

LHC-Computing

Christian Zeitnitz
Bergische Universität Wuppertal

- Status der Tier-Zentren
- Planung ab 2015
- Finanzierung



Computing-Ressourcen

■ Weltweite zugesagte („pledged“) Ressourcen

- CPU: ca. 200.000 CPU Kerne
- Disk: knapp 200 PByte

■ Deutsche Ressourcen

- Anteile am WLCG:
 - Tier-1: 15%
 - Tier-2: 10%
- Anteile in D
 - Tier-1: 60%
 - Tier-2: 40%

Ohne Tier-3 und NAF

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	

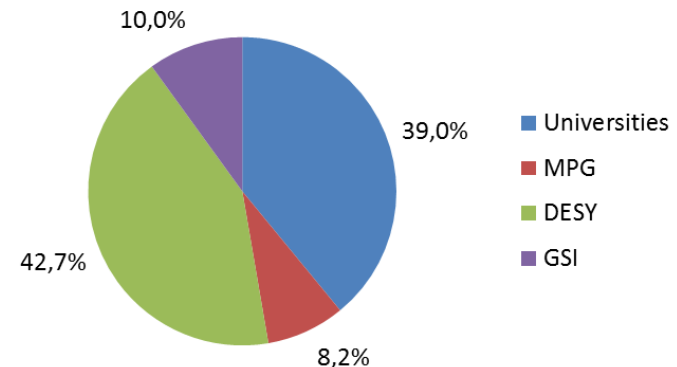
Anmerkung: 1 CPU Kern = 10 HEPSpec06

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

■ Tier Standorte

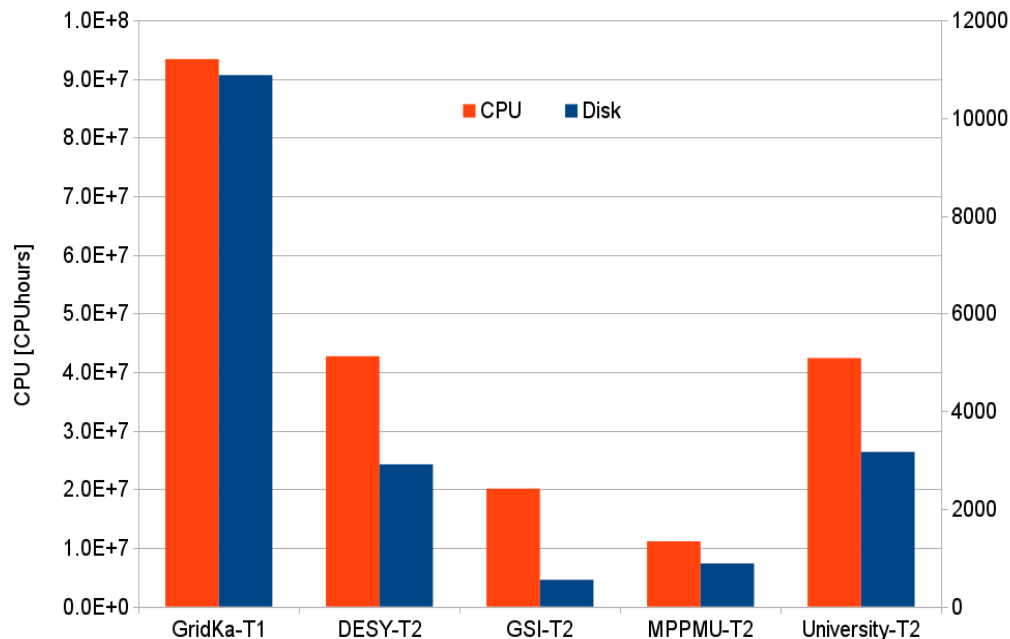
- Tier-1: KIT
- Tier-2
 - HGF: DESY, GSI
 - MPI-M
 - Universitäten: AC, FR, GÖ, M, W

German Tier-2 CPU Resources 2012



Nutzung der Ressourcen in D

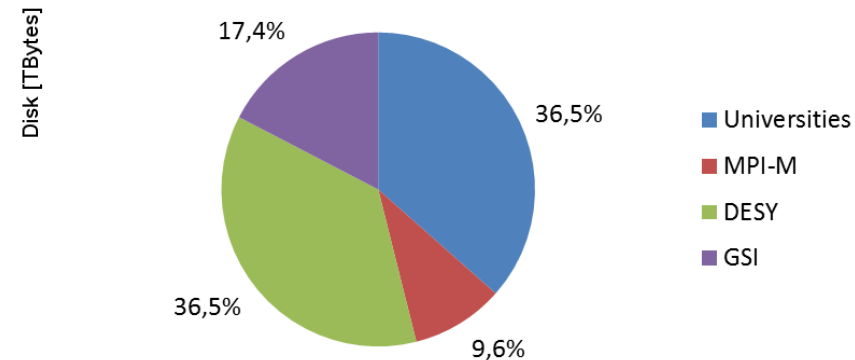
- Deutsche Zentren sehr zuverlässig und erfolgreich
- Tier-1: wichtig für zentral geplante Aufgaben und Datenverteilung
- Tier-2: wichtig für Datenanalyse
- Extrem hohe Auslastung des Tier-1 und der Tier-2
- Tier-2 liefern deutlich mehr CPU Leistung als versprochen
 - Analyse benötigt mehr Rechenleistung als geplant



Genutzte Ressourcen 7/2012-6/2013

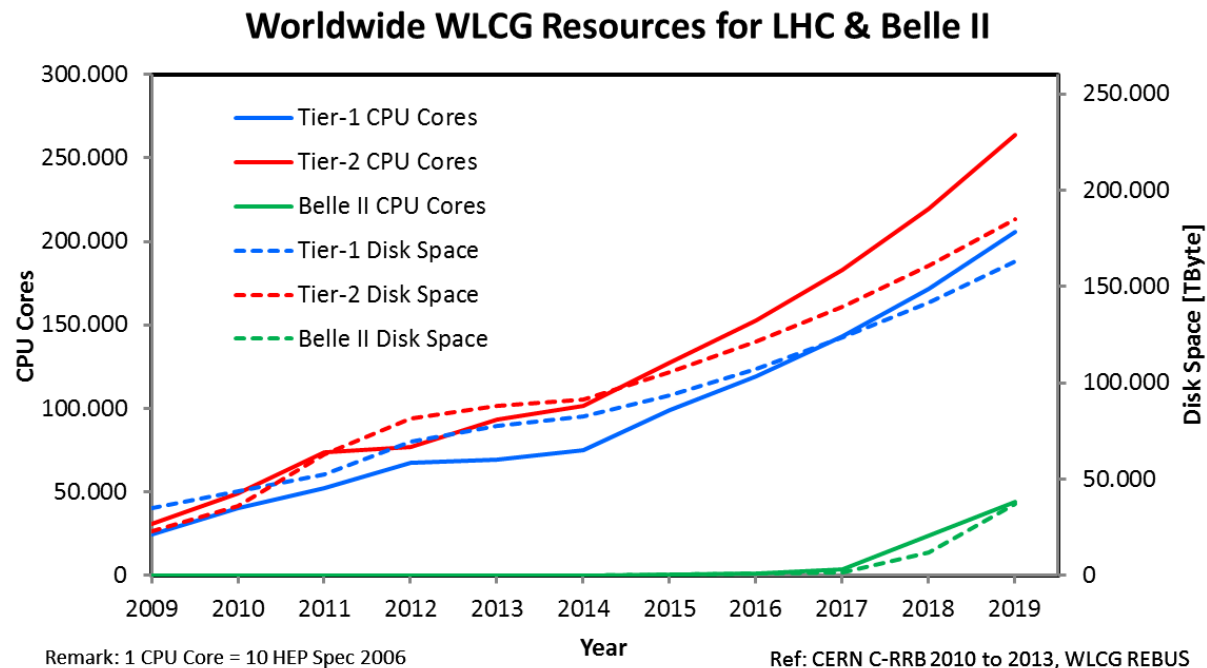
- Insgesamt über 200 Mio. CPU Std
- Tier-1: 46%, Tier-2: 54%

CPU shares of German Tier-2



Weltweite Ressourcenplanung bis 2019

- Ausschließlich installierte bzw. „pledged“ Ressourcen
 - 2014/15: Computing-Resources Review Board (C-RRB) des CERN Okt. 2013
 - 2016-19: 20% Anstieg pro Jahr für CPU (15% für Disk)
 - Abschätzung der Ressourcen berücksichtigt bereits Verbesserungen in den Computing-Modellen und bei der Software



- Annahmen
 - Deutsche Anteile an Tier-1 und Tier-2 bleiben konstant
 - Ressourcensteigerungen werden durch die Technologieentwicklung ausgeglichen
 - Eine NAF entspricht von den Ressourcen einem Tier-2

Entwicklung des LHC-Computings

- Computing-Modelle ändern sich
 - Netzwerk erlaubt große Datentransferraten, zwischen beliebigen Tier-Zentren → Daten werden zu den Jobs kopiert
 - Gesamtsystem wird weniger hierarchisch
 - Computing wird dadurch komplexer und weniger gut planbar
 - Optimierung der Ressourcennutzung sehr wichtig
- Computing Technologie
 - Technologische Anpassung (Hard- und Software) notwendig
 - Nutzung von Vektorisierung, Parallelisierung in normalen CPUs und GPUs
 - Nutzung von Cloud und HPC Systemen
 - Optimierung der Komponenten der Datenspeicherung und Verteilung, Überwachungssysteme
- Effizienzsteigerung um ca. Faktor 4 notwendig
 - Ressourcen wachsen deutlich langsamer als die Datenmengen!

Planungen für 2015-2019 in D

- Tier-1 (GridKa)
 - Ressourcenerhöhung entsprechend der Anforderungen (C-RRB)
 - Betriebskosten und Personal: Eigenmittel (Teil des HGF PoF III Antrags)
 - Investitionsmittel: ca. 16M€ über 5 Jahre
 - Nicht Teil von PoF III, sondern separater Antrag notwendig
- Tier-2 und NAF (9 Standorte)
 - Ressourcenerhöhung entsprechend der Anforderungen (C-RRB)
 - Betriebskosten und Personal
 - HGF Zentren DESY, GSI: Eigenmittel (Teil des HGF PoF III Antrags)
 - MPI M und Universitäten: Eigenmittel
 - Universitäten: Eigenmittel
 - Investitionsmittel: ca. 16M€ über 5 Jahre
 - Nicht Teil von PoF III, sondern separater Antrag notwendig
- Personal für experimentenspezifischen Betrieb
- Personal für die technologische Weiterentwicklung und Anpassung

Bisherige Finanzierung

- Tier-1
 - Investitionsmittel: ca. 4 M€/Jahr (BMBF)
 - Betriebskosten und Personal: ca. 3,8 M€ (HGF PoF I und II)
- Tier-2 und NAF
 - Mittelwerte über 2008-12
 - Investitionsmittel: ca. 2,9M€/Jahr (sehr heterogene Finanzierung)
 - Betriebskosten: ca. 2,6 M€/Jahr (DESY, GSI, MPI M, Universitäten)

■ Personalkosten

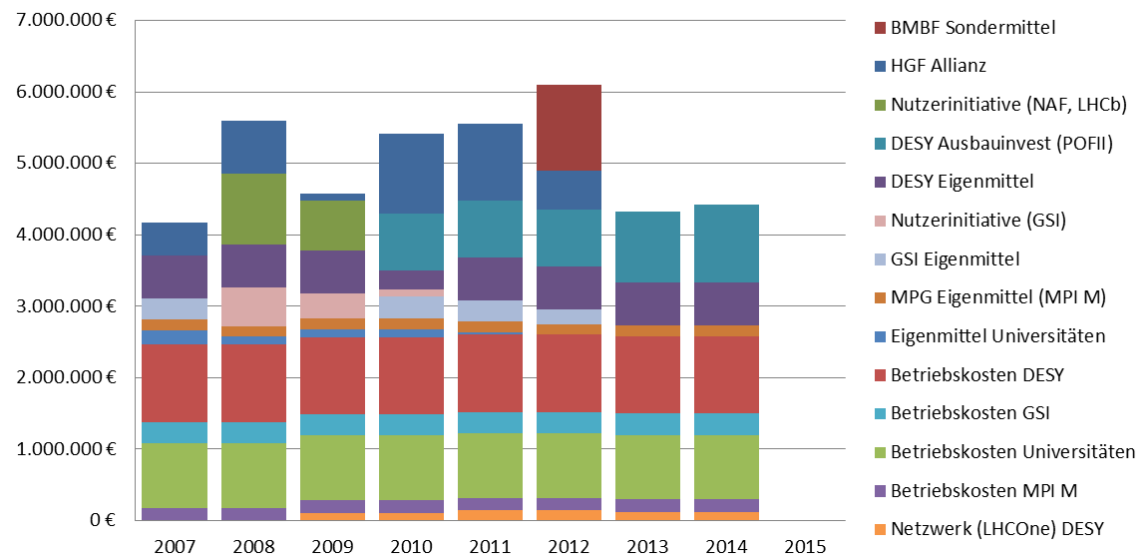
○ Experimentspezifisch

- 19,5 FTE
- Finanziert durch:
BMBF, DESY, KIT, GSI, MPG

○ GRID Tools Entwicklung

- ~10 FTE
- Finanziert durch:
Terascale Allianz, DESY,
KIT, GSI, BMBF

LHC Tier-2 und NAF Invest & Basisbetriebskosten



Finanzierung ab 2015

- Ressourcenanforderungen ab 2015 des Tier-1 GridKa und der 7 Tier-2 in Deutschland
 - 2015: +30% bzw. 25%
 - 2016-19: +20%/Jahr (CPU) und +15% (Disk und Tape)
- Konstantes Budget pro Jahre erlaubt diesen Zuwachs
- Mögliche Finanzierung der Hardware (2015-19: ca. 32M€)
 - Investitionsantrag für ein LK2-Großgerät bei der HGF durch KIT, DESY und GSI
 - Beinhaltet auch Mittel für universitäre Tier-2

Falls nicht erfolgreich, muss eine andere Lösung gefunden werden!
- Übernahme der Betriebskosten durch die beteiligten Institutionen
- Weitere Kosten durch Infrastruktur (z.B. Netzwerk LHC OPN , LHCOne) und Personal

Zusammenfassung

- Die deutschen Tier-1, Tier-2 Zentren und die NAFs sind essenziell für den wissenschaftlichen Erfolg der deutschen Gruppen bei der Analyse der LHC-Daten
 - **Ohne Computing gibt es keine Analyse der Daten in Deutschland!**
- Die deutschen Tier-Zentren gehören zu den zuverlässigsten weltweit
- Anforderungen an Zentren ab 2015 sind bekannt
 - Computing-Resources Review Board des CERN im Oktober 2013)
- **Finanzierung der notwendigen Hardware ist zur zeit unklar**
- Benötigtes Personal
 - Technischer Betrieb – Finanzierung durch betreibende Institutionen
 - Experimentspezifischer Betrieb (BMBF, HGF, MPG)
 - Technologische Weiterentwicklung - unklar
 - Verbesserung und technologische Anpassung der Software - unklar

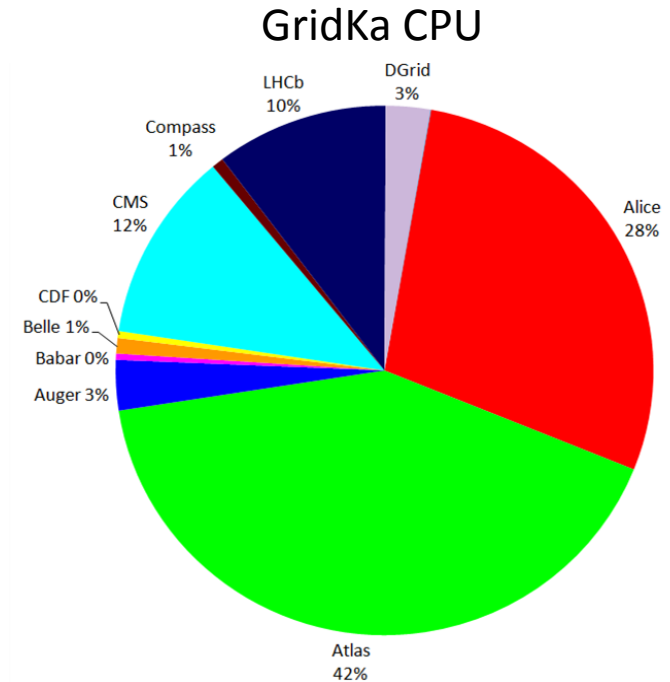
Weitergehende Informationen

1. [Computing in der Hochenergiephysik in Deutschland Bestandsaufnahme und Aussicht](#)
Bericht des erweiterten GRID Project Boards der Terascale Allianz im Auftrag des KET:
2. [Kurzfassung: Computing in der Hochenergiephysik \(Oktober 2013\)](#)

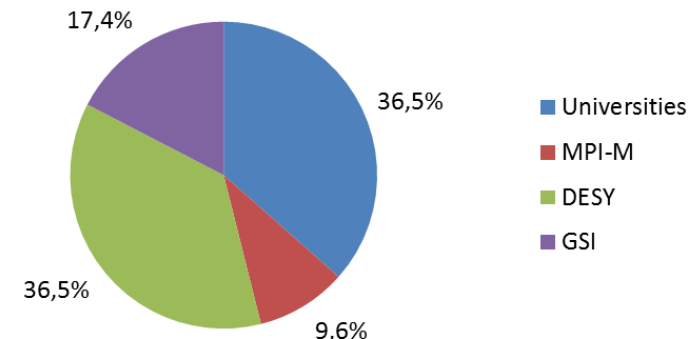
Backup

Betrieb der deutschen Tier-Zentren

- GridKa CPU Nutzung
 - LHC Experimente nutzen 90% der Ressourcen
 - Über 90 Mio. CPU Std
- Tier-2
 - Über 110 Mio. CPU Std
 - Faktor 2 mehr als zugesagt
- Hinzu kommen die National Analysis Facilities am DESY und der GSI
- Computing-Modelle ändern sich
 - Netzwerk erlaubt große Datentransferraten, zwischen beliebigen Tier-Zentren
 - Daten werden zu den Jobs kopiert
 - Computing wird dadurch komplexer und weniger gut planbar



CPU Usage of German Tier-2



Personal

- Technisches Betriebspersonal ist bereits in die Betriebskosten eingerechnet
- Experimentspezifisches Personal
 - Betrieb der Experimentsoftware, Jobüberwachung, Datenmanagement und Unterstützung der Nutzer
 - Übernahme von zentralen Computingaufgaben im Experiment
 - Finanzierung (wie bisher)
 - BMBF (2015-18): 12 FTE
 - ATLAS (7), CMS (4), LHCb (1)
 - KIT (1,5), DESY (4,5), GSI (1) und MPG (0,5)
 - Insgesamt 19,5 FTE

Personal für Weiterentwicklung

- Datenmenge wird bis 2019 um Faktor 10 steigen. Computing Ressourcen nur um Faktor 2,5
 - Optimierung um ca. Faktor 4 auf verschiedenen Ebenen bereits eingeplant
- Finanzierung der Weiterentwicklung der Computing-Technologie und der benötigten Dienste
 - Nutzung neuer Technologien und Optimierung des Betriebs
 - Finanzierung bisher : HGF Allianz „Physics at the Terascale“ (bis 2015) und Eigenmittel DESY, KIT und GSI

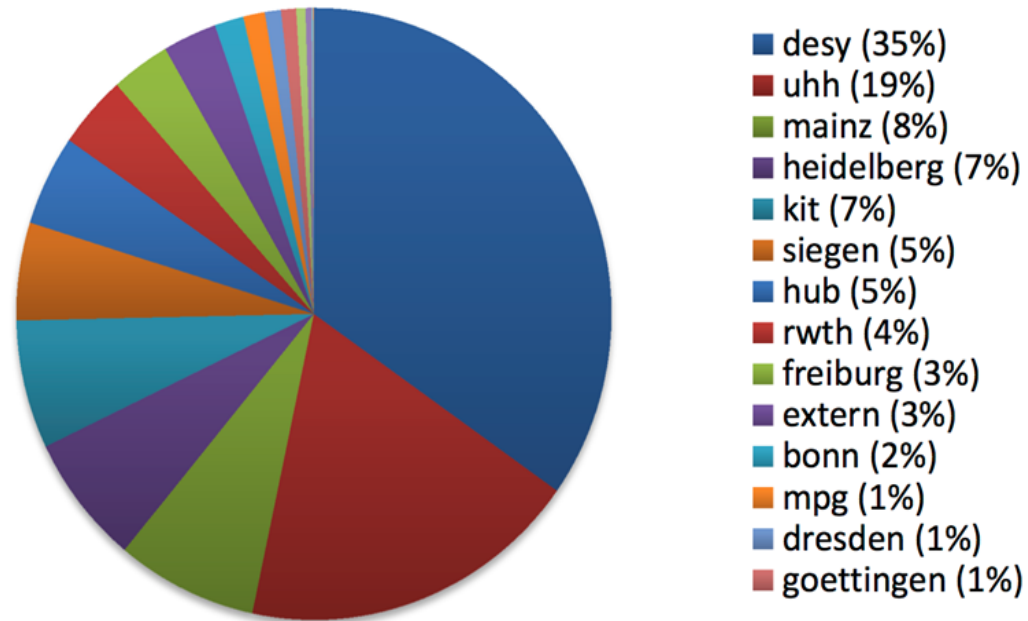
Bedarf etwa 10 Stellen

- Anpassung der Rekonstruktions-, Analyse- und Simulationssoftware an neue Technologien
 - Optimierung der Software für moderne Prozessoren und Nutzung von GPUs
 - Reine Servicearbeit ohne direkten Bezug zur Analyse!

Bedarf etwa 7 Stellen

NAF Nutzung am DESY

NAF usage by institutes
1.10.2010-1.4.2013 / Lustre and CPU combined



Starke Nutzung durch die Uni Hamburg kommt durch die Einbindung der eigenen Ressourcen in die NAF zustande

„Pledged“ LHC-Ressourcen 2013

■ Weltweite Ressourcen

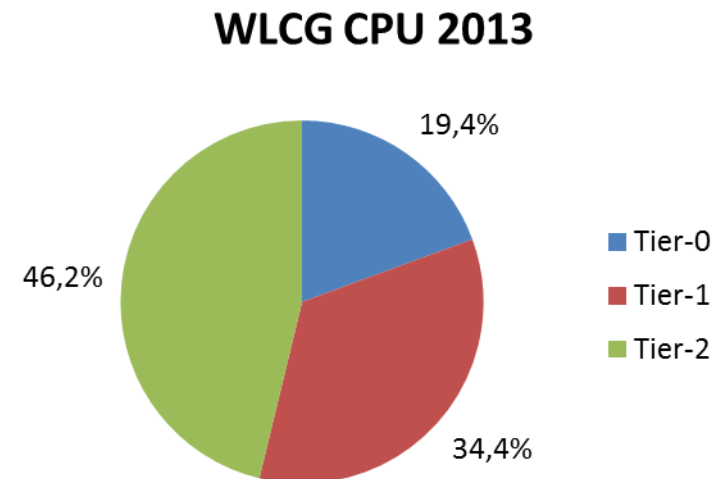
	Tier-0	Tier-1	Tier-2
CPU Kerne	39.200	69.500	93.400
Disk (PByte)	31	78	88
Tape (PByte)	68	107	

1 CPU Kern = 10 HEPSpec06

- Summe CPU: ca. 202.000 CPU Kerne
- Summe Disk: ca. 197 PByte

■ Aufteilung auf die Zentren

- Deutsche Anteile
 - Tier-1: ca. 15%
 - Tier-2: ca. 10%

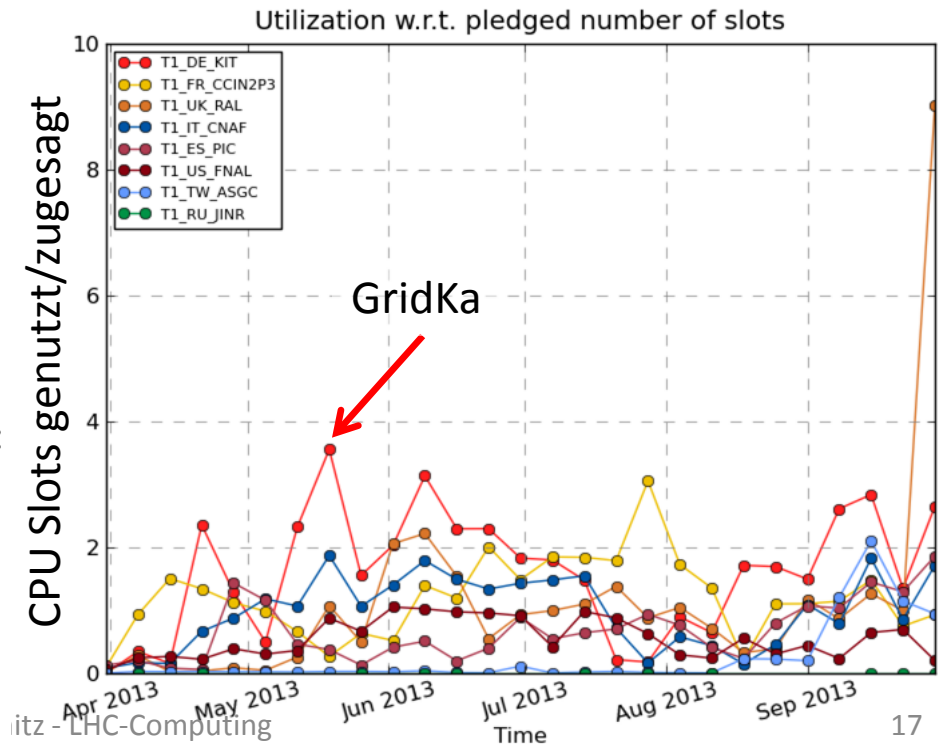
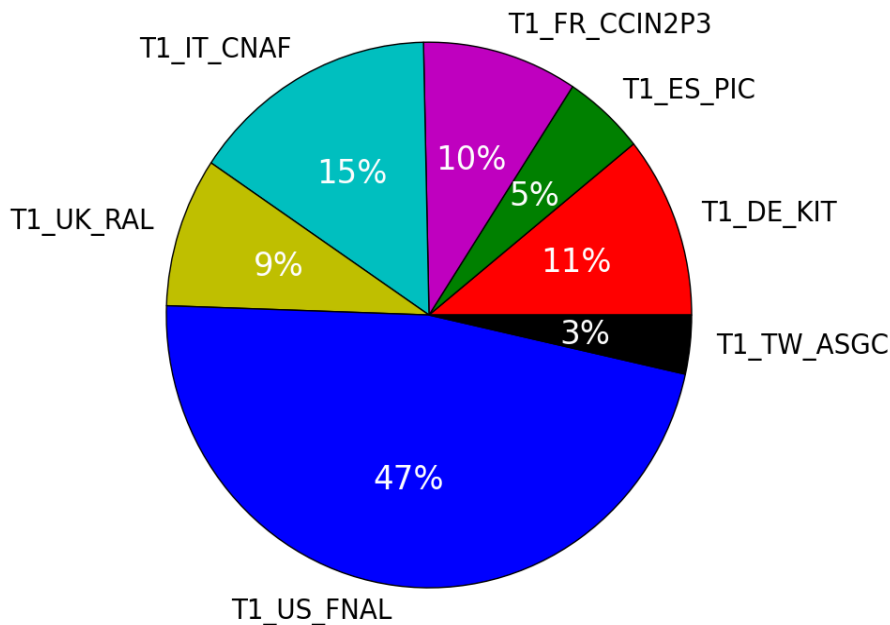


Tier-1 Betrieb

■ Beispiel: CMS

- Gespeicherte Daten an den Tier-1 Zentren (auf Tape)
 - Sehr zuverlässige Speicherung der wichtigen Datensätze
 - Deutscher Anteil: 11%
- Genutzter Anteil der CPU Ressourcen
 - Im Mittel etwas oberhalb der zugesagten Ressourcen

Custodial Tape Storage Overview



Tier-1/2 Betrieb

- Beispiel: ATLAS
 - An GridKa angeschlossene Tier-2 (auch Tschechien, Schweiz, Polen, Slowakei und Österreich)
 - CPU Zeit (Okt 2012-Sept 2013)
 - Zentrale Produktion: 118 Mio. CPU Std
 - Analyse: 32 Mio. CPU Std
 - Notwendige Datentransfers: ~4,5 PBytes/Monat (knapp 1,7 GByte/Sekunde)