Fermilab Perspectives in Computing and Data Management

Victoria White,

Associate Lab Director for Computing Science and Technology/CIO

May 16, 2011- Data Preservation Workshop

U.S. DEPARTMENT OF ENERGY

Fermilab

Scientific Computing at Fermilab

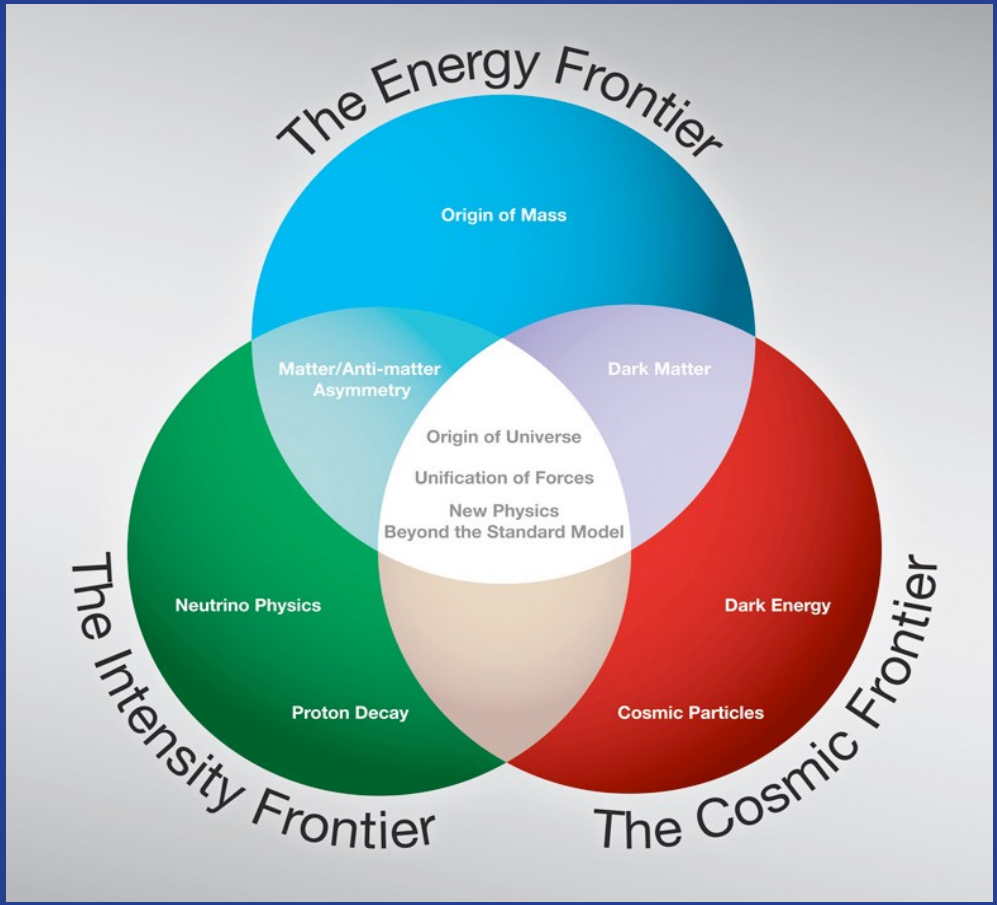
- Scientific computing at Fermilab provides the computing facilities, expertise, partnership and support for the lab's scientific research programs
  - High Throughput Computing for our data intensive sciences
  - High Performance Computing for simulation sciences
  - Cyber Infrastructure in support of science
  - Partnership and technical expertise
  - Education and Outreach

Fermilab Scientific Program:

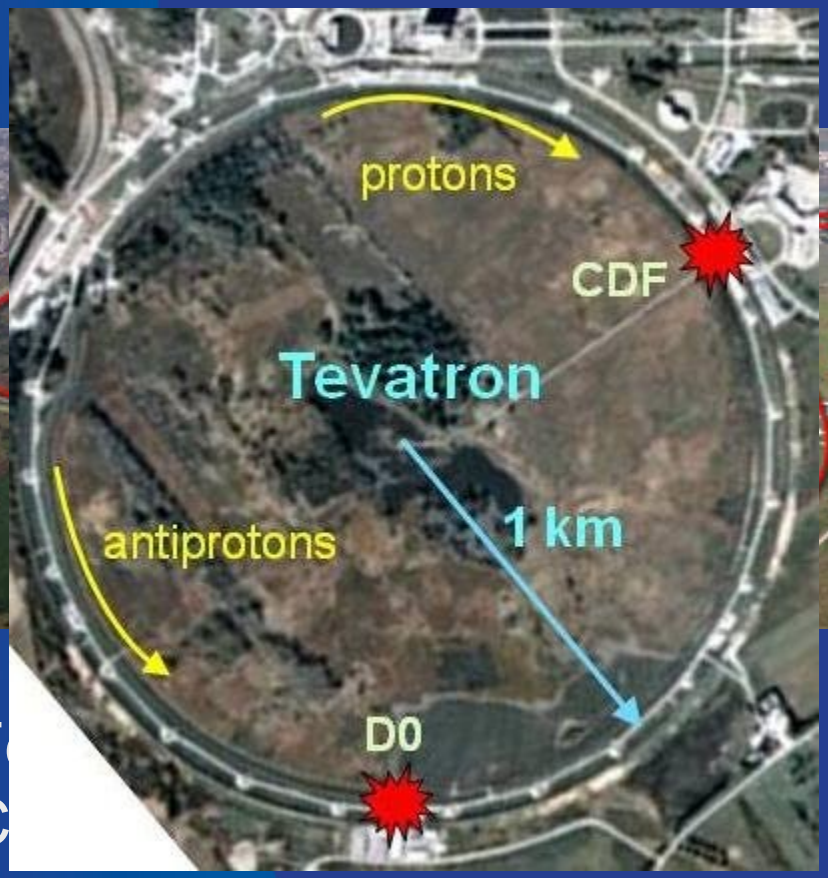
basic research at the frontiers of high energy physics and related disciplines.

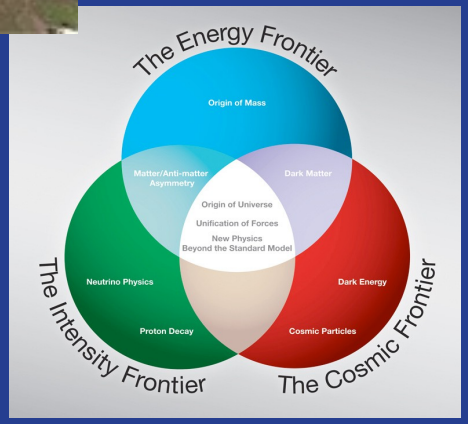
Built on:

Accelerators,
Detectors,
Computing



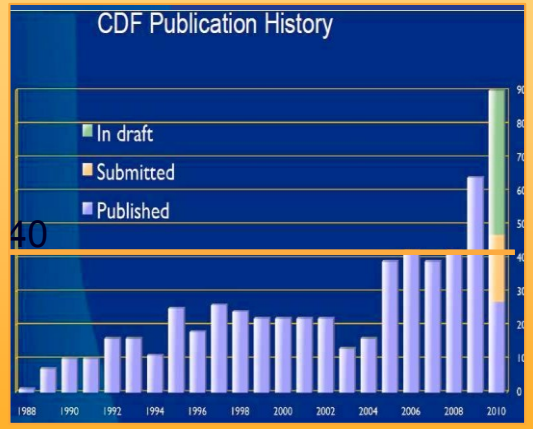
Accelerators for research into the nature of energy and matter (particle physics)



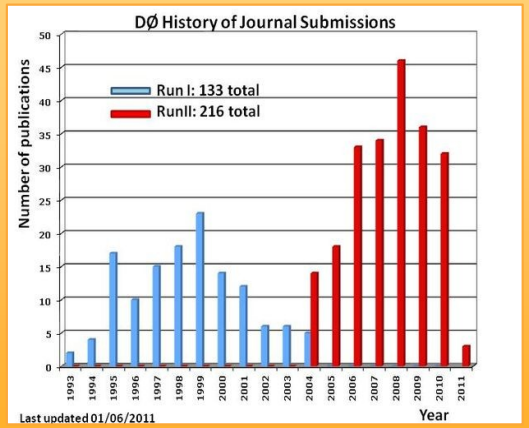


CERN – Large Hadron Collider (LHC) in Geneva Switzerland

Tevatron in the news: looking ahead



CDF Publications each year



D0 Publications each year

CDF and D0 expect the publication rate to remain stable for several years.

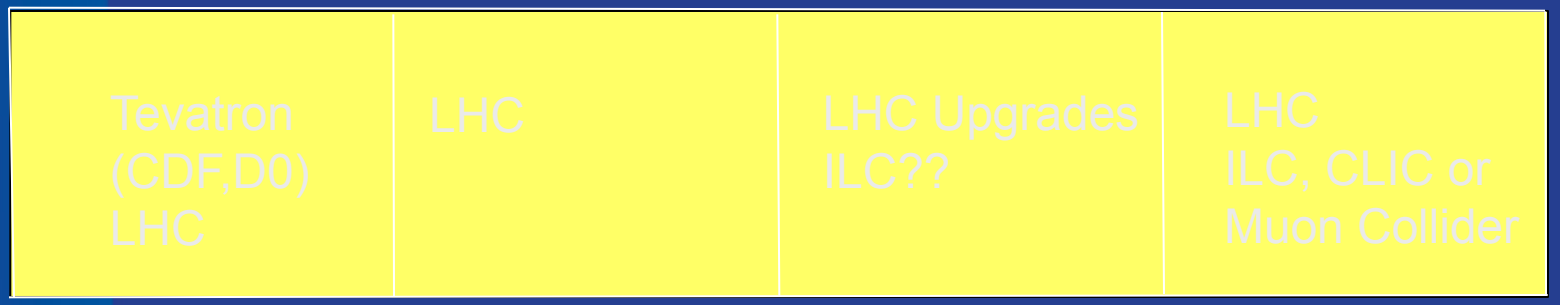
Analysis activity:

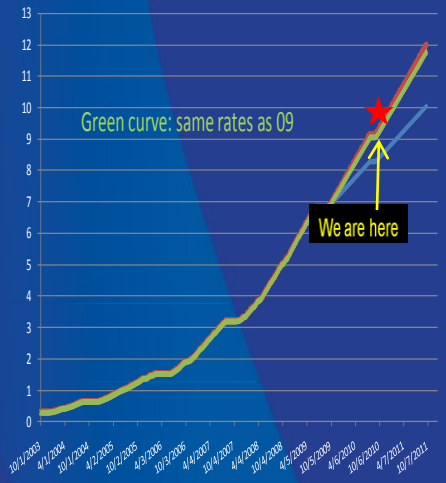
- Expect > 100 (students+ postdocs) actively doing analysis in each experiment through 2012.
- Expect this number to be much smaller in 2015 though data analysis will still be on-going.

Tevatron “Data Preservation” note

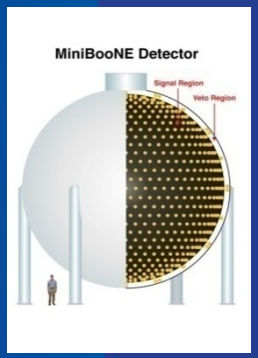
- Collaborations are still strong
- All the data management systems and access to conditions data is working well
- All the codes are maintained and authors are not yet far removed
- Many distributed sites can help with computation – for simulation and analysis
- BUT – full reprocessing of the data today is still prohibitively costly
  - Selective partial reprocessing is in the plans for both CDF and Dzero

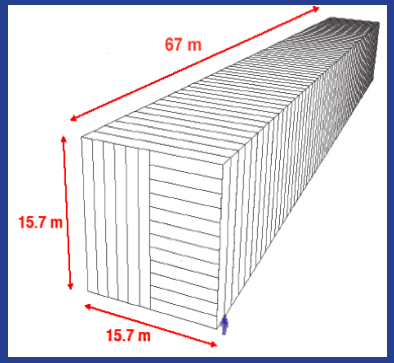
Fermilab Energy frontier roadmap

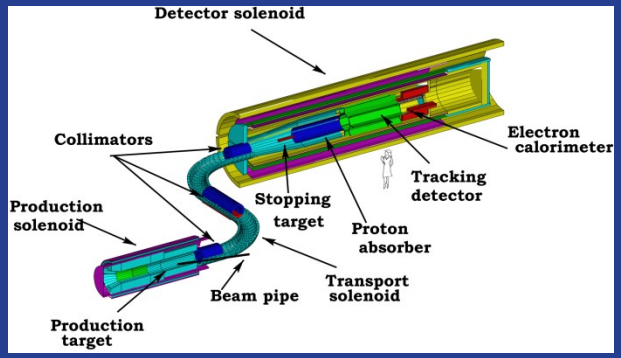


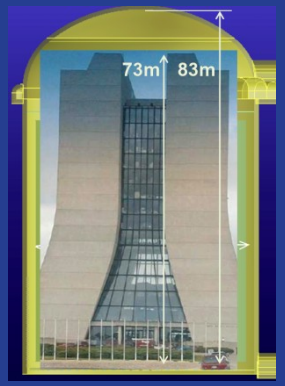


Intensity frontier roadmap

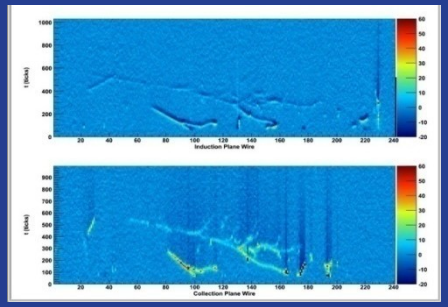


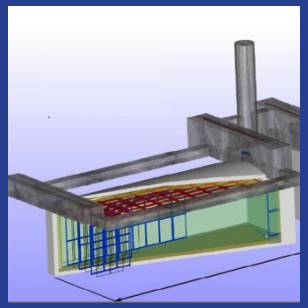




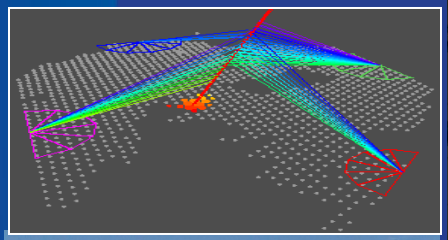


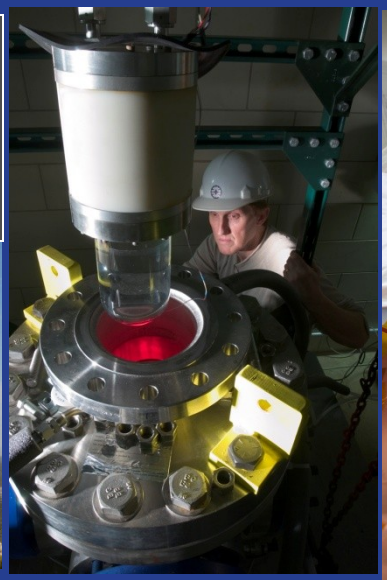
Now	2013	2016	2019	2022
MINOS MiniBooNE MINERvA SeaQuest	NOvA MicroBooNE g-2 SeaQuest	LBNE Mu2e	Project X+LBNE μ, K, nuclear, … ν Factory ??	

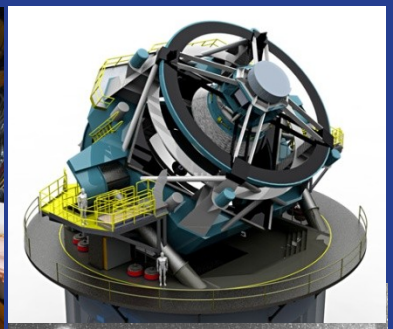




Cosmic frontier roadmap for Dark Matter (DM) and Dark Energy(DE)



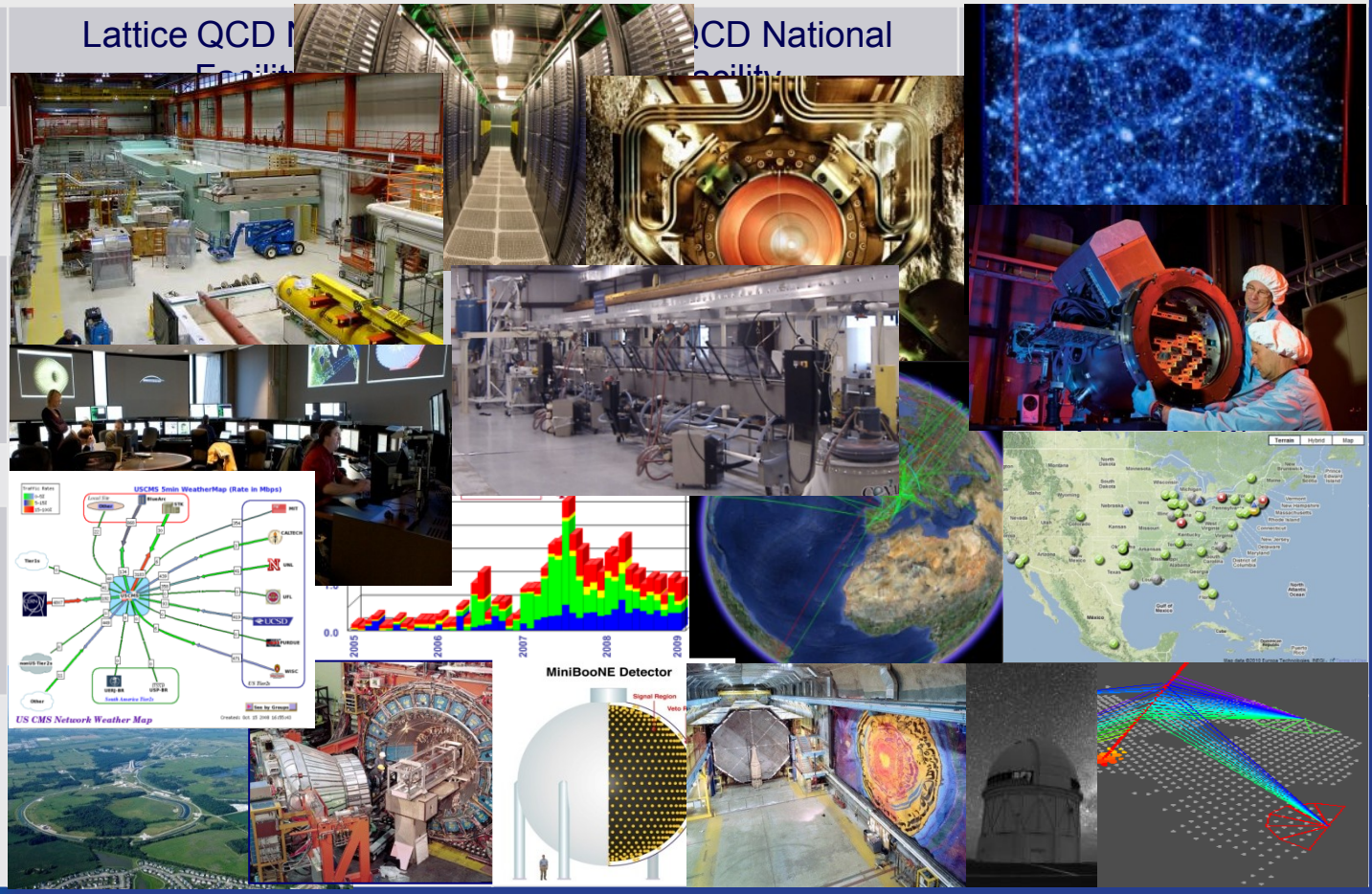





Now	2013	2016	2019	2022
DM: COUPP ~ 10 kg DE: SDSS P. Auger	DM: ~100kg DE: DES P. Auger Holometer?	DM: ~1 ton DE: LSST WFIRST?? BigBOSS??	DE: LSST WFIRST??	

Science using Large Scale User Facilities

Large Scale User Facilities: Skill	Energy Frontier	Intensity Frontier	Cosmic Frontier
Theory	Lattice QCD N … CD National …		
Accelerator Technologies			
Advanced Instrumentation			
Simulation			
Data Analysis & Distributed Computing			
Systems Integration, Operations, Project Management			



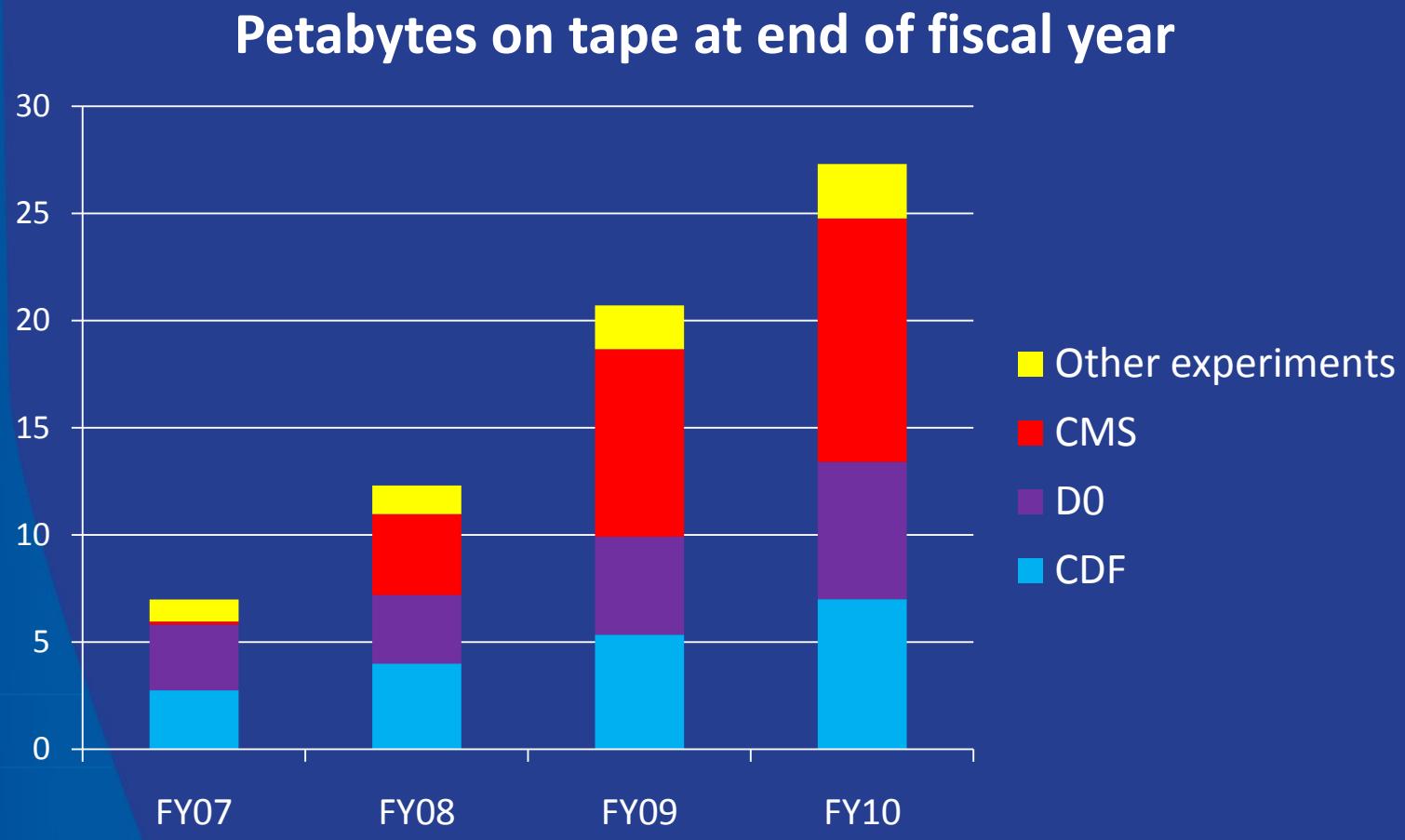
The Fermilab Scientific Program

	Applications	Type of Computing	Computing Facilities
EXPERIMENT	• Detector simulation • Event simulation • Event processing • Data analysis. • DAQ software triggers	High Throughput and Small Scale Parallel (<= number of cores on a CPU)	• Fermilab campus grid (FermiGrid) • Open Science Grid (OSG) • World Wide LHC Computing Grid (WLCG) • Dedicated clusters • FermiCloud
COMP SCI	• Accelerator modeling • Lattice Quantum ChromoDynamics (LQCD) • Cosmological simulation	Large Scale Parallel High Performance Computing	• Local "mid-range" HPC clusters • Leadership class machines: NERSC, ANL, ORNL, NCSA etc.
	• Data acquisition and event triggers	Custom computing	Custom, programmable logic, DSPs, embedded processors.

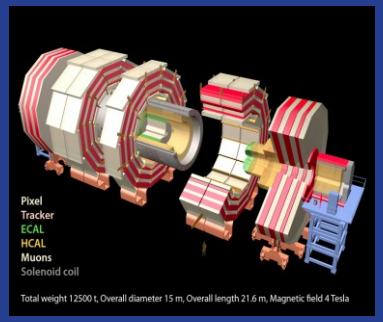
Computing required for experiment data (on all frontiers)

- Triggers to reduce and select data for recording
- Reconstruct Raw data -> physics summary data
- Analyze reconstructed data
- Create simulated (MC) data needed for analysis
- Reprocess data and regroup processed data
- Store and distribute data to collaborators worldwide
- Software tools & services and expert help at times (e.g. detector simulation, generators, code performance)
- Long-term curation of data and preservation of analysis capabilities after experiment ends
- Software frameworks, algorithms and performance tools
- Support for Collaboration on a national and worldwide scale

Data Storage at Fermilab - Tape



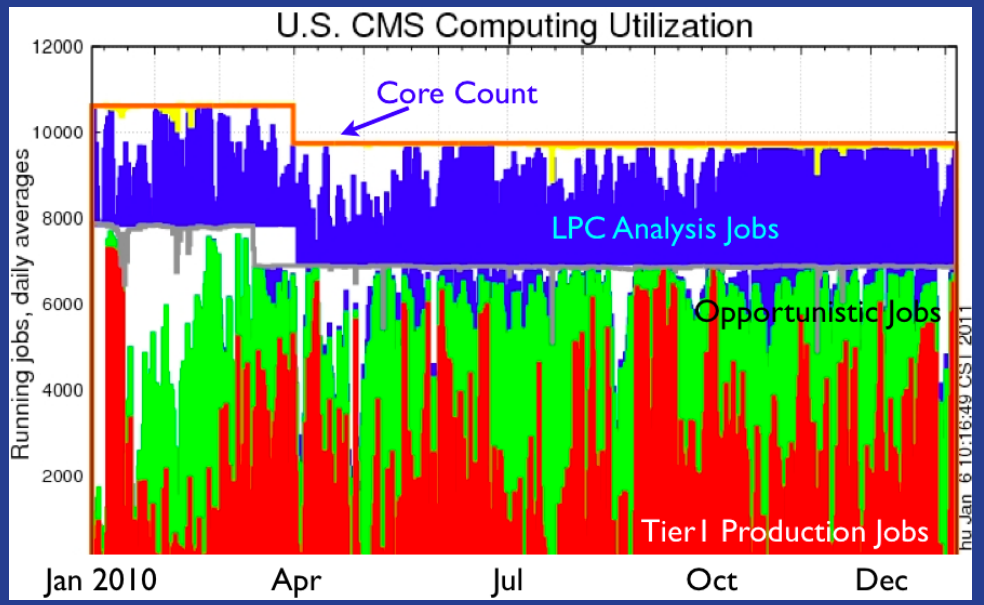
CMS Tier 1 at Fermilab



- The CMS Tier-1 facility at Fermilab and the experienced team who operate it enable CMS to reprocess data quickly and to distribute the data reliably to the user community around the world.

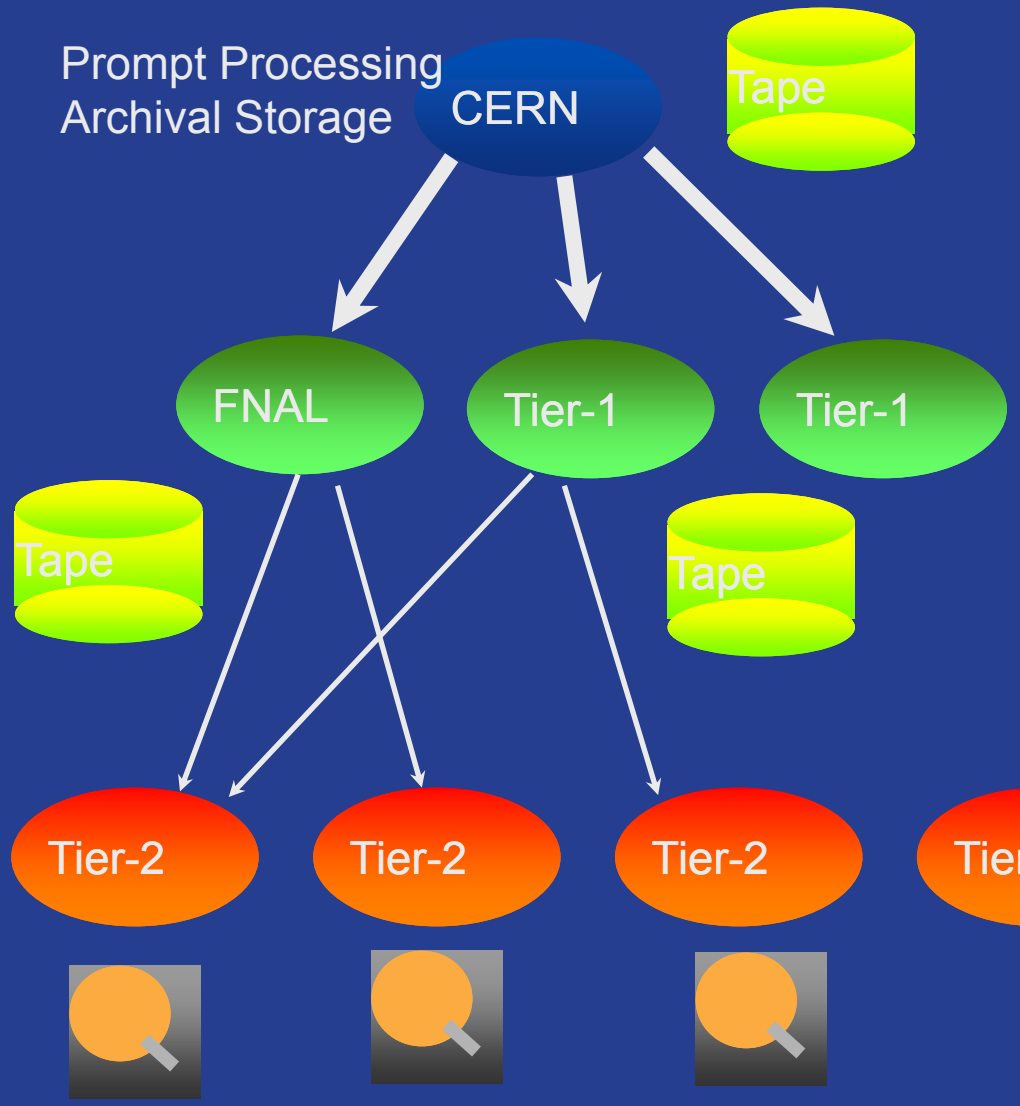
Fermilab also operates:

- LHC Physics Center (LPC)
- Remote Operations Center
- U.S. CMS Analysis Facility



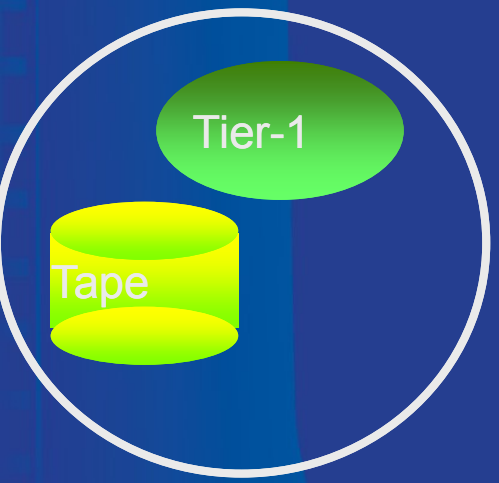
Today: data processing and data

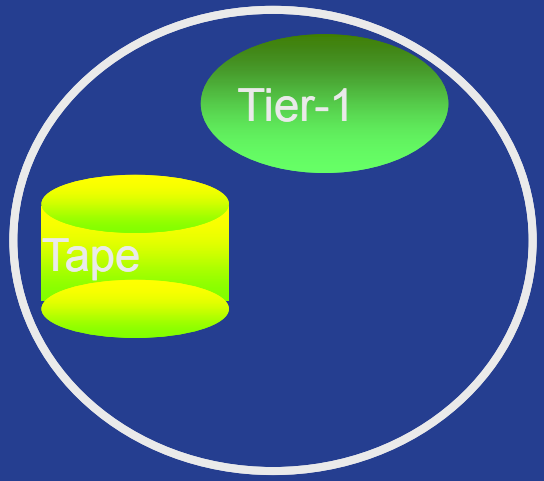
- In modern distributed computing systems the bulk of the processing is located away from the archives

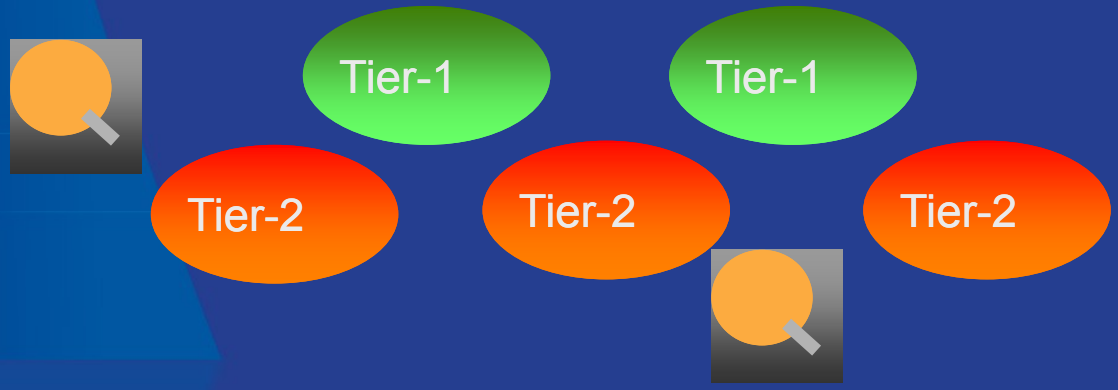


More Efficient Networking

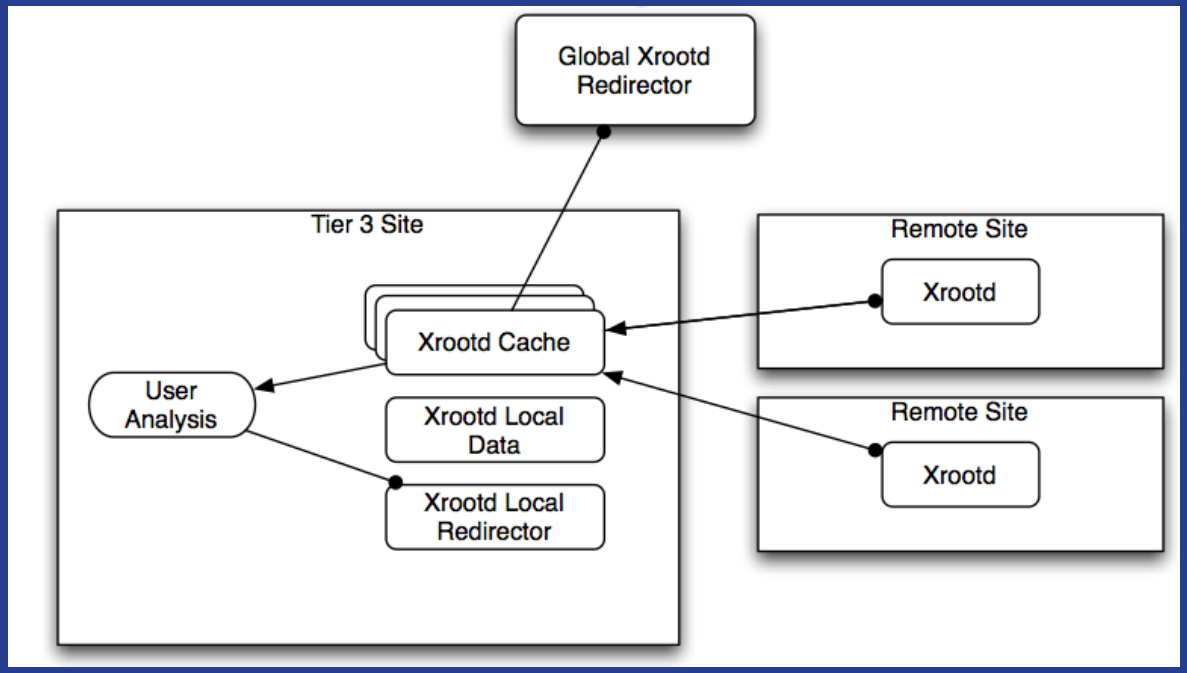
- In the presence of next generation networking and network aware applications, sites could be treated as less independent
  - Benefits of centralized computing combined with distributed







Any Data, Anywhere, Any time: Early Demonstrator



- Root I/O and Xrootd demonstrator to support the CMS Tier-3s and interactive use
- Cost? Value? - will have to be quantified

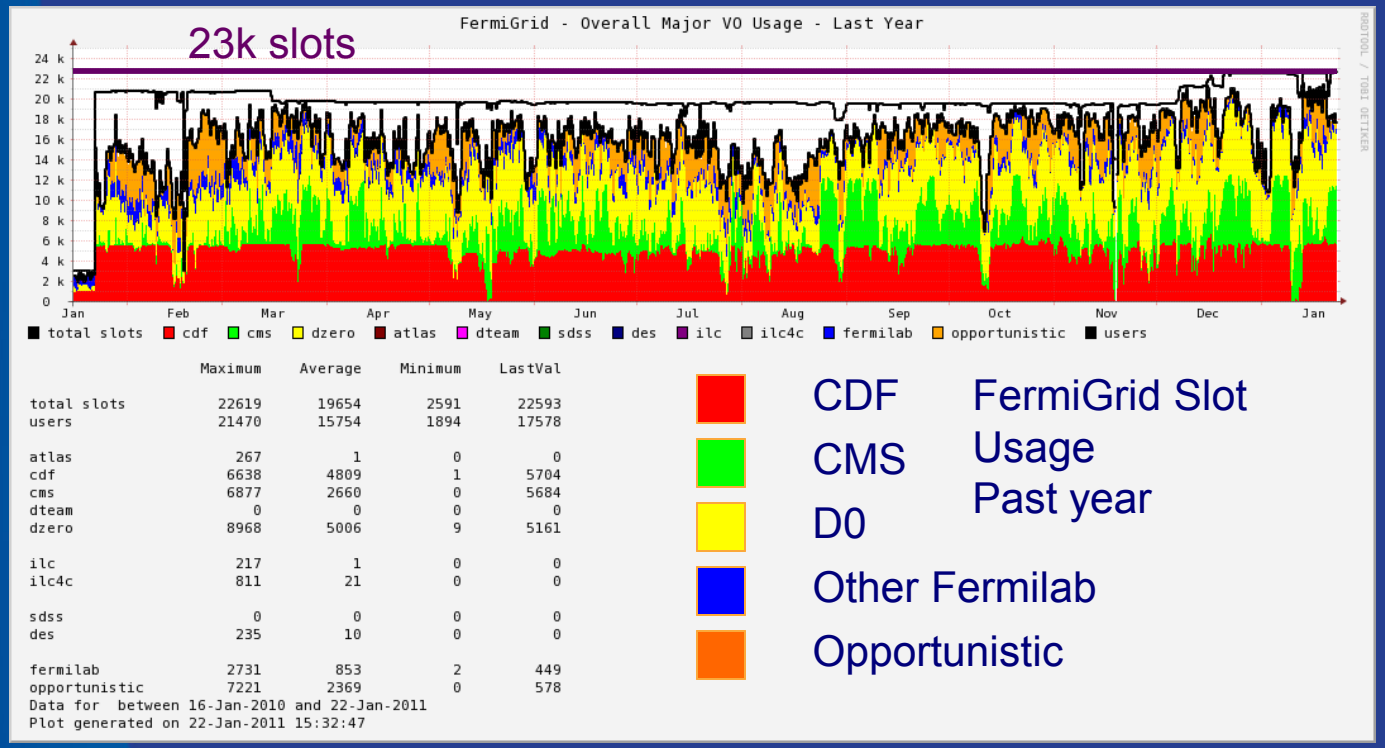
Open Science Grid (OSG)

- The Open Science Grid (OSG) advances science through open distributed computing. The OSG is a multi-disciplinary partnership to federate local, regional, community and national cyberinfrastructures to meet the needs of research and academic communities at all scales.

The US contribution and partnership with the LHC Computing Grid is provided through OSG for CMS and ATLAS

- Total of 95 sites; ½ million jobs a day, 1 million CPU hours/day; 1 million files transferred/day.
- It is cost effective, it promotes collaboration, it is working!

FermiGrid – campus grid and gateway to OSG

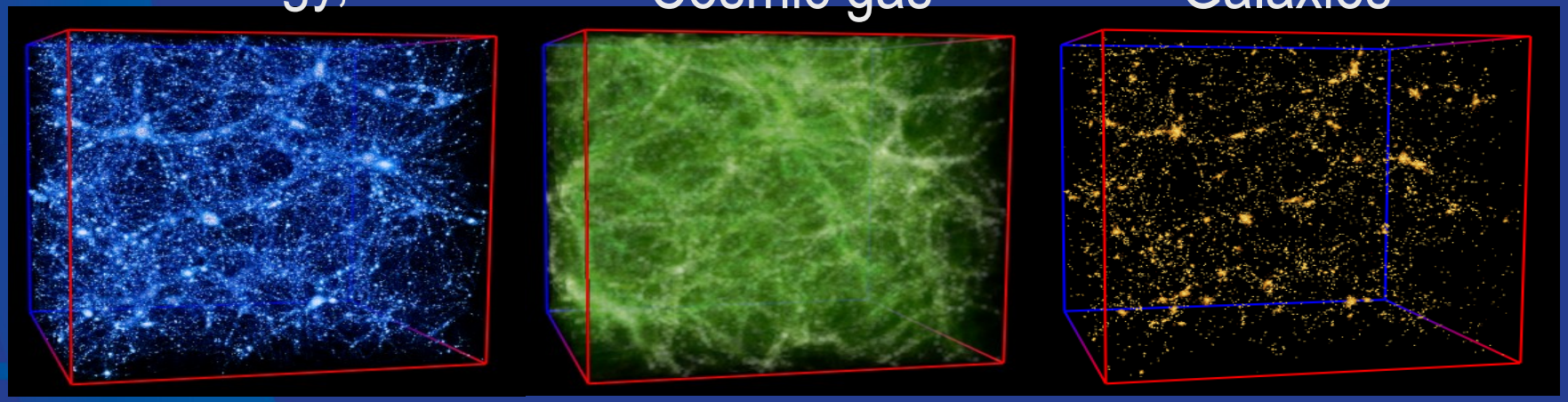


http://fermigrid.fnal.gov

Computing for Theory and Simulation Science – needs HPC

- Lattice Gauge Theory calculations (LQCD)
- Accelerator modeling tools and simulations
  - Fermilab leads the COMPASS collaboration
- Computational Cosmology:

Dark energy, matter | Cosmic gas | Galaxies



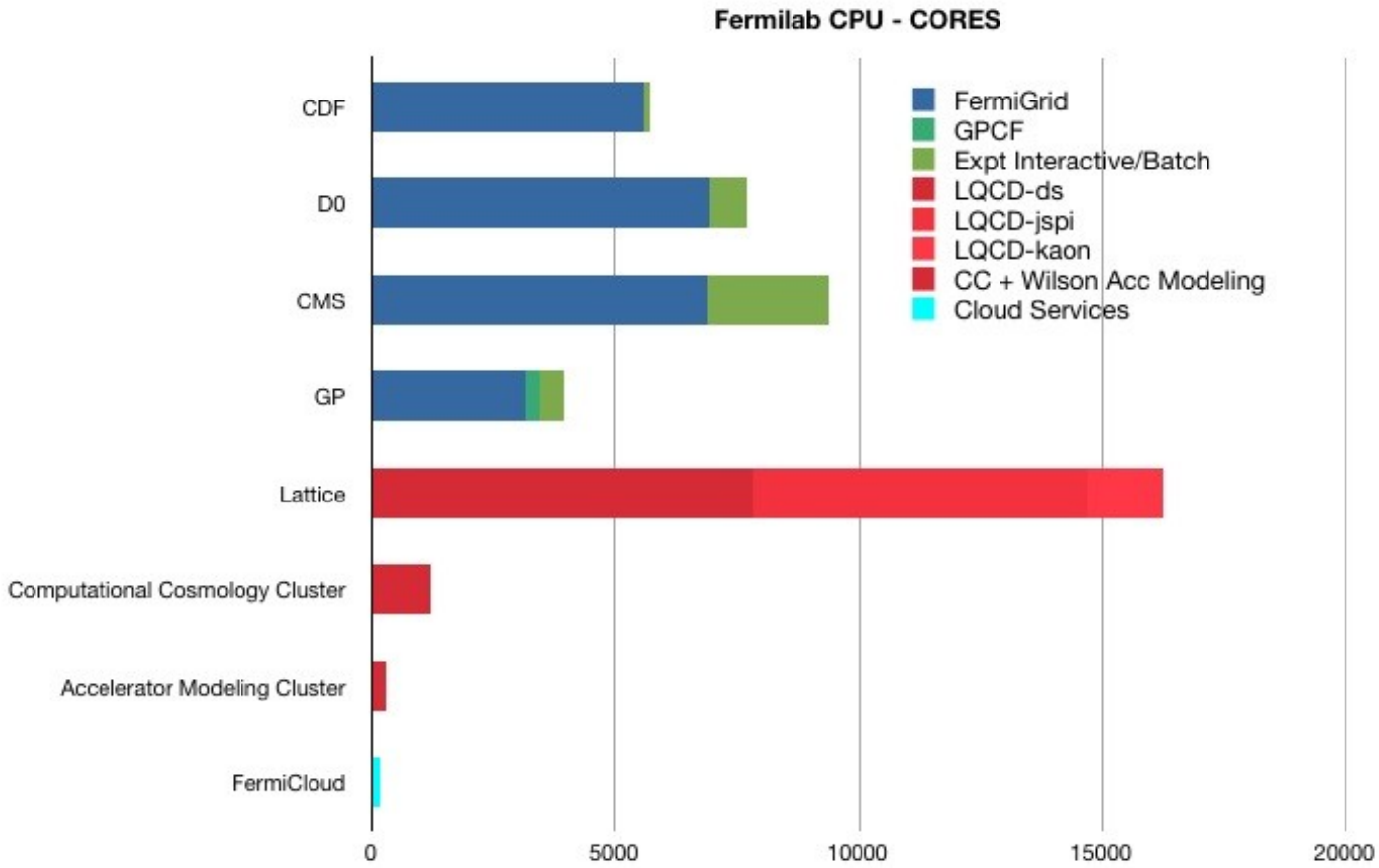
Simulations connect fundamentals with observables

Lattice Gauge Theory: significant HPC computing at Fermilab

- Fermilab is a leading participant in the US lattice gauge theory computational program funded by Dept of Energy (OHEP, ONP, and OASCR).
- Program is overseen by the USQCD Collaboration (almost all lattice gauge theorists in the US)
  - USQCD's PI is Paul Mackenzie of Fermilab.
- Purpose is to develop software and hardware infrastructure in the US for lattice gauge theory calculations.
  - Software grant through the DOE SciDAC program of ~ $2.3 M/year.
  - Hardware and operations funded by the LQCD Computing Project of ~$3.6M/year.

http://www.usqcd.org/

FNAL CPU – core count for science



Fermilab Computing Facilities

- Feynman Computing Center (FCC)
  - High availability services – e.g. core network, email, etc.
  - Tape Robotic Storage (3 10000 slot libraries)
  - UPS & Standby Power Generation
  - ARRA project: upgrade cooling and add HA computing room - completed

- Grid Computing Center (GCC)
  - High Density Computational Computing
  - CMS, RUNII, Grid Farm batch worker nodes
  - Lattice HPC nodes
  - Tape Robotic Storage (4 10000 slot libraries)
  - UPS & taps for portable generators

- Lattice Computing Center (LCC)
  - High Performance Computing (HPC)
  - Accelerator Simulation, Cosmology nodes
  - No UPS

Computer Centers

Feynman Computing Center (FCC)

- 3 computer rooms
- 17891 ft^2 24" raised floor (typical)
- 660kVA UPS for all computing systems
- 1.5 MWatt standby diesel generator
- <4kW/cabinet (typical)
- Security system
- Fire detection/suppression systems
- Under floor water detection system

Lattice Computing Center (LCC)

- 2 computer rooms
- 2841 ft^2 12" raised floor (typical)
- No UPS / ~850kVA total capacity for computing
- Tap boxes to connect portable generators
- 4kW & <8kW/cabinet (typical)
- Security system
- Fire detection/suppression systems

Grid Computing Center (GCC)

- 4 computer rooms / 2 communication rooms
- 10384 ft^2 24" & 36" raised

EPA Energy Star award 2010

- Security system
- Fire detection/suppression systems (including VESDA)

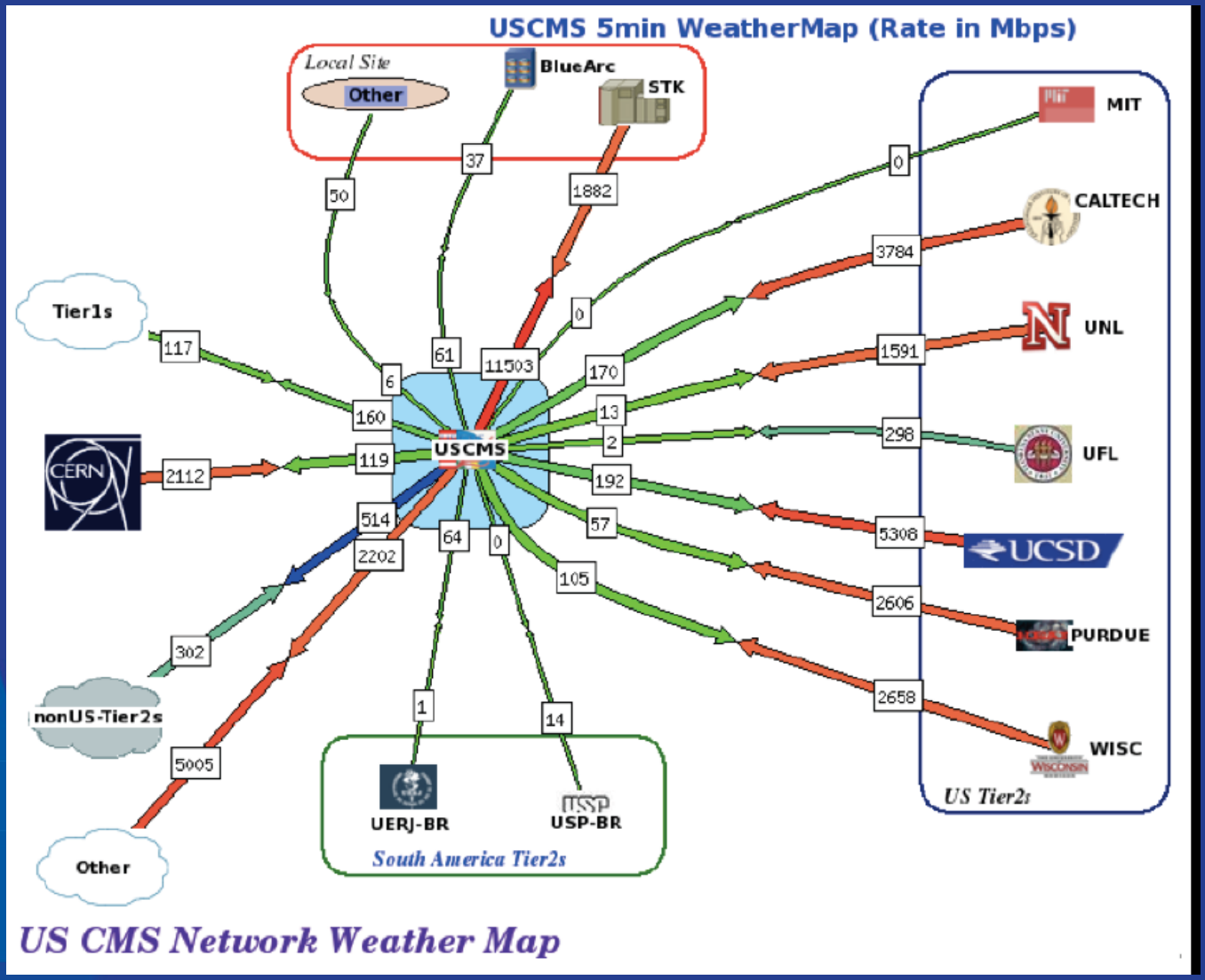
WH8 Fiber Central Server Room

- 1 server room
- 792 ft^2 12" raised floor (typical)
- 20kVA UPS for computing systems
- Standby generator
- <3kW/cabinet (typical)
- Security system
- Fire detection/suppression systems

WH5NW Server Room

- 1 server room
- 490 ft^2
- 18kVA UPS for computing systems
- <3kW/cabinet (typical)
- Security system
- Fire detection/suppression systems

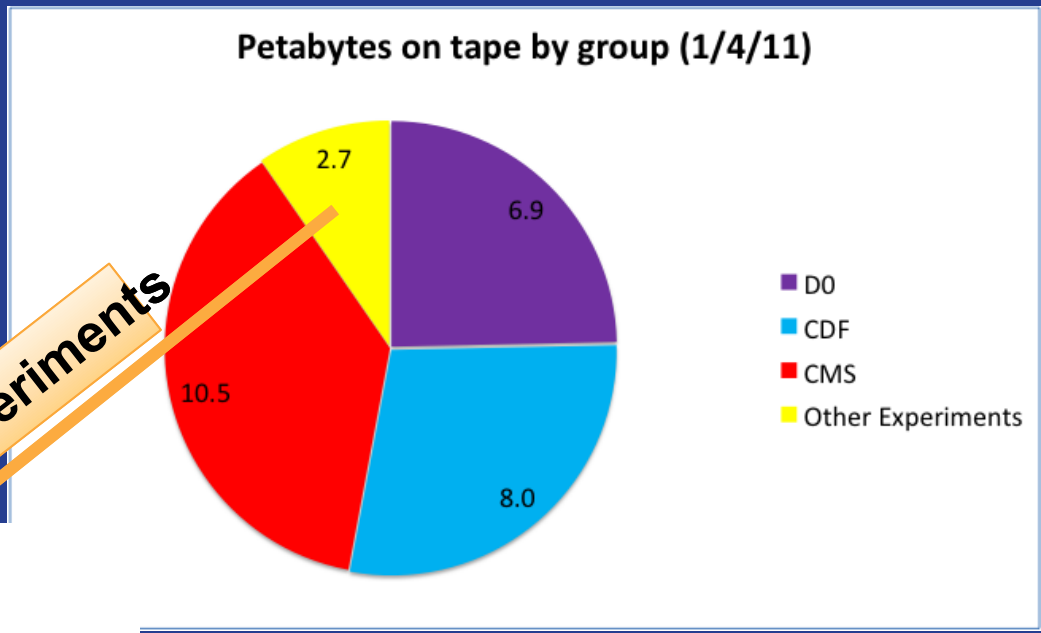
Reliable high speed networking is key



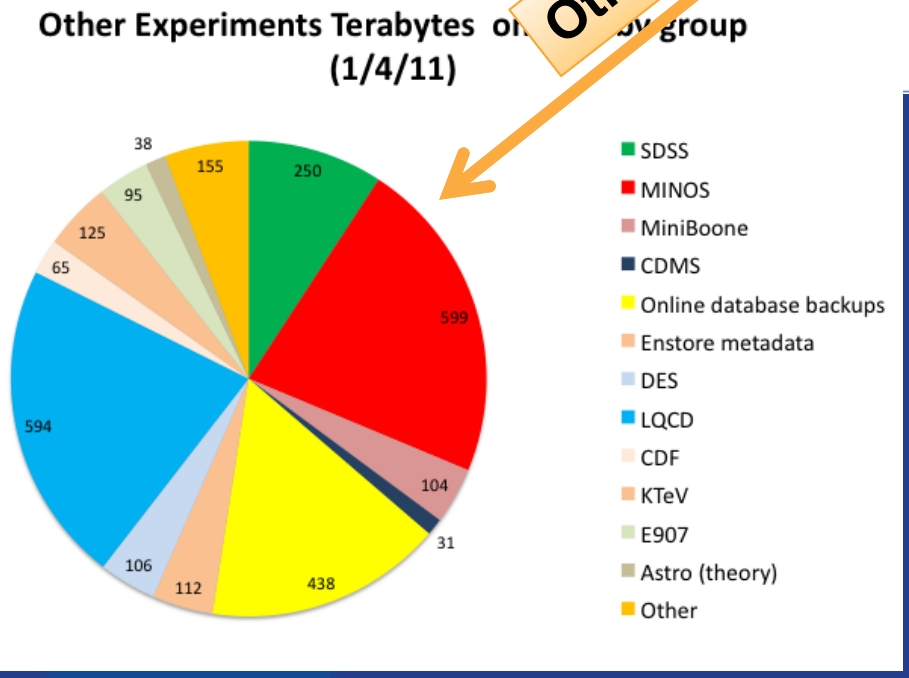
Large and growing datasets for all scientific programs: continuous migration to denser media

- Mass Storage (tape)
  - 6 ORACLE/StorageTek SL8500 Libraries.
  - Total of 60,000 slots (tapes)
  - 4 in GCC, 2 in FCC
    - Allows for geographical distribution of data
  - 141 tape drives
    - Primarily LTO4 (800 Gbytes/tape)
    - LTO5 and T10000C coming online
  - 26 Petabytes of stored data

Data on tape - total



Other Experiments



Data lives a long time (and is migrated to new media many times)

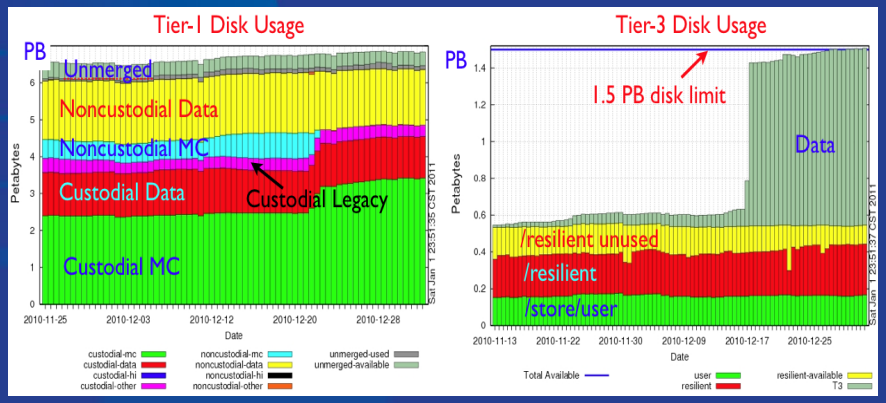
COMPUTING Data Storage Tape Services		Funding Type =====>	Facilities Operations Operations	Program Specific	
	Example Tape Metrics (as of 1/27/2011)		Shared, Common, Core Services	Additional or Targeted Capabilities	
User	Terabytes	Library Slots Used	FY11 FTE	Library slots purchased	Tape Drives Purchased
Core Service:			8.94	$	$
CMS	10,121	15,423	4.18	34,680	$
CDF	7,560	13,160		12,150	$
Dzero	6,491	10,222		9,500	$
LQCD	567	1,020		$	
Intensity frontier				700	
MINOS	554	1,381		$	
Scientific Database Ba	524	931			
SDSS	227	482		L	
KTEV	114	166		L	
DES	97	166		$	$
MiniBooNE	95	192		L	
MIPP	85	166		L	
CDMS	29	49		L	
ILC	16	25			
MINERvA	15	29		$	
Nova	10	18			
Theory Group	8	59		L	
AUGER	7	28		L	
Mu2e	4	6			
All others	79	140			

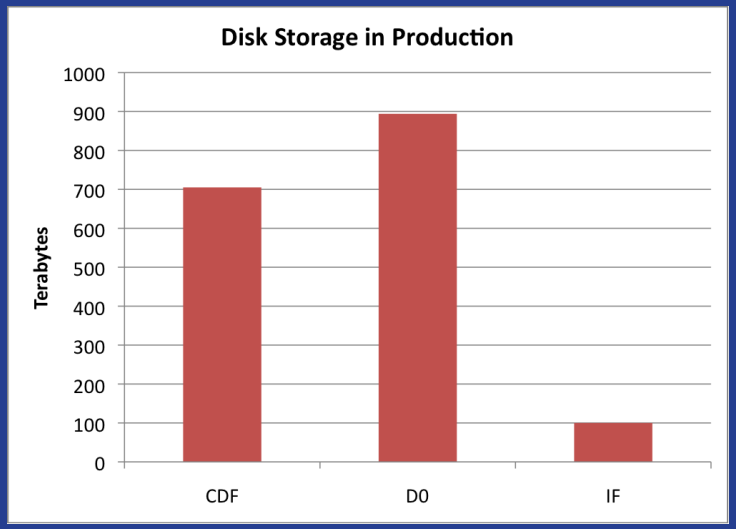
L- legacy tape
$ -contributes funding

User	Terabytes	Library Slots Used
ASTRO	36	52
CHARMONIUM	0	3
COUPP	1	3
DONUT	0	1
E791	0	1
FERMIGRID	0	1
FOCUS	2	8
HYPERCP	10	19
NEES	4	8
NUSEA	0	2
NUTEV	0	1
SCIBOONE	7	13
SELEX	18	28
TOTAL OTHER	79	140

Disk Storage Services

- Large cache storage for D0, CDF, CMS (1, 1, 7 PB)
- BlueArc storage area network (1.3 PB)
- Lustre (distributed parallel I/O used on Lattice QCD and Cosmology clusters and CMS in test)
- AFS – legacy system





FermiCloud: Virtualization likely a key component for long term analysis

- The FermiCloud project is a private cloud facility built to provide a testbed and a production facility for cloud services
- A private cloud—on-site access only for registered Fermilab users
  - Can be evolved into a hybrid cloud with connections to Magellan, Amazon or other cloud provider in the future.
- Unique use case for cloud - on public production network, integrated with the rest of the infrastructure.

Data Preservation and long-term analysis: general considerations

- Physics Case
- Models
- Governance
- Technologies

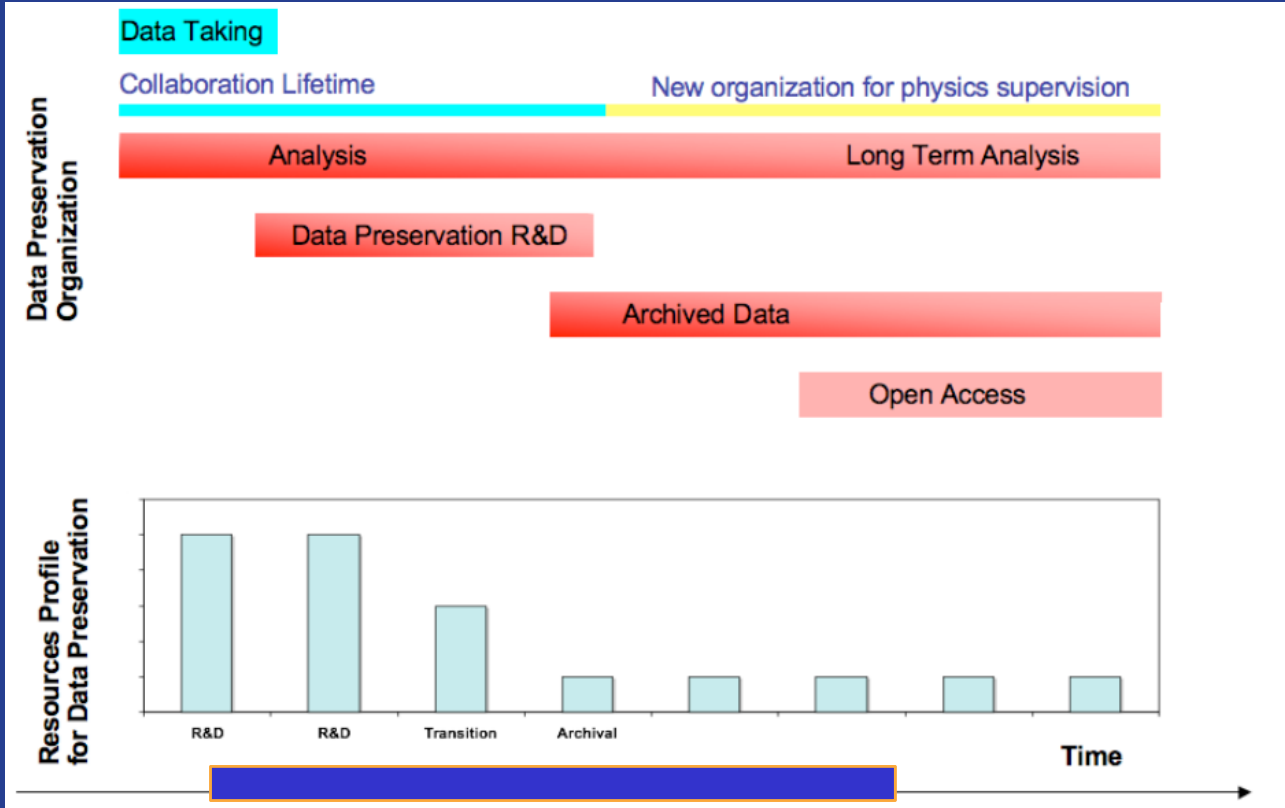
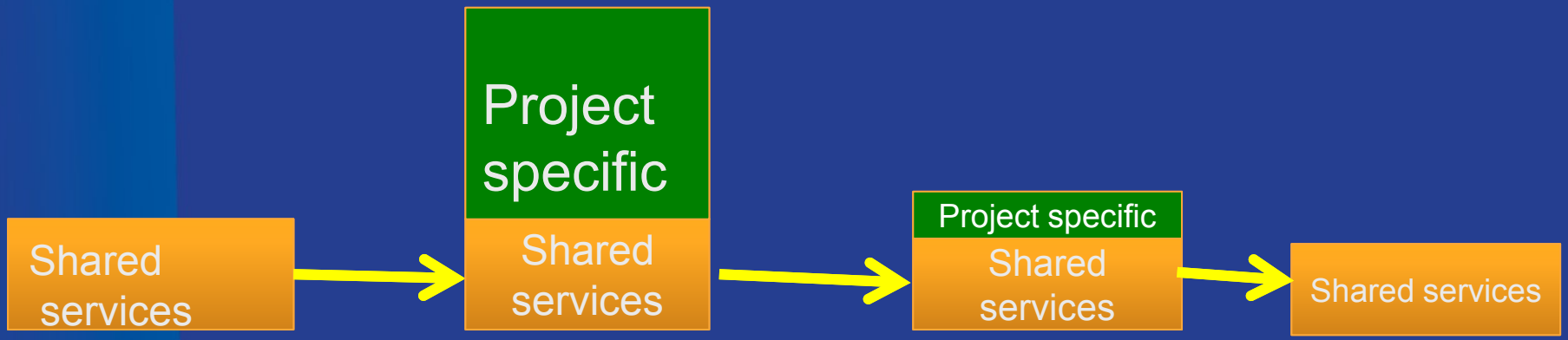


Figure 1: A possible model for data preservation organisation and resources presented as the milestones of the organisation and the resources evolution as a function of time.

Experiment/Project Lifecycle and funding



Early Period
R&D, Ideas, Simulations
LOI, TDR, Proposals

Mature phase
Construction, operations, analysis

Final data-taking and beyond
Final analysis, data preservation and access

Summary thoughts: tradeoffs and value

- Need to build Data Preservation MODELS – just like we have computing models, risk registers, ROI (return on investment) models
  - In the end it is about the value of data and the value of
    - A) doing the upfront work to make data accessible and usable – up to being "open access"
    - B) doing the end-game work to keep the codes, databases, data management systems, workflows and analysis tools alive
  - Value is a function of cost; probability and scientific impact of extracting new science; interests and capabilities of scientists/students/the public to extract new science from old data
- Technology is not the main problem – need the value proposition to be easy to articulate.