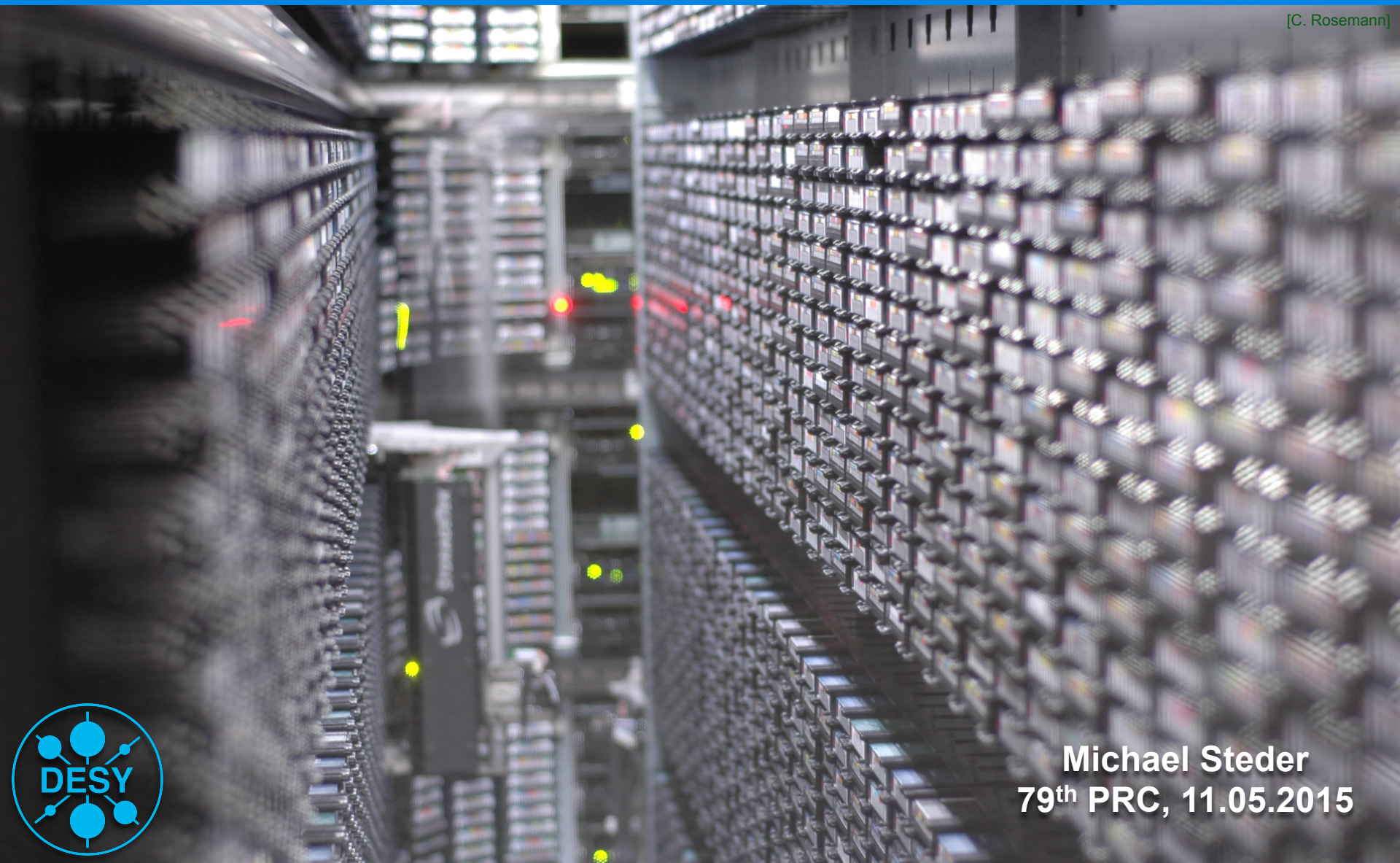
DESY Computing.



[C. Rosemann]

Michael Steder

79th PRC, 11.05.2015

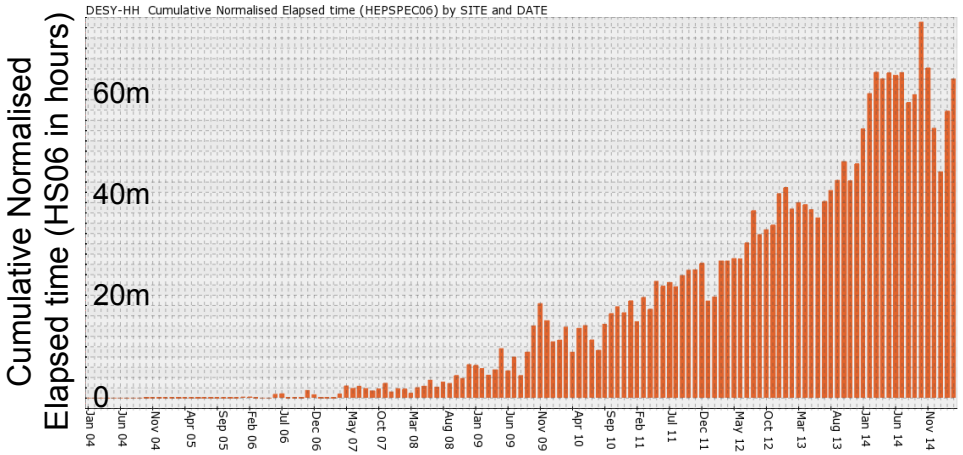
Outline

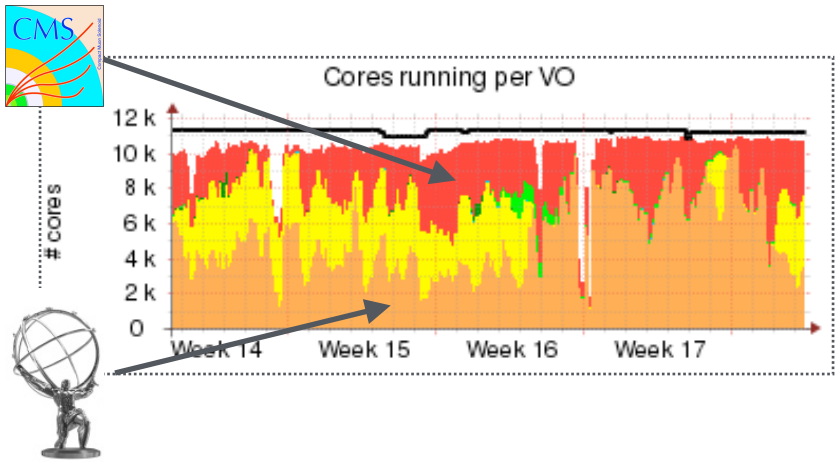
[UBA]

- „Bread & Butter“
  - NAF 2.0 & Grid
- Belle II Computing (at DESY)
- Horizon2020/Cloud Projects

Grid@DESY (since 2004)

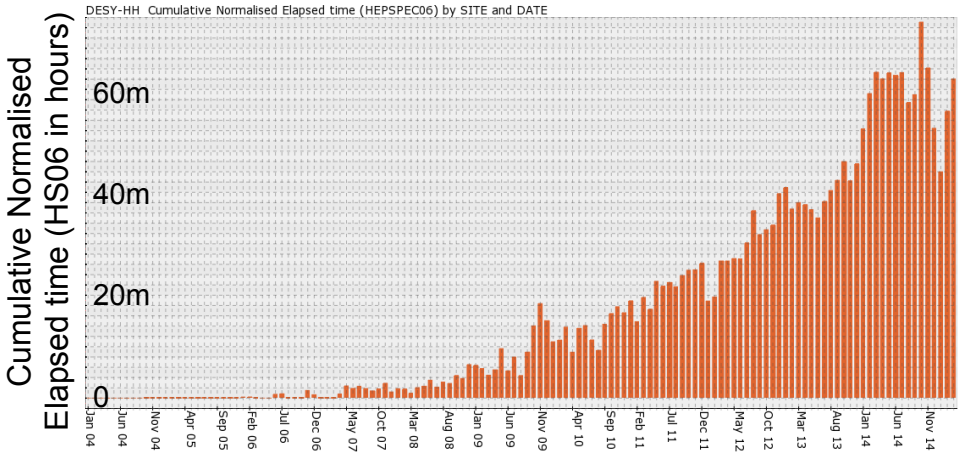
- Home of 10 VO's
  - Tier-0/1: ILC & Tier-2: Atlas, CMS, LHCb, Belle2
  - Opportunistic model on federated ressources
- Multi-core jobs since 1/2015
- dCache-based SE's
  - 10 PB disk & >10PB tape
  - on latest or „golden" release respectively

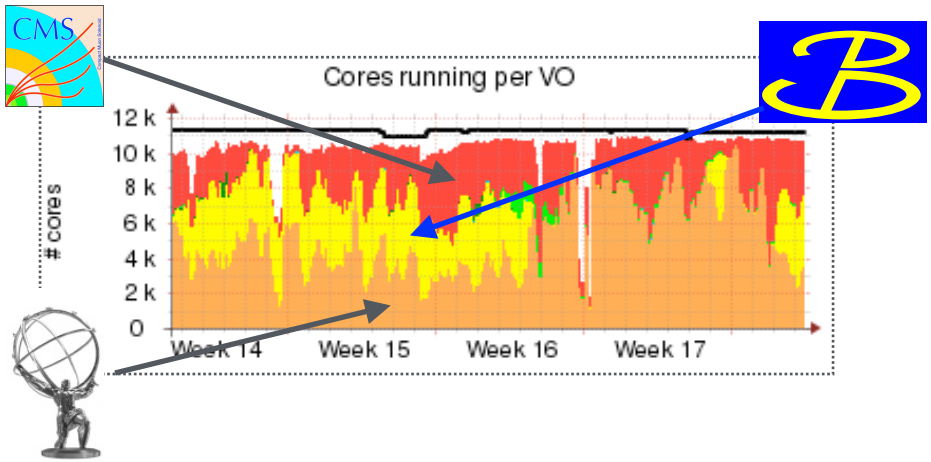




Grid@DESY (since 2004)

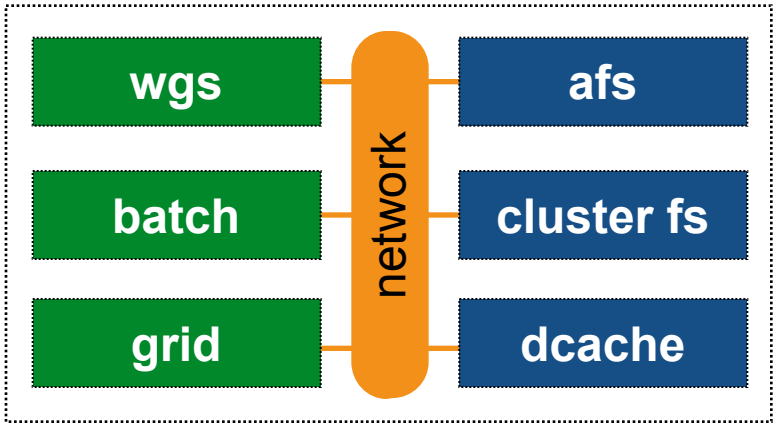
- Home of 10 VO's
  - Tier-0/1: ILC & Tier-2: Atlas, CMS, LHCb, Belle2
  - Opportunistic model on federated ressources
- Multi-core jobs since 1/2015
- dCache-based SE's
  - 10 PB disk & >10PB tape
  - on latest or „golden" release respectively

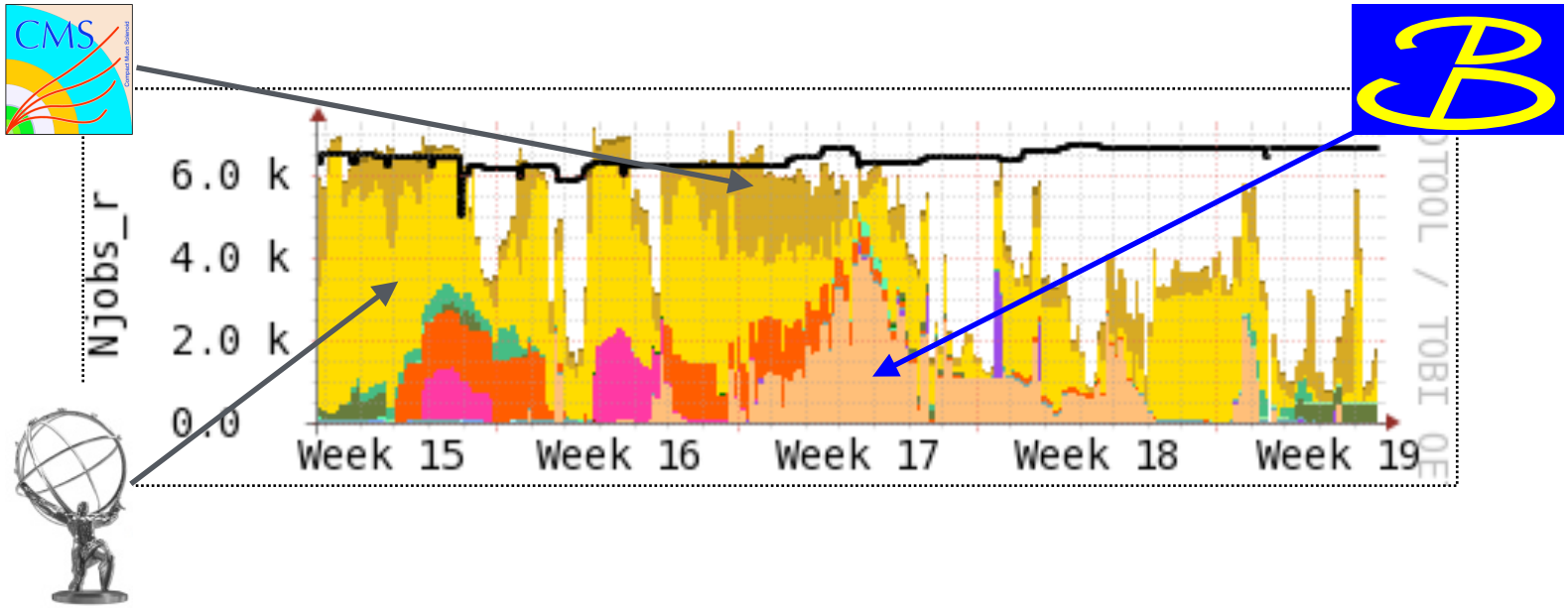




National Analysis Facility (since 2007)

- NAF 2.0
  - Complementary to Grid
  - Opportunistic usage
  - Federated Ressources
- Since 2015 HERA experiments

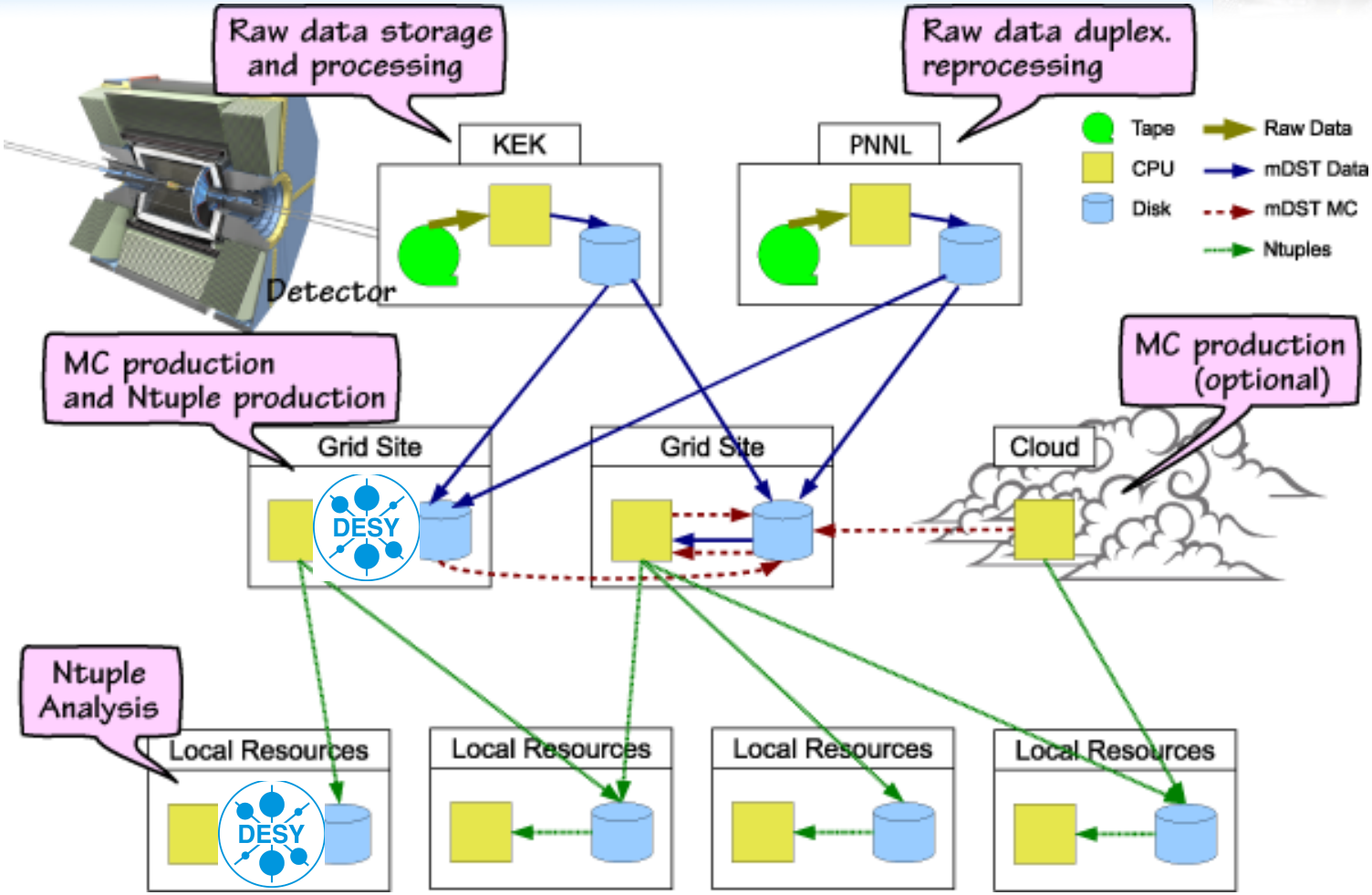




> Physics at Belle II
→ T. Ferber (after lunch)

Belle II Computing Model

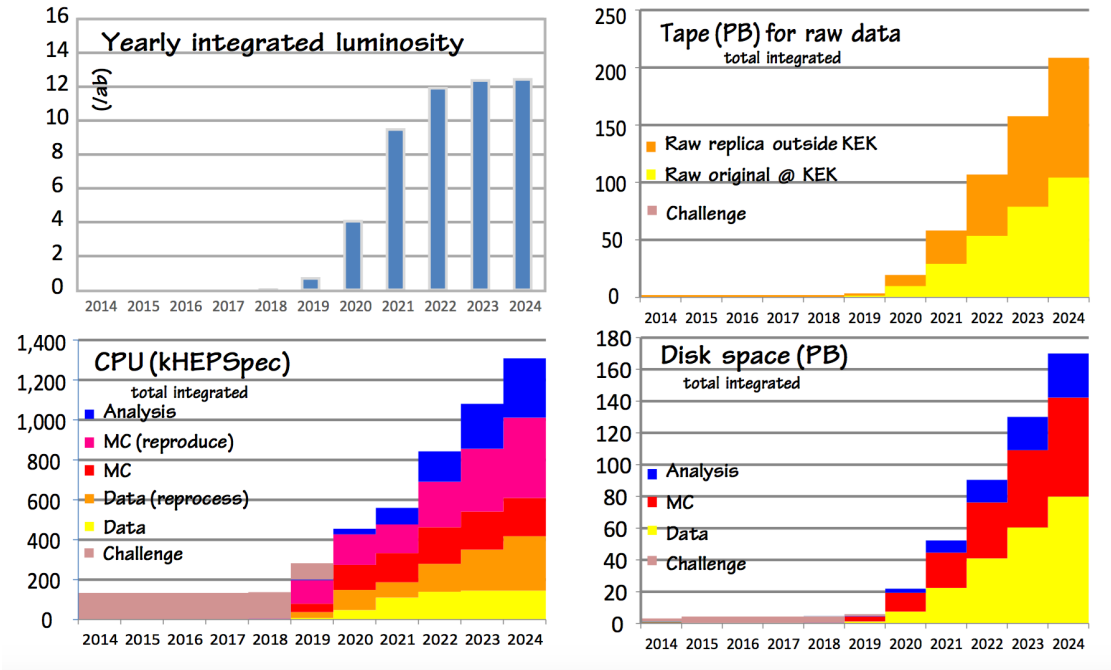
[T. Hara, KEK]



> DESY: Tier2 / MC Production Site

Belle II Computing Requirements

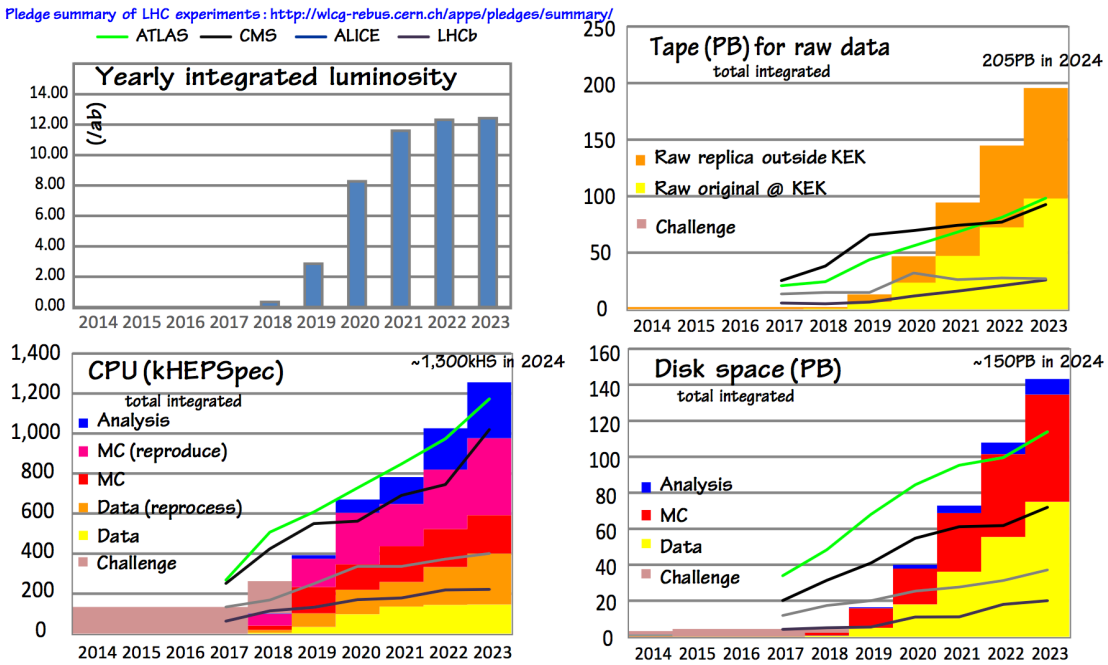
- First Collisions
  - Mid of 2017
- First Physics
  - End of 2018



[T. Hara, KEK]

Belle II Computing Requirements

- First Collisions
  - Mid of 2017
- First Physics
  - End of 2018
- Similar or higher demand wrt LHC
  - Small event size
  - Enormous rate



[T. Hara, KEK]

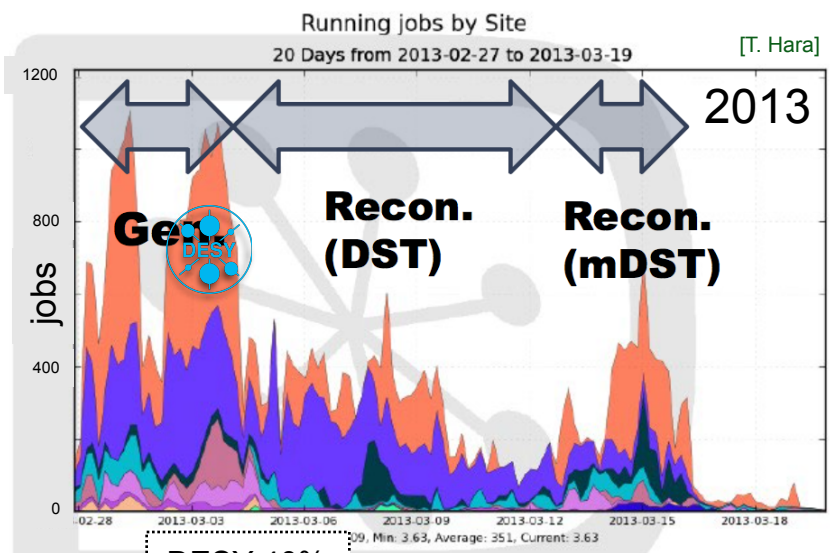
Experiment	Event Size [kB]	DAQ rate [Evt/s]	DAQ rate [kB/s]
Belle II (max. Lumi)	300	6.000	1.800
Alice (PbPb)	50.000	100	4.000
Alice (pp)	2.000	100	200
ATLAS	~700	600	400
CMS	~1.000	500	several 100s
LHCb	55	4.500	250

LHC experiments as seen in 2011/2012 runs

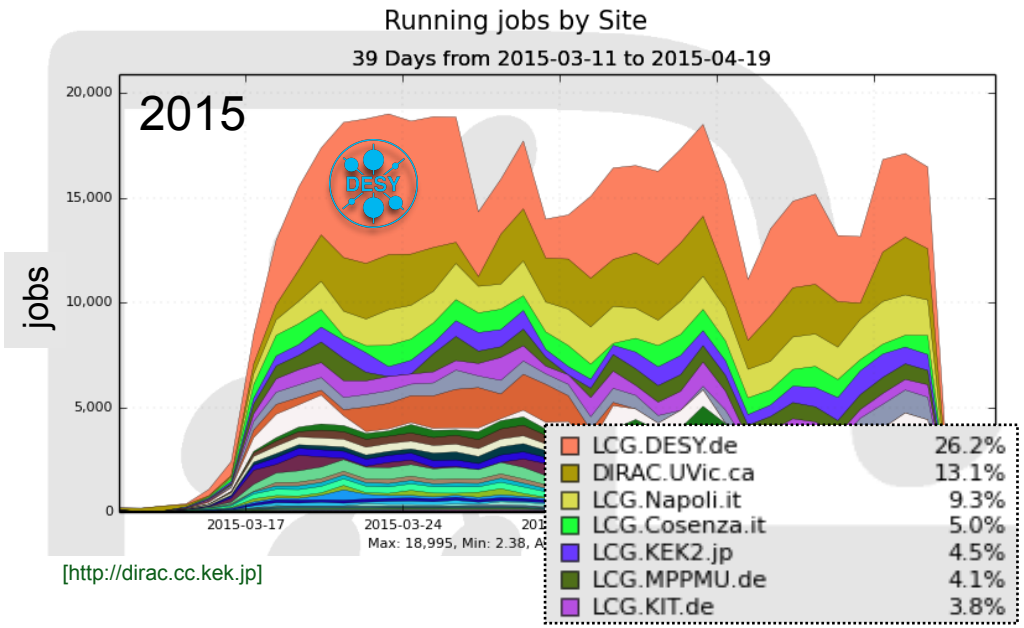
Another Challenge for HEP Computing

- 1st MC campaign in March 2013
- 5 campaigns until now
  - Latest: ongoing

- DESY contribution acknowledged by Belle II Collaboration
  - About ¼ - ⅓ of overall kHS
  - Owing to opportunistic usage (and LHC shutdown)

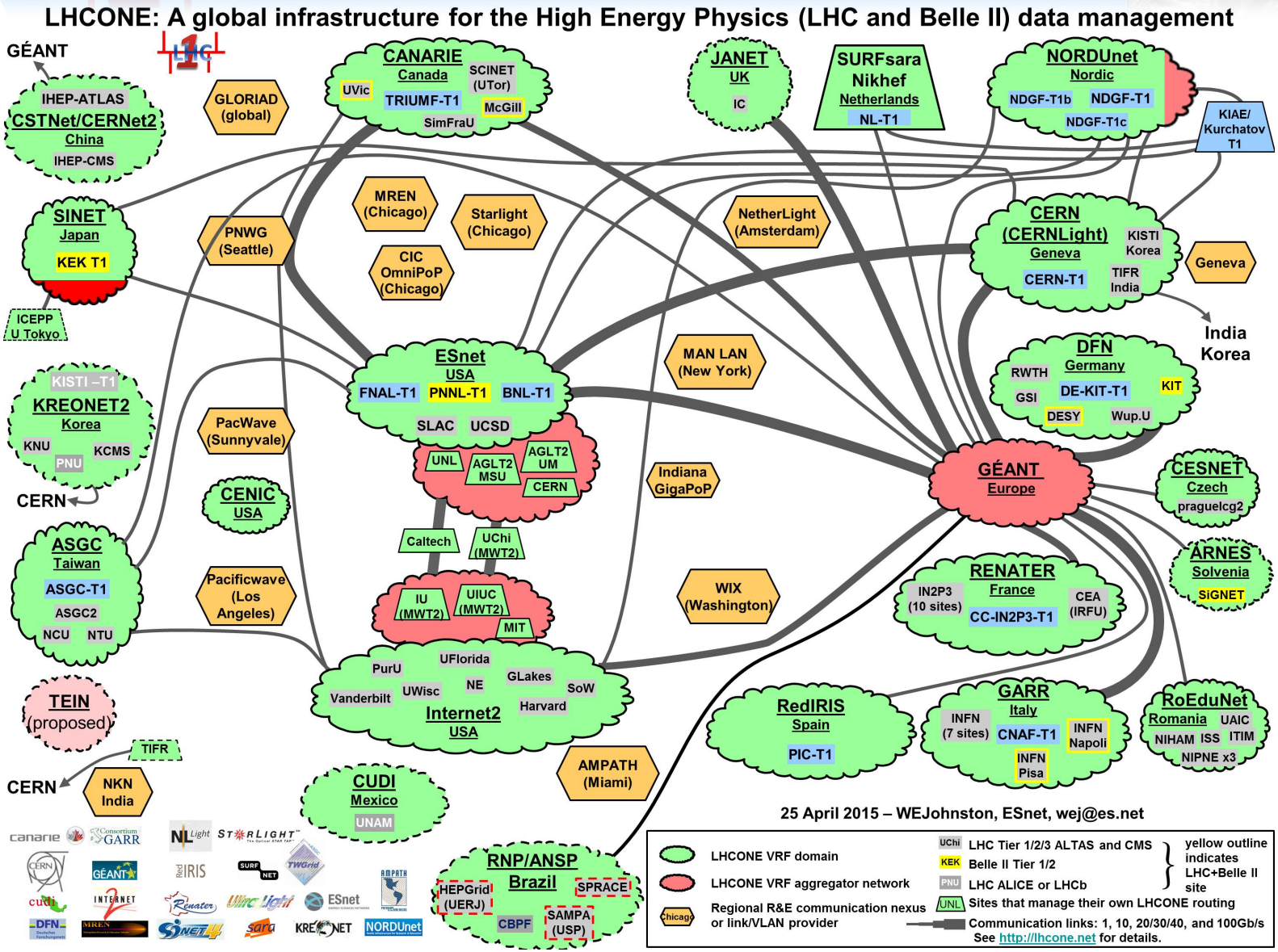


DESY 40%
KEK 36%



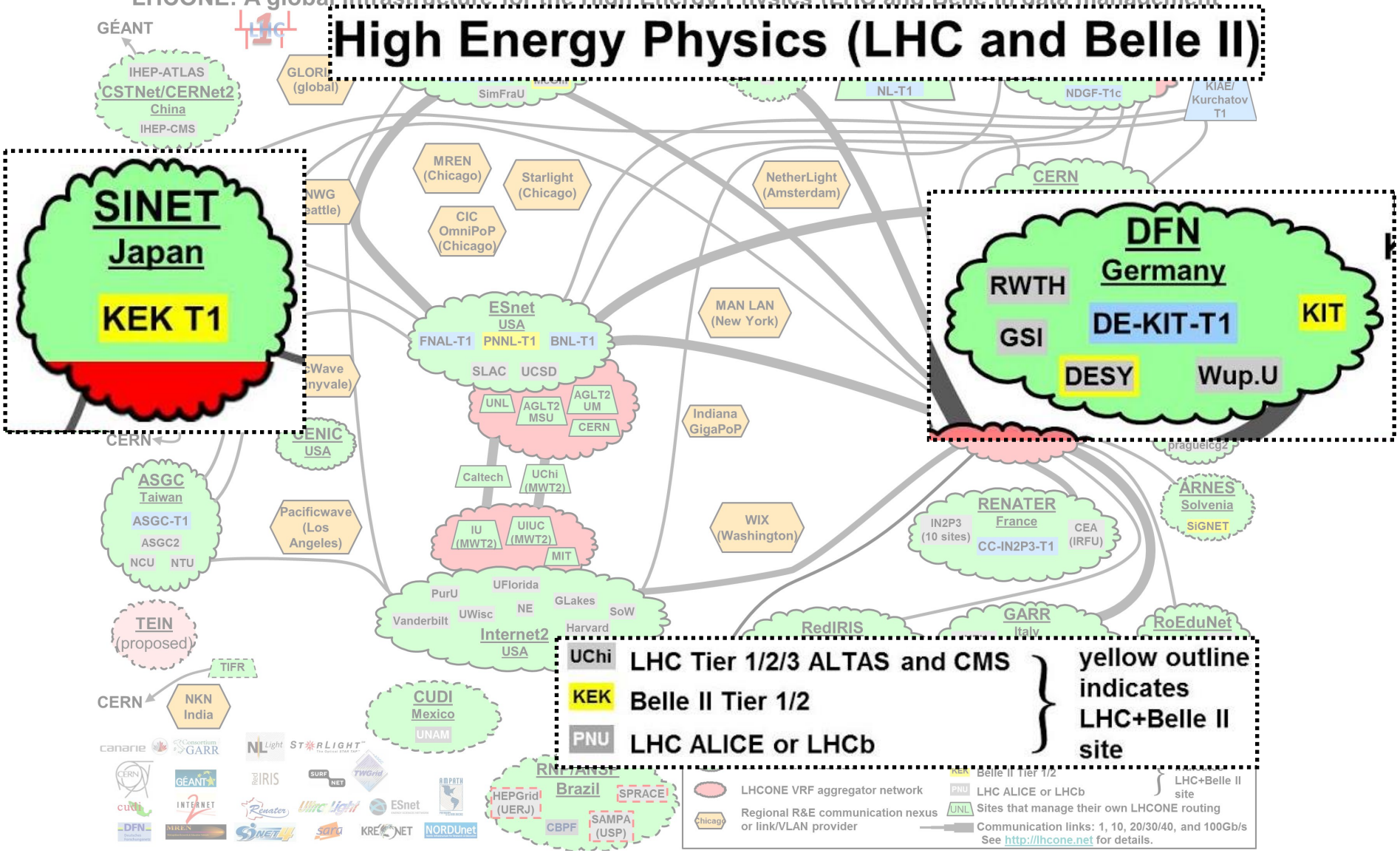
[http://dirac.cc.kek.jp]

A Network for HEP



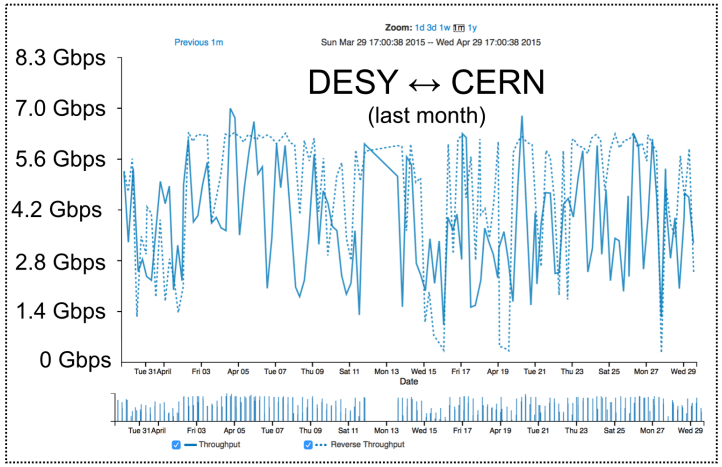
A Network for HEP

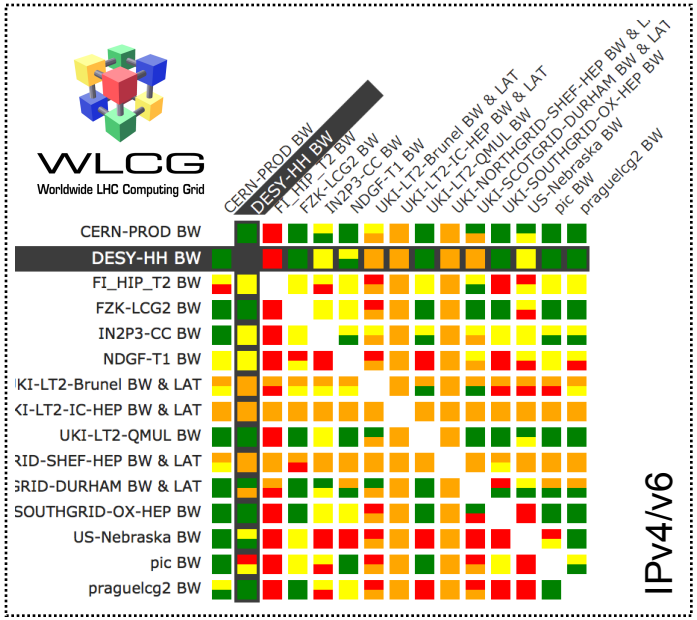
LHCONE: A global infrastructure for the High Energy Physics (LHC and Belle II) data management

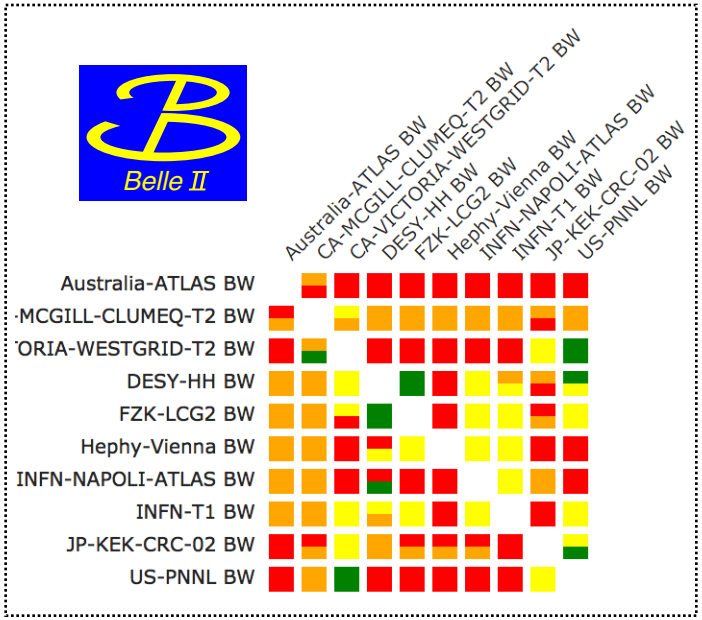


Still some way to go…

- Automated Network Performance Monitoring
- Schedule for Data Challenges
  - Trans-Atlantic / -Pacific
  - Last May/June 2014





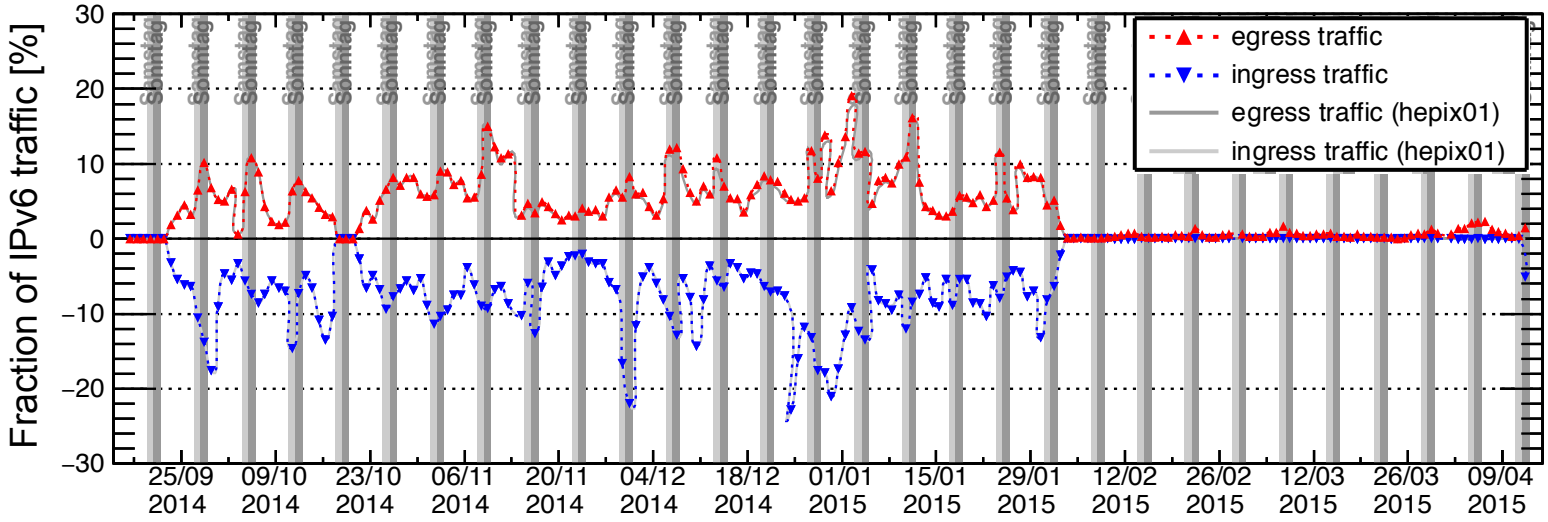


Throughput >= 900Mbps | Throughput < 900Mbps | Throughput <= 500Mbps | Unable to retrieve data | Check has not yet run

[http://maddash.aglt2.org/maddash-webui/index.cgi?dashboard=Dual-Stack%20Mesh%20Config]

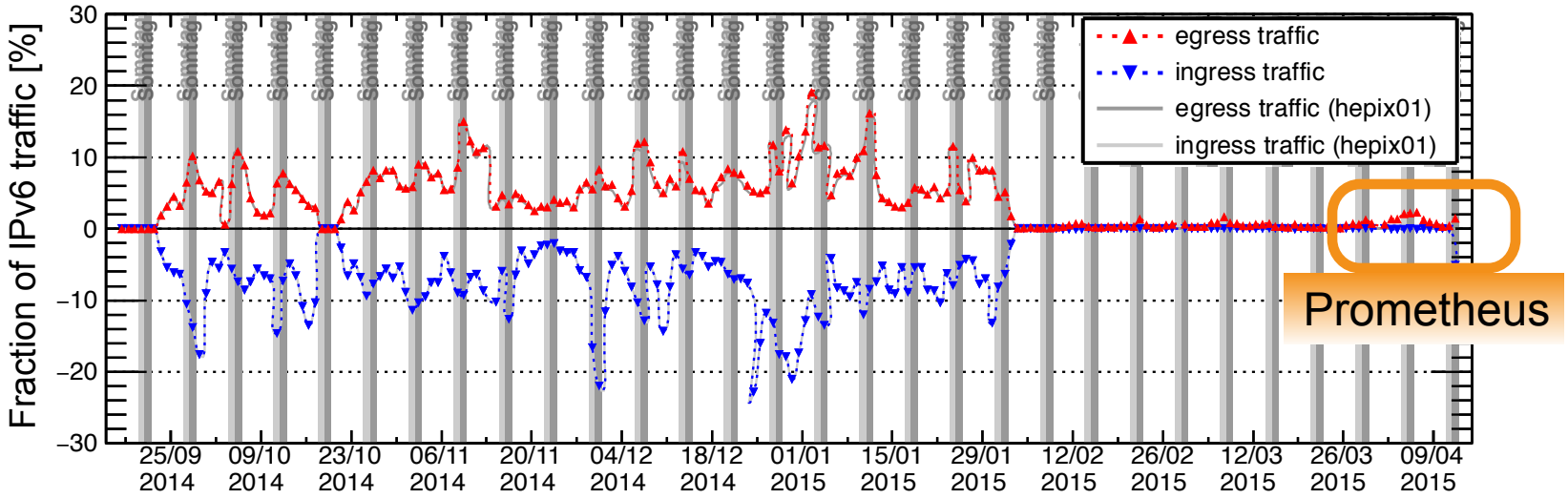
Make your choice — 10^{38} addresses

- Member of HEPIX IPv6 Group
  - founded April 2011
- Central DESY IT services are IPv6 enable
- DESY IPv6 traffic still negligible
  - „Only“ test-bed traffic



Make your choice — 10^{38} addresses

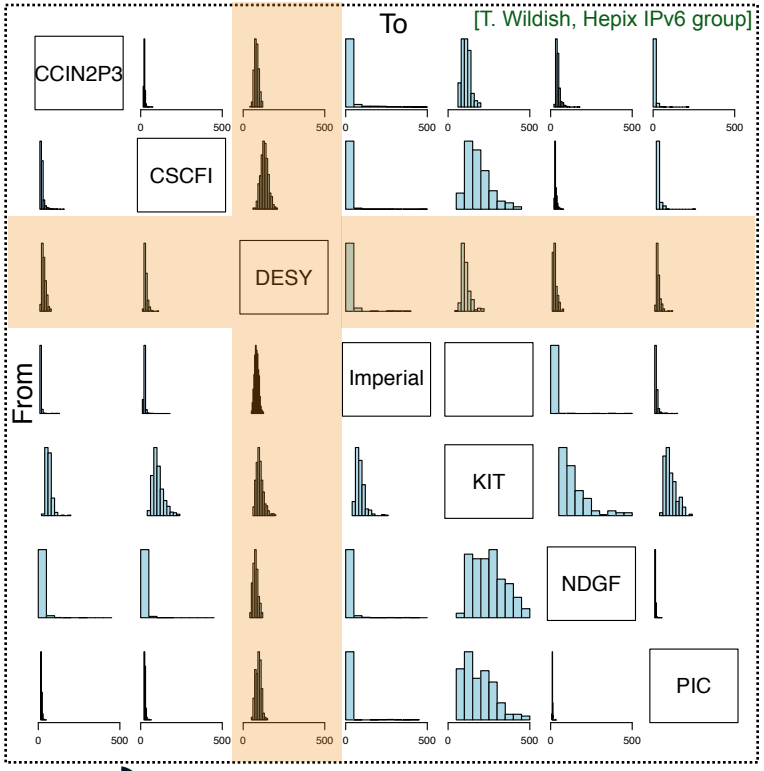
- Member of HEPIX IPv6 Group
  - founded April 2011
- Central DESY IT services are IPv6 enable
- DESY IPv6 traffic still negligible
  - „Only“ test-bed traffic



Make your choice — 10^{38} addresses (2)

- DESY joined FTS3 IPv6 test-bed
  - Just in time for CHEP `15

- prometheus.desy.de
  - running latest dCache version
  - functional testing of
    direct web access (http(s)), WebDAV(s), xrootd, dcap, gsidcap, gsiftp



FTS3 Time to Transfer [s]

Be ahead of the times!

- HEPIX IPv6 group
  - Survey and test IPv6 compliance
- Steady progress

CERN Accelerating science Sign in Directory

HEPiX IPv6 Working Group

Home | WLCG applications | WLCG IPv6 Task Force | Testbed | Knowledge base | perfSONAR probes | Contact

Home

IPv6 compliance of WLCG applications

Software Component	Type	Used by Experiment	Version	IPv6 Compliance
AliEN	LHC Experiment Application	ALICE		
ARC CE	Middleware	ATLAS, CMS		YES
ARGUS	Middleware	ALICE, ATLAS, CMS, LHCb		Unknown
BDII	Middleware	ATLAS, CMS, LHCb	EMI 2	YES
BestMAN	Middleware	ATLAS, CMS		YES
CASTOR	Middleware	ALICE, ATLAS, CMS, LHCb		NO

Be ahead of the times!

- HEPIX IPv6 group
  - Survey and test IPv6 compliance
- Steady progress, but not yet done…

CERN Accelerating science Sign in Directory

HEPiX IPv6 Working Group

Home | WLCG applications | WLCG IPv6 Task Force | Testbed | Knowledge base | perfSONAR probes | Contact

Home

IPv6 compliance of WLCG applications

Software Component	Type	Used by Experiment	Version	IPv6 Compliance
AliEN	LHC Experiment Application	ALICE		
ARC CE	Middleware	ATLAS, CMS		YES
ARGUS	Middleware	ALICE, ATLAS, CMS, LHCb		Unknown
BDII	Middleware	ATLAS, CMS, LHCb	EMI 2	YES
BestMAN	Middleware	ATLAS, CMS		YES
CASTOR	Middleware	ALICE, ATLAS, CMS, LHCb		NO

The Indigo DataCloud Project

- **IN**tegrating **D**istributed data **I**nfrastructures for **G**lobal Expl**O**itation
  - Coordination: INFN

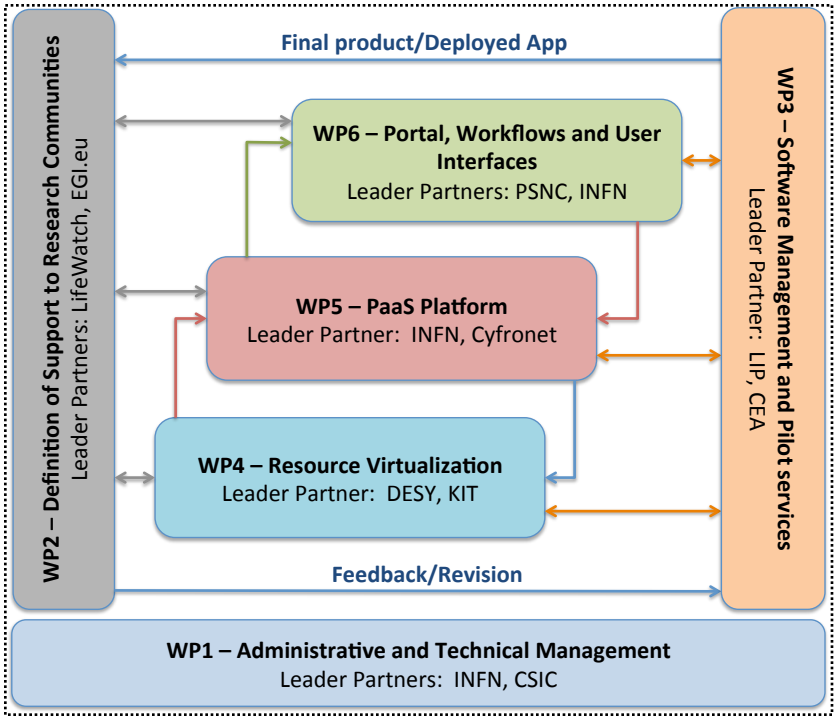
- Pan-european Project
  - 22 institutes
    - Biological and Medical Science
    - Social science, arts and humanities
    - Environmental and earth science
    - Physical science
  - 4 industrial partners

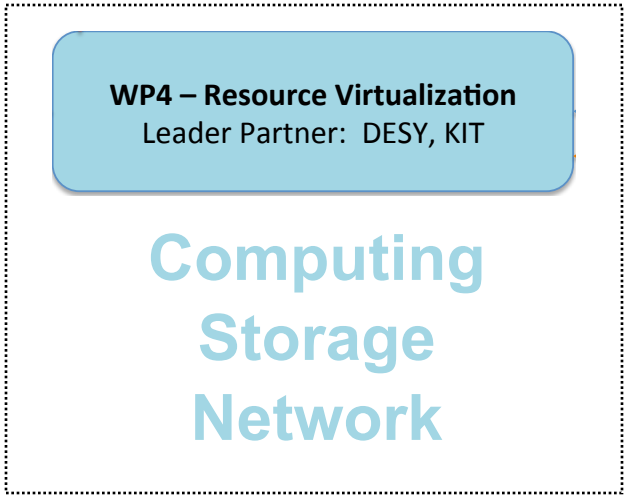
INDIGO DataCloud

- Approved within ‚Horizon2020‘
  - „EINFRA-1“ call
  - 11M€ funding (duration 30 months)

The Indigo DataCloud Project

- Standardised and reliable access to Cloud services
- DESY & KIT: Leading „WP4 - Resource Virtualisation“
  - IaaS — „Infrastructure as a Service“





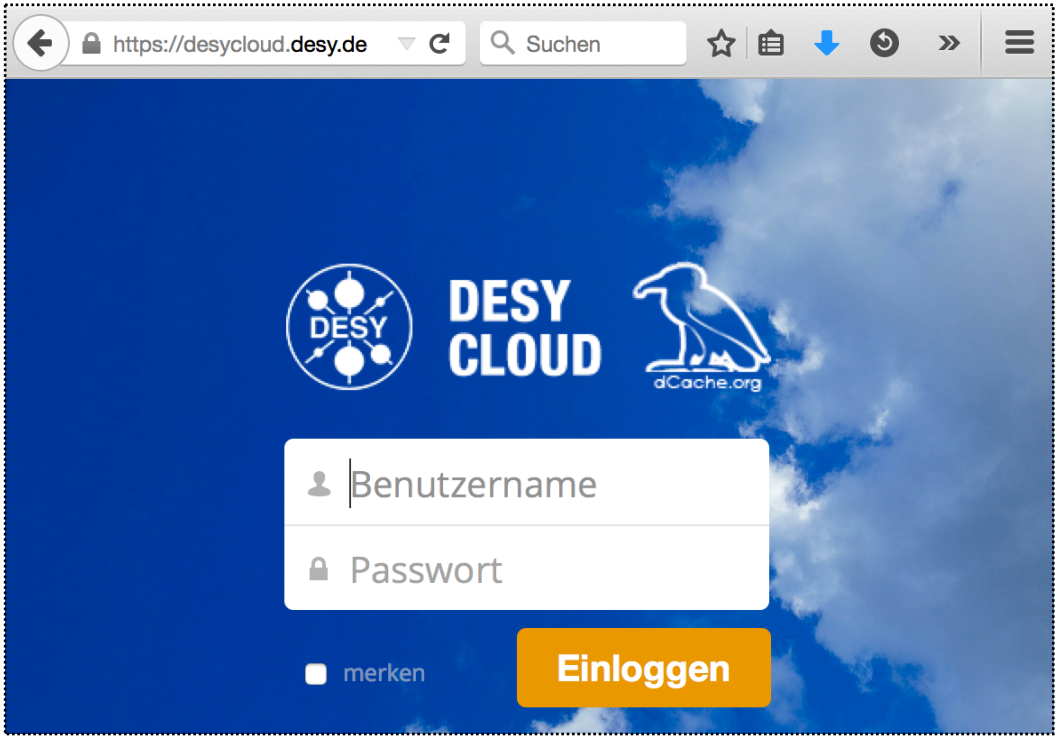
Cloud Storage - an incomplete list

Cloud Storage - an incomplete list

> „Free“ and convenient, unfortunately not compliant at all!

Talk is cheap…

- Provide an alternative cloud storage!
  - Easy to use
  - Hosted at DESY
  - Integrated in DESY infrastructure



Talk is cheap…

- Storage
  - Powerful managed storage
    combination of SSDs, HDDs and tapes
  - No syncing or sharing option

dCache.org

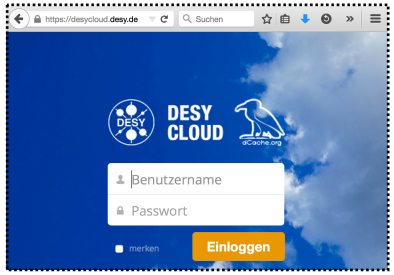
⊕

- Front-end
  - Open Source project
    Assumes managed storage

ownCloud

=

- dCache ⊕ ownCloud → DESYCloud
  - final stage of pilot operation
    219 users, 2×10^6 files, 2.4TiB



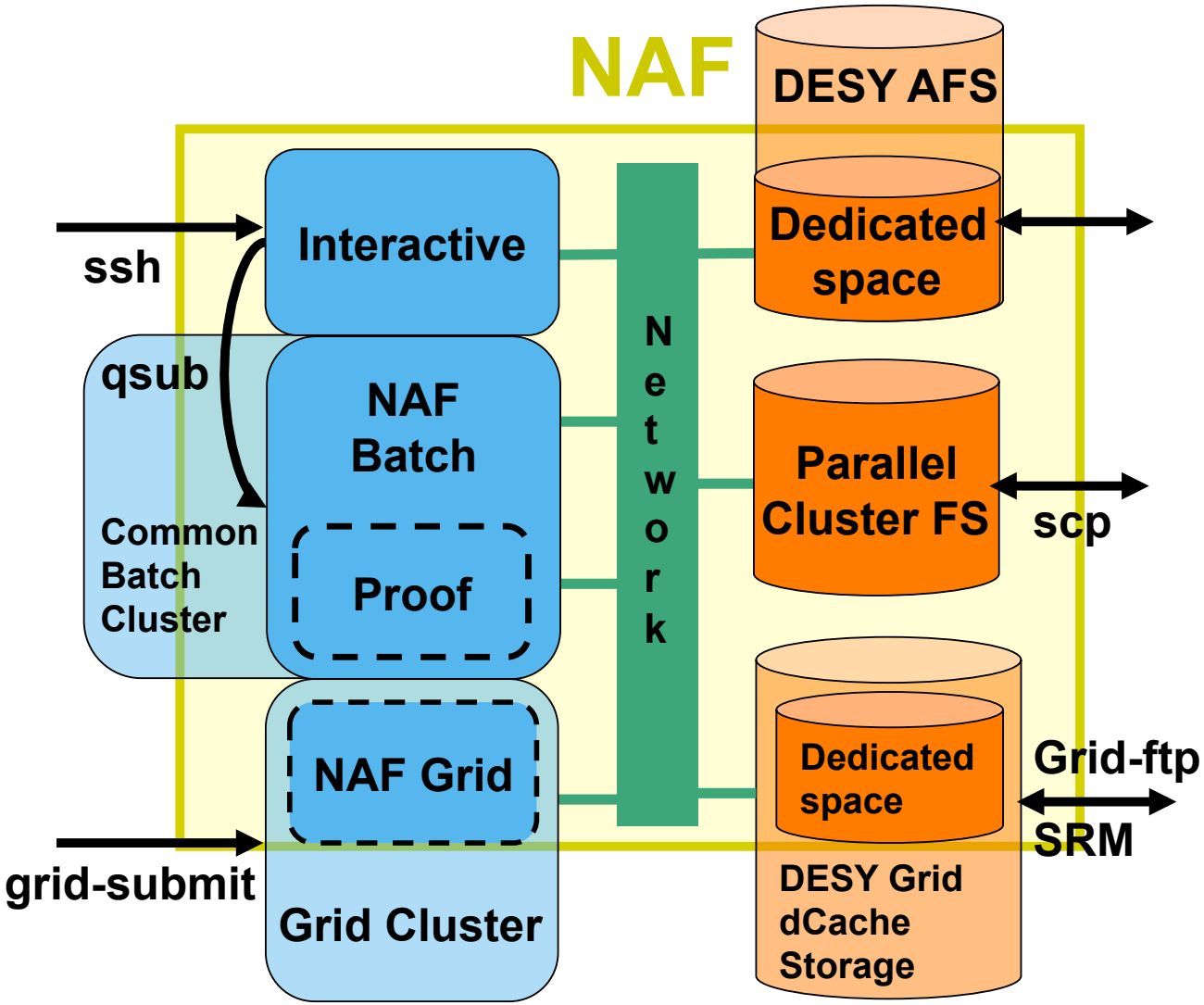
Summary.

- NAF 2.0 & Grid (continue to) run smoothly
- Make user's life as convenient as possible
  - DESY cloud, Atlassian Tools, …
- Contribute to pan-european projects
  - HEPIX IPv6, Cloud Computing Projects
- Be ready for challenges to come
  - End of LHC LS1, Start-up of Belle II, Growing PhotonSciences demands, …

BACKUP

"Jim was told that he could back up his data by making an image of his computer."

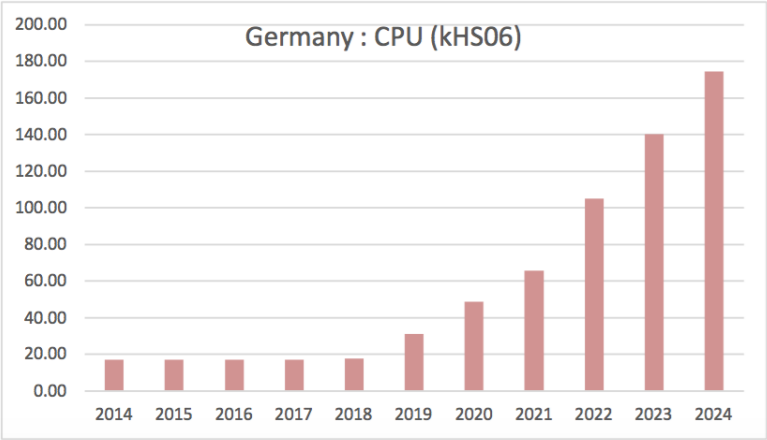
NAF 2.0

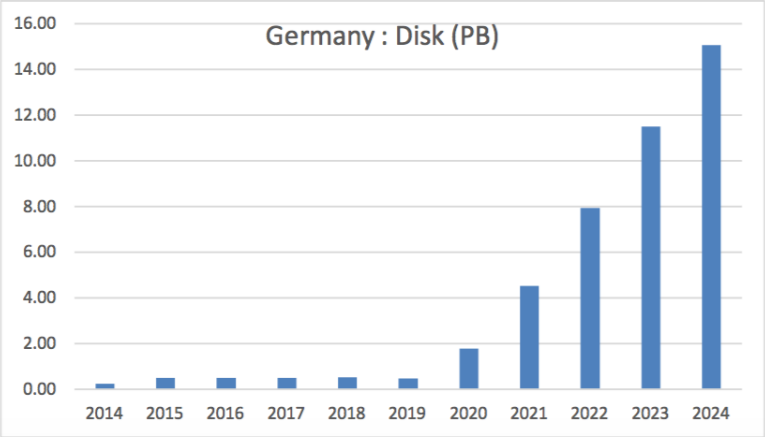

NAF
DESY AFS
Dedicated space
ssh
Interactive
qsub
NAF Batch
Common Batch Cluster
Proof
Network
Parallel Cluster FS
scp
NAF Grid
Dedicated space
Grid-ftp
SRM
grid-submit
Grid Cluster
DESY Grid dCache Storage

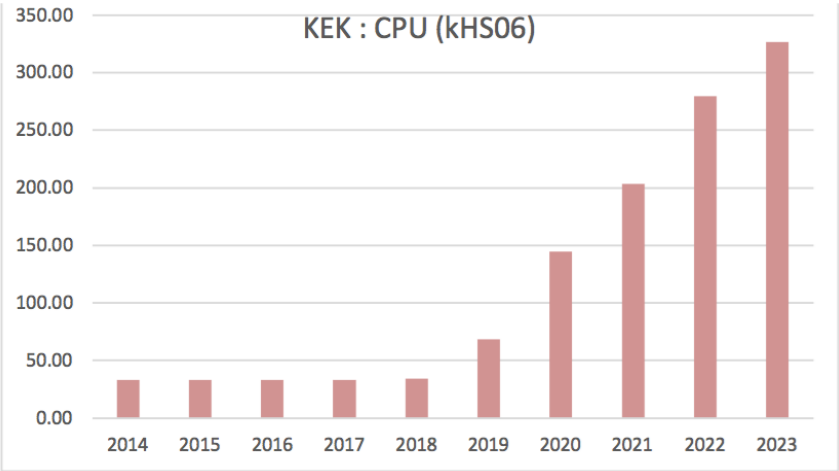
Site or Tier 1 fraction	Germany	This can be a site name as given on the Sites sheet or a numbe
Data faction		If this is empty the data fraction will be equal to the number given on the Sites sheet or 0 for numberical Tie
Used Tier 1 facrion	0.13	*The fraction of MC production and user analysis done at the site*
Used data fraction	0.40	*The fraction of mDST data stored at the site*

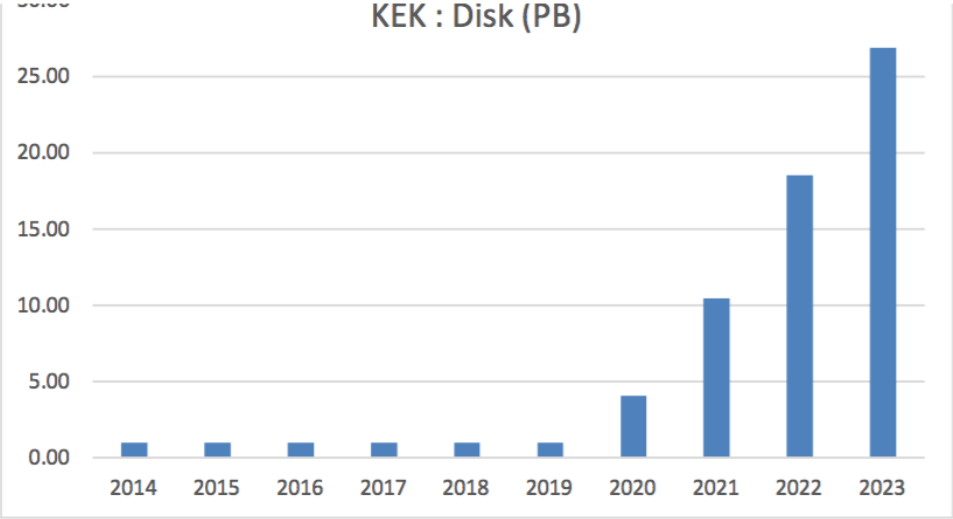
Regional Center resources

Year	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024
Tape (raw data) [PB]	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.88	8.75	13.81	18.90
Disk (data) [PB]	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.10	0.69	2.07	3.80	5.60	7.41
Disk (MC) [PB]	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.20	0.77	1.44	2.29	3.17	4.05
Disk (analysis) [PB]	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.34	1.01	1.85	2.73	3.61
Disk (challenges) [PB]	0.00	0.13	0.26	0.52	0.52	0.52	0.52	0.13	0.00	0.00	0.00	0.00	0.00
Disk [PB]	0.00	0.13	0.26	0.52	0.52	0.52	0.53	0.48	1.80	4.52	7.95	11.50	15.07
CPU (raw pro) [kHepSPEC]	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8.85	11.13	11.55	11.61
CPU (raw repro) [kHepSPEC]	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	10.14	22.89	36.12	49.43
CPU (MC) [kHepSPEC]	0.00	0.00	0.00	0.00	0.00	0.00	0.42	20.26	45.27	36.05	51.45	63.62	75.15
CPU (an.) [kHepSPEC]	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.53	3.60	10.79	19.83	29.22	38.65
CPU (ch.) [kHepSPEC]	0.00	8.18	17.32	17.32	17.32	17.32	17.32	10.52	0.00	0.00	0.00	0.00	0.00
CPU [kHepSPEC]	0.00	8.18	17.32	17.32	17.32	17.32	17.75	31.31	48.88	65.83	105.29	140.51	174.83
CPU [# of cores 14.72HS/core]	0.00	555.70	1,176.77	1,176.77	1,176.77	1,176.77	1,206.02	2,126.90	3,320.37	4,472.04	7,153.03	9,545.75	11,877.24
WAN (raw) [Gbit/s] (inward)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.34	2.95	3.06	3.08
WAN (data) [Gbit/s] (inward)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.10	0.40	0.51	0.76	0.95	1.13
WAN (MC) [Gbit/s] (outward)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.29	0.19	0.27	0.34	0.40
WAN (MC) [Gbit/s] (inward)	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.27	0.58	0.39	0.55	0.68	0.79
WAN (ch.) [Gbit/s](in/out)	0.00	0.21	0.42	0.42	0.83	0.83	0.62	0.62	0.00	0.00	0.00	0.00	0.00
WAN [Gbit/s] (outward)	0.00	0.21	0.42	0.42	0.83	0.83	0.63	0.76	0.29	0.19	0.27	0.34	0.40
WAN [Gbit/s] (inward)	0.00	0.21	0.42	0.42	0.83	0.83	0.63	0.99	0.98	3.24	4.25	4.69	5.00



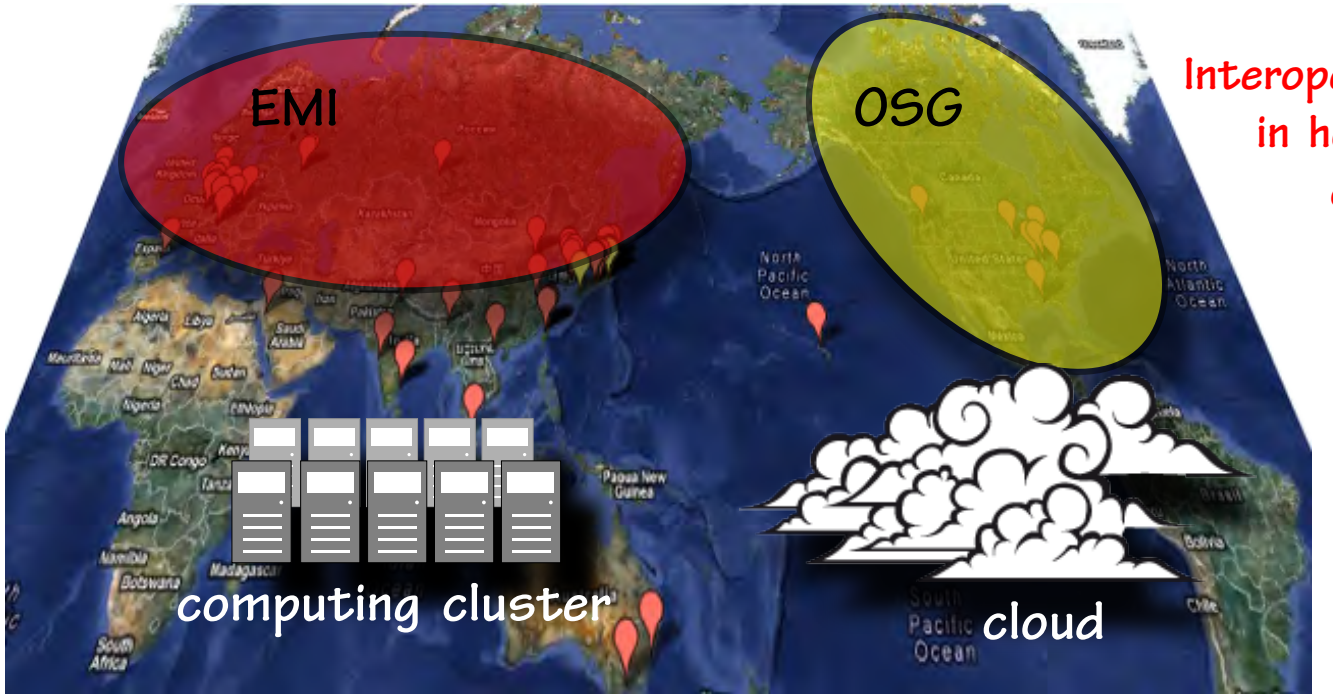



KEK : CPU (kHS06)
350.00
300.00
250.00
200.00
150.00
100.00
50.00
0.00
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023


KEK : Disk (PB)
25.00
20.00
15.00
10.00
5.00
0.00
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023

DIRAC

- Distributed Infrastructure with Remote Agent Control (developed by LHCb)
  - → Pilot jobs
  - → Modular structure that enabled it possible to submit jobs to different backends.

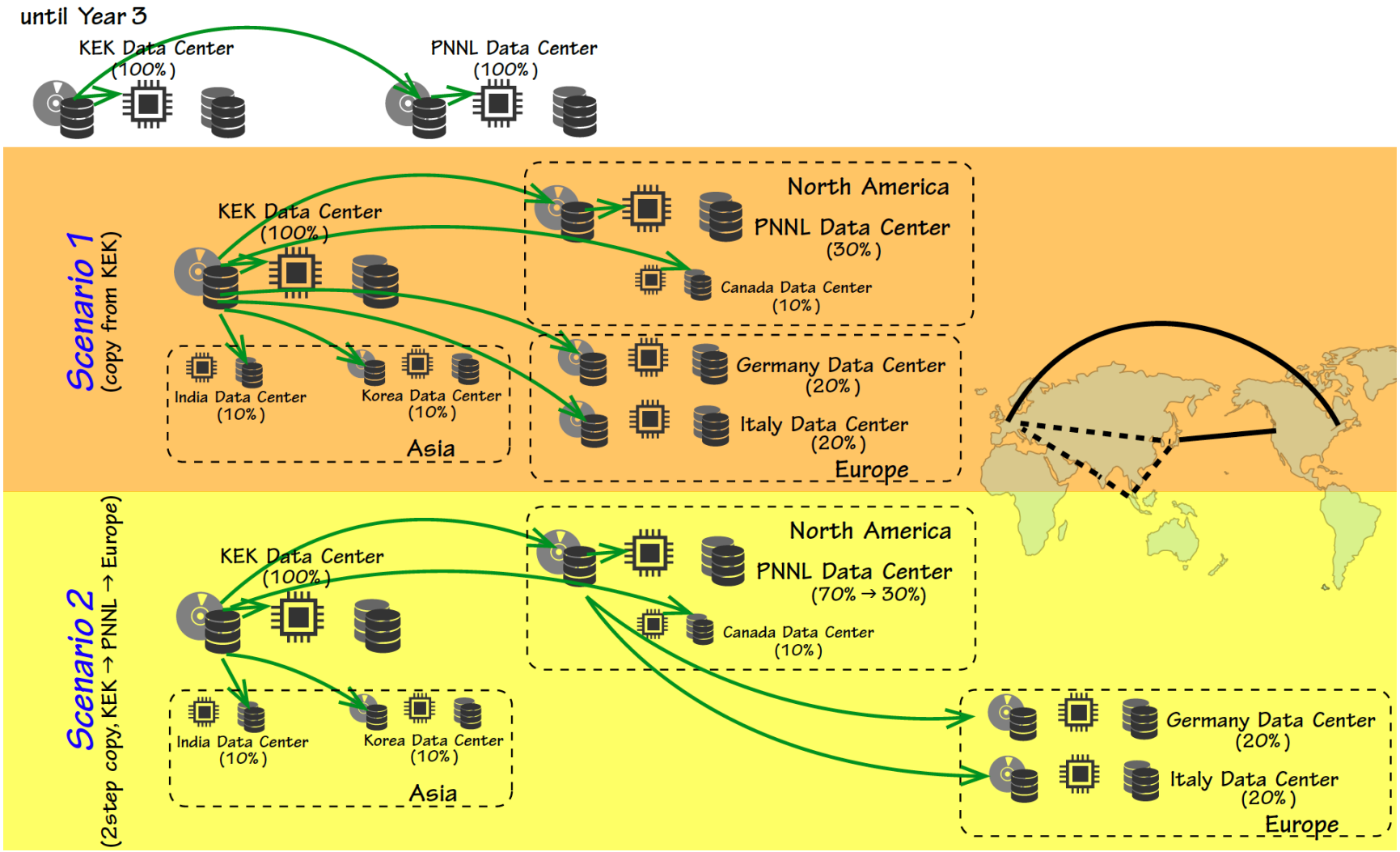


Interoperability in heterogeneous computings

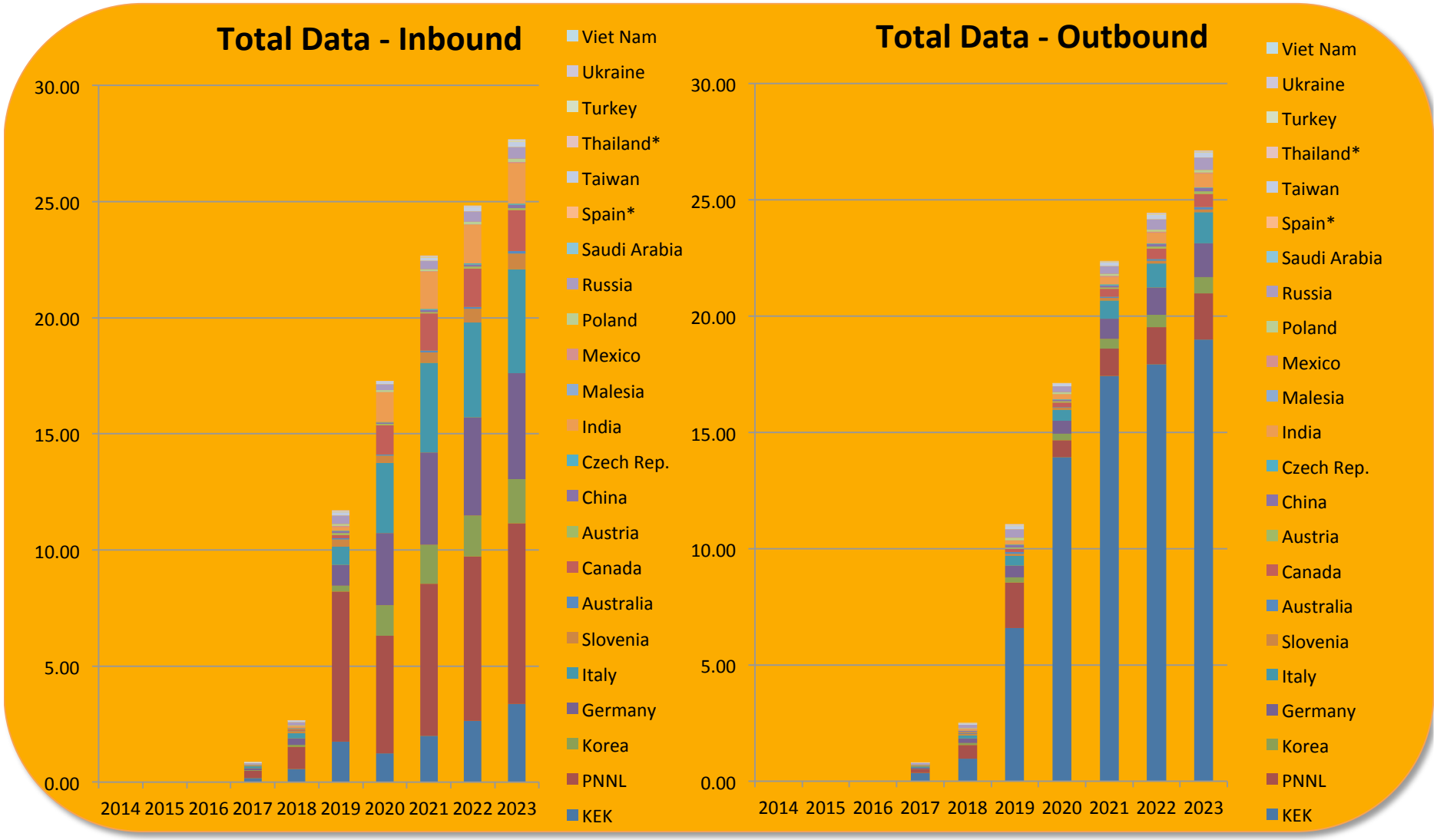
[T. Hara]

Belle II - Raw Data Replication

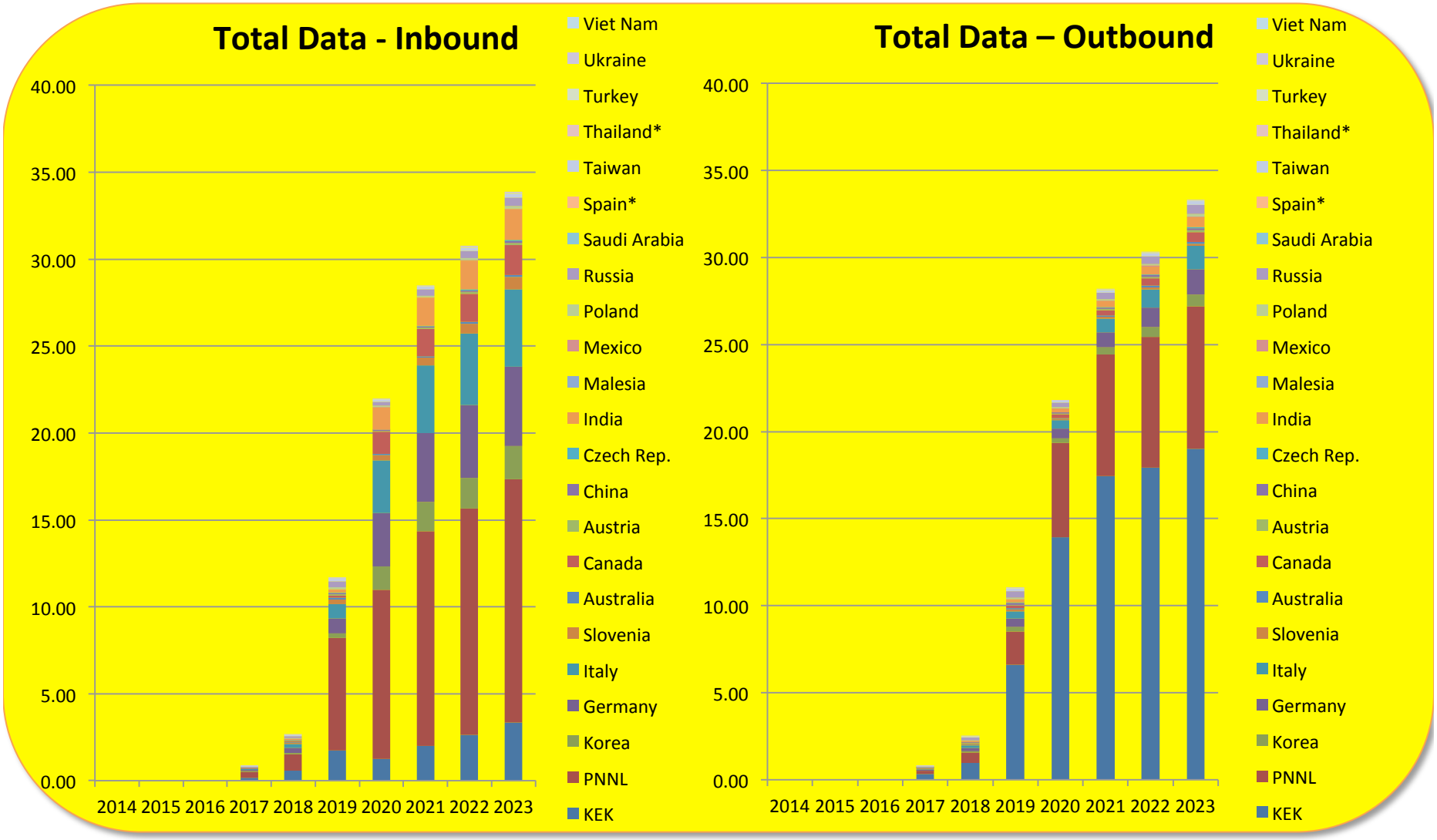
[M. Schram, PNNL]



Belle II - Network Bandwidth - Scenario 1



Belle II - Network Bandwidth - Scenario 2



Indigo - WP4

Computing	Storage	Network
Providing support for container	Defining interfaces and implementing QoS support for storage systems	Evaluation of available SDN features and operability.
Improving the on-demand compute capabilities through improved orchestration and scheduling	Providing access to the same storage through various standard access protocols.	Using SDN to configure local networks and meet PaaS needs.
		Manage local virtual Networks.
Common Subtasks		
Authentication, Authorization and Identity Management (AAI)		
Service Discovery and Monitoring		

Table 6: Breakdown of WP4 into specific and common tasks