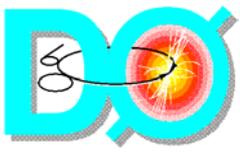


DO Requirements and Budget

Amber Boehnlein

Shank Review

Sept 14, 2004



Computing Contributions

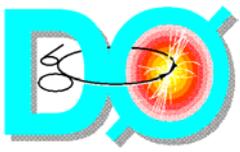
Use the FNAL equipment budget to provide very basic level of functionality

- ◆ Databases, networking and other infrastructure
- ◆ Primary Reconstruction
- ◆ Robotic storage and tape drives
- ◆ Disk cache and basic analysis computing
- ◆ Support for data access to enable offsite computing

Estimate costs based on experience or need for replacements

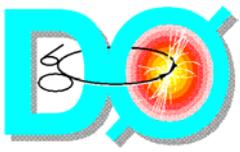
Remote Contributions

- ◆ Monte Carlo production takes place at remote centers
- ◆ Reprocessing (or primary processing)
- ◆ Analysis at home institutions
- ◆ Contributions at FNAL to project disk and to CLuED0
- ◆ Collaboration-wide analysis



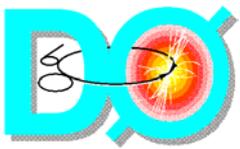
Virtual Center

- For the value basis, determine the cost of the full computing system at FNAL costs, purchased in the yearly currency
 - ◆ Disk and servers and CPU for FNAL analysis
 - ◆ Production activities such as MC generation, processing and reprocessing.
 - ◆ Mass storage, cache machines and drives to support extensive data export
- Assign fractional value for remote contributions
 - ◆ Merit based assignment of value
 - ◆ Assigning equipment purchase cost as value (“Babar Model”) doesn’t take into account life cycle of equipment nor system efficiency or use.
 - ◆ While shown as a predictor, most useful after the fact
- Not included as part of the value estimate yet
 - ◆ Wide Area Networking, Infrastructure, desktop computing, analysis



Tools for projections

- Over time, DO has developed a set of spreadsheets used for making projections and calculating value.
 - ◆ Original version—Chip Brock in 2002 for first Bird Review
 - ◆ Without them, couldn't have developed concepts for Virtual Center
- Typically use either data rate or total number of events collected as underlying quantity.
- Use past years to predict hardware futures
- Infrastructure usually budgeted on a replacement cycle, except for networking, which has a component that scales.
- Spreadsheets are available
 - ◆ Gives some indication of how the assumptions influence the outcome
 - ◆ Typically some differences between final documentation and this presentation—we often tweak the assumptions
 - ◆ Typically differences between budget request and actual purchase request
 - ◆ Typically differences in planned schedule and actual fact for reprocessing
 - ◆ Can evaluate the tools we have in place that guide decisions through the year
 - ◆ If you find any mistakes, let me know! Constant program of improving them



Accumulation Estimates

| data assumptions | | | | | 2005 | 2006 | 2007 | 2008 | 2009 |
|------------------|----------------|---------|------|--|------|------|------|------|------|
| rates | average event | 16 | Hz | | 16 | 30 | 30 | 30 | 30 |
| | weekly average | | | | 25 | 60 | 60 | 60 | 60 |
| | raw data rate | 5 | MB/s | | | | | | |
| | Geant MC rate | 1.65344 | Hz | | 1.60 | 3 | 4 | 3 | 4 |
| | PMCS MC rate | 0 | Hz | | 8 | 8 | 8 | 8 | 8 |

data samples (events)

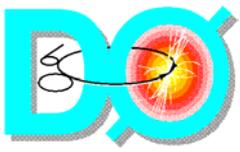
| | Current | 2005 | 2006 | 2007 | 2008 |
|------------------|----------|----------|----------|----------|----------|
| events collected | 1.00E+09 | 5.05E+08 | 9.46E+08 | 9.46E+08 | 9.46E+08 |
| total events | | 1.50E+09 | 2.45E+09 | 3.40E+09 | 4.34E+09 |

TAPE data accumulation (TB)

| | | | | | |
|---------------------|-----|-------|-------|-------|-------|
| Yearly storage (TB) | 757 | 525 | 697 | 763 | 830 |
| total storage (TB) | 757 | 1,282 | 1,979 | 2,742 | 3,572 |

disk data accumulation (TB)

| | | | | | |
|------------------------------------|----|-----|-----|-----|-----|
| Yearly storage (TB) | 45 | 51 | 96 | 96 | 96 |
| adjusted for format change in 2005 | 0 | 43 | 0 | 0 | 0 |
| Yearly adjusted storage (TB) | 45 | 95 | 96 | 96 | 96 |
| total storage (TB) | 45 | 140 | 236 | 332 | 428 |

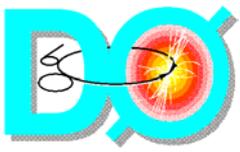


Mass storage

- DO uses STK powerhorn silos and an ADIC AML/2
- In 2004, evaluated and put LTO I drives into production—purchasing 10 more.
- Have 20 9940b drives
- During busy periods, as much as 2/3 of files consumed can be transferred from tape.
- In 2005, buy very few drives, LTOLs, upgrade the 5 year old mover nodes, buy drives for evaluation
- In early FY2006, deploy 500 Gb drives, and migrate data to recover slots
- Will likely activate the second arm in the AML/2
- Currently sharing the AML/2 with SDSS 4500 mounts/day



Mounts per day in AML/2, past 4 years.



Primary Production

Primary Reconstruction Cost Estimate

| Year | 2005 | 2006 | 2007 | 2008 |
|-------------------|-----------|-----------|-----------|-----------|
| Average Rate | 16 | 30 | 30 | 30 |
| efficiency | 80% | 80% | 80% | 70% |
| contingency | 20% | 20% | 20% | 20% |
| Reco time | 55 | 80 | 80 | 80 |
| Required CPU | 628320 | 1713600 | 1713600 | 1958400 |
| Existing system | 344947 | 436170 | 1248642 | 1219671 |
| Nodes to purchase | 92 | 293 | 75 | 85 |
| Node Cost | \$202,147 | \$644,279 | \$165,787 | \$186,248 |

Rate increase planned as part of the upgrade

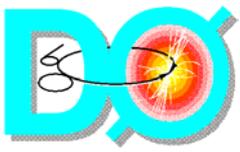
Calculation uses SpecInts

Using measured reco performance, luminosity profile, and preliminary

Indications of reco speed-up to guess at average time/event

2005: 16 Hz yearly average—25Hz weekly, how large a backlog is tolerable?

Amber Boehnlein, FNAL

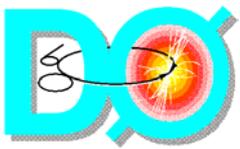


Estimated Analysis Costs

Fileservers: Cheap IDE for SAM cache, use more expensive infotrend disk for project space (where the users keep their results)

| | | 2004 | 2005 | 2006 | 2007 | 2008 |
|--------------------------|--|------------|------------|------------|------------|------------|
| Cache Data Volume (TB) | | 45 | 95 | 96 | 96 | 96 |
| contingency | | 40% | 100% | 100% | 120% | 150% |
| # to retire | | 0 | 0 | 0 | 18 | 24 |
| years volume (# servers) | | 18 | 24 | 18 | 10 | 8 |
| replacements | | 0 | 0 | 0 | 3 | 7 |
| #purchase | | 18 | 24 | 18 | 13 | 15 |
| #owned | | 18 | 42 | 60 | 55 | 46 |
| Cost | | \$ 288,000 | \$ 384,000 | \$ 288,000 | \$ 208,000 | \$ 240,000 |
| project disk volume (TB) | | 12 | 24 | 25 | 25 | 25 |
| Cost | | \$ 68,000 | \$ 85,000 | \$ 68,000 | \$ 85,000 | \$ 68,000 |

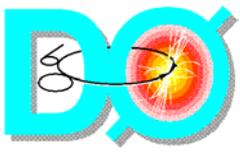
Typically have 3 year warranty on equipment, retirements taken into account. Do not have good model for cache space, size for disk resident samples, add factor. Assume need more cache as years go by as some hapless student(s) will be several versions behind



FNAL Analysis Cost

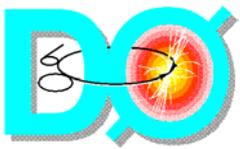
| Year | 2005 | 2006 | 2007 | 2008 |
|-------------------|-----------|-----------|-----------|-----------|
| Average Rate | 2.49E+03 | 4.05E+03 | 5.62E+03 | 7.18E+03 |
| efficiency | 70% | 70% | 70% | 70% |
| contingency | 20% | 20% | 20% | 20% |
| Analysis time | 0.5 | 0.5 | 0.5 | 0.5 |
| Required CPU | 1014992 | 1653220 | 2291449 | 2929677 |
| Existing system | 430248 | 592609 | 1151749 | 1540218 |
| Nodes to purchase | 190 | 243 | 185 | 159 |
| Cost | \$417,132 | \$534,926 | \$406,376 | \$350,311 |

Typically have 3 year warranty on equipment, retirements taken into account. 70% efficiency is current CPU/Walltime ratio, analysis time is measured, and routinely spin through 850 M events per week. 20% contingency for non-SAM work—root based analysis, usually.



Infrastructure Costs

- Usually stable—not this year!
- Networking ~200K
 - ◆ 10 G uplinks (postponed)
 - ◆ Networking to support new nodes
 - ◆ \$40K to finish DAB upgrade started last year
- Domino replacement parts \$60K (postponed)
 - ◆ Code builds and distribution
 - ◆ Interactive login cluster
 - ◆ NIS Slave
- DOWorld replacement \$15K
- Dobbin replacement (farm i/o) \$50K (postponed)
- DO2KA replacement-NetApp NFS server appliance + linux NIS server and disk \$100K (postponed)
- Replacement disk for database machine \$20-\$70K
- Enstore mover nodes \$50K



Cost Estimate-Sept 2004

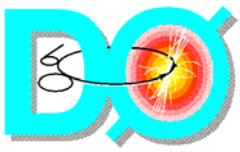
| | Purchased 2003 | Purchased 2004 | Purchase 2005 | Purchase 2006 | Purchase 2007 | Purchase 2008 |
|---------------------|-------------------|-------------------|------------------|------------------|------------------|------------------|
| FNAL Analysis CPU | \$470,000 | \$277,000 | \$417,132 | \$534,926 | \$406,376 | \$350,311 |
| FNAL Reconstruction | \$200,000 | \$370,000 | \$454,269 | \$717,742 | \$443,490 | \$362,546 |
| File Servers/disk | \$111,000 | \$350,000 | \$357,000 | \$356,000 | \$293,000 | \$276,000 |
| Mass Storage | \$280,000 | \$254,700 | \$40,000 | \$600,000 | \$300,000 | \$100,000 |
| Infrastructure | \$244,000 | \$140,000 | \$547,000 | \$200,000 | \$200,000 | \$200,000 |
| FNAL Total | \$1,305,000 | \$1,391,700 | \$1,815,402 | \$2,408,667 | \$1,642,867 | \$1,288,856 |

The guidance in 2002 was \$2M, cut to \$1.5 M. In 2003, \$1.5M, cut to \$1.35M (\$0.05M off the top, \$0.1M for Wideband tax.)

Added replacing mover nodes to infrastructure relative to document

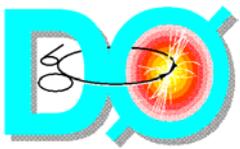
We did not add a “tax cost” to the price of the nodes, and probably should consider doing so. (\$535/node in FY2004)

Amber Boehnlein, FNAL



Reprocessing & MC

- Resources to reprocess needed will vary as a function of amount of data to process, how quickly it needs to be done, and speed of Reco
- Reprocessing is constrained by release cycle, analysis timescales and availability of remote resources
- Usually considered not to be a steady state event, but something that we plan for.
- MC production is steady state.
 - ◆ Try to estimate MC needs as a fraction of the data collection rate.
 - ◆ Using a fast parameterized MC in production has always been part of the plan.
 - ◆ Geant based simulation is being tuned and corrected to better model the data—most data generated to date will have to be regenerated
 - ◆ We do overlay min-bias events over the geant simulation, which adds a data handling component, beyond the simple store.



Value Estimate-Sept 2004

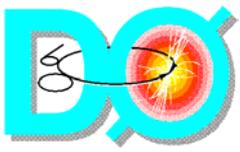
| | | Estimated Value | | | |
|-----------------------------|--|--------------------|--------------------|--------------------|--------------------|
| | | 2005 | 2006 | 2007 | 2008 |
| FNAL Analysis CPU | | \$724,054 | \$833,811 | \$817,048 | \$738,631 |
| FNAL Reconstruction | | \$820,089 | \$1,087,730 | \$773,295 | \$543,752 |
| File Servers/disk | | \$560,000 | \$688,000 | \$528,000 | \$560,000 |
| Mass Storage | | \$1,182,000 | \$1,201,000 | \$1,501,000 | \$1,501,000 |
| FNAL Infrastructure | | \$0 | \$0 | \$0 | \$0 |
| MC | | \$128,353 | \$170,152 | \$160,390 | \$85,056 |
| Reprocessing | | \$1,792,632 | \$3,317,845 | \$3,245,506 | \$2,940,120 |
| Virtual Center Total | | \$5,207,128 | \$7,298,539 | \$7,025,239 | \$6,368,560 |

This reflects the full value of doing all D0 computing in one year
In current year dollars—legacy systems are worth what it would cost
to replace them.

Refinements continue—Infrastructure currently valued at \$0

We no longer calculate yearly “cost” for remote centers—not a relevant
concept for many places.

Amber Boehnlein, FNAL



Conclusions

- **The DO computing model is successful**
 - We have developed tools to enable us to target computing spending at FNAL
 - We use metrics from SAM and system monitoring to provide estimators.
- **Use Virtual Center Concept to calculate the “value” that remote computing give the collaboration.**
- **DO continues to pursue a global vision for the best use of resources by moving towards interoperability with LCG and OSG**
- **DO computing remains effort limited—a few more dedicated people could make a huge difference.**
- **Short budgets, needs for continued construction projects and aging computing infrastructure is a serious cause for concern**