

Überwachung und Performance des DESY Tier 2-Zentrums für CMS

Benedikt Mura¹, Birgit Lewendel², Christian
Autermann¹, Christian Sander¹, Christoph Wissing², Florian Bechtel¹,
Hartmut Stadie¹, Peter Schleper¹, Roger Wolf¹ und Yves Kemp²

DPG Tagung Freiburg
März 2008



¹ Institut für
Experimentalphysik,
Universität Hamburg

² Deutsches Elektronen-
Synchrotron (DESY)

Einleitung

- CMS Datennahme beginnt im Sommer
- Computing-Zentren müssen bereit sein, die Daten
 - zu übertragen
 - zu speichern
 - zu prozessieren
- Verschiedene Aktivitäten finden statt um dies zu testen und zu demonstrieren

Inhalt

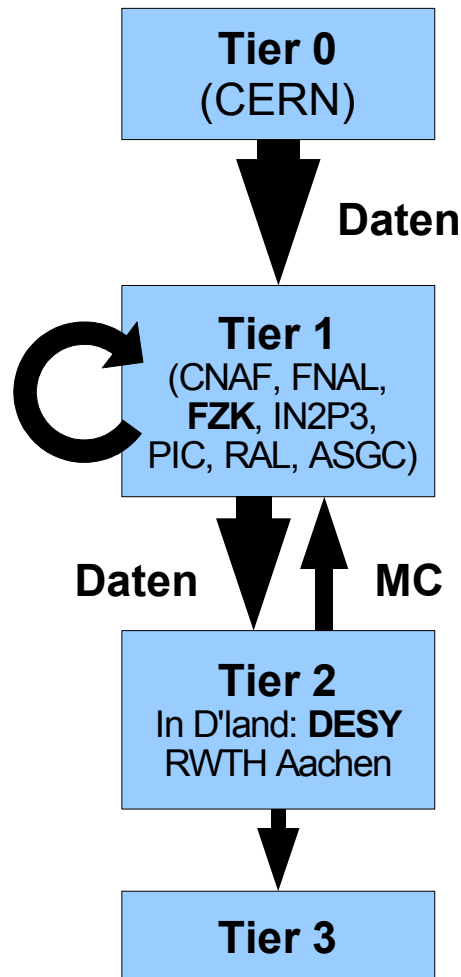
- Das CMS Computing Modell
- Das DESY Tier 2 für CMS
- Funktionstests und Monitoring
- Performance
 - bei *Site*-Tests
 - im CSA '07

WLCG & CMS Computing Modell

Worldwide
LHC Computing Grid

Wesentliche Aufgaben

Tier Struktur



- **Tier 0**
 - Online Rekonstruktion, Daten Archivierung
 - Zentrale Grid/Computing Services
- **Tier 1**
 - Re-rekonstruktion, Daten- & MC-Archivierung
 - Produktion selektierter Datensätze
 - Zentrale Grid/Computing Services
- **Tier 2**
 - Monte Carlo Produktion
 - Speicherung von Analyse-Datensätzen
 - Physikanalyse
- **Tier 3:** Physikanalyse

Das DESY Tier 2 für CMS

Grid-Computing @ DESY

- gLite basierte Grid-Infrastruktur
- Anwendungen verschiedener Experimente
 - CMS, Atlas, Zeus, H1, ILC, Icecube u.a.m.
 - „Durchschnittliches“ Tier 2 für CMS
- Worker Nodes
 - 820 CPU Kerne für alle VOs (ca. 1/3 CMS Anteil)
 - 2 GB RAM pro Kern
 - Betriebssystem Scientific Linux 4, 32 Bit
 - ~ 1100 kSpecInt2k

CMS spezifische Dienste

- Datenbank Cache (squid)
- Datenmanagement Werkzeug (PhEDEx)

Zusammenarbeit DESY/Uni HH

Ressourcen für CMS

	CPU (kSI2k)	Speicher (TB)
2007	300	50
2008	600	170
2009	1000	340
2010	1800	530

Funktionstests & Überwachung

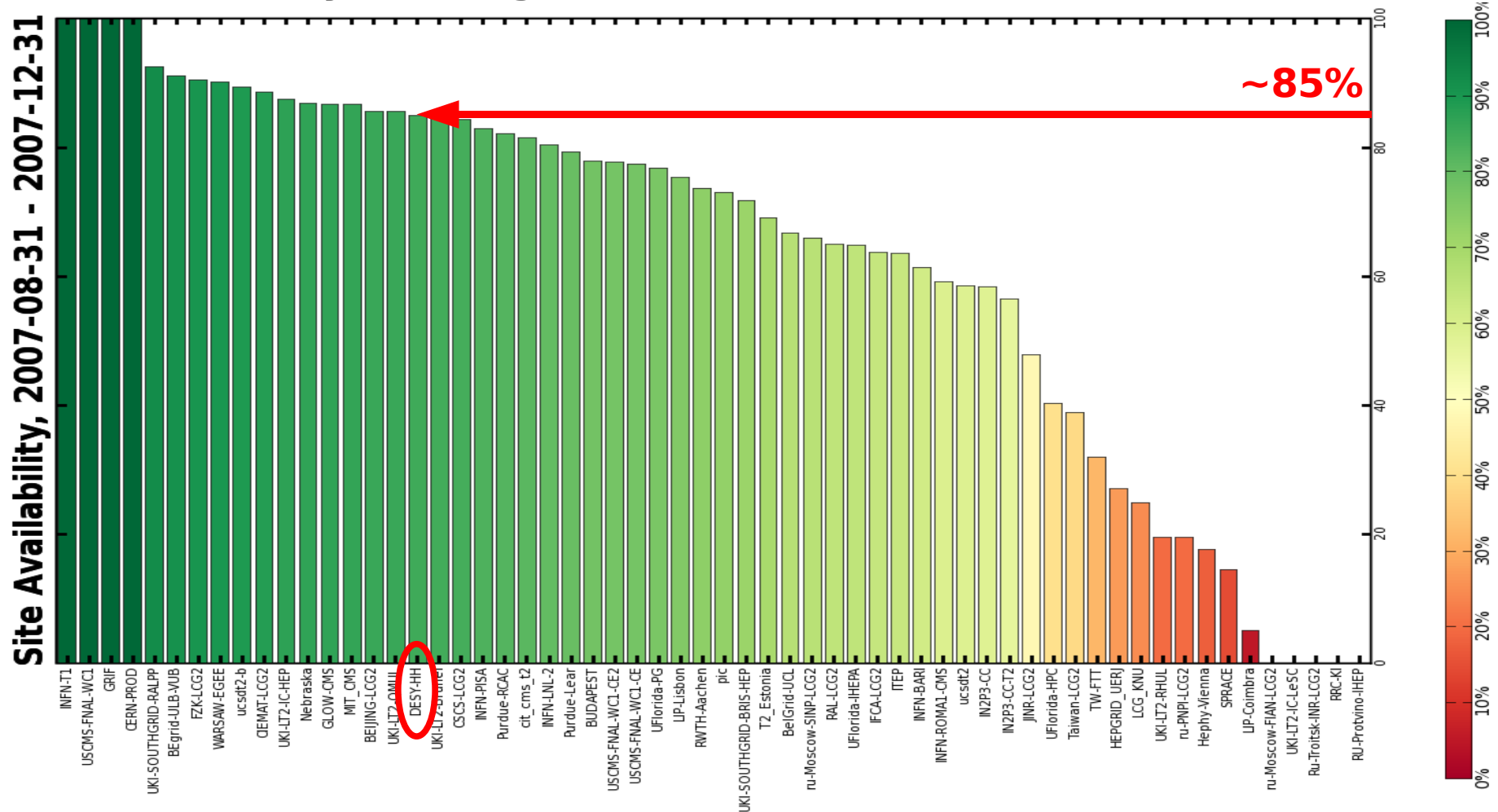
- CMS weite, zentralisierte und automatisierte Tests überprüfen ständig z.B.
 - (1) Konfiguration/Erreichbarkeit der Grid-Dienste
 - (2) CMS-Job Ausführung (CRAB)
 - (3) Datentransfers
 - Monitoring der Ergebnisse
 - Übersicht der Resultate auf Webseiten verfügbar
 - Häufige Überprüfung (Uni HH, Monitoring Schichten)
 - Schnelle Rückmeldung an Administratoren bei Fehlern
- **Verbesserung der Verlässlichkeit**

Site Availability Monitoring

- Satz von Skripten, läuft etwa alle 4h auf allen Sites
- Mehrere CMS spezifische Tests für z.B.
 - CMS-Software Installation
 - Monte Carlo Produktions-Jobs
 - Datenbank Zugriffe
- Schlägt ein kritischer Test fehl, ist die *availability* bis zum nächsten erfolgreichen Test gleich Null.
- Aus Ergebnissen werden offizielle *Rankings* generiert

Site Availability Rangliste

Site Availability Ranking: Tier 1 & Tier 2, September bis Dezember '07



- Position im oberen Drittel
- Rate > 90% wünschenswert

CSA07

- CMS **C**omputing, **S**oftware and **A**nalysis Challenge 2007
 - Test des Computing Modells mit 50% der erwarteten Datenrate bei LHC Betrieb mit niedriger Luminosität
 - Beiträge der Tier 2-Zentren
 - Datentransfertestes (*Link Commissioning*)
 - Verbindung zwischen Tiers zum Transfer von MC Datensätzen mussten erst „commissioned“ werden (s.u.)
 - Monte Carlo Produktion & Transfer zum Tier 1
 - Speicherung gefilterter Datensätze (*Skims*)
 - (Individuelle Daten Analyse)

Link Commissioning Anforderung im CSA07

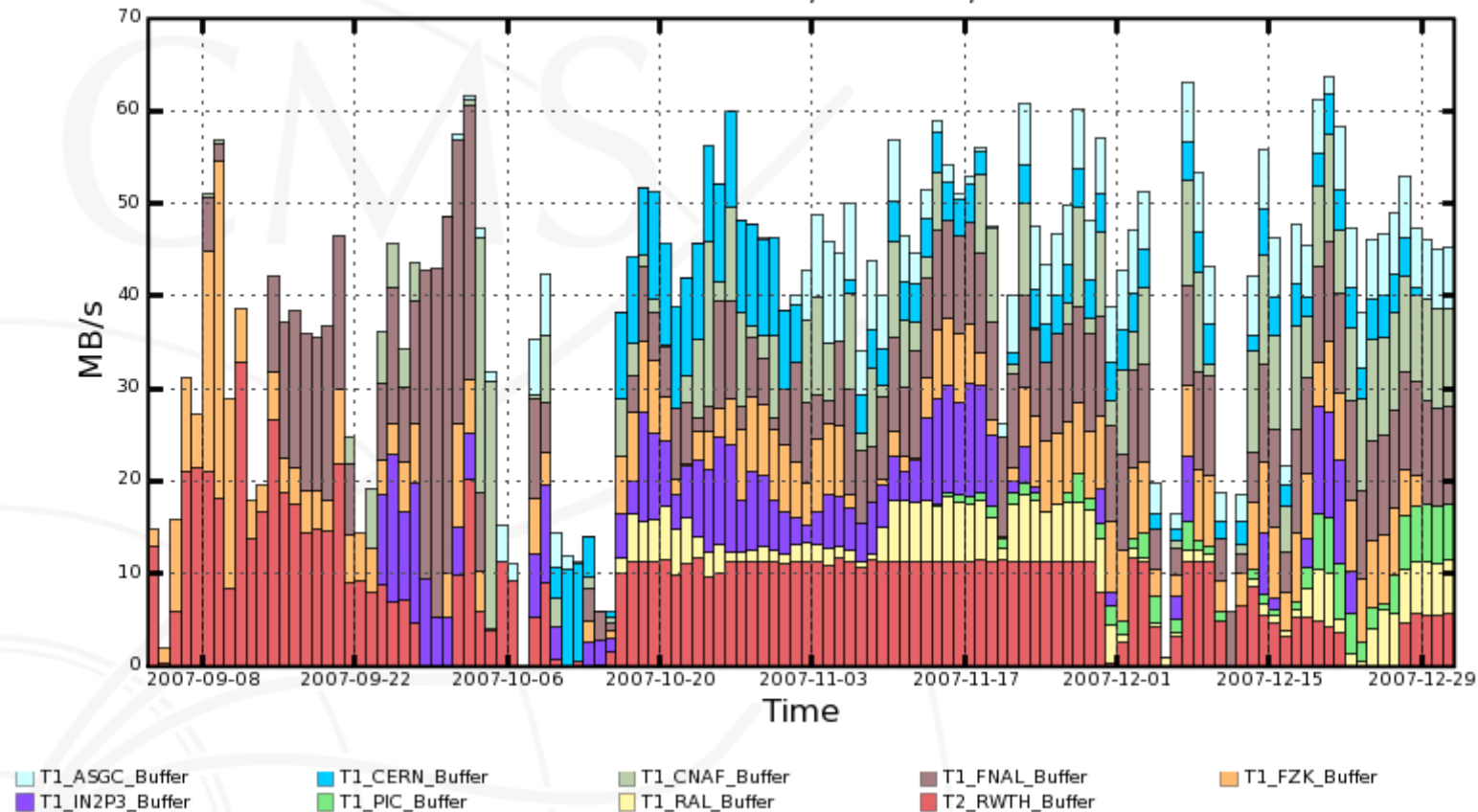
Link aktiviert bei: >300 GB/Tag an 4 von 5 Tagen und > 1.7 TB insgesamt
Deaktiviert bei: 7 Tage am Stück mit < 300GB/Tag
(300 GB/Tag $\approx$ 3.5 MB/s)

Link Commissioning: Rate

Datentransfertest zum DESY, September-Dezember '07:

CMS PhEX - Transfer Rate

17 Weeks from 2007/35 to 2008/00 UTC



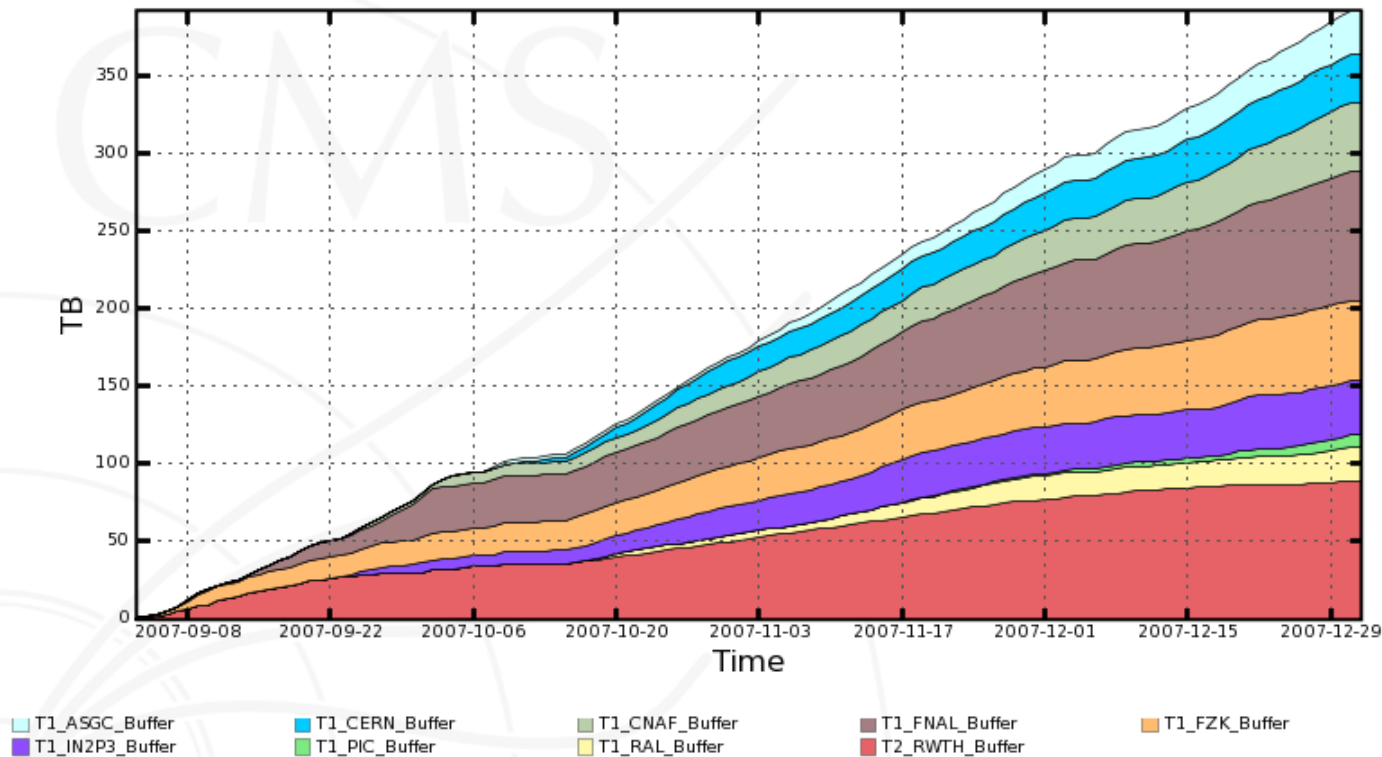
- Tests mit **allen Tier 1 Zentren** und RWTH Aachen
- Durchschnittlich 40 MB/s für alle 9 Verbindungen
- Rate pro Verbindung: ca. 5-10 MB/s → Anforderungen erfüllt

Link Commissioning: Volumen

Datentransfertest zum DESY, September-Dezember '07:

CMS PhEx - Cumulative Transfer Volume

17 Weeks from 2007/35 to 2008/00 UTC

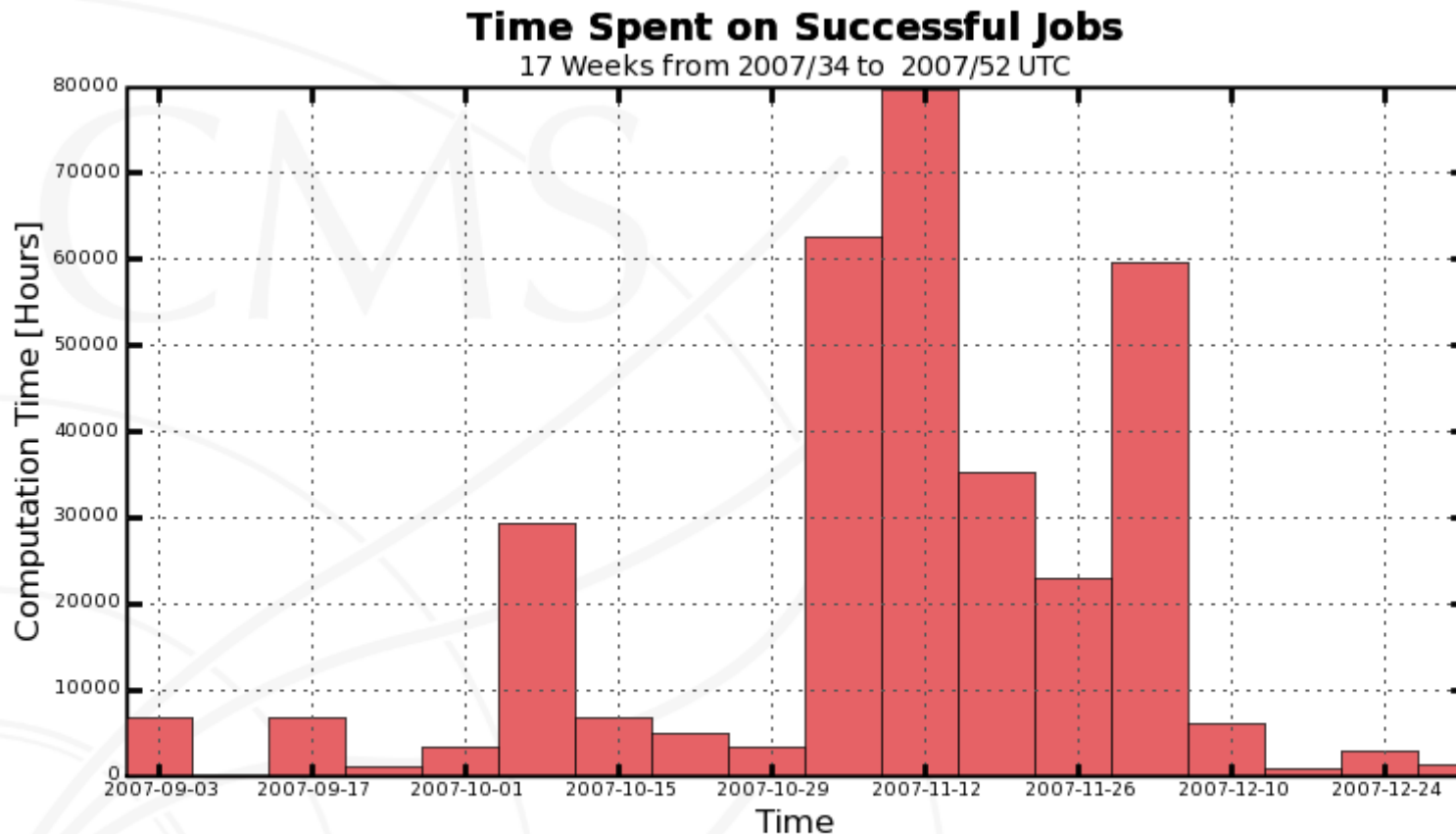


- Durchschnitt: 392 TB Daten in 17 Wochen ≈ 3.2 TB/Tag ≈ 0.3 Gb/s
- Computing Modell Annahme: ~ 5 TB/Tag $\rightarrow$ Bleibt zu demonstrieren, ABER
 - Zusätzliche Transfers von „echten“ Datensätzen
 - Netzwerkverbindung und Storage System noch nicht ausgelastet
 - $\rightarrow$ genügend Steigerungspotential vorhanden

Monte Carlo Produktion I

Monte Carlo Produktion, September-Dezember '07:

Verbrauchte Rechenzeit (erfolgreiche Jobs)



- Durchschnitt: 19558h/Woche $\approx$ 115 CPUs ständig im Einsatz
 - Maximal 80000h/Woche $\approx$ 476 CPUs
- Großer Teil der CPU Zeit wird für Monte Carlo Produktion genutzt

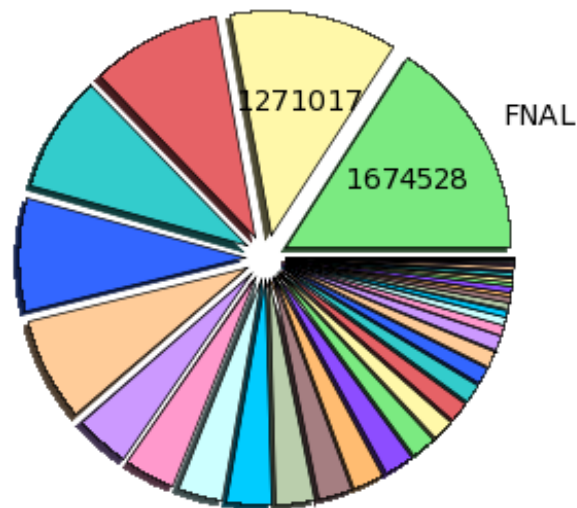
Monte Carlo Produktion II

Monte Carlo Produktion, September-Dezember '07:

Anteile der verschiedenen Zentren

Time Spent on Successful Jobs (Sum: 10551752 Hours)

17 Weeks from 2007/34 to 2007/52 UTC
Wisconsin



Am DESY produziert:

- etwa 12 TB MC Daten
- ca. $3,2 \cdot 10^7$ Ereignisse

(Computing-TDR: $\sim 10^8$ Ereignisse/Tier 2/Jahr)

Zentren nicht gleichmäßig ausgelastet, da verschieden Produktions- Teams zugeteilt



10. DESY (332497h; ca. 3%)

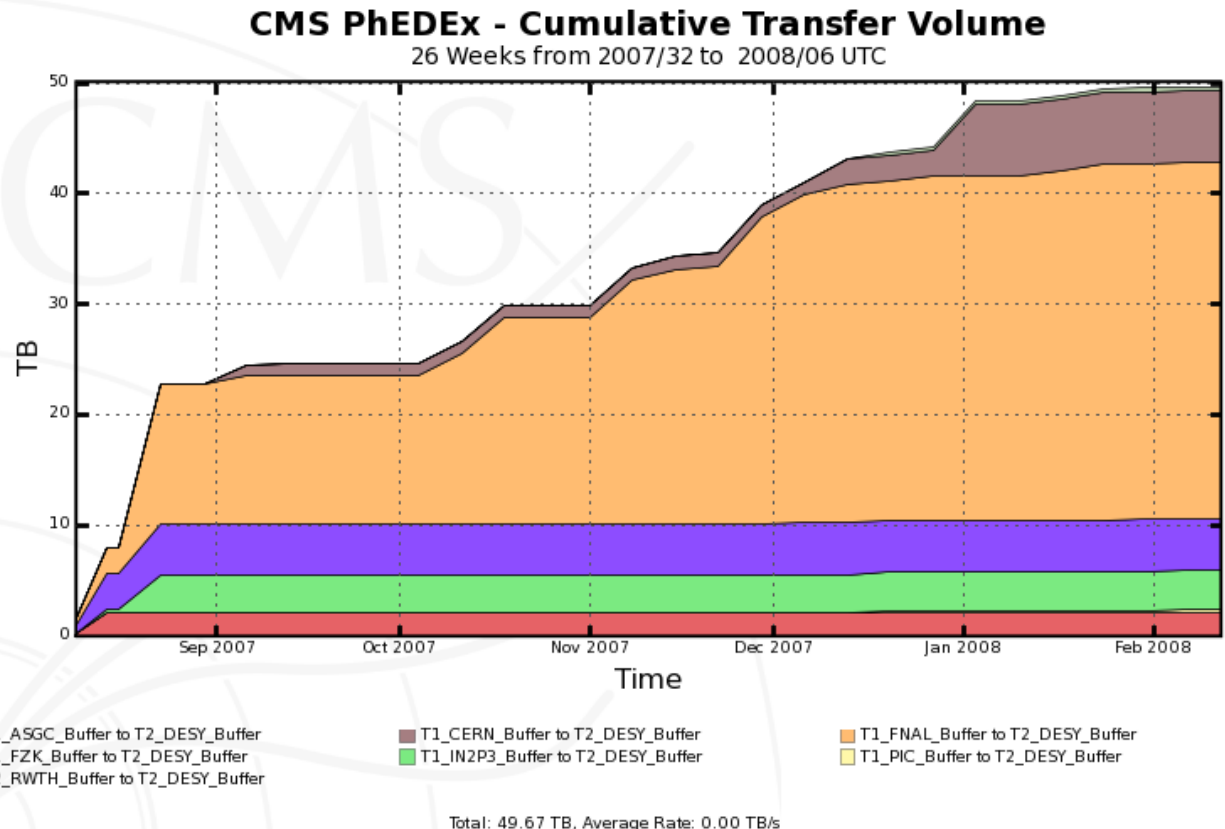
FNAL (1674528)	Wisconsin (1271017)	MIT (984366)	IN2P3 (890827)	Florida (861314)
CERN (791432)	Purdue (423022)	Pisa (389161)	Nebraska (345458)	T2_DESY (332497)
UCSD (297595)	INFN (252822)	Spain_CIEMAT (238526)	Caltech (231366)	WARREN (185543)
Legnaro (170382)	PIC (154617)	FZK (151824)	T2_Belgium (137301)	Belgium_UCL (122105)
London_IC_HEP (108435)	Cornell (70556)	Estonia (56391)	London_Brunel (51218)	Vienna (46630)
BUDAPEST (41349)	RWTH (37481)	Bari (35921)	FIU-PG (29424)	Taiwan (23574)
CSCS (23371)	RAL (22037)	London_RHUL (21494)	ASGC (20547)	Rome1 (16701)
SPRACE (12236)	RutherfordPPD (11266)	TTU (6862)	LIP-Lisbon (5390)	GRIF (2749)
Spain_IFCA (2171)	LIP-Coimbra (149)	unknown (79)	Bristol (0)	PNPI (0)

Speicherung von Datensätzen

- Gesamtes Transfervolumen August-Dezember: ~50 TB
- CSA07 Datensätze: ~23 TB (von ca. 107 TB)
- Physik-Analyse auf zahlreichen Datensätzen, insbesondere Top-Physik Skims

DESY beheimatet zudem

- ~0.6 TB private Benutzerdaten
- temporäre Dateien aus der MC Produktion



Zusammenfassung & Ausblick

- Ständiges Monitoring stellt Funktionalität des Tier 2 sicher
 - Gute Ergebnisse bei *Site Availability* Tests
- Im CSA07
 - Erfolgreiches *Link Commissioning*
 - Umfangreiche Monte Carlo Produktion
 - Speicherung und Analyse von Physik-Datensätzen
- Anforderungen an ein Tier 2 erfolgreich erfüllt

Ausblick

- CSA08 im May stellt neue Herausforderungen an Computing Infrastruktur
- CSA08 finaler Test vor den ersten Daten?!