

# **D0 Computing and Operations Planning Budget 2003-2007**

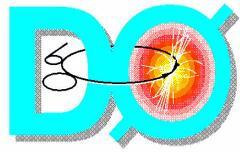
**Amber Boehnlein**

**For the D0 Computing Planning  
Board**

**Director's Computing Review**

**June 6, 2002**

Amber Boehnlein, FNAL



# General Approach

## DO has a highly successful computing model in place

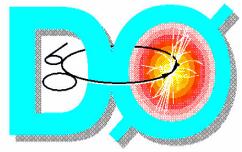
- Sequential Access by MetaData (SAM) catalogs and manages data access
- Robotic storage with reliable drives and media
- Large SMP machine provides high I/O capacity and user access to large amounts of data
- FNAL Reconstruction production farm
- Basic software infrastructure in place
- Fruitful collaboration with the Computing Division on joint projects.
- MC generation performed at collaborating institutions

## Next steps:

Scale up systems to meet needs

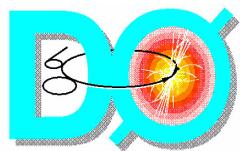
Introduce commodity computing for analysis

Expand computing roles at collaborating institutions



# Analysis Computing

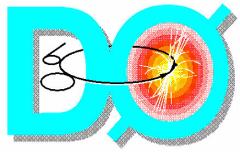
- The size of D0mino was determined by scaling up Run I usage to estimate needs.
  - ◆ Analysis usage on D0mino and CLuED0 currently estimated to be O(350) 500 MHz processors
- D0mino processors slow—augment with commodity computing as a backend.
- Estimate overall analysis needs
  - ◆ Assume that FNAL provides at sufficient analysis computing to generate derived data sets and to support analysis
  - ◆ Collaborating institutions to supply the rest either in desktop computing, CLuB or in a more sweeping model.



# Analysis Patterns

Use current access and analysis patterns to make an estimate of needs

- ◆ **Physics group coordinated efforts**
  - ▲ Derived data sets by skimming through data sets (DST or TMB)
  - ▲ Picked event samples of raw data for re-reco studies
  - ▲ Specialized reprocessing of small data sets
- ◆ **Physics topic analysis includes generation of test samples, trigger simulation, background studies, efficiency studies**
- ◆ **User level analysis primarily takes place on skimmed data samples on high level tier**
- **Assign three usage classes to make estimate of the computing needed for analysis**

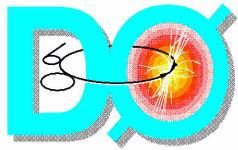


# Analysis Assumptions

Assume that analysis breaks down into usage categories

	Jobs	Data Set (%)	Duration	Processing Time(500 MHz)
Long	6	30%	12 weeks	5 sec/event
Medium	50	10%	4 weeks	1 sec/event
Short	150	1%	1 week	0.1 sec/event

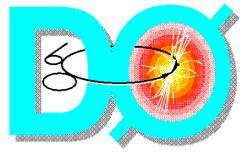
This leads to an estimate (rounded) of 4 THz for analysis CPU for the Run 2a data sample.



# Analysis Computing

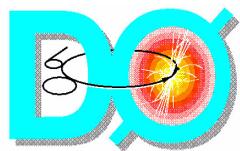
Total Data Sample:	7.89E+09												
Offline Efficiency:	70%												
Contingency:	0%												
Analysis Type:	Short	Medium	Long										
Time/event:	0.1	1.0	5.0										
% of Data Sample:	1%	10%	30%										
Duration (Days):	7	30	90										
Number of Jobs:	150	50	6										
Total Event Fraction:				10%	10%	20%	30%	30%			Total		
Analysis Type	THz CPUs at End of Run	FY03, 3GHz Nodes		FY04, 4GHz Nodes		FY05, 6GHz Nodes		FY06, 10GHz Nodes		FY07, 15GHz Nodes		Target	
		No. Nodes	Cost	No. Nodes	Cost	No. Nodes	Cost	No. Nodes	Cost	No. Nodes	Cost	No. Nodes	Cost
Short	1.40	31	77,500	23	57,500	31	77,500	28	70,000	18	45,000	131	327,500
Medium	10.87	244	610,000	183	457,500	156	390,000	234	585,000	234	585,000	1051	2,627,500
Long	6.52	146	365,000	110	275,000	93	232,500	140	350,000	140	350,000	629	1,572,500
Total:	18.79	421	1,152,500	316	865,000	280	750,000	402	1,105,000	392	1,055,000	1811	4,927,500

Calculation uses Run 2 estimated collected event sample through 2007



# Infrastructure Costs

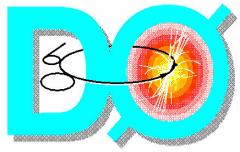
- Database – machines and disk
- Networking
  - ◆ Expand links between buildings, FCC
  - ◆ Additional switches for DAB, farms
  - ◆ D0 to FCC upgrade to 10 Gb backbone upgrade '06, Rewiring D0 for Gb to desktop in '06
- Linux build machines and disk
- Web servers
- I/O machines for small SAM stations



# Infrastructure Costs

<u>Infrastructure Costs</u>	2003	2004	2005	2006	2007	Total
<b>Databases:</b>						
<b>Server upgrades</b>	\$60K	\$60K	\$0	\$25K	\$25K	
<b>Disk and controllers</b>	\$60K	\$20K	\$10K	\$10K	\$10K	
<b>DB system replacement</b>			\$300K			
<b>Software</b>	\$50K	\$0	\$50K	\$0	\$50K	
<b>DB totals</b>	<b>\$170K</b>	<b>\$80K</b>	<b>\$360K</b>	<b>\$35K</b>	<b>\$85K</b>	<b>\$730K</b>
<b>Networking</b>	\$120K	\$120K	\$100K	\$500K	\$100K	\$940K
<b>Build Machines/web servers</b>						
<b>Build Machines/web servers</b>	\$60K	\$60K	\$60K	\$60K	\$60K	\$300K
<b>Datahandling servers</b>	\$50K	\$50K	\$50K	\$50K	\$50K	\$250K
<b>Total, fixed cost</b>	<b>\$400K</b>	<b>\$310K</b>	<b>\$570K</b>	<b>\$645K</b>	<b>\$295K</b>	<b>\$2,220K</b>

Not included, but important is FNAL Connectivity

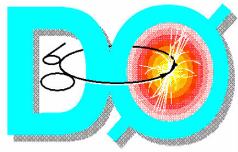


# Robotic Storage

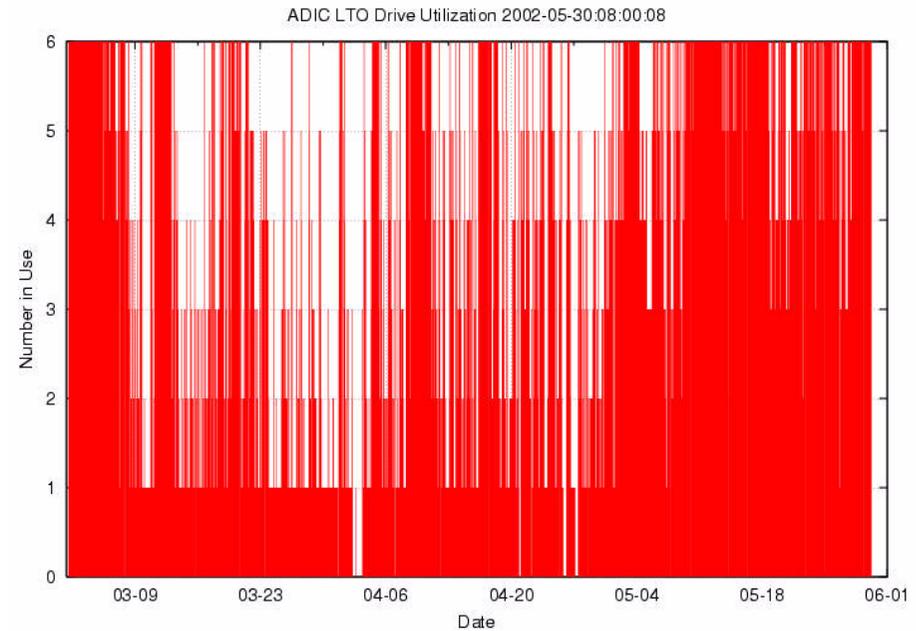
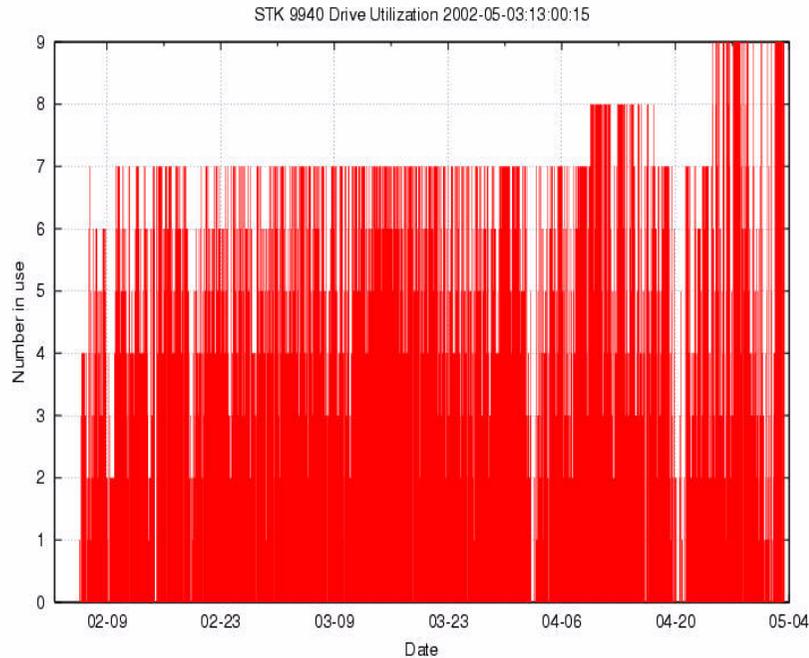
- DO has 1 STK silo with 9 9940 drives
- DO has an ADIC/AML2 with 6 LTO drives
  - ◆ Raw data going to STK,
  - ◆ MC to LTO
  - ◆ RECO output to LTO as a test

The 1.5 PB robotic storage needed at FNAL for Run II a can be accommodated by the two STK silos and the AML2 with the current generation of drives and media

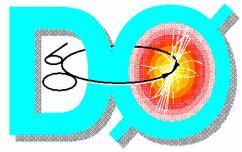
- Plan to purchase STK 9940B drives in 2003 and 2004
- Plan to purchase second STK silo in 2003
- For long term plan to purchase silos and drives through 2007



# Current Drive Usage

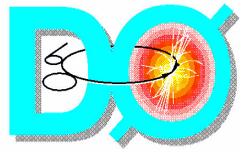


- **DO relies on low latency tape transfers**
  - ◆ **Online**
  - ◆ **Production**
  - ◆ **User access to DST and Raw data**
- **Purchase 15 drives in 2003 and 2004, assume STK**



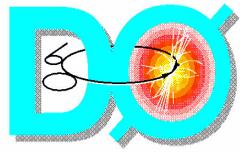
# Disk Estimates

- Plan for sufficient SAM cache and TMB storage on D0mino
  - ◆ All 2002 D0mino project disk additions supplied by the Institutions
  - ◆ Assume that model continues for project space
  - ◆ Supply additional 15 TB cache per year (\$150K estimated)
- Around 2005, put in an allocation for commodity disk—projection for network attached disk is around \$1/GB in 2005 to address replacement of D0mino's disk cache.



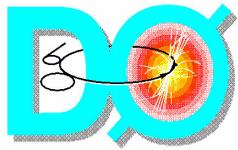
# Upgrading D0mino

- Replacing D0mino in kind with another large SMP is cost prohibitive
- Investigating solutions
  - ◆ Linux back end on D0mino
  - ◆ CLuED0/CLuB
  - ◆ Evaluate I/O performance and needs
  - ◆ Evaluate experience of other experiments
- D0 has flexibility thanks to SAM



# Backup facility

- Currently do not have a workable model for user back ups
- Two primary consumers
  - ◆ Project disk archive
  - ◆ User driven backups of small samples
- Clearly a need, but not clear how best to accomplish. What is the cost benefit?
- What are the requirements?
- Allocate \$100K to address this using existing robotics

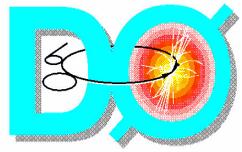


# Total Cost Estimate

<b>DØ Total Cost Estimate (assuming institution contributions)</b>						<b>Total</b>
	<b>2003</b>	<b>2004</b>	<b>2005</b>	<b>2006</b>	<b>2007</b>	<b>2003-2007</b>
<b>Infrastructure Analysis</b>	<b>\$400,000</b>	<b>\$310,000</b>	<b>\$570,000</b>	<b>\$645,000</b>	<b>\$295,000</b>	<b>\$2,220,000</b>
<b>Including Institution Contributions</b>	<b>\$1,152,500</b>	<b>\$865,000</b>	<b>\$1,152,500</b>	<b>\$1,025,000</b>	<b>\$680,000</b>	<b>\$4,875,000</b>
<b>FNAL CLuB Contribution</b>	<b>\$50,000</b>	<b>\$50,000</b>	<b>\$50,000</b>	<b>\$50,000</b>	<b>\$50,000</b>	<b>\$250,000</b>
<b>Reconstruction</b>	<b>\$225,000</b>	<b>\$325,000</b>	<b>\$575,000</b>	<b>\$150,000</b>	<b>\$200,000</b>	<b>\$1,475,000</b>
<b>Disk cache</b>	<b>\$150,000</b>	<b>\$150,000</b>	<b>\$200,000</b>	<b>\$200,000</b>	<b>\$200,000</b>	<b>\$850,000</b>
<b>Robotic storage</b>	<b>\$75,000</b>	<b>\$0</b>	<b>\$150,000</b>	<b>\$150,000</b>	<b>\$150,000</b>	<b>\$525,000</b>
<b>Tape drives</b>	<b>\$450,000</b>	<b>\$450,000</b>	<b>\$300,000</b>	<b>\$600,000</b>	<b>\$600,000</b>	<b>\$2,400,000</b>
<b>Backup facility</b>	<b>\$100,000</b>					
<b>Sum</b>	<b>\$2,602,500</b>	<b>\$2,150,000</b>	<b>\$2,997,500</b>	<b>\$2,820,000</b>	<b>\$2,175,000</b>	<b>\$12,745,000</b>

Overall cost estimate including institution contributions for analysis and data reprocessing. Does not include project disk or Monte Carlo

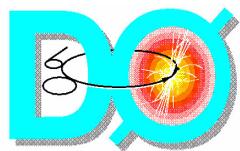
Amber Boehnlein, FNAL



# FNAL Guidance

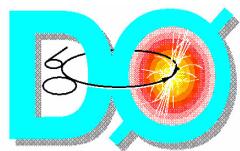
- The guidance to the experiments is \$2M/year
- The estimated needs exceed this
- Use the FNAL equipment budget to provide basic level of functionality
  - ◆ Database and other infrastructure
  - ◆ Reconstruction farm
  - ◆ Robotic storage and tape drives
  - ◆ Disk cache
  - ◆ Basic analysis computing to generate derived data sets

**Rely on the collaborating institutions to supply additional analysis computing and secondary reprocessing.**



# Institution Contributions

- Monte Carlo production takes place at regional centers now and in the future.
- Investigating compute intensive operations in addition to MC generation for regional centers, such as secondary reprocessing.
  - One-half of analysis computing to be on CLueD0, Club and other institution resources.
    - ◆ CLueD0 desktop cluster, administered by DO collaboration members, contributions by institutions
    - ◆ Institutions can provide project disk on D0mino or disk servers on CLuB
    - ◆ Anticipated that institutions will contribute to CLuBs, theCLueD back end or CAB [Central analysis backend]

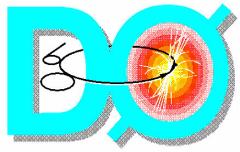


# FNAL Cost Estimate

<b>DØ Cost Estimate, FNAL contributions</b>							<b>Total</b>
	<b>2002</b>	<b>2003</b>	<b>2004</b>	<b>2005</b>	<b>2006</b>	<b>2007</b>	<b>2003-2007</b>
<b>Infrastructure Costs</b>	<b>\$400,000</b>	<b>\$400,000</b>	<b>\$310,000</b>	<b>\$570,000</b>	<b>\$645,000</b>	<b>\$295,000</b>	<b>\$2,220,000</b>
<b>Analysis CPU</b>	<b>\$400,000</b>	<b>\$635,000</b>	<b>\$462,500</b>	<b>\$412,500</b>	<b>\$610,000</b>	<b>\$597,500</b>	<b>\$2,717,500</b>
<b>FNAL CLuB Contribution</b>	<b>\$30,000</b>	<b>\$50,000</b>	<b>\$50,000</b>	<b>\$50,000</b>	<b>\$50,000</b>	<b>\$50,000</b>	<b>\$250,000</b>
<b>Reconstruction</b>	<b>\$400,000</b>	<b>\$162,500</b>	<b>\$230,000</b>	<b>\$395,000</b>	<b>\$150,000</b>	<b>\$200,000</b>	<b>\$1,137,500</b>
<b>Disk cache</b>	<b>\$0</b>	<b>\$150,000</b>	<b>\$150,000</b>	<b>\$200,000</b>	<b>\$200,000</b>	<b>\$200,000</b>	<b>\$850,000</b>
<b>Robotic storage</b>	<b>\$400,000</b>	<b>\$75,000</b>	<b>\$0</b>	<b>\$150,000</b>	<b>\$150,000</b>	<b>\$150,000</b>	<b>\$525,000</b>
<b>Tape drives</b>	<b>\$200,000</b>	<b>\$450,000</b>	<b>\$450,000</b>	<b>\$300,000</b>	<b>\$600,000</b>	<b>\$600,000</b>	<b>\$2,400,000</b>
<b>DØmino Memory</b>	<b>\$150,000</b>						
<b>Backup facility</b>		<b>\$100,000</b>					
<b>Sum</b>	<b>\$1,980,000</b>	<b>\$2,022,500</b>	<b>\$1,652,500</b>	<b>\$2,077,500</b>	<b>\$2,405,000</b>	<b>\$2,092,500</b>	<b>\$10,250,000</b>

Spending within guidance meets basic needs  
 DØmino maintenance is not included

Amber Boehnlein, FNAL



# Conclusions

- **The D0 computing model is successful**
  - ◆ **CD Joint projects**
  - ◆ **SAM, an integrated data handling system, enables flexibility in the allocation of resources and effective use disk cache and robotic storage.**
- **Scale up current solutions**
- **Add commodity computing for analysis to supplement SMP**
- **Use FNAL Computing budget to provide base for infrastructure, robotic storage, reconstruction and analysis**
- **Use collaborating institution resources for project disk, analysis computing, MC generation and secondary reconstruction.**