



Belle II distributed computing with DIRAC

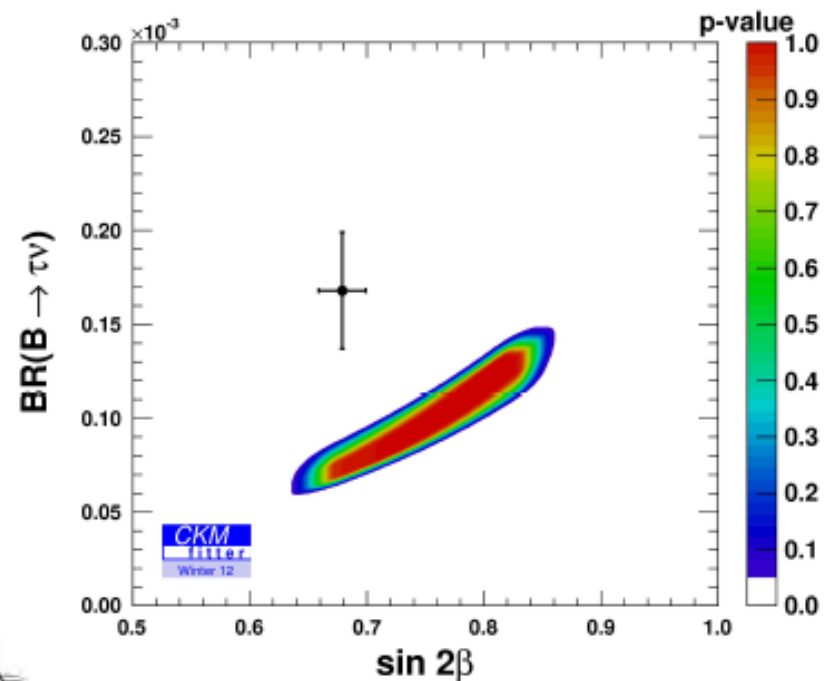
Hideki MIYAKE (KEK)

May 26, 2014 DIRAC user workshop

Goal of Belle II experiment

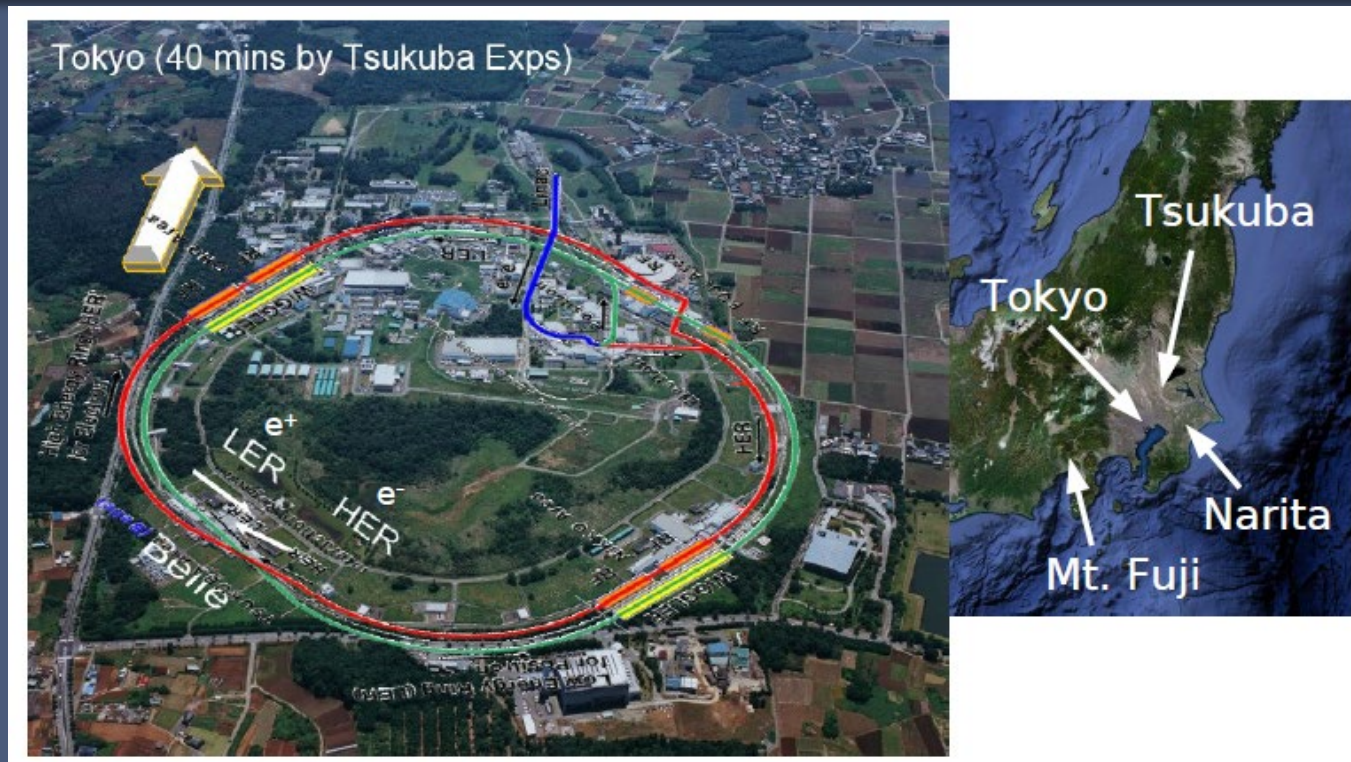
- ✓ Confirmation of KM mechanism of \mathcal{CP} in the Standard Model
- x \mathcal{CP} in the SM too small (by many orders of magnitude) to generate observed baryon asymmetry in the universe

- Need sources of \mathcal{CP} beyond the SM

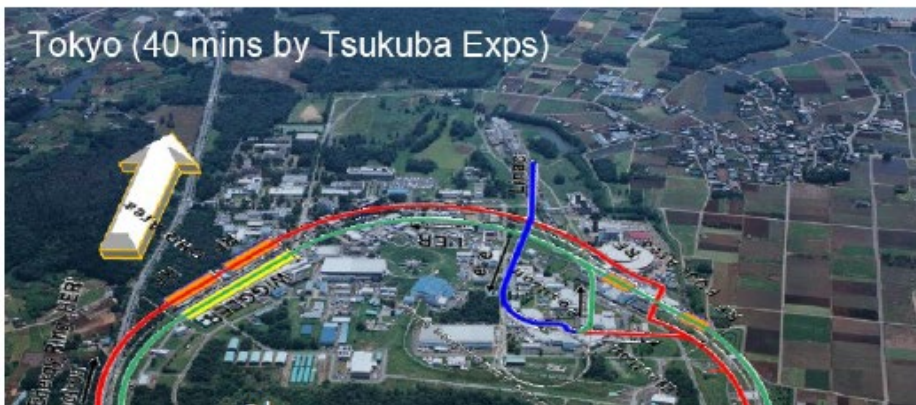


- Super B factory
Complementary to LHCb

Belle II experiment



Belle II experiment

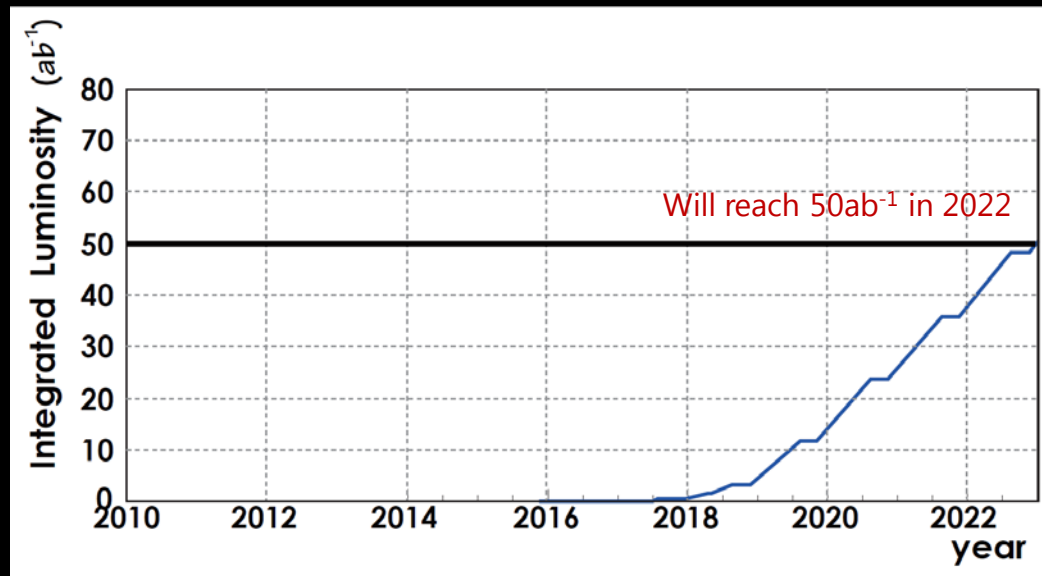


Challenging performance

x50 larger integrated luminosity than Belle

Comparable data rate to all LHC experiments

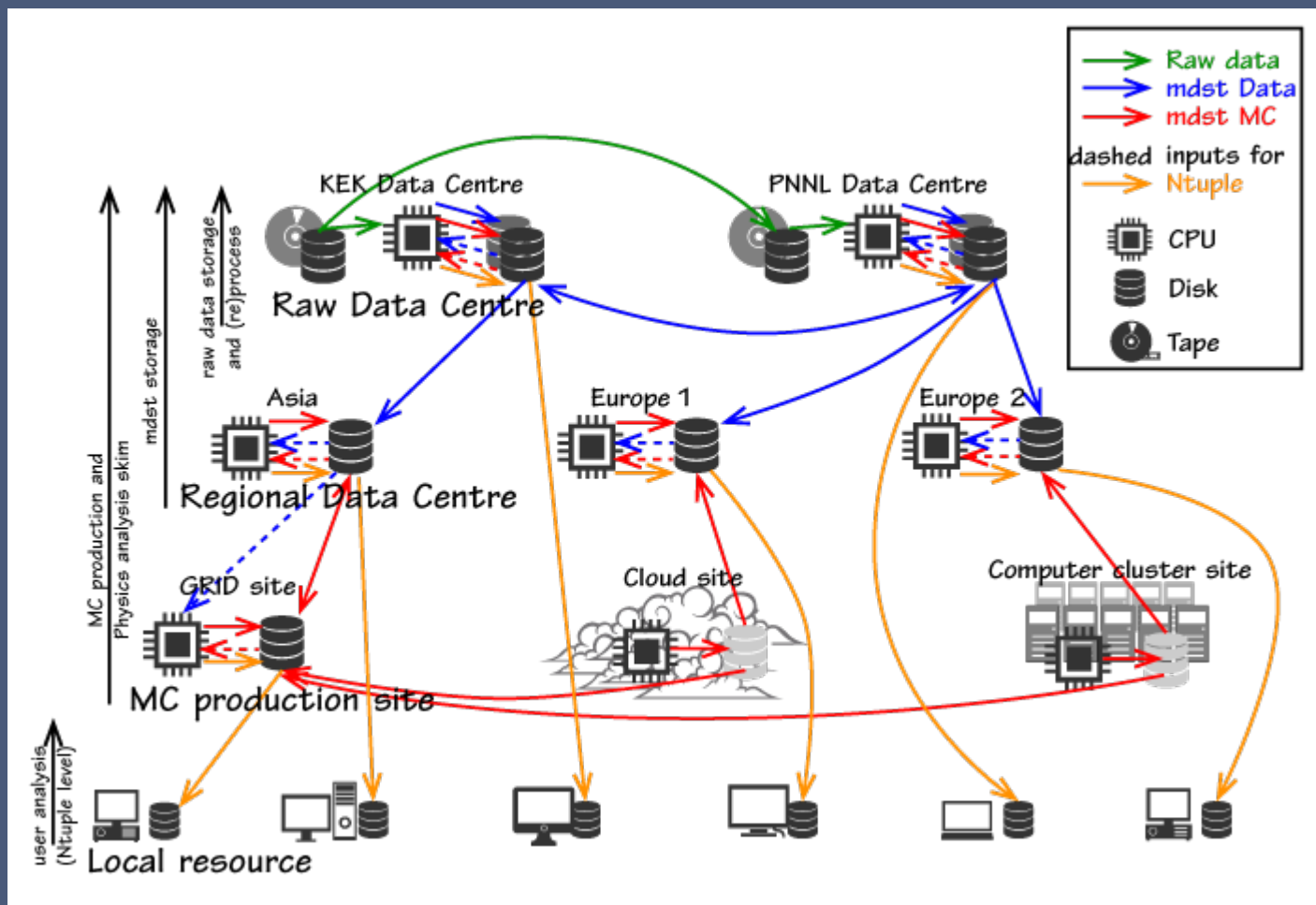
Experiment	Event Size [kB]	Rate [Hz]	Rate [MB/s]
<i>High rate scenario for Belle II DAQ:</i>			
Belle II	300	6,000	1,800
<i>LCG TDR (2005):</i>			
ALICE (HI)	12,500	100	1,250
ALICE (pp)	1,000	100	100
ATLAS	1,600	200	320
CMS	1,500	150	225
LHCb	25	2,000	50



Belle II must handle large amount of data

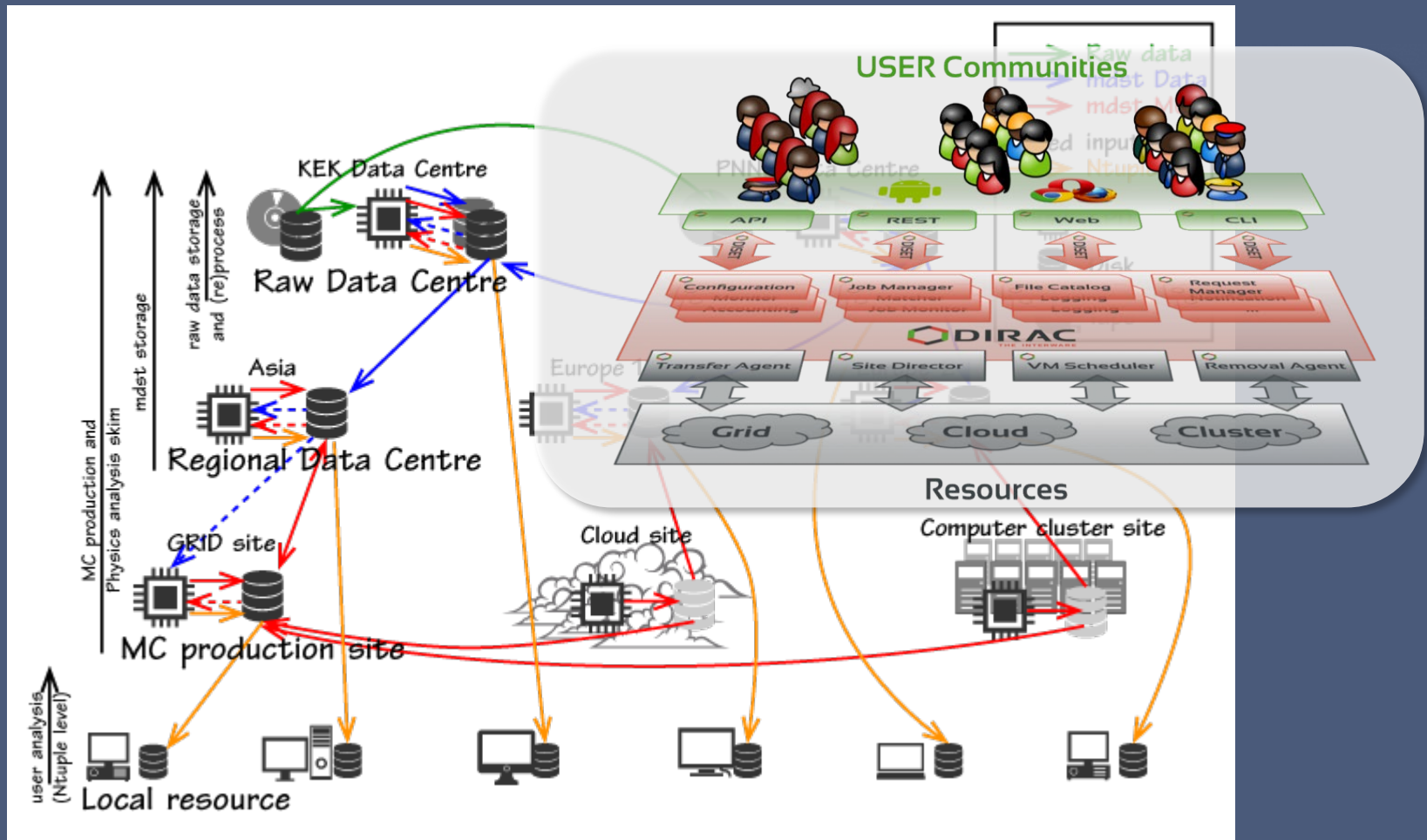
- Tape ~ 200PB
- Disk ~ 200PB (incl. MC)

Belle II computing model



As of May 26

Belle II computing model



DIRAC matches well with our computing model!

As of May 26

DIRAC servers at Belle II

In addition to main production setup, several development setups are deployed

- **Belle-KEK (production)**
 - KEK master
 - KEK slave (SiteDirector, AccountingDB, etc.)
 - PNNL (SiteDirector)
 - UVic (SiteDirector)
- **Belle-IHEP** → Belle-KEK-Development
- **Belle-PNNL (production@PNNL)** → Belle-PNNL-Development
- **Belle-Krakow**

KEK plans system renewal in 2015 and DIRAC will be also replaced

AMGA catalogue

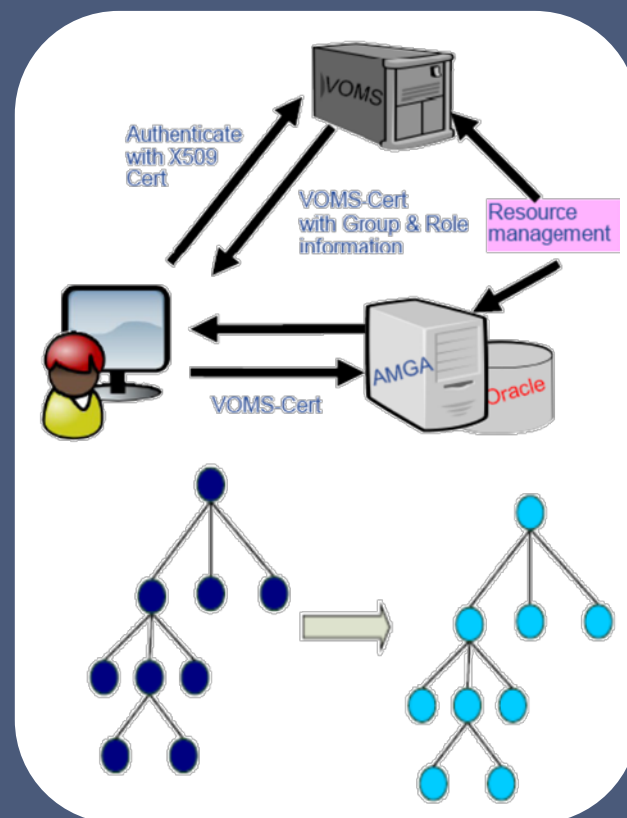
- ARDA Metadata Grid Application
 - Metadata server for GRID environment

Metadata: data of data

LFN, run range, software version...

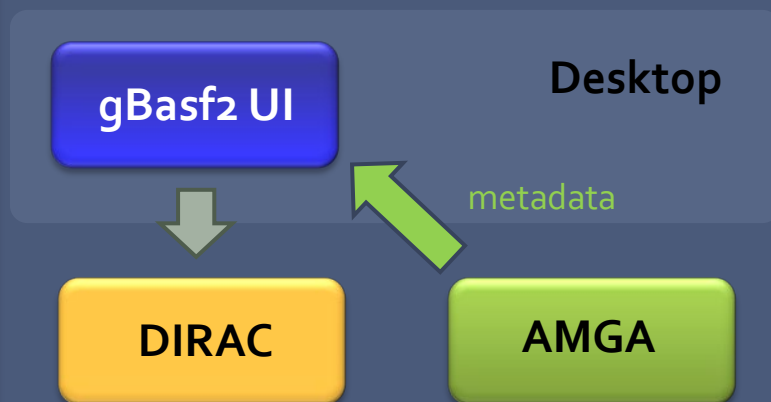
- Main feature
 - Integration with GRID security
 - Secure connection using SSL
 - Replication of data
 - Asynchronous and hierarchical

e.g. replication of specific data set or run period for a GRID site

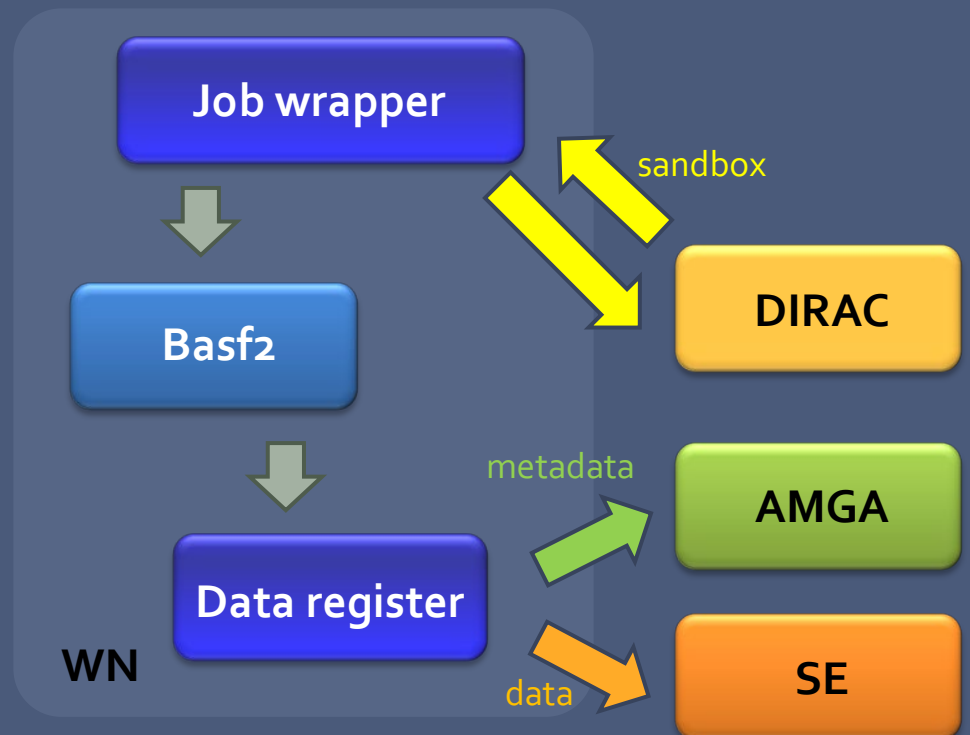


gBaf2 (BelleDIRAC)

- GRID UI dedicated for Belle II
 - Based on DIRAC API (UI + job wrapper on WN)
 - Not only job submitter but a collection of job and data management tools
 - Provide **transparent user experience** of Belle II standard analysis framework (Baf2)



```
% basf2 analysis.py
% gbasf2 -s analysis.py
```

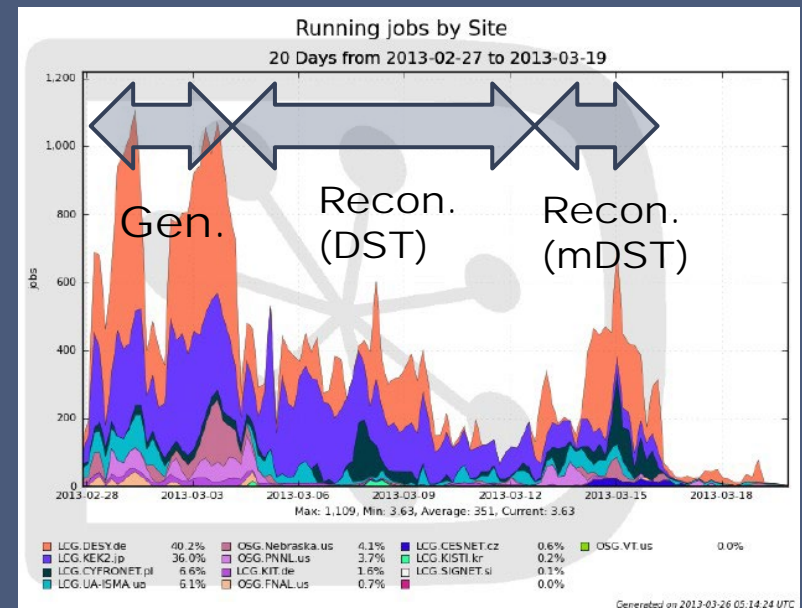


MC production campaign

- Good exercise to learn practical knowledge about bottle-neck, operation know-how...
- Belle II held two times MC mass productions in 2012
 - 1st MC campaign
 - 2nd MC campaign
- Based on the experience, we are performing 3rd iteration of MC mass production campaign
- Each job requires ~1GB input files (simulate beam background) which is provided by SE, and takes a few hours to be processed

First MC production campaign

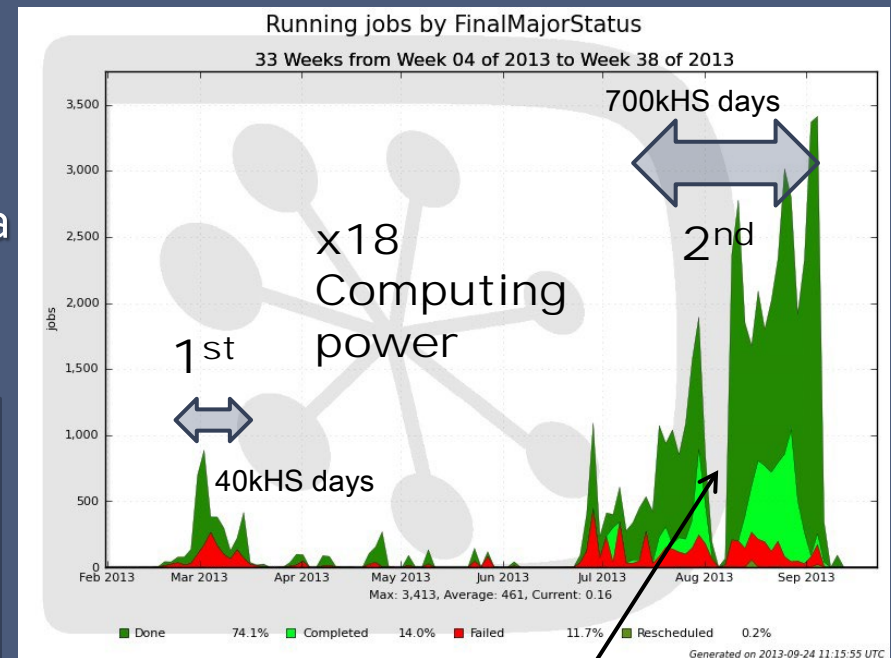
- The 1st iteration of MC mass production using Belle II software on GRID
 - Feb. 28th ~ Mar. 20th, 2013
- The main goal
 - Find possible bottle-necks at everywhere
- Two stages
 - Event generation and detector simulation
 - Reconstruction
- 60M events resulted in 190 TB data (raw level format: DST → high level format: mDST)



- **20% failure rate**
 - Metadata registration
 - Input data download
 - Output data upload
 - Application errors

Second MC production campaign

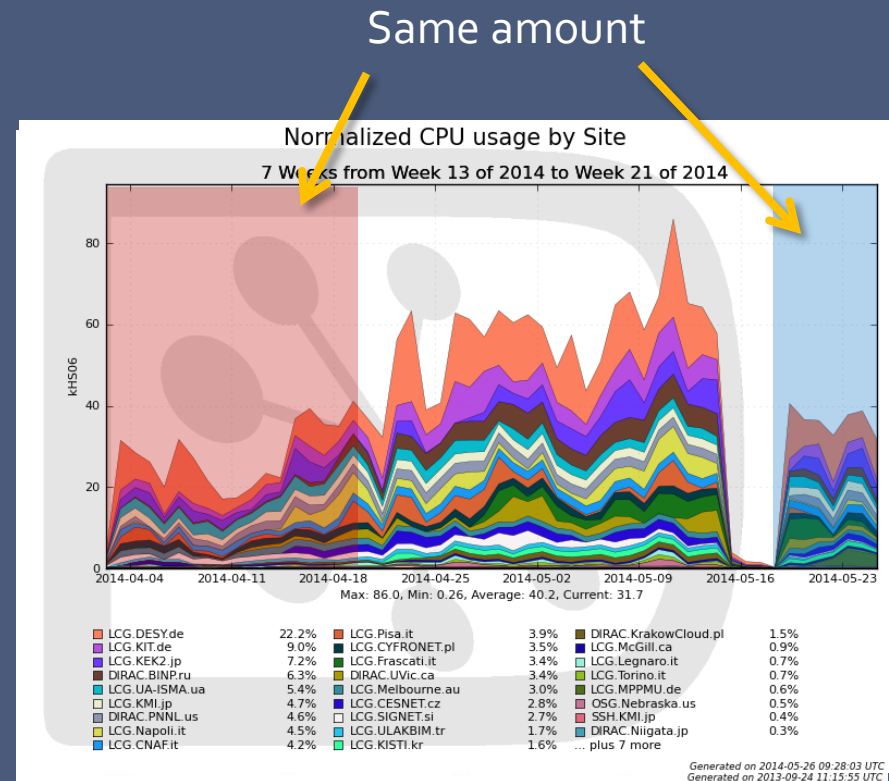
- The 2nd iteration of MC has finished
 - July 23rd ~ Sep. 8th, 2013
- More realistic situation
 - Event generation + reconstruction
 - Background mixing
- 560M events resulted in 8.5 TB data
 - mDST format
- **10% failure rate**
 - Getting decreased through production
 - Final failure rate ~ a few percent
 - No application crash



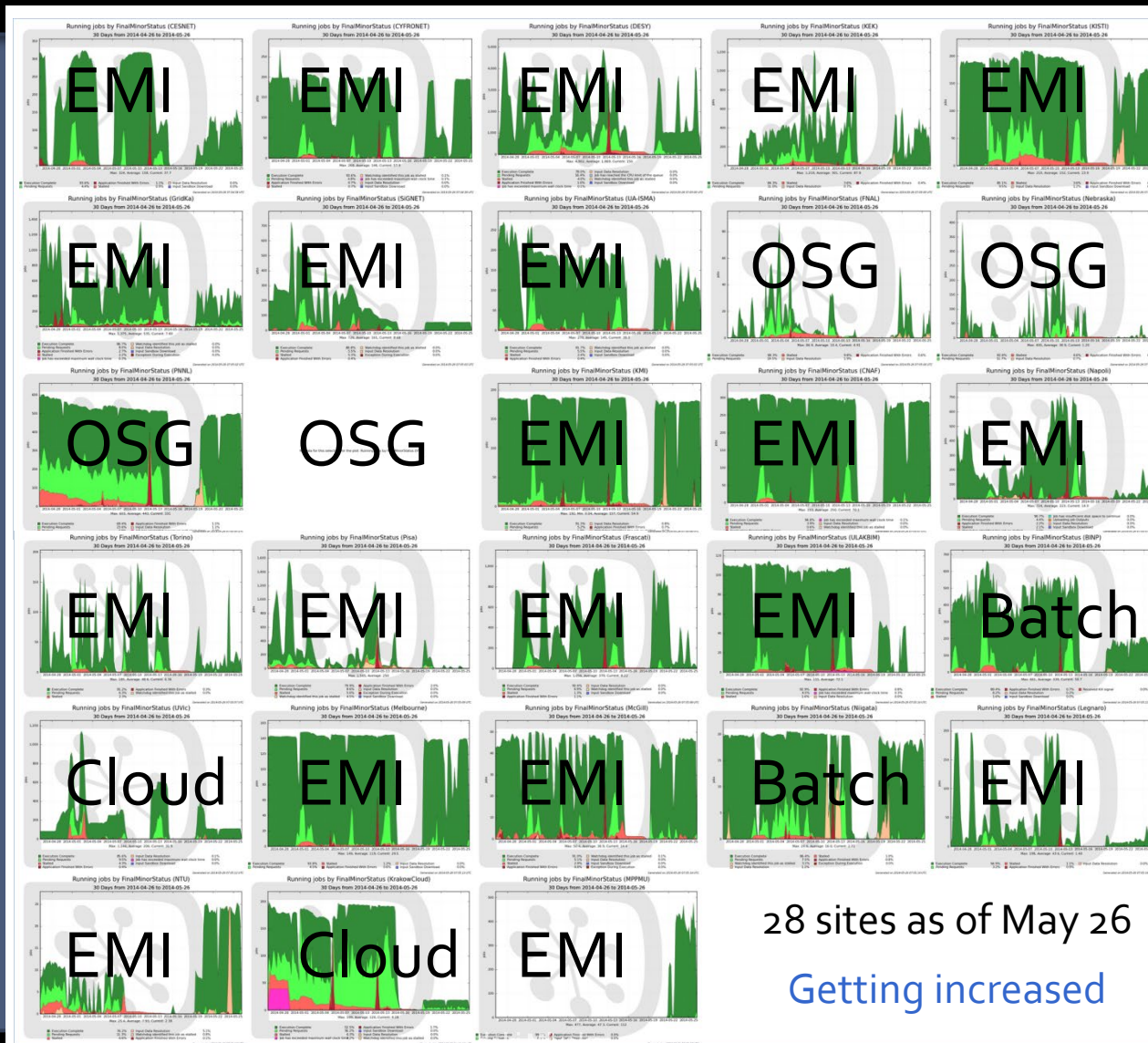
KEK shutdown

Third MC production campaign

- The 3rd iteration of MC is ongoing
 - Apr. 3rd ~ , 2014
- Various resources
 - GRID, cloud, batch system
- More than 4 billion events $\sim 900\text{fb}^{-1}$
 $\sim 86\text{K HepSPEC at the maximum}$
- **5% failure rate at the beginning**
 - Site specific trouble
 - Or wrong VOMS server configuration
 - Finally error rate $\sim 1\%$



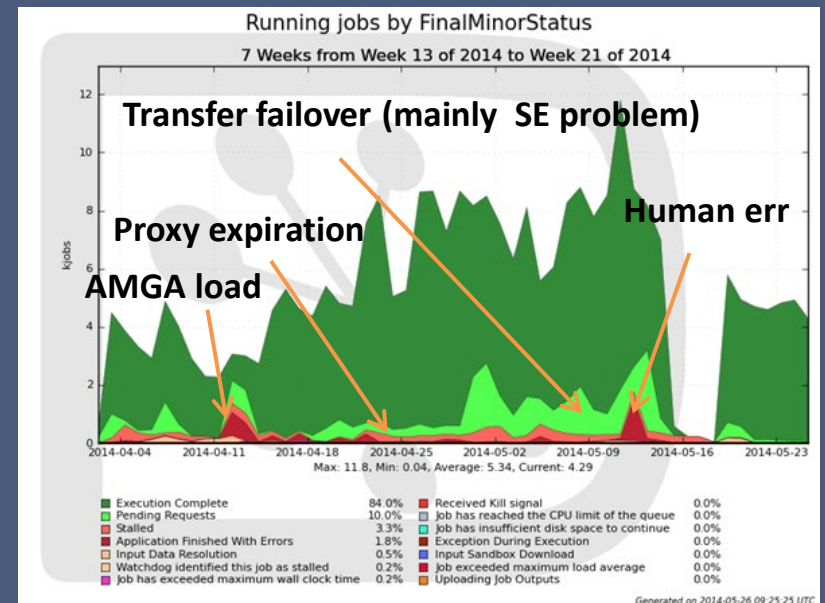
Contributing sites



28 sites as of May 26
Getting increased

Dominant job failures

- At the beginning of the production, job failure happened due to pilot abort (proxy expiration) and AMGA load burst
- Solved by VOMS server reconfiguration (gives long life proxy) and dedicated BelleDIRAC agent to register AMGA metadata (AMGA proxy)
- After that almost no significant error happens except human error or any site specific failure
- Input data inaccessible by SE trouble
 - Allow local data access inside DownloadInputData plugin



BelleDIRAC development plan

- gBaf2 refactoring
 - Redesign code structure and command syntax
- gBaf2 server side
 - Hide logic from client side so that response and security are improved
- Belle II production system
 - Workload management (complicated workflow is handled by Baf2 software)
 - Data transfer

Summary

- Belle II distributed computing system is based on DIRAC and (g)basf2, our own software framework
- We have performed two MC mass production campaigns in last year and running one just for now.
- BelleDIRAC is being actively developed
 - Next major milestone is production system implementation