

Distributed Computing Models for SuperBelle Experiment

Takanori Hara (KEK)
on behalf of SuperBelle Computing group



The background features a light green gradient with a central dark teal band. White decorative elements include stylized flowers, leaves, and circular patterns. A large, faint white graphic of a stylized eye or lens is centered behind the text.

Introduction

Current Belle & Future Super Belle

What is B-Factor?

To explain the dominance of matter in the universe

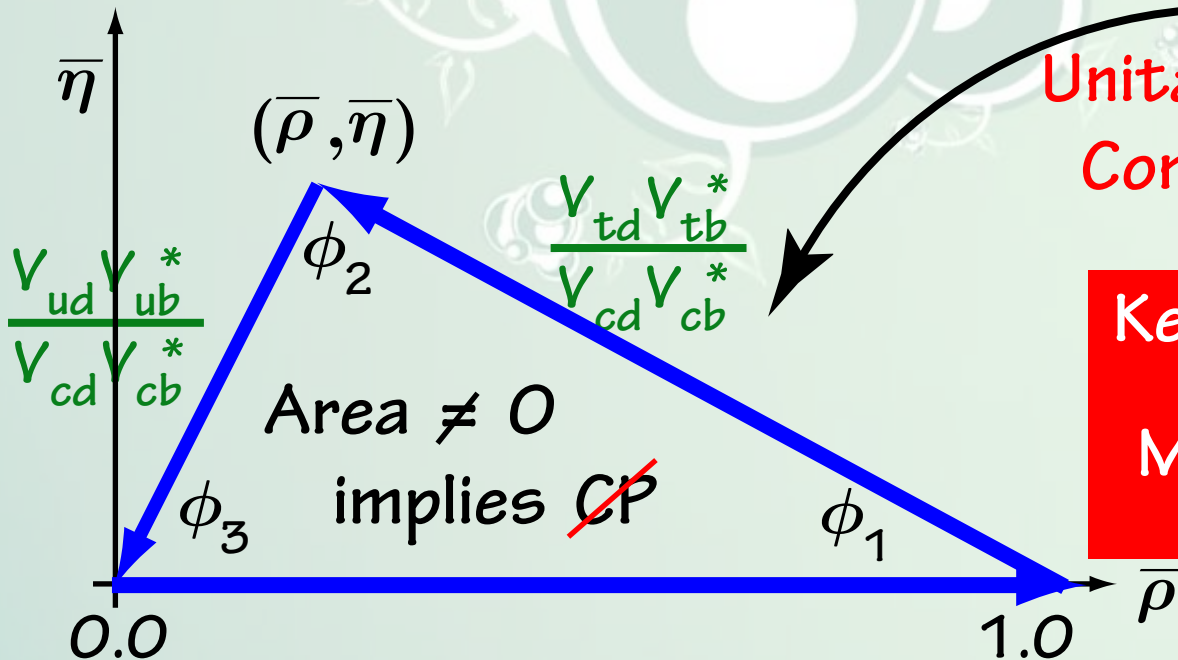
the Sakharov conditions

- . Departure from thermal equilibrium.
- . Baryon number violation.
- . **CP-symmetry violation.**

In the SM of particle phys,
Kobayashi-Maskawa matrix
(= quark mixing matrix)

$$\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}$$

**Unitarity
Constraints**

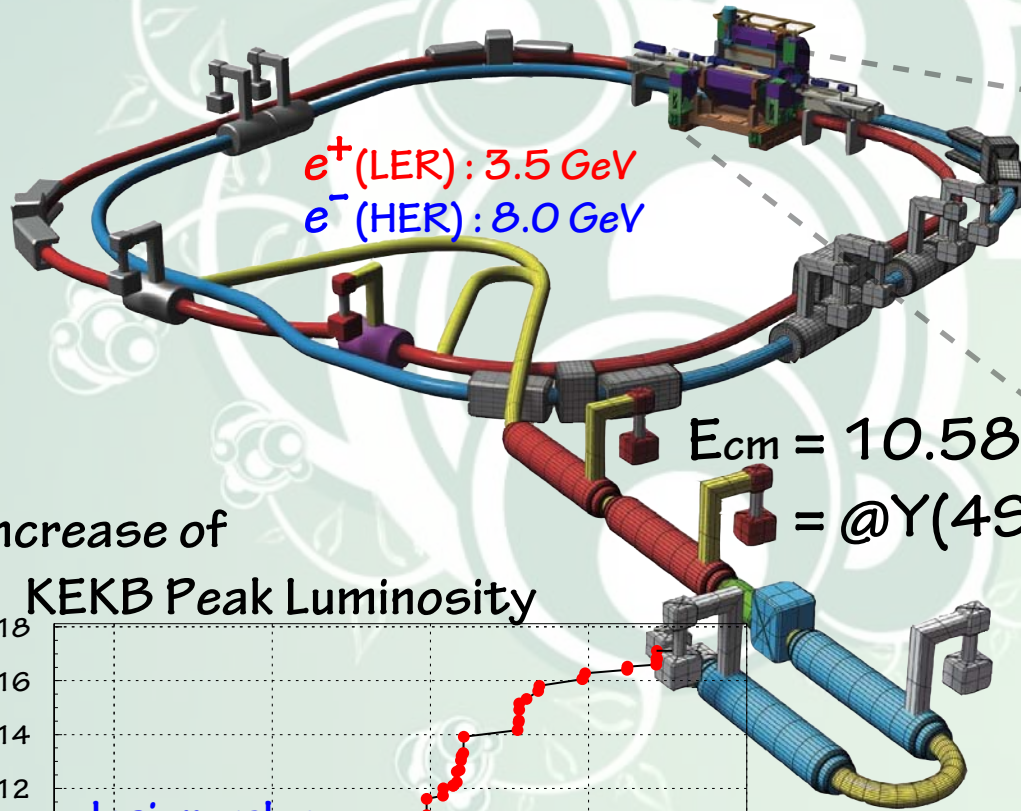


Key objective of B-factories :

Measurements of the $\begin{cases} 3 \text{ sides} \\ 3 \text{ angles} \end{cases}$

Vast numbers of B-mesons !!

KEKB Accelerator + Belle Detector



~13 nations
 ~400 persons

tracking/vertexing

1.5T Solenoid

Drift Chamber

e/γ detection

CsI calorimeter

Silicon Strip Detector

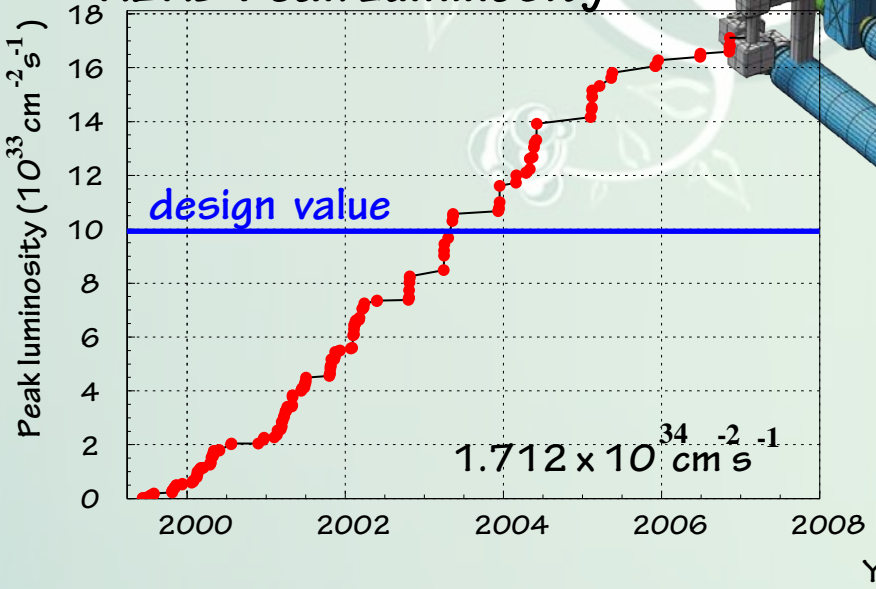
$\pi/K/p$ identification

Cherenkov detector

μ/K_L identification

Resistive Plate Chamber

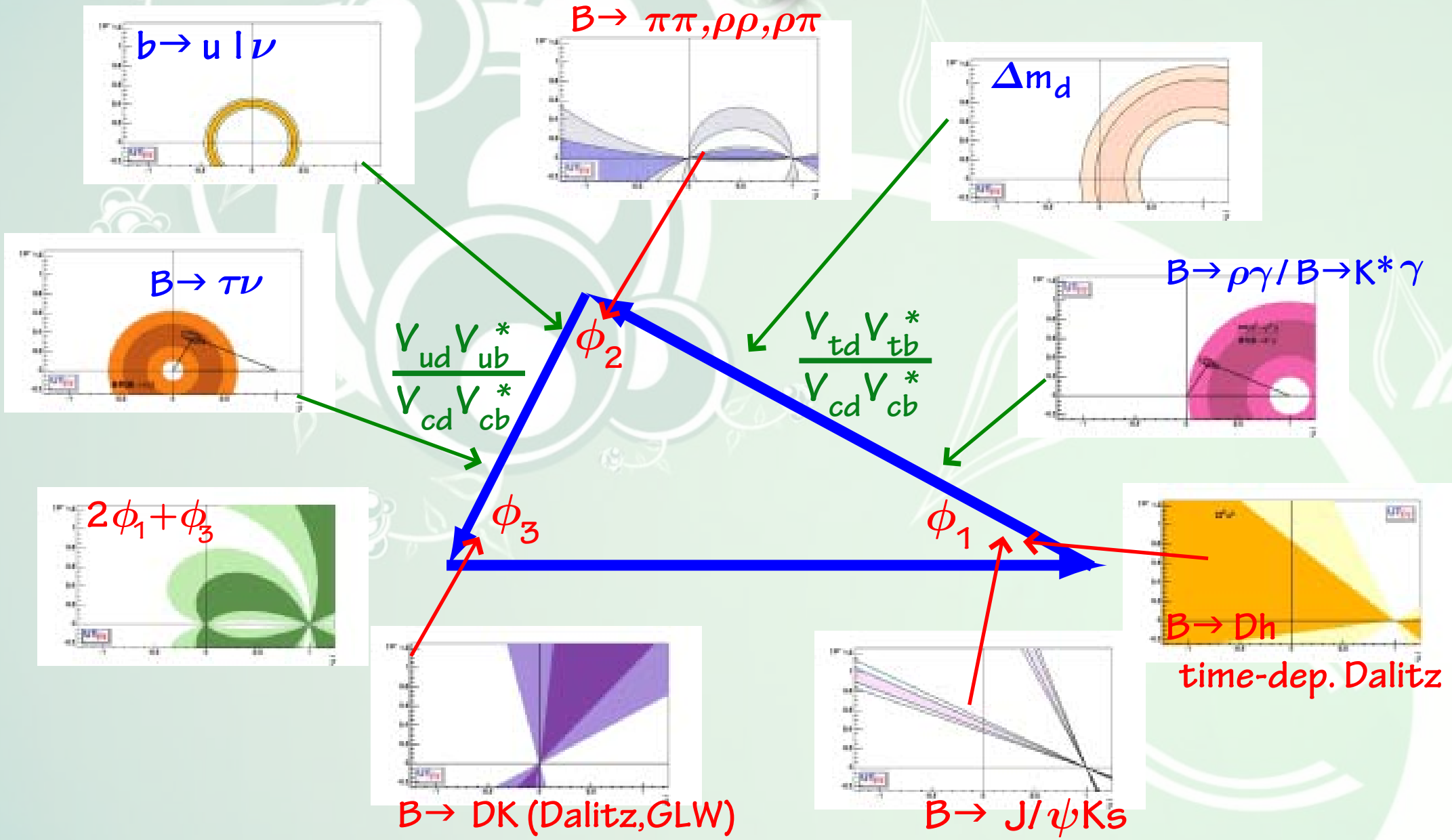
Increase of
 KEKB Peak Luminosity



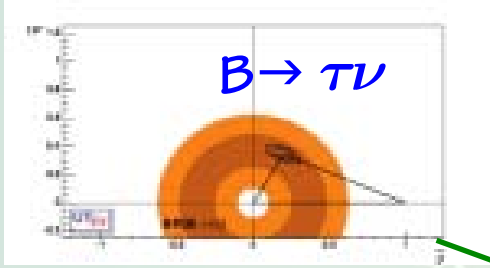
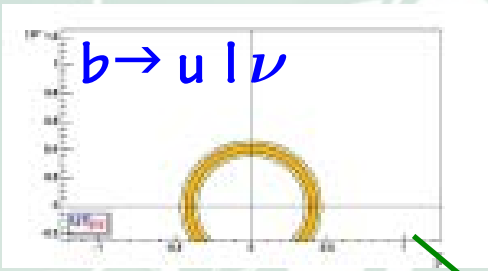
Total $\sim 0.9 \text{ ab}^{-1}$
 (as of 2009/Mar/30)

Experiment runs through 2009 fiscal year

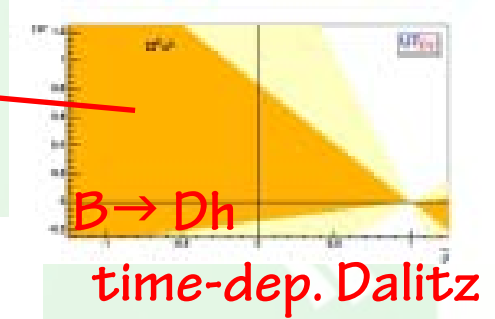
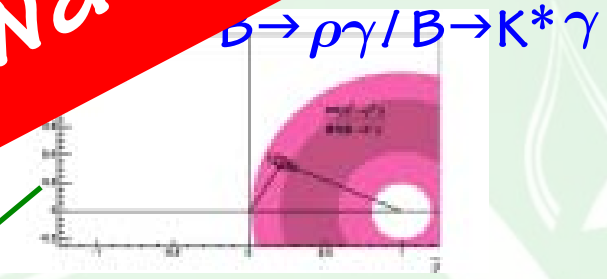
Remarkable Physics Results



Remarkable Physics Results



$$\frac{V_{ud} V_{ub}^*}{V_{cd} V_{cb}^*}$$



Confirm Kobayashi-Maskawa theory!

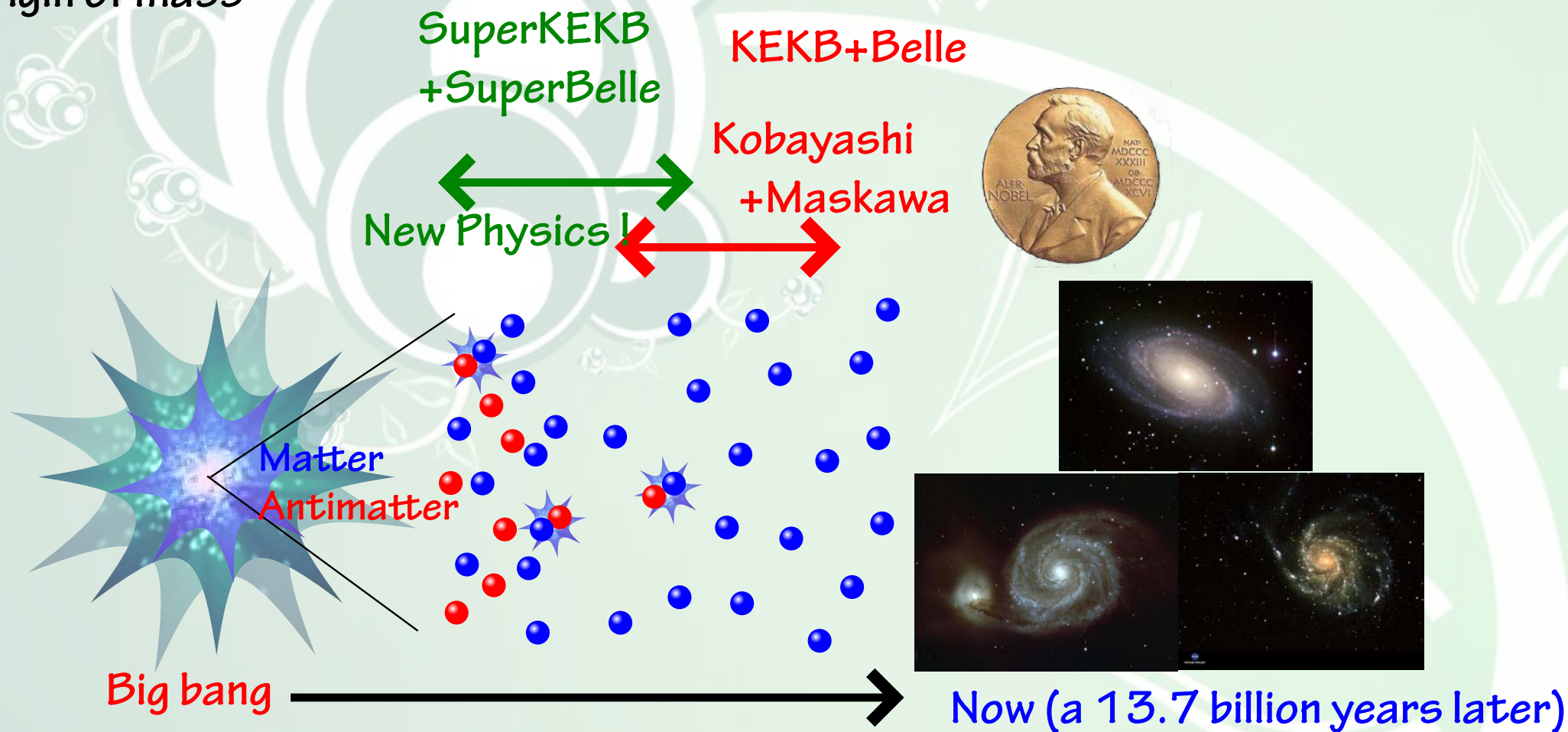
ϕ_2

ϕ_1

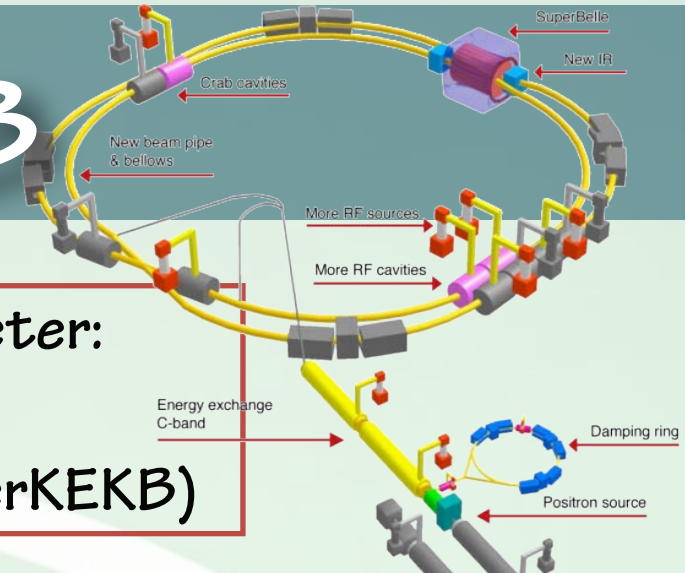
B-factory Next Generation

Matter excess of Universe
Dark matter
Origin of mass

In order to see CP violation,
in the early stage of Universe



SuperKEKB



Stored current:

1.7 / 1.4 A (e^+ / e^- KEKB)

→ **9.4 / 4.1 A** (SuperKEKB)

Beam-beam parameter:

0.059 (KEKB)

→ **>0.24** (SuperKEKB)

$$L = \frac{\gamma_{\pm}}{2e r_e} \left(1 + \frac{\sigma_y^*}{\sigma_x^*} \right) \frac{I_{\pm} \xi_{\pm y}}{\beta_y^*} \left(\frac{R_L}{R_y} \right)$$

Vertical β at the IP:

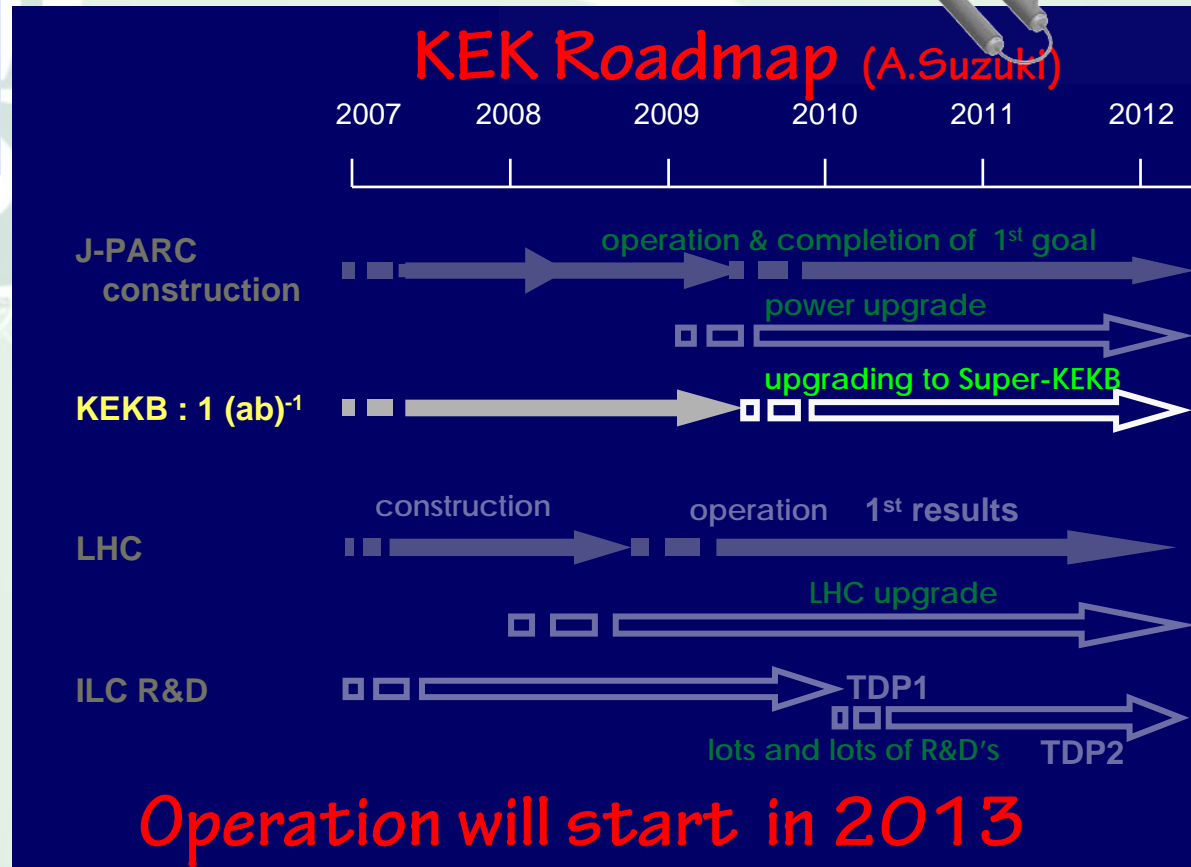
6.5 / 5.9 mm (KEKB)

→ **3.0 / 3.0 mm** (SuperKEKB)

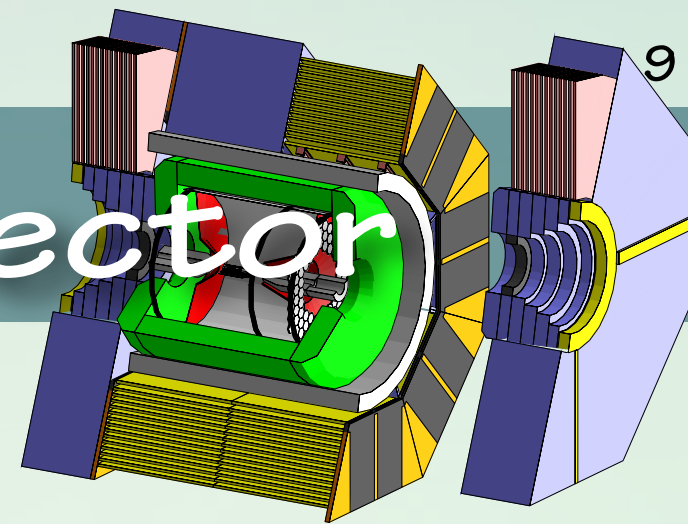
Luminosity:

$0.17 \times 10^{35} \text{ cm}^{-2} \text{ s}^{-1}$ (KEKB)

$8 \times 10^{35} \text{ cm}^{-2} \text{ s}^{-1}$ (SuperKEKB)



SuperBelle Detector



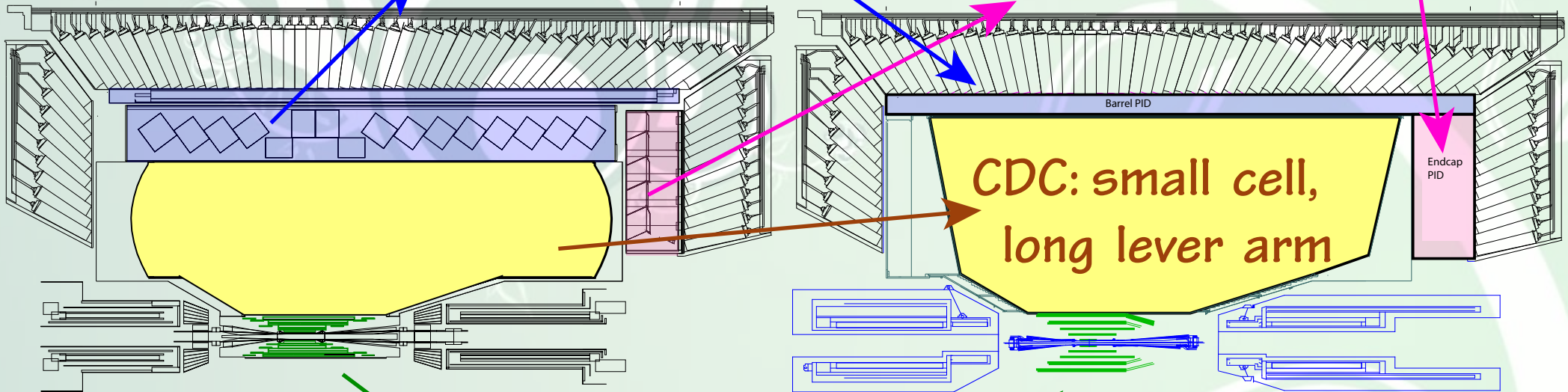
Beam BG tolerant
low material budget
better performance

Belle

SuperBelle

Barrel: TOF+ACC → TOP

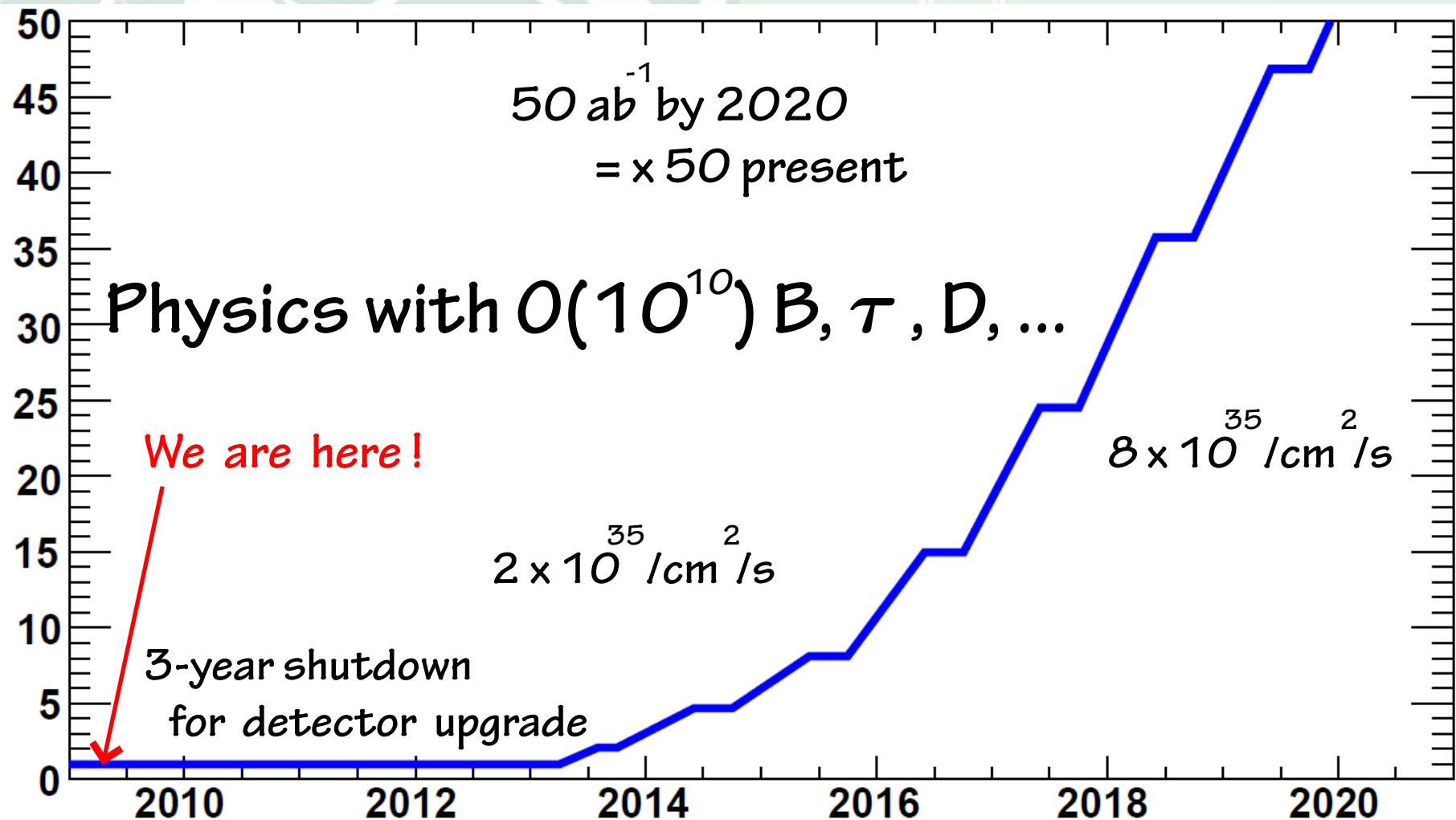
EndCAP: EACC → A-RICH



SVD: 4 lyr → 2-lyr PXD + 4-lyr DSSD

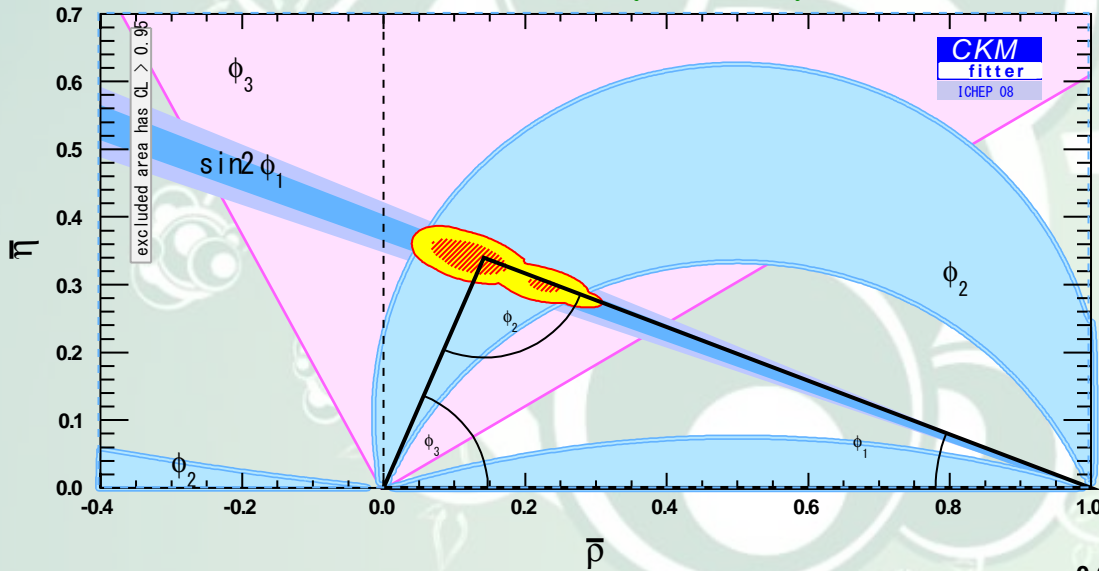
ECL: waveform sampling, pure CsI for end-caps
KLM: RPC → Scintillator + SiPM (end-caps)

Luminosity Prospect



Physics @ $50ab^{-1}$

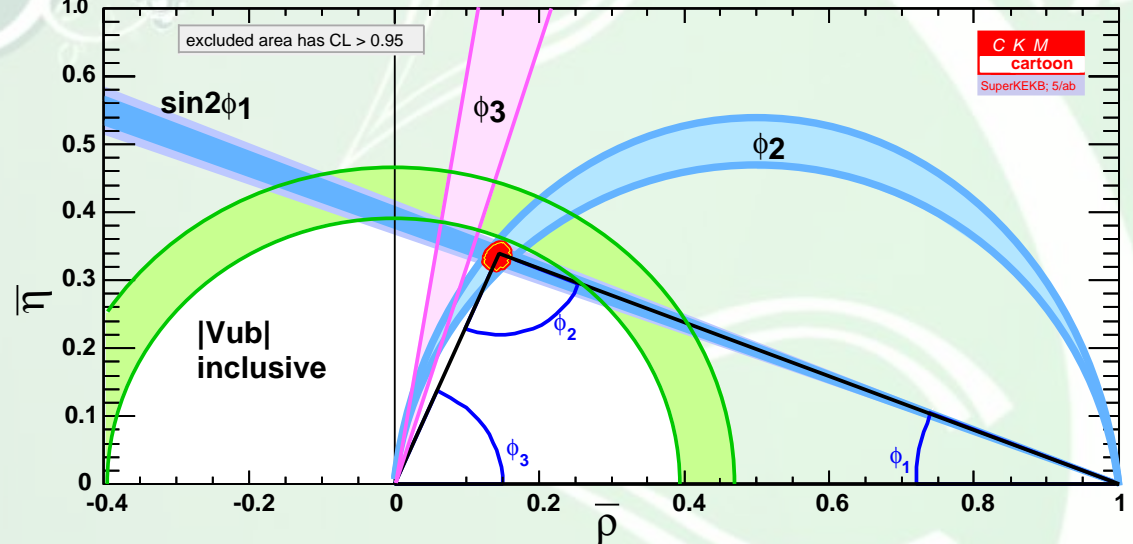
2008 ($1ab^{-1}$)



- . ϕ_1 is like a needle.
 - . ϕ_2, ϕ_3 allowed regions can be reduced drastically
- 1%-precision measurement
 → search for new CP phases (beyond the SM)

Other Features

- Charged Higgs
- LFV ($\tau \rightarrow \mu \gamma$)
- New resonances

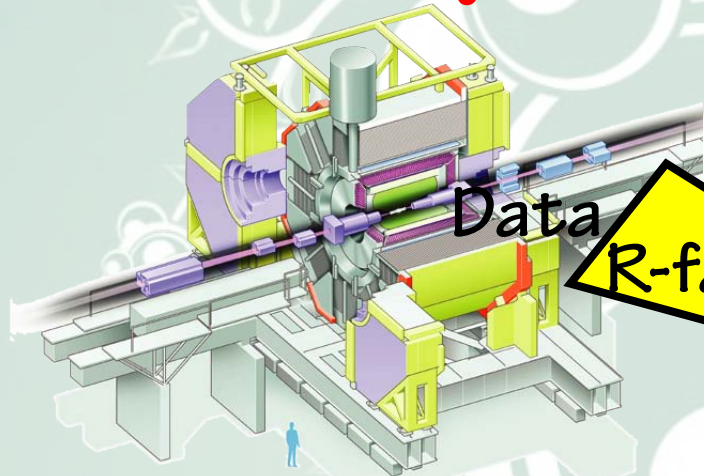


201X ($50ab^{-1}$)

Current Belle Computing

Belle Data Distribution

Only 1 ab^{-1} data for 10 years



Data

R-farm

PC-clusters

Nagoya U., VPI, Korea
Vienna, U. Tokyo, LPHE, ...

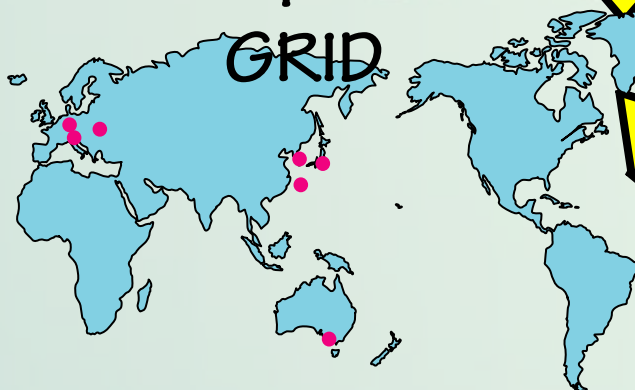


MC-prod.



MC-prod.

GRID

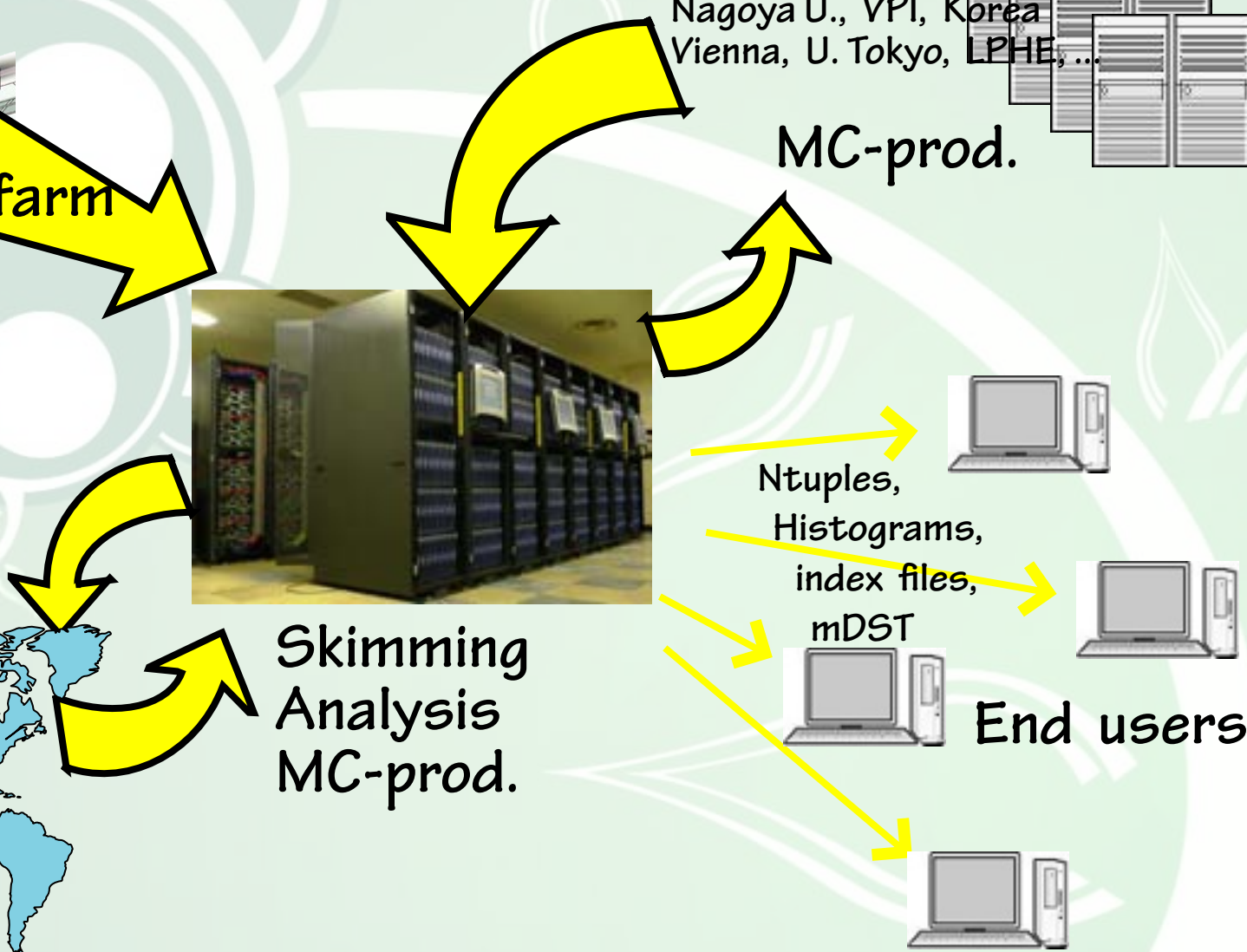


Skimming
Analysis
MC-prod.

Ntuples,
Histograms,
index files,
mDST



End users



Current Belle Computing

B computer is upgraded in Mar. 2009.

Year (contract) Performance	1999- (4years)	2001- (5years)	2006- (6years)	2009- (continued)
CPU [SI2k]	~100 (WS)	~1200 (WS+PC)	~42500 (PC)	~115200 (PC)
Disk [TB]	4	9	1000	1500
Tape [TB]	160	620	3500	3500
Workgroup server [#hosts]	3+9	11	80 (+16 file servers)	80 (+26 file servers)
User work- station [#hosts]	28WS + 68X	23WS + 100PC	128PC	128PC



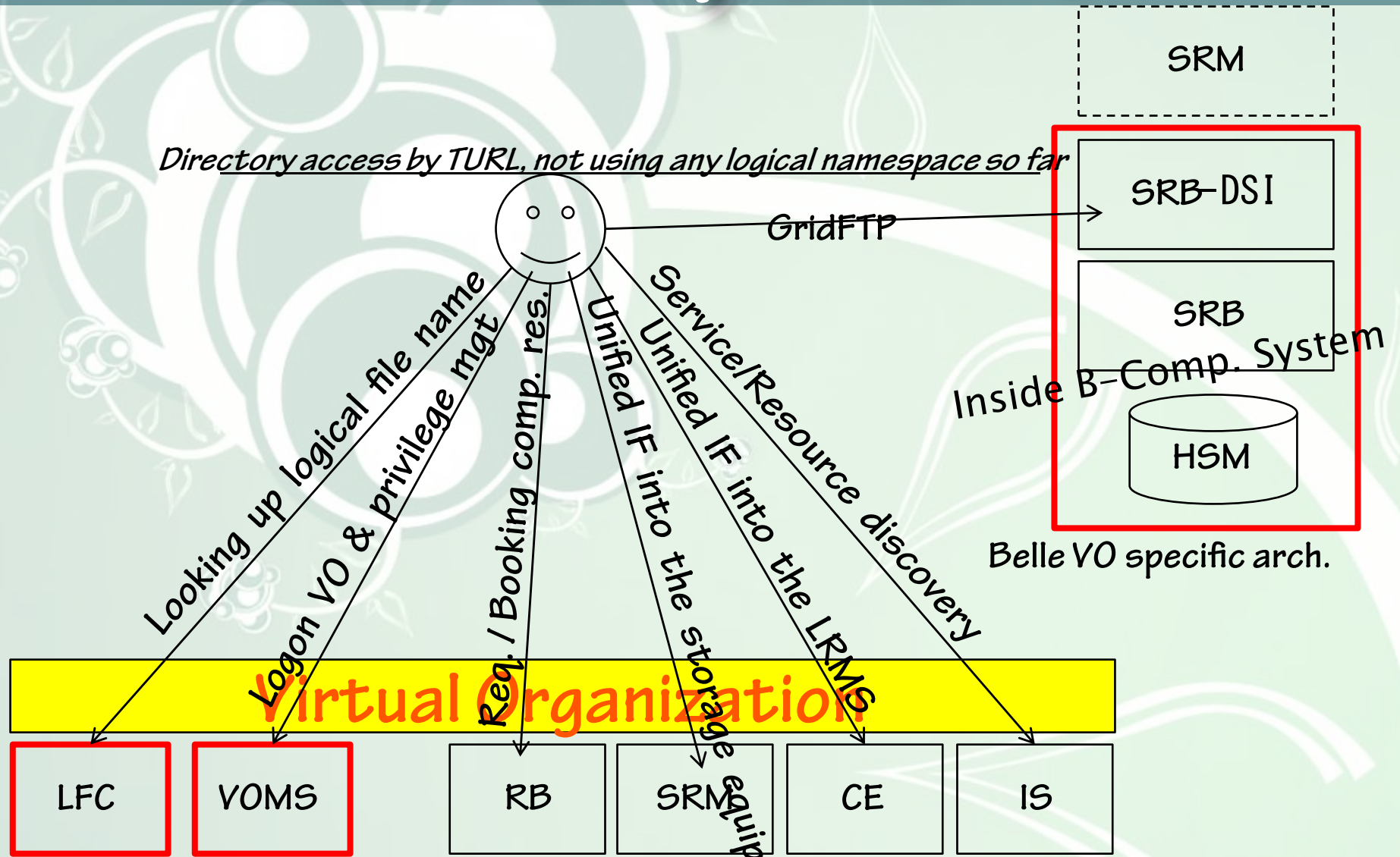


Net One Systems
DELL™
SONY



Net One Systems
DELL™
SONY

VO Component



LCG Deployment @ KEK

JP-KEK-CRC-01 (KEK-01)


- ▶ Production in GOC since Nov 2005
- ▶ Mainly operated by KEK staffs
- ▶ Site Role:
 - Practical operation for KEK-2
 - Getting' started for university groups
- ▶ Resource and Component:
 - SL- 3x or SL - 4x
 - gLite - 3.X
 - CPU: 14
 - Storage: ~7TB for disk and DSI for HSM, HPSS
 - Fully functional services
- ▶ Supported VO:
 - belle apdg ail g4med dteam ops ppj ilc calice
 - naokek

*In plan to upgrade
In Q1 of FY2009*

JP-KEK-CRC-01 (KEK-02)

- ▶ Production in GOC since early 2006
- ▶ Site operation:
 - Manabu and Kohki
- ▶ Site Role:
 - More stable services based on KEK-1 experiences.
- ▶ Resource and Component:
 - SL- 3x or SL - 4x
 - gLite - 3.X
 - CPU: 48
 - Storage: ~1TB for disk
 - Fully functional services
- ▶ Supported VO:
 - belle apdg ail g4med dteam ops ppj ilc calice
 - naokek

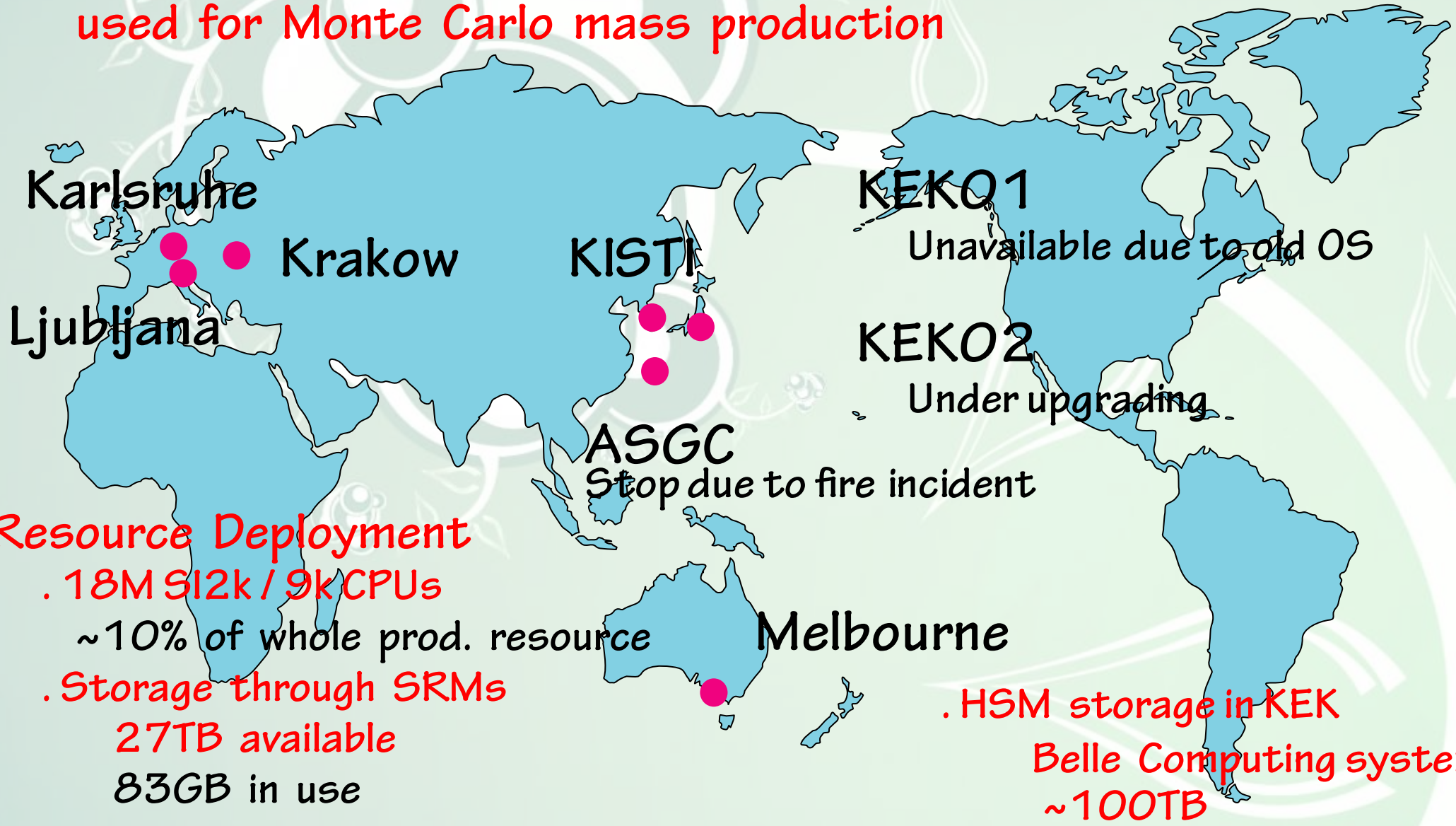
Fully upgrade and up soon!



- 10WNs x 8CPUs x ~4kSI2K
- Storage capability on demand basis
 - HPSS virtually works as the backend disk of SE
 - ~200USD/1TB
- VM (Xen) technologies are widely supported in whole site
 - Higher availability & more robustness
- Old blade servers (B - Comp) are now being integrated with KEK-2
 - 250 x 2CPUs

Belle VO Sites

used for Monte Carlo mass production



Resource Deployment

. 18M S12k / 9k CPUs

~10% of whole prod. resource

. Storage through SRMs

27TB available

83GB in use

. HSM storage in KEK

Belle Computing system

~100TB

Belle Data Distribution

Only 1 ab^{-1} data for 10 years

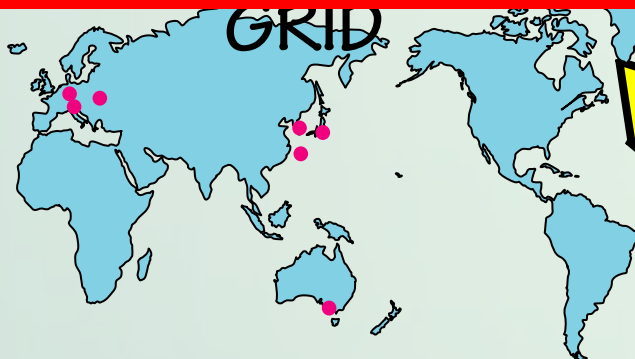
PC-clusters

Nagoya U., VPI, Korea
Vienna, U. Tokyo, LPHE, ...

Works very well, so far !!

However,

What happens in the SuperBelle ??



GRID

Skimming
Analysis
MC-prod.

End users



SuperBelle Computing Model



Expected Event Rate

Early stage (a couple of years): $2 \times 10^{35} \text{ cm}^{-2} \text{ s}^{-1}$

$\int L = 4 \text{ ab}^{-1} / \text{yr}$ (including run-time factor: $\sim 2/3$ yrs)

$B\bar{B}$: $4 \text{ ab}^{-1} / \text{yr} \times 1 \text{ nb} = 4 \times 10^9 \text{ events/yr}$

$B\bar{B}, \text{udsc}, \tau\tau, \text{calib} = \sim 20 \times 10^9 \text{ events/yr}$

$$\sigma(B\bar{B}) = \sim 1 \text{ nb}$$

$$\sigma(\text{udsc}) = \sim 3 \times \sigma(B\bar{B})$$

$$\sigma(\tau\tau) = \sigma(B\bar{B})$$

c.f.1 Current Belle: $\int L = \sim 1 \text{ ab}^{-1} / 10 \text{ yrs}$

c.f.2 Atlas $\sim 2 \times 10^9 \text{ events/yr}$

After a couple of years

$\int L \geq 12 \text{ ab}^{-1} / \text{yr}$

$B\bar{B}, \text{udsc}, \tau\tau, \text{calib} \geq \sim 60 \times 10^9 \text{ events/yr}$

+ BG, Bhabha, mu-pair, etc.

$$\sigma(\text{habhabha}) = 44 \times \sigma(B\bar{B})$$

Expected Data Size

Raw Data Size

PXD	400 kB <i>under evaluation</i>
SVD	8 kB
CDC	18 kB
TOP	1-6 kB
ARICH	4-7 kB
ECL	12 kB
KLM	2-4 kB
TRG	unknown (not negligible)
Total	40-60 kB (w/o PXD)

~100kB (w/ PXD)???

c.f. Current Belle

Raw Data = 30 kB/evt

DST = 140 kB/evt

mDST = 15 kB/evt

MC: only mDST

empirically

30×10^6 events per 1 fb^{-1}

Assuming 50kB/evt

and same B event purity as now,

Raw Data size for 4 ab^{-1} requires

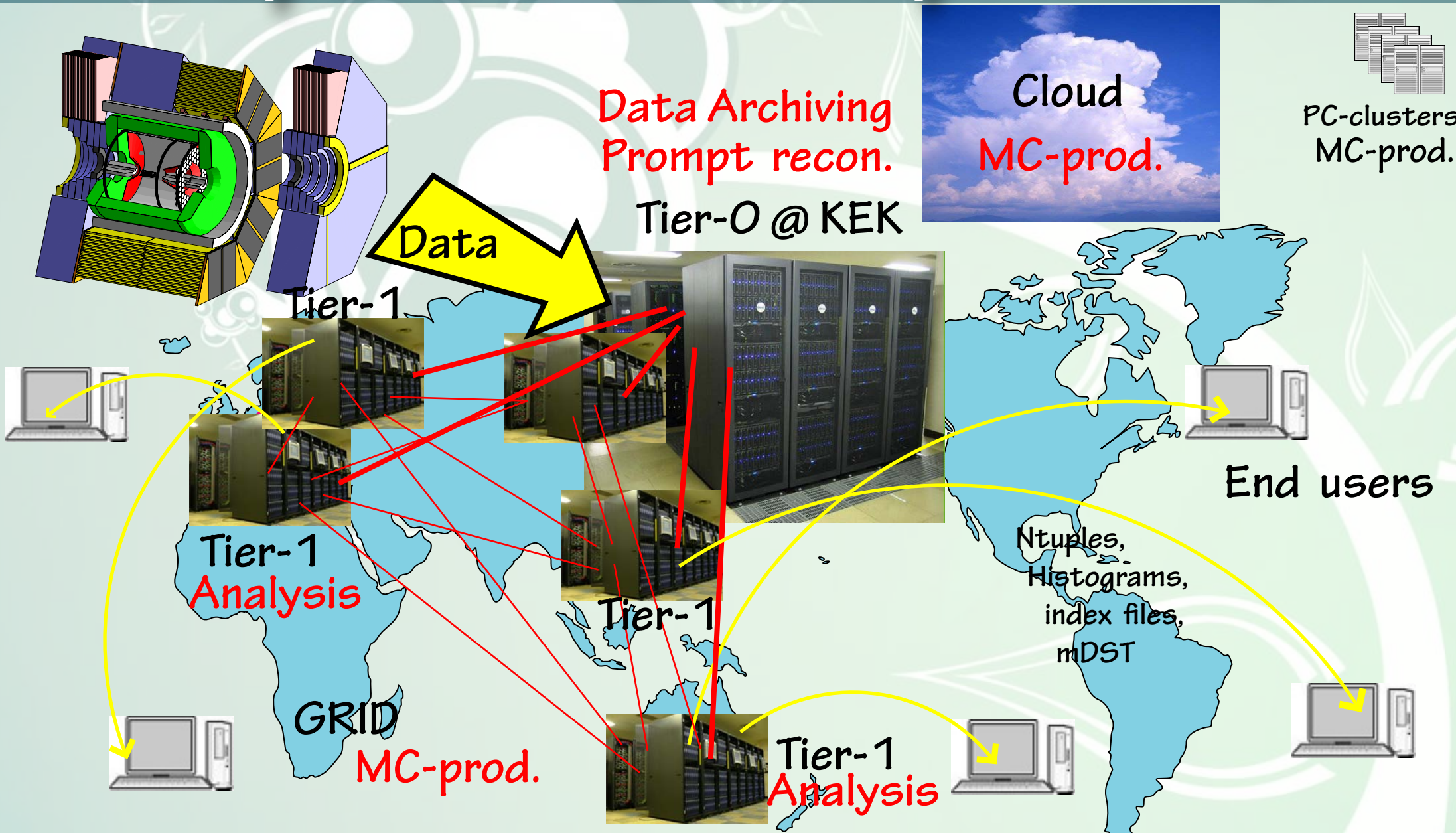
$\sim 6 \text{ PB/yr}$ storage

+ DST, mDST(Data/MC)

→ $\text{Totally } \sim 15 \text{ PB/yr}$

Besides, Luminosity will increase !!

SuperBelle Comp. Model



the Advent of the “Cloud”

We need **at least 3 x data or more MC** events for analyses
It **consumes a large fraction of HEP CPUs**
HEP CPUs is **at peak demand before international conferences**
Running/maintenance costs
to keep the HEP CPUs is not negligible

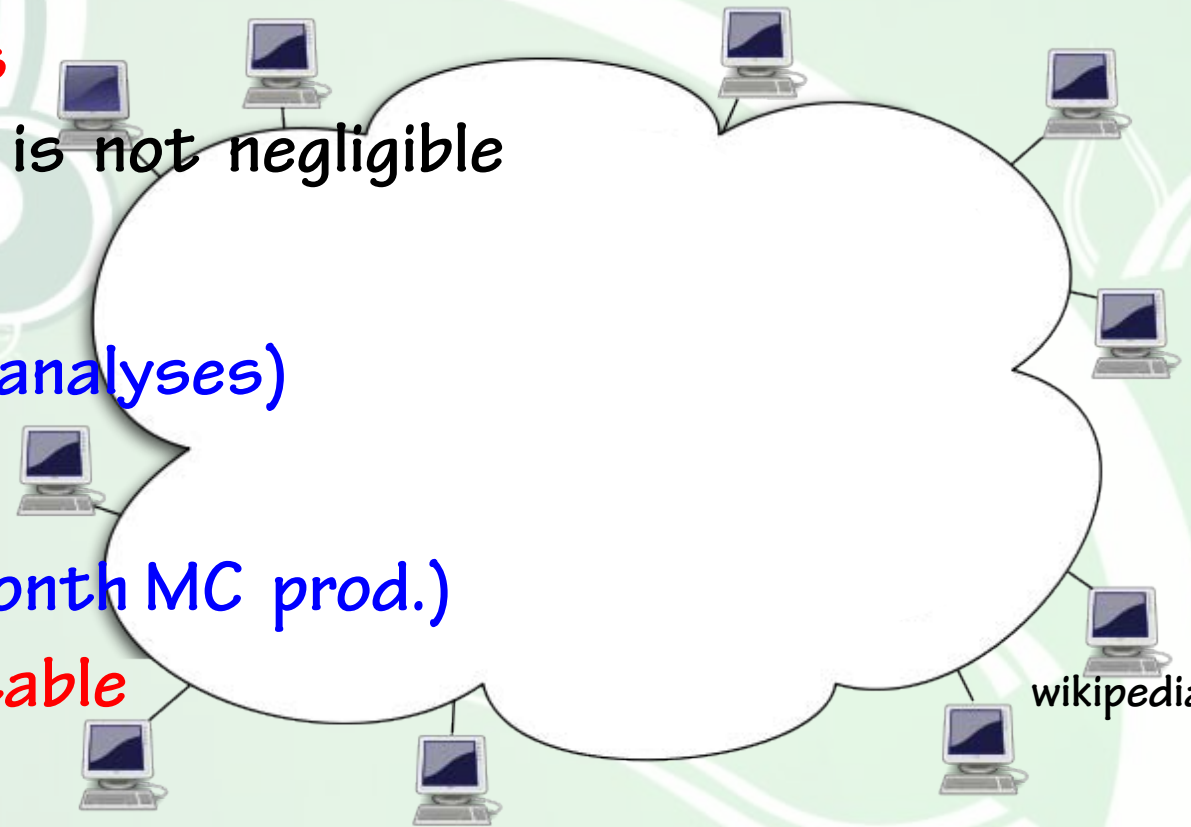
Constant use (e.g. recon., analyses)

GRID is suitable

Temporal use (e.g. a few month MC prod.)

Comercial Cloud is suitable

- . Pay per use
- . Flexible resource deployment



wikipedia

the Advent of the "Cloud"

We need **at least 3 x data or more MC** events for analyses
 It **consumes a large fraction of HEP CPUs**
 HEP CPUs is **at peak demand before international conferences**

Running/maintenance costs

to keep the HEP CPUs is

Constant use

GRID is

Temporal use

Comercial C

We have started the tests on Amazon EC2
please check Martin's talk on Apr. 21st

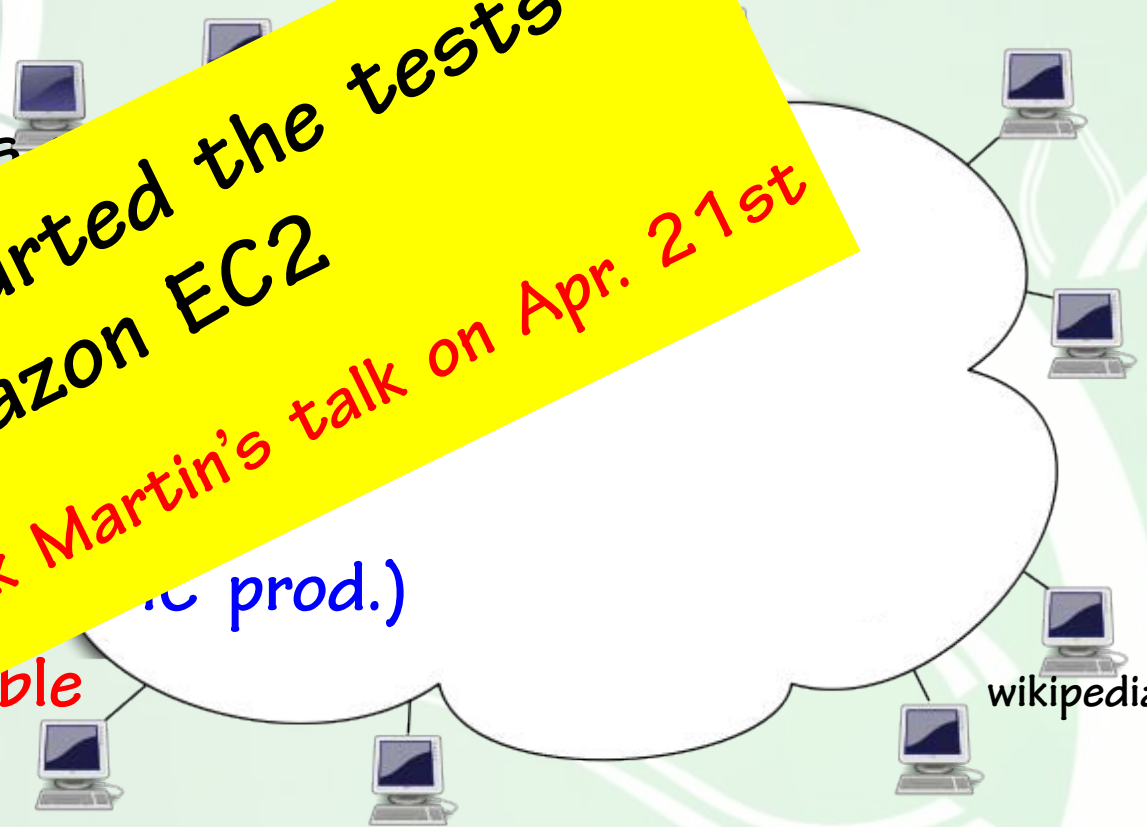
(c prod.)

able

. Pay per

. Flexible resource deployment

wikipedia



Data Handling

- . to **Set up and maintain a storage system** for raw, calibration, dst, skimmed data, index files, MC and ntuples at KEK
- . to **Develop a system for file storage at remote sites**
- . to **Implement a global data catalog** that allows for an easy selection of datasets
- . to **Implement and maintain a system for data transfer** to and from remote sites
- . to **Estimate tape and disk storage and bandwidth requirements** for the SuperBelle computing model

Proposal for DH-System

Basic Policies

Use concepts of metadata and projects

Keep system as simple as possible

Use grid (gLite) services as much as possible

- . No **SAM station** at each site,
but **just one central service**
(SAM: Sequential data Access via Metadata: *used in CDF and DO*)
- . No automatic transfer of files to jobs,
but **managed transfer of datasets**
- . **Submit jobs to sites where the data is**
instead of copying the data to the job

Proposal for DH-System

Terminology

CE : Compute Element

coordinates access to computing resources

SE : Storage Element

coordinates access to storage resources

LFC : LCG File Catalogue

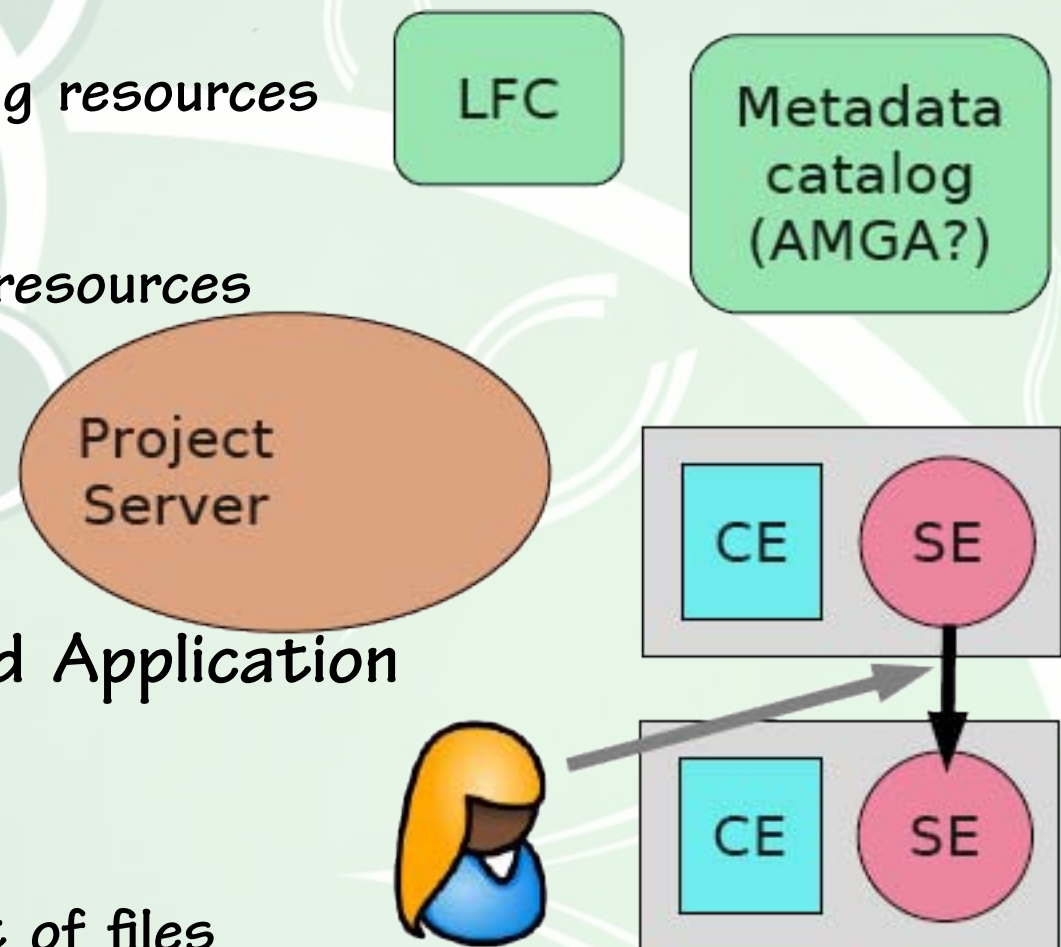
*keeps track of file locations
on the GRID*

AMGA : ARDA Metadata Grid Application

keeps track of data about files

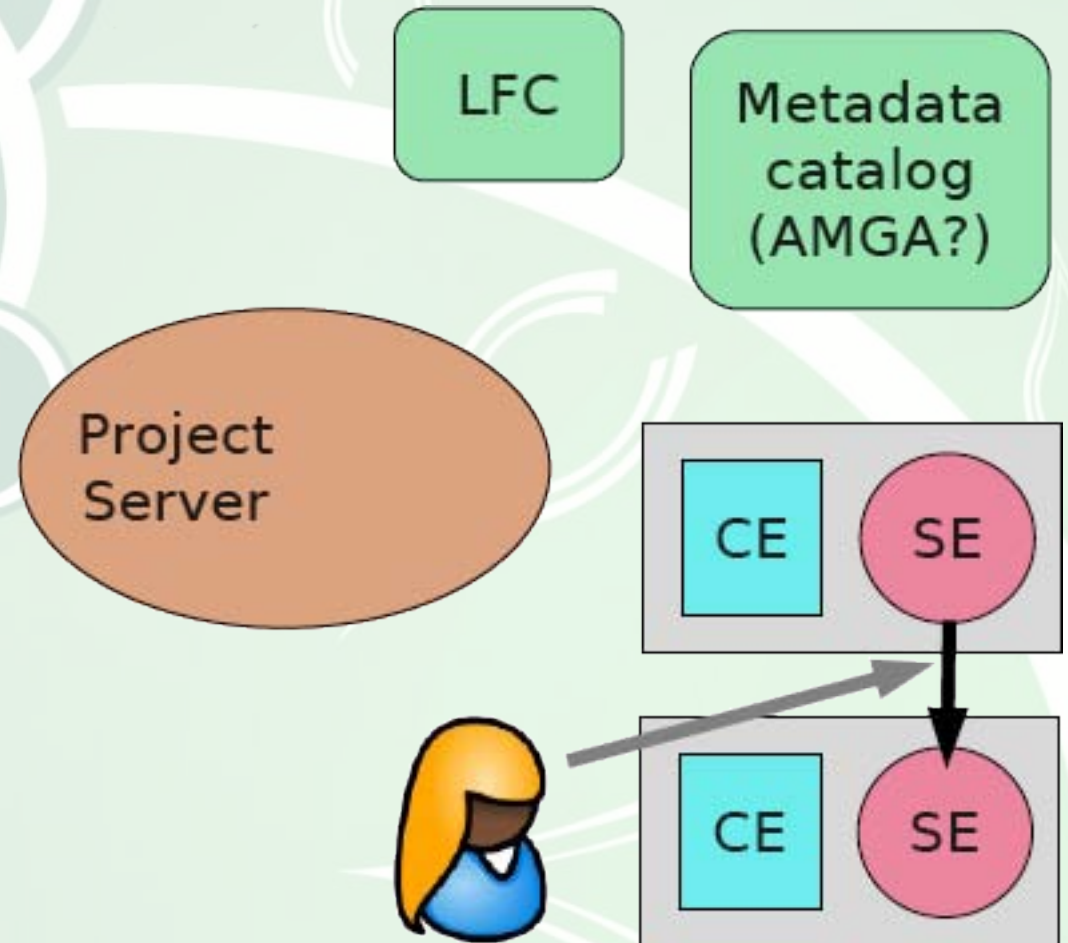
Project server

arranges the delivery of the set of files



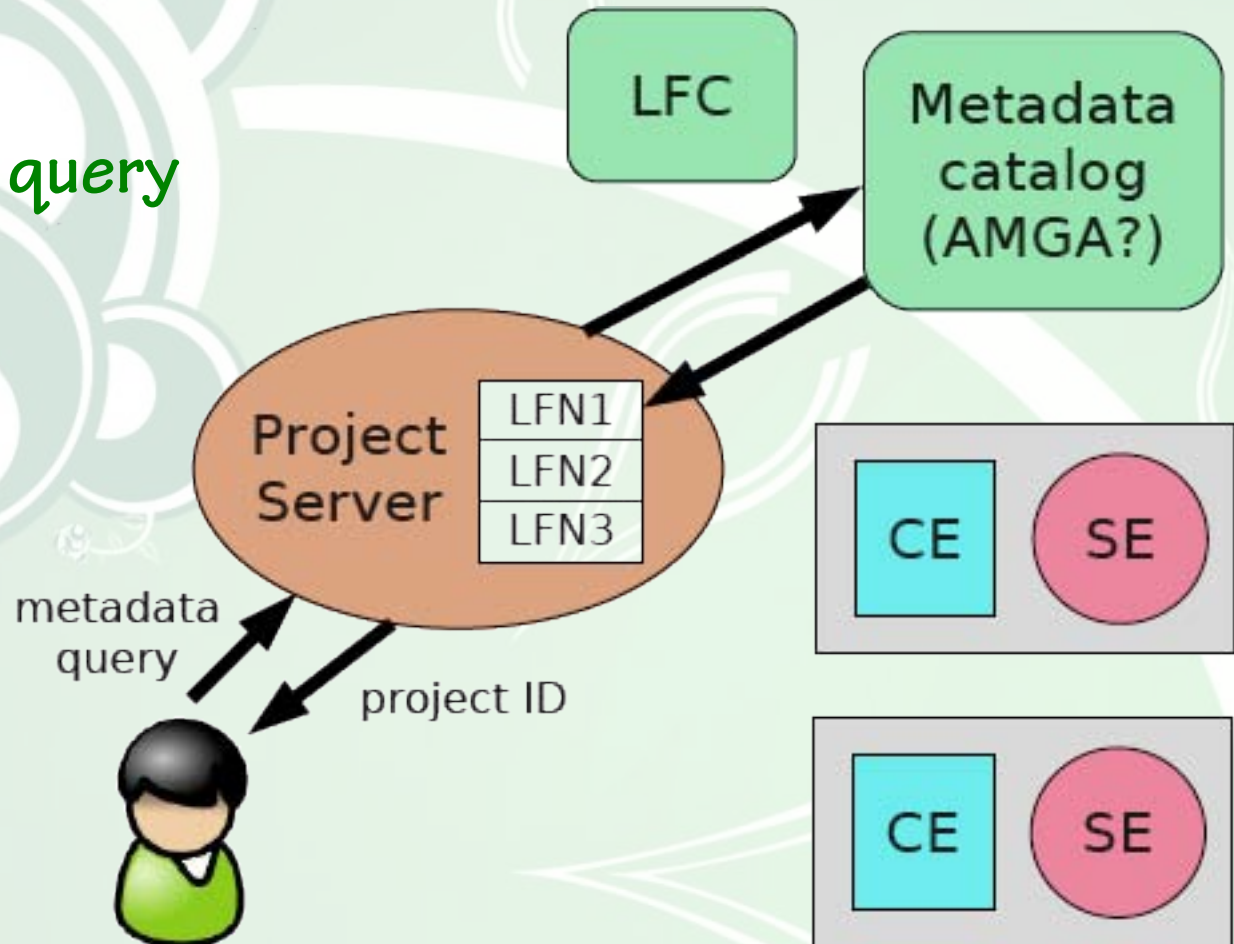
Proposal for DH-System

. Site manager transfers datasets



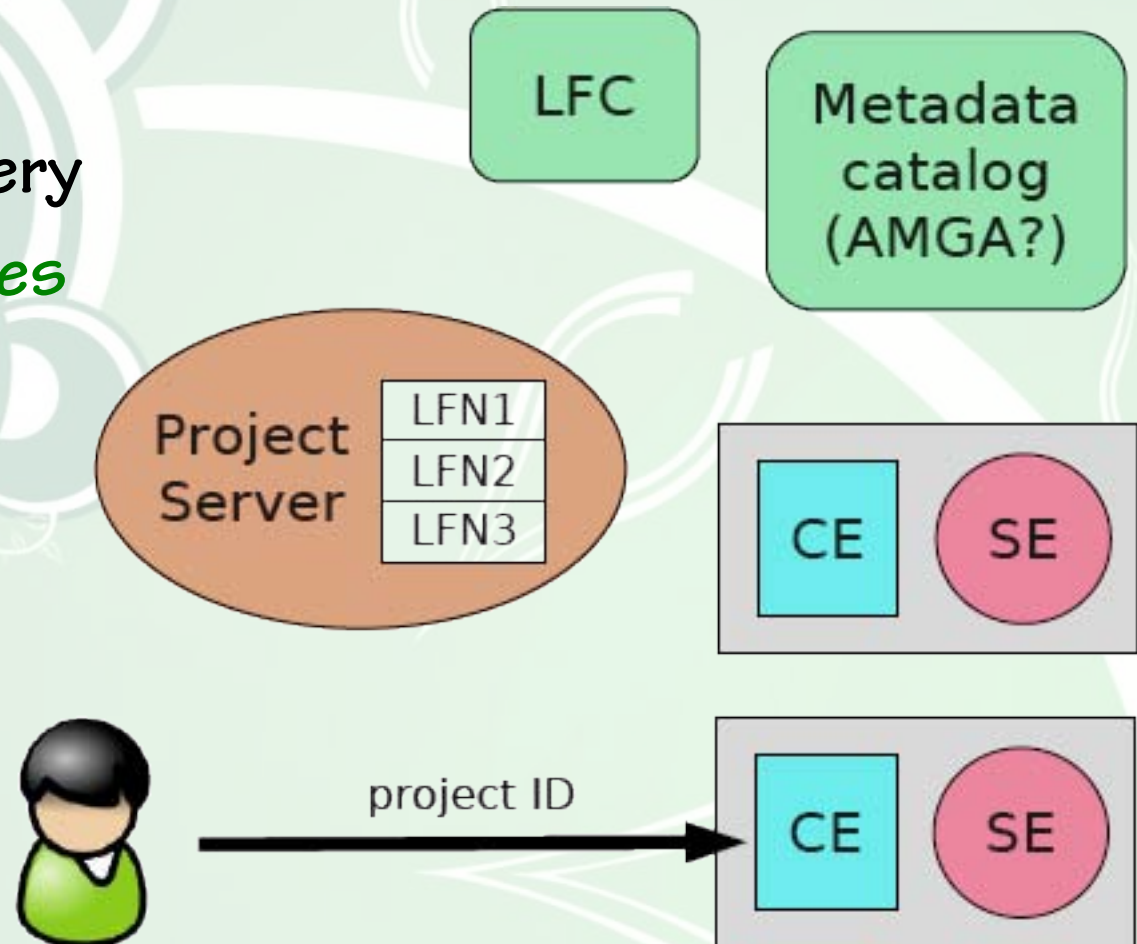
Proposal for DH-System

- . Site manager transfers datasets
- . User starts project using metadata query



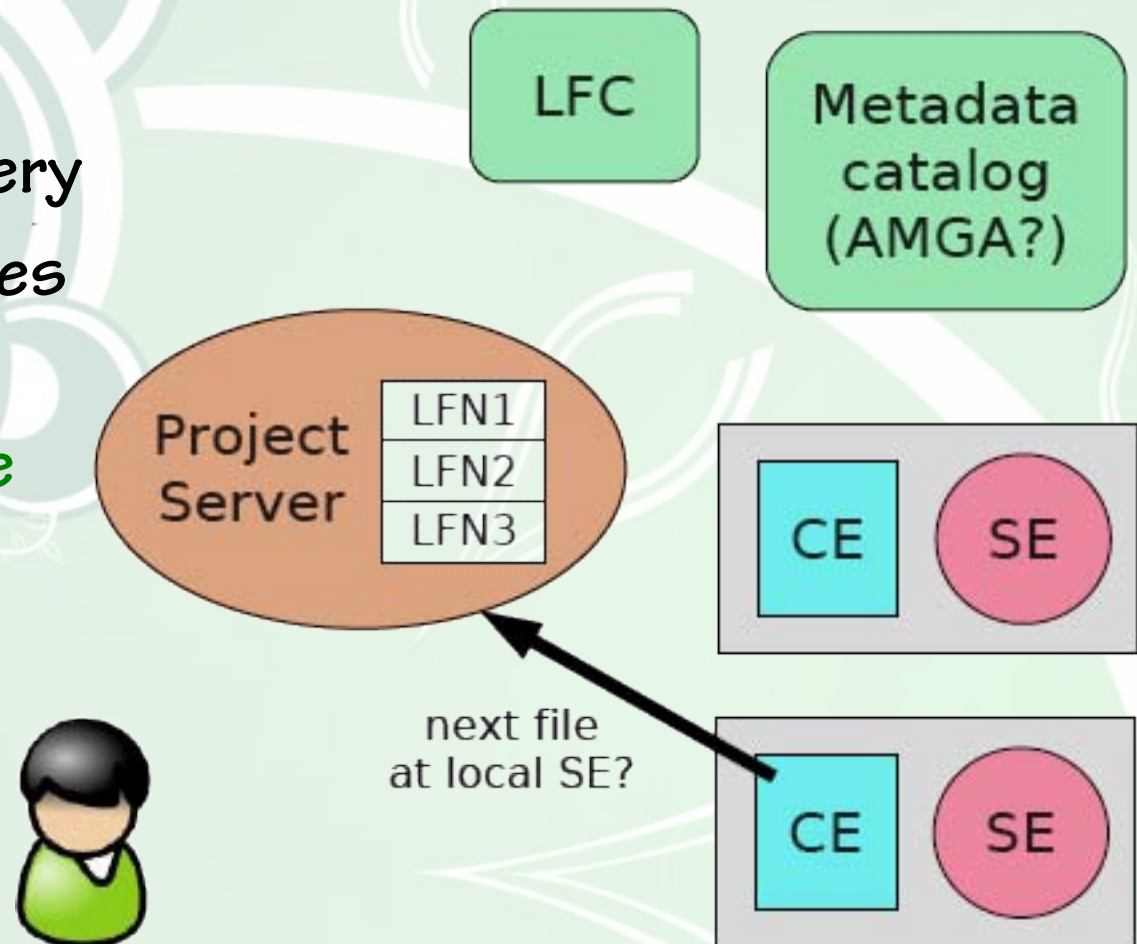
Proposal for DH-System

- . Site manager transfers datasets
- . User starts project using metadata query
- . User submits jobs to sites where the data is



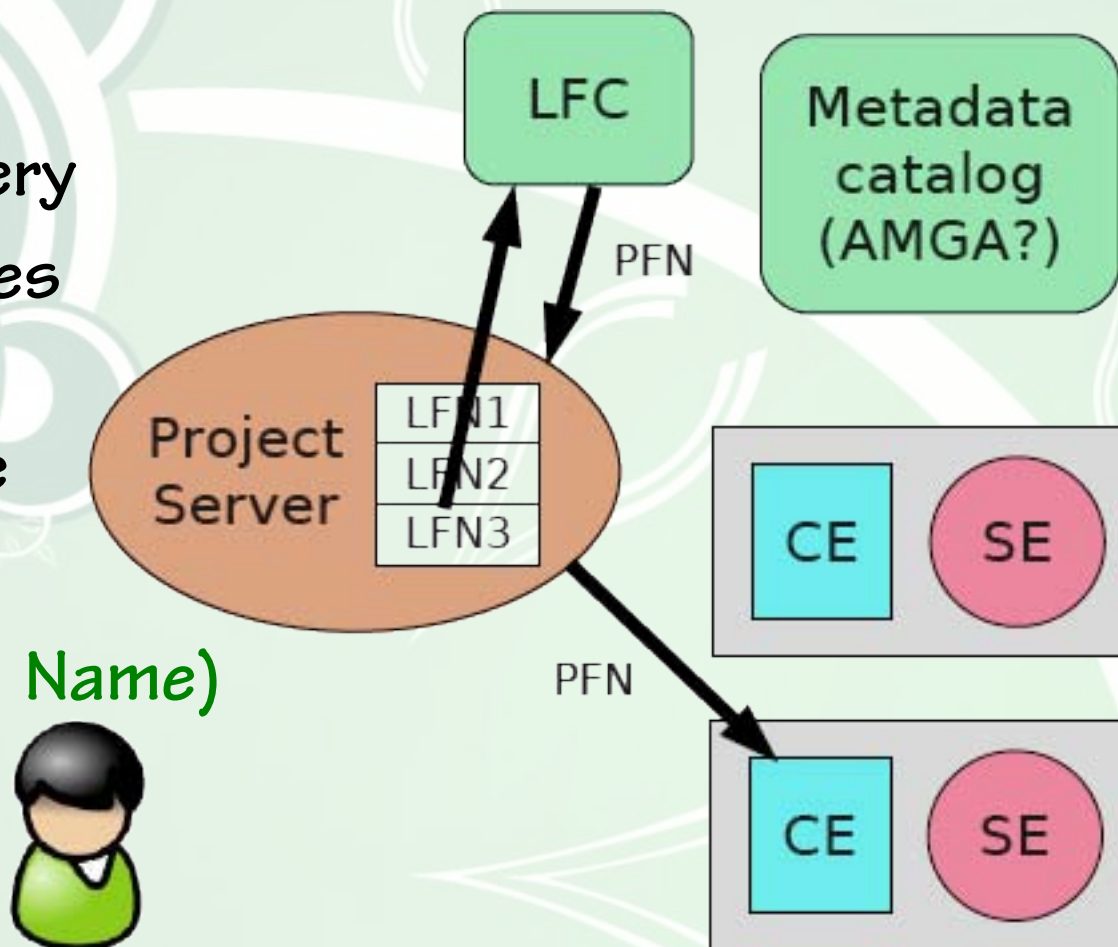
Proposal for DH-System

- . Site manager transfers datasets
- . User starts project using metadata query
- . User submits jobs to sites where the data is
- . **Job asks for the next file**



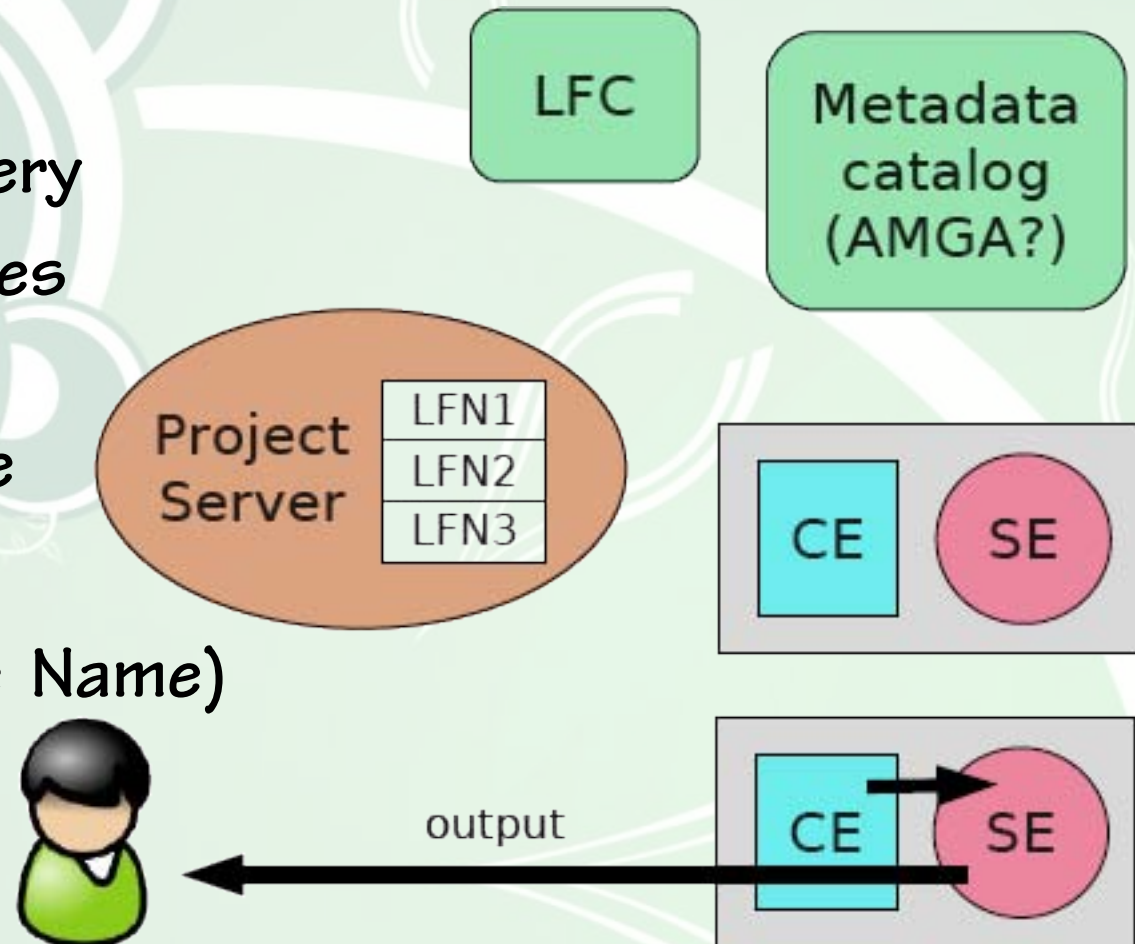
Proposal for DH-System

- . Site manager transfers datasets
- . User starts project using metadata query
- . User submits jobs to sites where the data is
- . Job asks for the next file
- . Project server returns next PFN (Physical File Name)



Proposal for DH-System

- . Site manager transfers datasets
- . User starts project using metadata query
- . User submits jobs to sites where the data is
- . Job asks for the next file
- . Project server returns next PFN (Physical File Name)
- . **Output stored at local SE**



Proposal for DH-System

. Site manager transfers datasets

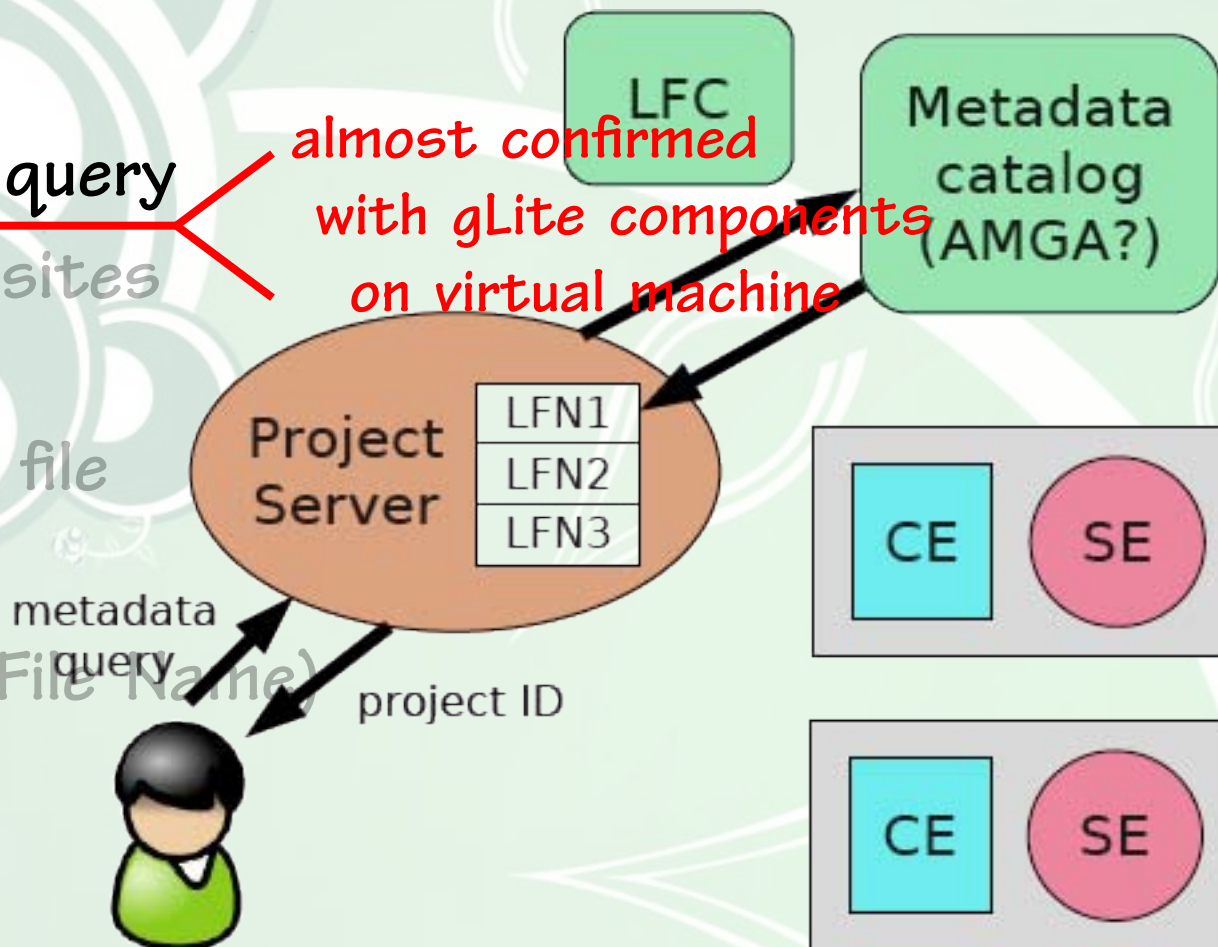
. User starts project
using metadata query

. User submits jobs to sites
where the data is

. Job asks for the next file

. Project server returns
next PFN (Physical File Name)

. Output stored
at local SE



Other issues

- ▶ **Computing model ?**
- ▶ **Analysis Framework**
- ▶ **Data persistency**
- ▶ **Data distribution**
- ▶ **Language (C++, Python)?**
- ▶ **Database**
- ▶ **Documentation**
- ▶ **web**
- ▶ **Release coordination**

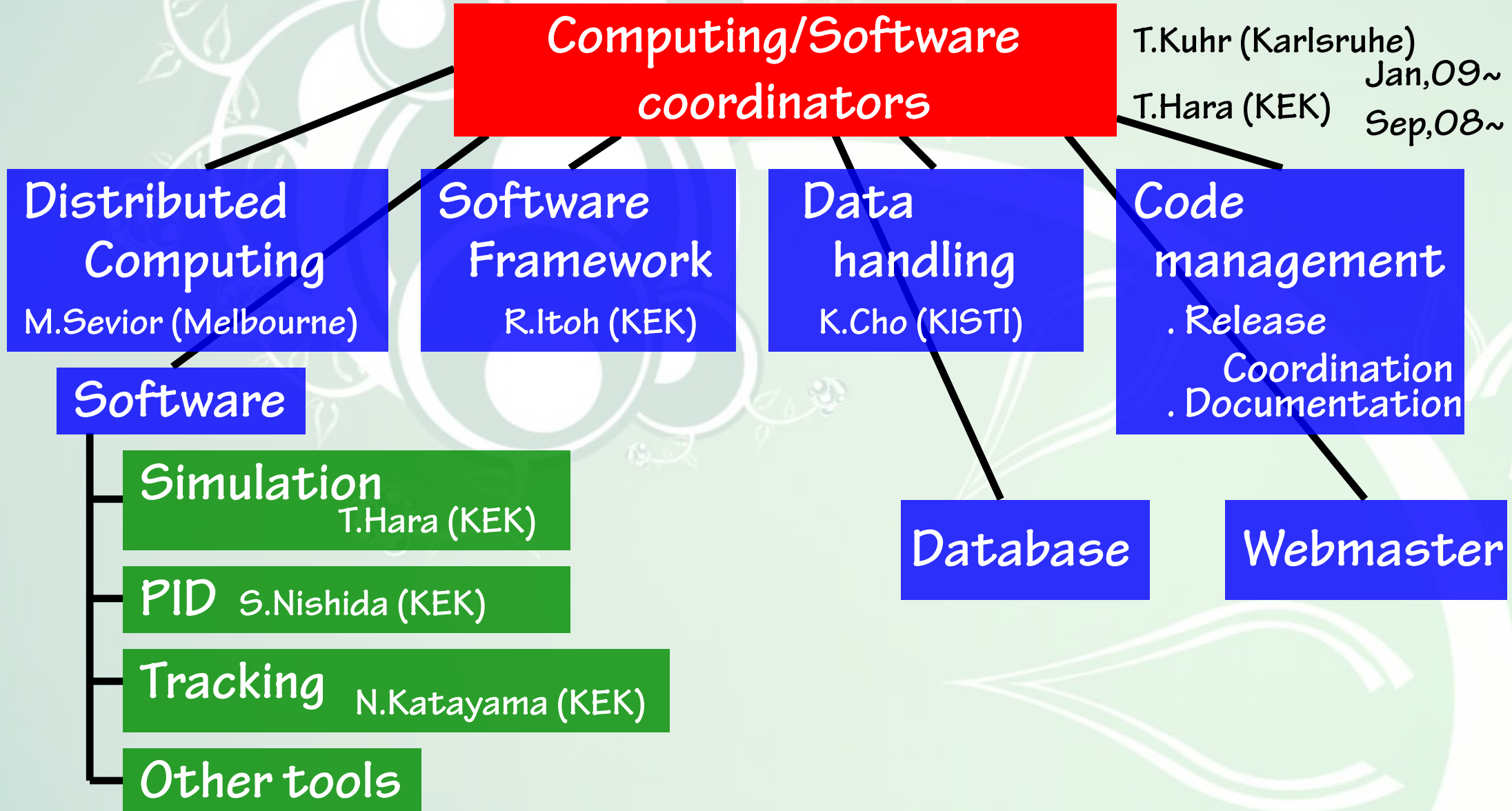
▶ **Software**

- ▶ **Simulation**
- ▶ **Tracking**
- ▶ **PID**
- ▶ **Other tools for analysis**

⋮

Software/Computing group has many things to do ...

Organization is being formed



Summary

KEKB/Belle

has accumulated $\sim 1 \text{ ab}^{-1}$ data

runs through 2009 fiscal year!

computing system works very well so far

Data center @ KEK : reconstruction, user analyses, MC prod....

PC clusters @ remote sites, GRID : MC prod.

SuperKEKB/SuperBelle

needs to tackle with 10~100 x present Belle data

needs a smart Distributed computing model

needs a smart Data handling system

database, maintenance, framework, data model, etc...

We are aiming for the 2013 start

(suppl. budget 5M\$ was allocated to start upgrade!)