Computing @ Belle II

Thomas Kuhr

Terascale Alliance Workshop

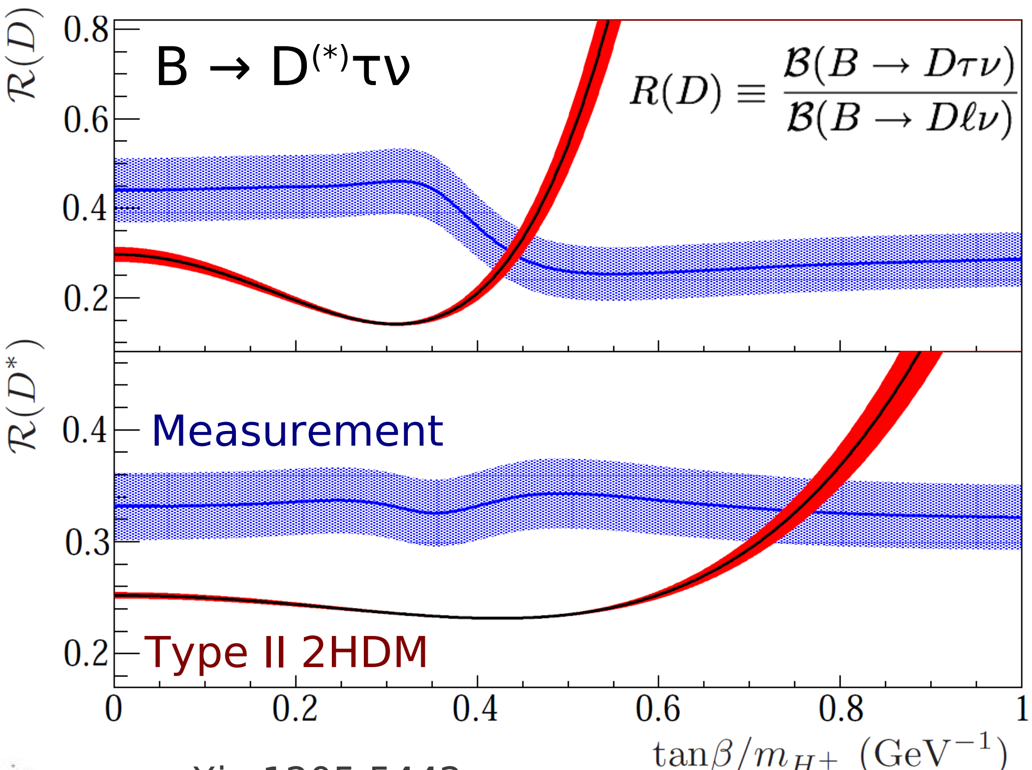
03.12.2013

Physics Objective of Belle and Belle II

- ✔ Confirmation of KM mechanism of $\not{C}\not{P}$ in the Standard Model

- ✗ $\not{C}\not{P}$ in the SM too small (by many orders of magnitude) to generate observed baryon asymmetry in the universe

- ➔ Need sources of $\not{C}\not{P}$ beyond the SM

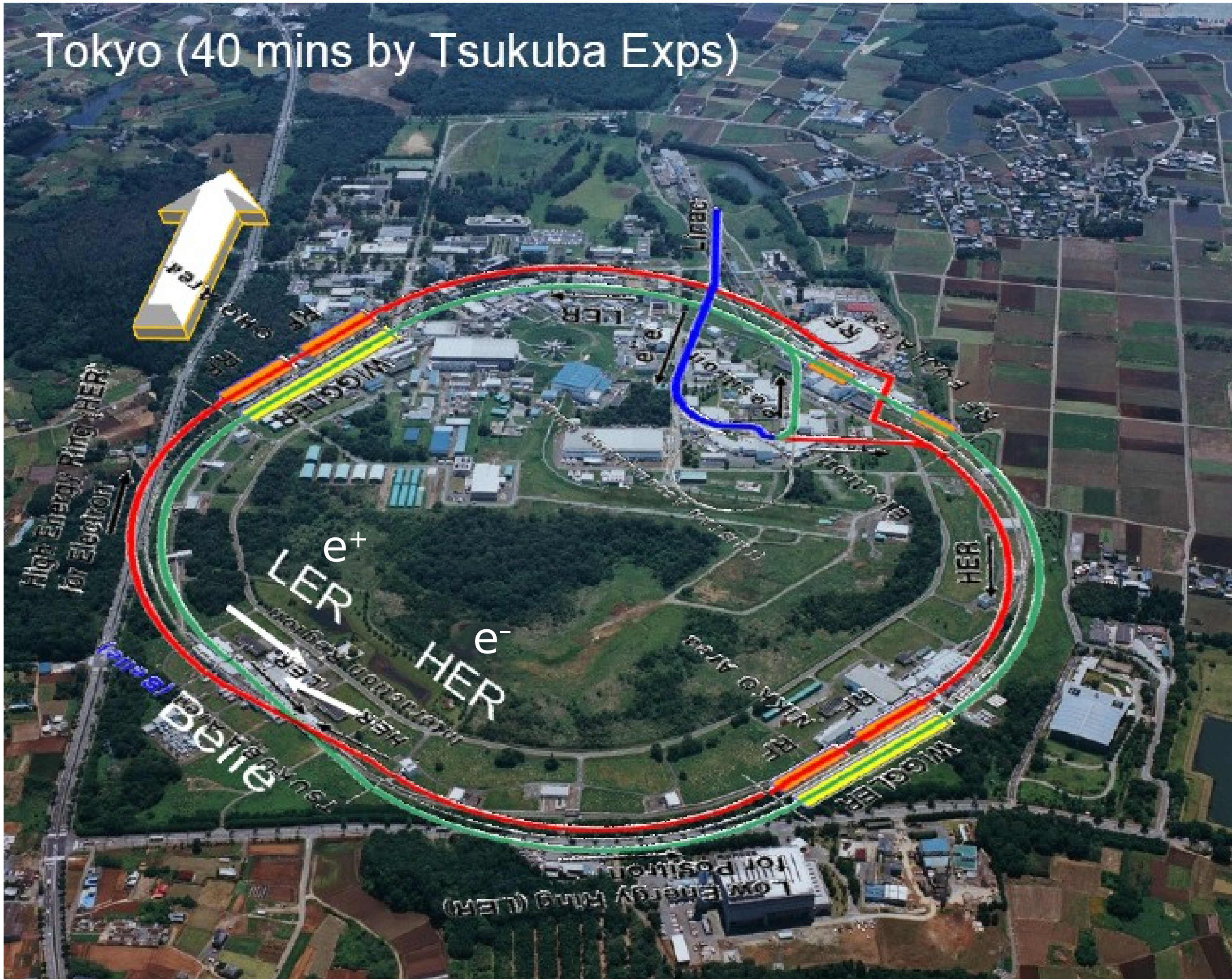


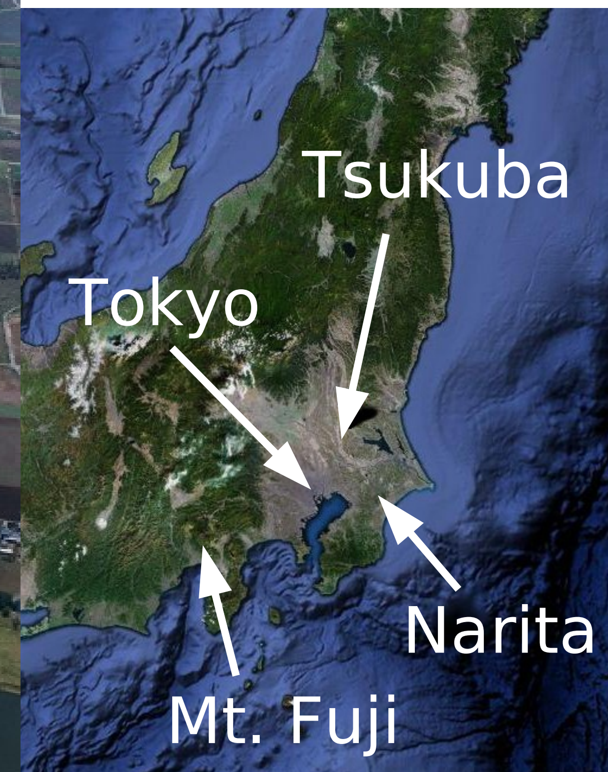
arXiv:1205.5442

➔ Super B factory

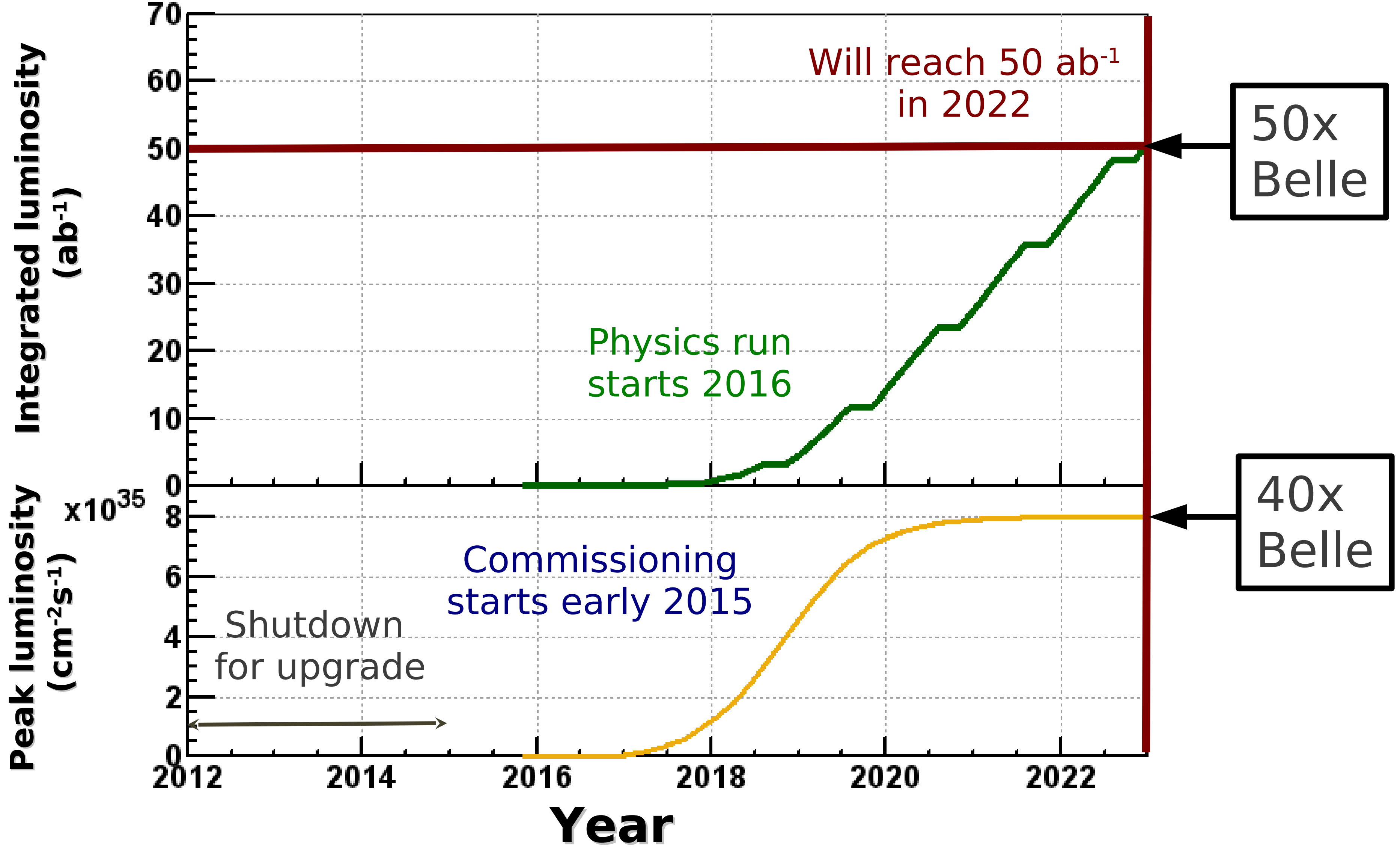
Complementary to LHCb

KEK Site


Tokyo (40 mins by Tsukuba Exps)
e+
LER
e-
HER
Belle
Linac
HER
Low Energy Ring (LER) for Positron
High Energy Ring (HER) for Electron


Tsukuba
Tokyo
Narita
Mt. Fuji

Projection of Luminosity at SuperKEKB



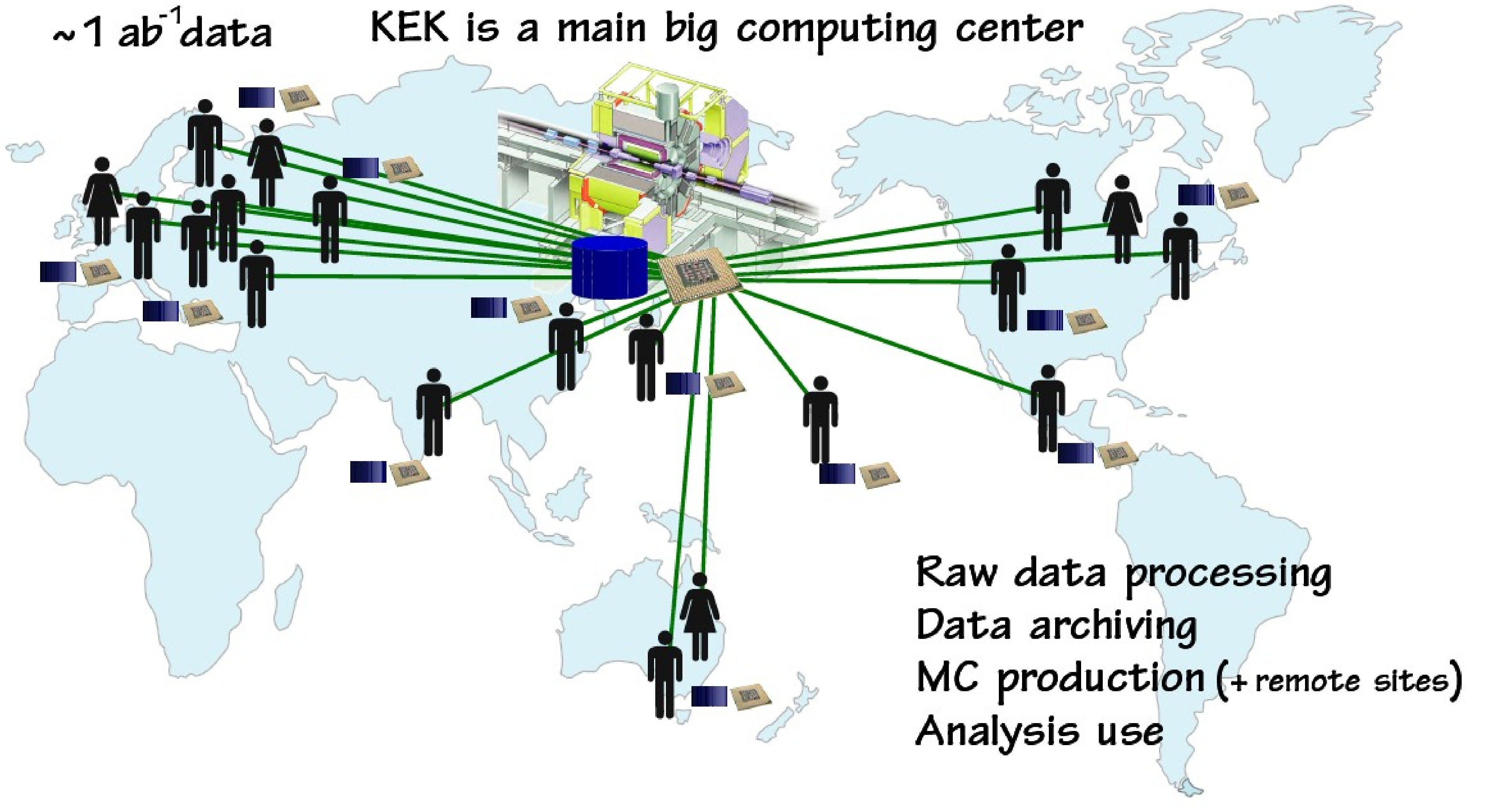
Estimated Data Rates

Experiment	Event Size [kB]	Rate [Hz]	Rate [MB/s]
High rate scenario for Belle II DAQ:			
Belle II	300	6,000	1,800
LCG TDR (2005):			
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

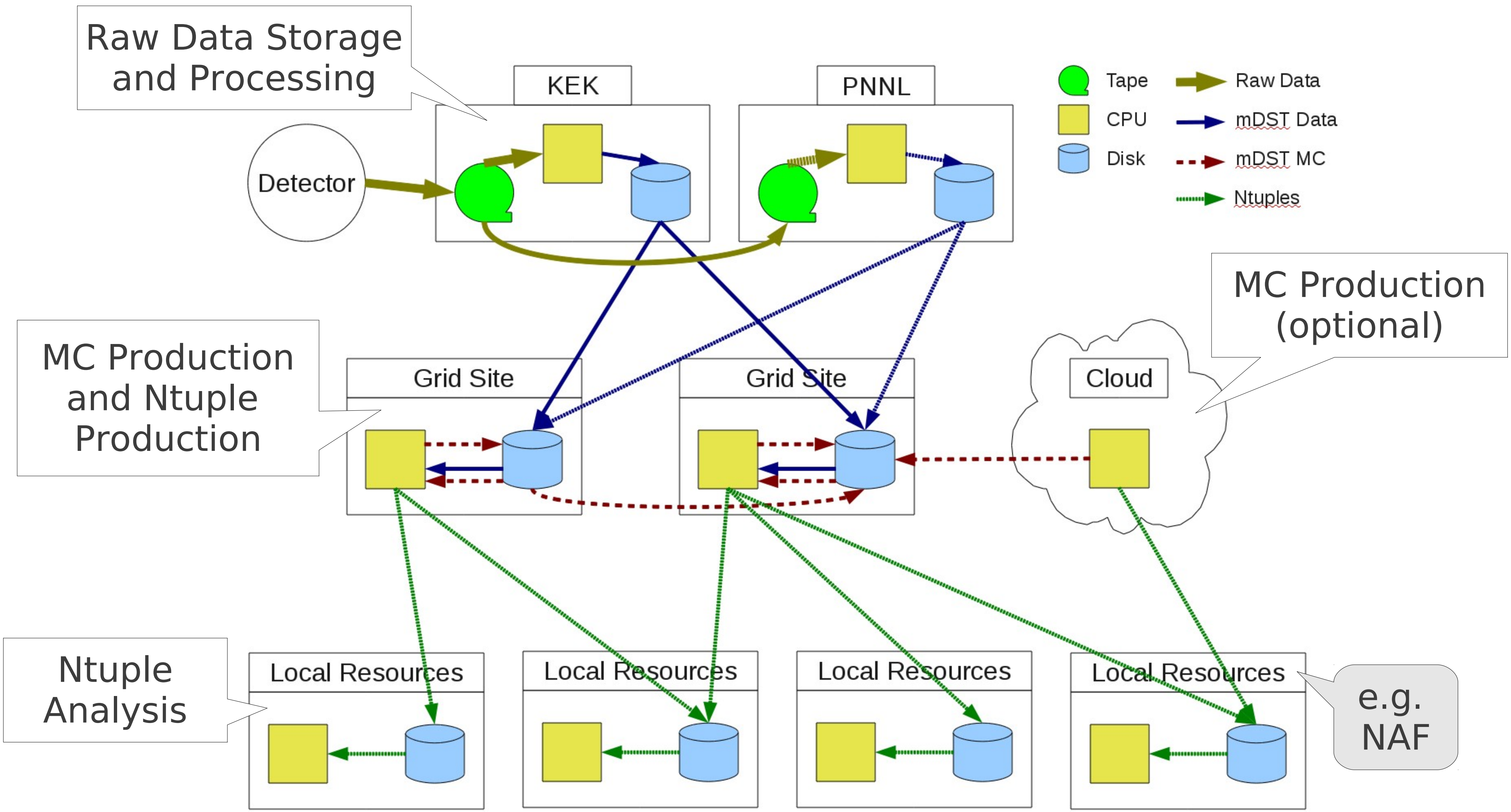
Belle II Collaboration

~600 members
99 institutions
from 23 countries

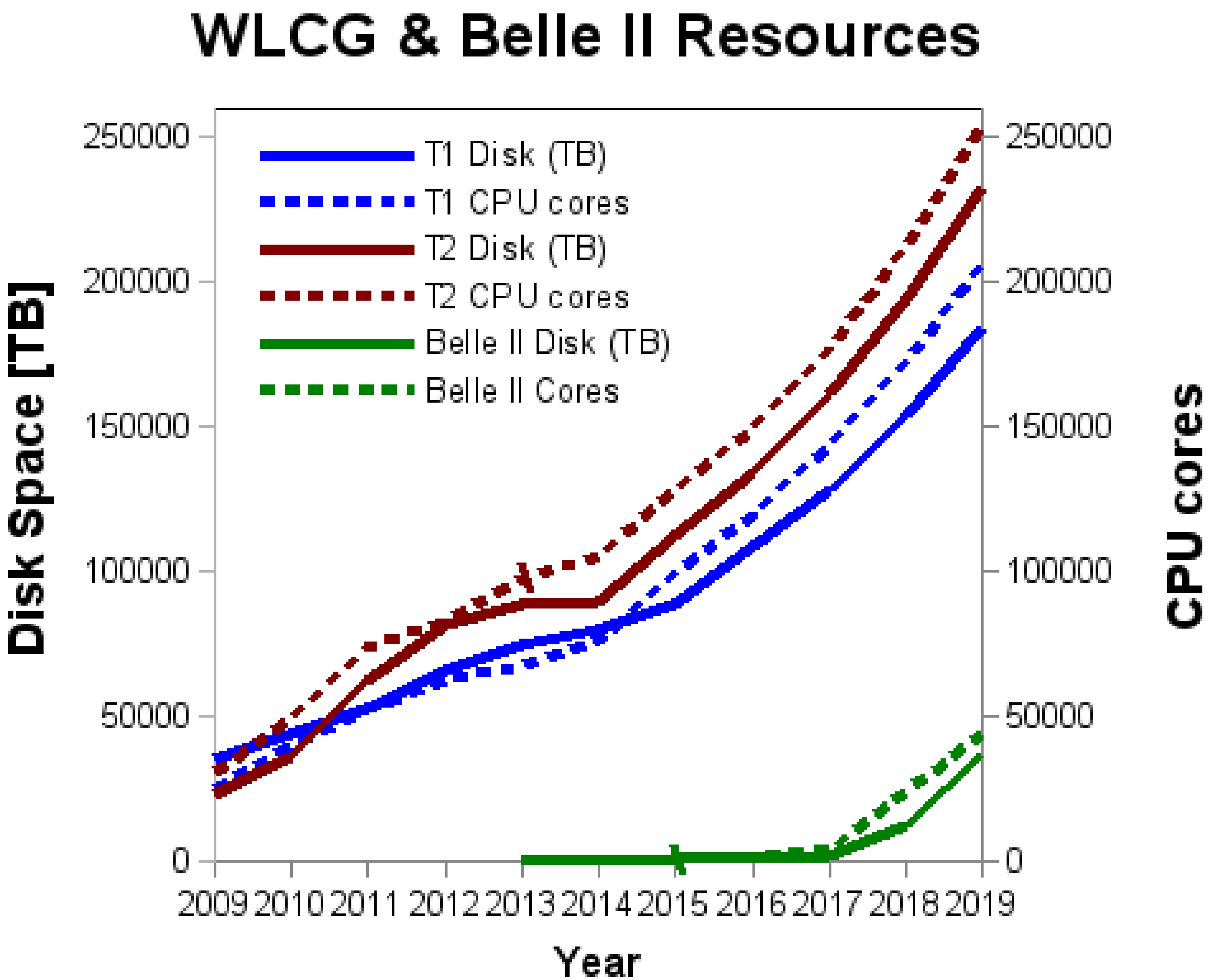
Belle (I) Computing Model

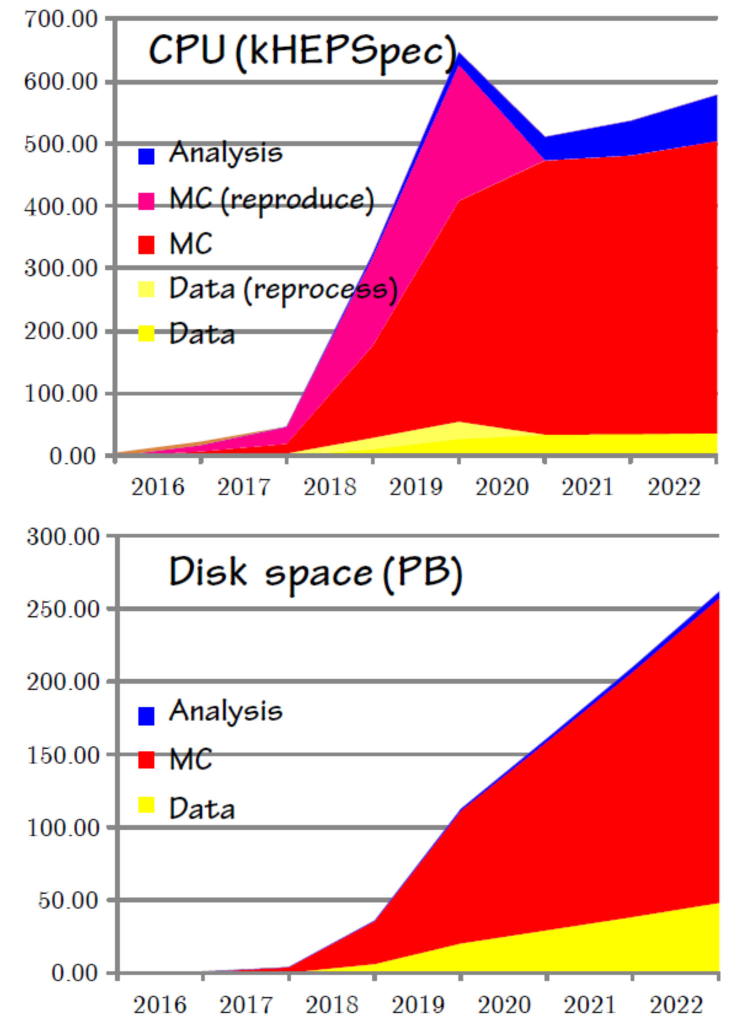


Belle II Computing Model



Resource Estimates





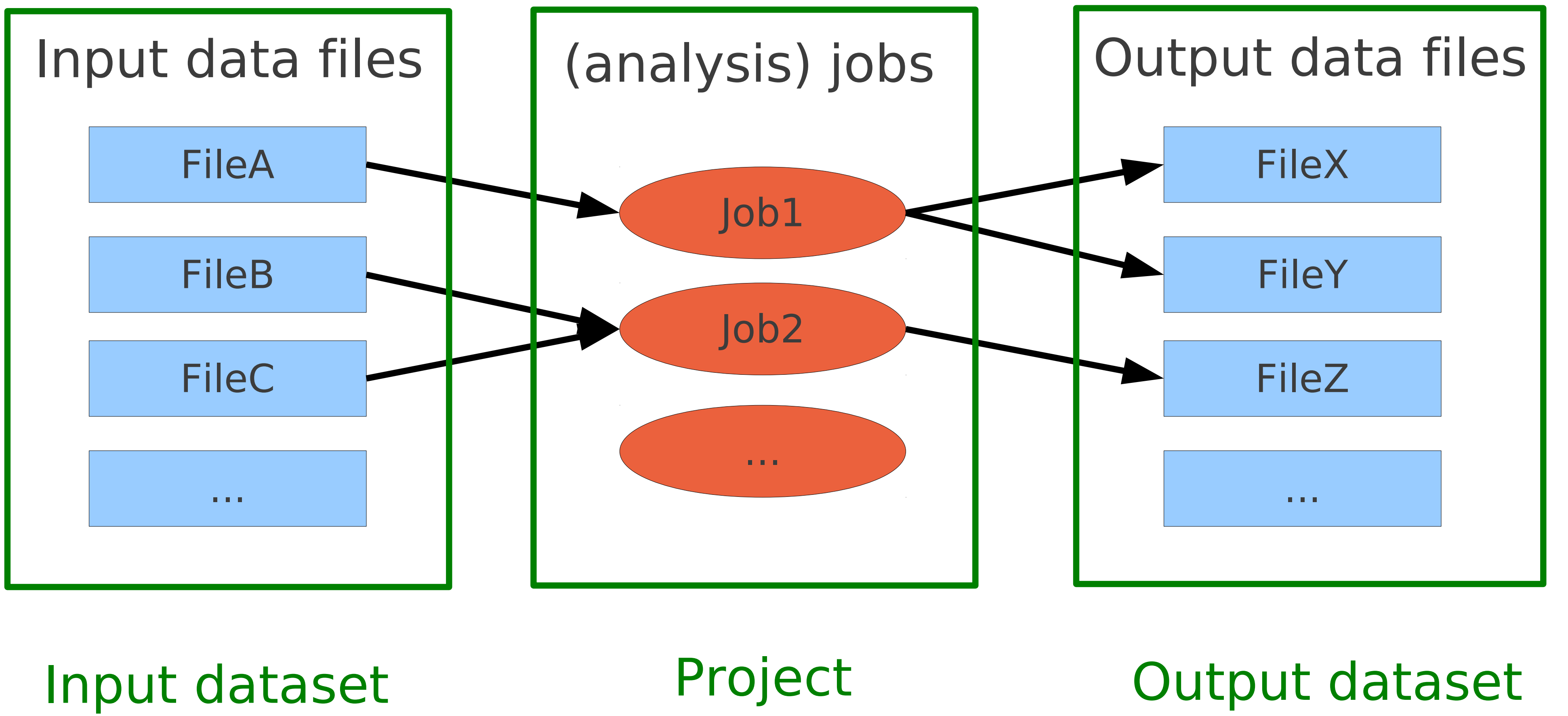
Distributed Computing System

- Based on existing, well-proven solutions plus extensions for Belle II
  - → DIRAC for job management
  - → AMGA for metadata
  - → CVMFS for software distribution
  (thanks to CERN and Steve Traylen for providing the Stratum-0 server, and to GridKa for the stratum-1 server)

root/svn/trunk/grid/BelleDIRAC

FrameworkSystem/	**4326** (9 months ago) by *myco*: basic sites management service for BelleDIRAC
Web/	**5519** (4 months ago) by *hideki*: remove unused AMGA API
WorkloadManageme…	**6098** (2 months ago) by *hideki*: fix a bug
gbasf2/	**6647** (2 weeks ago) by *hideki*: fix unnecessary AMGA initialization
README	**4325** (9 months ago) by *myco*: init files for BelleDIRAC distribution
__init__.py	**6348** (6 weeks ago) by *hideki*: release for 2nd MC campaign

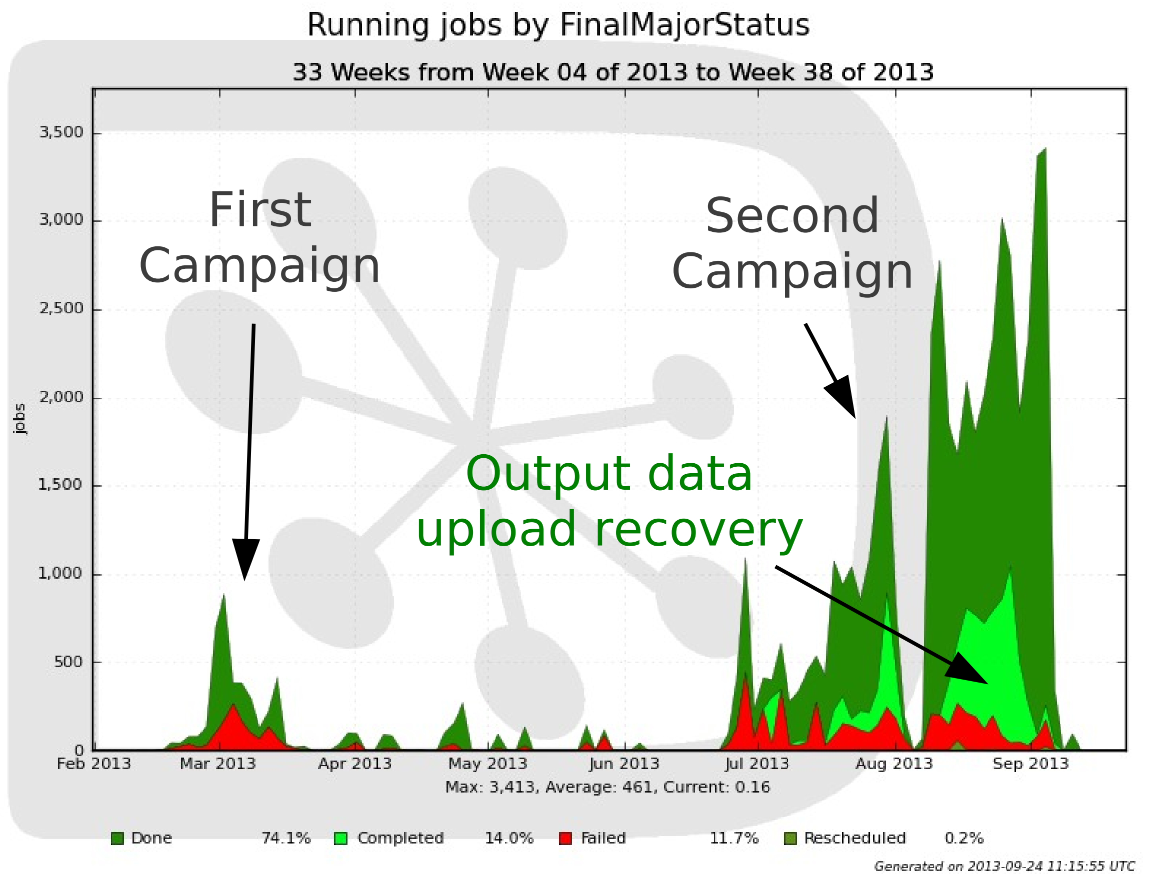
Workflow Abstraction



- Don't deal with single files and jobs, but with datasets and projects

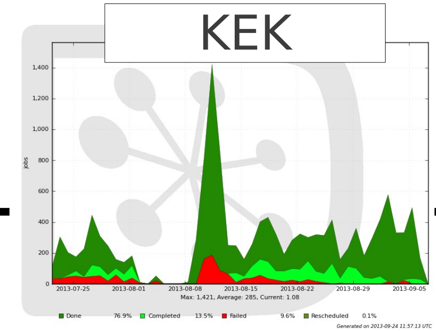
Second MC Production Campaign

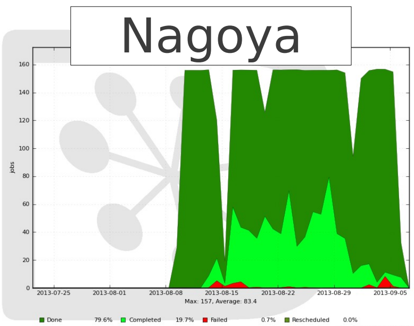
- July 23 – September 8, 2013
- Simulation and reconstruction, with background mixing → mdst data
  - 630k jobs, 700 kHS*days
  - 560M events, 8.5 TB of output data

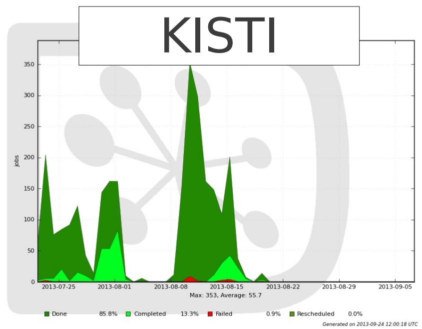


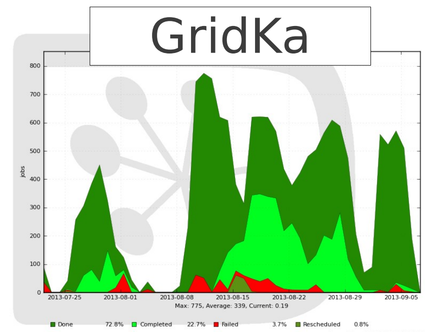
- ~10% → 1% failure rate: site configuration/ downtime, proxy expiration, server load, human errors
- No crash of offline software

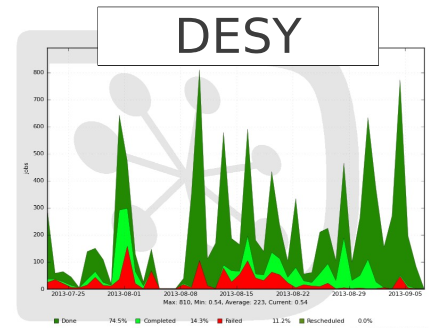
Contributing Sites

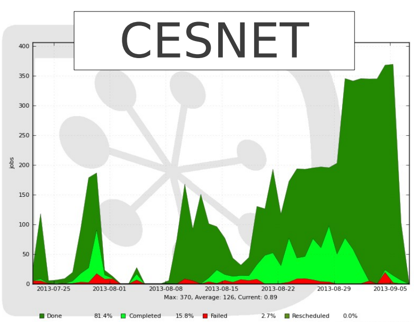


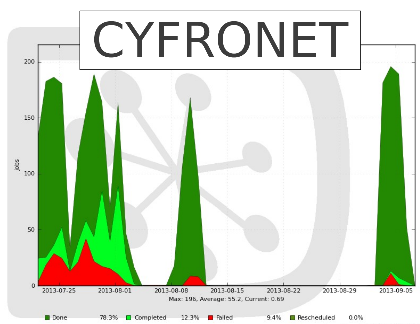


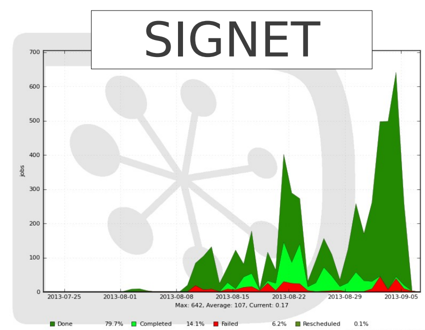


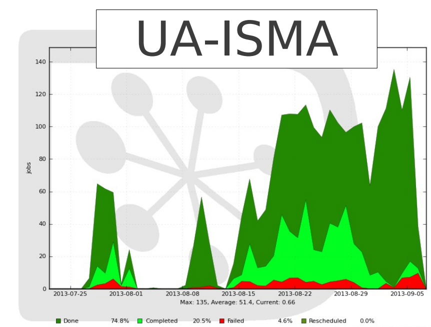


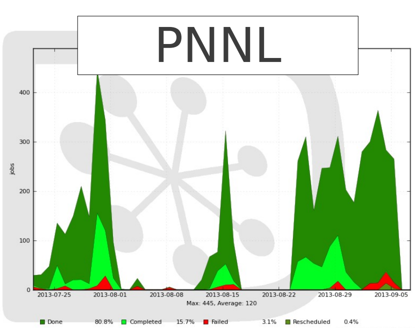


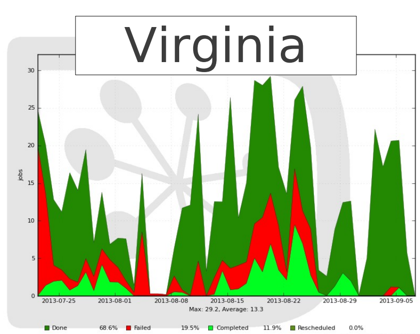


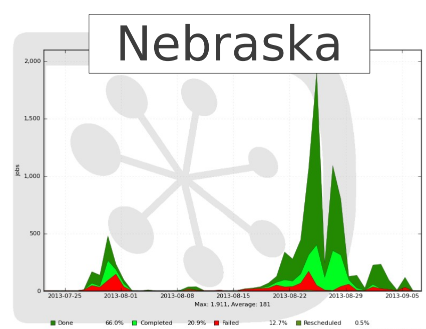


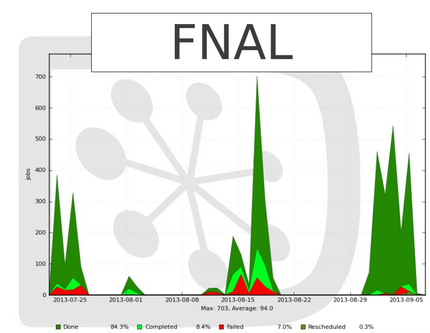




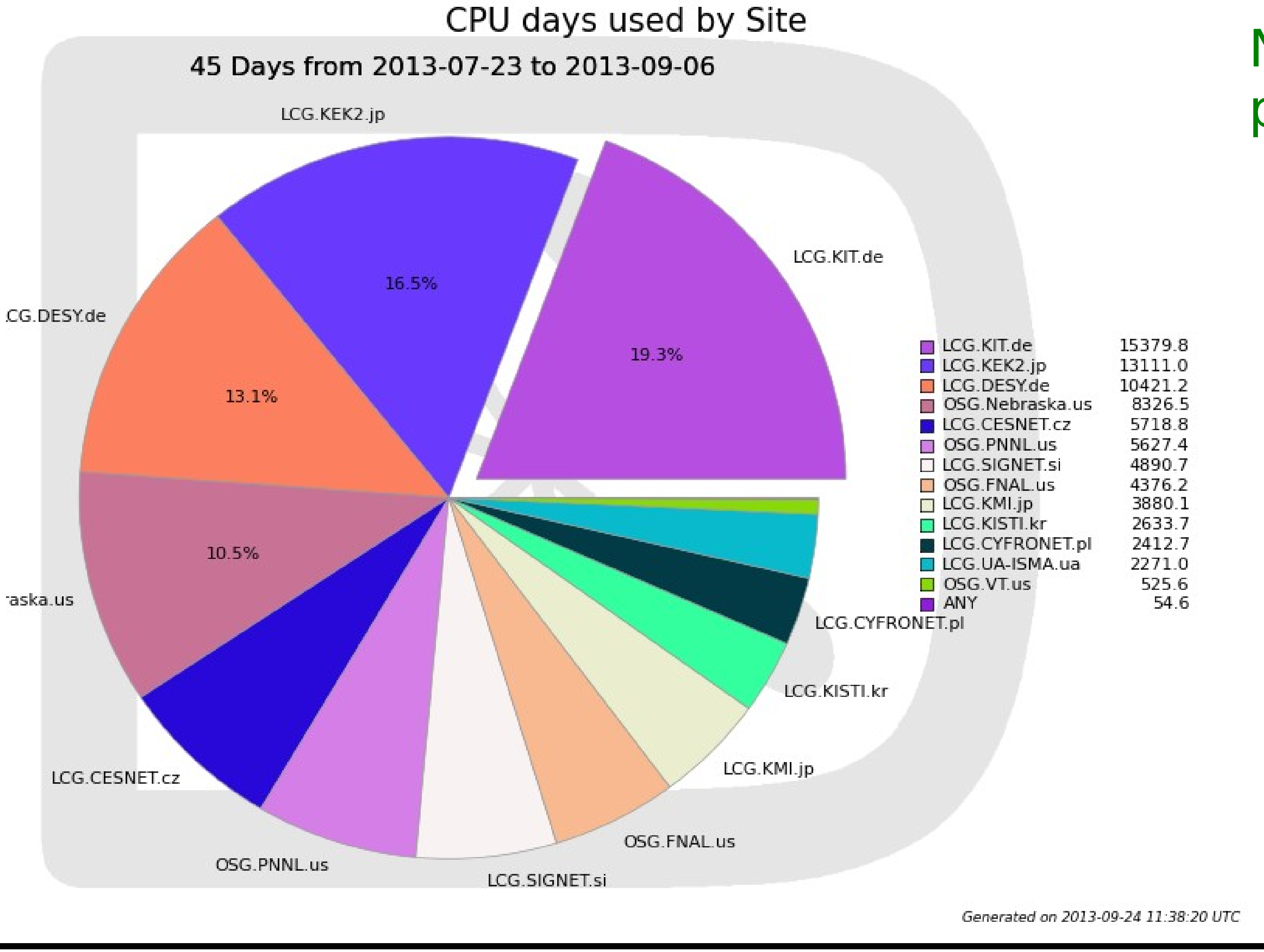








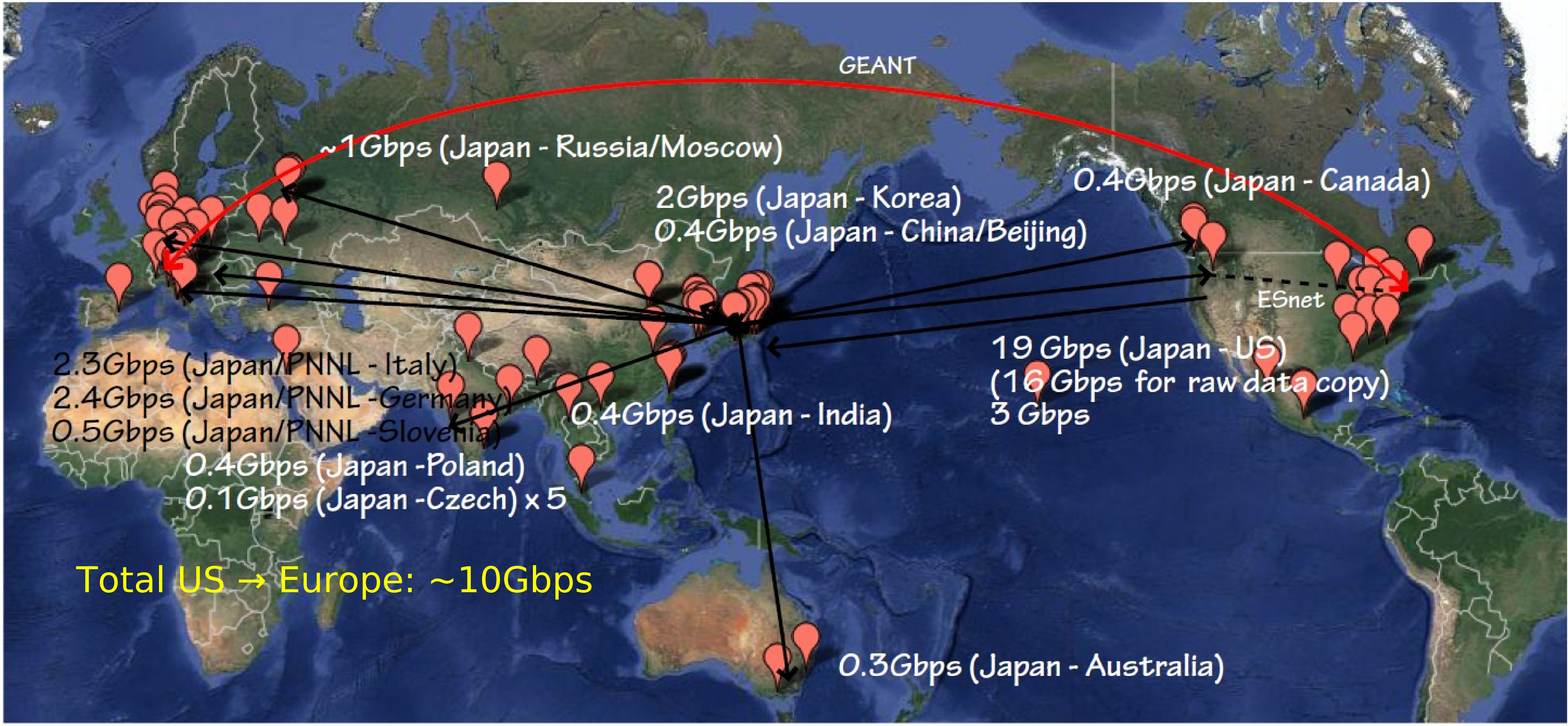
Contributing Sites



New or planned sites

- Italy
- Canada (cloud)
- Australia (cloud)
- Cracow Cloud
- Turkey
- India
- Munich

Network Bandwidth Estimates for 2022

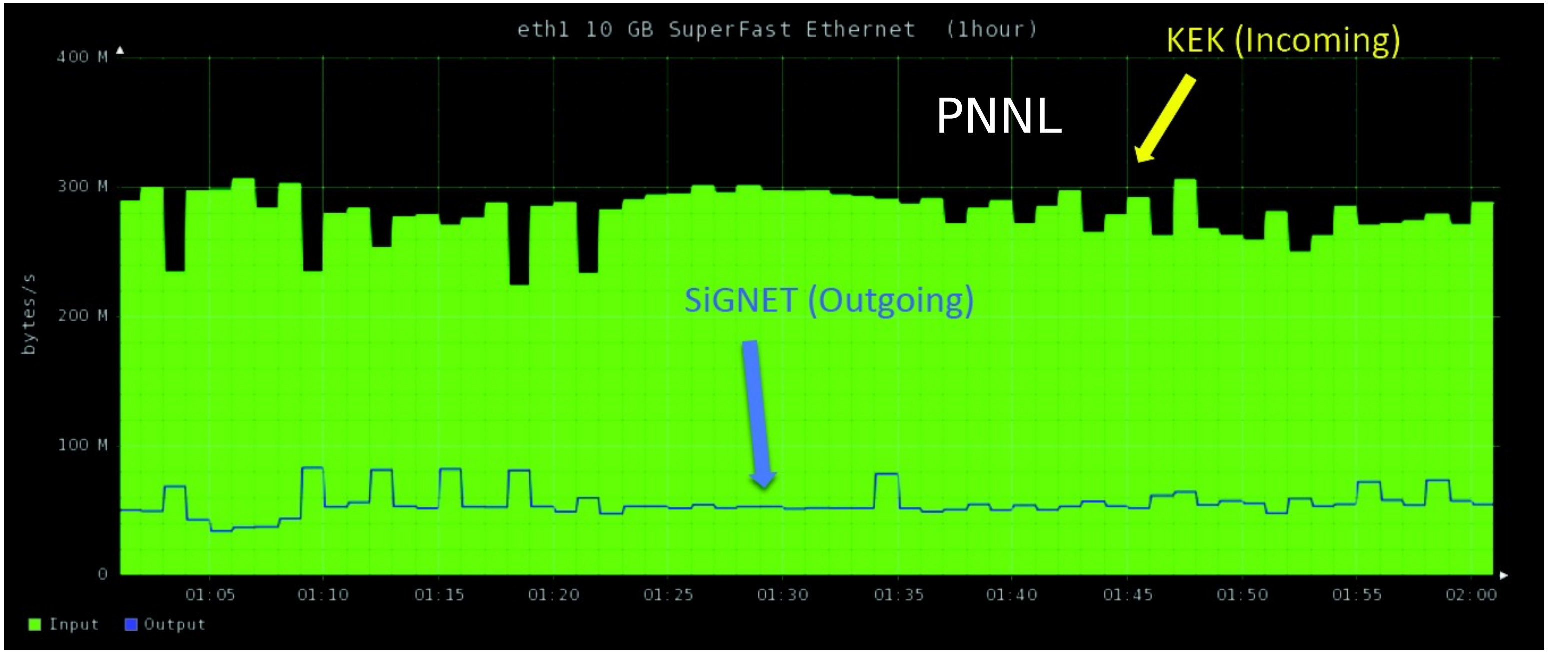


Japan/PNNL - XX : mdst transfer from Japan and/or PNNL + data transfer between XX and other sites
Japan - XX : data transfer between XX and Japan + other sites

Data Challenge

- Network connection between sites is essential
  - Raw data from KEK to PNNL
  - Mdst data and MC between sites world wide

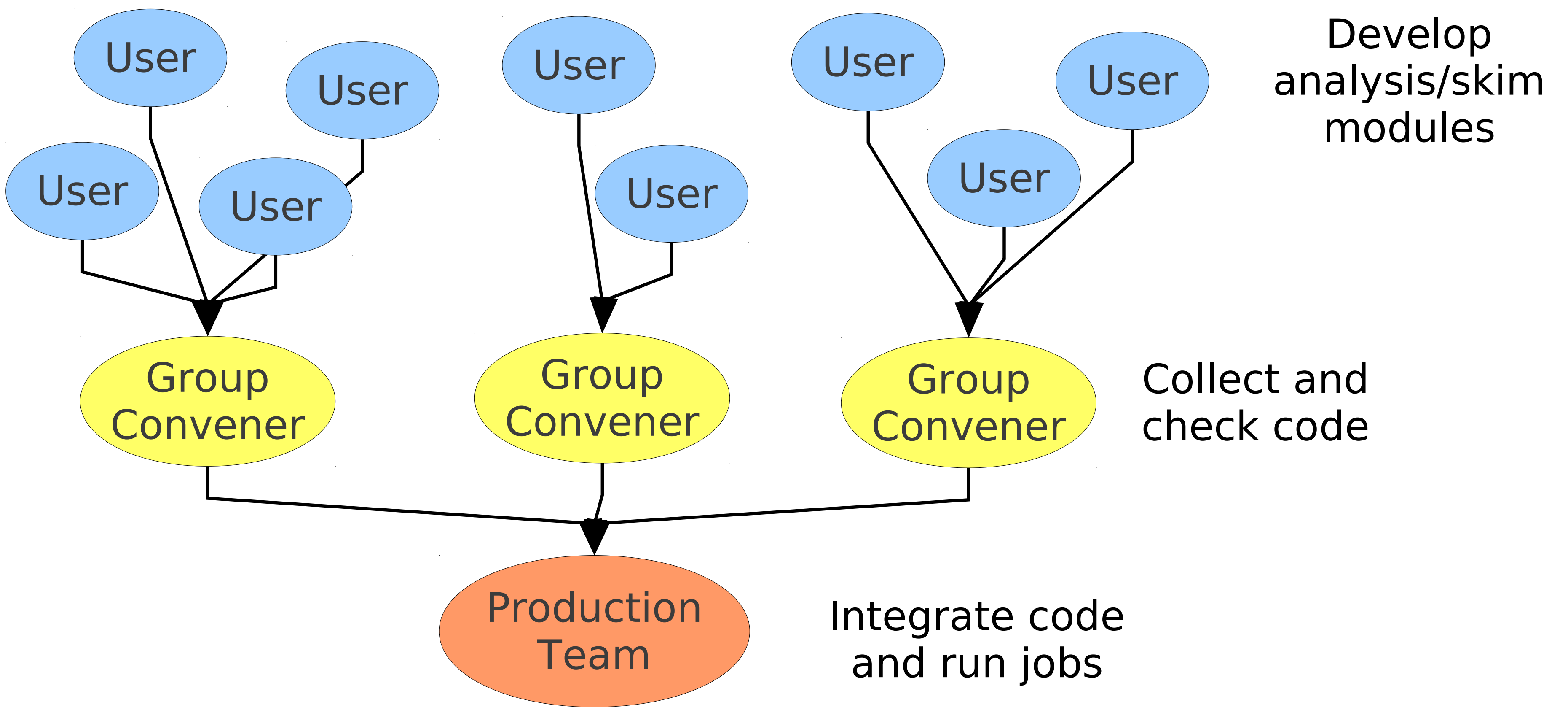
- Transfer tests between different sites in May 2013 with FTS2 server at GridKa



- Test with FTS3

Organized Analysis

- Problem: inefficient resource usage by many users
- ➔ Limit resources per user, but maintain free access to data
- ➢ Offer high-performance organized analysis as a service

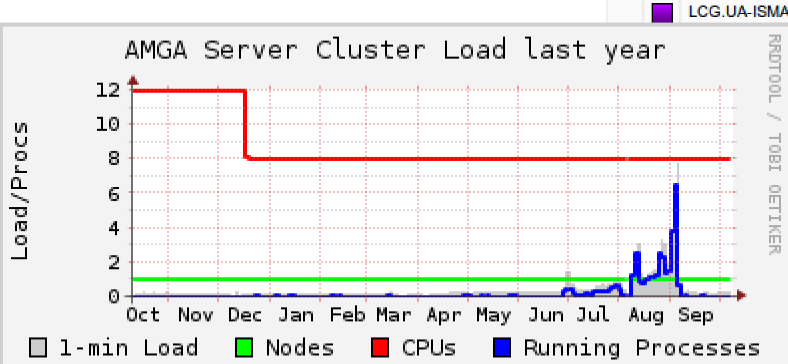


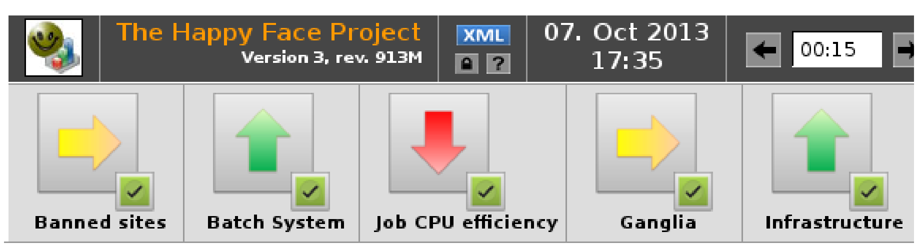
Monitoring

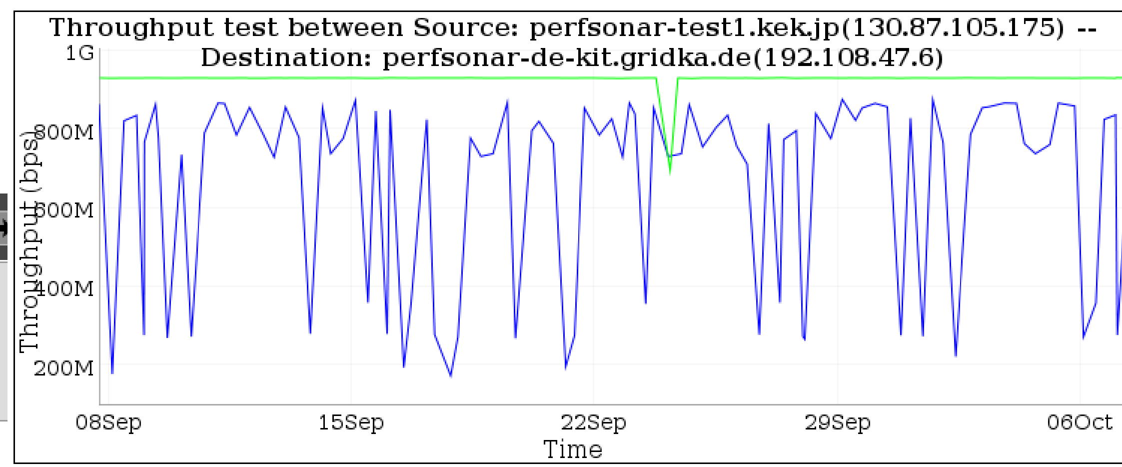
BelleII AMGA Server (belle2-amga)

Host	Status	Services	Actions
amga.pnl.gov	UP	1 OK	
belledh.kisti.re.kr	UP	1 OK	
can68.cc.kek.jp	UP	8 OK	
can69.cc.kek.jp	UP	4 OK	

Site	CE	Status	PilotJobEff (%)	PilotsPerJob	Waiting	Scheduled	Running	Done	Aborted
	Multiple	Bad	0	0	0	0	0	0	20
LCG.CESNET.cz	Multiple	Fair	62.54	1	0	0	150	42	115
LCG.CYFRON...	Multiple	Good	88.31	1	0	0	0	1133	150
LCG.DESY.de	Multiple	Fair	76.73	1	0	0	1051	225	387
LCG.KEK2.jp	Multiple	Good	85.09	1	0	155	224	1128	264
LCG.KISTI.kr	Multiple	Poor	41.68	1	0	64	16	4005	5716
LCG.KIT.de	Multiple	Good	99.29	1.01	0	2042	558	4292	50
LCG.KMI.jp	ncream01.hepl....	Good	100	1	0	80	159	1808	0
LCG.SIGNET.si	creamce.ijs.si	Poor	26.55	1	0	100	53	1	426
LCG.UA-ISMA.ua	gl-ce.isma.khar...	Bad	14.85	1.12	0	4	91	174	1543
	Multiple	Fair	74.63	1.01	0	133	864	612	1677
a....	Multiple	Poor	53.92	1	0	1484	491	916	13949
	osg.pnl.gov	Bad	21	1	0	0	0	30	15281
	valiant.phys.vt.e...	Poor	51.57	1	0	66	80	58	417

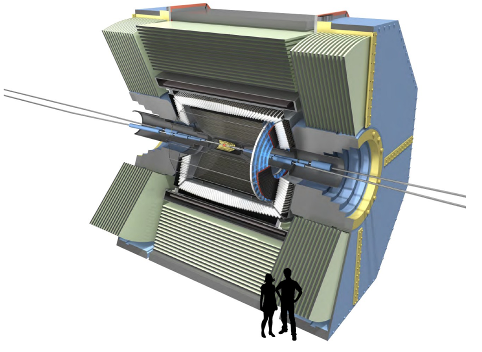






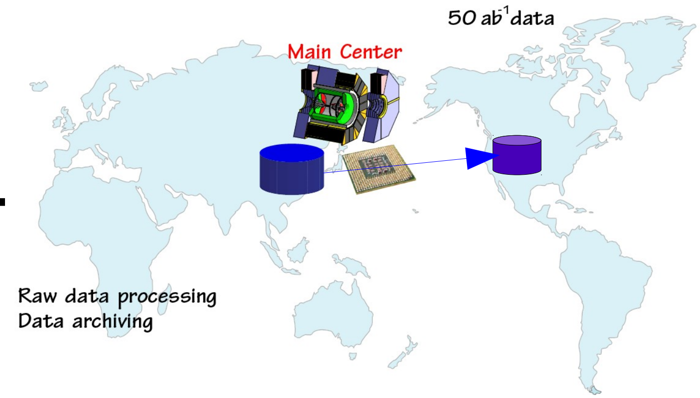
Summary

- Belle II will search for New Physics with O(50) times more data than current B factories
- Huge data volume is a challenge for the computing
  - Similar, but bit simpler computing model than WLCG
  - Distributed computing system based on existing technologies and infrastructures
  - Workflow abstraction with projects and datasets
- First two MC production campaigns this year
  - ✔ Belle II distributed computing system works!
- Next steps:
  - MC campaign with more (cloud) sites
  - Further automatize and harden the system
  - Exercise user analysis on the grid



Backup

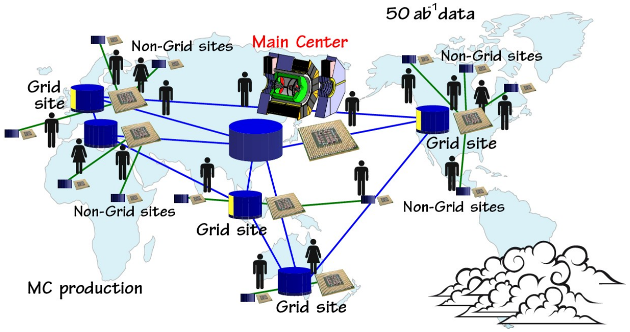
Computing Tasks



Raw data processing

- Tape as storage medium
- Store and process at KEK, replication to just one remote site

➢ Simpler than LCG model



Monte Carlo Production

- 6 times the real data size
- Produced in managed way, (almost) no input data needed

➔ Well suited for a distributed environment, including cloud

Physics Analysis

- Random, uncoordinated access → Store input data on disk

➢ Ntuple analysis on local resources for fast turn-around

Projects

```
[ccx13] ~ $ gb2_project_summary -g belle_mcprod
Project                Owner    Status    Done/Fail/ Run/Wait Submission Time(UTC) Duration
B2Kstargamma_BGx1_s1 tkuhr    Good      1000/     0/     0/     0 2013-08-14 14:41:57  06:47:32
B2Kstargamma_BGx0_s1 tkuhr    Good      1000/     0/     0/     0 2013-08-14 14:45:15  05:18:30
```

- Job submission
  - gbasf2 -r 1000 -s B2Kpi.py -p B2Kpi_s01
- Job monitoring
  - gb2_project_summary
  - gb2_project_analysis --Project B2Kpi_s01
  - gb2_job_status --Project B2Kpi_s01 --Status=failed
- Rescheduling of failed jobs
  - gb2_job_reschedule --Project B2Kpi_s01
- Job output
  - gb2_job_output --Project B2Kpi_s01 --Status=failed

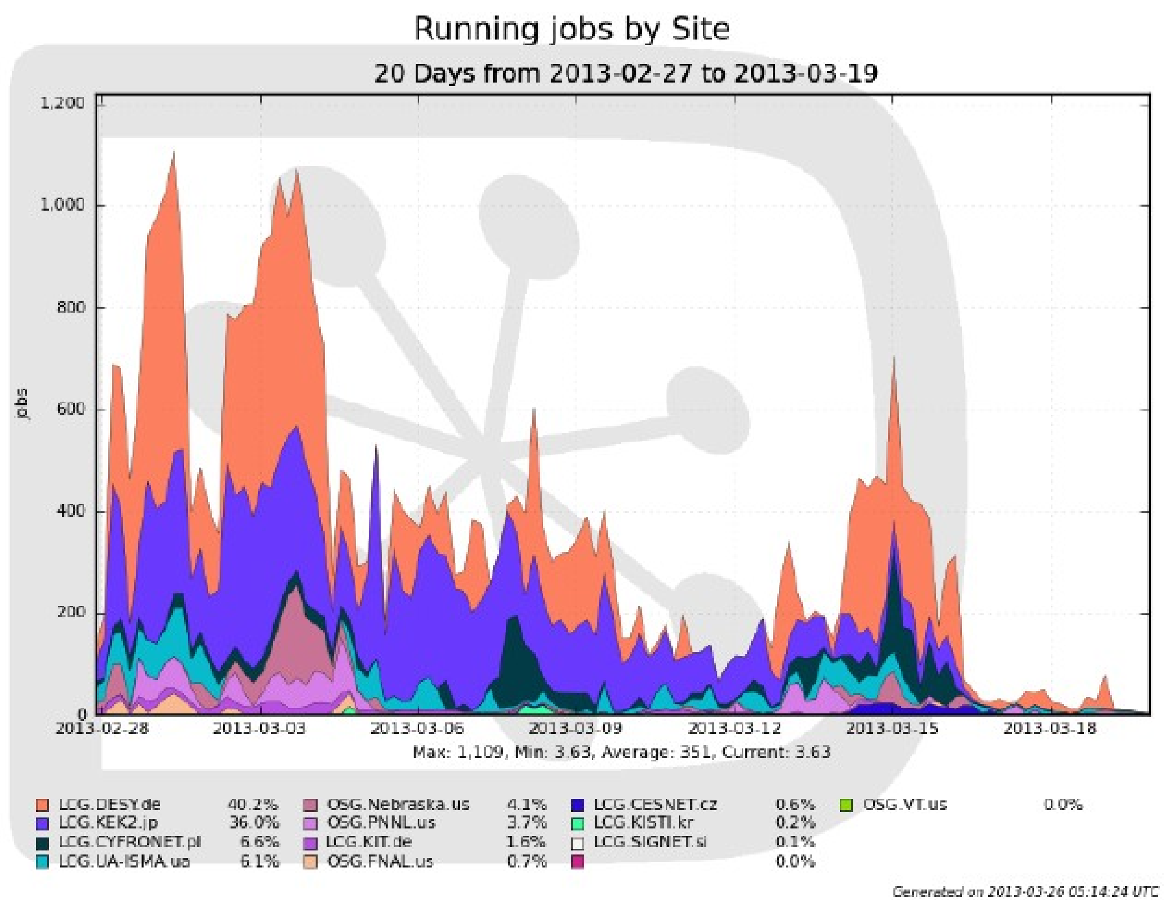
```
[ccx13] ~ $ gb2_project_analysis --Project testneb_b1
100 jobs are selected.

Project testneb_b1 summary:
Done (60)
  Execution Complete (60)
    Done (60)
       OSG.Nebraska.us:  60
Completed (25)
  Pending Requests (25)
    Done (25)
       OSG.Nebraska.us:  25
Failed (15)
  Application Finished With Errors (6)
    Exit Status 1 (6)
       OSG.Nebraska.us:   6
  Job stalled: pilot not running (9)
    Preparing to upload (1)
       OSG.Nebraska.us:   1
    Registering (1)
       OSG.Nebraska.us:   1
    Running (3)
       OSG.Nebraska.us:   3
    Selecting SE (1)
       OSG.Nebraska.us:   1
    Uploading (3)
       OSG.Nebraska.us:   3
```

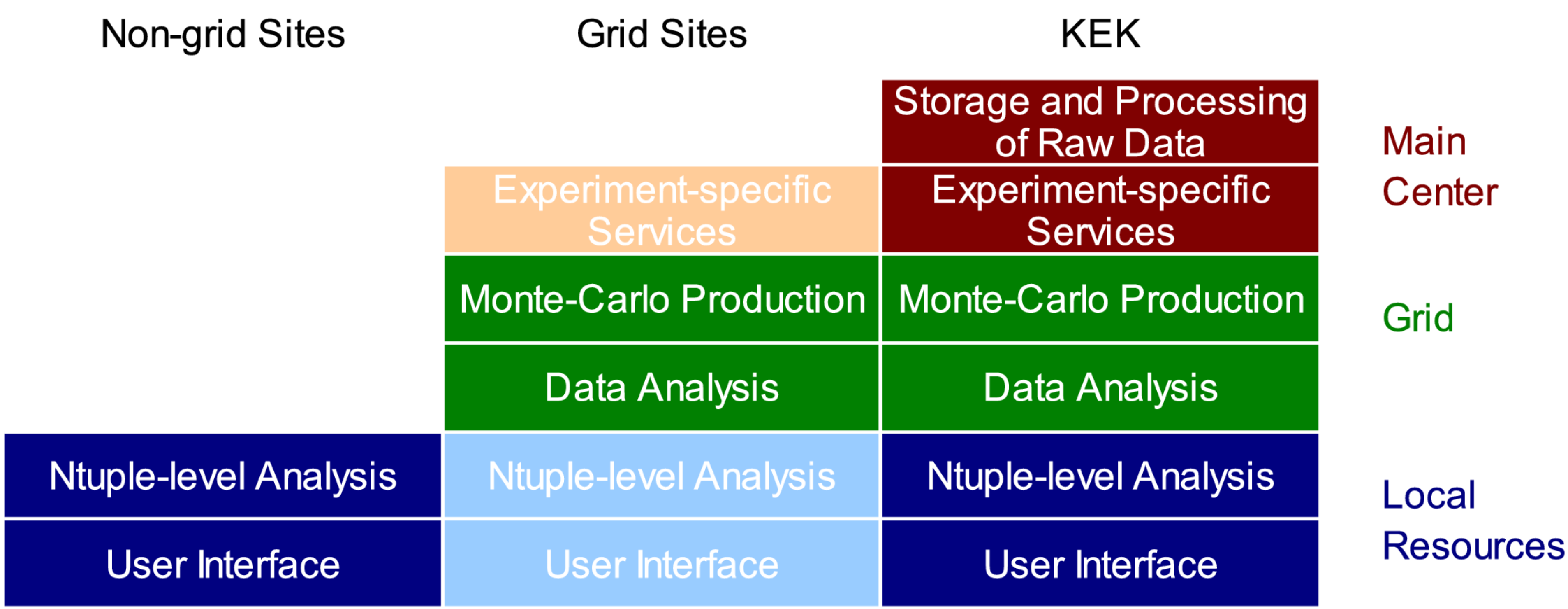
First MC Production Campaign

- February 28 – March 19, 2013
- 1st stage: event generation and detector simulation
  → raw data
- 2nd stage:
  reconstruction
  - 240k jobs,
    40 kHS*days
  - 60M events,
    190 TB of output data

- ~20% failure rate:
  metadata registration,
  input data download,
  application errors



Tasks of Computing Facilities



(Commercial) Cloud Computing

- Resource demands vary with time
- Fair-share can solve this issue only to some extent

➔ Cloud computing allows to buy resources on demand
  - Well suited to absorb peaks in varying resource demand