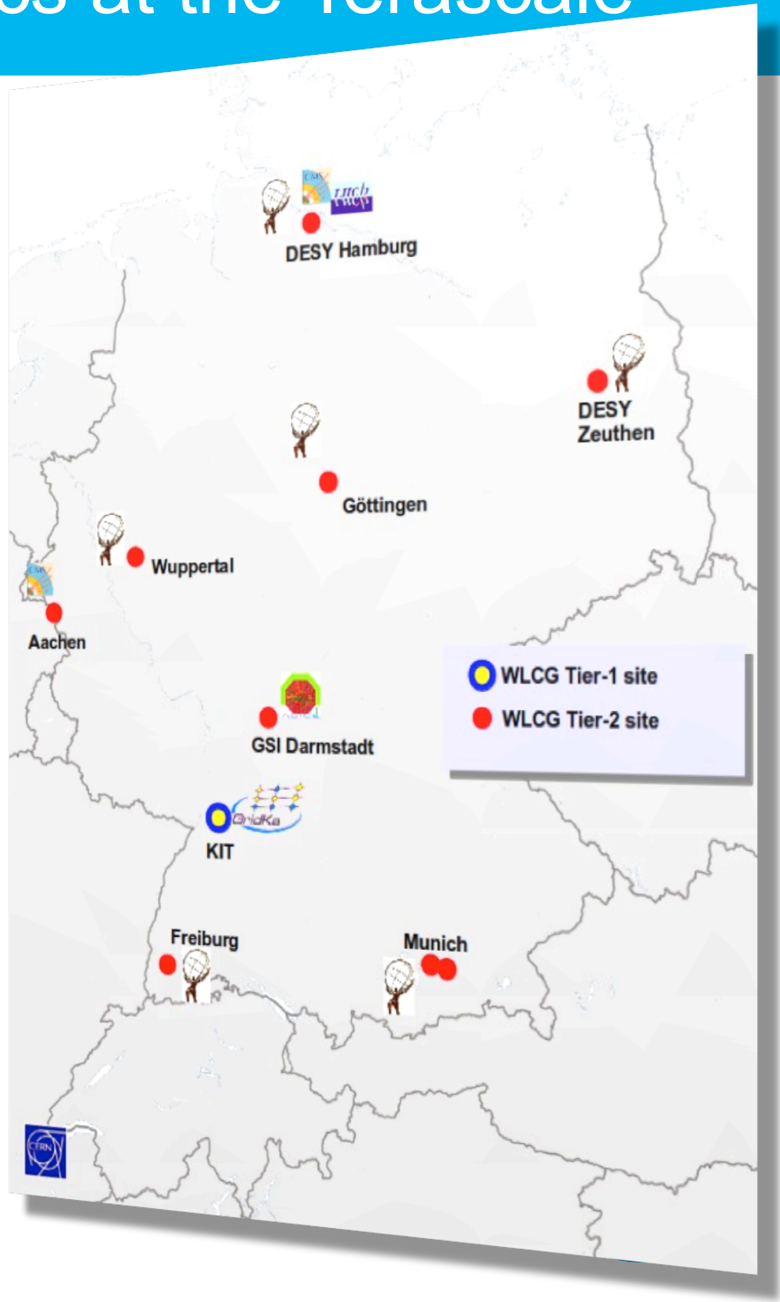
7th Annual Workshop of the Helmholtz Alliance “Physics at the Terascale”

Next generation computing for HEP data analysis in Germany

Matthias Kasemann / DESY
- on behalf of the Computing Project Board

HELMHOLTZ
GEMEINSCHAFT



Computing session yesterday - jointly with GridKa TAB

- > Experiments reports
- > GridKa T1 report
- > Networking
- > Project reports

Computing parallel session: Computing Project and GridKa technical Advisory Board

Conveners: Matthias Kasemann (DESY/CMS) , Thomas Kress (RWTH Aachen, IIIb)

Location: KIT, Campus Nord (SCC Hoersaal)

14:00 **Evolution of experiments computing: ATLAS** *15'*
Speaker: Johannes Elmsheuser (LMU Muenchen)
Material: Slides

14:15 **Evolution of experiments computing: CMS** *15'*
Speaker: Christoph Wissing (DESY)
Material: Slides

14:30 **Experiments computing: Belle-2** *15'*
Speaker: Thomas Kuhr (KIT Süd)
Material: Slides

14:45 **GridKa T1 report, status and outlook** *30'*
Speaker: Andreas Petzold (KIT)
Material: Slides

15:15 **Evolution of experiments computing: ALICE** *15'*
Speaker: Peter Malzacher (GSI)
Material: Slides

15:30 **Coffee Break** *25'*

15:55 **Networks for HEP data analysis: the global picture** *20'*
Speaker: Yves Kemp (DESY)
Material: Slides

16:15 **Storage developments in D and internationally** *30'*
Speaker: Paul Millar (DESY)
Material: Slides

16:45 **Monitoring for analysis jobs and computing infrastructure** *30'*
Speaker: Jordi Nadal
Material: Slides

17:15 **New ways using computing resources: workflow developments** *20'*
Speaker: Johannes Elmsheuser (LMU Muenchen)
Material: Slides

PHYSICS AT THE TERA SCALE
Helmholtz Alliance

Organizational: Computing Organization for the Terascale

- The “Grid Project Board” changed to “Computing Board” to reflect the extension to all computing aspects for HEP data analysis in Germany.
  - Additional representatives from all German T1 and T2 sites for LHC analysis as well as from other experiments included (ATLAS, CMS, LHCb, Alice, Belle-II, …)
- This enables the Computing board as a communication and coordination forum for German Computing for HEP Data analysis
  - We were asked to help the “Kommittee für ElementarTeilchenphysik” on computing matters
- In 2013 we had one workshop in Wuppertal (June 5th) to discuss the status of computing for data analysis and to exchange ideas and steer future directions.
  - Preparation of the KET report: “**Computing in der Hochenergiephysik in Deutschland, Bestandsaufnahme und Aussicht”** delivered in September, made public by KET, available at: https://www.ketweb.de/content/e199639/e223266/GridPB_HEP_Computing_211113.pdf

German computing in WLCG context

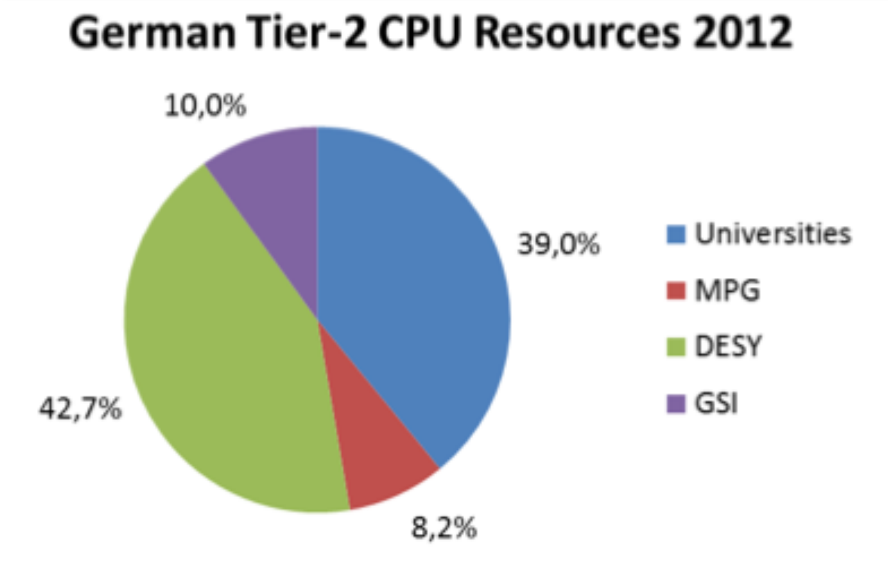
- **>** T1 center in Karlsruhe:
  - About 15% of global
  - About 60% of DE
  - Largest T1 center
  - Very reliable

Ort	CPU		Plattenplatz	
	Kerne	Anteil	TByte	Anteil
GridKa	10658	60%	9885	56%
DESY	2995	17%	2852	16%
MPG	578	3%	670	3%
GSI	700	4%	550	4%
Universitäten	2737	16%	3836	22%
Summe	17668		17793	

Anmerkung: 1 CPU-Kern = 10 HEPSpec06
Quelle: WLCG - http://wlcg-rebus.cern.ch/apps/pledges/resources/

Tabelle 2: Zugesagte („pledged“) Computing-Ressourcen in Deutschland 2012

- **>** T2 centers at DESY, GSI, MPI-M, AC, FR, GÖ, M, WU
  - About 10% of global
  - About 40% in DE
  - Rank 4 globally *after US, UK, IT*
  - Very reliably

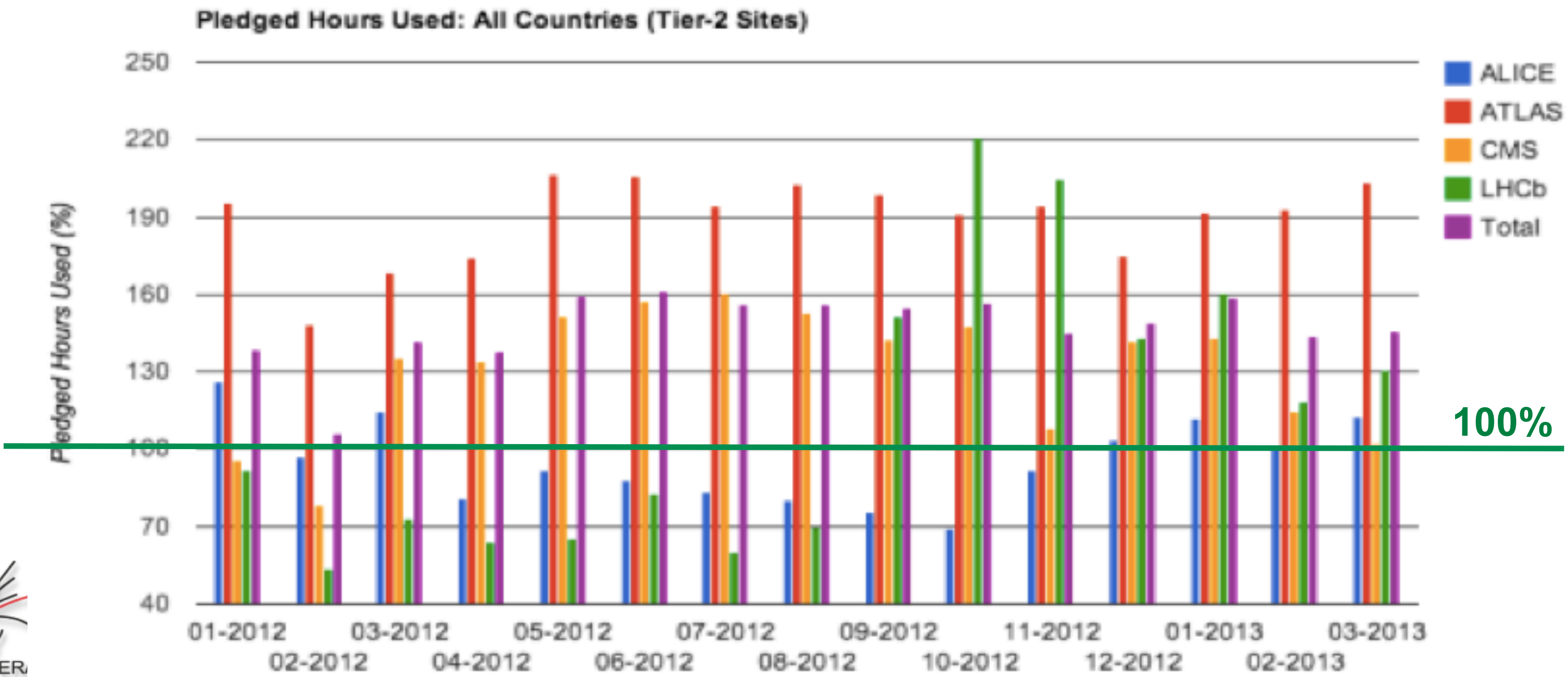


Computing resource usage in Germany

> GridKa is used by all 4 LHC experiments as well as from Auger, Babar, Belle (II), CDF, Compass und Dzero.

- This makes it the biggest T1 world-wide.
- Usage dominated by LHC

> The Tier-2 centers are used at >100% level globally as well as in DE



Computing resource usage in Germany (2)

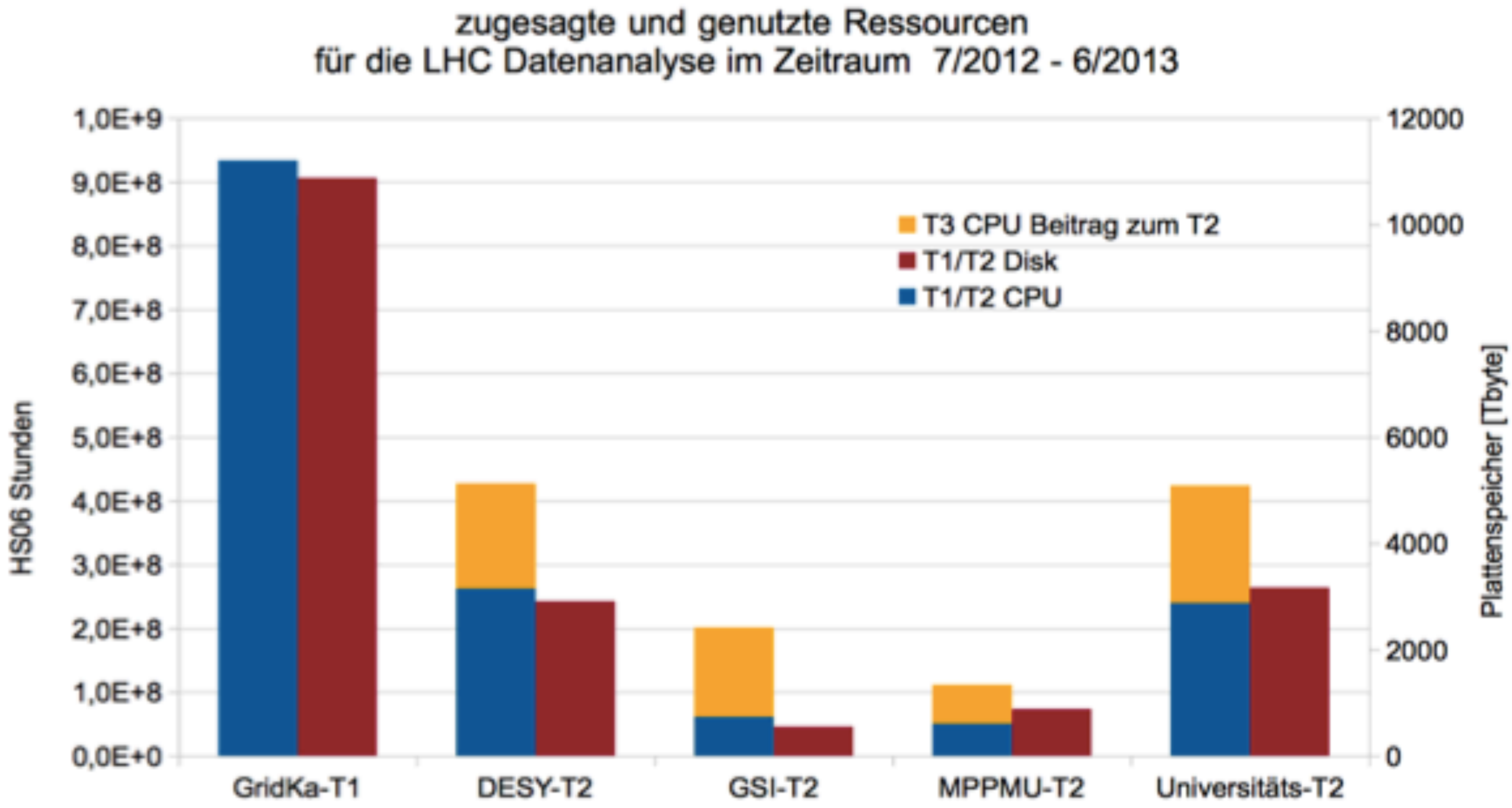
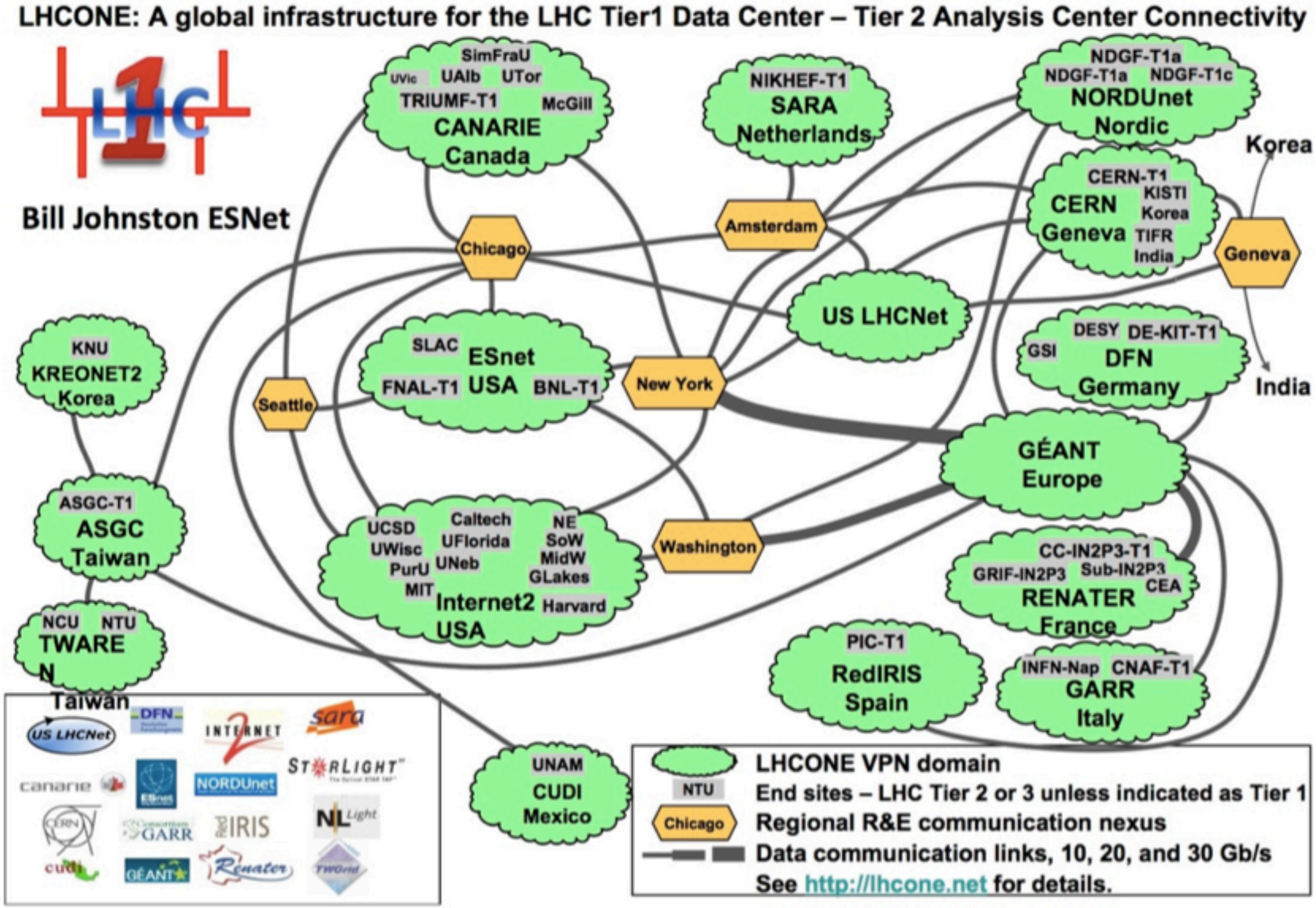


Abbildung 2: Nutzungs- und Ressourcenstatistik für die LHC-Datenanalyse an den Standorten de deutschen Tier-1 und Tier-2 Zentren

> T2 resources pledged are too low:

- much more CPU delivered than pledged (from T3-like resources)

LHCONE: A global Infrastructure: networking T1-T2



ATLAS developments

- Work on the analysis model
- New workload management
- New data management
- Federated storage using xrootd
- Opportunistic compute resources:
  - Clouds
  - HPCs

Challenges of Run 2

Trigger rate: from 550Hz to 1kHz:

- Therefore, more events to record and process

Luminosity increase: event pile-up from 25 to 40:

- so more complexity for processing and +20% event size

Flat resource budget:

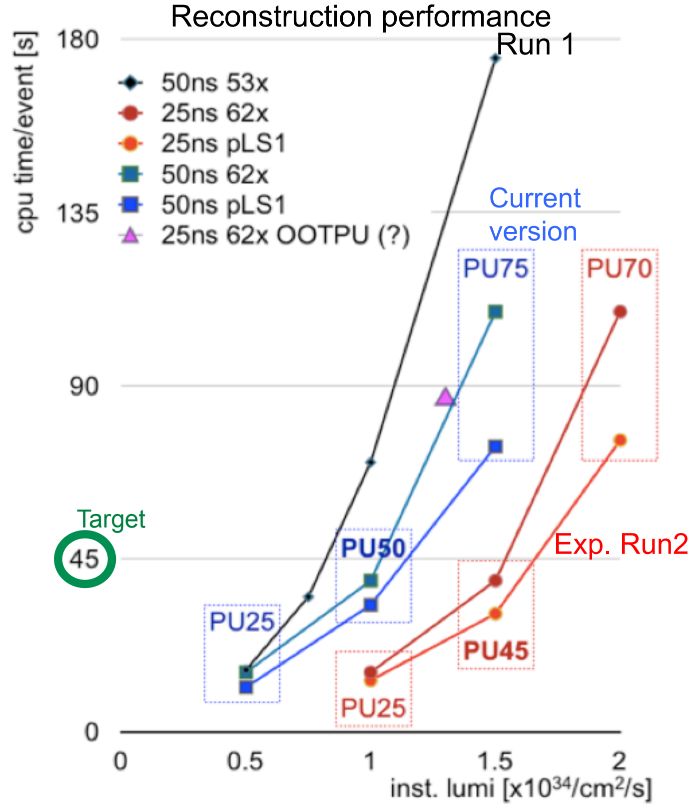
- For storage, CPUs and network (apart for Moores law)
- For operations manpower
- The LHCC recommends that the hypothesis of flat future resources be removed from the assumptions; instead physics motivated needs should be stated.

The ATLAS Distributed Computing infrastructure needs to evolve in order to face those challenges

CMS developments

Challenges for Computing in 2015 and beyond

- Higher Luminosity and more complex events in 2015 and beyond
  - Factor 2-3 in CPU increase
- Keep trigger threshold at 2012 level
  - Logging rate increases to ~1kHz
  - ~ Factor 2 increase
  - Turns directly into CPU and storage requirement
- Big effort spent in speeding up the reconstruction software

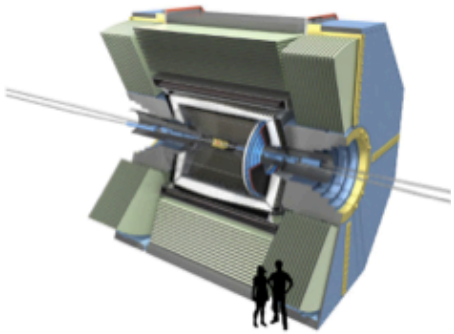


- Use resources more efficiently
  - Fewer re-reconstr.
  - Fewer/faster MC
  - Use T1 for reconstruction
  - Reduce data replication
- Improve software performance
  - Smaller memory footprint
  - Thread-safe code for efficient use of multi-core infrastructures
- Use of WAN data access: reduce storage needs

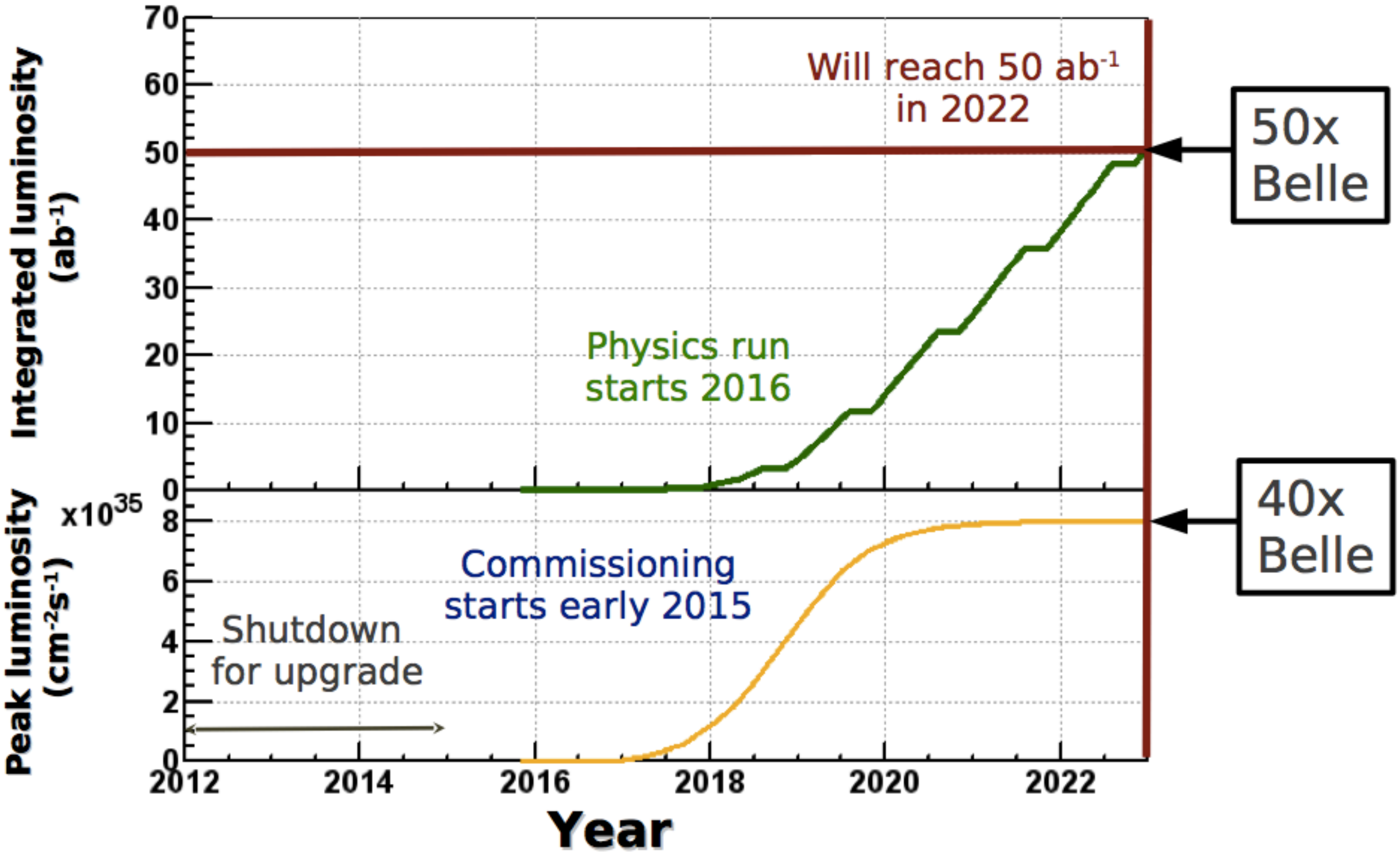
Prepare for Cloud computing and opportunistic computing

Summary

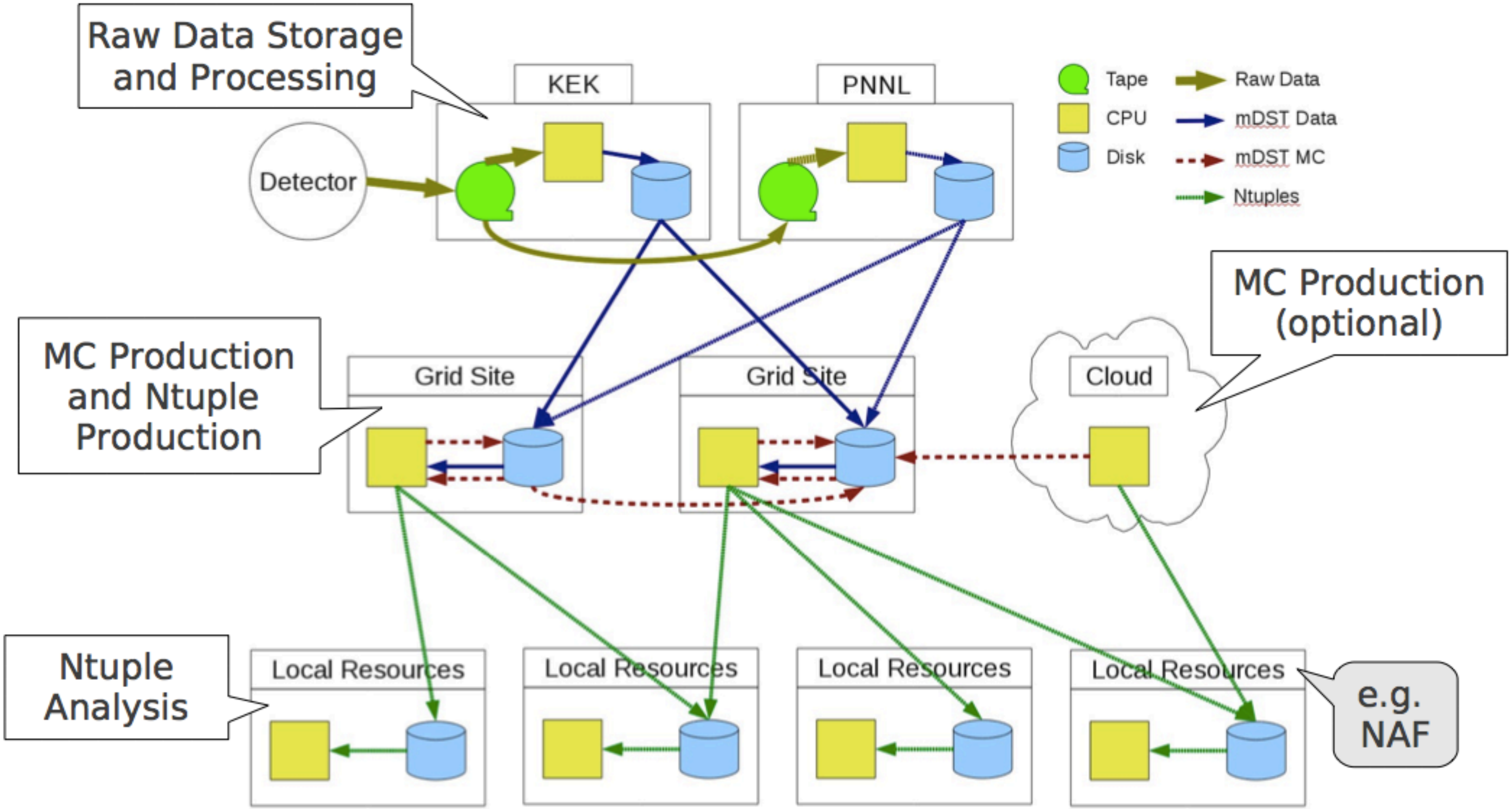
- Belle II will search for New Physics with O(50) times more data than current B factories
- Huge data volume is a challenge for the computing
  - Similar, but bit simpler computing model than WLCG
  - Distributed computing system based on existing technologies and infrastructures
  - Workflow abstraction 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



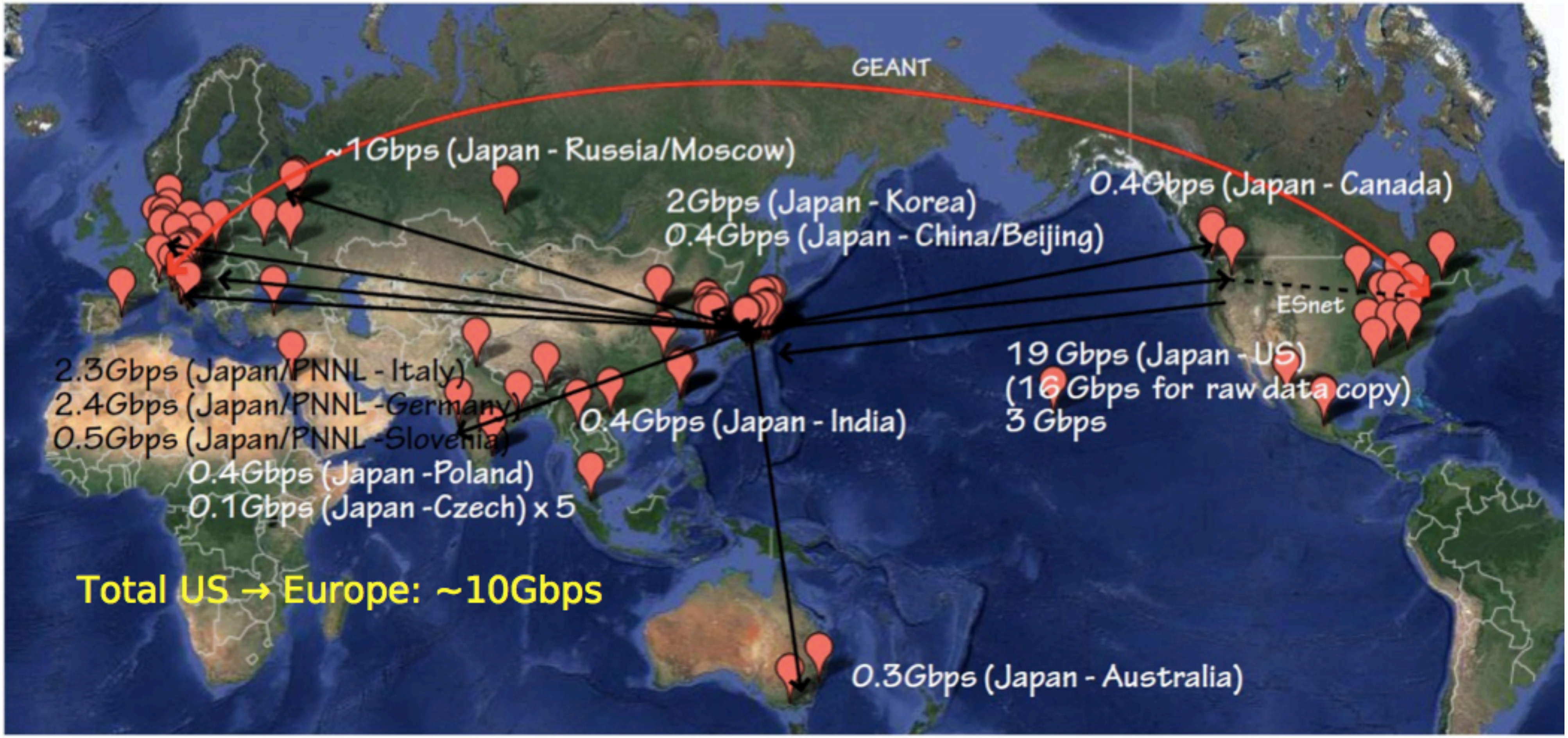
Projection of Luminosity at SuperKEKB



Belle II Computing Model

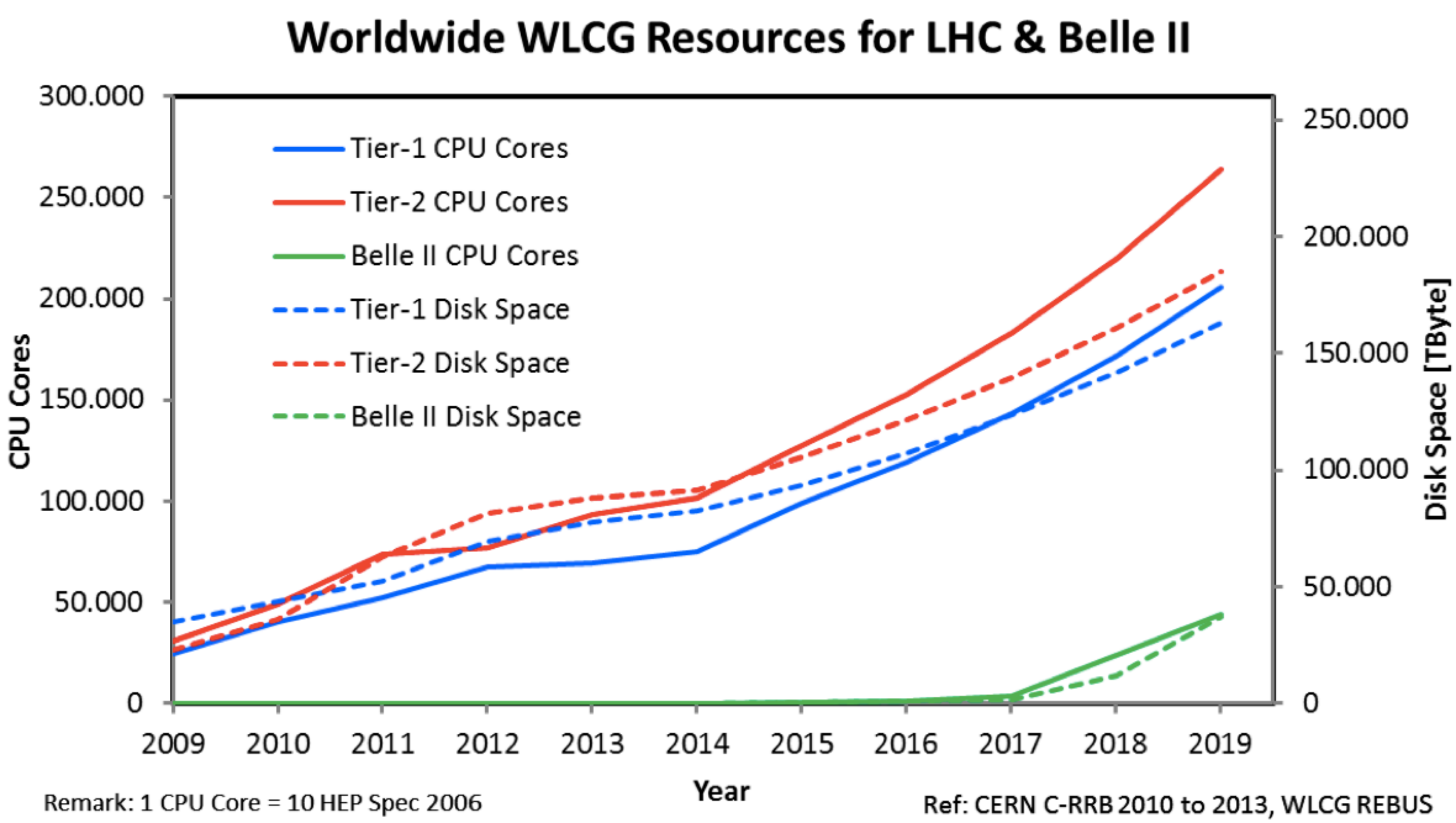


Network Bandwidth Estimates for 2022



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

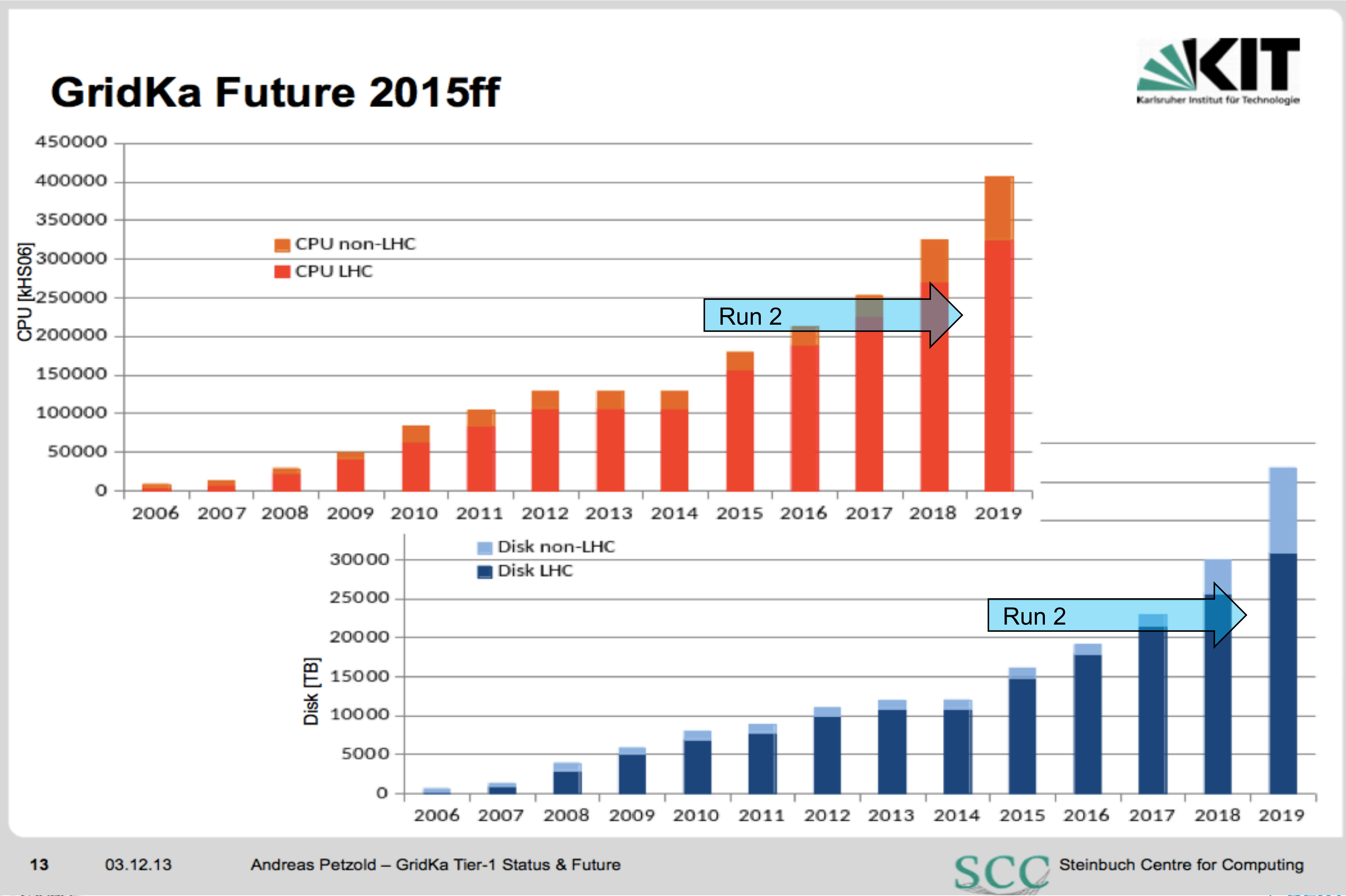
Computing for LHC run-2 and Belle-II



- Estimates for 2015: ++30% for T1, ++25% for T2
- Estimates for '16-19: ++20% for CPU, ++15% for disk per ann.

Under the assumption that experiments achieve large performance improvements handling trigger rates and large pile-up.

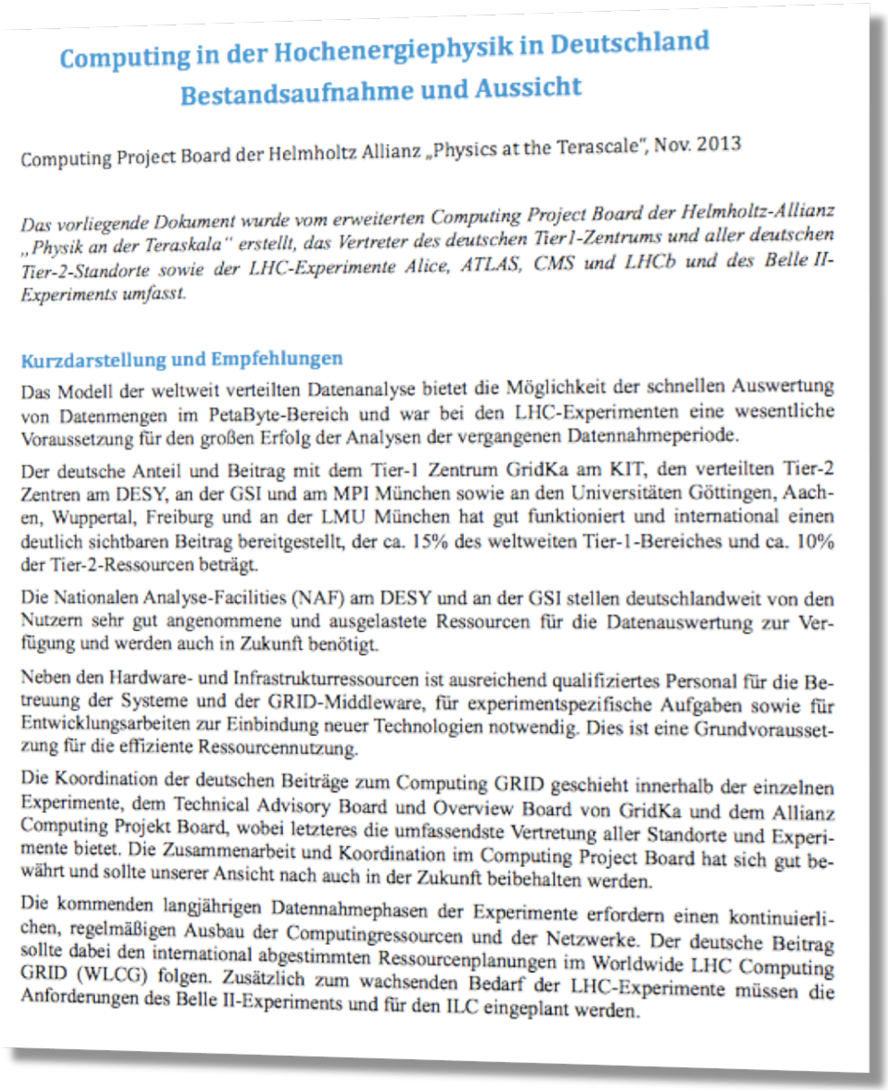
Resource planning for GridKa



Background of Computing Report for KET: Financing

Main Recommendations:

- LHC Computing Grid is successful. Germany provides ~15% T1 and ~10% T-2 of global resources. This needs to continue at about this level, similarly the NAF at DESY (ALTAS, CMS, LHCb) and at GSI (ALICE).
- Qualified personnel is required to operate systems and middleware as well as experiment specific services. Developments are required for new computing architectures to maintain good efficiencies.
- Coordination bodies in Germany are: experiments, GridKa-TAB and –OB and Terascale Computing Board, they should work closely together.
- LHC run-2 as well as Belle-II require a continuous upgrade of compute resources and networking.



Computing in der Hochenergiephysik in Deutschland
Bestandsaufnahme und Aussicht

Computing Project Board der Helmholtz Allianz „Physics at the Terascale", Nov. 2013

Das vorliegende Dokument wurde vom erweiterten Computing Project Board der Helmholtz-Allianz „Physik an der Teraskala" erstellt, das Vertreter des deutschen Tier1-Zentrums und aller deutschen Tier-2-Standorte sowie der LHC-Experimente Alice, ATLAS, CMS und LHCb und des Belle II-Experiments umfasst.

Kurzdarstellung und Empfehlungen

Das Modell der weltweit verteilten Datenanalyse bietet die Möglichkeit der schnellen Auswertung von Datenmengen im PetaByte-Bereich und war bei den LHC-Experimenten eine wesentliche Voraussetzung für den großen Erfolg der Analysen der vergangenen Datennahmeperiode.

Der deutsche Anteil und Beitrag mit dem Tier-1 Zentrum GridKa am KIT, den verteilten Tier-2 Zentren am DESY, an der GSI und am MPI München sowie an den Universitäten Göttingen, Aachen, Wuppertal, Freiburg und an der LMU München hat gut funktioniert und international einen deutlich sichtbaren Beitrag bereitgestellt, der ca. 15% des weltweiten Tier-1-Bereiches und ca. 10% der Tier-2-Ressourcen beträgt.

Die Nationalen Analyse-Facilities (NAF) am DESY und an der GSI stellen deutschlandweit von den Nutzern sehr gut angenommene und ausgelastete Ressourcen für die Datenauswertung zur Verfügung und werden auch in Zukunft benötigt.

Neben den Hardware- und Infrastrukturressourcen ist ausreichend qualifiziertes Personal für die Betreuung der Systeme und der GRID-Middleware, für experimentspezifische Aufgaben sowie für Entwicklungsarbeiten zur Einbindung neuer Technologien notwendig. Dies ist eine Grundvoraussetzung für die effiziente Ressourcennutzung.

Die Koordination der deutschen Beiträge zum Computing GRID geschieht innerhalb der einzelnen Experimente, dem Technical Advisory Board und Overview Board von GridKa und dem Allianz Computing Projekt Board, wobei letzteres die umfassendste Vertretung aller Standorte und Experimente bietet. Die Zusammenarbeit und Koordination im Computing Project Board hat sich gut bewährt und sollte unserer Ansicht nach auch in der Zukunft beibehalten werden.

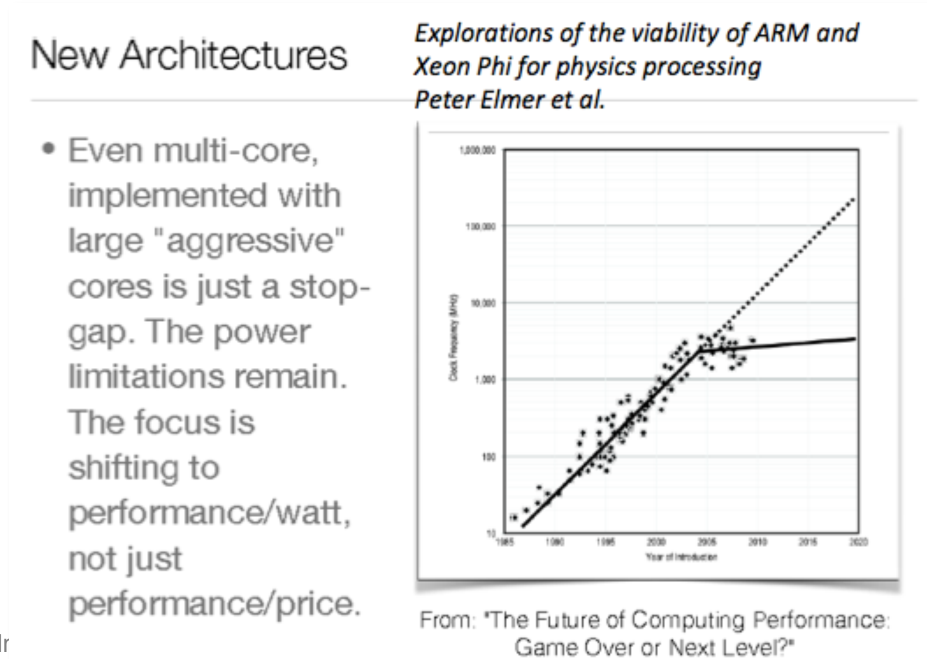
Die kommenden langjährigen Datennahmephasen der Experimente erfordern einen kontinuierlichen, regelmäßigen Ausbau der Computingressourcen und der Netzwerke. Der deutsche Beitrag sollte dabei den international abgestimmten Ressourcenplanungen im Worldwide LHC Computing GRID (WLCG) folgen. Zusätzlich zum wachsenden Bedarf der LHC-Experimente müssen die Anforderungen des Belle II-Experiments und für den ILC eingeplant werden.

Conclusion and outlook (1)

- The LHC computing models are evolving:
  - High bandwidth networking is available in most countries
    recent record: demonstrated 100 Gb from KIT to SC conference in Denver
  - Data placement is much more dynamic than originally foreseen in the “Monarc-model”
  - Fewer static data replication allows more efficient use of disk storage

    => This requires more elaborated workload and data management tools

- The computing technology is changing
  - Multi-core computing
  - Thread safe programming
  - Porting vs re-implementing existing frameworks and algorithms
  - Optimize software for
    - ARM architecture
    - Xeon Phi
    - GPUs
  - Give-up ‘one-size-fits-all’ approach
    => increased development and certification effort



PHYSICS AT THE TERA SCALE Helmholtz Alliance

Conclusion and outlook (2)

- The big success of the LHC physics programme was also possible because of the excellent performance of the LHC Computing Grid.
- German sites rank among the most-used and most-efficient sites in the WLCG.
- "Physics at the Terascale" has significantly contributed to the Grid infrastructure in Germany: T2 at University sites and NAF.
  - We are prepared to continue at coordination level
- Computing resource estimates for LHC-run2 were reviewed in C-RRB
  - Much reduced relative to run1 – fewer compute cycles and storage per fb^{-1} will be possible
  - Belle-II computing needs are comparable to ATLAS or CMS
  - **No funding could be secured yet for German T1 and T2 resource replacements and upgrades.**