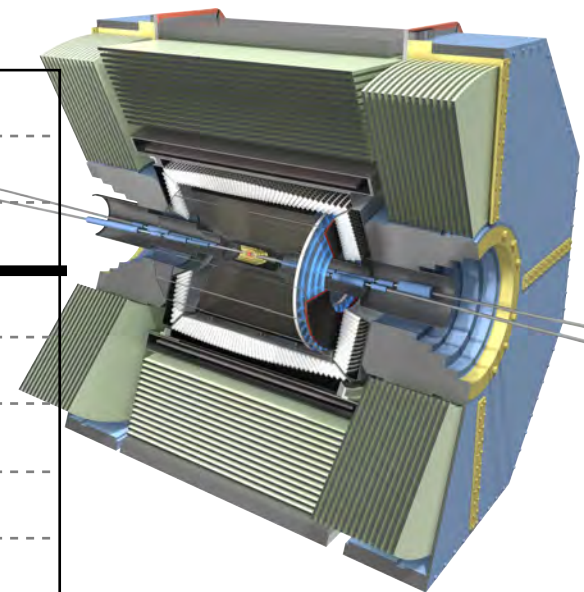
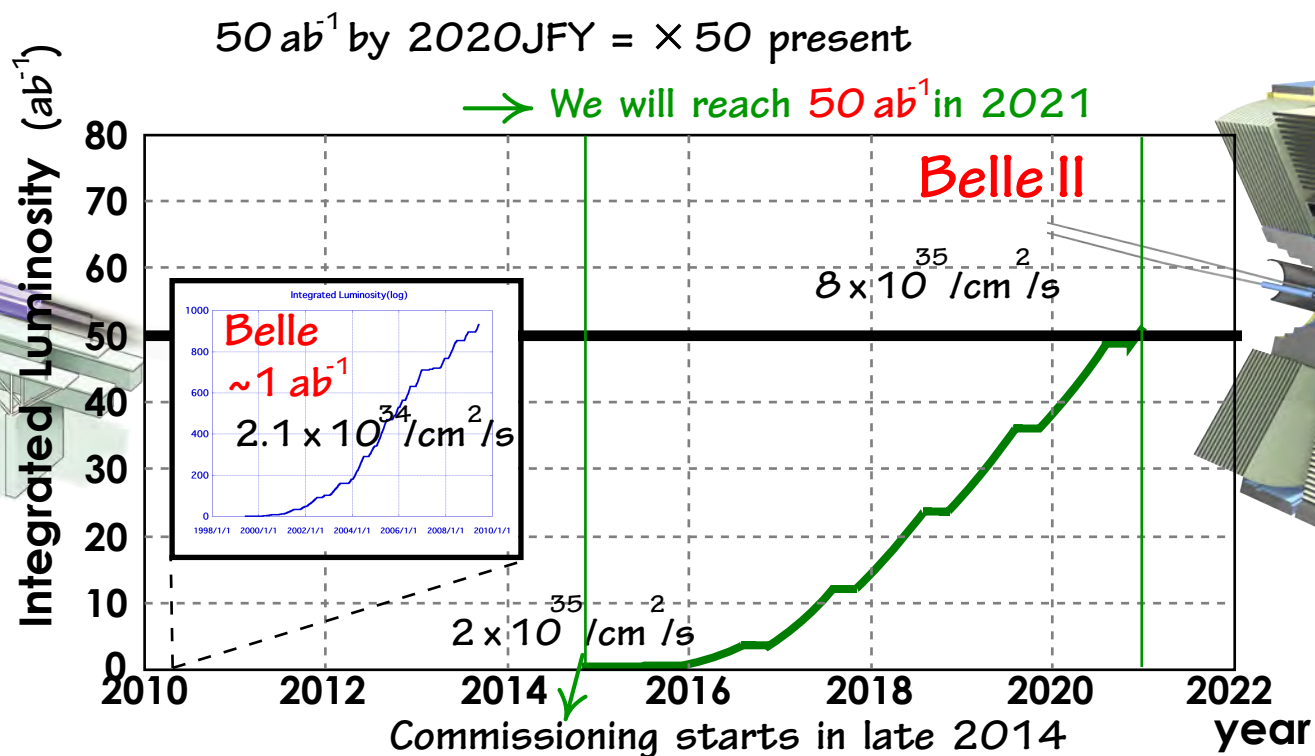
Preserve all RAW and mDST, at least.

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

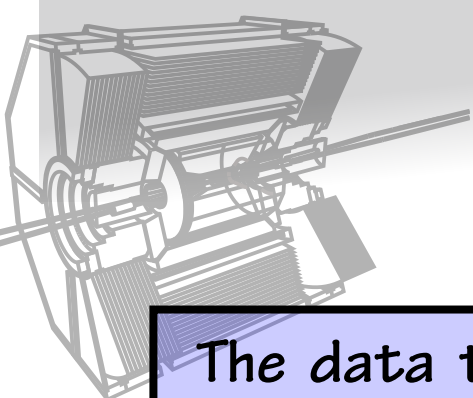
*Belle I data will be used by the Belle community
until the time the statistics of Belle II > Belle I data set*



Belle
(operation:
1999-2010)



HOW LONG ?



The data taking of the Belle experiment finished in 2010
But in late 2014, SuperKEKB accelerator commissioning starts
(no delay because of the earthquake)

- . We have to work on updating the Belle detector
- . We will have a new computing system for Belle II experiment
(including resources for Belle I + J-PARC)

We are aware of the need for Belle data preservation

Belle $\Upsilon(4S)$ data will be superseded in 2015-16?

It is important to validate Belle II initial results

$\Upsilon(nS)$ data are unique data sets

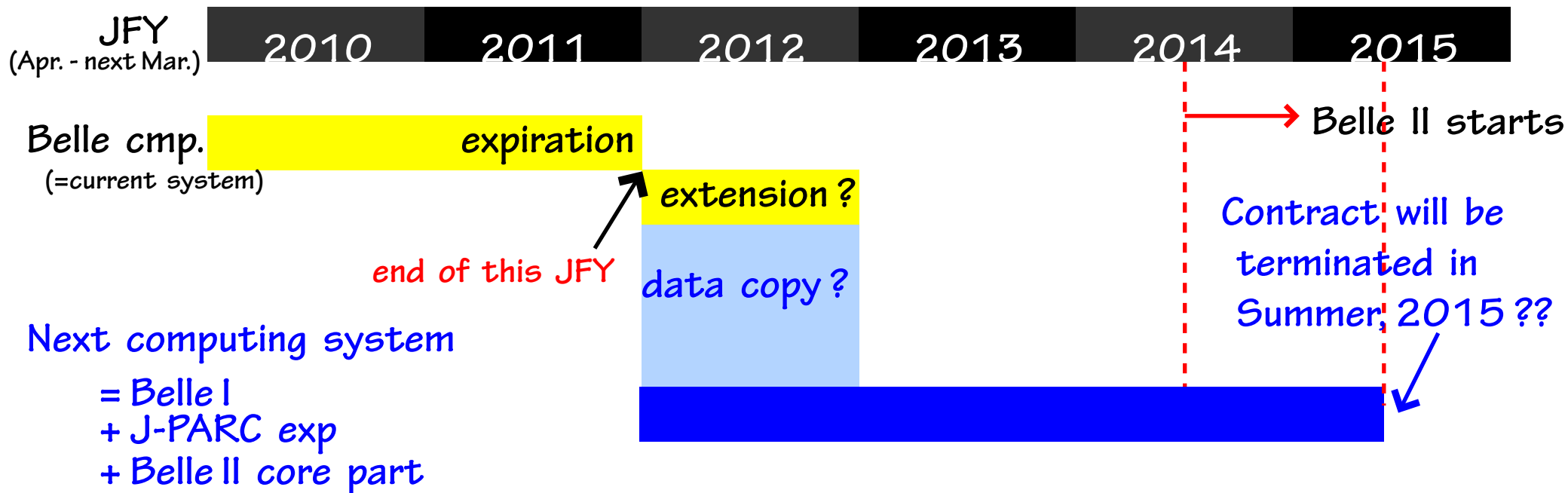
Belle data should be preserved at least until 2016-17?

Until that, the Belle Data will be stored on storage
within the Belle community

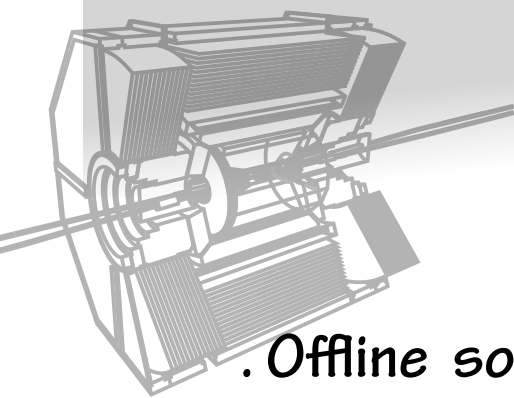
NEXT COMPUTER SYSTEM

Year (contract) specification	1999- (4years)	2001- (5years)	2006- (6years)	2009- (continued)	2012- (until 2015/8?)
CPU [SI2k]	~100 (WS)	~1200 (WS+PC)	~42500 (PC)	~115200 (PC)	~30000 (for physics analysis)
Disk [TB]	4	9	1000	1500	1000 (for mDST w/ new-tracking)
Tape [TB]	160	620	3500	3500	3000

(for Belle I)



STATUS AND PROSPECT



- . **Offline software : almost frozen**
event generator, simulation, reconstruction, ...
- . **DST/MC production : not planned** new idea of reconstruction may reactivate
the offline soft/DST/MC prod. activities
- . **Retention of knowledge : not well documented** this is a bad habit of Belle....
- . **Computing resources : will be included in the next comp. system**
requested CPU/storage for Belle I analysis

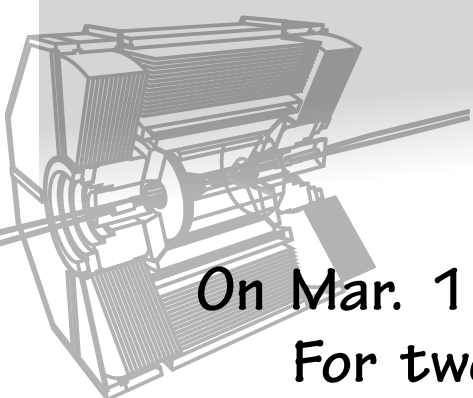
However, we need to consider

how to transfer the Belle Data on the current system
to the next system

how to establish the B-comp equivalent system
in the next system

setting of database,
treatment of file catalog
library compatibility, etc...

SUMMARY



On Mar. 11, 2011, KEK was hit by a big earthquake
 For two weeks after the earthquake, we could not turn on B-comp.
 Then, the situation of electricity got better gradually,
 and from Apr. 28, we could start almost the normal operation.

(we expect the current level of operation even in summer season)

During this two months, the availability of B-comp was limited.

However, we coordinated “official physics skimming” and performed.

For this summer and/or the end of this JFY, we may have
 a shutdown of B-computers for a certain period. Therefore
 it is important to have alternative Belle data centers outside KEK

Policy is not changed from the last DPHEP workshop

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

Till then, the Belle Data will be stored on storage
 within the Belle community at least until 2016-17?

Next computing system will be installed at the end of this JFY

Earthquake

Data Preservation

Pacific Northwest National Laboratory

Facts & Figures for FY2010

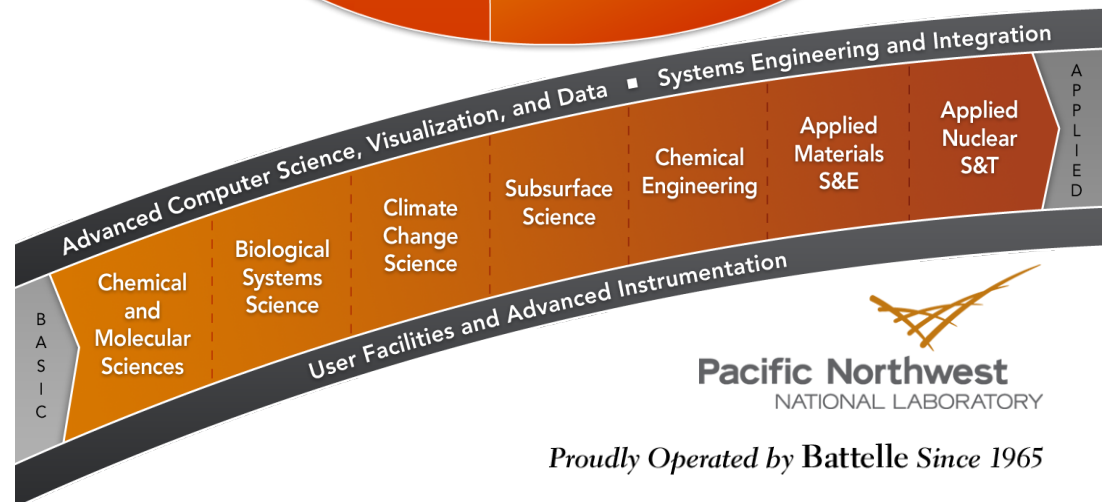
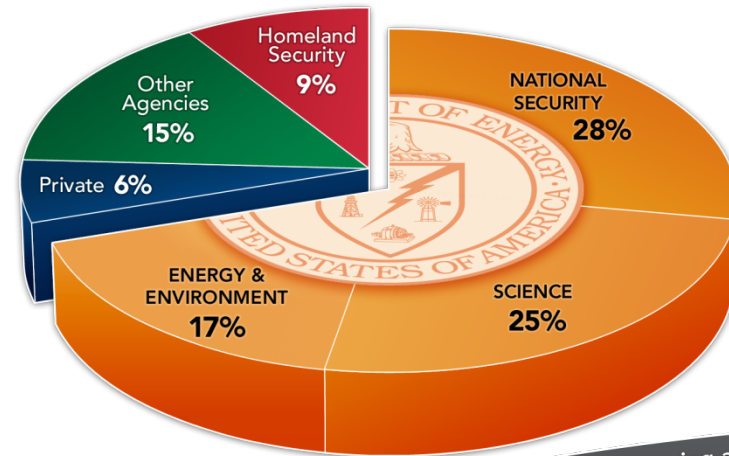


- ▶ Largest DOE Office of Science Laboratory
- ▶ > 4,900 staff with 3,000 technical staff
- ▶ \$1.1 billion in R&D expenditures

▶ Among top 1% of research institutions in publications and citations in:

- Chemistry
- Geosciences,
- Physics
- Engineering
- Biology and Biochemistry
- Environment/Ecology
- Materials science
- Clinical medicine
- Microbiology

1 ▶ 930 peer-reviewed publications



Pacific Northwest
NATIONAL LABORATORY

Proudly Operated by Battelle Since 1965

PNNL lead (only) DOE laboratory on Belle and Belle II experiments

- ▶ June 2010 – SuperKEKB and Belle II upgrades approved!
- ▶ Sept 2010 – PNNL joined Belle and Belle II experiments
- ▶ Jan 2011 – HEP announced intent to support Belle II
 - Broad role for PNNL on US Belle II effort
 - project management, detector R&D, electronics, engineering
 - scientific leadership, Belle II computing center
- ▶ April 2011 – KEK/DOE/PNNL agree to provide substantial computing for Belle
- ▶ May 2011 – Belle Executive Board Endorsed PNNL plan
- ▶ July 2011 – Belle Computing at PNNL – “plans to go live”
 - Full deployment: ~1000 cores, 1 PB disk, for up to 400 Belle users
 - In progress: Hardware configuration, software installation/testing, data transfer, new hardware acquisition, Belle user access plan
 - Effort Supported by DOE-SC-HEP and PNNL

More Details on Belle Computing at PNNL



Pacific Northwest
NATIONAL LABORATORY

Proudly Operated by Battelle Since 1965

PNNL Resources for Belle Computing

- ▶ PNNL will reconfigure cluster EMSL/NW-ICE for Belle Computing
 - 7 racks of 27 compute nodes with 8 cores
 - Priority use of 756 cores
- ▶ PNNL re-use storage
 - >1 PB of tape
 - ~150 TB of disk
- ▶ 3% of 163 TF Chinook Super Computer
 - Monte Carlo Generation



PNNL and DOE Support for Belle Computing



- ▶ **Computational Sciences Facility**
 - 10,000 square feet
 - Up to 10 MegaWatts
 - Geothermal Cooling
- ▶ **New hardware from PNNL be housed in CSF**
 - Storage: ~900 TB disk
 - Servers: ~250 cores

Network Optimization/Data Transfer Statistics

- ▶ Identified KEK-PNNL data transfer pathway and configured jumbo frames (9000 MTU)
- ▶ Qualified pathway using 10G KEK workstation (outside KEK firewall)
 - Achieved [5.6 Gbps] data transfer rate
- ▶ Qualified pathway using 1G KEK workstation (inside KEK firewall)
 - Achieved [711 Mbps]
- ▶ Qualified data transfer capability between KEK and PNNL data transfer nodes (DTN)
 - Achieved [81.3 MB/s]
- ▶ Total transfer to date > 25 TB of HadronBJ mDST
 - Completed transfer of Y(5S) HadronBJ mDST
 - Initiated transfer of Y(4S) HadronBJ skim
- ▶ Actual average transfer rates April 28 to May 8 from KEK to data transfer node (DTN) at PNNL achieved using a single session scp
 - 1.7 TB/day, 70.8 GB/hour, 1.18 GB/min, 20 MB/s
 - Note transfer rate > 60 MB/s on evening, weekend



Pacific Northwest
NATIONAL LABORATORY

Proudly Operated by Battelle Since 1965

Hardware Status Overview

▶ Accomplished

- Setup a basic belle server (belle-test)
- Validation of Exp45 Monte Carlo produced
- Mounted DTN to belle-test via 10 G
- Hardware allocated to setup a belle production server (belle.pnl.gov).
- Belle.pnl.gov running Rocks 5.4 cluster software with default CentOS5.
- Hardware and OS for this system (including 8 compute + 2 login nodes) has been configured.

▶ In progress

- Testing of 10 node cluster
- Test of queuing software

▶ This month

- Provide user access to US users for system testing
- Procure 900 TB disk storage

▶ Next month

- Scale up to full NW-ICE
 - 1512 cores
- Generate MC on Chinook SC

▶ This summer

- Deploy new hardware in CSF



Pacific Northwest
NATIONAL LABORATORY

Proudly Operated by Battelle Since 1965

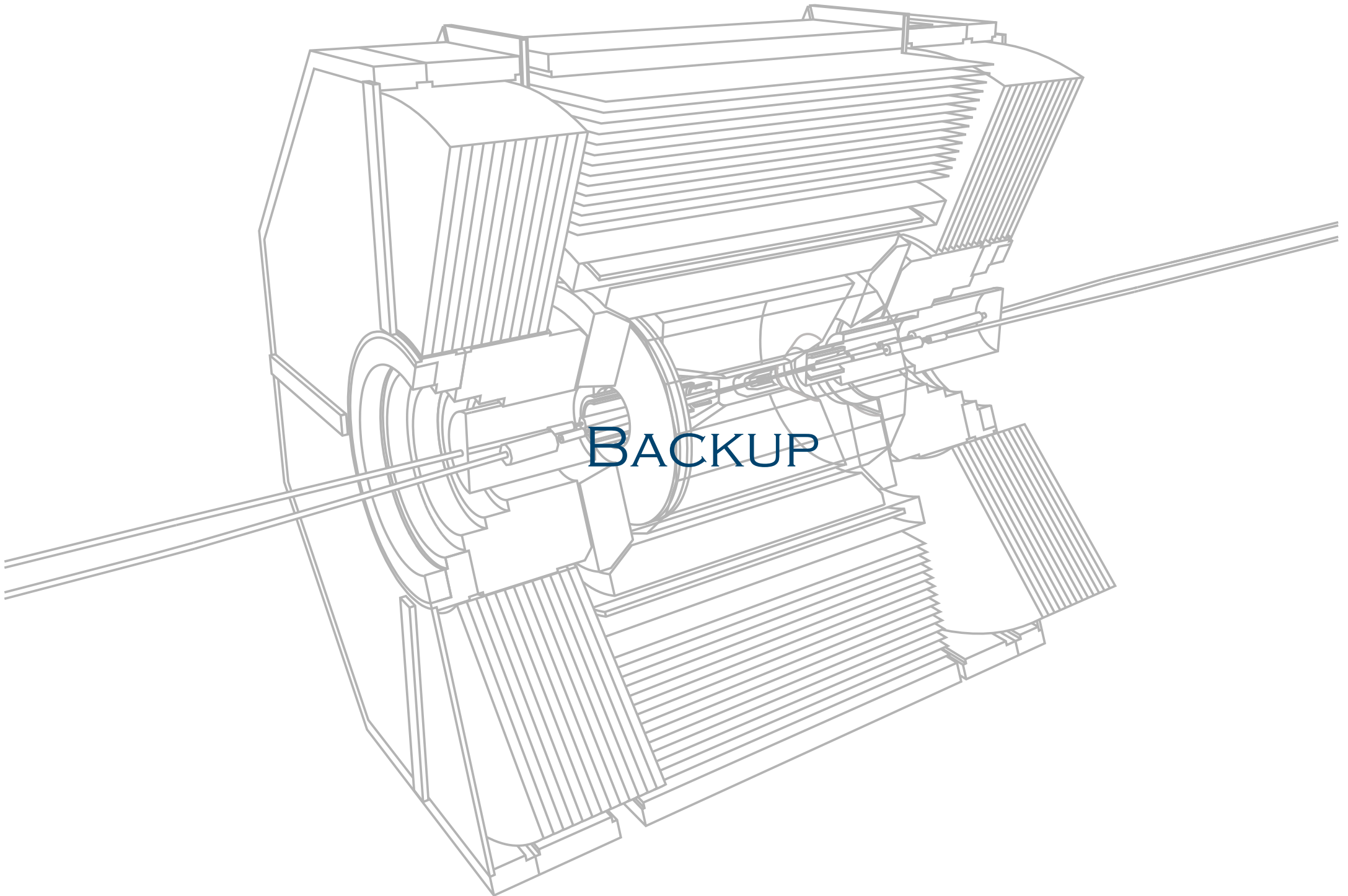
Summary

- ▶ Expect PNNL computing resources to be tested, deployed, and available to all Belle users in July
 - Existing PNNL resources will remain in EMSL
 - New resources will be located in CSF
- ▶ We plan to provide computing resources for Belle through the computing transition at KEK at the end of JFY11/start of JFY12.
- ▶ Proposal to evolve PNNL Belle computing into Belle II grid site is in progress
- ▶ Many thanks to Hara-san, Nishida-san, Karim Trabelsi, Soh Suzuki-san and Leo Pillionen for technical advice



Pacific Northwest
NATIONAL LABORATORY

Proudly Operated by Battelle Since 1965



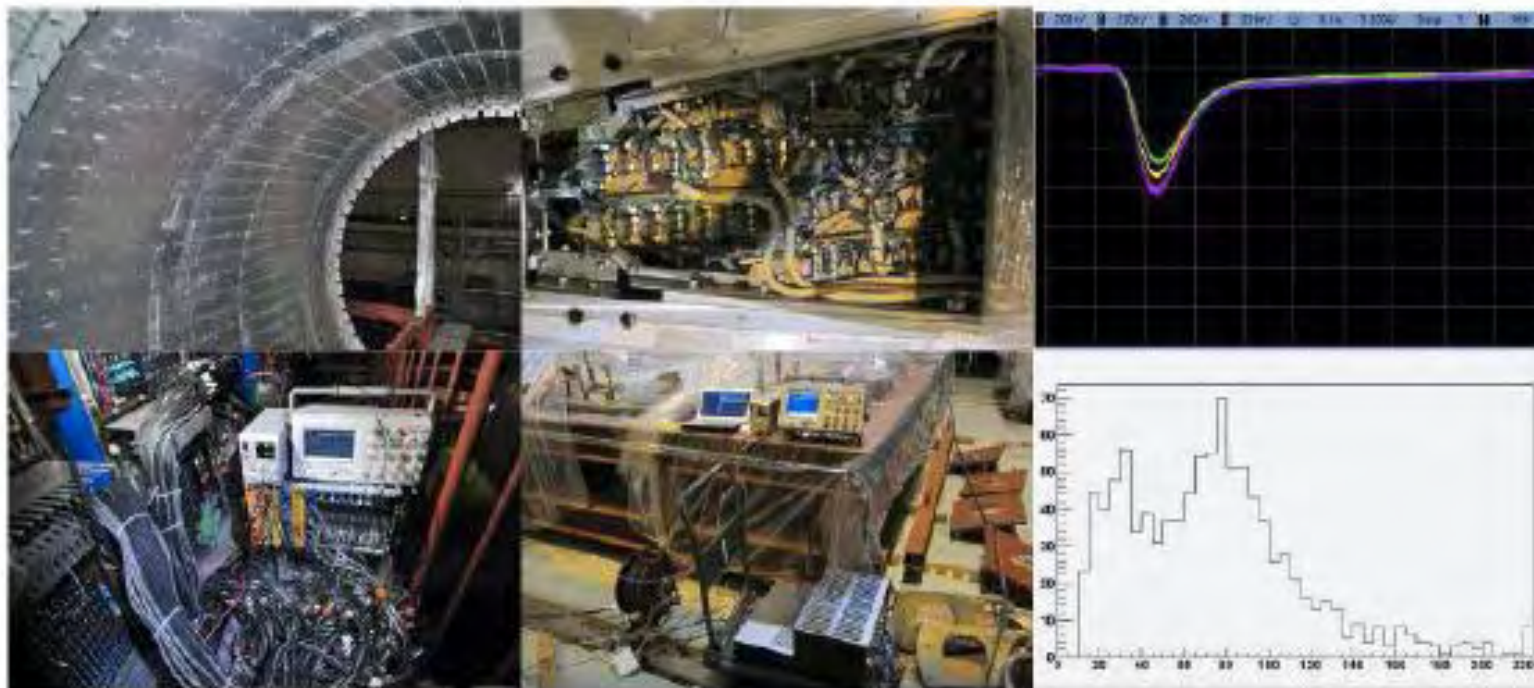
BACKUP

ECL Calorimeter

- ❑ Dryair circulation recovered on 16th March
- ❑ Barrel structure looks fine
 - ◊ measurement of barrel geometry will be performed soon
- ❑ Endcaps were on the B4 Floor. No damage observed
- ❑ Signal from barrel counters confirmed with LED and cosmic
 - ◊ 920 out of 6624 channels checked, no dead channel
 - ◊ complete cosmic test when enough DAQ boards are produced by the end of JFY2011



簡易乾燥空氣供給器



KLM RPC

- Butane wormer was broken
- Gas pressure test was done
seems no glass broken
- Needs to dry further to
supply HV

