# CMS Computing

## "International ICFA Workshop on HEP Networking, Grid and Digital Divide Issues for Global e-Science"

## Daegu, Korea, May 23 - 27, 2005


LATBauerdick/Fermilab

U.S. CMS Software and Computing Manager

CMS Computing Coordinator

# CMS Computing Status and Drivers

- CMS Data Model, Computing Model, Computing TDR
- CMS produced about 90M events in the last year or so using LCG2, GRID3/OSG and "local" computing resources
  - Even quite complex computing production (such as digitization with pile-up) is being run on Grid3 and LCG now
  - simulated data are being served to CMS physicists now for analysis.
    - The data is being analyzed where it is located, CRAB analysis job Grid system

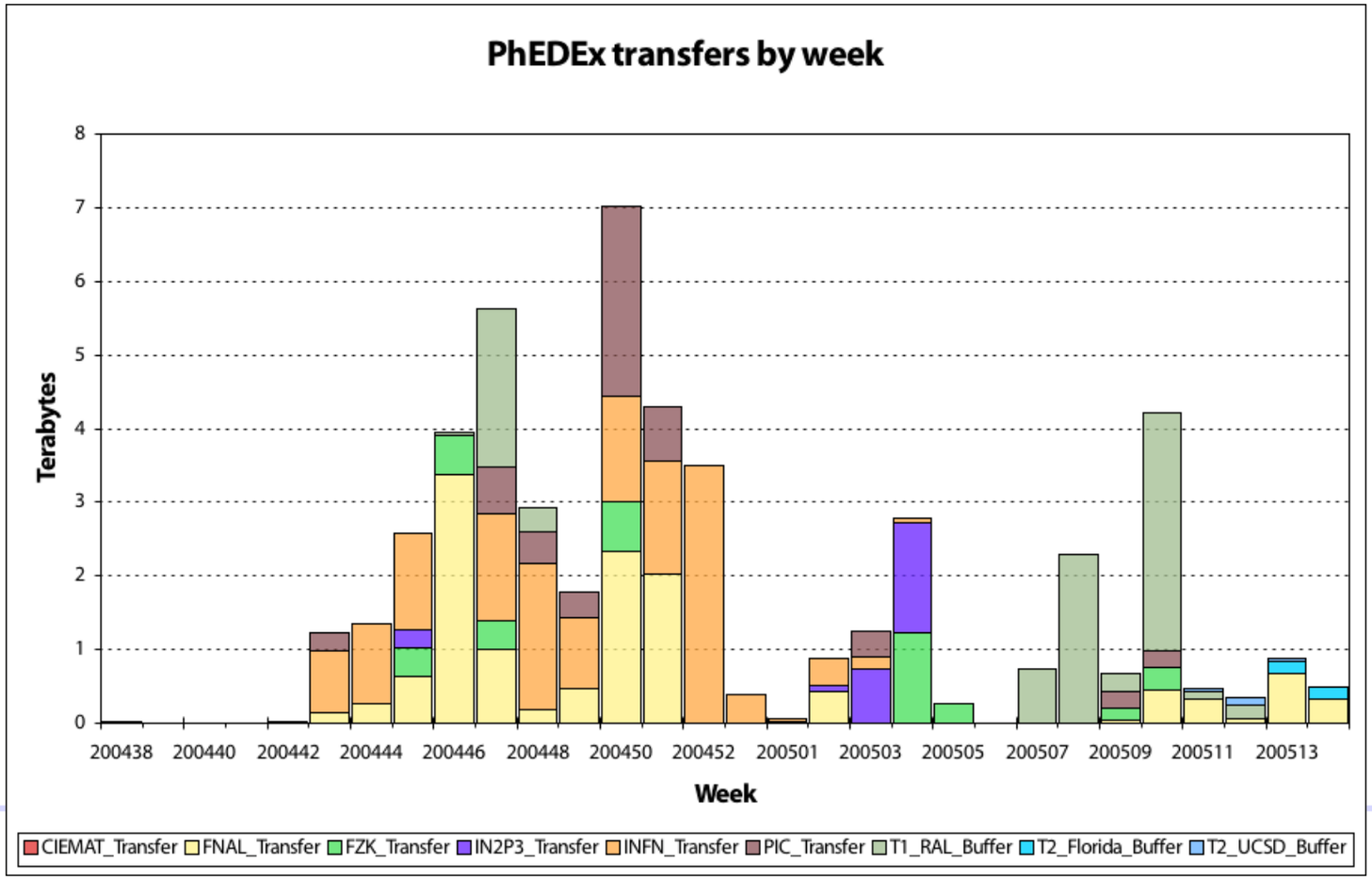
- next goals: implement computing baseline components, framework, and integrate application services with Grid environment
  - analysis data processing in the distributed environment
  - placement of datasets, configuring and running jobs, serving data, managing & monitoring the system, implementing collaboration policies & priorities, secure, open & friendly environment

# CMS Data Transfers Operations and Testing

CMS data management

## Current operational data

- Production: ~70 TB known, ~150 TB total replicated
- SC2: 1.6 PB — 1.6M replicas of 40 files (!)
- Test instances: 2 x testbed, integration test, castor test



# Elements of Computing Baseline

- Structured Data, Structured Grids

- Data Granularity
  - LHC Triggers cut deeply into physics;
    - Data always needs to be considered in its trigger context
      - Split annual O(2PB) raw data into O(50) (40TB) trigger determined datasets

- Data Tiers
  - RAW, Reconstructed, Analysis, Tag
    - Keep Raw and Reconstructed close together (initially at least)
      - Custodial RAW+Reco distributed over Tier-1s (one copy somewhere)
    - Analysis Data, Full copy at each Tier-1, partial copies at many Tier-2
- Computing Tiers
  - CMS-Tier0: Close connection to Online, highly organized,
    - Tier-1: Data Custody, Selection, Data Distribution, (Analysis), Re-Reco
      - Tier-2: Analysis Data under Physicist "control", MC production

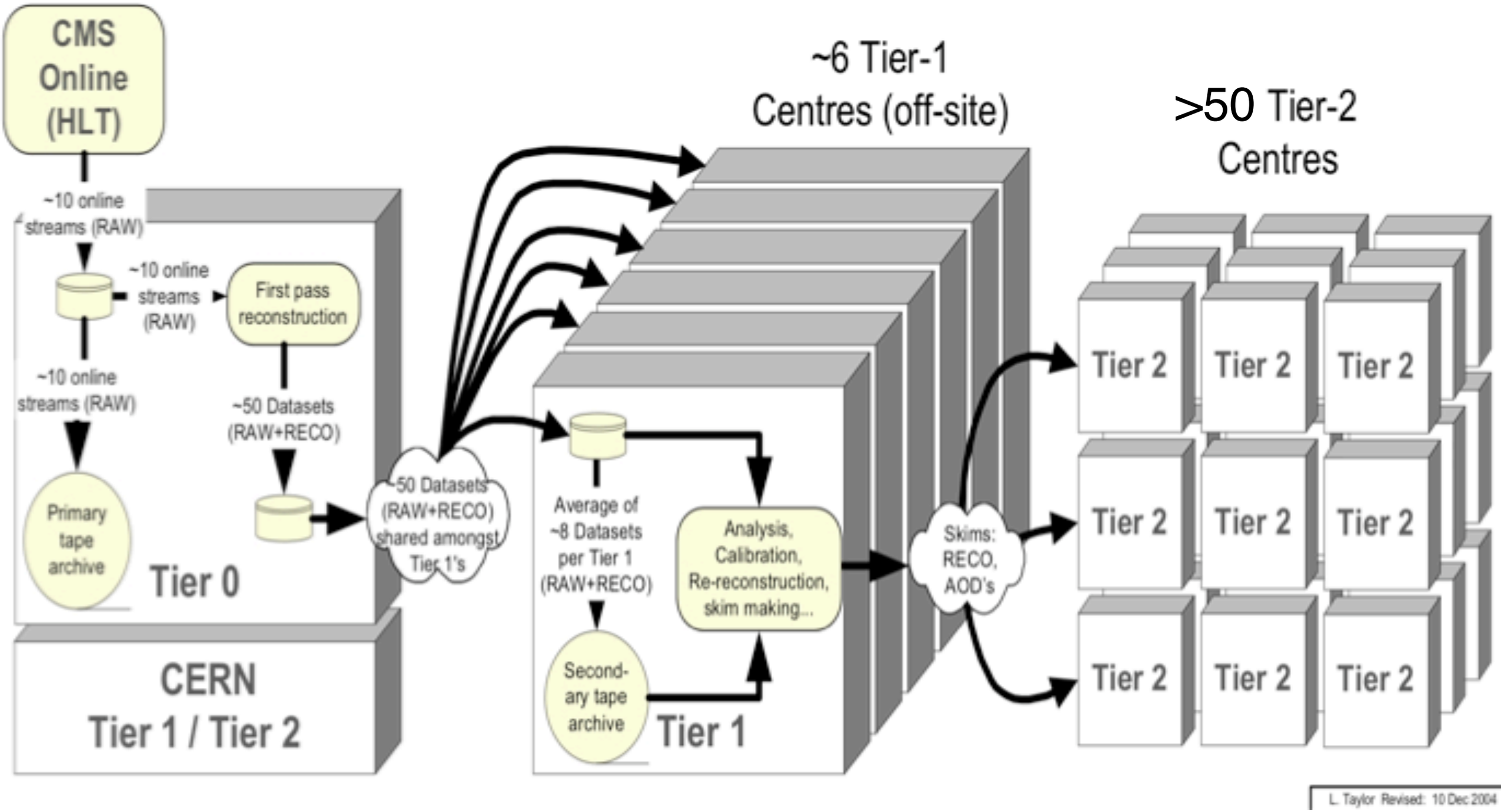
# Roles of Tier-0/1/2 Centers

- Roles ~well defined through task of Tier-0, Tier-1, Tier-2 in processing and managing of structured CMS data
- bulk processing of datasets at Tier-0/1
  - Tier-1s act together across CMS to keep key datasets available
    - managing datasets: bookkeeping, transferring, hosting, accessing, producing, ....
- users and physics groups extract "analysis datasets" to Tier-2s
  - Tier-2s working with local communities and also cross-CMS groupings
  - WM tools interacting with DM services for physics analysis jobs

- require a strong physics organization taking responsibilities
  - understanding of (set of ~ orthogonal) datasets based on trigger paths
  - understanding of all the relevant workflows
  - ability to prioritize and assign resources to physics/detector topics
    - through assigning datasets and "local users" to Tier-2 sites

# Data Flows through the Tiers

- CMS CM defines roles for Tier-1, Tier-2 and requirements



# CM — Tier-1 Resources

## Tier-1 Specifications (Per Tier-1, of 6+1 in total)

LHCC Review of Computing Resources for the LHC experiments

| | | | *Eff Factors* |
|---|---|---|---|
| CPU scheduled | 1199 | kSI2K | *85.00%* |
| CPU analysis | 929 | kSI2K | *75.00%* |
| Disk | 1121 | Tbytes | *70.00%* |
| Active tape | 1837 | Tbytes | *100.00%* |
| Data Serving I/O Rate | 800 | MB/s | |

**WAN**

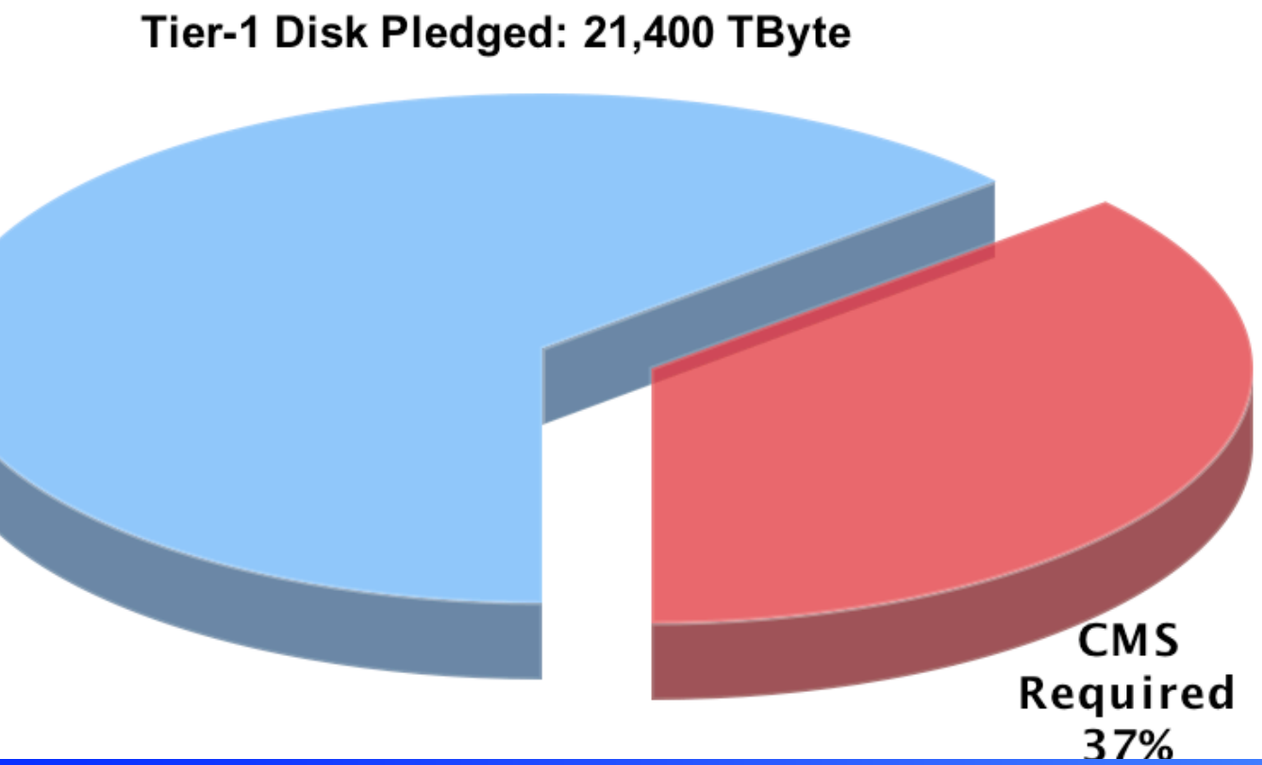
| | Raw Rates | Safety | Headroom | Totals |
|---|---|---|---|---|
| | Gb/s | Factor | Factor | Gb/s |
| MC Simu/and Reco from Tier2 | 0.1 | 2 | 2 | 0.5 |
| FEVT/AOD from Tier 0 | 0.7 | 2 | 2 | 2.2 |
| AOD Versions from ReReco | 1.0 | 2 | 2 | 3.0 |
| Total Incoming | | | | **5.7** |
| Event Serving to Tier-2s | 0.9 | 2 | 2 | 3.5 |
| Total Outgoing | | | | **3.5** |

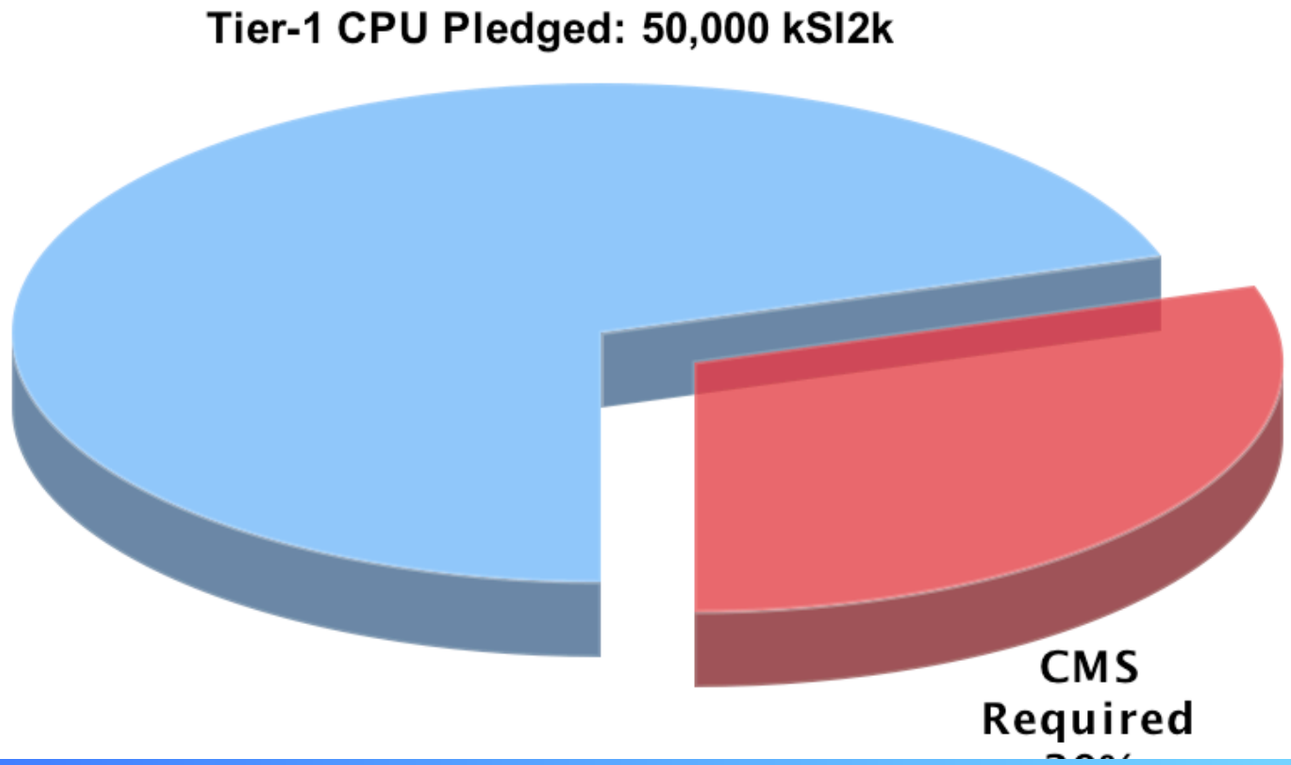
David Stickland

Jan 2005

# Tier-1 Resources for 2008

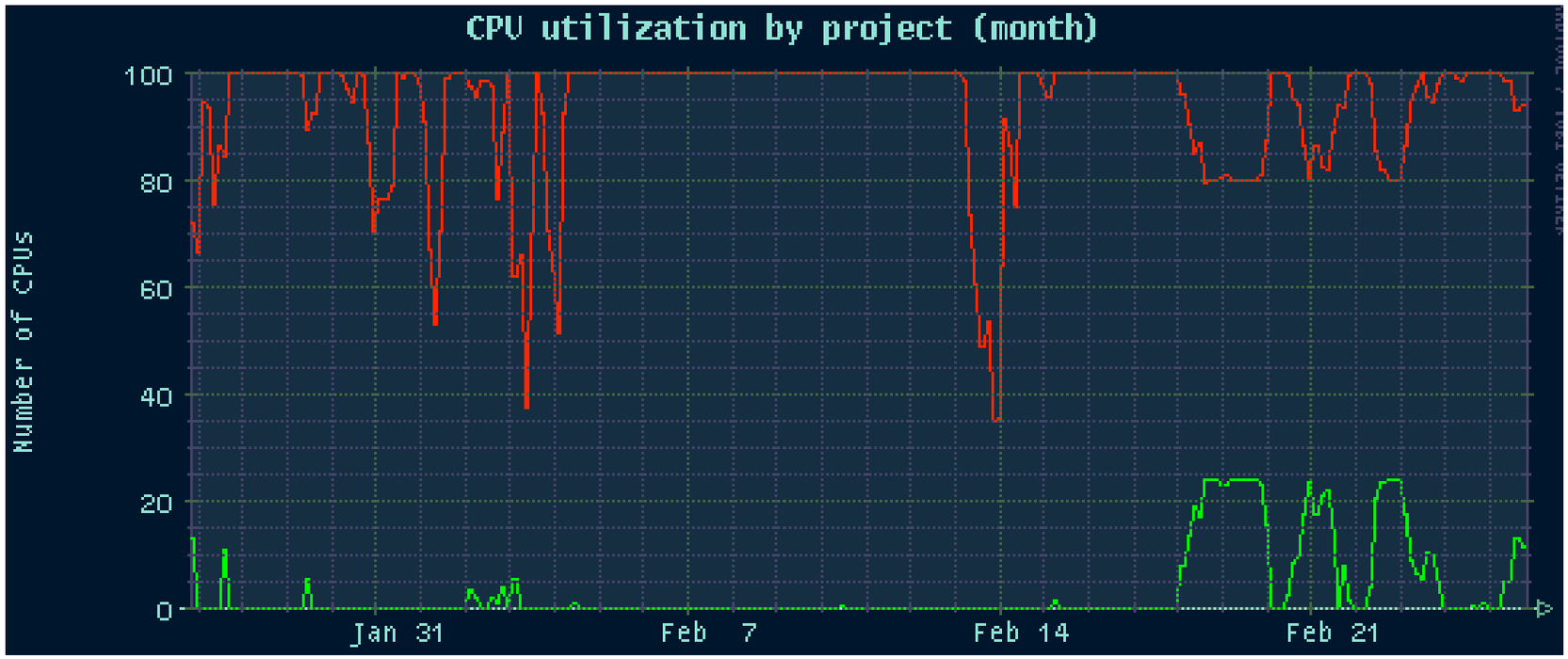
- CMS requirements from CM ~ commensurate with total resources
  - however, resources "pledged to CMS" would have 20-30% shortfalls
- Tier-1 resource shortfalls would have serious effect on physics
- >50 Tier-2 sites signed up for CMS — vital importance for physics!
  - expect a strong and important Tier-2 program, giving opportunity for countries that do not contribute to Tier-1 — 40% of CMS authors!



# Need to Support Analysis Computing

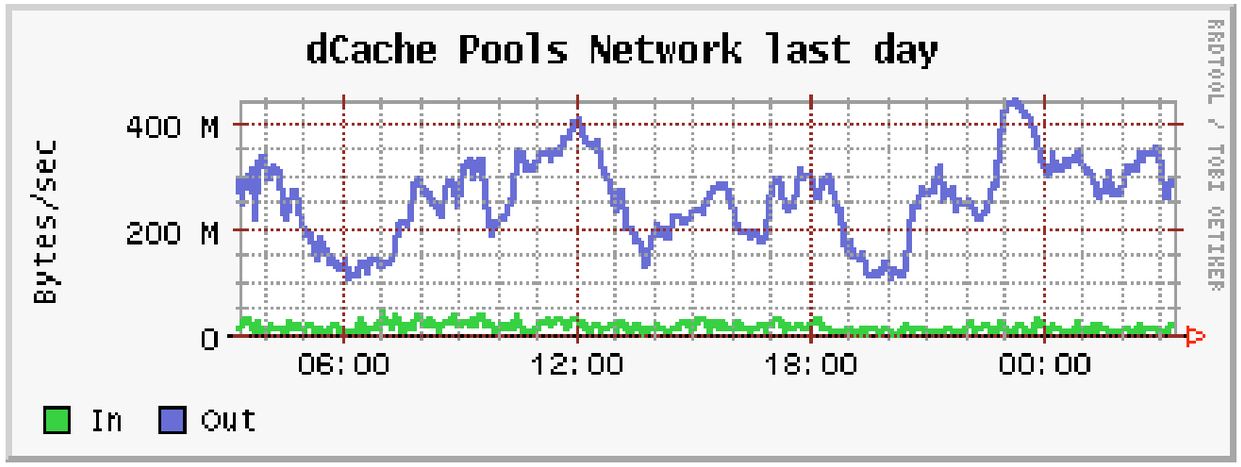
- Need major facilities at CERN (CMS-CAF) and elsewhere
- Fermilab: successful User Analysis Facility for LHC Physics Center users
  - Over 200 registered users, about 20 are active at any given time
  - 56 systems, about 100 batch slots -> extending to 300



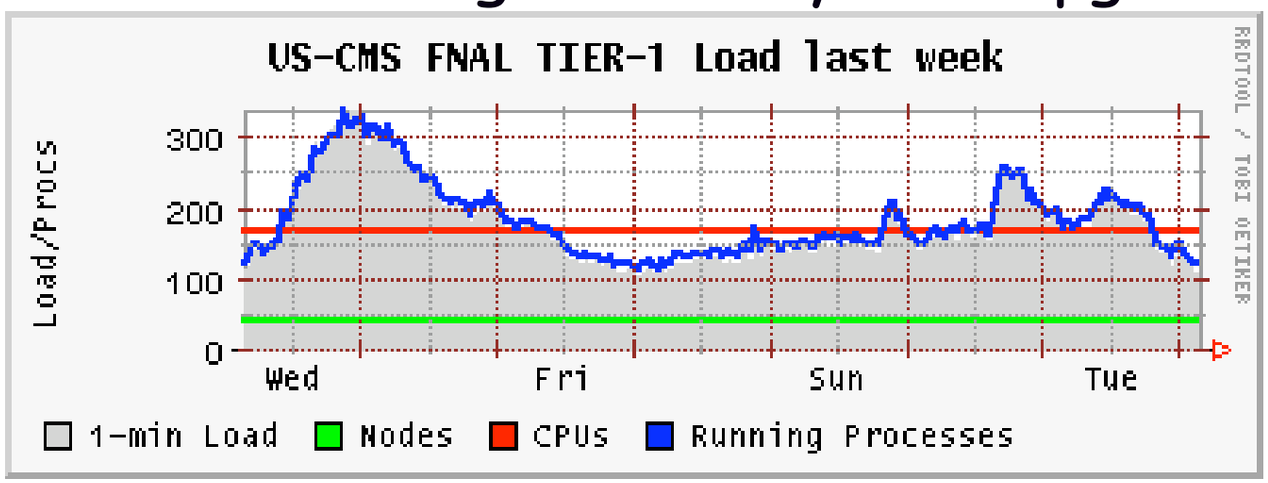

- CMS Software installed, datasets served, user scratch space
  - Good examples are published, CMS101 Tutorial has been offered twice, responsive support from the facility staff

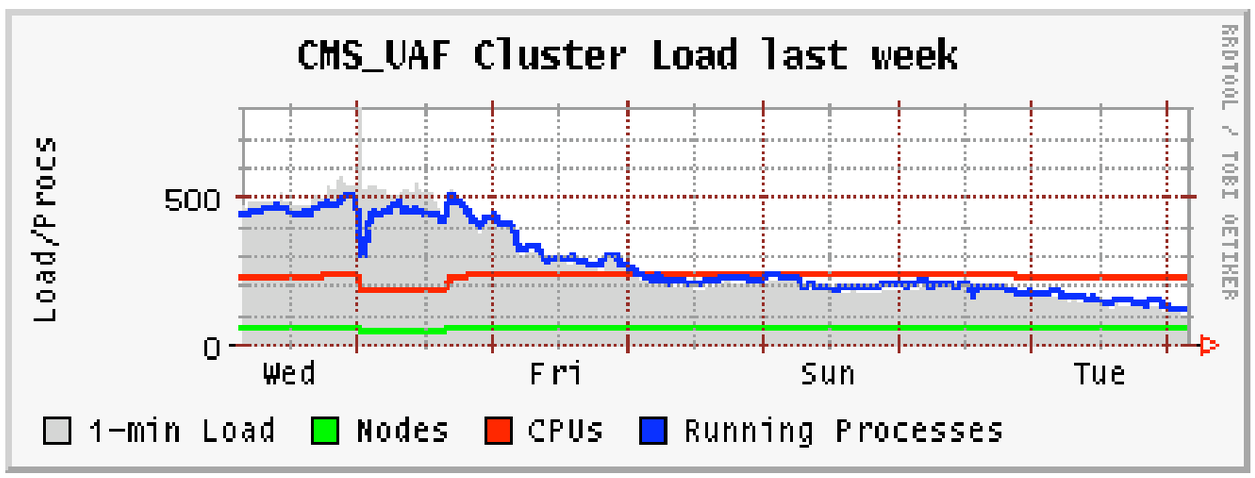
# example at Fermilab: Facilities Heavily Used

- Access to 250 CMS Datasets through DCache
  - Typically have a few thousand open file handles,
  - Processed nearly 4000 file restores over the weekend




- Production Nodes and Analysis Nodes are both heavily subscribed
  - Need to get facility CPU upgrade done ASAP



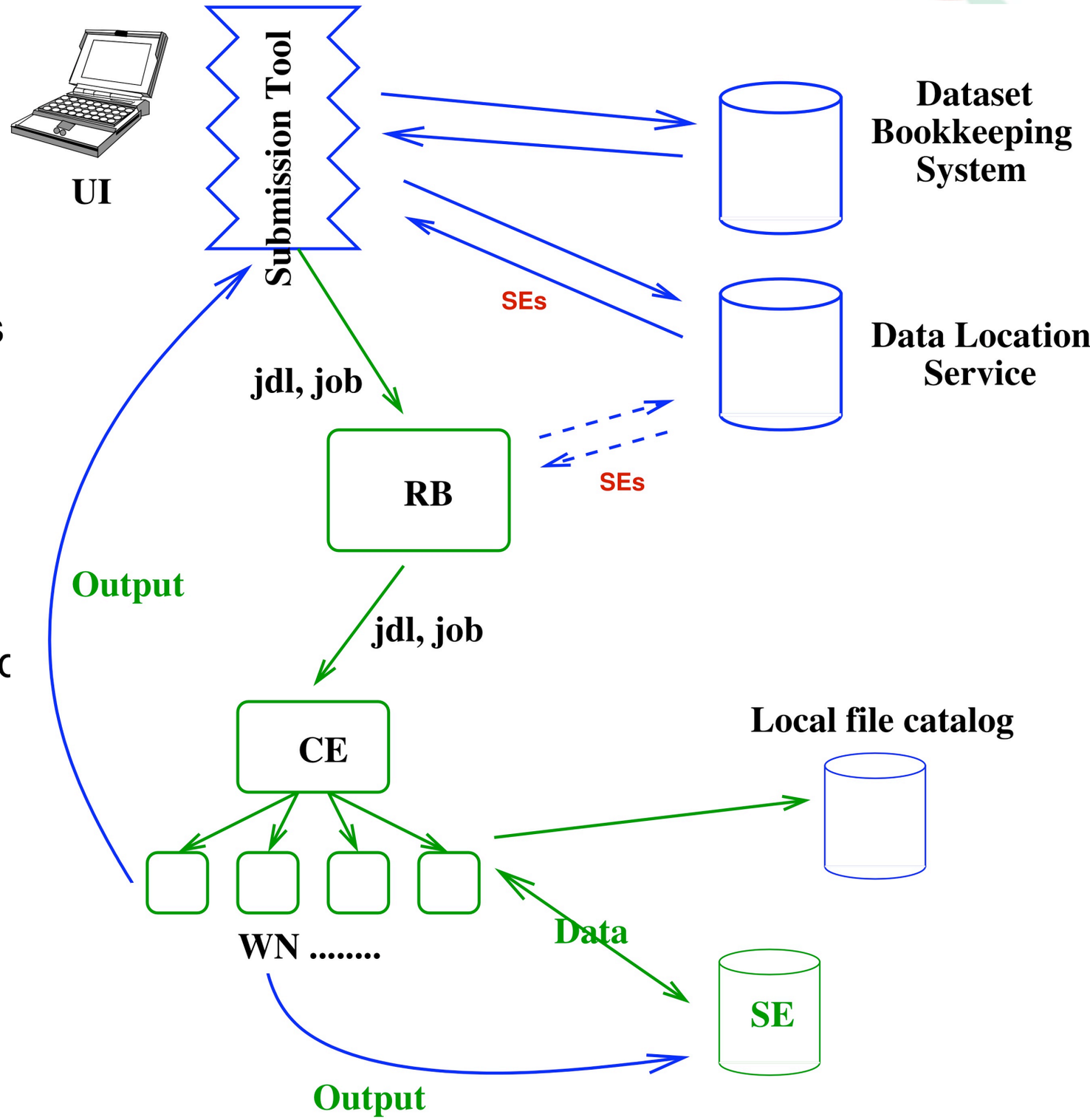



# Drivers for Technical Work

- Main driver: get data to the T1/T2 sites and make it accessible
  - then help users to run analysis on it -- including PROOF etc
- focus areas
  - storage and data access
  - configuration, software setup, stability of operation
  - job scheduling, scalability, etc in particular for MC production
- need to instrument and control the environment
  - operations scenarios, providing integrated view on WLCG
  - have instrumentation technologies like MonaLisa, GridICE etc
    - need user interface "dashboard", and some interaction with user tools
  - allow to exert some global control of environment
    - expressing & enforcing policies, quota & access rights, accounting & bookkeeping, etc are immature or non-existing in the Grid environment

# CMS Computing Baseline Components

- DataSet Bookkeeping System
  - What data exists
- Data Location Service
  - Where is it
- File Placement and Transfer Service
  - PhEDEx, reliable transfer services
- Local File catalogs
  - Site specific
- Data Access and Storage Systems
  - SRM on SE
  - Access POSIX-like
  - CMS tools to allow CMS policy anc
- Monitoring and job-tracking
  - Mona-Lisa, GridIce, Boss
  - Information for priority changing
- Workflow support
  - Configuration control
- Job preparation
  - CRAB (or son of CRAB)..
- etc



# CPT-Computing Integration Activity

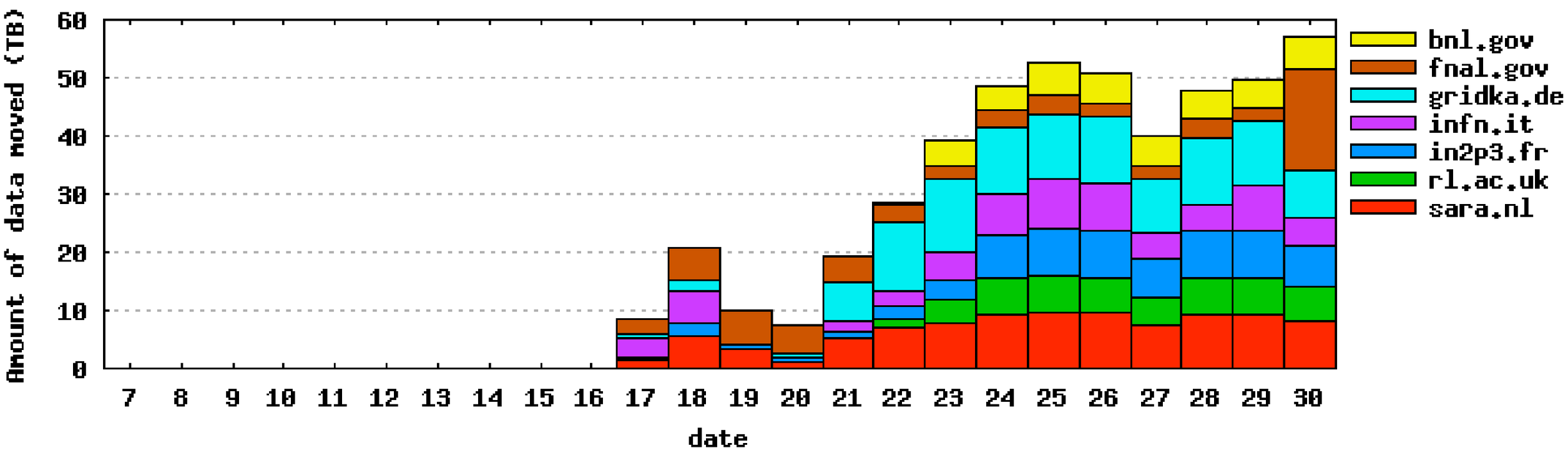
- activity coordinated between the computing tasks / subprojects
  - and some of the software tasks -- to be discussed and defined
- goal to evolve to the overall baseline **system** -- CMS and LHC Grid
  - integration will guide the development of components and workflows
    - understand, define, implement the fundamental workflows
    - understand and develop interfaces of components
  - reference is computing baseline, changes are "controlled" by project
  - prepare for getting system ready -- DC06 and beyond
  - provide CMS participation in LCG integration (SCs)
- involve broad fraction of CMS, including regional centers
  - developers respond to needs "uncovered", incl. spawning new projects
- concrete milestones with specific metrics, about every 3 months
  - functionality milestones to develop system
  - integration milestones to deploy systems
  - release to production environment LCG/OSG

# CMS Computing Integration Milestones (very preliminary)

- 2005-06 CIM-1 "initial integration of baseline components"
- 2005-09 CIM-2 "computing systems ready for SC3"
- 2005-12 CIM-3 "computing systems ready for cosmic challenge"
- 2006-03 CIM-4 "computing systems ready for SC4"
- 2006-06 CIM-5 "computing systems ready for DC06"
- 2006-09 CIM-6 "lessons learned DC-06"
- 2006-12 CIM-7 "integration of Tier-1 and Tier-2 regional centers"
- 2007-03 CIM-8 "CMS-Tier-0 and CMS-CAF ready for pilot run"
- 2007-06 CIM-9 "computing systems ready for pilot run, start of pilot run"

# Service Challenges

- Done: SC1, SC2 — "Robust File Transfers" and "Throughput"
  - develop and integrate the basic components and services
    - reliable and automated high-bandwidth data transfers CERN —> Tier-1 centers
    - 5 Tier-1 centers together 50TB/day, 500MB/sec throughput milestone achieved
    - basic services being debugged and made robust, then moved into production use




- Next SC3, SC4 — Experiment end-to-end challenge

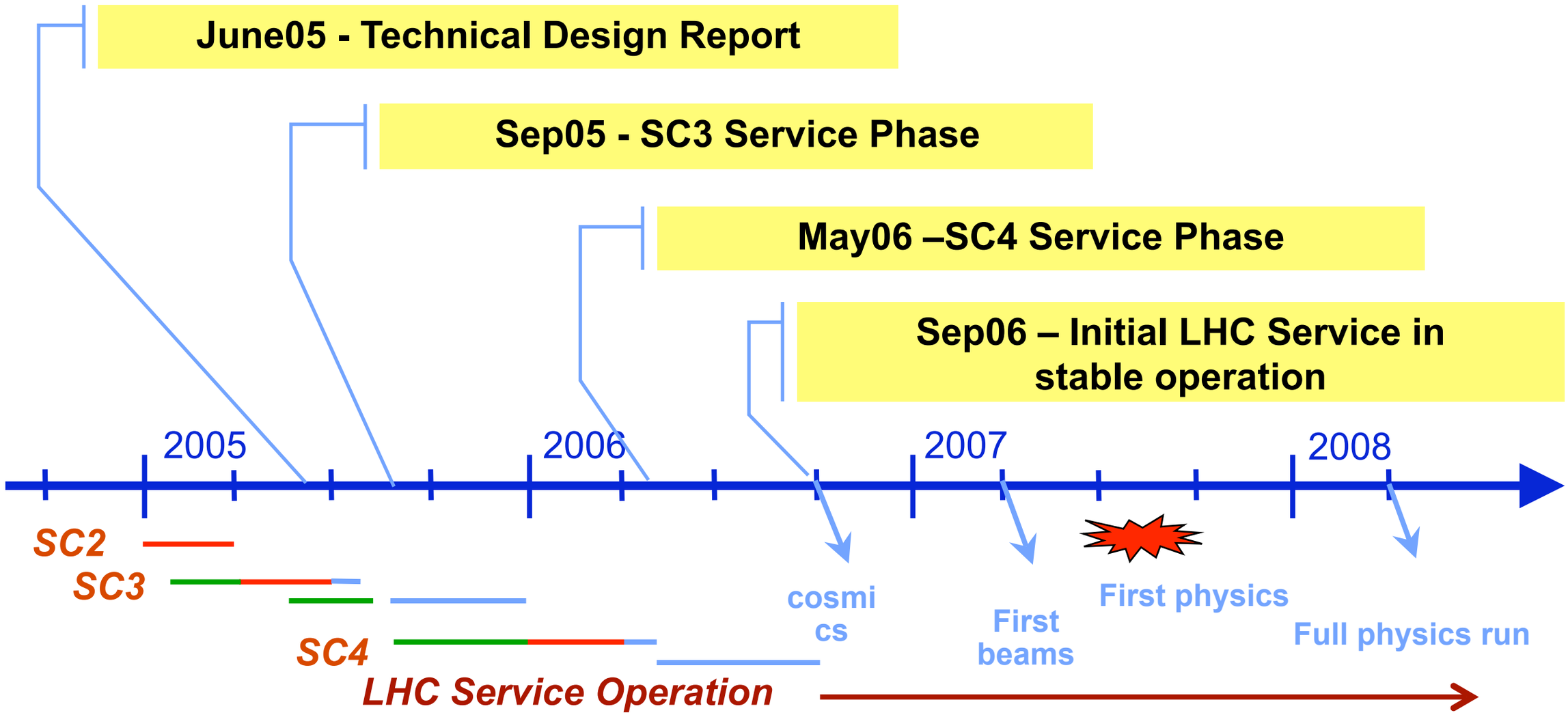
# CMS Service Challenge 3 Scenario

- Exercise roughly realistic scenario, but small amount of data
  - Data produced centrally and distributed to Tier 1 centers (MSS)
  - Strip jobs at Tier 1 produce analysis datasets
    - Approximately 1/10th of original data, also stored in MSS
  - Analysis datasets shipped to Tier 2s for analysis, published locally
  - Tier 2 sites produce MC data, ship to Tier 1 MSS
    - May not be the local Tier 1
- Transfers between Tier 1 sites
  - Analysis datasets, 2nd replica of raw
- Volumes + sites
  - 50 TB from CERN T1 to at least two T1s, plus smaller ones (~10 TB)
  - 5 TB to T2s, at least one per T1
  - 5-10 TB T1/T1 analysis dataset transfers
  - 50 TB T1/T1 2nd raw replica transfers

# Services for SC3

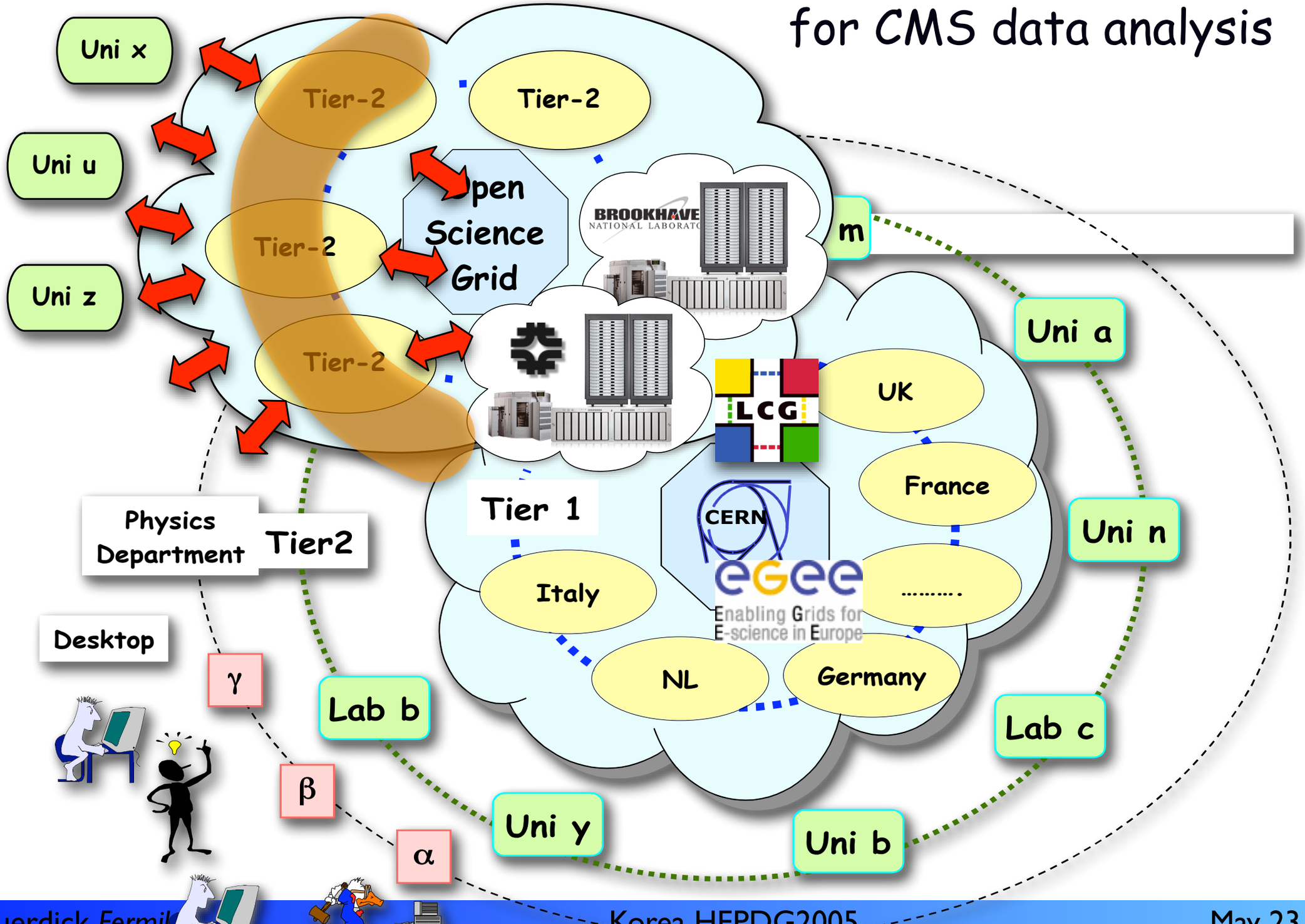
- Data transfers for dataset placements
  - PhEDEx; to be installed at each participating site
  - Underlying transfers expected to be either SRM or globus-url-copy
- File catalogue
  - POOL API, local, relational; MySQL default; LFC, Globus RLS option
- General environment
  - Data serving infrastructure -- dCache, Castor or xrootd
  - Computing element, job submission (UI)
  - Output harvesting for transfers (CMS agents, use UI-type machine)
  - CMS software installation and publishing into the information system
  - Bookkeeping / monitoring databases for production
  - Above-mentioned file catalogue
  - PubDB or successor

# Key dates for Services



# Tier-2 centers University computing

- Tier-2 interface Universities and Tier-1s/Tier-0 for CMS data analysis



LCG

> 100 sites

Open Science Grid

> 40 sites

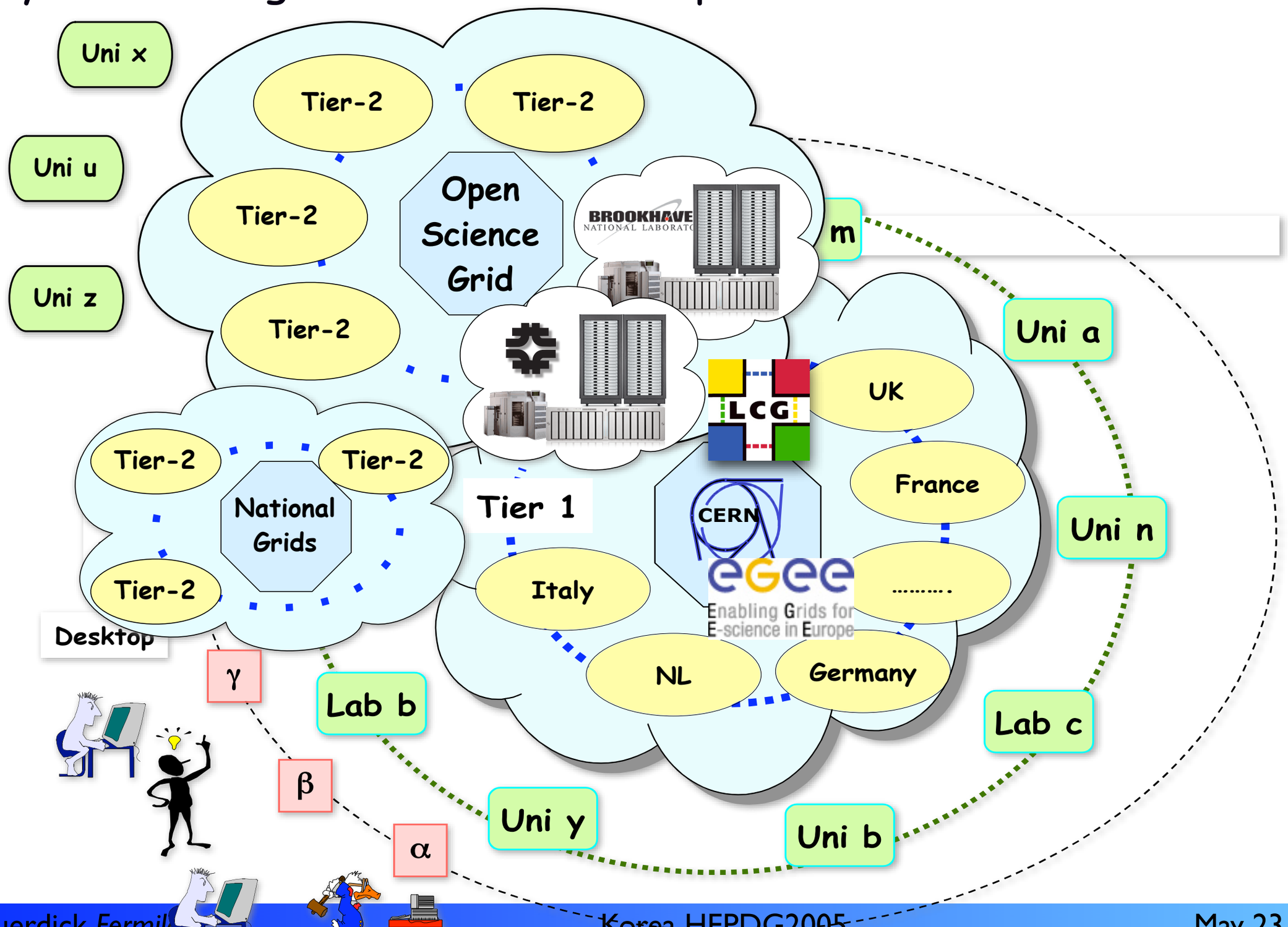
**Also NorduGrid**

# US CMS Tier-2s in the Open Science Grid

- Focussed CMS Tier-2s within broad OSG Infrastructure
- additional development effort for CMS integration (DISUN)

# Korea to Participate!

- By federating its resources and partner with LHC Grids



# Communities of Scientists Working Together...