Überwachung und Performance des DESY Tier-2-Zentrums für CMS

Benedikt Mura[1], Birgit Lewendel[2], Christian Autermann[1], Christian Sander[1], Christoph Wissing[2], Florian Bechtel[1], Hartmut Stadie[1], Peter Schleper[1], Roger Wolf[1] und Yves Kemp[2]

DPG Frühjahrstagung Freiburg
März 2008

[1] Institut fürExperimentalphysik,Universität Hamburg
[2] Deutsches Elektronen-Synchrotron (DESY)

GEFÖRDERT VOM
Bundesministerium für Bildung und Forschung

Einleitung

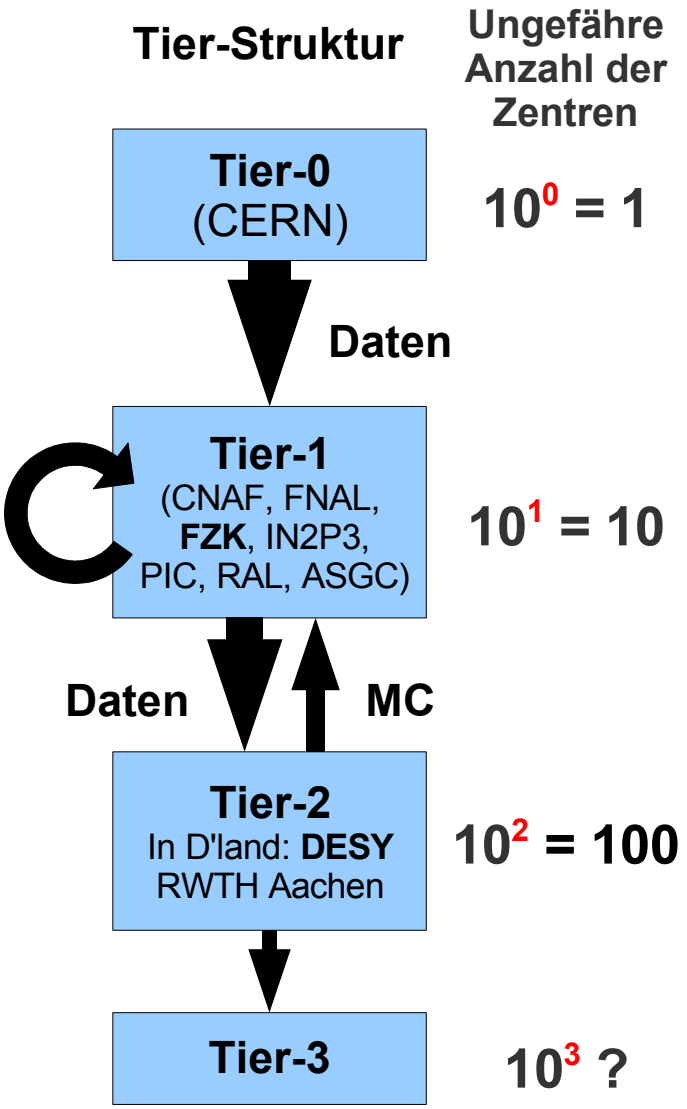
- CMS Datennahme beginnt im Sommer
- Computing-Zentren müssen bereit sein, die Daten
  - zu übertragen
  - zu speichern
  - zu prozessieren
- Verschiedene Aktivitäten finden statt, um dies zu testen und zu demonstrieren

Inhalt

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

WLCG & CMS Computing-Modell

Worldwide LHC Computing Grid



Wesentliche Aufgaben

- **Tier-0**
  - Online-Rekonstruktion, Datenarchivierung
  - Zentrale Grid-/Computing-Dienste
- **Tier-1**
  - Re-Rekonstruktion, Daten- & MC-Archivierung
  - Produktion selektierter Datensätze
  - Zentrale Grid-/Computing-Dienste
- **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
- Knoten
  - 820 CPU-Kerne für alle Gruppen (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
- Überwachung der Ergebnisse
  - Übersicht der Resultate auf Webseiten verfügbar
  - Häufige Überprüfung (Uni HH, Monitoring-Schichten)
  - Schnelle Rückmeldung an Administratoren bei Fehlern

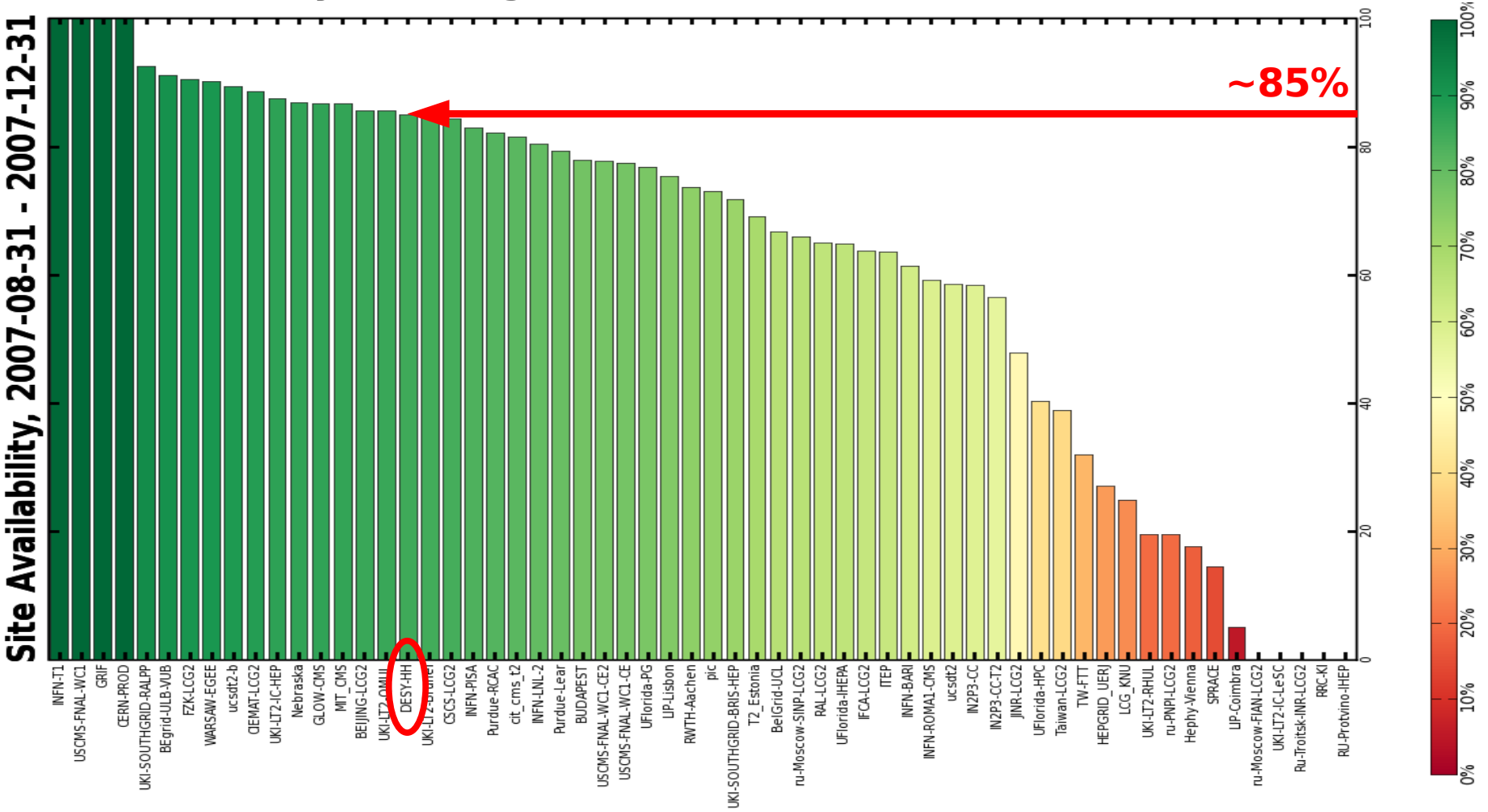
→ Gewährleitung der Verfügbarkeit

Site Availability Monitoring

- Test der Verfügbarkeit (*availability)* des Zentrums
  - Satz von Skripten, läuft etwa alle 4h auf allen Sites
  - Mehrere CMS-spezifische Tests für z.B.
    - CMS-Software Installation
    - Monte-Carlo-Produktions-Jobs
    - Datenbankzugriffe
  - Schlägt ein kritischer Test fehl, ist die Verfügbarkeit bis zum nächsten erfolgreichen Test gleich Null.
  - Ergebnisse dienen als Qualitätsmassstab für die Zentren
    - Erstellung von *availability* Ranglisten

Site Availability: Rangliste

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



- Position im oberen Drittel
- Rate > 90% wünschenswert

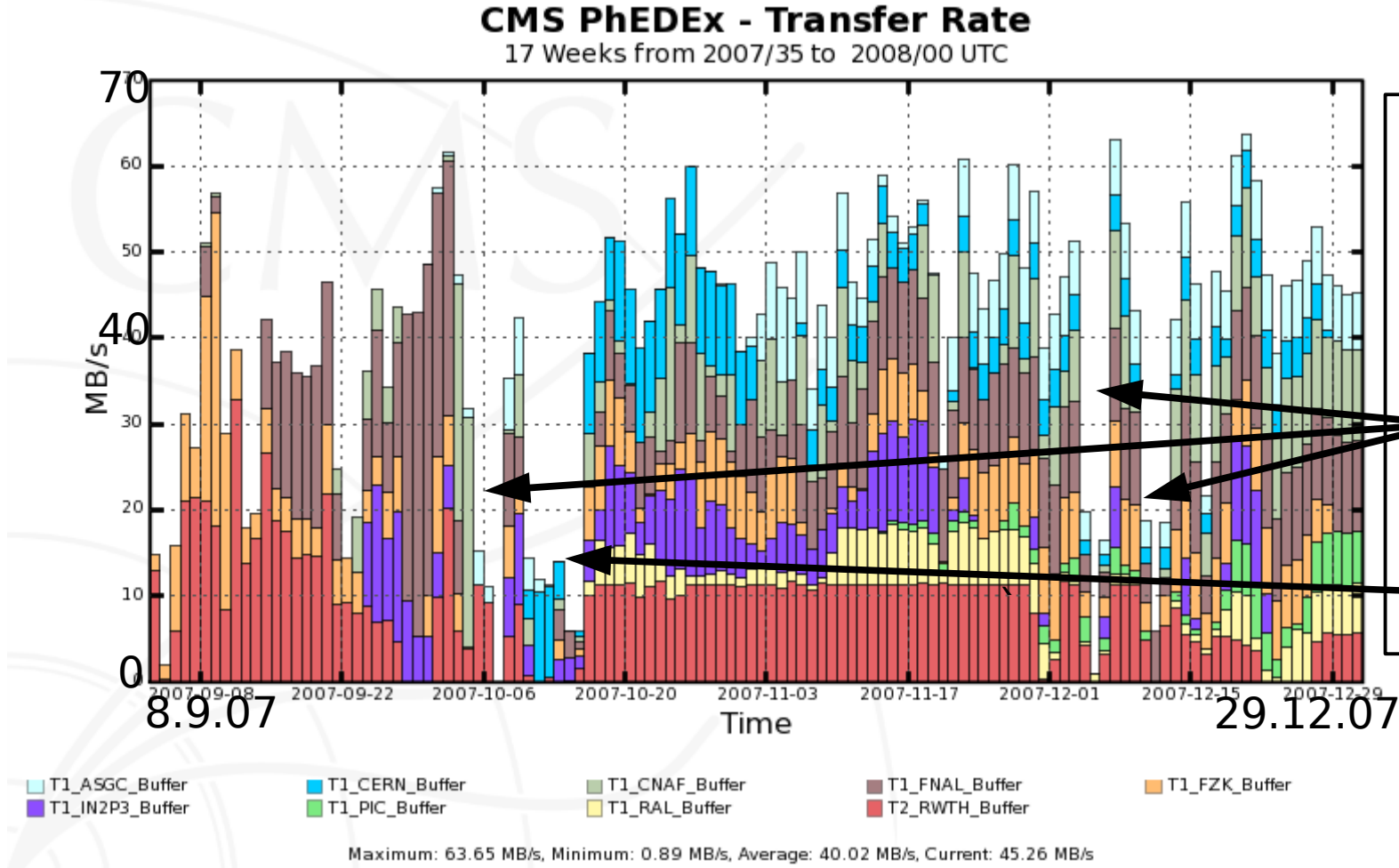
CSA07

- CMS **C**omputing, **S**oftware and **A**nalysis Challenge 20**07**
  - Test des Computing-Modells mit 50% der erwarteten Datenrate bei LHC Betrieb mit niedriger Luminosität
  - Beiträge der Tier-2-Zentren
    - Datentransfertests (*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 Datenanalyse)

***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 ≈ 3.5 MB/s)

Link Commissioning: Rate

Datentransfertest zum DESY, September-Dezember '07:

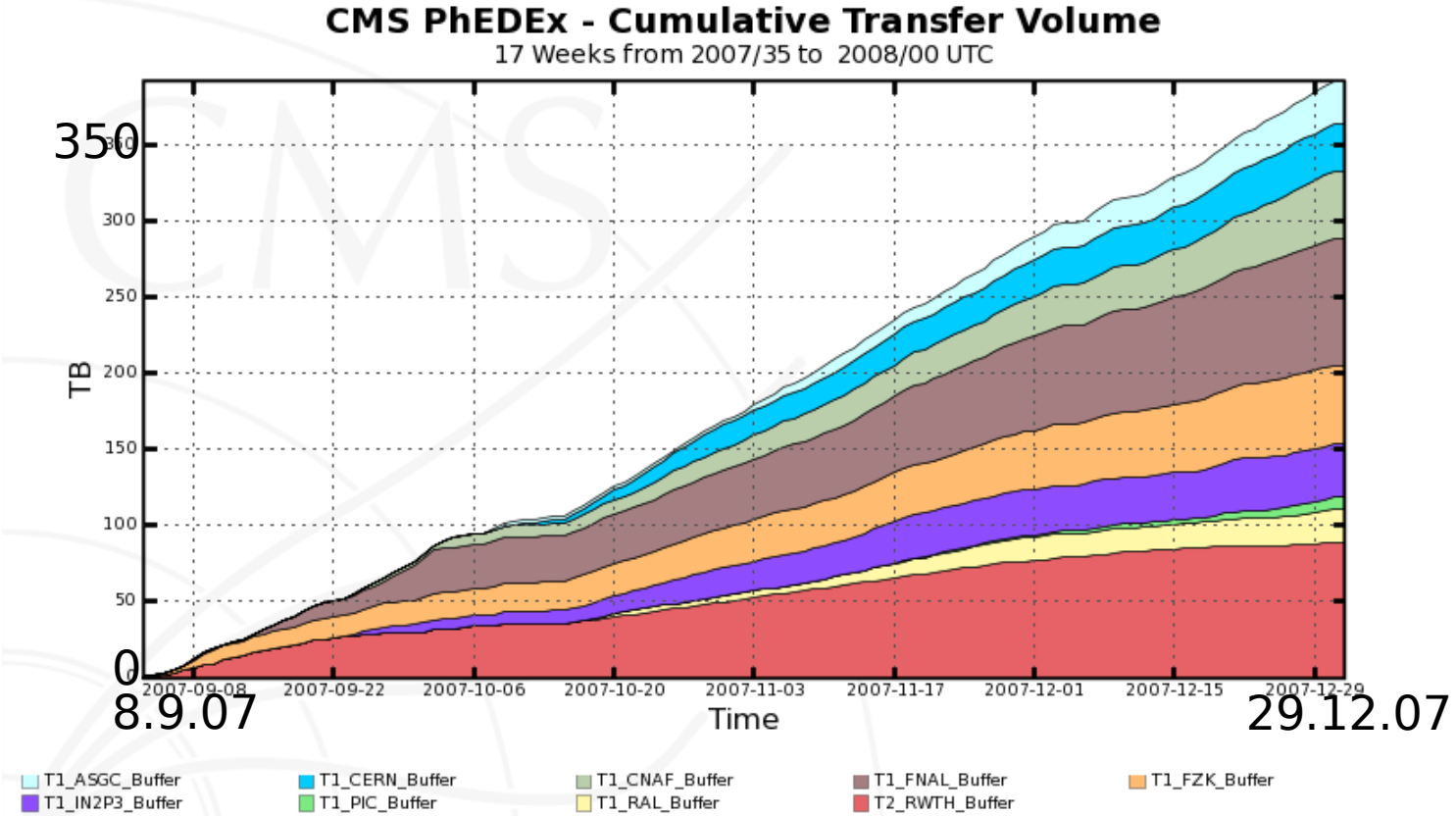


- Komplexes System mit vielen Fehlerquellen
- Längere Störungen durch
  - Probleme bei zentralen Diensten
  - Probleme am Massenspeicher

- 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:



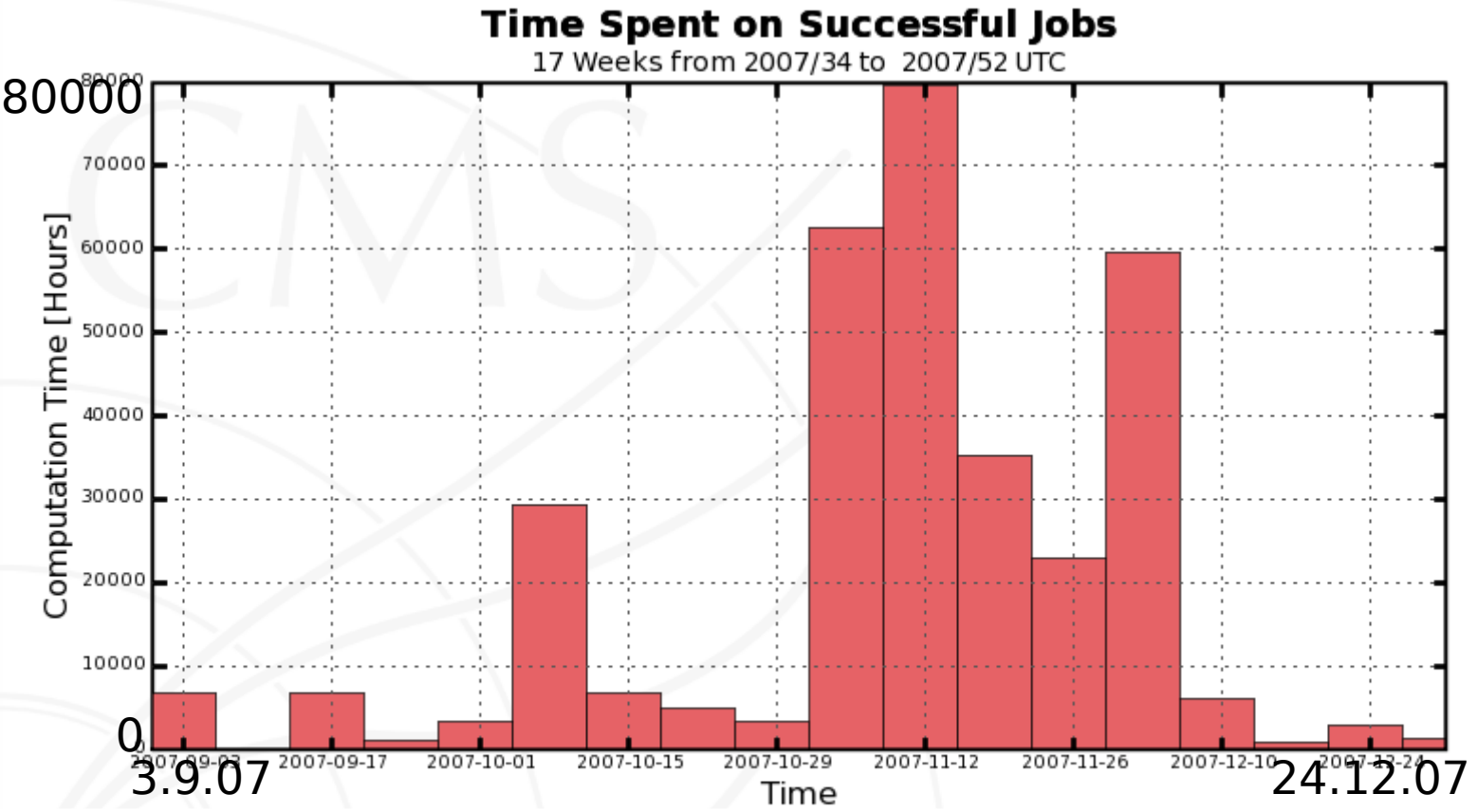
- Durchschnitt: 392 TB Daten in 17 Wochen ≈ 3.2 TB/Tag ≈ 0.3 Gb/s
- Computing-Modell Annahme: ~ 5 TB/Tag → Bleibt zu demonstrieren, ABER
  - Zusätzliche Transfers von „echten“ Datensätzen
  - Netzwerkverbindung und Massenspeicher-System noch nicht ausgelastet

  → genügend Steigerungspotential vorhanden

Monte-Carlo-Produktion I

Monte-Carlo-Produktion, September-Dezember '07:

Verbrauchte Rechenzeit (erfolgreiche Jobs)

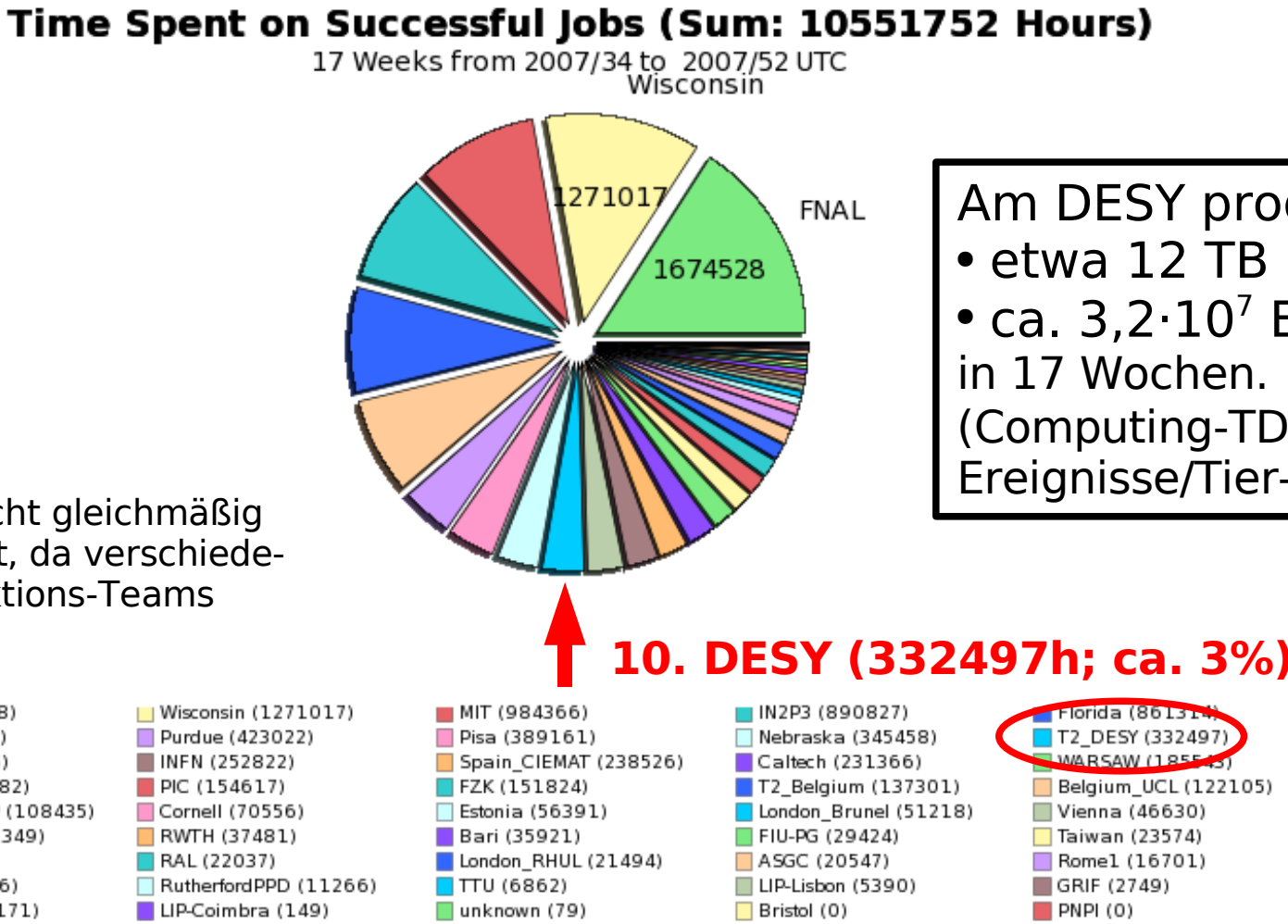


- Durchschnitt: 19558h/Woche ≈ 115 Prozessorkerne ständig im Einsatz
- Maximal 80000h/Woche ≈ 476 CPUs
- Großer Teil der CPU-Zeit wird für Monte Carlo-Produktion genutzt

Monte-Carlo-Produktion II

Monte-Carlo-Produktion, September-Dezember '07:

Anteile der verschiedenen Zentren



Zentren nicht gleichmäßig ausgelastet, da verschiedenen Produktions-Teams zugeteilt

Am DESY produziert:
- etwa 12 TB MC-Daten
- ca. $3{,}2 \cdot 10^7$ Ereignisse

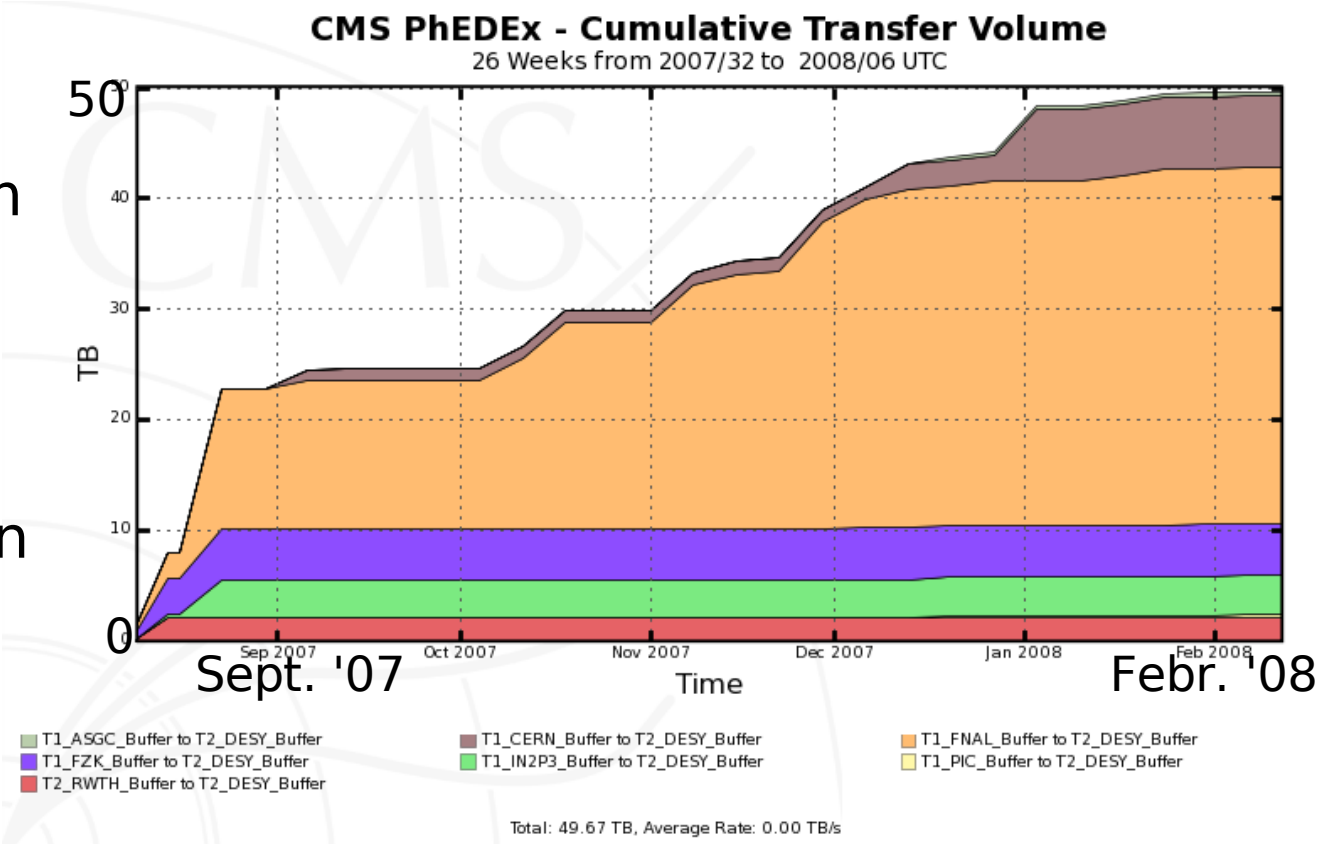
in 17 Wochen.
(Computing-TDR: ~10^8 Ereignisse/Tier-2/Jahr)

Speicherung von Datensätzen

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

DESY beheimatet zudem

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



Sept. '07 Febr. '08

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 Mai stellt neue Herausforderungen an Computing-Infrastruktur (100% Test)
- CSA08 finaler Test vor den ersten Daten?!