CDF Run 2 Computing Requirements

Kevin McFarland
Run 2 Computing Review
June 4, 2002

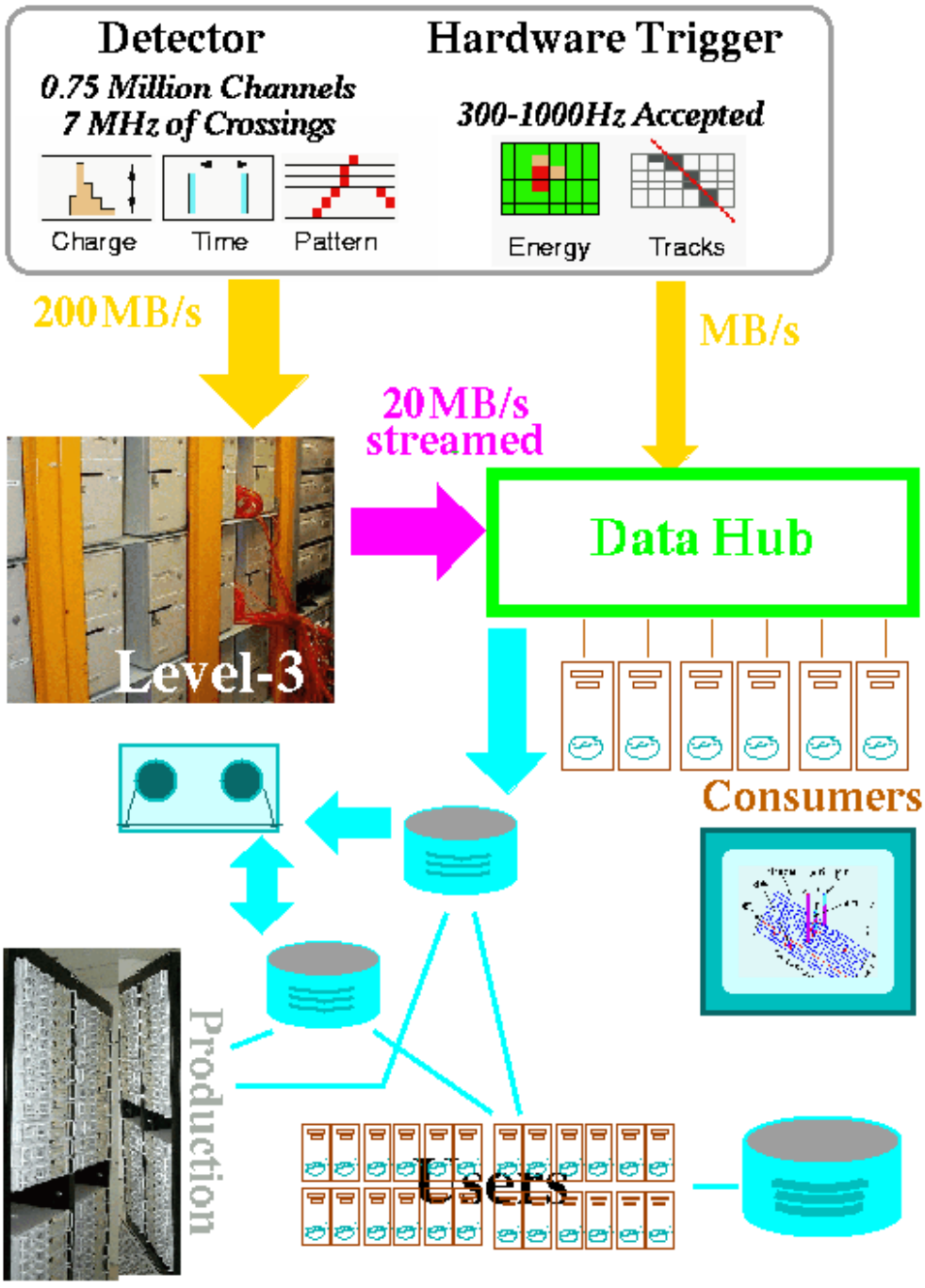
---

1. History and Inputs
2. Model Overview
3. Summary of Needs

Planning to Plan

- Original Long Range Planning effort: CDF 4100, March 1997
  - ↪ Scope of plan 2fb^{-1}, 2 years (1999-2001)
  - ↪ Things we know now:
    - ⋆ We will take longer to reach 2fb^{-1}
    - ⋆ CPU and Disk cheaper by time of 2fb^{-1}
    - ⋆ CPU needs are much greater than envisioned
      - · Si detector, event "I/O", exclusive reconstruction
- Current Planning Effort:
  CDF CAF and DH Reviews (Oct '01-Feb '02)
  - ↪ CAF review documents:
    - ⋆ CDF5743: CAF Benchmarking
    - ⋆ CDF5787: CDF Physics Needs Assessments
    - ⋆ CDF5802: CAF Report
  - ↪ Input from CDF "physics" (analysis) groups drives this
    - ⋆ Usage pattern for archive is different than originally planned in March '97
      - · Informed by early Run 2 analysis
    - ⋆ Large needs for permanent disk-resident data
    - ⋆ Significant non-detector data: ntuples, MC, etc.

CDF DAQ/Analysis Architecture



Computing Model in Broad Brush

Production, Analysis, Monte Carlo largely central
Large needs imply a network-based system

- Data archive network attached (Enstore, dCache)
  - ↪ Long-term directions are more "GRID-inspired" approaches, e.g. SAM
  - ↪ dCache is transport engine (central systems)
    (Jeff Tseng's talk, Rick St. Dennis' talk)
- Disk caches and static disk are network attached
  - ↪ Many few-TB scale RAID fileservers
- CPU provided through "PC pile" approach
  - ↪ Familiar from successful Level-3, Production farms
  - ↪ Small number of SMPs (Linux) as interactive nodes in initial stages
    (Frank Wuerthwein's talk)
- Heavy emphasis on central networking
  - ↪ We believe expansion of our existing infrastructure will be sufficient to cover this
    (Rob Harris' talk)
- Tight collaboration with FNAL-CD
  (Don Petravick's talk, Garzoglio/Pordes talks)

Luminosity Profile

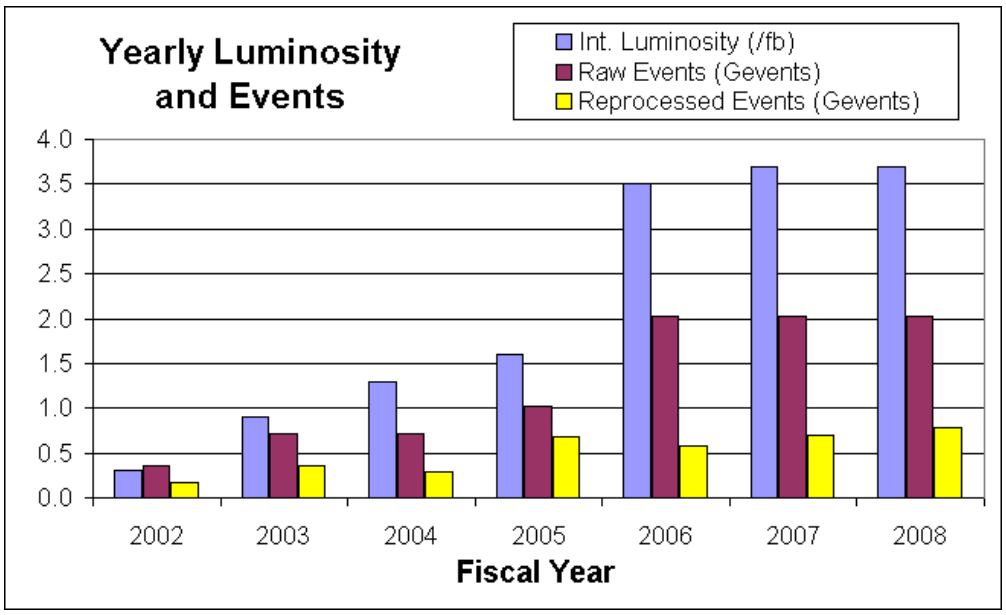
- FNAL Run 2 Luminosity profile(April 2002)

FY	Total (fb^{-1})	Yearly (fb^{-1})	Peak ($\times 10^{32}$)
2002	**0.3**	**0.3**	**0.3**
2003	**1.2**	**0.9**	**1.0**
2004	**2.5**	**1.3**	**1.4**
2005	**4.1**	**1.6**	**4.0**
2006	**7.6**	**3.5**	**4.0**
2007	**11.3**	**3.7**	**4.0**
2008	**15.0**	**3.7**	**4.0**

- From $\mathcal{L}$ to Raw Events
  - ↪ Physics trigger cross-section is assumed to be the Run "IIa" ($2\mathrm{fb}^{-1}$) design value
    - ⋆ Assumes high p_T (Higgs, top, SUSY) and low p_T (B) are on the menu throughout Run II

Luminosity Profile (cont'd)

- ↪ Peak raw data rate is (roughly) 20 MB/s peak, must increase to 70 MB/s in FY05
  - ⋆ *Note that full rate capability is used during FY02-03 commissioning, regardless of luminosity*
  - ⋆ Average rate is $1/3$ of peak rate (accelerator duty factor)
  - ⋆ In FY02, FY05 (commissioning, "Run IIb" shutdown), assume $1/6$ duty factor
  - ⋆ Assumed growth of event size with $\mathcal{L}$ is modest (FY05 transition to 132 ns bunch spacing)



  - ↪ Reprocessing occurs *rarely*
    - ⋆ Very different than, e.g., BaBar strategy
- Needs scale *roughly* with peak or integrated luminosity

How Do Requirements Scale?

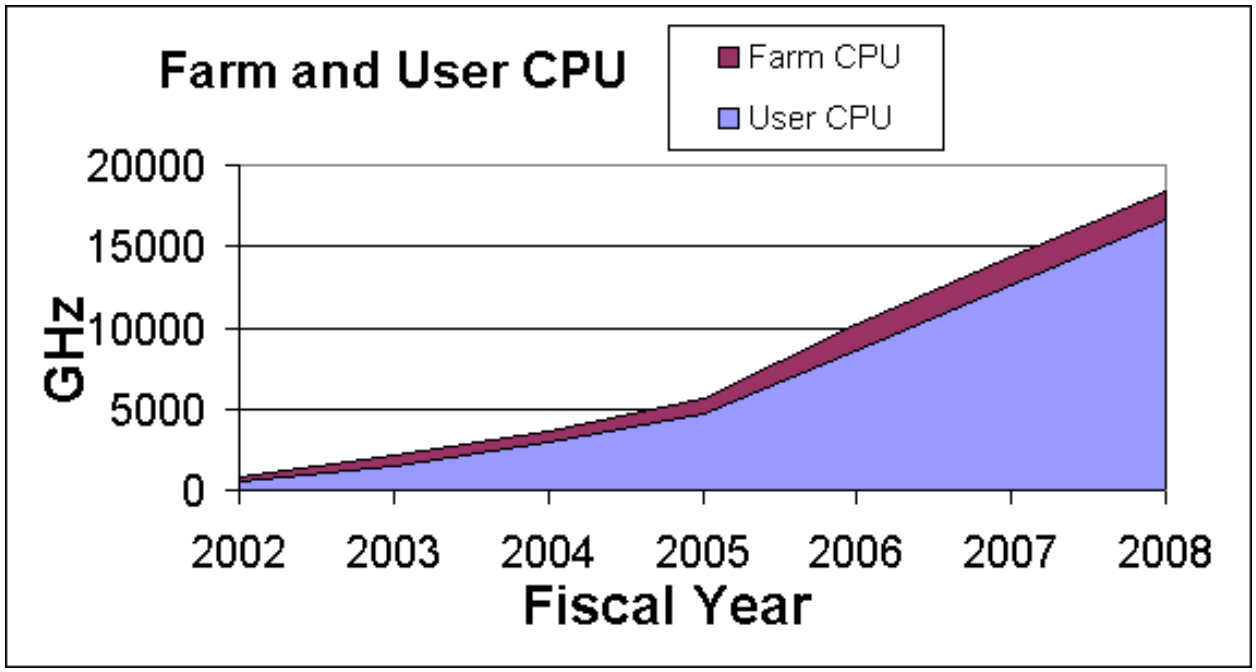
User CPU Bottom-up estimate, 1.1 THz/fb^{-1}

- Includes $1/3$ contingency, 10% for R&D needs
- Where does this come from?
  - ↪ $2\mathrm{fb}^{-1} \Rightarrow 2 \times 10^9$ events
  - ↪ Peak Requirement: "Light" analysis (e.g., ntuple creation) by 200 users of $\sigma = 5$ nb per day
    - ⋆ A $\sigma = 5$ nb sample is, e.g., loose $W \to e\nu$ samples 5×10^6/fb
    - ⋆ In other words, equivalent of full data set/day, roughly 25×10^3 events/sec
  - ↪ Analysis benchmarks: "TrackTest", "Stnmaker"
    - ⋆ Track test: $b\bar{b}$ MC sample, sift through track list looking for $D^+ \to K^+\pi^+\pi^-$, $B^+ \to \phi K^+$
    - ⋆ Stnmaker: $t\bar{t}$ MC sample, fill production quantities into extensive ntuple
    - ⋆ Each takes 0.1 GHz-ms per event

Requirements (cont'd)

Production CPU: Estimate driven by current experience (3–5 GHz-sec/event)

- RAW production stable; uncertainty reflects not knowing how code operates at high $\mathcal{L}$
- Monte Carlo bottom-up estimate: 1.2×10^6 events/day/fb^{-1}



Requirements (cont'd)

"Static" Disk needs are large
(bottom-up estimates of CDF Physics groups, CDF 5914)

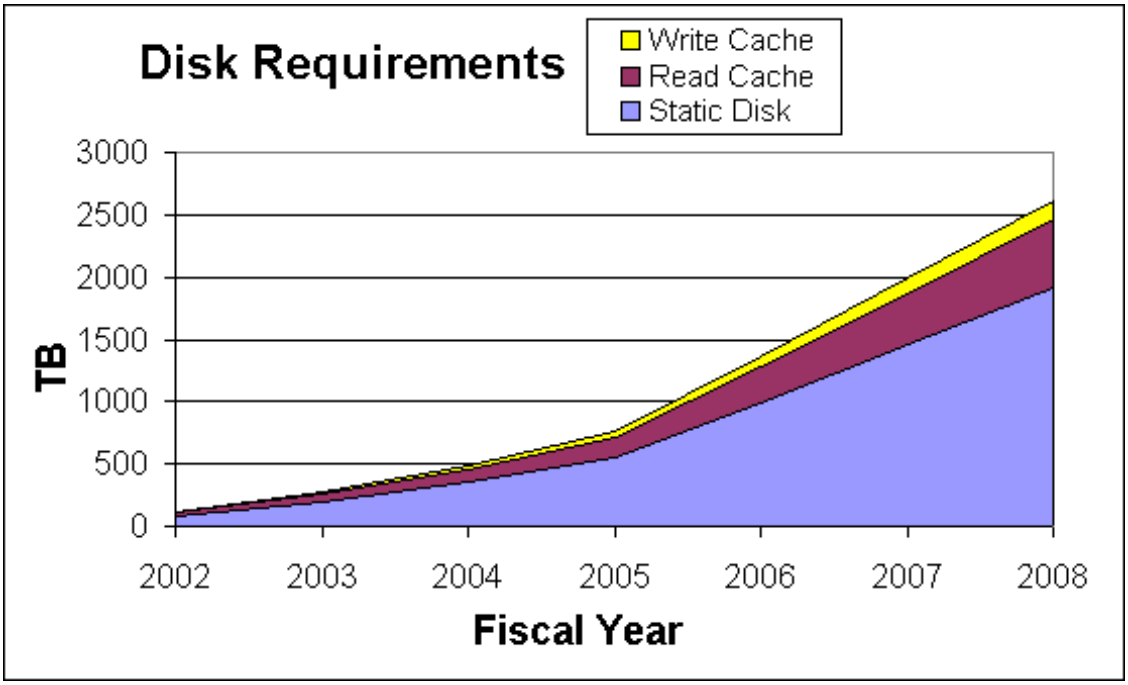
- Monte Carlo and "User" data, one static instance
- Produced output ("PADs", single instance) 70% static
  ↪ Assume 100 5nb datasets $\Rightarrow$ 150 TB/2fb^{-1}
- Usage is 60%–20%–20% for PADs–MC–User

Read/Write Disk Cache: smaller than static

- Rough model: $35/10$ TB/fb^{-1}, or $7\%/2\%$ of archive size

Disk I/O dominated by reading this static disk by User CPU, 1.0 MB/s/GHz

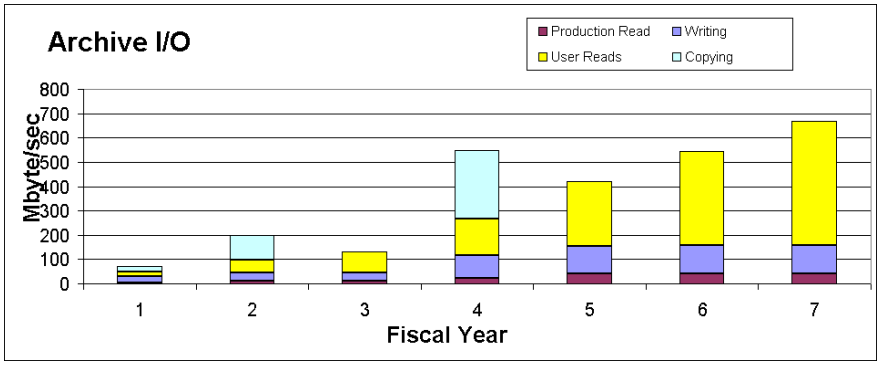
- Well below expected fileserver performance



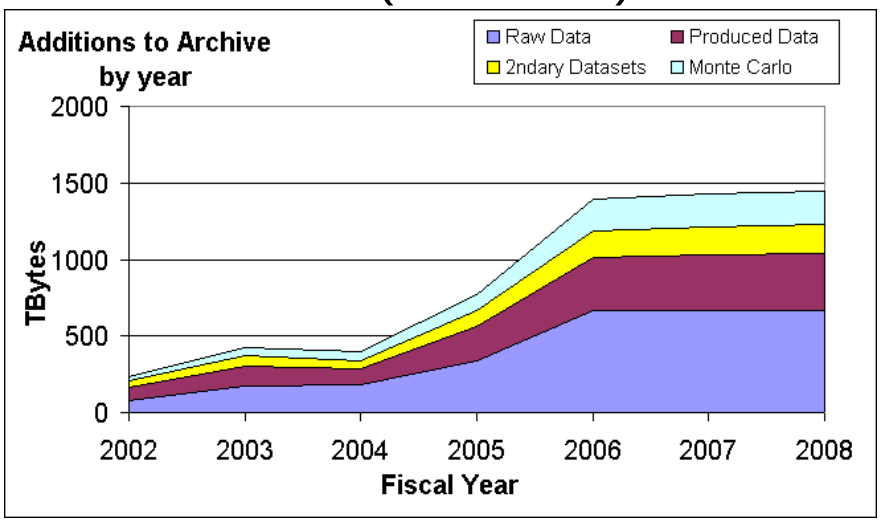
Requirements (cont'd)

Archive I/O includes raw, PADs, 2ndary, MC

- Writing: RAW and PADs dominate
  (very predictable with luminosity)
- Copying to new archives: important in FY03, FY05
  (see Rob Harris' technology evolution model)
- Read: User analysis

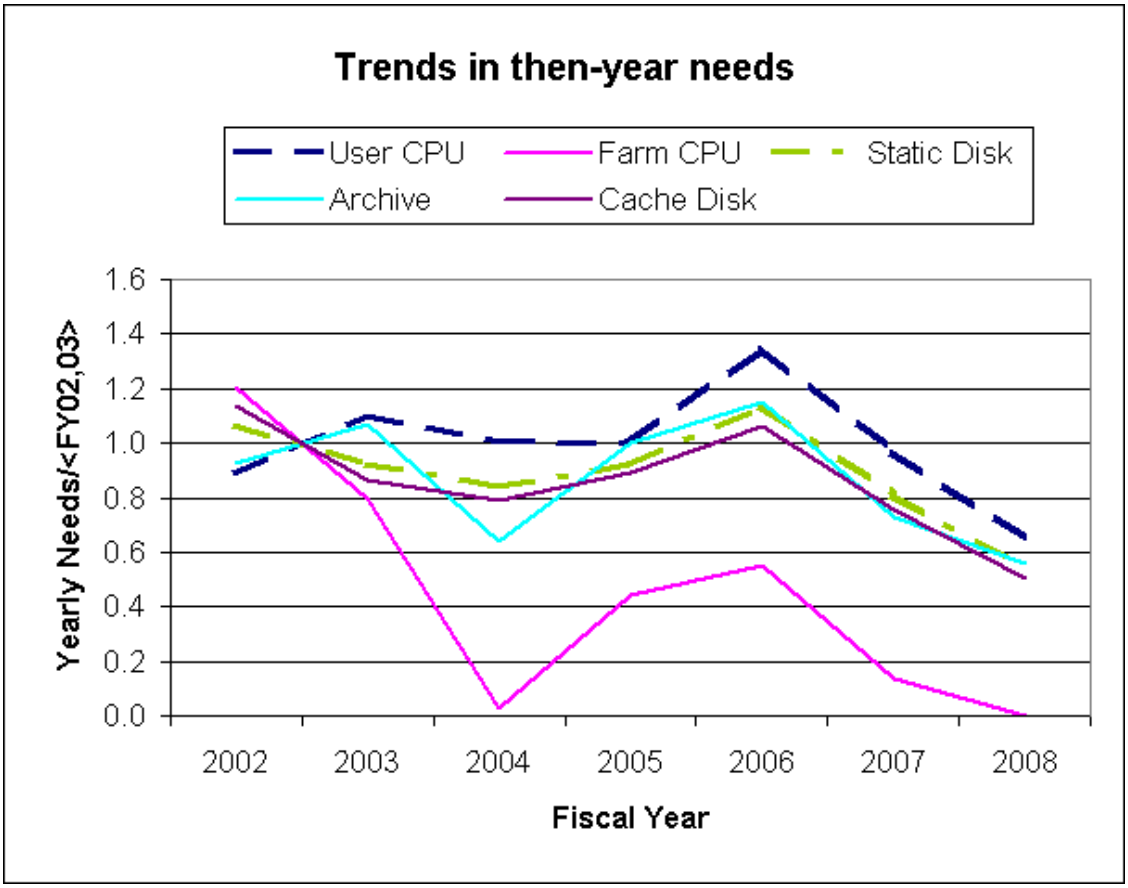


Archive Volume (as above)



Integrated Needs and Purchasing

- Archive media, User CPU, Static disk are most significant constant needs
- Three or four year obsolesence cycles assumed
- Archive capacity, I/O can be accomodated with discrete upgrades in FY03, FY05
- Long-term extrapolation works, assuming 18 month doubling of price/performance improvements
- Costs are ramping down for FY07, 08



Conclusions

- "Needs" have changed significantly from 1997 estimates
  - ↪ $2\text{fb}^{-1} \rightarrow 15\text{fb}^{-1}$, $\times 2$ instantaneous $\mathcal{L}$
  - ↪ 1999-2001 $\rightarrow$ 2002-2008
  - ↪ CPU and disk usage, primarily in user analysis
- Phenomenal advances in network bandwidth, storage and CPU capacity provide a solution
  - ↪ Moved from large SMP, direct-attached resource model
  - ↪ To network-based services from commodity components
- Total project resources (2008)
  - ↪ 20 THz of CPU
  - ↪ 2.5 PB of disk
  - ↪ 6 PB of tape, 1 GB/s to/from tape
- Subsequent talks
  - ↪ Technology (Tseng, Wuerthwein, St. Dennis)
  - ↪ Pricetag & Upgrade Model (Harris)