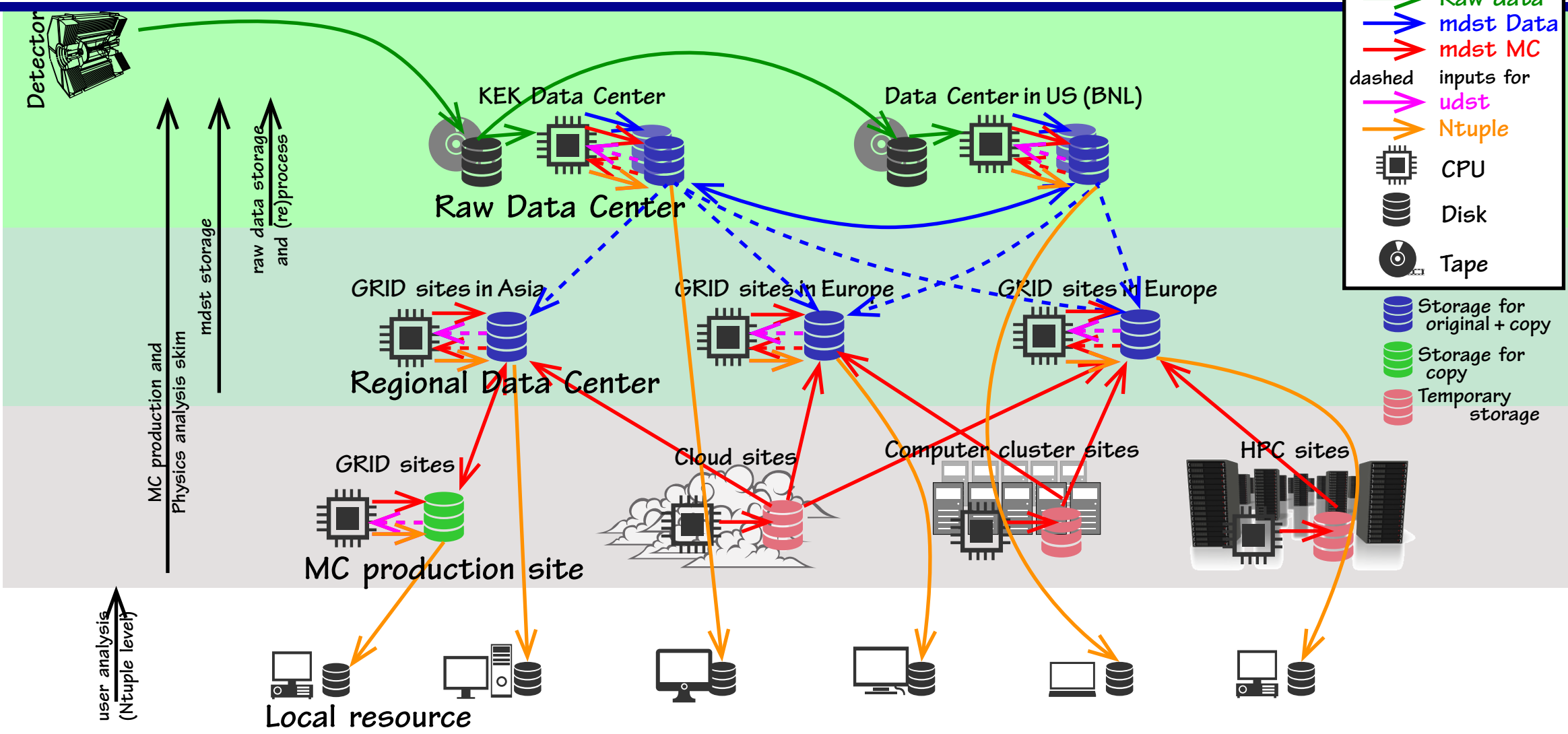
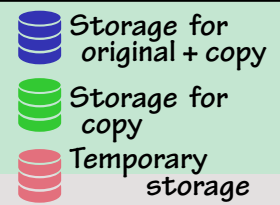
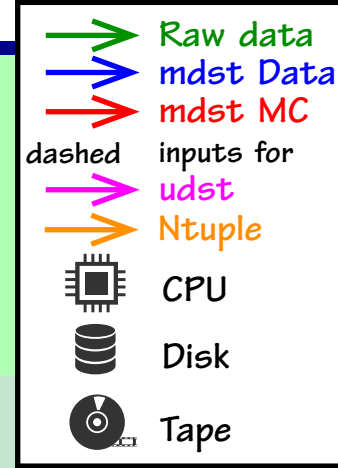


Review of the Belle II distributed computing system



Takanori HARA (KEK IPNS)
takanori.hara@kek.jp

Belle II Computing Model until the end of March 2021



Belle II Distributed Computing Structure

Human

Software interface

- + Interware extension
- + Analysis user interface

BelleDIRAC

Production Manager

Data Manager

End Users

Interware

- + management system



Cyberinfrastructure

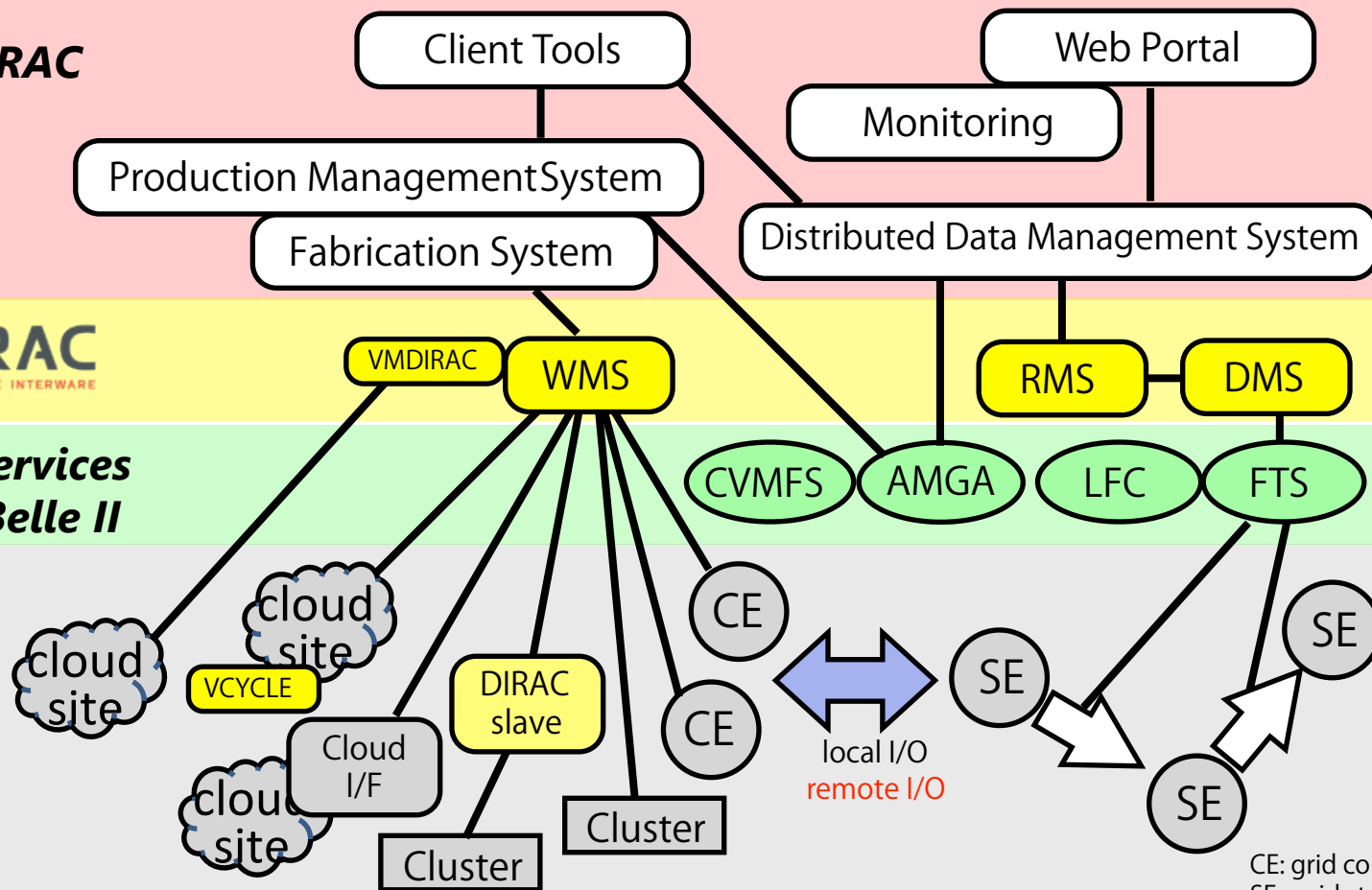
- + Services

**GRID services
for Belle II**

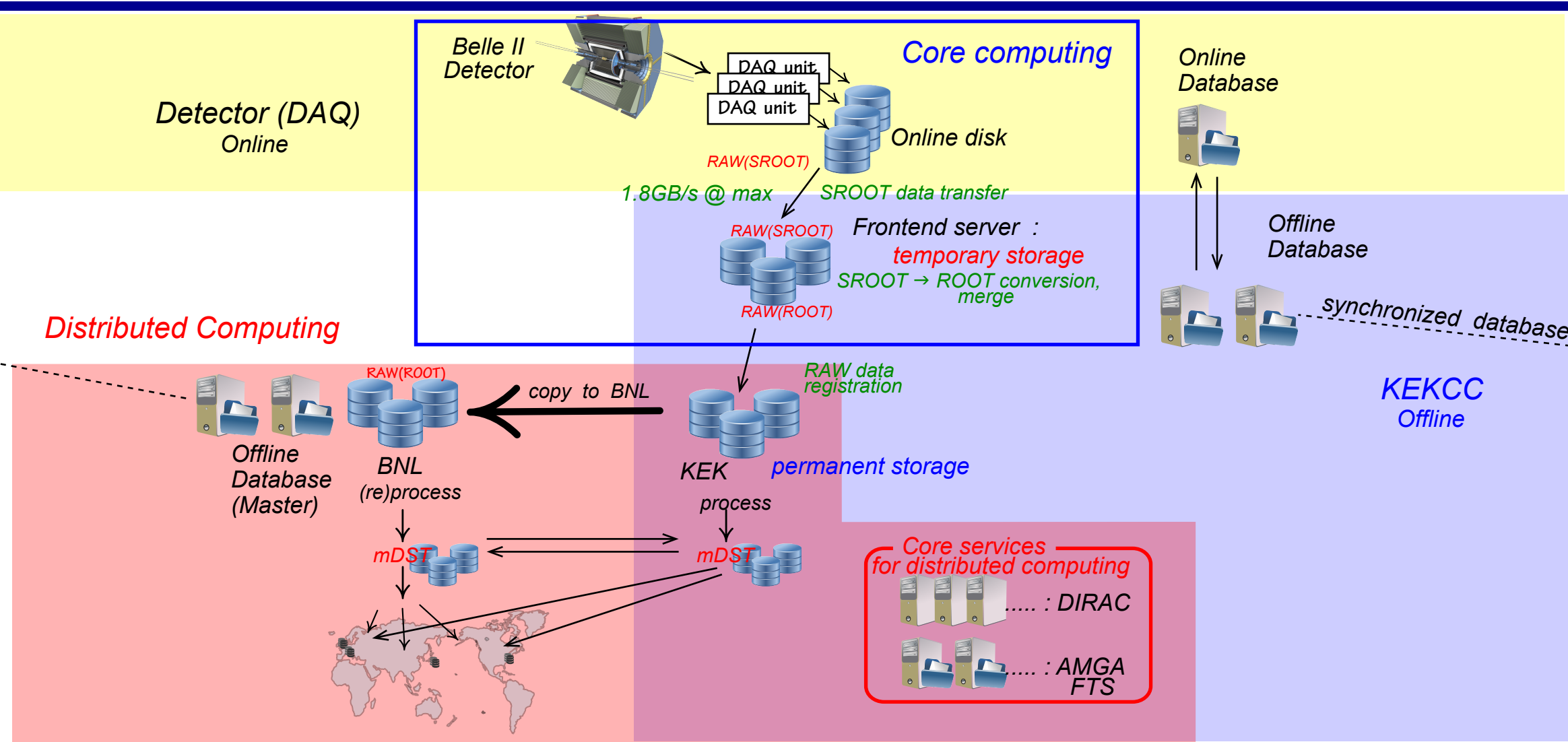
Platform

- + GRID Middleware
 - + OS
 - + Hardware
 - + Network
- Infrastructure

Sites

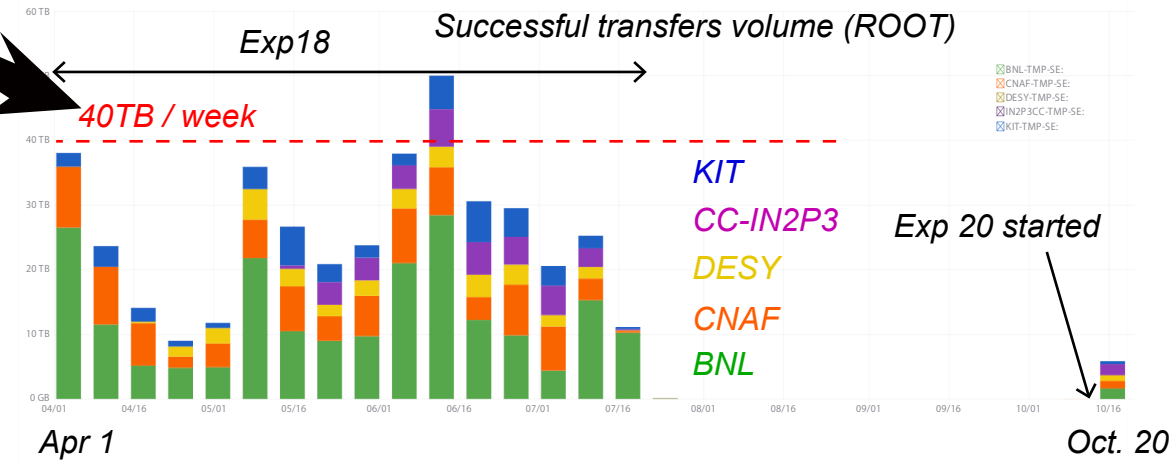
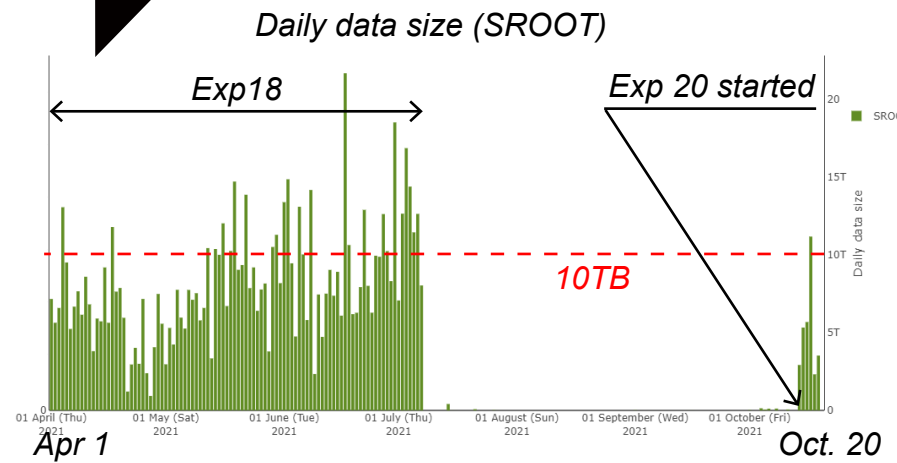
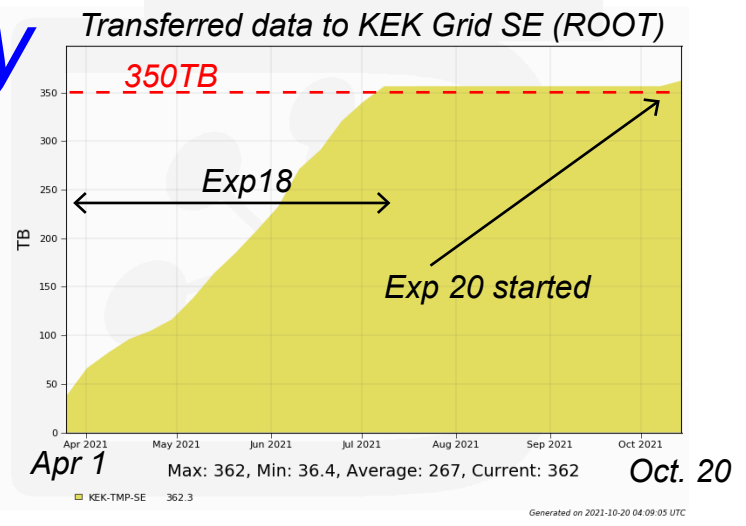
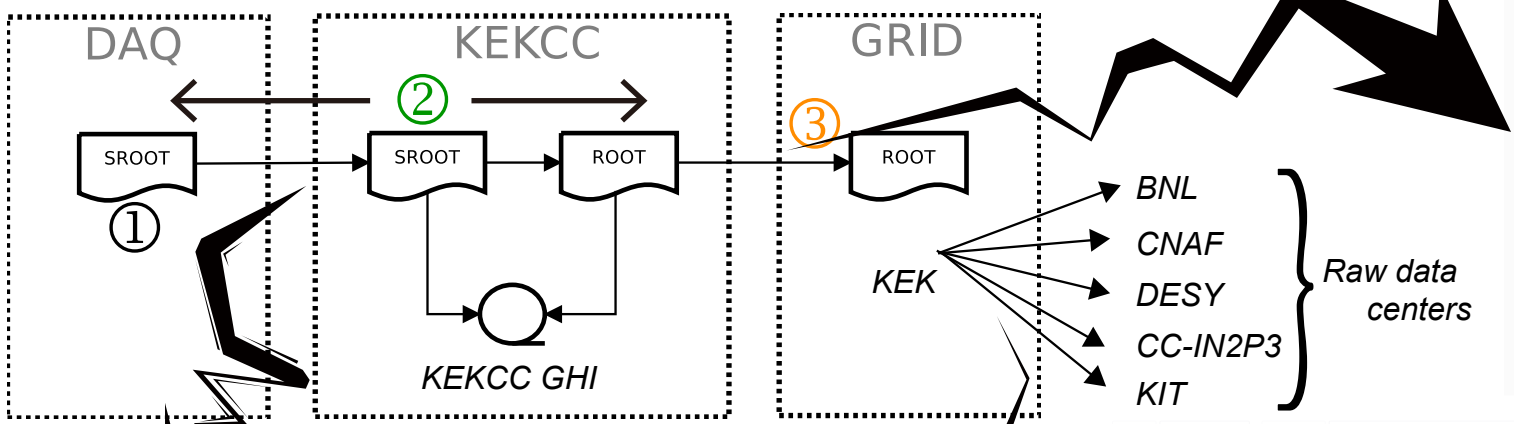


RAW data flow



Core computing

current workflow *operated smoothly*

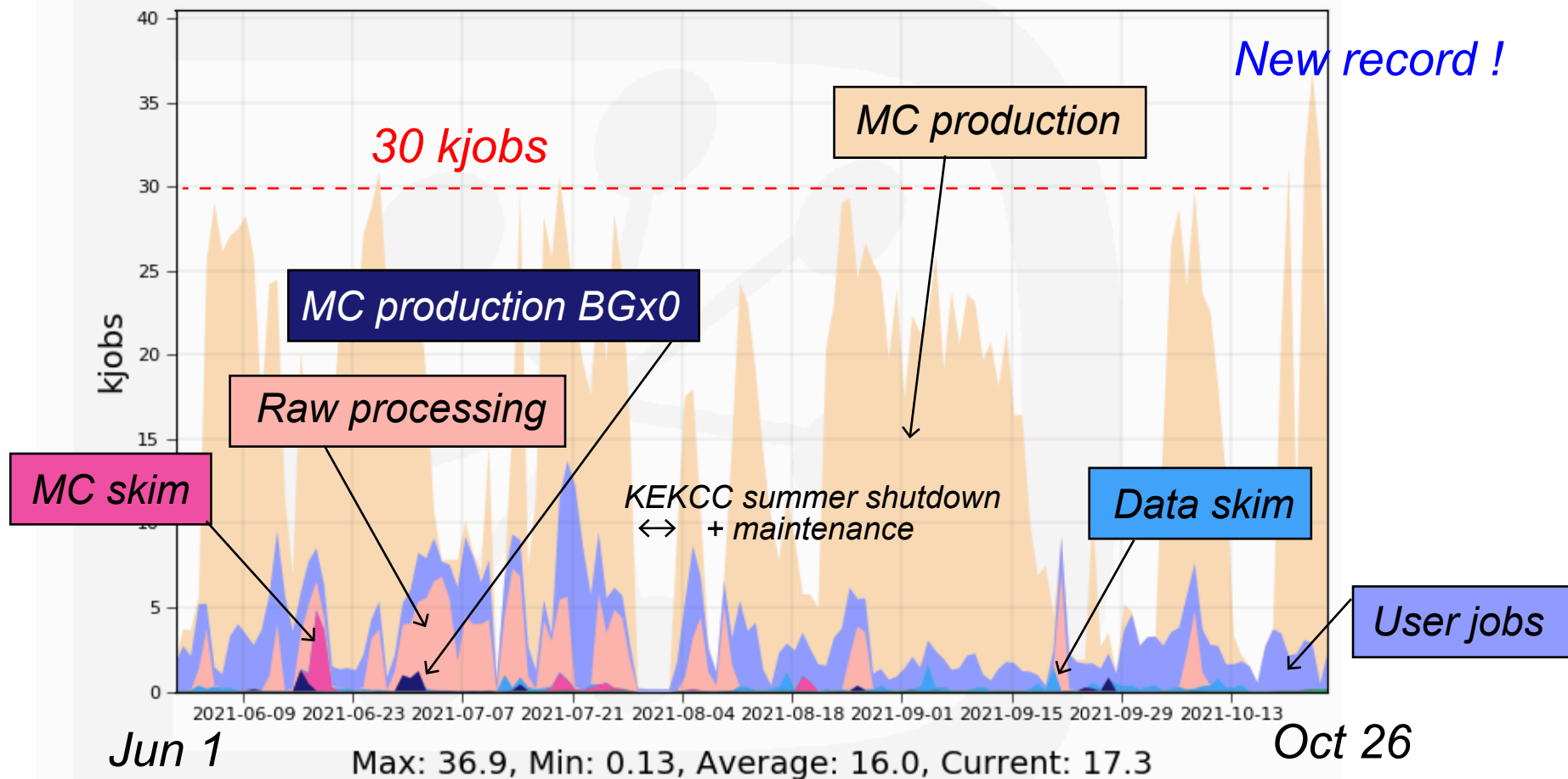


Distributed computing

Basically, smooth operation

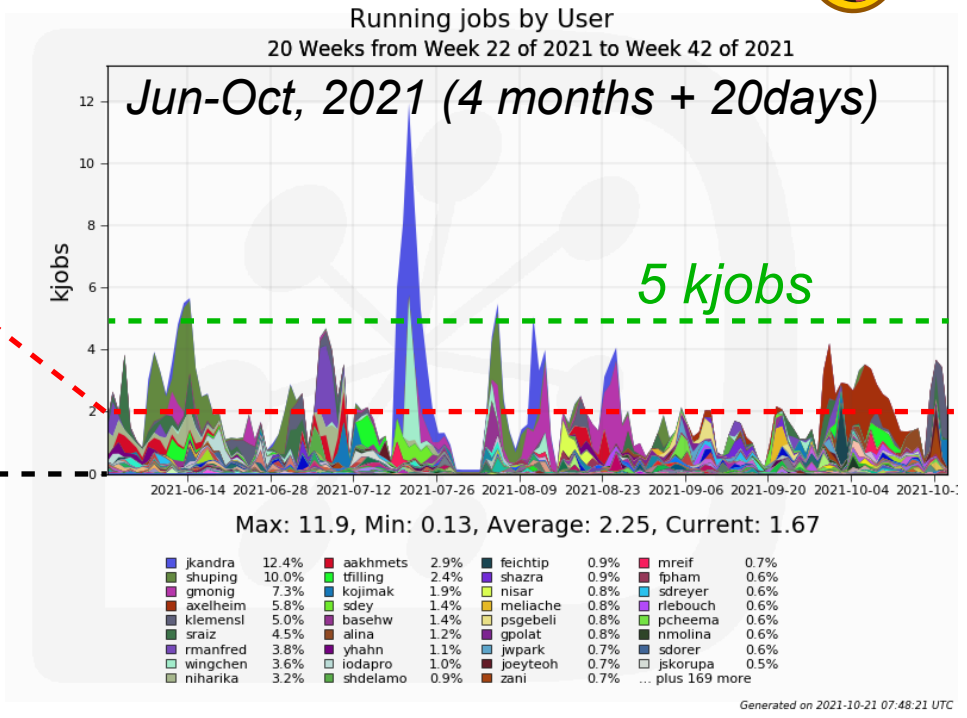
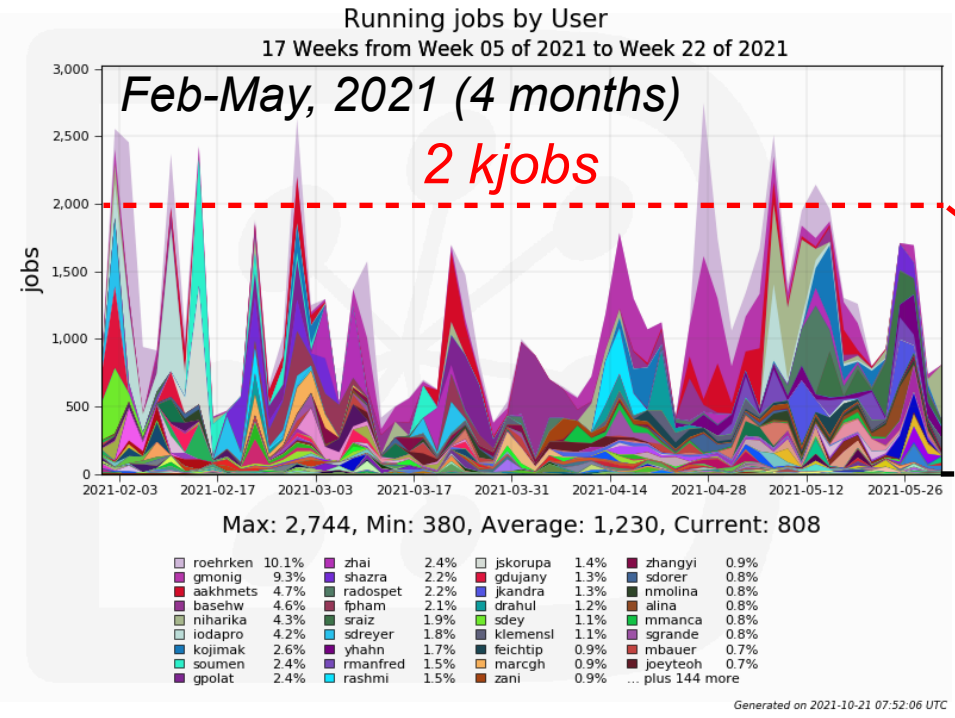
Running jobs by JobType

21 Weeks from Week 22 of 2021 to Week 43 of 2021



Users' activity on Distributed computing

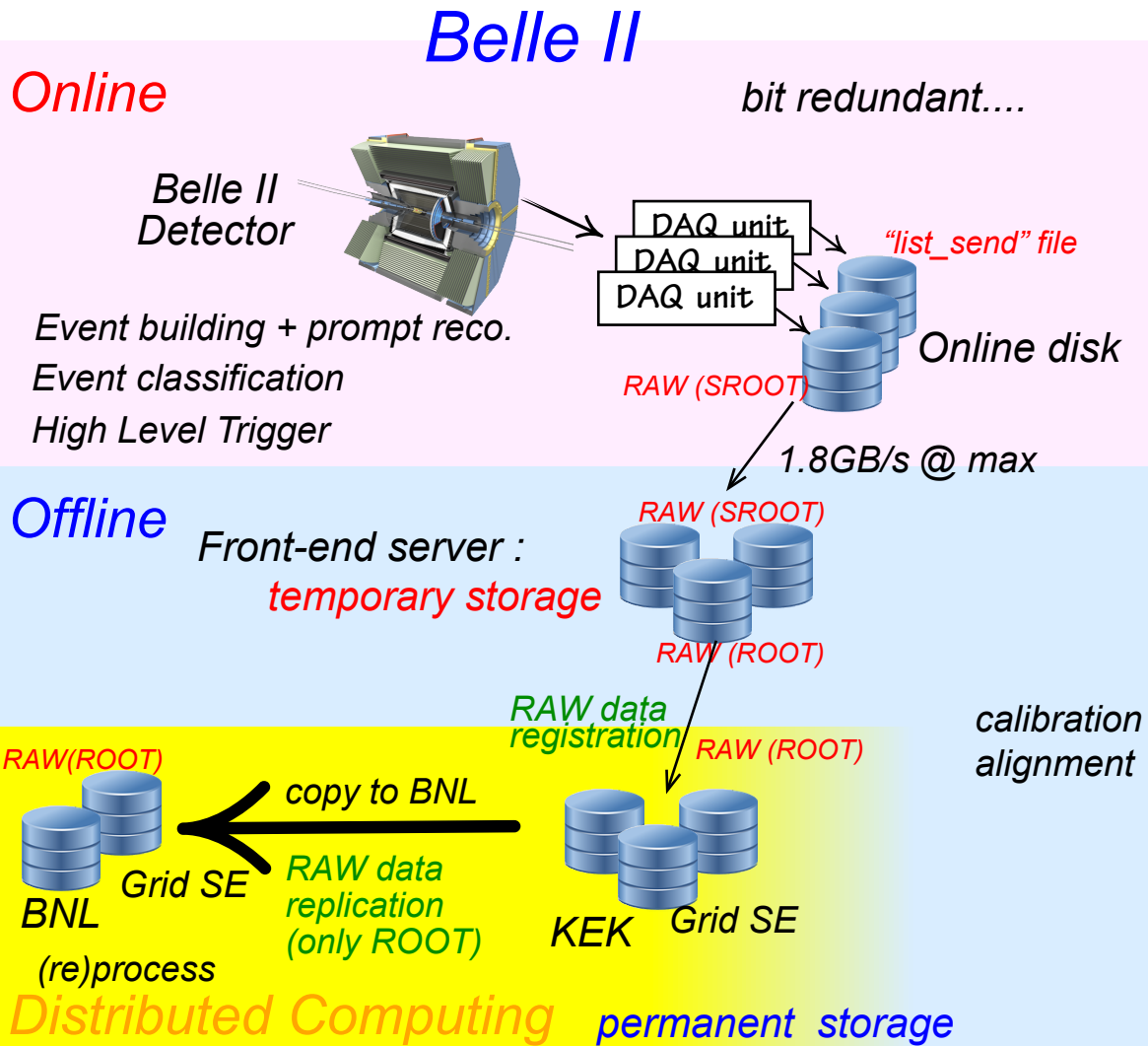
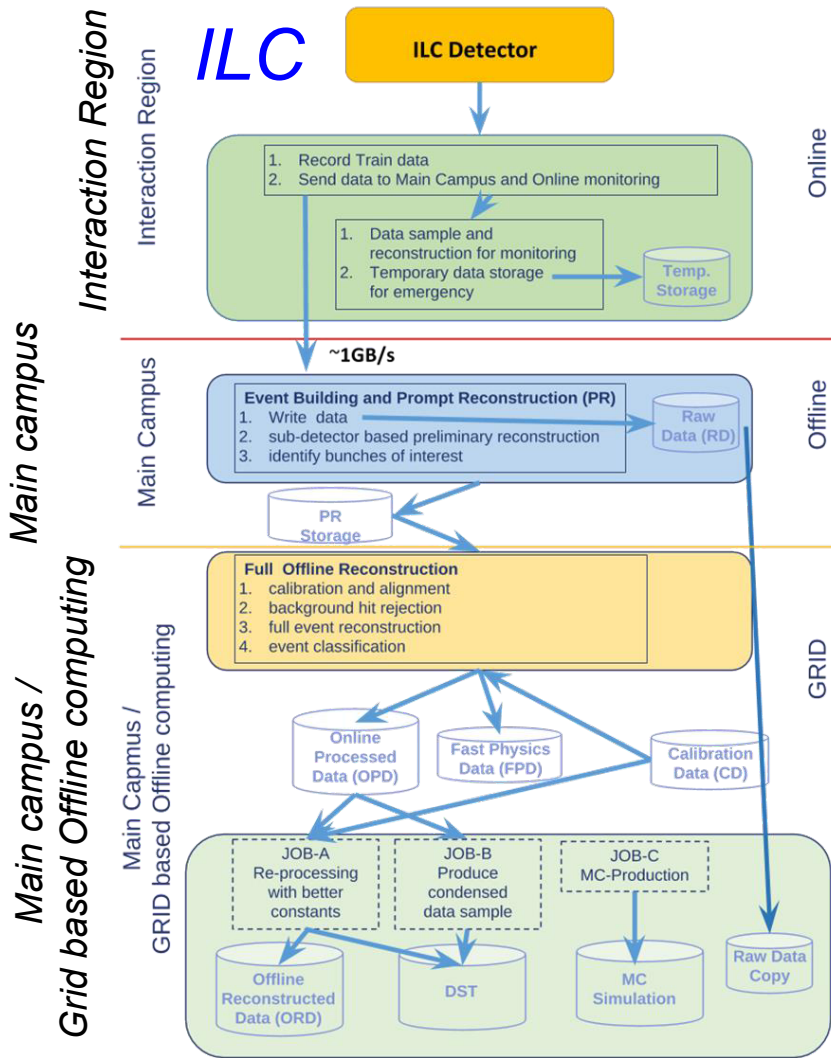
Users became more active on their analysis with Distributed computing before 😄



Basically smooth operation

Comparison : ILC vs Belle II

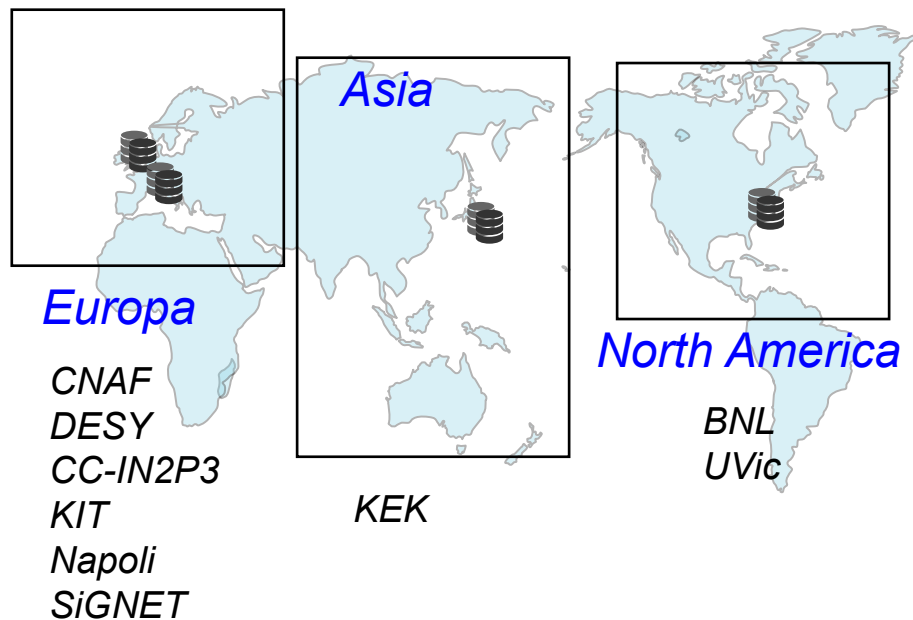
ILCComputing-EDMS1130485.A.1.1.pdf



the gap between ideal and reality

the regional data centers

One big center in each area



to reduce the random network traffic

Each region was expected to have one full set of data

Users can analyze the data locally

but in reality...

Data handling was not easy

Computing resources ↓

two replicas in the world

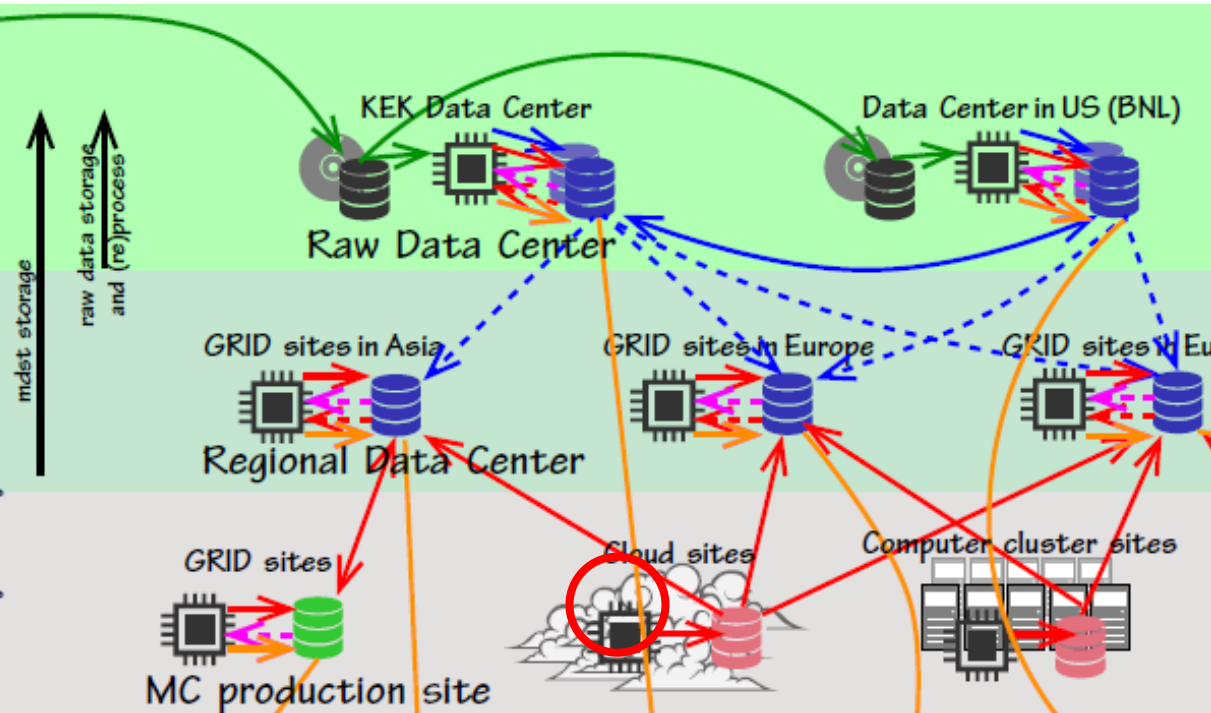
Network environment ↑

RAW data : Japan 100%

(plan) : USA 30%, Italy 20%, Germany 20%, Canada 15%, France 15%

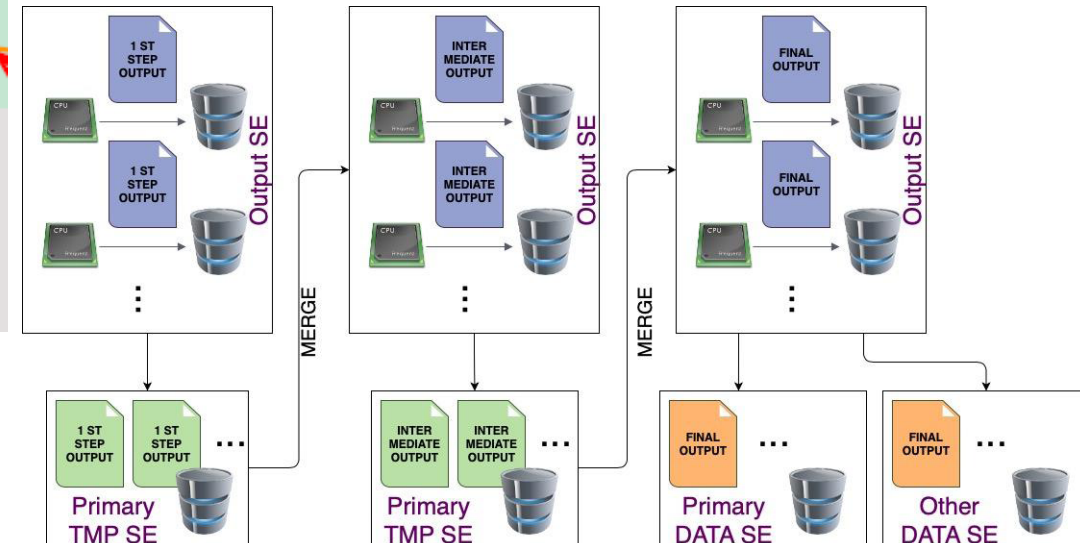
the gap between ideal and reality

MC output was planned to be stored on the SEs where it was produced



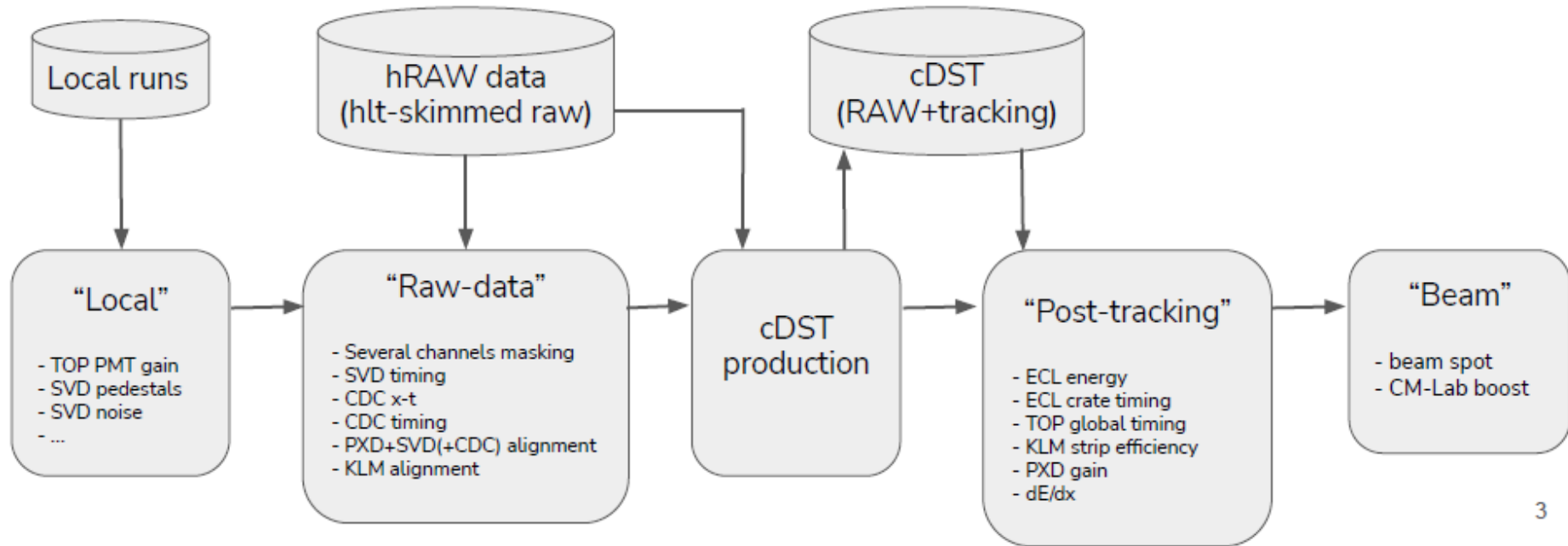
to reduce the network traffic
but in reality...

need to collect the output files to one SE
and merge them there
(until the file size is enough to large)



the gap between ideal and reality

Computing resources for calibration was not included in resource estimation



2030 approx projections for a full recalibration (6 mo/year of data taking):

- 5 PB of tape
- 2 PB of disk (split the re-calibration in 4-5 chunks)
- ~300000 kHS06xhr CPU

model is evolving time by time

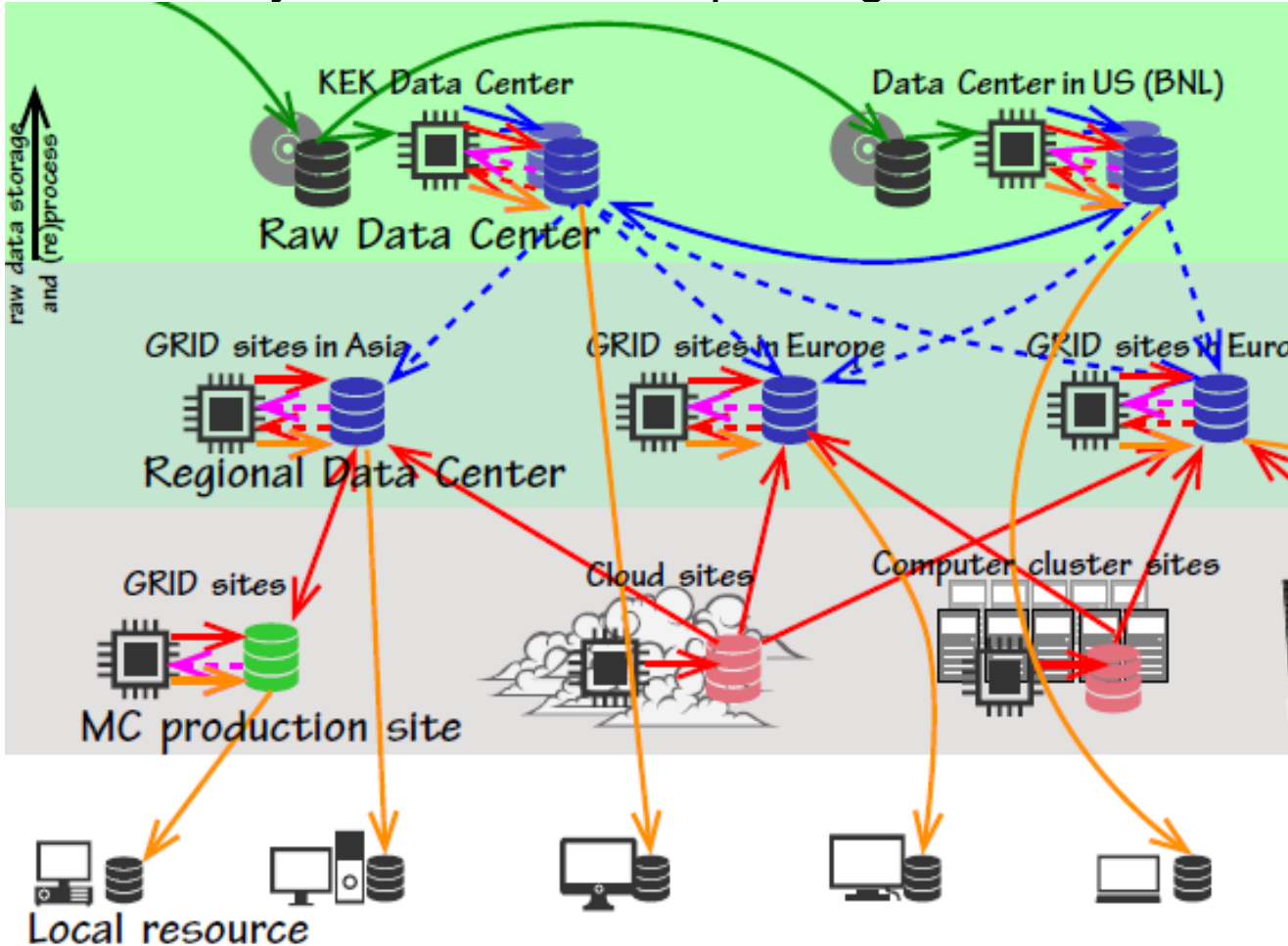
necessary computing resource keeps changing

it is likely to happen in the early stage of the experiment ...

unofficial

the gap between ideal and reality

Analysis users still keep using KEKCC as main analysis farm



Once Ntuple level outputs are created, users are expected to download them to their local computing resource in their home institute and continue their analysis there...

but in reality...

*Many users still keep using KEKCC for their analysis.
→ affecting the KEKCC CPU / storage*

*One reason is
“Belle” data is still important ...*

Network environment for Belle II

30% of sites on LHCONE covering more that 80% of Computing and Storage Resources

70% of sites General IP

5 Sites on LHCOPN

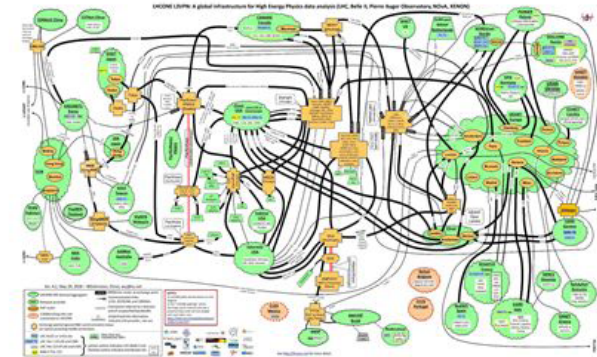
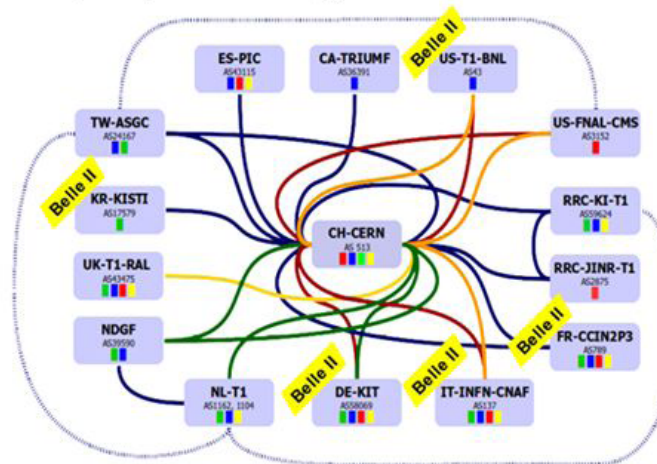
100G Global Ring
runned by SINET

LHCOPN optical
infrastructure that can
be used without
jeopardizing resources

LHCONE L3 VPN
Connecting all the major
Data Centers



SINET5 will be upgraded
to SINET6 from 2022



MoUs for Belle II distributed computing

Memorandum of Understanding

**for Collaboration in the Deployment and Exploitation of the
Belle II Computing Grid**

*Canada (IPP), U.S.A. (BNL), Austria (HEPHY),
Australia (Melbourne U. CoEPP),
Czechia (Charles U.), France (CC-IN2P3),
Germany (KIT, MPP, DESY), Italy (INFN),
Japan (KEK) signed*

*maintained by
Belle II Computing Steering Group*

<https://wiki.kek.jp/display/belle2grid/MoU>

MULTILATERAL

MEMORANDUM OF UNDERSTANDING

**For the Deployment, Operation and Security
of the Belle II Computing Grid**

*Canada (IPP), U.S.A. (BNL), Austria (HEPHY),
Australia (Melbourne U. CoEPP),
Germany (KIT, DESY),
Italy (CNAF, Tier2-federation),
Japan (KEK),
Slovenia (JSI), Korea (KISTI) signed*

*maintained by
KEK Computing Research Center*

Summary

- ◆ *ILC computing model is very similar to the Belle II computing model*

Structure, middleware, software (basf2 [Belle II] = basf [Belle] + ILC software framework)

- ◆ *We expect the ILC computing system will be established with*

More modern technologies

experience on other experiments such as ATLAS, CMS, LHCb, ALICE, Belle II, etc...

many infrastructures, software, middleware are already available

many experiments already established their own computing model

→ but this does not mean that the ILC computing model can be easily established

you must face many issues in technology development, operation, security, certificate, etc. etc.

- ◆ *Communication with Computing sites and Network providers is important*

- ◆ *Knowledge Transfer is also important*