A multi-purpose computing center: FNAL

Stephen Wolbers
Fermilab/CD/Scientific Facilities & CDF

Workshop on Data Preservation and Long Term Analysis in HEP

DESY, January 26-28, 2009

Outline

- Description of Fermilab Computing Center(s)
  - Facilities
  - Services
  - User communities
- Scientific Data
  - Components and architecture
  - Scientific Data Storage
  - Data Center Responsibilities
- Open Issues

Scope

- The scope and focus of this talk is the Fermilab Computing Facility.
- Not covered, or at least not covered well, are issues such as:
  - Data GRIDs
  - Cloud computing and storage
  - Commercial storage (Amazon S3, Google, others)
  - Data integrity, authorization, access rates
  - Novel data storage technologies
    - Holographic storage
    - Solid state disks
    - Disk farms as archival storage

Description of Fermilab Computing Center(s) - Facilities

- Three buildings, ~11 rooms, UPS, generator
  - Does not include accelerator, detector, some small computing rooms elsewhere on site.

GCC

LCC

FCC

Facilities: Feynman Computing Center

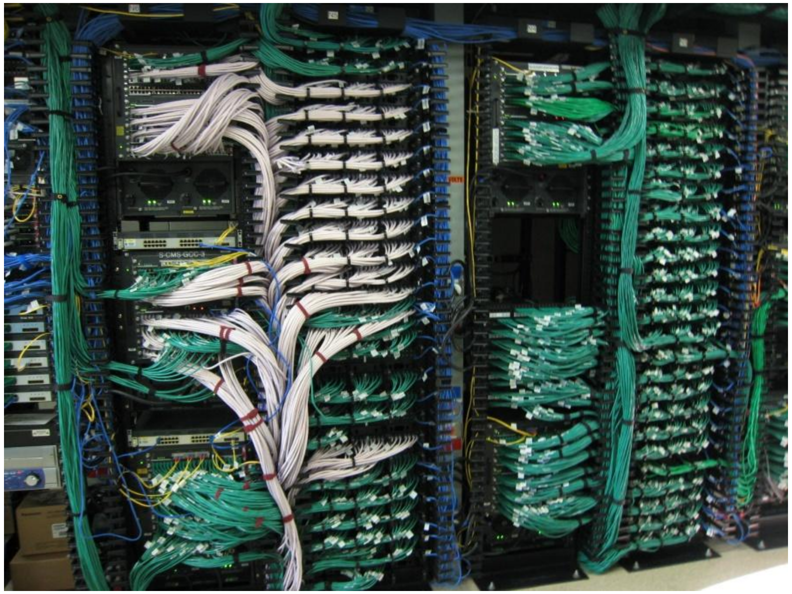
- Feynman Computing Center (FCC)
  - 3 rooms
  - 19000 ft^2 (1765 m^2)
  - ~700 kW UPS
  - 1.5 MW Diesel Generator
  - Uptime 99.99% (Tier 1 Data Center)
- FCC Home for:
  - Servers that require high availability
  - Part of the robotic storage (9 data robots + 1 central file backup system)
  - Web, email, backup, business systems
  - Networking
  - Many miscellaneous servers and computing systems.

Facilities: Feynman Computing Center

Facilities: Grid Computing Center

- **Grid Computing Center (GCC)**
  - 6 rooms
  - 10388 ft^2 (965 m^2)
  - ~1 MW UPS (constantly growing)
  - Generators can be connected for planned outages.
  - Uptime 99.78%. No permanent generator backup in place at this time.
- **GCC Home for:**
  - Robotic Storage (3 data robots)
  - Compute Farms (Run 2, CMS, FermiGrid)
  - Lattice QCD and high performance computing
  - Volatile disk cache
  - Networking

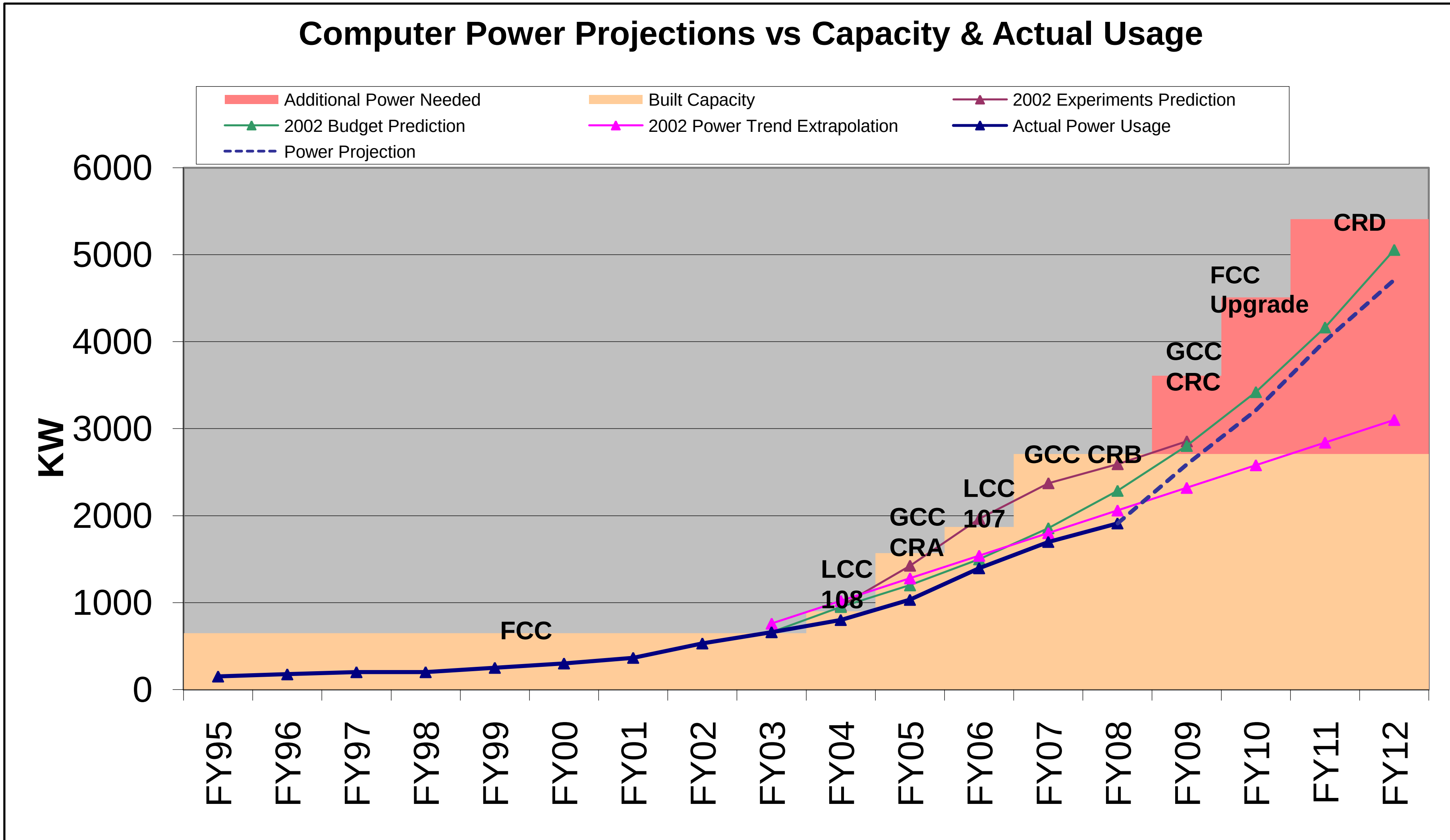
Facilities: Grid Computing Center



Facilities : LCC, the Future

- Lattice Computing Center (LCC)
  - 2 rooms
  - 4500 ft^2 (418 m^2)
  - 500 kW
  - No UPS, no generator
  - First few generations of Lattice QCD PC clusters
- Future Facilities & Upgrades
  - Upgrade to FCC for high availability computing
  - Upgrades/additions to GCC
  - UPS/Generators for GCC
  - Additional power, space and cooling driven by the needs of the program and the ability to get funding
  - Constant need! Working hard to increase and improve the facilities.

Scientific Computing at Fermilab power use history and projections



Types of data and services housed in the facilities

- Business Services
  - Professionally managed, including off-site storage, disaster recovery, etc.
- Core Scientific and Laboratory Services
  - email
  - Web services
  - Networking
  - Home areas
  - file servers
  - CVS repositories
  - Windows servers

  Approximately 1 Pbyte
  1 year retention
  TiBS backup system
- Scientific Computing Facilities
  - The main thrust of this talk
- Databases

  14 Pbyte
  Managed by Enstore

User Communities
(not an exhaustive list)

- The Computing Division
- The Laboratory
- Run 2: CDF and D0
- CMS Tier 1 and LPC centers
- Neutrino Program
  - MINOS, miniBooNE, NOVA, Minerva
  - SciBooNE, other small neutrino efforts
- HEP experiments analyzing data (KTeV, Selex, Focus, …)
- Sloan Digital Sky Survey
- DES + other Astrophysics including JDEM
- Lattice QCD
- Accelerator Simulation
- Test Beam Experiments
- Cosmological Computing Simulations Initiative
- ILC
- Project X
- Mu2e
- Theory

Scientific Data

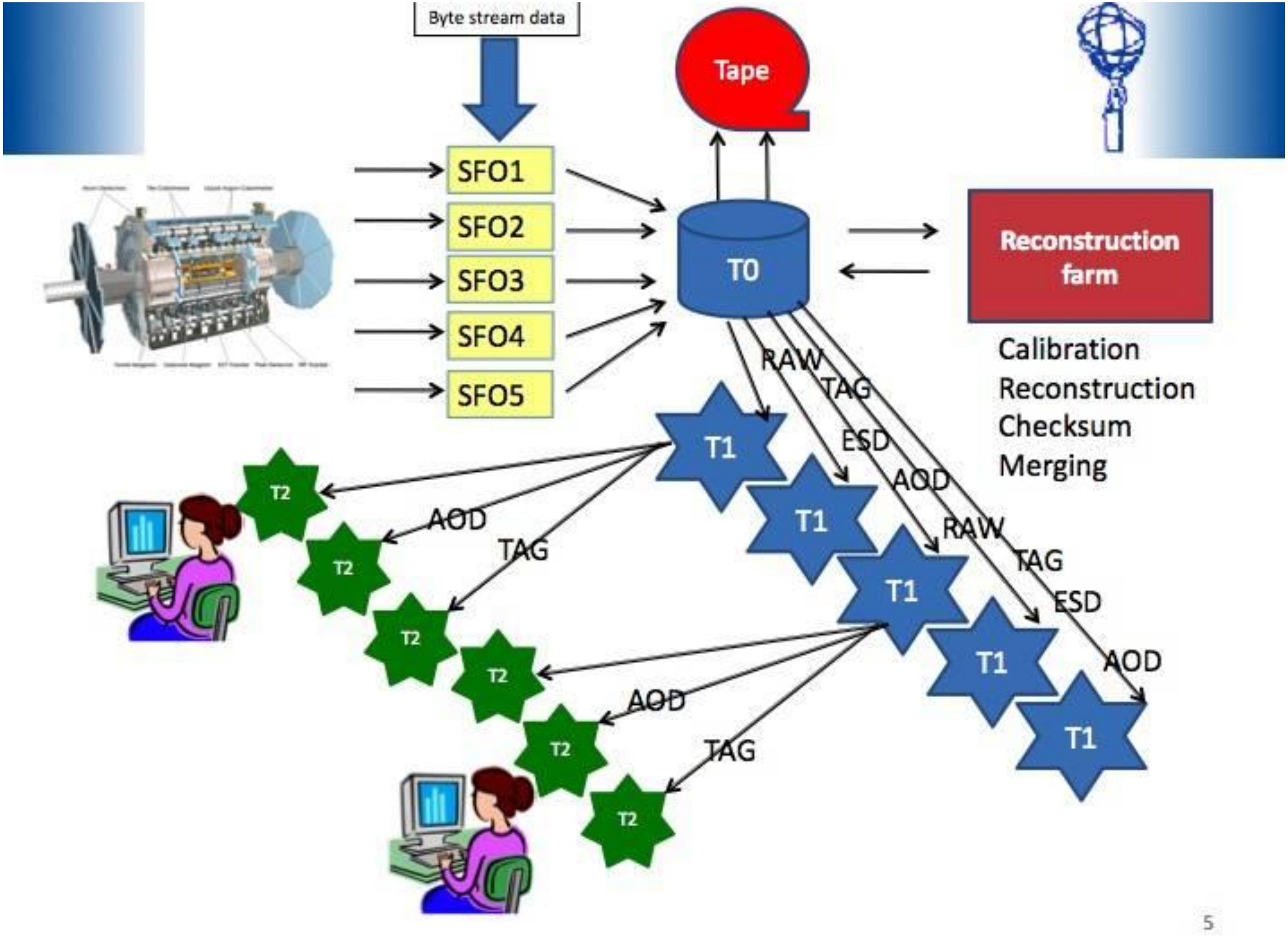
Scientific Data: Components

- Experimental data ("Raw")
- Derived datasets
  - Output of production
  - Selected datasets
  - "Ntuples"
- Simulated Data
- Metadata
- Databases
- Code
- Calibrations
- Other

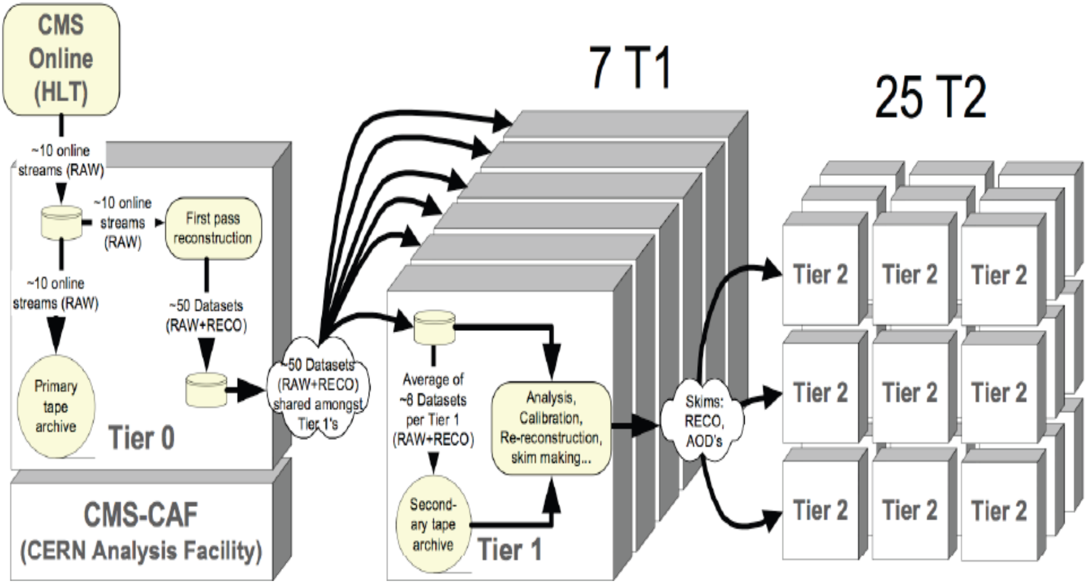
Scientific Data Architecture

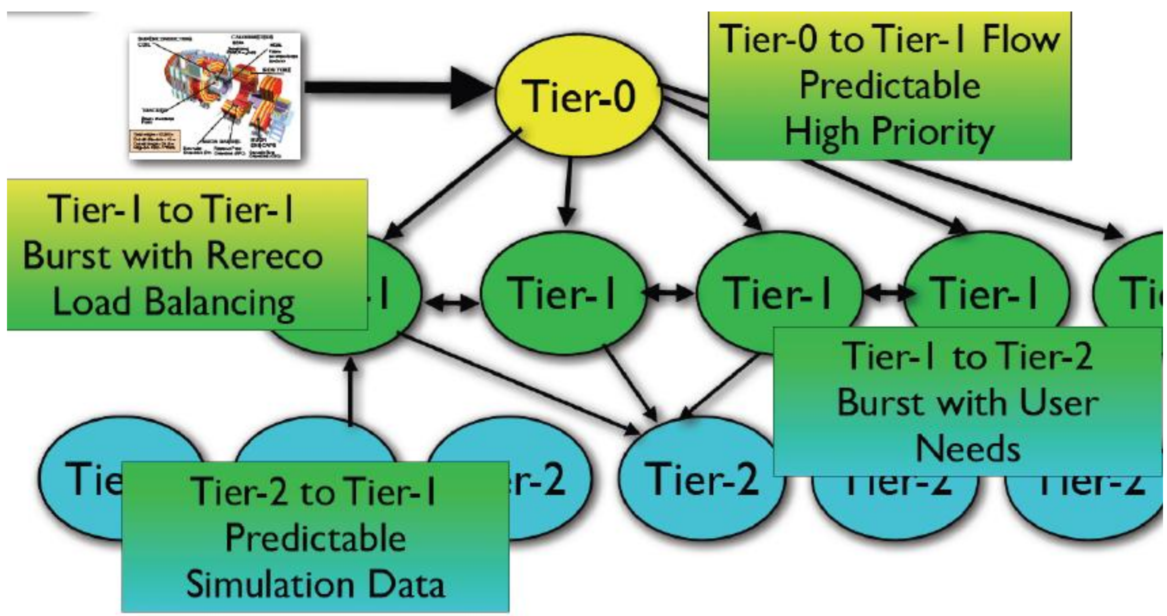
- The basic architecture:
  - Applications
  - Interfaces and middleware
  - Grid and data handling systems
  - Disk or other cache storage
  - Mass storage
  - Calibrations and databases
- All of this must function properly for the experiment to collect and analyze data
- Long term analysis likely uses a subset or even a highly modified architecture or model

ATLAS (borrowed from Jim Shank)



CMS Data Flow





Scientific Data: Archiving Evolution and examples

Scientific Data – I
pre-collider Run 2 (prior to 2000)

- Mix of fixed-target and collider experiments
- Huge amounts of data compared to the economical capability for robotic storage and management
- Emphasis on writing the maximum amount of data to any storage for offline analysis
  - Driven by the physics needs
  - Commodity tapedrives and media (8mm, DLT)
    - Large effort to manage unreliable devices
  - Consequences of minimal data loss:
    - Small change to results
    - Run experiments longer
    - Exploit the next generation of experiments

Scientific Data - I

Scientific Data - I

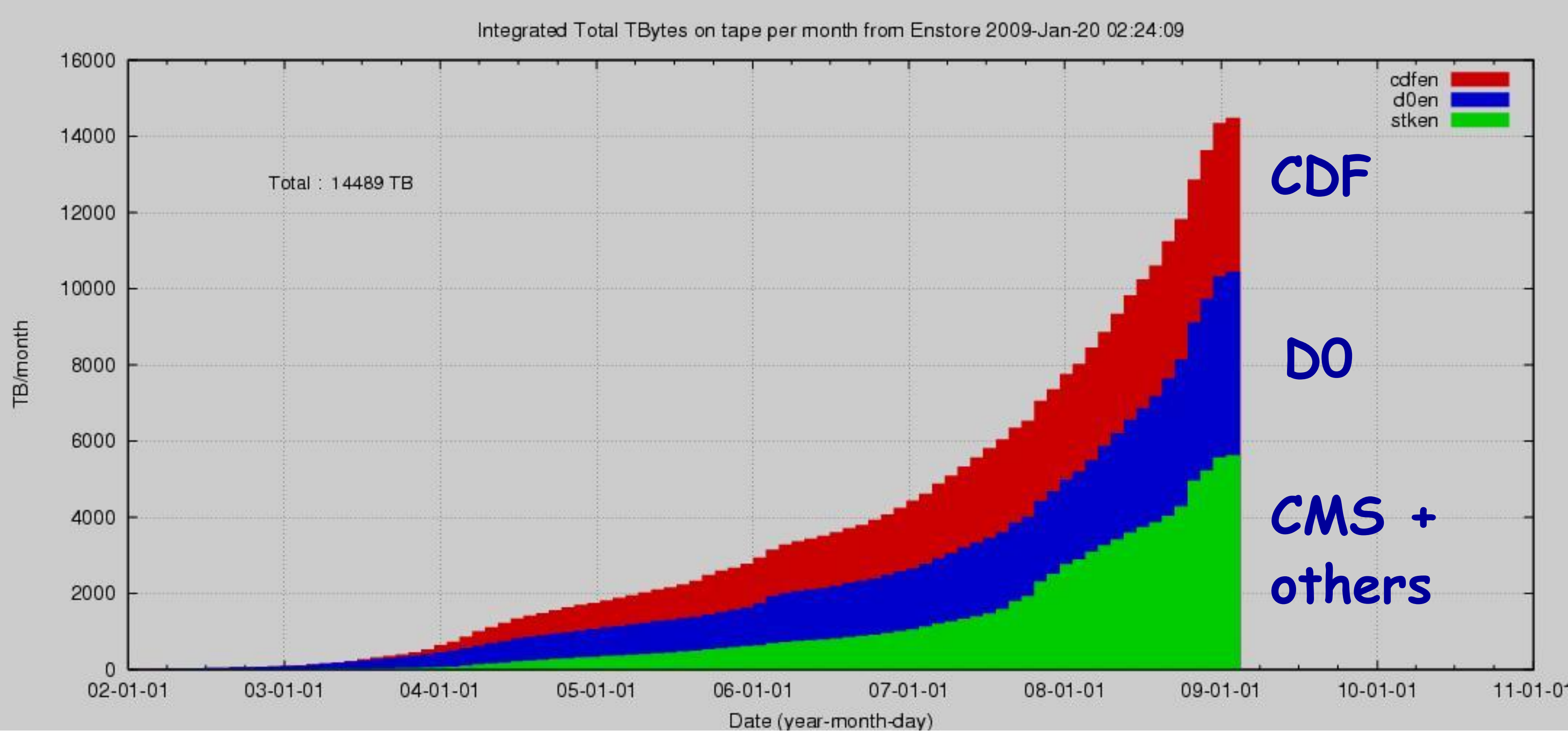
- > 500,000 volumes of media (non robotic)
- CD responsible for tape management, storage in tape vaults and off-site and access
- Experiments responsible for any necessary duplicates, derived datasets, internal data structure
  - CD provides and supports tools
- CD + experiments responsible for methods and software for data access, long-term storage, migration
- Best described as ad-hoc but effective and appropriate
  - It worked

Scientific Data – II
Run 2 era

- Run 2 Era – also robotic storage era at Fermilab
  - Earlier robotic storage was too limited for all of Fermilab's data
  - Limited to special datasets, especially those with a high frequency of access
- Run 2, robotic storage was planned for and grew to encompass all Run 2 data as well as all data from "small" experiments
  - This eventually included, MINOS, miniBooNE, SDSS, KTeV, parts of other experiments and projects

Scientific Data - II

Integrated Mass Storage Since 2002



Scientific Data - II

- **Commitment to Run 2 Data Storage:**
  - Long-term rapid access for production and analysis
  - "Low" data loss – mainly from tape damage
  - Data management (SAM, GRID)
  - Data access and mass storage software (dcache, SAM, Enstore)
- **Timescale**
  - The complete data taking period:
    - 2001-2010+
  - And some period for Run 2 analysis
    - Yet to be defined, but expected to last years beyond the end of data-taking

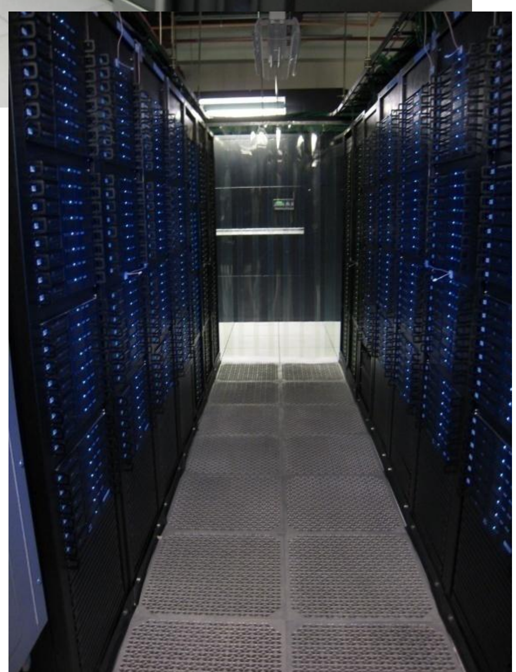
Scientific Data – III
(Sloan Digital Sky Survey)

- Sloan Digital Sky Survey (SDSS)
- Different from the standard HEP experiment.
- Covered by formal agreements with ARC (Astrophysical Research Consortium)
- Data taken on the mountain was copied and maintained in multiple locations (on tape).
  - Has since been copied into Enstore mass storage
- Formal data releases (DR1, DR2, …)
- Final data release (DR7):
  - http://arxiv.org/abs/0812.0649
  - Multiple mirrored copies.
  - Long-term commitment to serve and preserve the releases (5 years)
  - Fermilab + at least two other sites are hosts for the data

Scientific Data – IV
LHC era

- Fermilab is a Tier 1 and LHC Physics Center (LPC) CMS computing center
- There are certain responsibilities that go along with this in terms of data access and archiving
- Some of these responsibilities are spelled out in MOU's and other documents
- In addition to CMS, the Run 2, neutrino, Lattice, Astrophysics and other communities all use the shared mass storage and related facilities
  - In reality, Run 2 will dominate for a few years

Scientific Data – IV
LHC era



Scientific Data – IV

LHC era (for all experiments and projects)

- Scientific Data and File backups are now completely robotic
- Migration is in place to copy all data to higher density technologies as newer densities become available
  - No shelved tapes ever (one hopes)
  - Access "forever"
- This simplifies the issues of what to do with older data but adds some to the cost and operational complexity of the systems.
- A data policy is being developed to understand:
  - Data loss and risk analysis
  - Geographical placement/dispersal of data sets
  - Uniform framework for scientific data management

Data Center Responsibilities

- **Long-term data custodian**
  - Data preservation agreement with the experiments and other internal and external entities
  - Disaster planning and recovery
  - Risk analysis and preparation
  - Agency/regulation requirements
  - Minimize data loss and ensure efforts at data recovery whenever appropriate
  - Duplicate data when necessary
  - Data migration to higher density media automatically
- **Long-term data access**
  - Data format
  - File structure
  - Methods for access to data

Open Issues

- Lifetime of agreements and data preservation requirements
- Software and access method support over long periods of time
- Data format, metadata, file size, hardware and software migration over long time frames
- Experiment requirements
- Laboratory requirements
- What is right or desirable for the field, for the public
- Support load and the ability to provide the support over long periods
- New technologies for data storage and access
- GRID computing
- Cloud Computing
- Beyond LHC, international laboratories

Open Issues

- Access to data:
  - Experiment only
  - Other closely related experiments (CDF/D0, MINOS/NOVA)
  - Others affiliated with the same laboratory
  - Similar physics topics
  - All HEP
  - Public
- Data access methods and delivery rates
- Data integrity, security, and checks
- Instructions and Documentation
- Agreements and policies
- Custodians – libraries, SPIRES, Google?