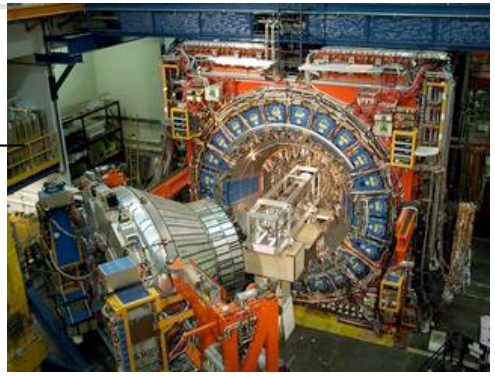
CDF Computing Plan and Budget for FY'2006

P.Murat, Fermilab

- Outline:
  - CDF computing overview
  - Summary of Run II review
  - Computing model and requirements
  - Budget for FY'2006 and projections for FY'2007-08

CDF Computing on one slide



Robotic Tape Storage System

Disk Cache

Reconstruction Farm

Databases
Data Catalog

Batch Analysis Farms
FermiGRID
GRID

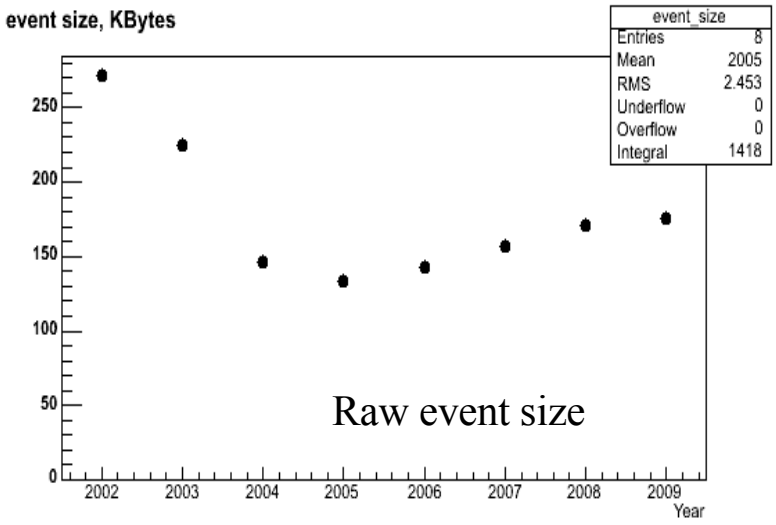
Analysis Disk Servers

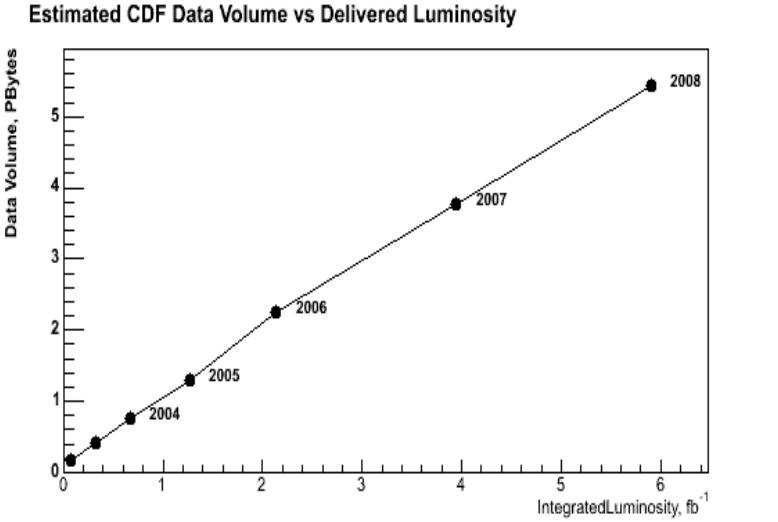
Central interactve login pool

- CPU: Interactive and Batch Analysis Facilities
- Storage Systems
- Databases
- Networking

Globally Distributed Structure of CDF Computing

- Upgraded CDFII detector needs world-wide distributed computing
- Major factor - upgrade of the data logging system :
  - 20 MBytes/sec --> 60 Mbytes/sec
- A lot of work on data compression ->
- Estimated total dataset size:
  - about 6 PBytes in 2008
- Plan presented at IFC'2003:
  - **50% of the CDF computing located off-site**
- FY'2005: about 40% of the CDF CPU resources wer outside the Fermilab





Run II Computing Review'2005

URL: http://cdinternal.fnal.gov/RUNIIRev/runIIMP05.asp

- Reviewed technical design, status of the operations and budget
- CDF presentations:
  - CDF Computing Strategy
  - Status of the CDF software
  - CDF offline operations
  - CDF Computing infrastructure and budget

Run II Computing Review: Summary of the Comments

(Final report still due)

CDF was commended for

- " new developments utilizing existing standard grid tools (Condor glide-in) to extend their functional environment to remote resources."
- "excellent progress on data handling and data processing"
- "achieved stability for their software going into "maintenance" mode, except for particular parts like forward tracking. CDF is monitoring the performance of algorithms as function of luminosities, at low lumi performance of the reconstruction software is very stable"
- "achieving a 6-week turn around for physics quality data-samples, which is an excellent success of data processing and validation, and their ability to prioritize work. There exists already good experience with 1-pass processing"
- **CDF data processing systems seem ready for reconstruction production of 1/fb luminosities**

Run II Computing Review: Summary of the Comments (II)

- "The committee heard that CDF wants to move a significant part of their data analysis running to remote sites, and proposes a model that is similar to how the LHC experiments plan to use their analysis Tier-2s, by moving specific datasets to remote sites for analysis."

Comments:

- We encourage CDF to explore this model, both on the technical side, to understand the implications for the CDF-CAF system, and on the management side, working with the funding agencies to ensure CDF to get "T2-like" resources outside Fermilab, in the US and in Europe.

CDF Computing Requirements

- Strategy:
  - Estimate total computing needs of the experiment
  - Divide the total between the Fermilab and collaborating institutions

- Some institutions continue to locate computing equipment at Fermilab
  - Not counted towards the total requirements if contributed with privileges

- Budget guidance
  - Assume approximately level funding of $1.5M per year from FNAL

CDF Computing: Cost Drivers'2006

- Increase in the data logging rate :
  - 20 MB/sec (2004) -> 60 MB/sec (2007)
  - Analysis dataset doubled in 2005
- Technological progress slower than projected:
  - Moore's scaling (x2 in 18 months) for CPU is not happening (observed 1.3/year, 2 times slower)
  - Doubling of the tape density (200GB/tape -> 400GB /tape, 30 MB/sec -> 60 MB/sec) did not happen in 2005

CDF Computing Model

- performance of the CDF detector:
  - Peak logging rate: 20 MBytes/sec (2002) –> 60 MBytes/sec (2007)
  - data logger upgrade in 2005-2006
- Computing requirements proportional to the total dataset size
- Cost model:
  - x1.3 increase in GHz/$ for CPU per year
  - X1.6 increase in TB/$ for disk per year
  - Retirement policy: 4-year old hardware gets retired
- Normalize to the 2005 data volume
- Planning: acquisitions made in the end of the fiscal year
  - budget'2006 accommodates the needs of FY'2007

Computing Requirements Summary

	2005	2006	2007	2008
Integrated luminosity, fb-1	1.4	2.2	3.8	6.1
Total N events, (x1e9)	2.0	3.4	5.7	9.2
Peak L3 rate, MB/sec	35	60	60	60
Tape volume, PB	1.3	2.2	3.8	5.7
Disk volume, PB	0.3	0.7	1.0	1.4
CPU needs total, Thz	6.5	10.1	17.6	26.5
Onsite CPU, Thz	4.1 (1.5)	6.4 (1.7)	8.3	12.7
Offsite CPU, Thz	2.4	3.7	9.3	13.8

In red: foreign contributions located at Fermilab

- Requirements model: computing needs proportional to the total dataset size
- In 2007-2008 ratio about 50% of the total CPU located offsite

Total Equipment Budget

	2006	2007	2008
CPU ($M)	1.98	1.9	1.75
Disk ($M)	0.42	0.32	0.2
Tape Drives, $M	0.15	0.4	0.33
Interactive computing	0.09	0.02	0.02
Databases	0.03	0.03	0.03
Miscelaneous	0.05	0.05	0.05
Networking ($M)	0.21	0.08	0.08
Total Cost ($M)	2.93	2.80	2.46

- the most expensive component is CPU
- CPU allocation strategy assumes 50% of the total CPU located offsite

Fermilab Equipment Budget

	2005	2006	2007	2008
CPU ($M)	0.7	0.71	1.06	0.88
Disk ($M)	0.51	0.42	0.32	0.2
Tape Drives+Slots ($M)	0	0.15	0.4	0.33
Interactive computing	0.1	0.09	0.02	0.02
Databases	0.05	0.03	0.03	0.03
Miscelaneous	0.05	0.05	0.05	0.05
Networking ($M)	0.14	0.21	0.08	0.08
Total Cost ($M)	1.55	1.66	1.96	1.59
Foreign contributions ($M)	0.23	0.13		

- $0.23M in 2005 and $0.13M in 2006 - contribution from Japan

Disk Requirements

	2005	2006	2007	2008
Disk Needs Total (TB)	343	710	1004	1446
Volume to retire (TB)		116	166	61
Volume to buy (TB)	367	410	608	525
Cost ($K)	520	400	340	180

- Requirements model: total disk volume proportional to the dataset size
- Cost model: $20K per 14 TBytes
- Moore's law(x2 increase in TB/$ in 18mos) followed better than for CPU
- Plan 2006 purchase to meet 2007 needs, retire 4-year old fileservers
- 2 components: disk cache (backed up by the tape) and analysis disk (static)
  - Working on optimization strategy

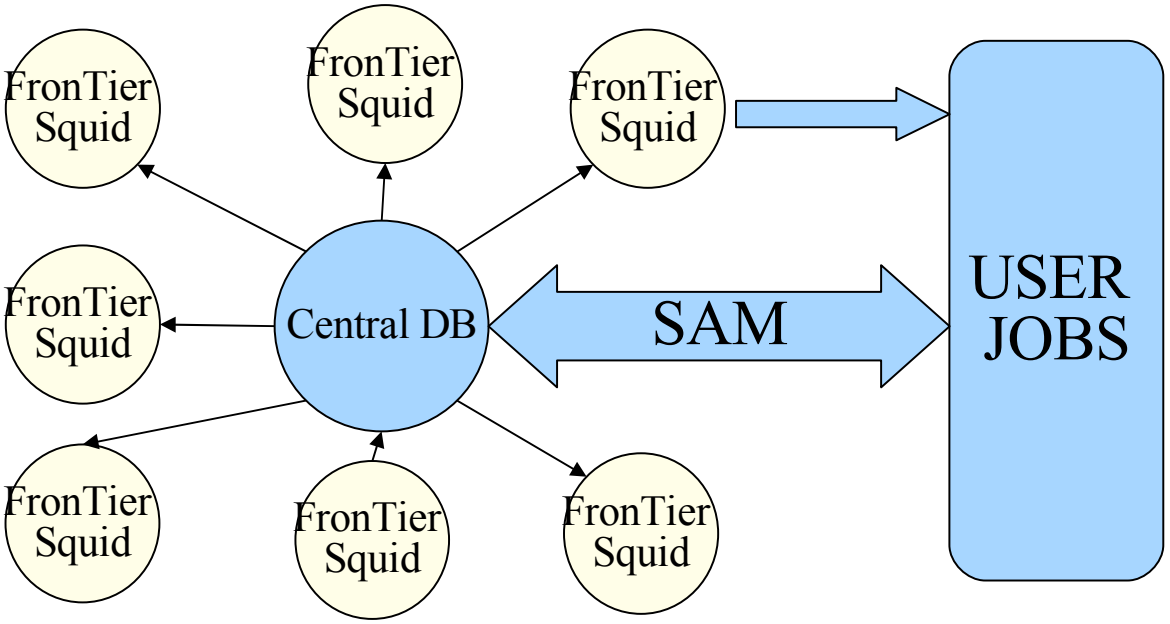
Tape Drive Requirements

	2005	2006	2007	2008
STK 9940B Drives	18	18	18	18
LTOIII Drives		5	10	5
Cost per drive ($K)		16	16	16
Cost, $M		0.08	0.16	0.08

- Recent decision: acquisition of a new tape robot in FY'2006, arrives in March (estimated cost of $400K)
- LT0III drives, in 2006 use 0.2 TB tapes
- Expect that one-time doubling the tape I/O bandwidth will cover the needs of the experiment through 2009

Interactive systems, DB, miscelaneous

- **DB infrastructure:**
  - Central server (ORACLE)
  - Replicas (FronTier squids)



	2005	2006	2007	2008
DB cost $M	0.05	0.03	0.03	0.03
Interactive Systems , $M	0.1	0.09	0.02	0.02
Misc. spendings, $M	0.05	0.05	0.05	0.05

- Moving to distributed databases model (Oracle+FronTier)
- Necessary hardware bought, maintenance stage
- Miscellaneous: not predicted hardware procurements

Tapes and Operating Budget

	2005	2006	2007	2008
Total Data Volume, PB	1.3	2.18	3.77	5.77
Added Data Volume, PB	0.54	0.88	1.59	2
Tapes, $M	0.18	0.21	0.24	0.30
Slots, $M		0.07	0.24	0.25

- Migration to a new tape technology delayed
  - use 200 GB tapes through FY'06
  - Will be buying tape drives capable of handling 400 GB tapes(60 MB/sec)
- Contingency 15% included

Summary

FY'2005 was a very successful year for CDF computing

- 1-pass reconstruction
  - **Impact on physics: reconstructed data available in 4-6 weeks**
  - significant savings on tape, reduced operational overhead
- Deployed SAM - distributed data handling system
- Unified architechture of the batch CPU farm for reconstruction, analysis and Monte Carlo production: CAF
- Significant steps towards GRID:
  - Deployed Glide-in technology allows CDF to use LCG and OSG sites
  - Work on fully GRID-ified solution for LCG (Italy/CDF) and OSG (CDF+CMS) in progress
- FY'06 CDF computing budget request :
  - Total estimated FY'06 cost about $2.9M
  - Estimated FNAL FY'06 equipment cost - $1.66M

Backup

CPU Requirements: "proportional" model

	2005	2006	2007	2008
CPU needs total, (Thz)	6.5	10.1	17.6	26.5
Retired CPU (Thz)	0	0.8	1.3	3.6
Single CPU clock (Ghz)	3.2	4.6	5.9	7.5
New nodes	320	320	480	400
Cost per node ($K)	2.2	2.2	2.2	2.2
Cost ($K)	704	704	1056	880
Onsite CPU	4.1	6.15	8.29	12.66
Offsite CPU needs	2.4	3.95	9.31	13.84

- CPU needs scale as the total dataset size
- Plan to have ~50% CPU resources off-site starting from 2007
- Budget "proportional" model