Computing model and budget

Rick Snider
Fermilab

Outline

- Computing model evolution
- Computing demand model
- Expenditures and budgets

CDF IFC Meeting
November 7, 2008

Issues that drive the computing model

- Computing demand
  - Raw data logging rate, total data volume
  - Complexity / sophistication of analysis
  - Number of people performing analysis / number of analyses
- Computing infrastructure and operations
  - Budget constraints
  - Evolving grid infrastructure, access policies
    - Access after LHC turn-on?
  - Number of people available for operations support

In general, the computing problem becomes more difficult with time due to increasing demand and declining effort.

Must evolve and adapt to meet these challenges.

Strategies to deal with change

- Manage demand via highly centralized, incremental data processing model
  - Allows most cost effective use of CPU
- Expand use of grid-based resources
  - Leverages effort used to create common tools, support common resources
- Simplify systems, automate operational procedures
  - Reduce cost of systems and effort required to run them
- Increase uniformity of infrastructure
  - Both hardware and software

Infrastructure and operations changes

A number of significant changes over the past year or expected soon in pursuit of these strategies

- Infrastructure
  - Consolidation of on-site CPU resources under Fermigrid
  - Retirement of aging hardware
  - Migration of data to higher density tape technology
- Operations
  - Improved MC processing model: luminosity profile scaling
    - Saves factor of 5 in processing relative to run-based scaling
  - Calibration automation improvements
    - Eliminating manual steps required for each 6-12 week production cycle
  - Production processing improvements
    - Reducing time to get data to tape and recover from processing errors

Infrastructure and operations changes

A number of significant changes over the past year or expected soon in pursuit of these strategies

- Infrastructure
  - Consolidation of on-site CPU resources under Fermigrid
  - Retirement of aging hardware
  - Migration of data to higher density tape technology
- Operations
  - Improved MC processing model: luminosity profile scaling
    - Saves factor of 5 in processing relative to run-based scaling
  - Calibration automation improvements
    - Eliminating manual steps required for each 6-12 week production cycle
  - Production processing improvements
    - Reducing time to get data to tape and recover from processing errors

Also reduces computing demand

Computing demand model

- Used in the budget and planning process
  - Estimate resources required in the future
  - Evaluate the cost of different scenarios
- CPU demand model
  - Two components
    - Production activities
      - Reconstruction, data reduction, Monte Carlo simulation
      - Completely centralized / coordinated processing
      - Demand scales with data logging rate
    - Analysis
      - Decentralized, largely uncoordinated activity
      - Demand scales with total data volume (at worst)
      - The number of people / analyses

Computing demand model

- Used in the budget and planning process
  - Estimate resources required in the future
  - Evaluate the cost of different scenarios
- CPU demand model
  - Two components
    - Production activities
      - Reconstruction, data reduction, Monte Carlo simulation
      - Completely centralized / coordinated processing
      - Demand scales with data logging rate
    - Analysis
      - Decentralized, largely uncoordinated activity
      - Demand scales with total data volume (at worst)
      - The number of people / analyses

This becomes increasingly important with time

Computing demand model

- Production demand model

  Essentially unchanged for past several years

  - Explicitly calculate CPU requirements
    - Measured CPU required for each type of process, type of data processed
    - Event processing overlaps, where appropriate
    - Measured luminosity dependence of reconstruction, ntupling

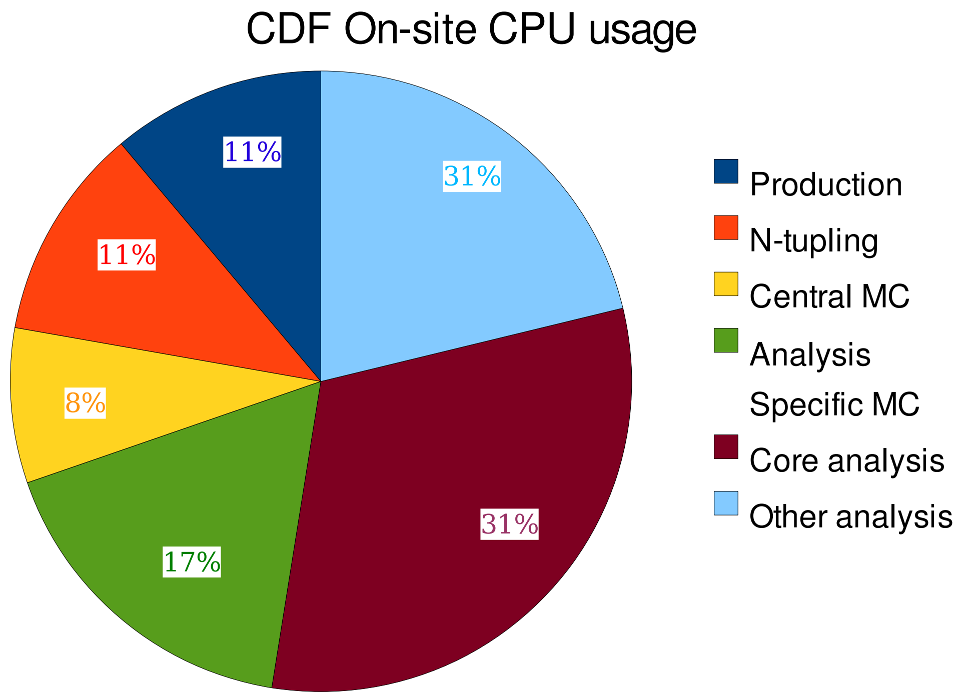
Computing demand model

- Analysis demand model

  New model starting this year

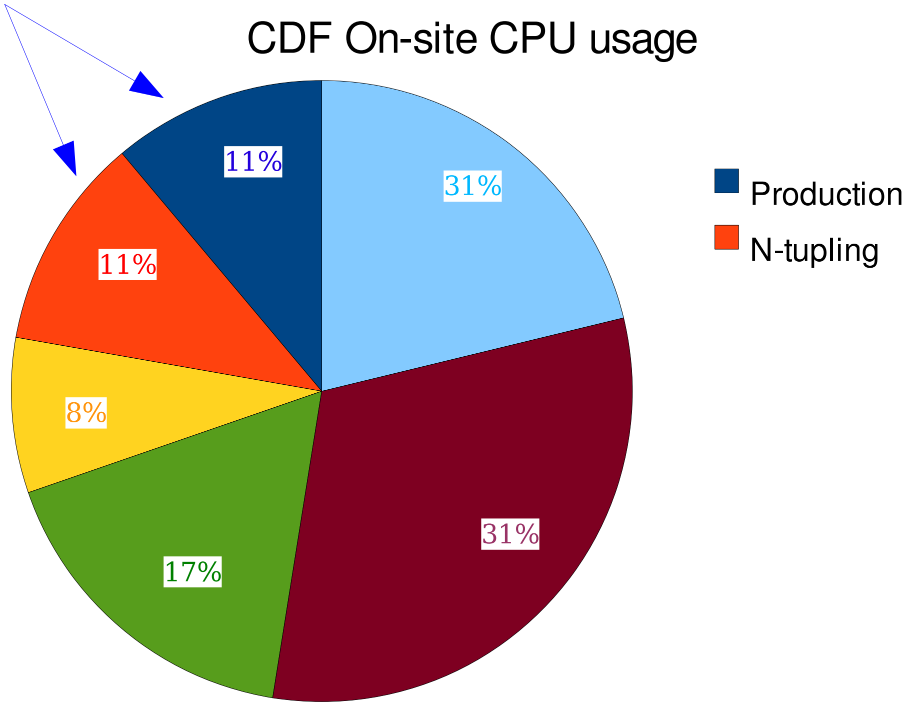
  - Separate analysis into two major categories
    - “Core” analyses
      (as defined in the Tevatron Collider Experiment Task Force Report, Dec., 2005)
    - “Other” analyses
  - Core analyses
    - Assume these are always fully staffed, so computing demand remains high
    - Some evolution in the analyses
      - More complex / sophisticated algorithms (e.g., matrix element methods)
      - Better procedures or more CPU efficient algorithms
    - All current data production activities needed to support core analyses
  - Other analyses
    - Staffed with remaining effort
    - Demand scales with the number of people working on non-core analyses

Measured on-site demand

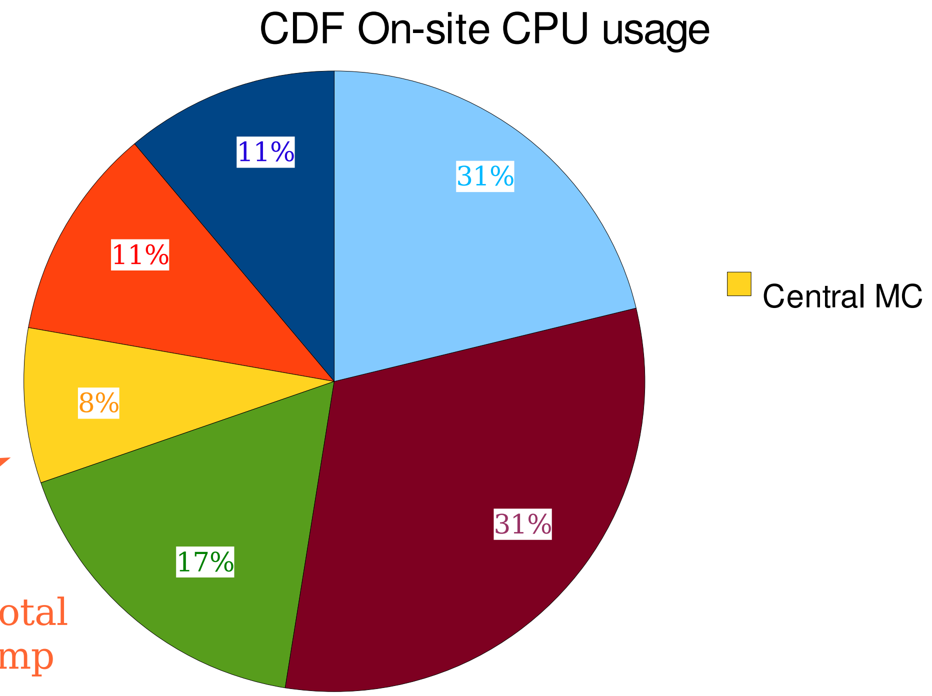


Measured on-site demand

Expect demand to remain approx constant through end of data taking. Then depends upon analysis needs.

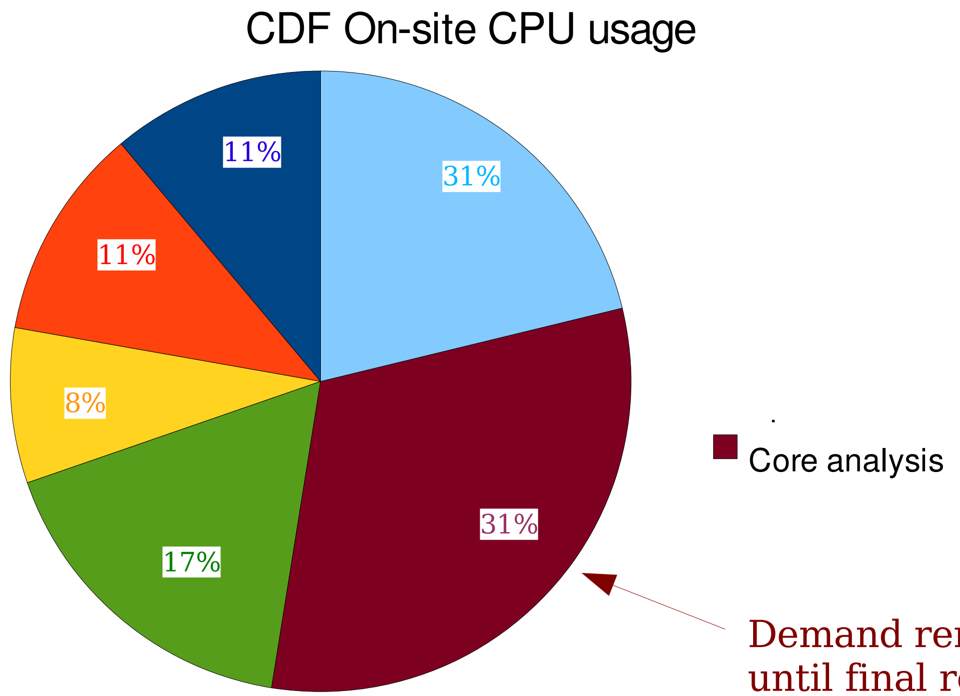


Measured on-site demand



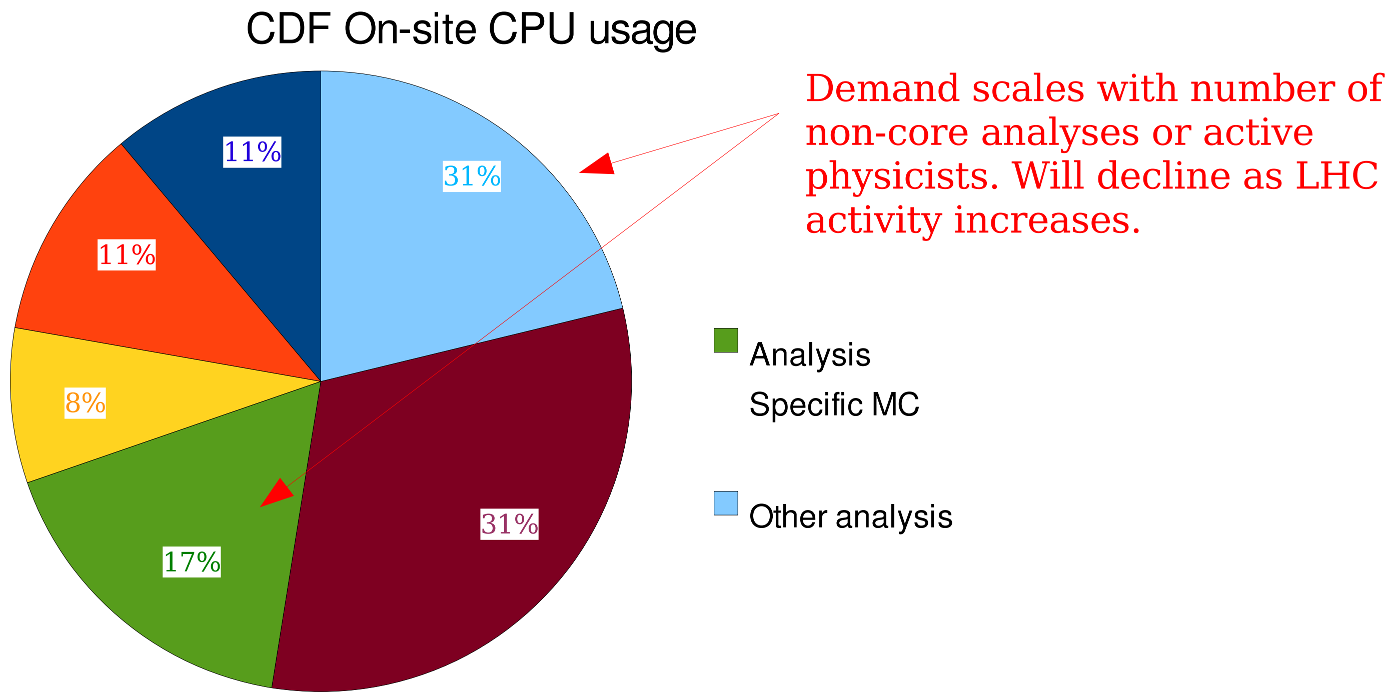
Requirements tied to size of total dataset. Expect demand to ramp down after end of data taking.

Measured on-site demand



Demand remains approx constant until final results are completed. Expect this to be at 2+ years after end of data taking.

Measured on-site demand



Computing projections and budget strategy

- Demand projections for FY2009
  - Existing CPU at Fermilab is sufficient to supply needs of the experiment
    - Assumes that access to existing off-site resources does not change significantly
  - Existing disk volume is adequate
    - Some fileservers are not well-matched to application
    - Optimize procurements to better match hardware to specific uses

Note that the data logging rate last year was 23 MB/s

Budgets and procurements for past few years assumed this would increase

Computing projections and budget strategy

- Budget strategy and priorities
  (continuation of strategy for FY2008 funds)

  1) Cover base costs and buy as much tape as needed
  2) Targeted disk and server upgrades
  3) Replace disk retirements
  4) Replace farm retirements

  For off-site computing

  1) Retain and extend access agreements with off-site resource providers

Major FY2008 Fermilab expenditures

- Equipment
  - CPU: $150k
    - Shifted budget away from CPU into storage and servers
    - Purchased replacements for retiring CPU
  - Data storage: $322k
    - Purchased 20 higher density LTO4 drives
      - Second consecutive year with new tape drive technology (unfortunate, but necessary)
  - Disk and servers: $350k
    - Replaced retirements of critical servers and fileservers
      - Disk volume remains approximately constant.
      - More smaller machines for high throughput applications.
- Operating
  - Tape: $155k
  - Data storage: $155k maintenance
- Total computing expenditures: ~$1.3 M

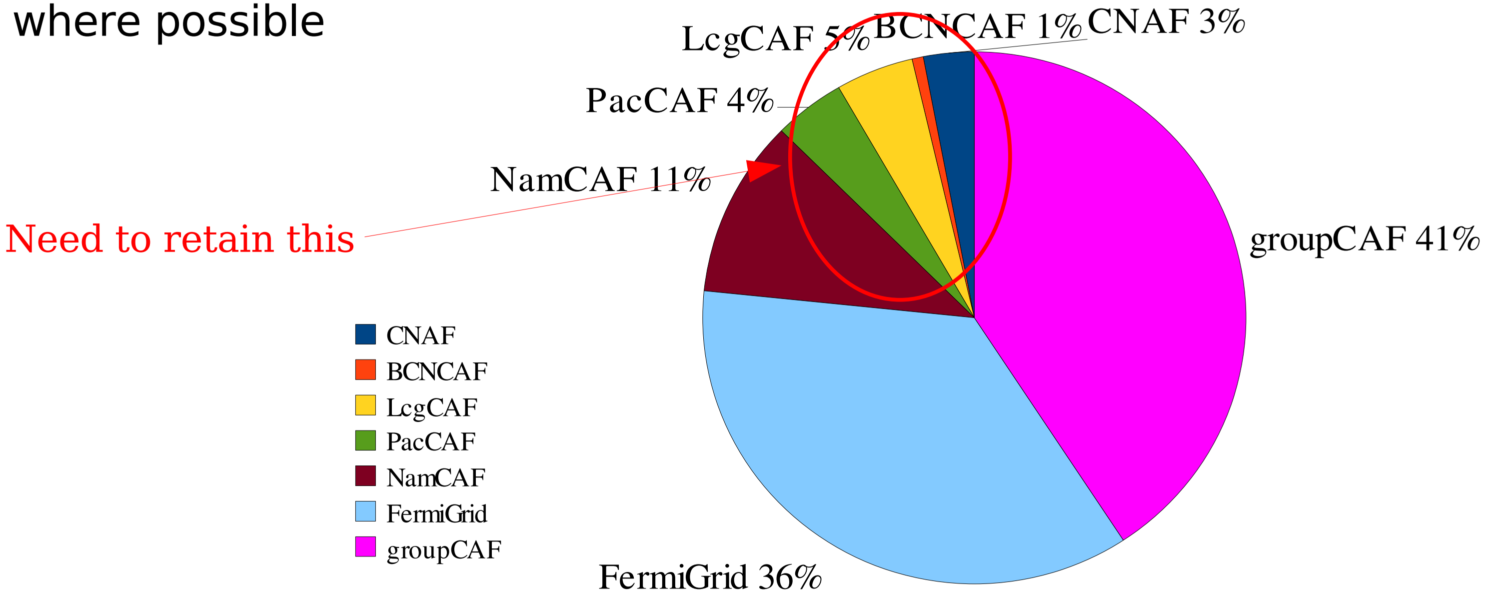
Major items in FY2009 Fermilab budget request

- Equipment
  - CPU: $130k to replace retirements
  - Disk: $390k to replace retirements (small increase in volume)
    - $140k cache + $210k user scratch and project disk + $40 ntuple servers
  - Servers: $160k to replace retirements
    - Includes $50k for production DB server + $70k for service node replacements
- Operating
  - Tapes: $240k
    - $180k for migration from 9940B tapes (2 PB)(includes 2 PB migration from 9940B)
  - Data storage: $130k for robot maintenance
    - $68k will go away/decline with retirement of 9940B tape library
  - Server virtualization project: $60k
- Total budget request: ~$1.4 M

Total costs holding steady

Request to IFC

- Guarantee priority access to some fraction of LCG/OSG pools
  - ► Particularly important after start-up of LHC
- Provide access to LCG/OSG pools if not already allowed
  - ◈ Priority access to some fraction where possible



Summary

- The success of CDF physics program owes much to the strong support and contribution of the Computing Division and IFC to the CDF computing project.
  - CDF has had access to sufficient computing to achieve physics goals as evidenced by copious output of papers and conference results
  - This computing has been provided with a combination of Fermilab, off-site dedicated and opportunistic resources
- Our plan for continued success is to continue evolving and improving the computing model while building on this fruitful collaboration.

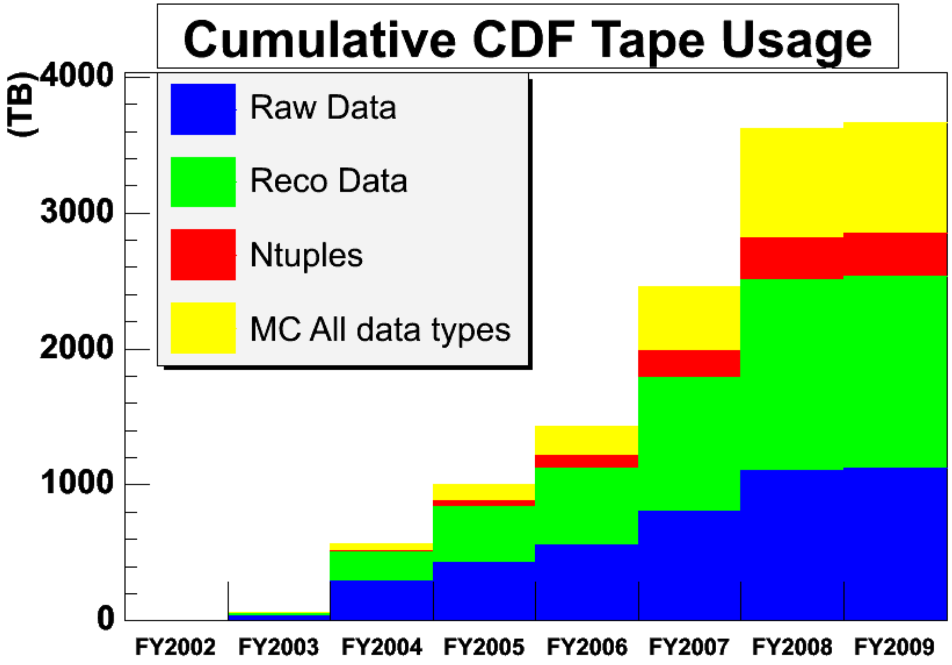
Backup

Core analyses

- Tevatron Collider Experiment Task Force Report (Dec, 2005)
  - Identified “core” analyses that formed the basis of the justification for extended running of the Tevatron:
    - Measurement of Δm_s or limit on B_s mixing;
    - Measurement of $\Delta\Gamma_s/\Gamma_s$;
    - Limit on the branching ratio of the process $B_s \rightarrow \mu^+\mu^-$;
    - High precision measurement of the W boson mass;
    - High precision measurement of the top quark mass;
    - Measurement of single top production cross-section;
    - Search for the Higgs boson both in the Standard Model and SUSY scenarios;
    - Searches for SUSY in the "golden" mode Gaugino-neutralino with tri-leptons;
    - Searches for SUSY in the “golden” mode Squark-gluino with multijets plus missing transverse energy;
    - Searches for high mass resonances in the e^+e^-, $\mu^+\mu^-$, $\gamma\gamma$ and jet-jet invariant mass spectra (sensitive to Large Extra Dimensions, Z' and other processes not present in the Standard Model);

Data volumes

- Data on tape
  - Total of 3.6 PB
  - Raw data
    - 7.9 billion events
  - Monte Carlo data
    - 4.6 billion events
    - Includes a combination of centrally produced MC and analysis-specific MC



CDF opportunistic computing strategy

- Opportunistic computing
  - Access to opportunistic computing until now has been adequate
    - Routinely obtained 1–2 THz during FY2007, FY2008
    - Assumes level of off-site priority access remains at current values
  - Reliance on opportunistic resources introduces risk
    - Resources will become increasingly tight as LHC startup approaches
    - Many competitors for opportunistic cycles
  - Risk mitigation
    - Secure agreements / guarantees for priority access to grid pools
      - Examples: CNAF T1, KISTI T2 centers
    - Reduce logging rate
      - Expected sustained demand for opportunistic cycles is small for logging rate < 25 MB/s
      - This is the regime in which we expect to operate for the foreseeable future

Presented at
2007 IFC meeting

Computing inventory

		Actual		Requirements		
	Fiscal Year	2007	2008	2008	2009	2010
CPU (THz)	Estimated requirement			15	17	18
	Fermilab	7.9	9.6	10	11	12
	On-site contributions	1.7	1.7	1.7	1.7	1.7
	Remote (dedicated)	1.6	1.6	1.6	2.3	2.3
	Opportunistic	1.7	1.7	1.7	2.0	2.0
	Total available	13	15			
Disk (PB)	Estimated requirement			1.0	1.3	1.5
	Fermilab	0.7	1.0	0.98	1.2	1.4
	On-site contributions	0.1	0.06	0.06	0.06	0.06
	Remote	0.1?	0.1?			
	Total available	0.9	1.2			
Volume on tape (PB)		2.6	—	4.1	5.7	7.3