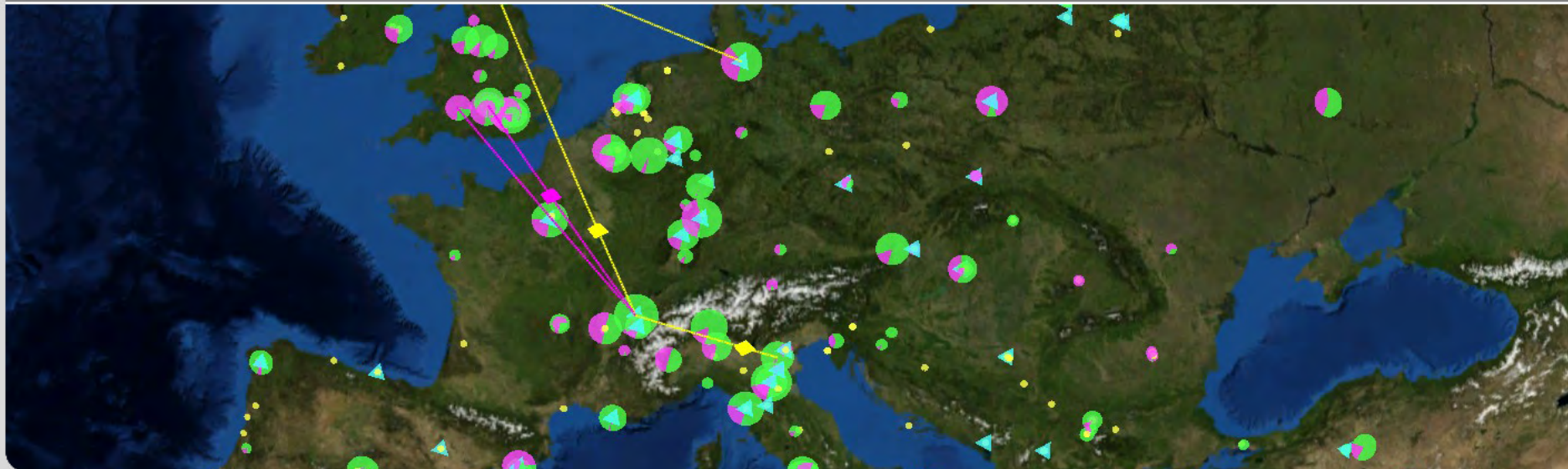


Grid-Computing in Deutschland: Status und Zukunft

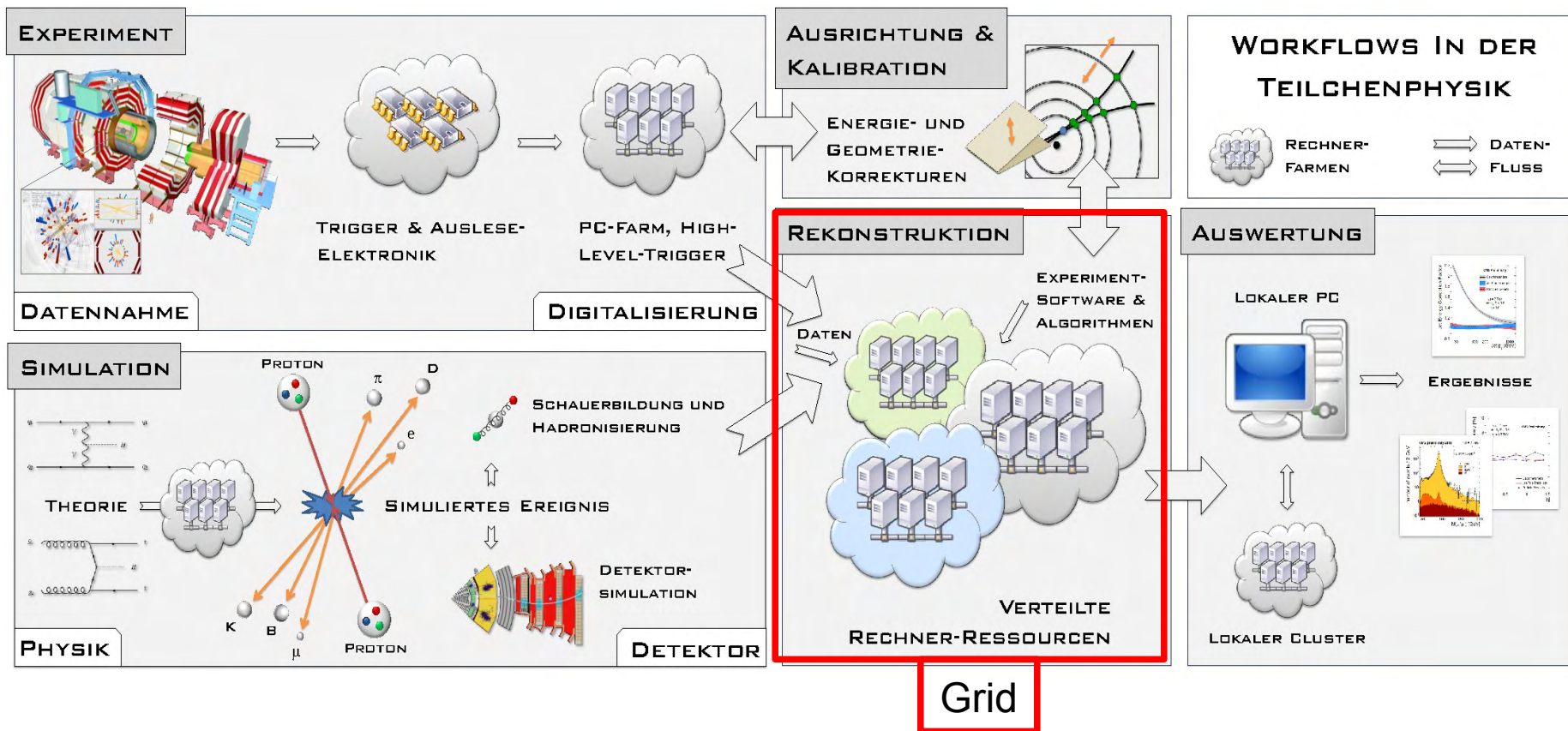
Günter Quast

Jahresversammlung der Teilchenphysik
Bad Honnef, 19.11.2011

Institut für Experimentelle Kernphysik



Computing-Aufgaben in der HEP



- CPU-intensive Simulation von Teilchenreaktionen und des Detektors
- Prozessieren (=Rekonstruktion) großer Datenmengen
- I/O-intensives Filtern und Verteilen von Daten
- Transfer zu lokalen PC-Clustern zur physikalischen Interpretation

WLCG: Grid mit hierarchischer Struktur

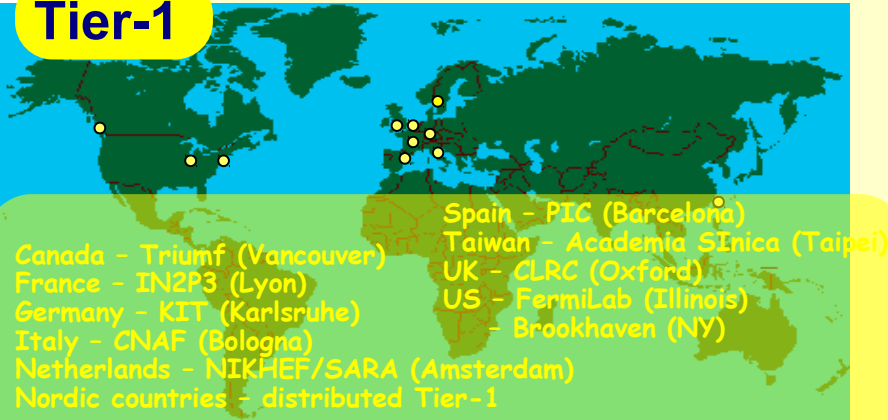
Tier-0

Standort des Beschleunigers

- Datennahme & erste Rekonstruktion
- Langzeit-Datenspeicherung
- Datenverteilung an T1 / T2



Tier-1



11 Tier-1 Zentren

- “quasi-online” im Datennahmeprozess
 - hohe Verfügbarkeit
- Massenspeicher: Disk und Band
 - Grid-basierte Datendienste
- Rekonstruktion und Daten-intensive Analyse
- Nationale Nutzerunterstützung

Tier-2

150 Zentren in 60 Föderationen in 35 Ländern

- Physik-Analyse durch Gruppen und Einzelnutzer & Zusammenarbeit mit T3 (= Instituts-Ressourcen)
- Monte-Carlo-Simulation

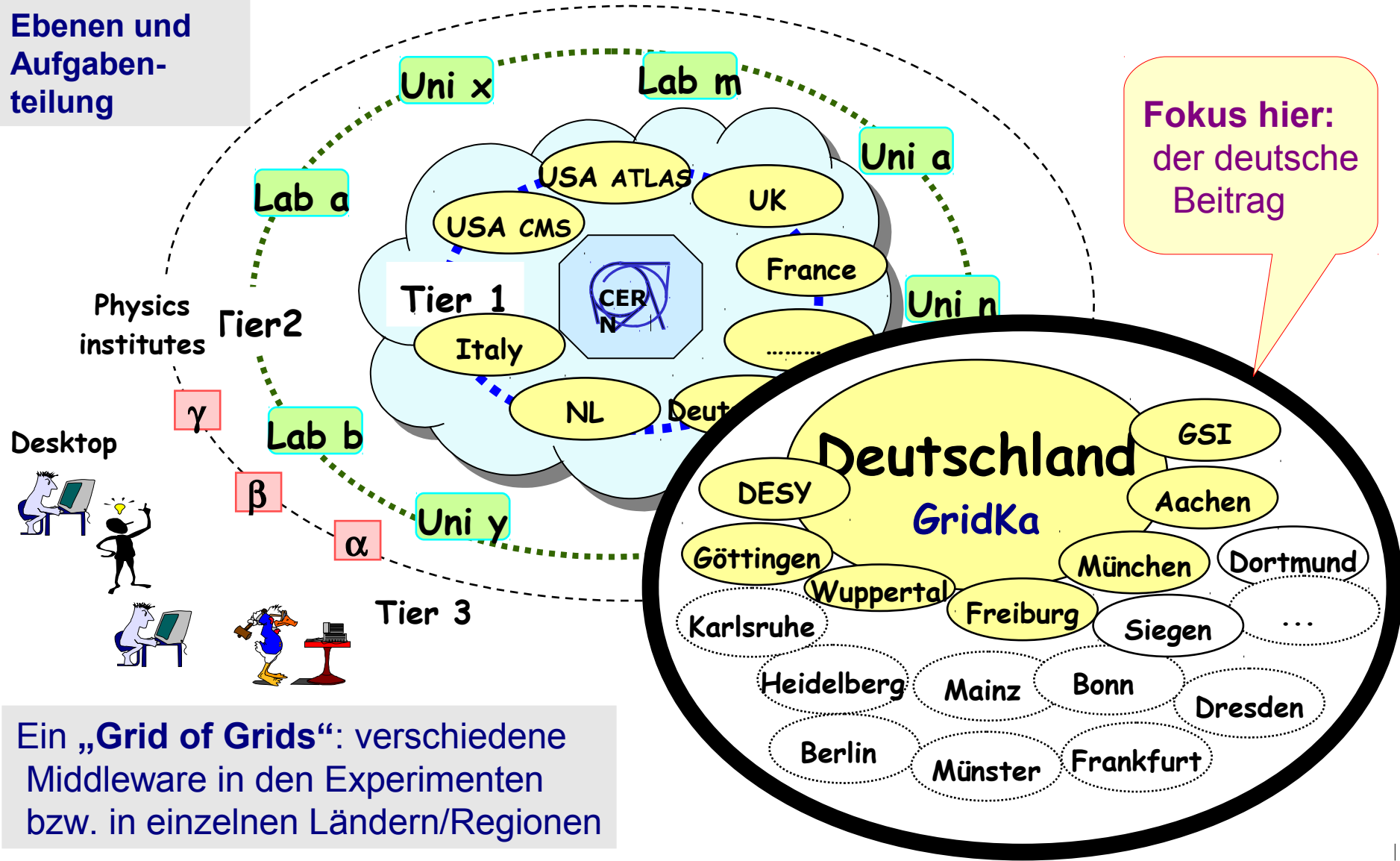
Tier-3

einige 100 Grid-fähige PC-Cluster an Instituten

- wo die Entdeckungen gemacht werden

Das weltweite LHC-Computing-Grid („WLCG“)

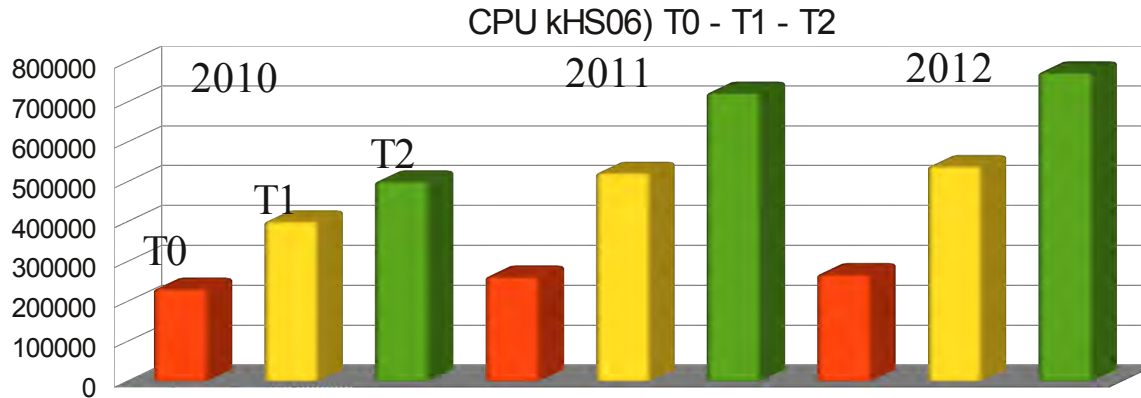
Ein Grid mit hierarchischen Ebenen und Aufgabenteilung



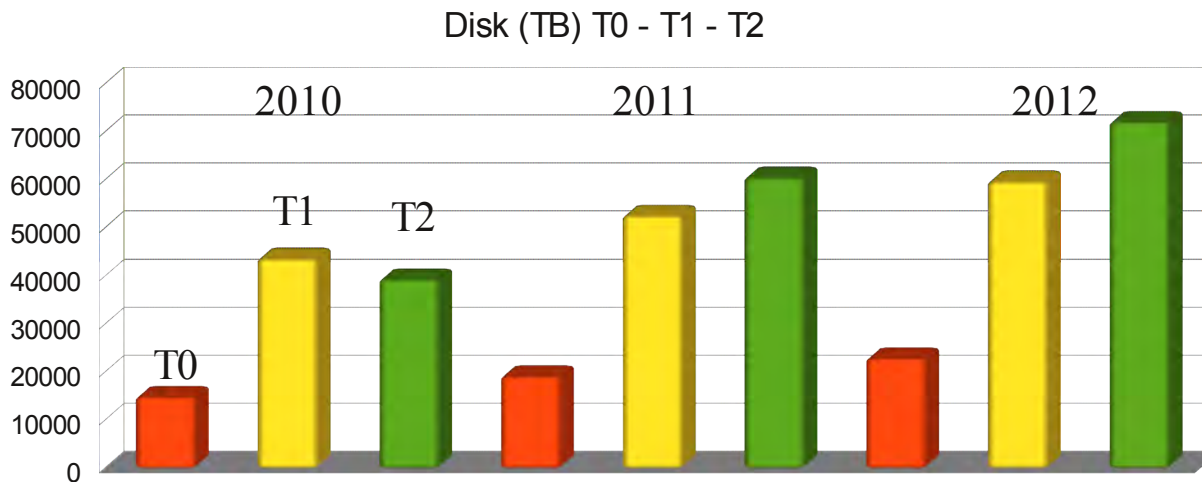
Ein „Grid of Grids“: verschiedene Middleware in den Experimenten bzw. in einzelnen Ländern/Regionen

Grid Computing Ressourcen 2010-2012

10 HS06 ≈
1 CPU-Kern



Gesamt-CPU 2011:
1500 kHS06
etwa äquivalent
150'000 CPU-Kernen



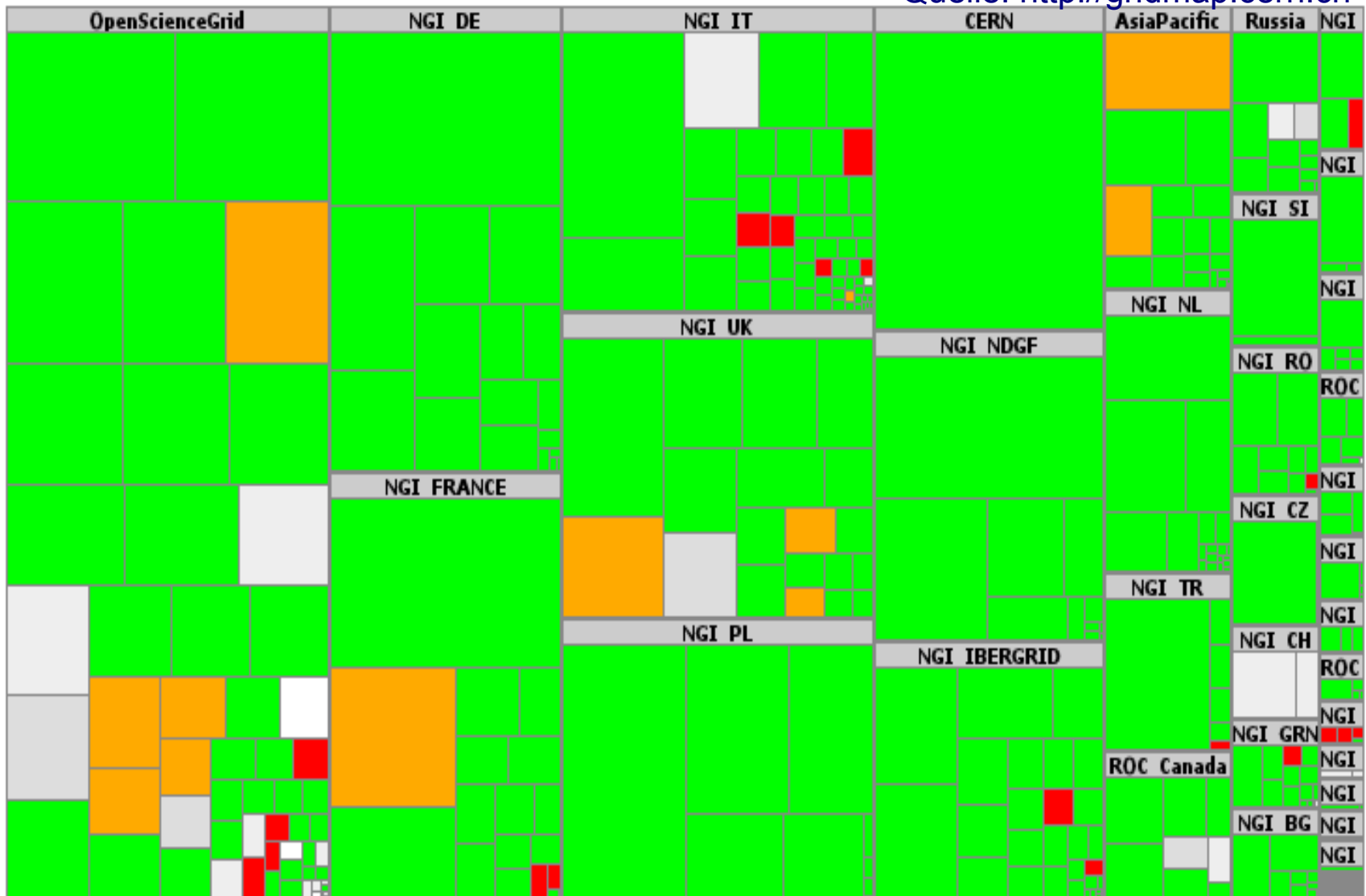
Gesamt-Disk 2011:
130 PB

Bandspeicher 2011:
130 PB
(40 PB am CERN,
0 @ T2s)

Das weltweit größte Wissenschafts-Grid

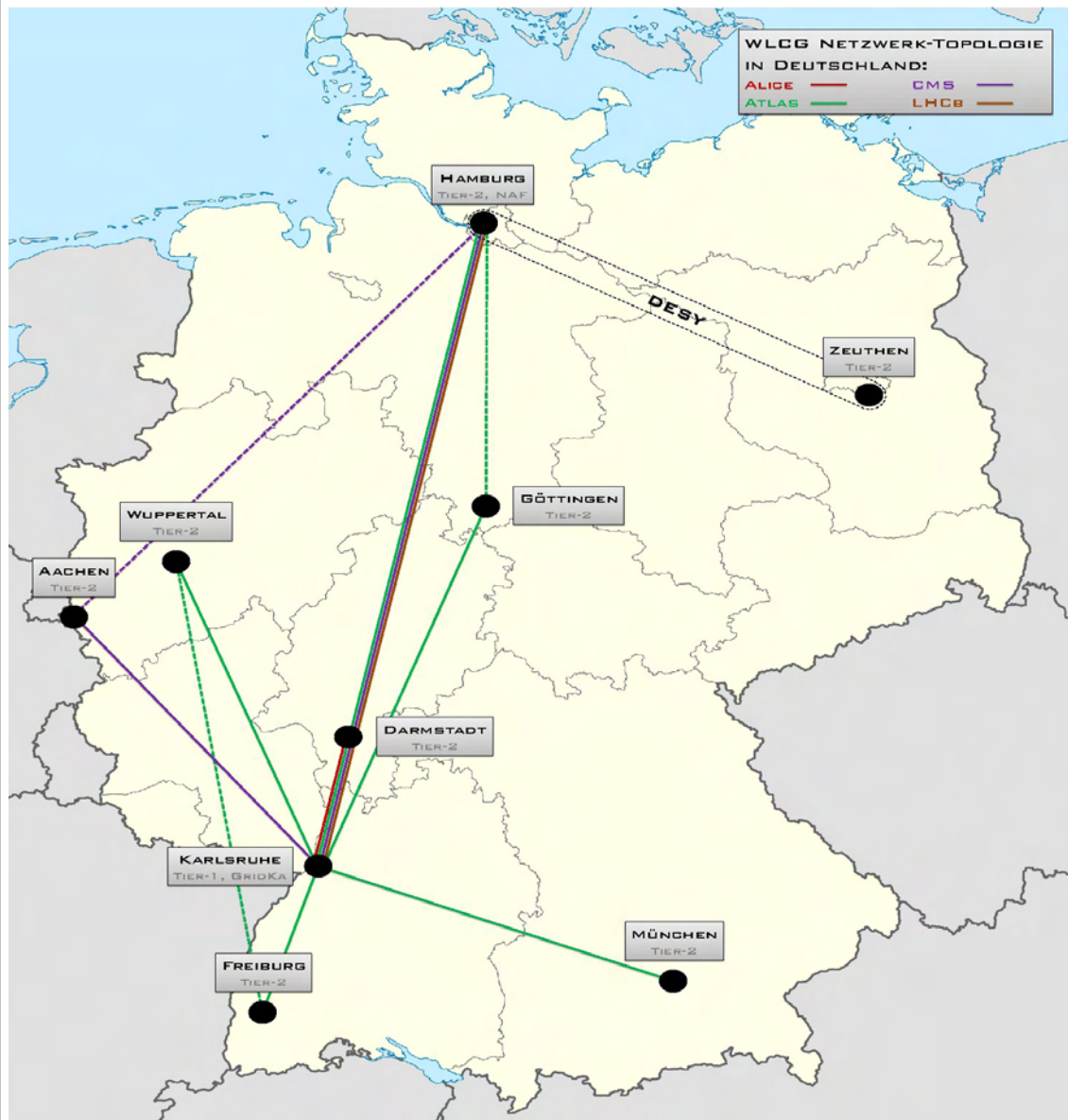
Grid Sites und deren Größe

Quelle: <http://gridmap.cern.ch>



Fläche entspricht **Rechenleistung** eines Zentrums im WLCG
Farbe reflektiert aktuellen **Status** (**S**ite **A**vailability **M**onitoring, **SAM** Test)

Grid-Struktur in Deutschland



Standorte und Funding

- GridKa T1 (alle LHC-Exp.)
HGF & BMBF
- DESY T2 (Atlas, CMS und LHCb)
HGF & BMBF
- Aachen T2 (CMS)
HGF-A Terascale & NRW
- Göttingen T2 (ATLAS)
Niedersachsen und HGF-A Terascale
- Freiburg T2 (ATLAS)
HGF-A Terascale und Ba.-Wü.
- München T2 (ATLAS)
MPG, HGF-A Terascale und Bayern
- Wuppertal T2 (ATLAS)
HGF-A Terascale und NRW
- GSI (ALICE)

Beiträge der Bundesländer i.W.
zu Betrieb von Hardware und
Betriebssystemsoftware

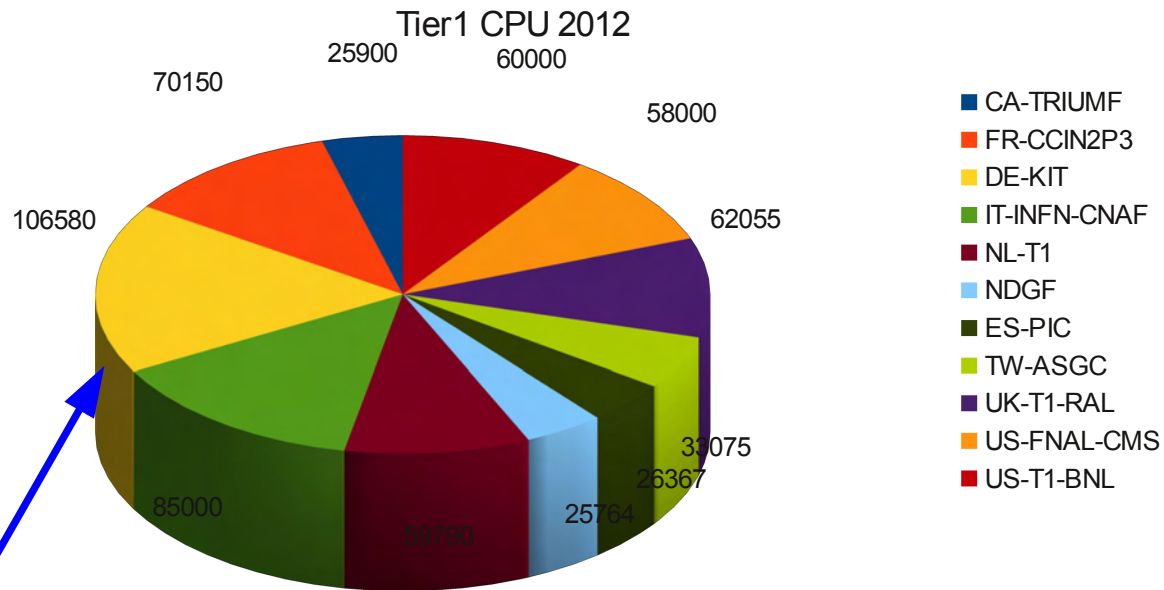
Tier-1: GridKa

WLCG Tier-1
2012

CPU (HS06)
553'000

Disk
67 PB

Band
103 PB

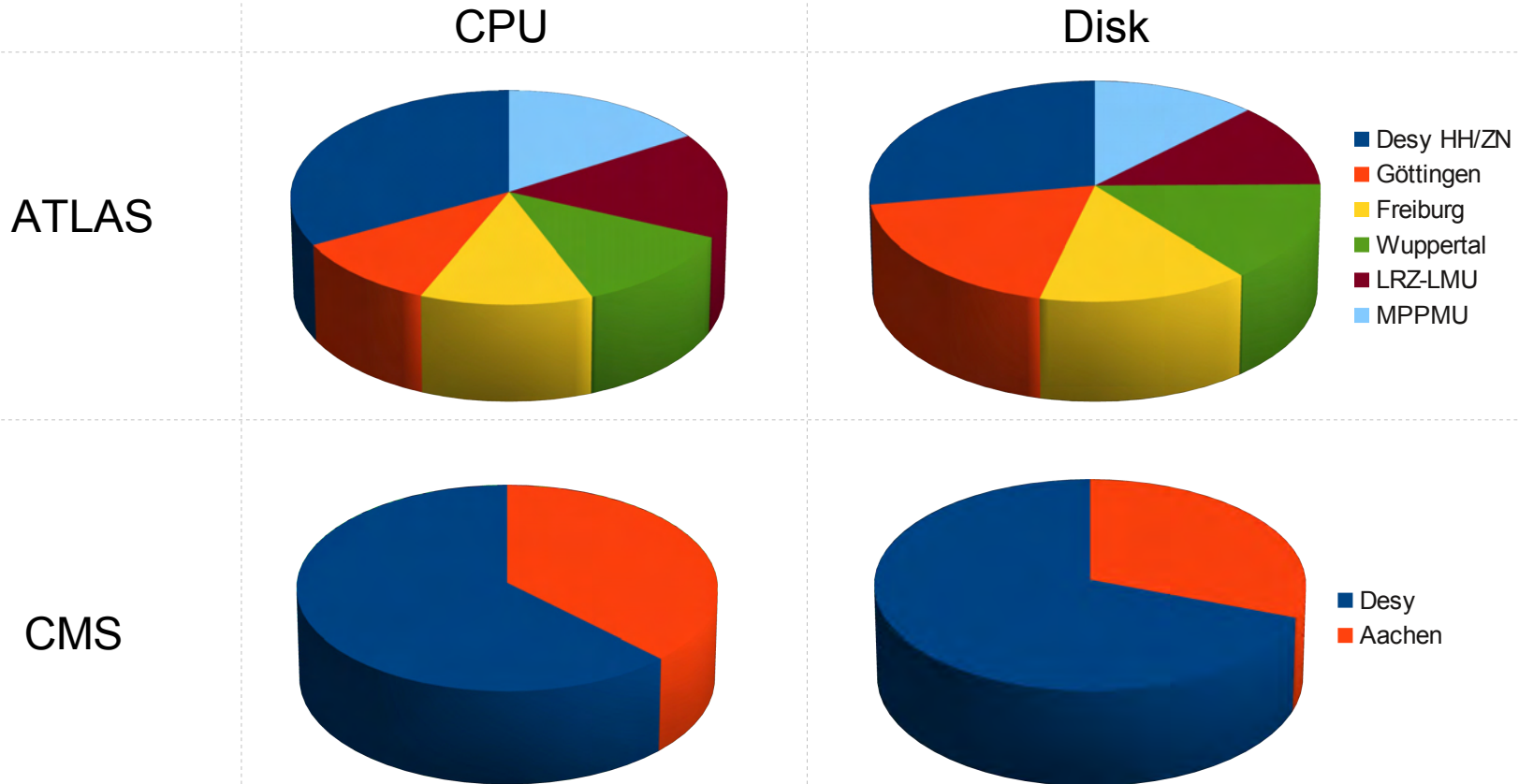


GridKa ist das größte Tier1 im WLCG und stellt etwa 15% der gesamten T1-Ressourcen

GridKa:	CPU (HS06)	%WLCG	Disk	%WLCG	Band	% WLCG
ALICE :	40000	25%	2,7 PB	25%	5,2 PB	25%
ATLAS:	32400	12.5%	3,4 PB	12,5%	4,5 PB	12,5%
CMS:	24000	10%	2,2 PB	10%	5,1 PB	10%
LHCb:	19200	17%	1,6 PB	17%	1,6 PB	17%

Aufteilung der deutschen Tier-2-Beiträge

DE T2	CPU (HS06)	%WLCG T2	Disk	%WLCG T2
ATLAS:	36000	13,6%	5,4 PB	11,5%
CMS:	24000	7,5%	2,0 PB	7,5%
LHCb:	3200	7%	0.2 PB	10 %



Ressourcen an den Universitäten decken im Jahr 2012
46% der CPU- und **52% der Disk-** Anforderungen ab.

Aufgabenteilung zwischen Zentren und Experimenten:

Zentren

- betreiben Hardware
 - pflegen Software bis zur Betriebssystem- und Middleware-Ebene
- (finanziert aus Mitteln des jeweiligen Standorts)

Experimente

- installieren und pflegen Experiment-spezifische Software
 - betreiben und verantworten Experiment-spezifische Dienste („VOBox“)
 - koordinieren Datentransfers und Datenmanagement
 - pflegen Datensätze, sorgen für Datenkonsistenz
 - unterstützen die lokalen Nutzer
 - fungieren als Ansprechpartner der Experimente und sind Teil der experimentellen Support-Teams
- (gefördert durch BMBF-Verbundforschung)

Koordinierende Gremien:

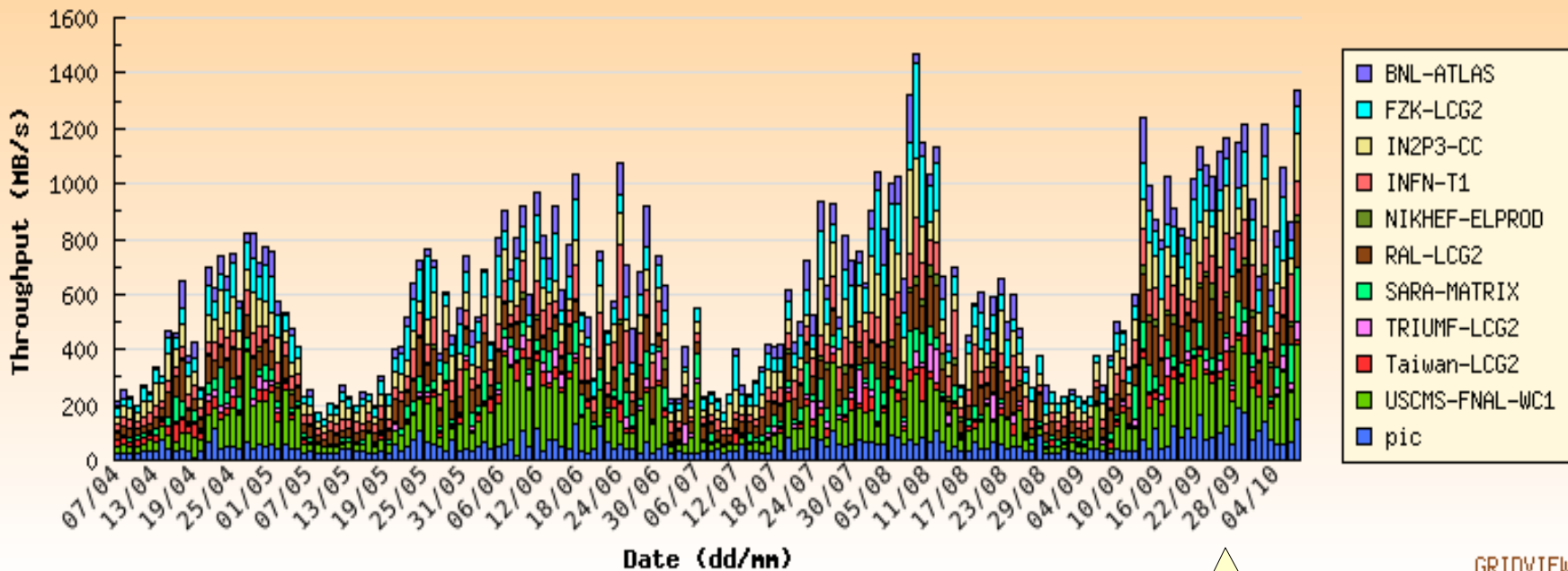
GridKa Technical Advisory Board (TAB), GridKa Overview-Board, Grid Project Board und Management Board der HGF-A „Terascale“

Funktioniert das Grid ?

Tier-1

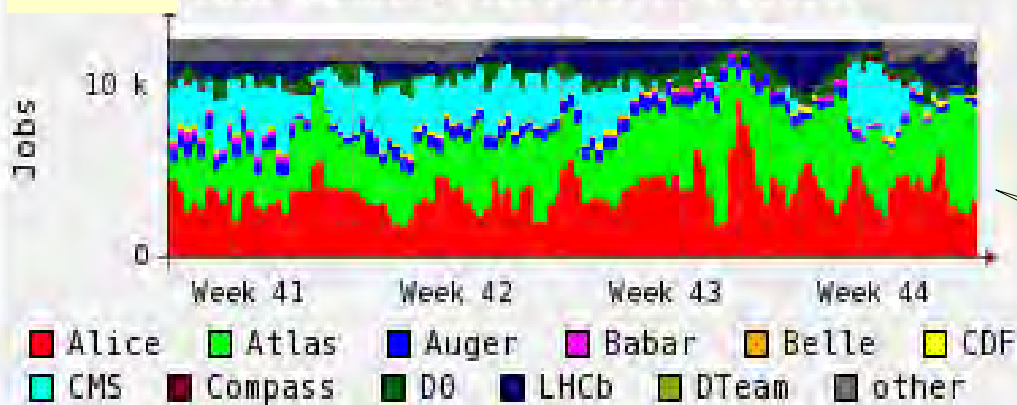
CERN Export

Averaged Throughput From 07/04/11 To 07/10/11
Site-wise Data Transfer From CERN-PROD To All T1 Sites



GRIDVIEW

GridKa Jobs running last month (total)



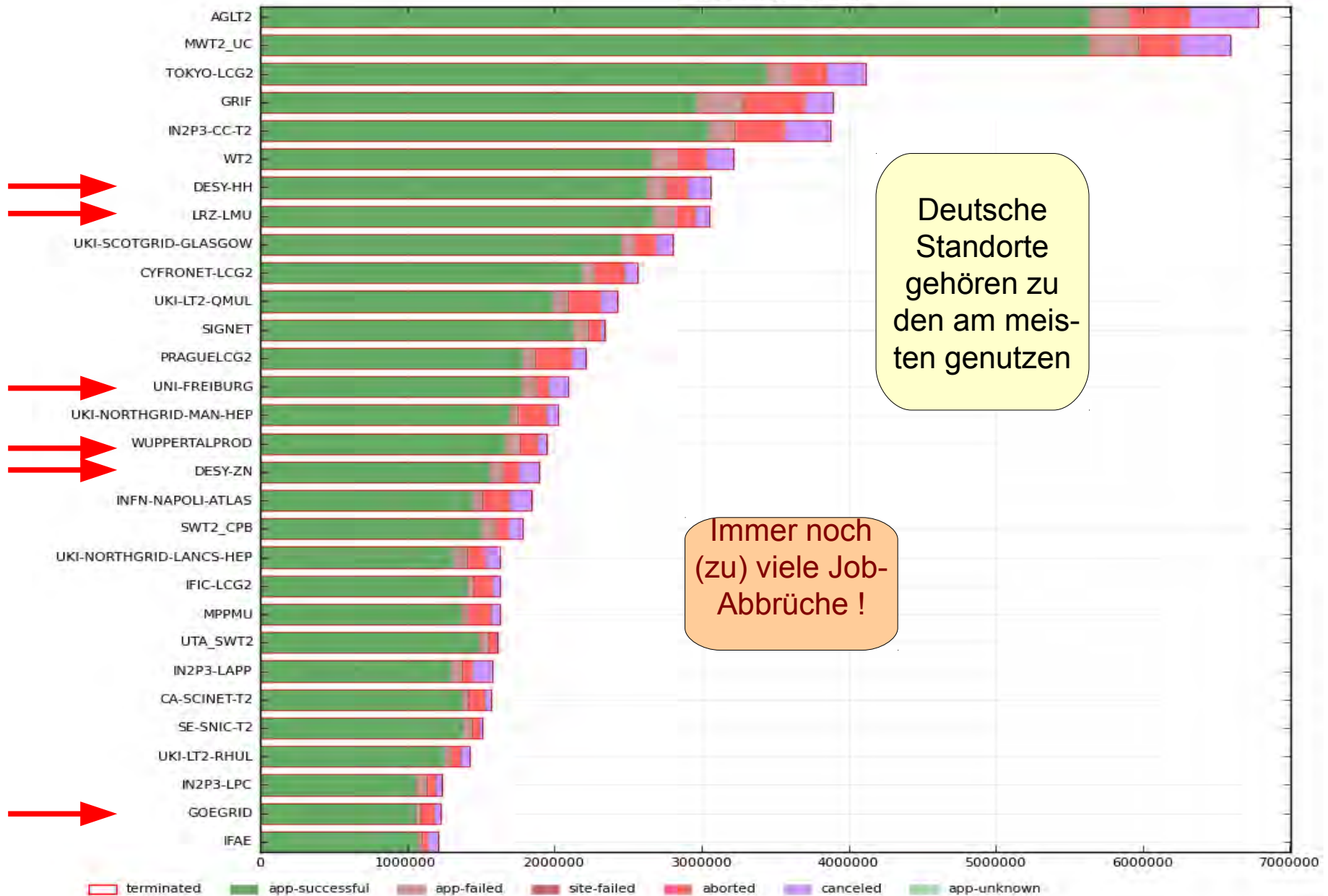
CERN exportiert
den Inhalt einer DVD
in 4 sec !

12'000 Job-Slots bei GridKa
immer gut ausgelastet

ATLAS Tier-2: Job-Statistik

Terminated Jobs per site

Jan. - Oct. 2011



Deutsche Standorte gehören zu den am meisten genutzten

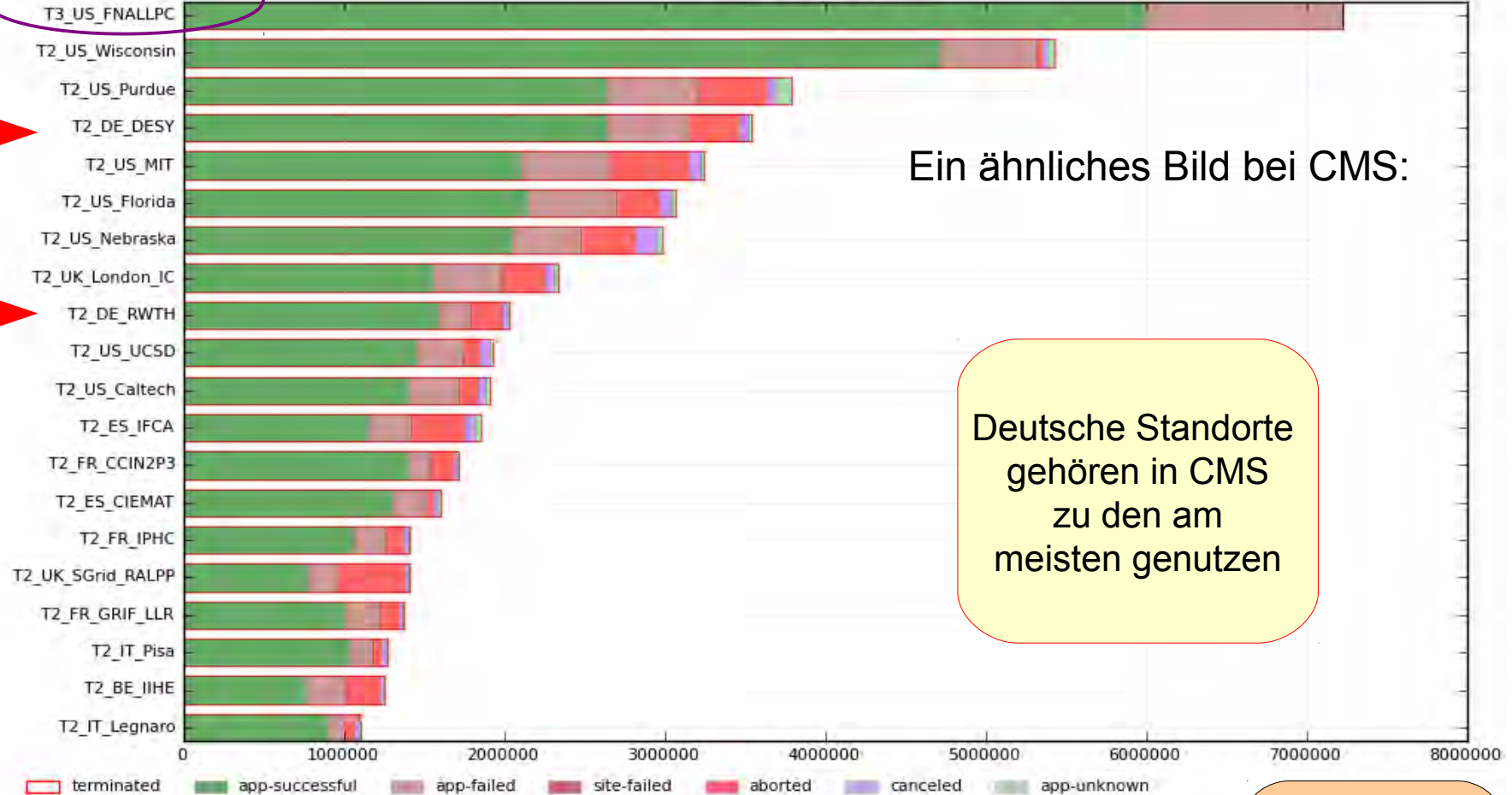
Immer noch (zu) viele Job-Abbrüche !

CMS Tier2: Job-Statistik

Terminated Jobs per site

Jan.- Oct. 2011

amerikanische „NAF“



Ein ähnliches Bild bei CMS:

Deutsche Standorte gehören in CMS zu den am meisten genutzten

Problem mit Job-Abbrüchen auch hier !

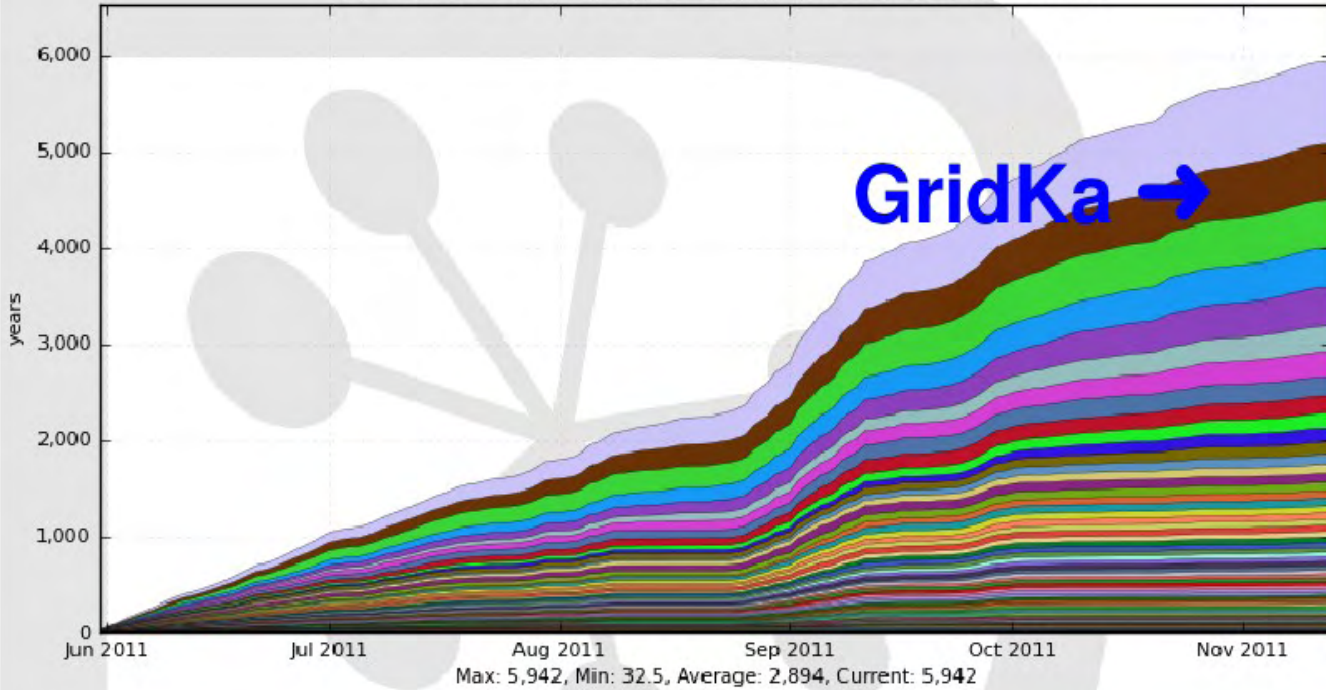
LHCb

LHCb hat spezielles Computing-Modell:

- Physikanalysen an 6 Tier1-Zentren
- Tier2 nur für MC (viele Ressourcen außerhalb WLCG-Pleges)

CPU used by Site

23 Weeks from Week 22 of 2011 to Week 46 of 2011



Größter T1-Beitrag:
GridKa

LCG.CERN.ch	862.0	LCG.PIC.es	156.5	LCG.TCD.ie	70.7
LCG.GRIDKA.de	579.4	LCG.SARA.nl	147.9	LCG.DESY.de	65.2
DIRAC.YANDEX.ru	499.5	LCG.Manchester.uk	124.7	LCG.USC.es	63.5
LCG.CNAF.it	409.2	LCG.UKI-LT2-QMUL.uk	106.7	LCG.GLASGOW.uk	61.0
LCG.IN2P3.fr	404.3	LCG.CNAF-T2.it	93.8	LCG.JINR.ru	59.2
LCG.RAL.uk	270.0	LCG.RAL-HEP.uk	91.1	LCG.Liverpool.uk	58.8
LCG.NIKHEF.nl	255.3	LCG.AUVER.fr	86.2	LCG.GRISU-UNINA.it	52.2
LCG.Krakow.pl	200.4	LCG.Pisa.it	81.1	LCG.NIPNE-07.ro	48.3
LCG.IN2P3-T2.fr	178.0	LCG.UKI-LT2-IC-HEP.uk	80.1	...	plus 68 more

Erheblicher Ressourcen-
Aufwuchs in 2012 vorgesehen
(wegen neuem „Charm-Trigger“)

Generated on 2011-11-14 14:22:27 UTC

Hat es funktioniert ?

Sollten die Anwender beantworten:

D. Charlton, **ATLAS**, EPS HEP 2011

Computing Grid Delivers Physics

Data preparation:

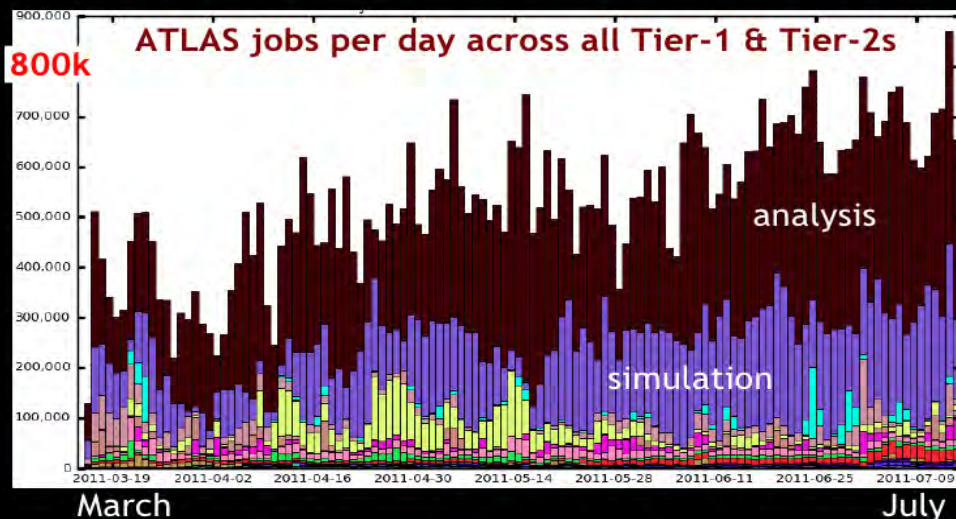
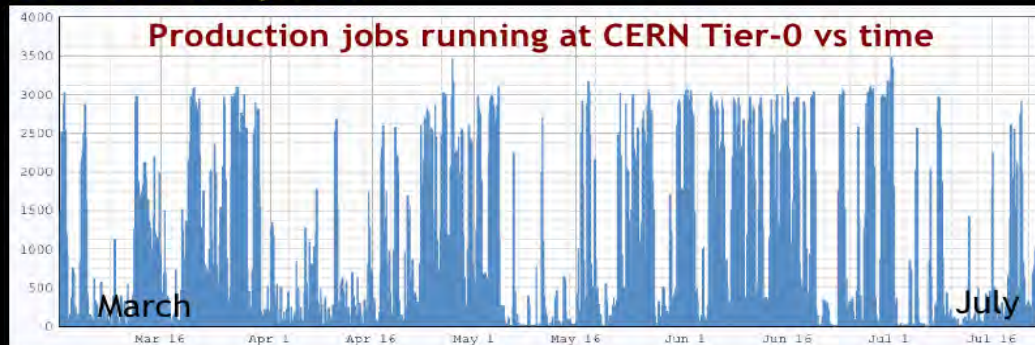
- First-pass reco. at Tier-0 within -2 days
- Calibration/DQ good for physics analysis
- Data analysable on Grid within -1 week

Tier-1 and Tier-2's process $\sim \frac{2}{3}$ M jobs per day

- simulation
- re-reconstruction (campaigns)
- group production (ntuples...)
- physics analysis

The high quality computing system allows us to show results on data taken until the end of June

Payback for the years of investment and hard work



**Report on the Mid Term Review of the
Helmholtz Alliance
"Physics at the Terascale"**

Hamburg, 30 November – 01 December 2009

■ ■ ■

B) GRID COMPUTING

Since the creation of the Alliance, the effectiveness of the German Tier-2 (T2) structures has improved measurably. There are two particular noteworthy aspects:

- 1) The funding, at least in part, of hardware resources at University T2 centres. This has leveraged additional 3rd party funding and explicitly supported the integration of these resources into a national, even international, infrastructure. This is vital for the LHC research programme and would have been impossible without the stimulus provided by the Alliance.

Over the remaining years of the Alliance, the focus on the data access and optimisation for analysis is appropriate. However, more effort should perhaps be focused directly on working towards longer term commitments or support for those aspects where the Alliance has shown clear added value and which can not currently be supported easily by other means,

e.g.:

- Facilitating T2 hardware investments integrated into a national or international infrastructure
- Exploiting existing expertise at DESY and KIT in direct support of the T2 and T3/NAF activities
- Providing high level strategic vision and coordination of the T1+T2+T3 infrastructure.
- Supporting and sustaining strategically important expertise such as that in large scale data management.

Situation in Deutschland

Tier-2 Ressourcen in Deutschland

T2-Ressourcen in Deutschland

Entsprechen in etwa dem **Anteil deutscher Autoren (M&O)**:
ATLAS: 11,5%
CMS: 7,2%

DE	2012		2013 plan	
	CPU (HS06)	Disk (TB)	CPU (HS06)	Disk (TB)
ATLAS				
Desy HH/ZN	12000	1500	12000	1500
Goettingen	3853	1000	3853	1000
Freiburg	4430	783	4430	783
Wuppertal	4430	783	4430	783
LRZ-LMU München	5780	670	5780	670
MPP München	5780	670	5780	670
Sum	36273	5406	36273	5406
CMS	CPU	Disk	CPU	Disk
Desy	14750	1350	14750	1350
Aachen	8875	600	8875	600
Sum	23625	1950	23625	1950

T2-Ressourcen WLCG gesamt (empfohlen durch RSCG & CCRB)

WLCG	2012		2013 plan	
	CPU	Disk	CPU	Disk
ATLAS	266000	47000	289000	53000
DE share	13,6%	11,5%	12,6%	10,2%
CMS	315000	26000	315000	26000
DE share	7,5%	7,5%	7,5%	7,5%

LHC shutdown 2013
 → kaum Zuwachs in 2013

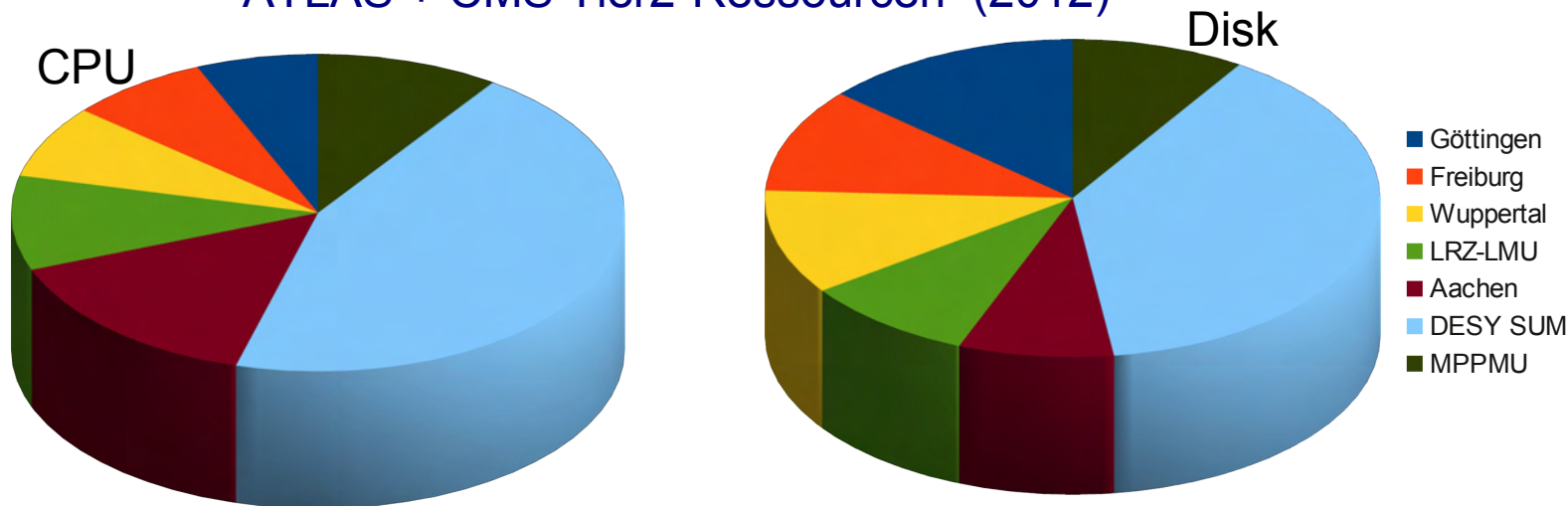
Quellen: GridKa TAB, HGF-A Grid-PB und <http://wlcg-rebus.cern.ch/apps/pledges/resources/>

Aufteilung der deutschen Tier-2-Beiträge

Insgesamt 2012:	CPU (HS06)	%WLCG T2	Disk (PB)	%WLCG T2
ATLAS:	36000	13,6%	5,4	11,5%
CMS:	24000	7,5%	2,0	7,5%

Derzeit nur kleiner Ressourcenzuwachs in 2013 vorgesehen

ATLAS + CMS Tier2-Ressourcen (2012)



Ressourcen an den Universitäten decken im Jahr 2012 **46%** der **CPU**- und **52%** der **Disk**- Anforderungen ab.

Ressourcen für nationale Nutzer an allen Grid-Standorten:

- NAF am DESY
- National Resources @ GridKa („NRG“, entstanden aus D-Grid-Initiative)
- Tier3-Installationen der universitären Gruppen
incl. Ressourcen aus D-Grid-Initiative, aufgestockt aus Eigenmitteln

Vorteile:

- Synergien beim gemeinsamen Betrieb von T2/T3
Tier3-Ressourcen mit Grid-Zugang für nationale Nutzer (des gleichen Experiments)
- priorisierter Zugang zu offiziellen T1/T2- Datensätzen für nationale Nutzer
- Plattenspeicher für nationale Nutzer mit Grid-Anbindung
- NAF @ DESY erlaubt interaktiven Zugang für alle deutschen Nutzer

Physik-Analyse im Tier3-Bereich profitiert stark von den universitären Tier2

↔ Bereitschaft der Uni-Gruppen zum Betrieb von Tier2-Hardware

Grid-Aktivitäten an deutschen Universitäten

teilweise unterstützt durch Projekte der HGF-A „Terascale“

- ermöglichten Diplom- und Promotionsarbeiten im Computing-Bereich
- waren erfolgreich und sichtbar
(Monitoring, Grid-Zugang, Virtualisierung, Nutzung von Cloud-Ressourcen, ...)
- sind ausgezeichnet durch besondere Nähe zur Datenanalyse
und damit zu den speziellen Nutzer - Anforderungen
- stellen zusätzliche Expertise bereit (Rechenzentren und Informatik-Fakultäten)
- bilden Computing-Experten für Wissenschaft und Industrie aus
(auch für CERN und die Grid-Standorte von HGF und MPG)
- sind wichtig für den deutschen Beitrag zur Weiterentwicklung des HEP-Computings

Ein großes Problem:

Sicherung der notwendigen T2-Ressourcen ab 2013
als nationale forschungspolitische Aufgabe

↔ Hardware für die universitären Tier2

zugesagter Eigenbeitrag der Universitäten

Betrieb und Personal zu Bedingungen wie HFG-A „Terascale“
(entspricht in etwa den Hardware-Investitionen bei 4 Jahren Betriebszeit)

Kostenabschätzung durch *Grid Project Board* der HGF-A für
Ersatz alter Hardware und Ressourcenerweiterung:

ca. 125'000€ - 150'000€ / Jahr und T2-Standort,

d.h Finanzierungsbedarf von $\sim 3/4$ M€ /Jahr

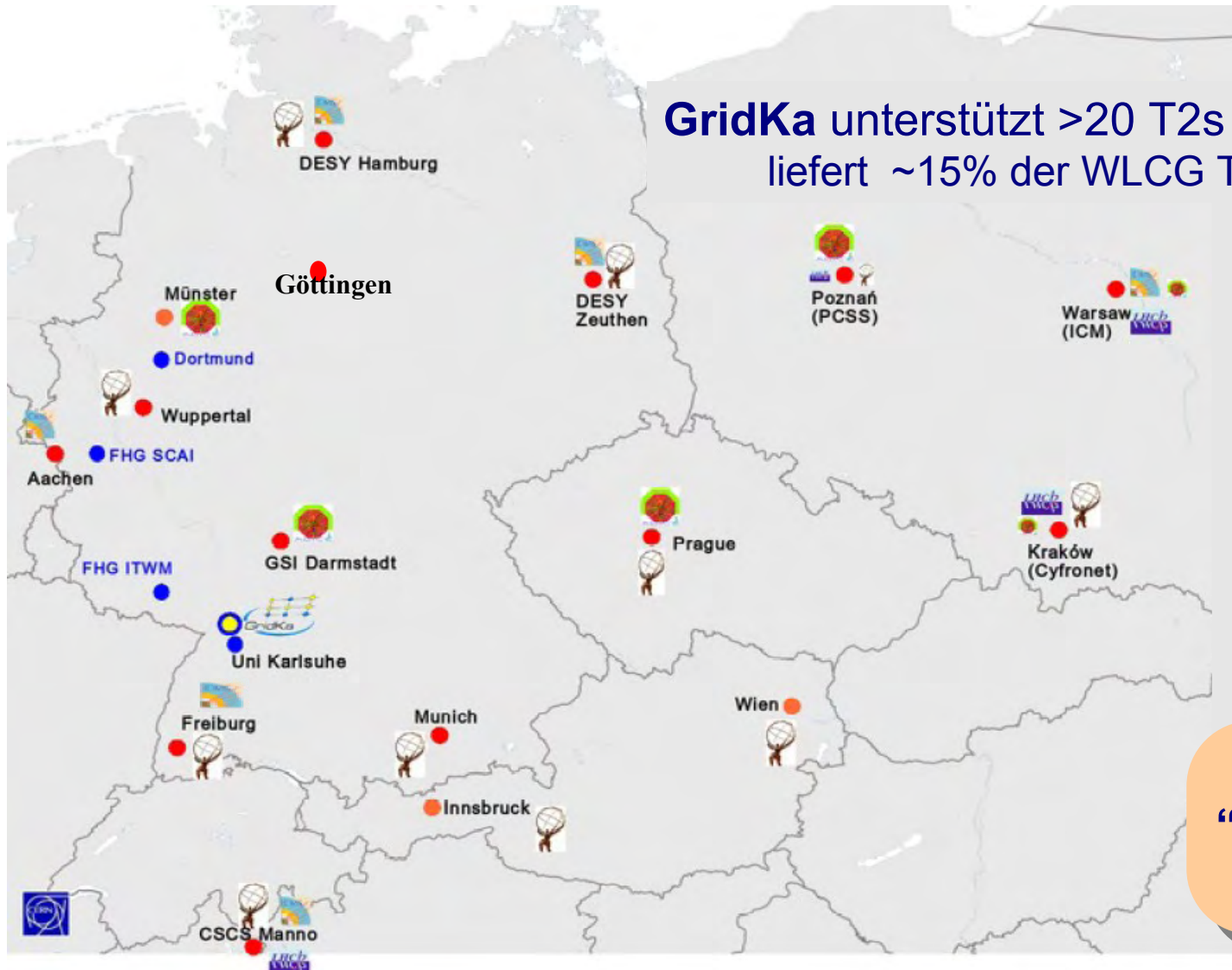
Zusammenfassung und Fazit

- **Grid-Computing für den LHC hat sich bewährt** -
trägt entscheidend zum Erfolg der LHC-Analysen bei.
- Deutschlands Beitrag zum WLCG ist derzeit quantitativ angemessen
und im internationalen Vergleich von hoher Qualität.
- Etablierte und funktionierende Arbeitsteilung beim technischen Betrieb der
Grid-Installationen und zur Abdeckung Experiment-spezifischer Aufgaben
**Personalmittel für Experiment-spezifische Aufgaben durch Verbund-
forschung sind auch in der nächsten Förderperiode notwendig**
Empfehlung des GridKa-TAB, Sitzung vom 10.11.2011
- **Beiträge der Universitäten zur Grid-Infrastruktur** von entscheidender Wichtigkeit
für die Integration der Analyse-Infrastruktur in Deutschland (T1-T2-T3/NAF)
- Lösung zur Finanzierung der **Hardware für die universitären Tier2** benötigt -
als Partner von HGF-Instituten zu Bedingungen wie in der HFG-A ?
Kostenabschätzung liegt vor.

Dank an die Kollegen des
GridKa Technical Advisory Board und des **Grid Project Board der HGF-A**
für die Unterstützung bei der Vorbereitung dieses Vortrags

Zusätzliches Material

GridKa - Deutschland und Umgebung



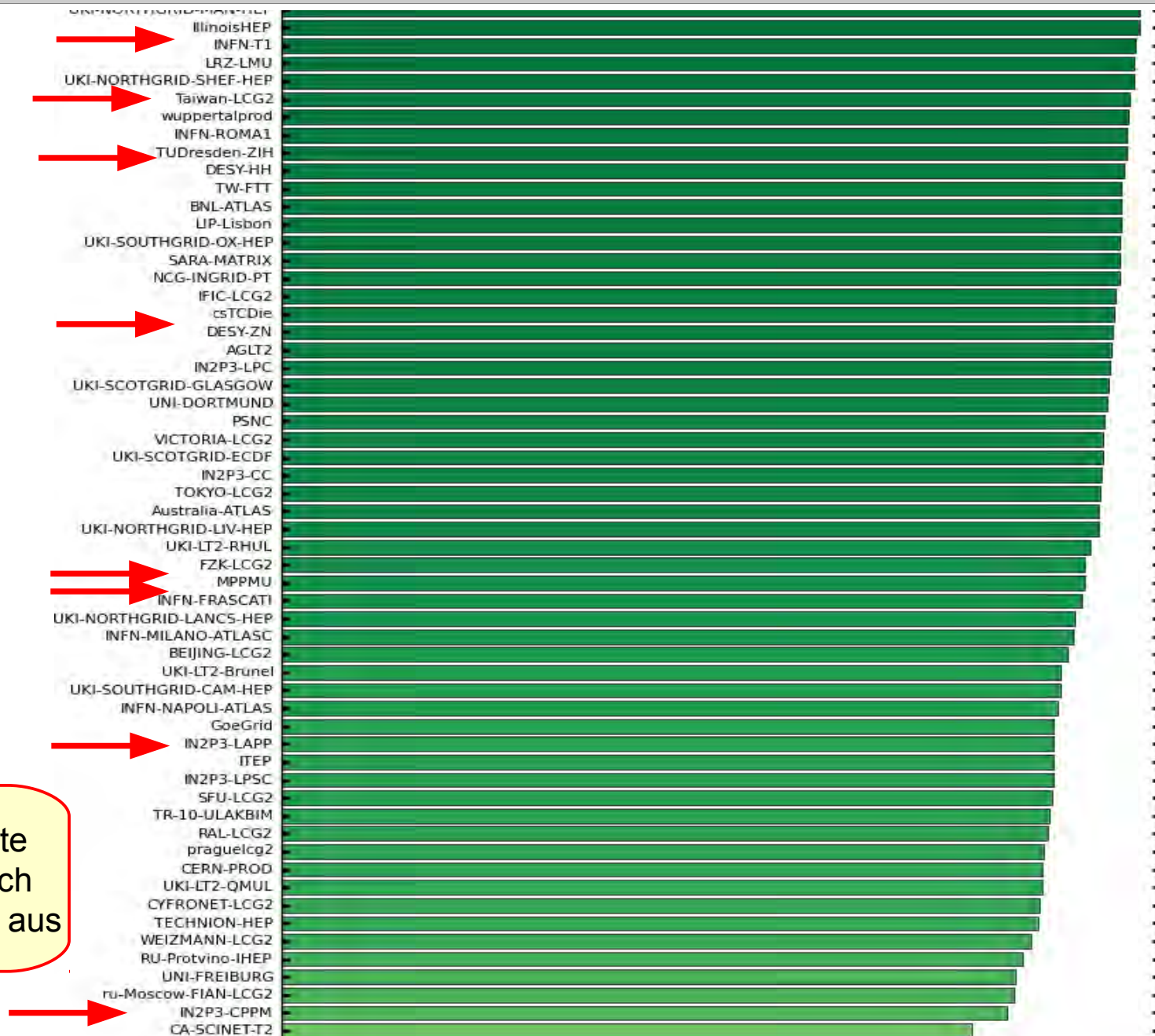
GridKa unterstützt >20 T2s in 6 Ländern,
liefert ~15% der WLCG T1-Ressourcen

Alice T2s
in Russ-
land

**komplexeste
“T1-T2 cloud”
in WLCG**

GridKa ist derzeit das größte T1 im WLCG

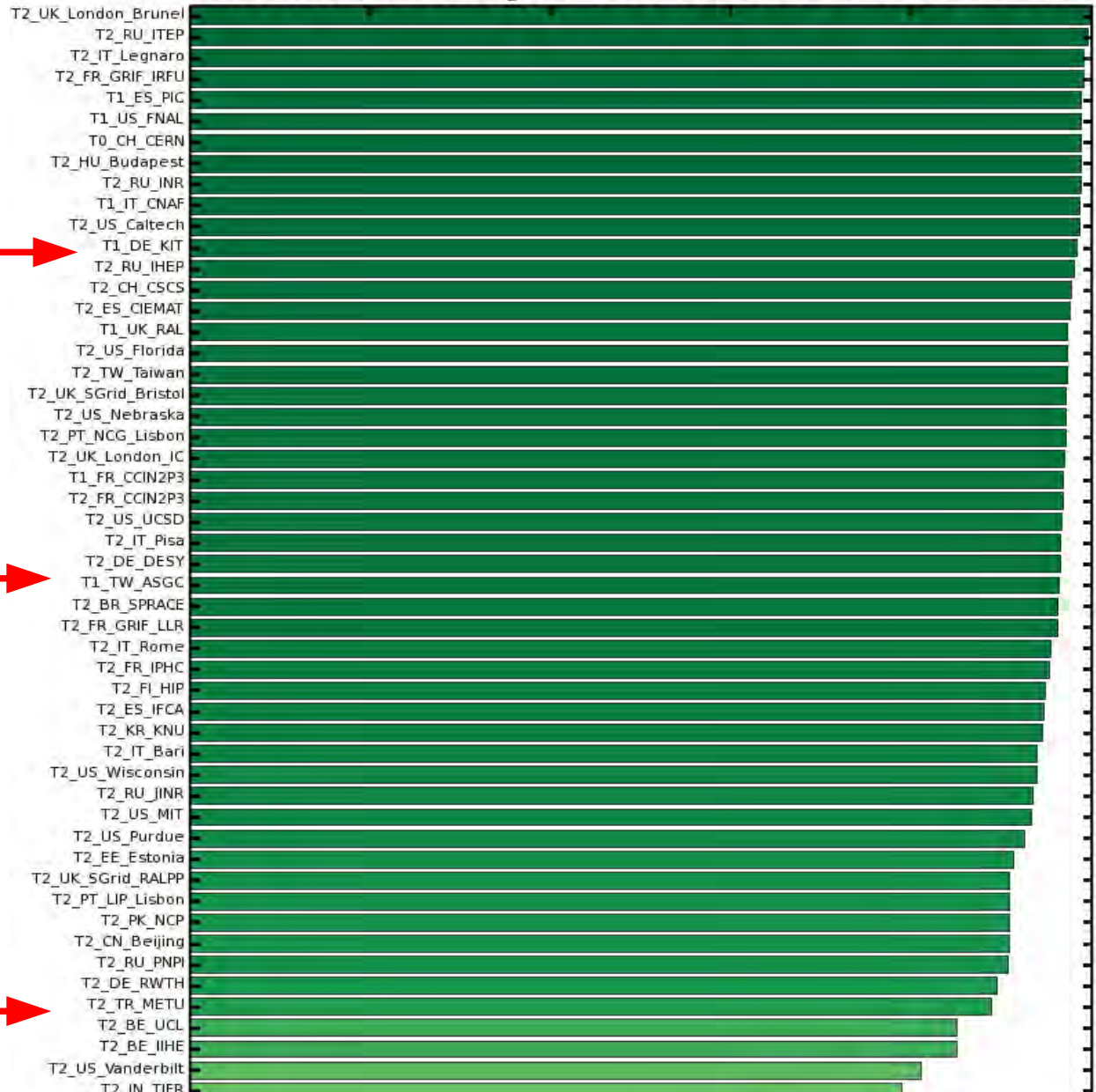
Verfügbarkeit der (ATLAS-)Zentren Jul.-Okt. 2011



Deutsche Standorte zeichnen sich durch hohe Verfügbarkeit aus

Verfügbarkeit der CMS-Sites

Site Availability, 2011-07-01 - 2011-10-30



Deutsche Standorte
zeichnen sich durch
hohe Verfügbarkeit aus

Hat es funktioniert ?

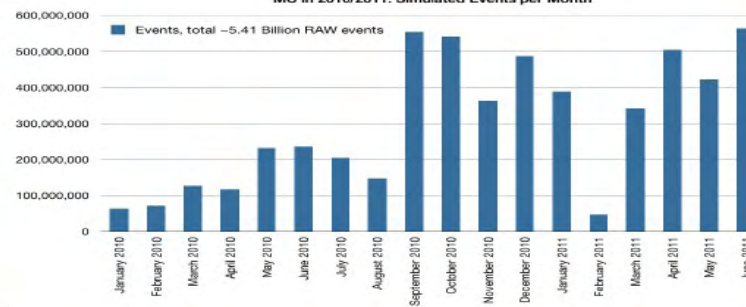
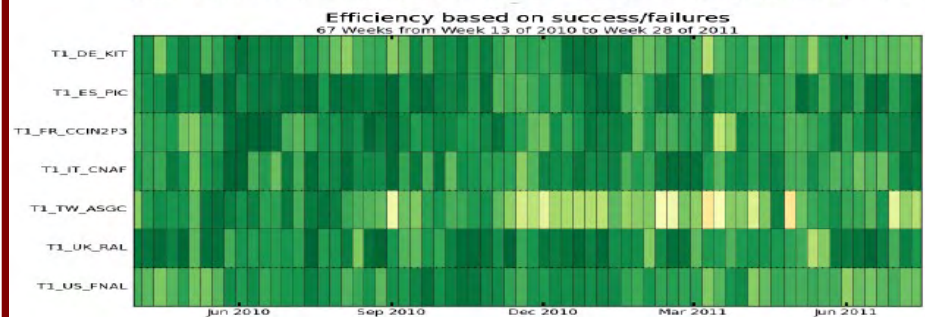
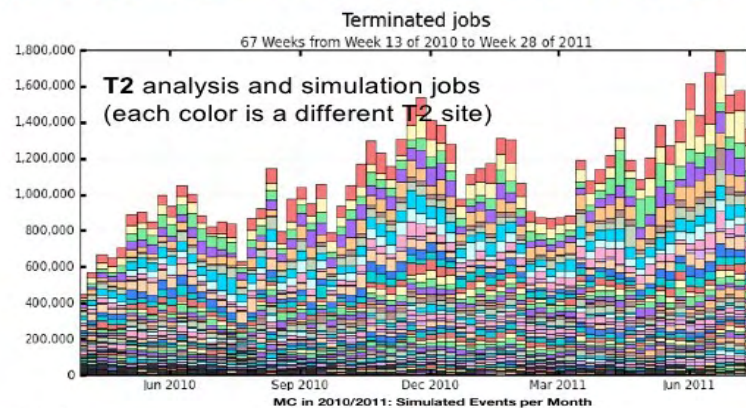
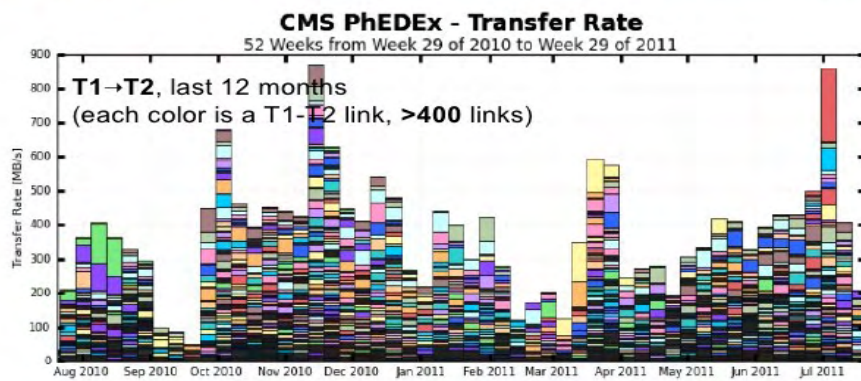
Sollten die Anwender beantworten:

G. Tonelli, **CMS**, EPS HEP 2011



Offline and Computing running smoothly

- Smooth **Tier-0** operation, keeping up with the data taking. Increase in **Tier-1** utilization, for reprocessing and skimming jobs; High usage of **Tier-2** for analysis, **>400 (800)** individual users per week (month). More than 5.4 Billions MC events.



G. Tonelli, CERN/INFN/UNIPI

HEP_2011_GRENOBLE

July 25 2011

36

Autorenanteil

Experiment: Autoren (2010)	D-Autoren (BMBF)	D-Autoren (DESY)	D-Autoren (MPI)	D-Gesamt
ATLAS: 1827	155 (8,5%)	29 (1,6%)	23 (1,3%)	207 (11,3%)
CMS: 1368	60 (4,3%)	39 (2,8%)		99 (7,2%)

Deutsche Beteiligungen an ATLAS und CMS, Quelle: CERN RRB , 11-13. Oktober 2010,

<http://indico.cern.ch/conferenceDisplay.py?confId=93082>.

Deutsche Tier-2-Ressourcen 2012

Tier2	Federation	Pledge Type	ATLAS	% req.	CMS	% req.	LHCb	% req.
ATLAS	DESY	CPU (HS06)	12000	5%				
	DESY	Disk (Tbytes)	1500	3%	<i>Hamburg und Zeuthen</i>			
	FR/WU	CPU (HS06)	8860	3%				
	FR/WU	Disk (Tbytes)	1566	3%	<i>Aufteilung FR:WU = 1:1</i>			
	Goettingen	CPU (HS06)	3853	1%				
	Goettingen	Disk (Tbytes)	1000	2%				
	MPG/LMU Munich	CPU HS06)	11560	4%	<i>Aufteilung MPG:LMU = 1:1</i>			
	MPG/LMU Munich	Disk (Tbytes)	1340	3%	<div style="border: 1px solid red; padding: 2px; display: inline-block;"> \sum CPU: 13% \sum Disk: 11% </div>			
<hr style="border-top: 1px dashed blue;"/>								
CMS	DESY- Aachen	CPU (HS06)			23625	8%	<i>Aufteilung DESY/RWTH = 2:1</i>	
	DESY-Aachen	Disk (Tbytes)			1950	8%		
<hr style="border-top: 1px dashed blue;"/>								
LHCb	DESY	CPU (HS06)					3200	7%
	DESY	Disk (Tbytes)					2	10%

Quelle: <http://wlcg-rebus.cern.ch/apps/pledges/resources/>