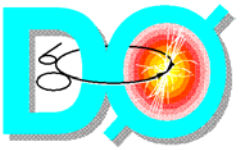


Run II Computing

Amber Boehnlein
FNAL/CD

For CDF and DO collaborations

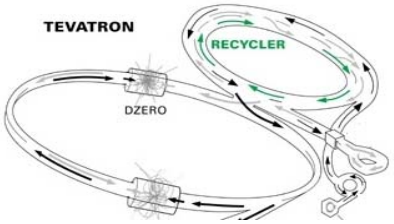
September 27, 2004

Computing Model

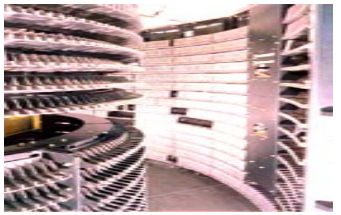
Remote Farms

Central Farms

Raw Data
RECO Data
RECO MC
User Data



Data Handling Services

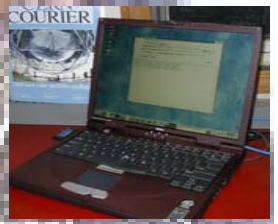


Central Storage

Remote Analysis Systems

Central Analysis Systems

User Desktops



Amber Boehnlein, FNAL

Global Collaboration

Remote Facilities

Canada: Toronto+, West Grid
US: San Diego, Rutgers, MIT
SAR* (UTA, Oklahoma +)
Michigan State
South America Sao Paulo*
Europe GridKA*, IN2P3*, INFN
Prague*, NIKHEF*,UK*
Asia: Japan, China, Korea, India*
Taiwan

CDF:
Institutional Clusters
DO:
CLuED0

Sequential Access Via Metadata & Job&Information Monitoring(*)

Storage
dCache (Desy/FNAL)
Enstore
Into STK or ADIC robots

Central Systems
CDF :
CDF Analysis Facility
Production Farm
DO:
Central Analysis <Backend>
Production Farm

Amber Boehnlein, FNAL

Vital Statistics

Vital Statistics	CDF	DO
Raw Data Size (kbytes/event)	205	250-300
Reconstructed Data Size (kbytes/event)	180	200 (20->60)
User formats	25-180	20-40
Reconstruction Time (Ghz-sec/event)	(5)10	50(120)
Monte Carlo Chain	fast	full Geant
user analysis times (Ghz-sec/event)	1 (3)	1
Peak Data Rate(Hz)	75(+)	50(+)
Persistent format	RootIO	DOom/dspack

Both collaborations continue to evaluate and evolve data formats in response to analysis needs and computing constraints

D0 computing has a strong production focus
CDF computing has a strong analysis focus

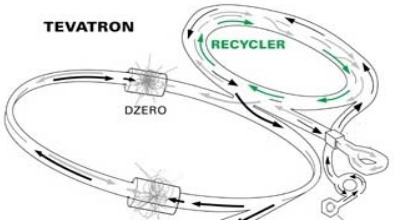
Amber Boehnlein, FNAL

Computing Model

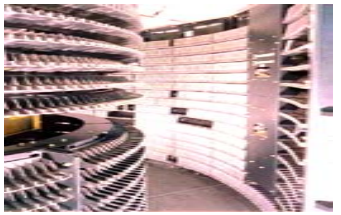
Remote Farms

Central Farms

Raw Data
RECO Data
RECO MC
User Data



Data Handling Services



Central
Robotic Storage

Remote Analysis Systems

Central Analysis Systems

User Desktops

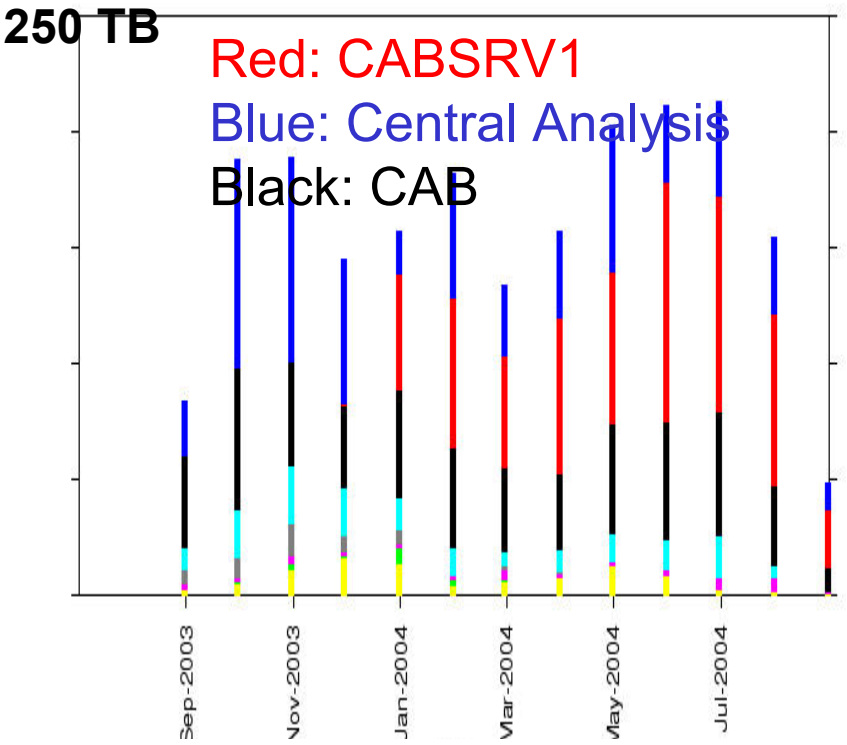
Amber Boehnlein, FNAL

SAM Data Handling

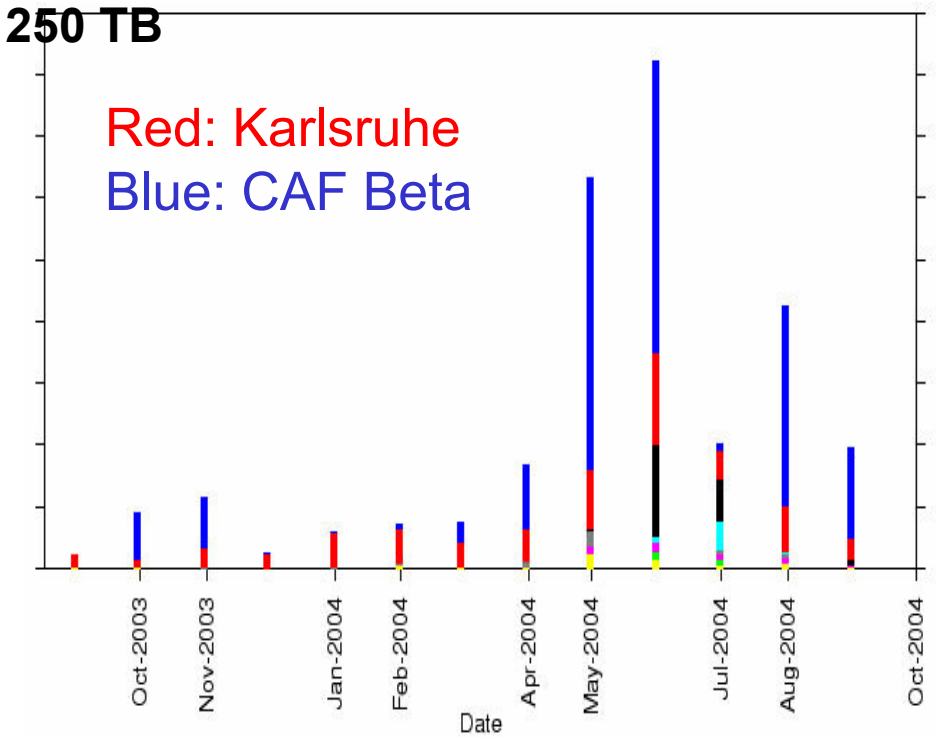
- Flagship CD-Tevatron Joint project—initial design work ~7 years ago, in production for DO for 4+ years, CDF remote for 1 year
- Provides transparent global access to the data
- Stable SAM operations allows for global support and additional development
- Services provided
  - Comprehensive meta-data to describe collider and Monte Carlo data.
  - Consistent user interface via command line and web
  - Local and wide area data transport
  - Caching layer
  - Batch adapter support (PBS, Condor, lsf, site-specific batch systems)
  - Optimization knobs in place
- Second Generation –Experience and new perspectives extend and improve functionality
  - Schema and DBserver updated in 2004
  - Introduction of SRM interface/dCache
  - Monitoring and Information Server prototype
    - move away from log file monitoring
    - Provide more real time monitoring

Amber Boehnlein, FNAL

SAM Performance



D0 ALL Stations GB/month



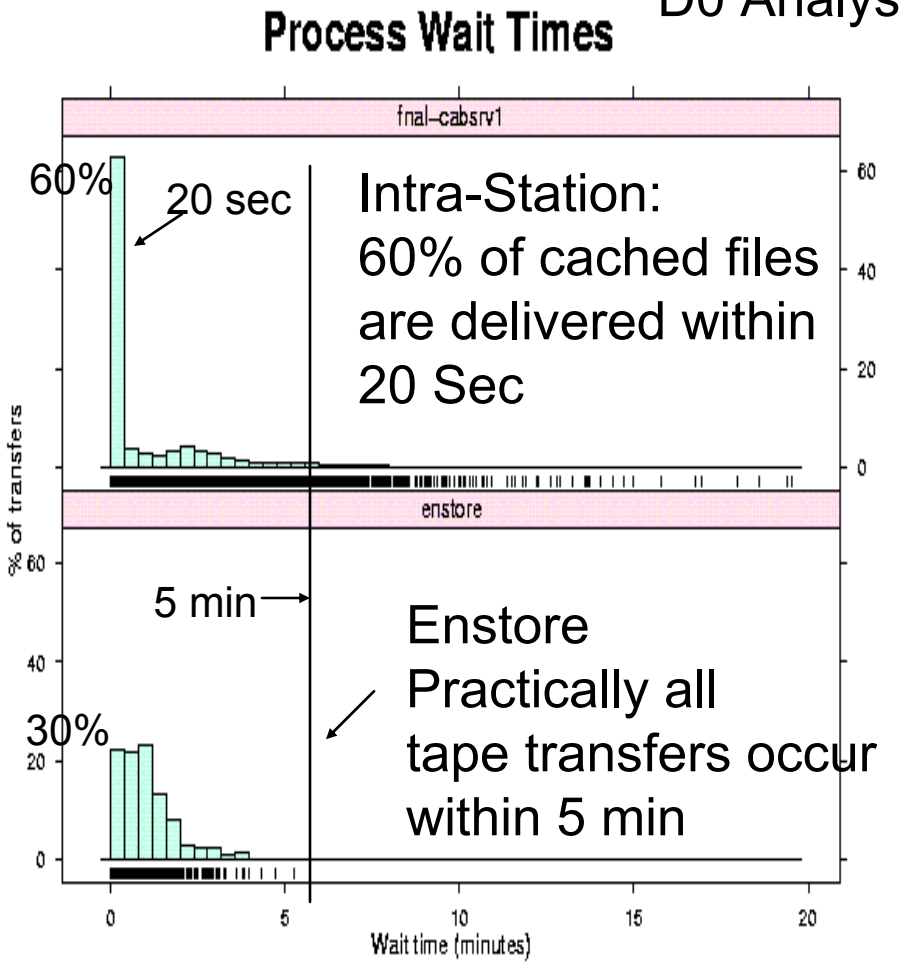
CDF-SAM GB/month

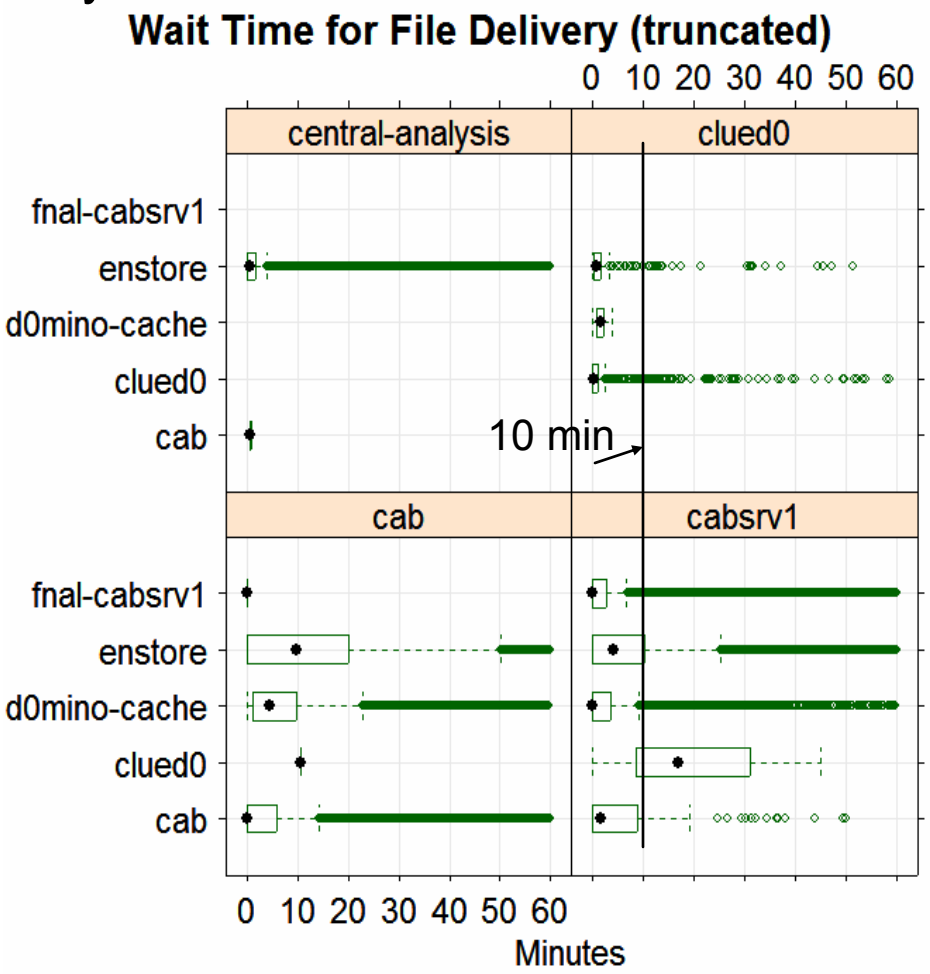
Oct 2003-Sept 2004
CDF: 1.5 PB 12B events
DO: 2.1 PB; 50B events

Amber Boehnlein, FNAL

DO SAM Performance

D0 Analysis systems



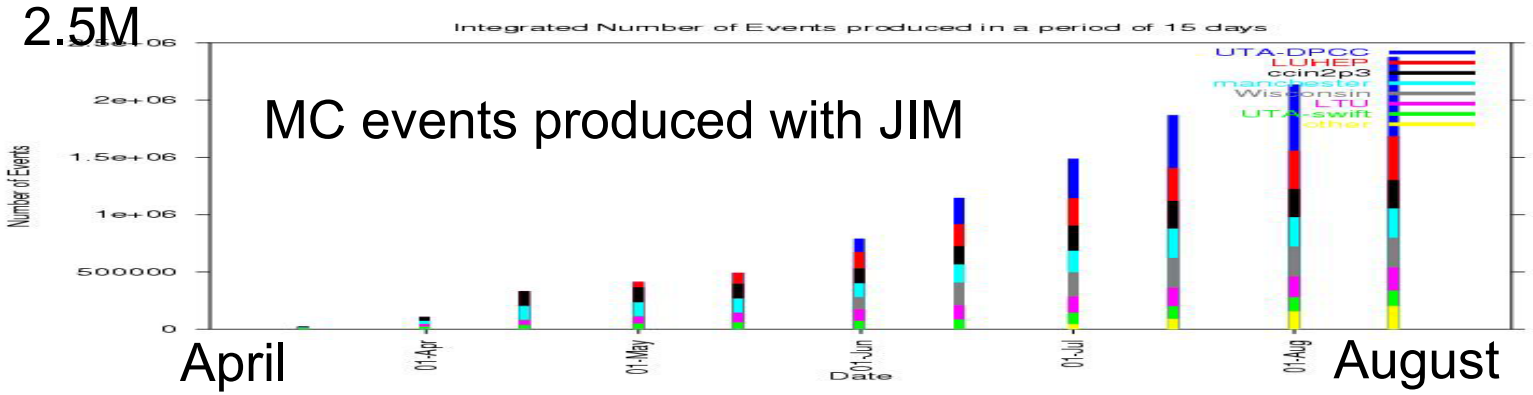


Before adding 20TB of Cache,2/3 transfers could be from tape.
Still robust!

Amber Boehnlein, FNAL

SAMGrid

- **SAMGrid project includes Job and Information Monitoring (JIM), grid job submission and execution package**
  - **JIM is in production for execution at 10 DO MC sites**
  - **Migration to VDT completed**
  - **Collaboration/discussions within the experiments on the interplay of LCG and Open Science Grid with SAMGrid efforts**
    - **Demonstration of use of sam_client on LCG site**
    - **University of Oklahoma runs Grid3 and JIM on a single gatekeeper**

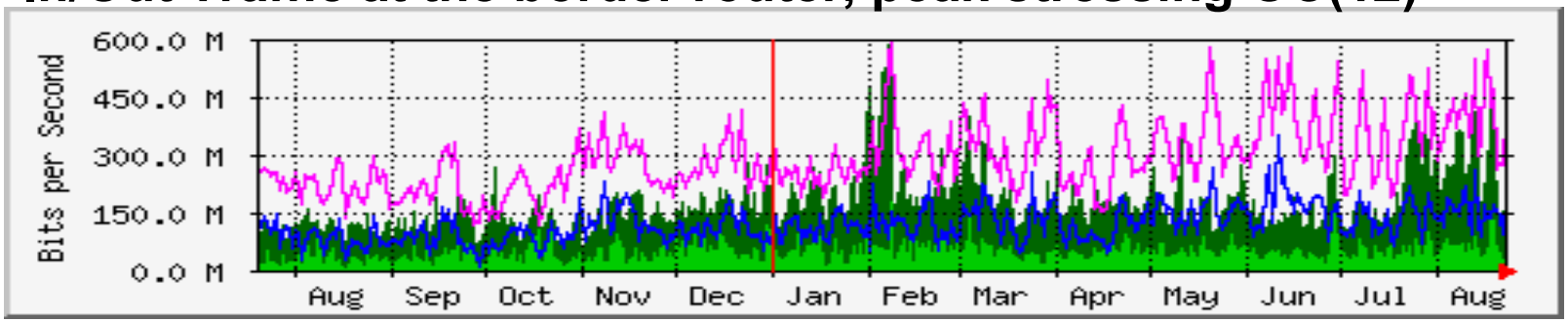


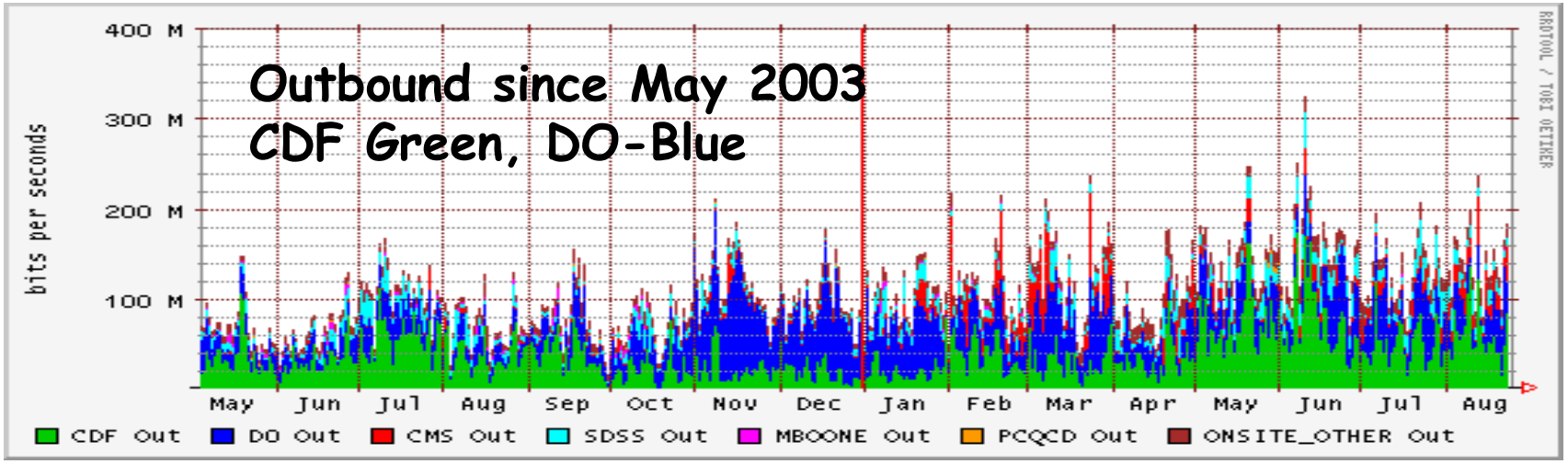
Amber Boehnlein, FNAL

Wide Area Networking

- OC(12) to ESNET, filling production link, anticipate upgrade
- R&D: Fiber link to Starlight

In/Out Traffic at the border router, peak stressing OC(12)

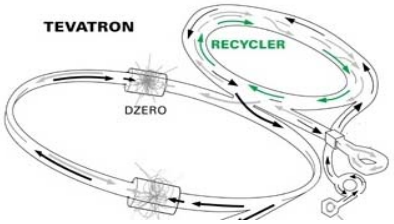




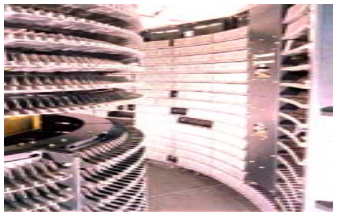
Amber Boehnlein, FNAL

Computing Model

Raw Data
RECO Data
RECO MC
User Data



Data Handling Services



Central Storage

Remote Analysis Systems

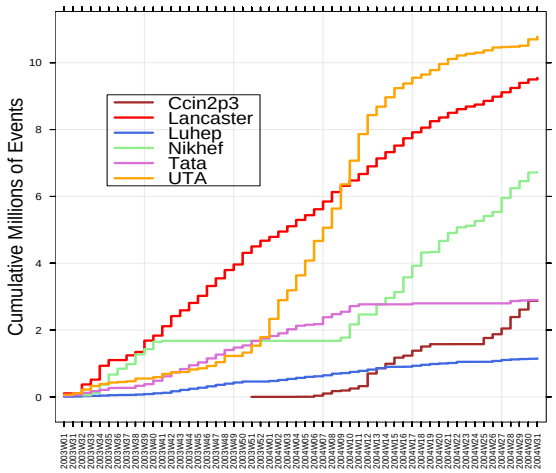
Central Analysis Systems

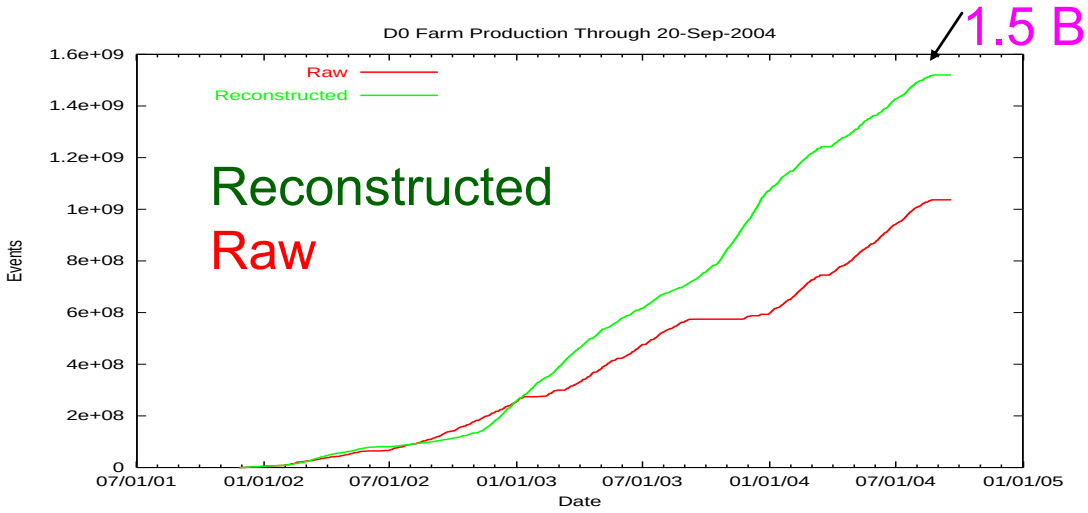
User Desktops

Amber Boehnlein, FNAL

DO Farm Production

- DØ Reconstruction Farm—18-20 M event/week capacity- operates at 80% efficiency—events processed within days of collection. 1.5 B events processed in Run II (1B events collected)
  - Successful remote re-reconstruction effort-100M events processed at IN2P3, NIKHEF, gridka, UK, and WestGrid (Canada)
- DØ Monte Carlo Farms—1 M event/week capacity-globally distributed resources. Running Full Geant, reconstruction and trigger simulation





MC production by Site

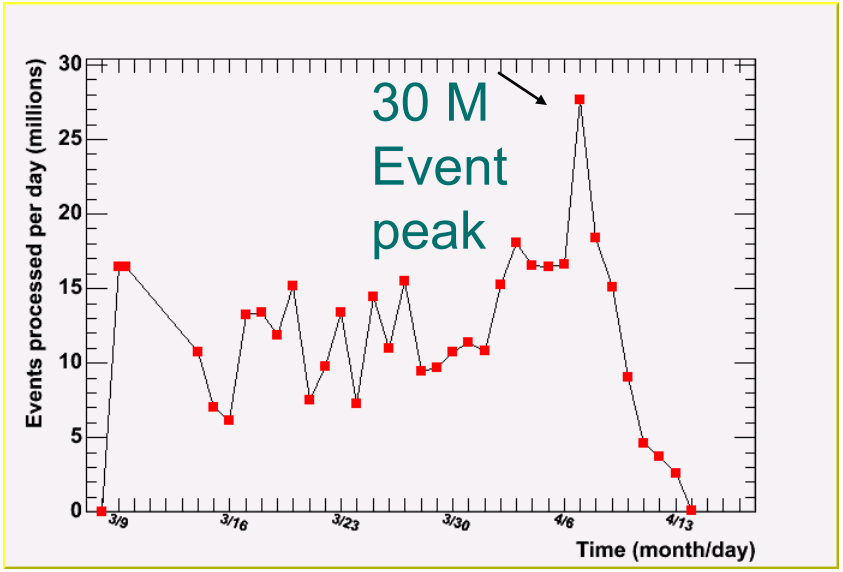
P14 Reprocessing Status as of 26-Apr-2004 (Remote sites only)

Processed Events	97619114	
Sites	fnal ccin2p3 gridka nikhef uk westgrid	

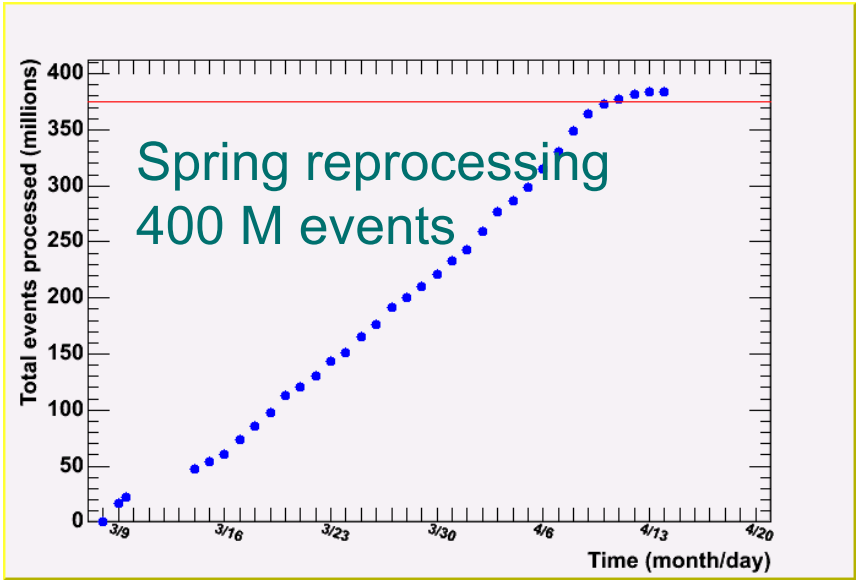
in, FNAL

CDF Farm Production

- CDF Reconstruction Farm sized to keep up with data collection and provide reprocessing capacity
  - Plan to integrate resources with CAF, move to using SAM for data handling by Dec 2004. Provides reprocessing buffer while maximizing availability of resources for CDF central analysis
  - "H" stream reprocessing to serve as prototype production with SAM



Events processed daily



Events processed daily

- CDF MC produced remotely at Toronto (60%), UK (30%), San Diego, Italy

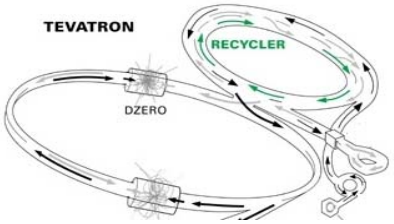
Amber Boehnlein, FNAL

Computing Model

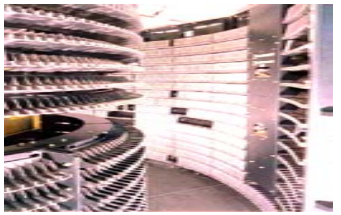
Remote Farms

Central Farms

Raw Data
RECO Data
RECO MC
User Data



Data Handling Services

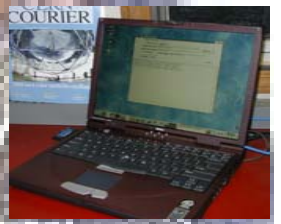


Central Storage

Remote Analysis Systems

Central Analysis Systems

User Desktops



Amber Boehnlein, FNAL

DCache

dCache—disk cache

Installed at FNAL: CDF, CMS, D0, LQCD, STK

Stabilized at highest load in May 2004

LAN interface – dcap

Linux, OSF1, IRIX, SunOS

WAN/Grid interface –

Ftp, GridFTP, Storage Resource Manager (SRM)

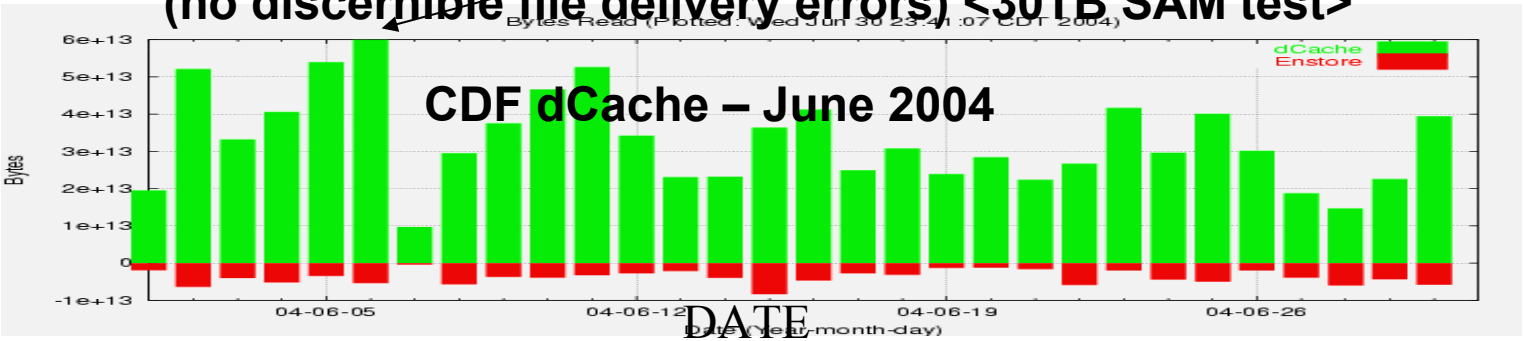
Direct dCache access at CDF

- 60 TB/day movement at peak
- Currently provides primary access to data on central systems

60 TB read by CDF clients on 06 June 2004

(no discernible file delivery errors) <30TB SAM test>

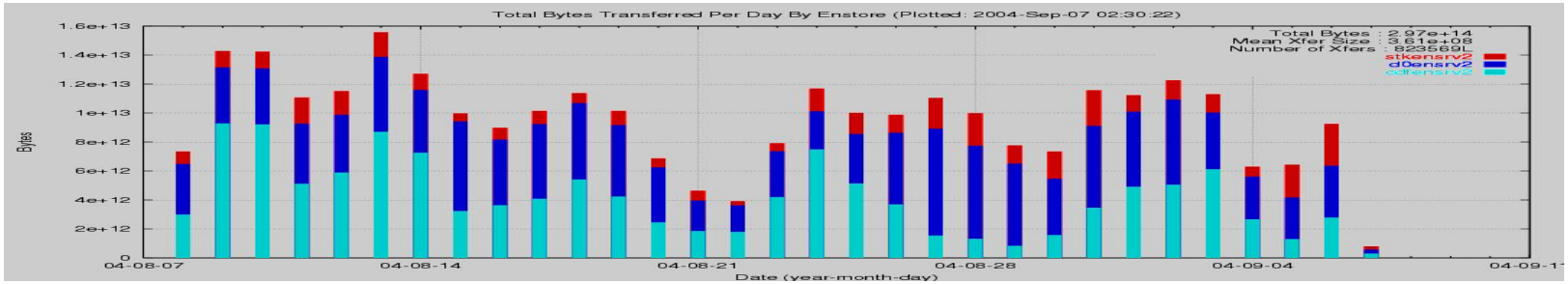
60 TB



nBytes Read Per Day

Central Robotics

20TB
At peak



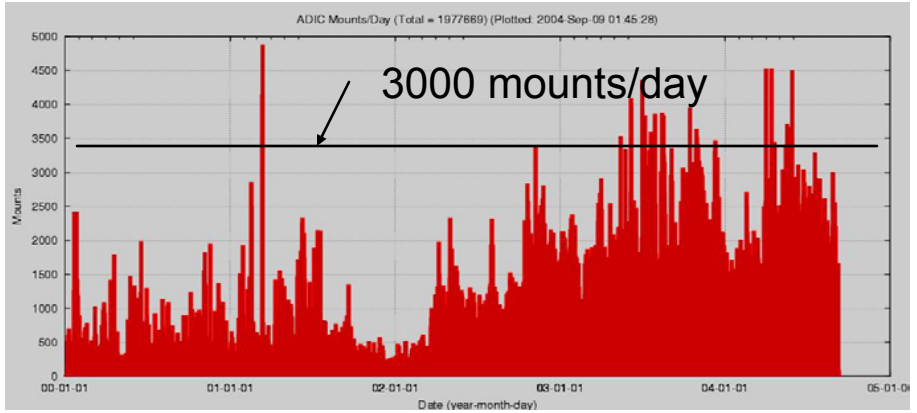
Daily Enstore traffic for CDF, DO, and other users

Data to tape, Sept 20, 2004
CDF 9940b ~ 1pb

DO 9940	565 TB
DO LTOI	175 TB
DO LTOII	70 TB
	800 TB Total

Diversity of robotics/drives maintains flexibility

Mounts/day on ADIC



Known data loss due to Robotics/Enstore for DO >10 GB
Somewhat larger for CDF due to a hardware problem

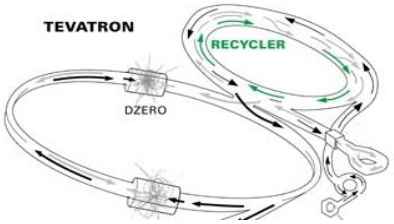
Amber Boehnlein, FNAL

Computing Model

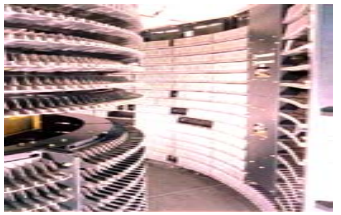
Remote Farms

Central Farms

Raw Data
RECO Data
RECO MC
User Data



Data Handling Services



Central Storage

Remote Analysis Systems

Central Analysis Systems

User Desktops

Amber Boehnlein, FNAL

Central Analysis

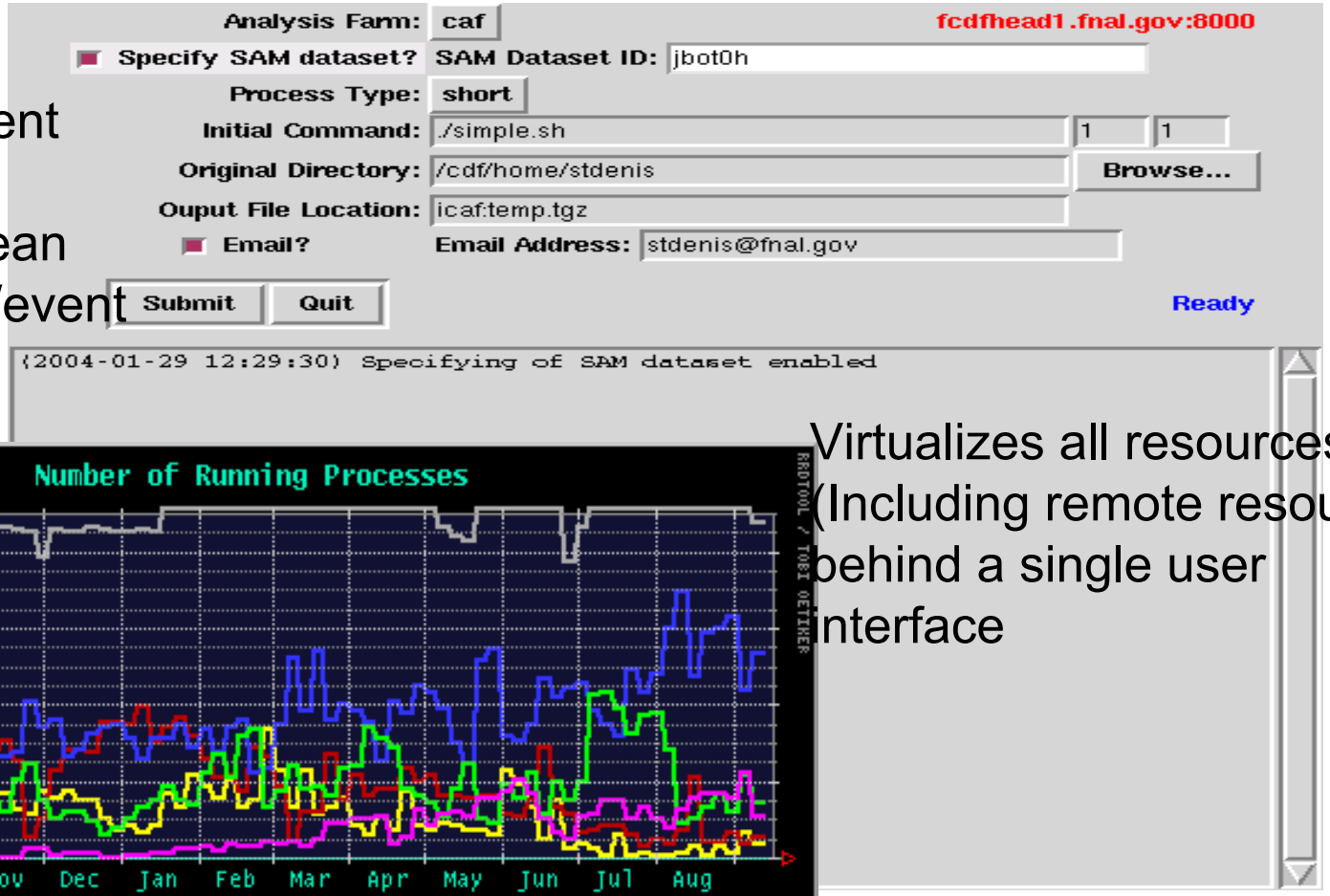
- Both experiments support peaks of 200 users
- Ntuple based analysis, some user MC generation
- DO supports post-processing "fixing" as a common activity (moving to production platform)
- B physics tends to be most cpu and event intensive—uses full framework/event size for CDF
- CDF Analysis Facility—Linux based system
  - 3.25 THZ with ~150 TB Cache and ~150 TB of group controlled space
- DO—Central Analysis Backend
  - ~2 THZ
  - Past year, short of cache, over-reliance on tape access.
  - Deployed 21 TB as SAM Cache on CABSRV1. 20 TB local disk cache and 70 TB user controlled space, primarily on CLuED0

CDF Central Analysis Facility

- **CAF is two farms—FBS and Condor—with single submission mechanism**

83% of jobs
Average
1Ghz*sec/event

17% have mean
Of 3Ghz*sec/event

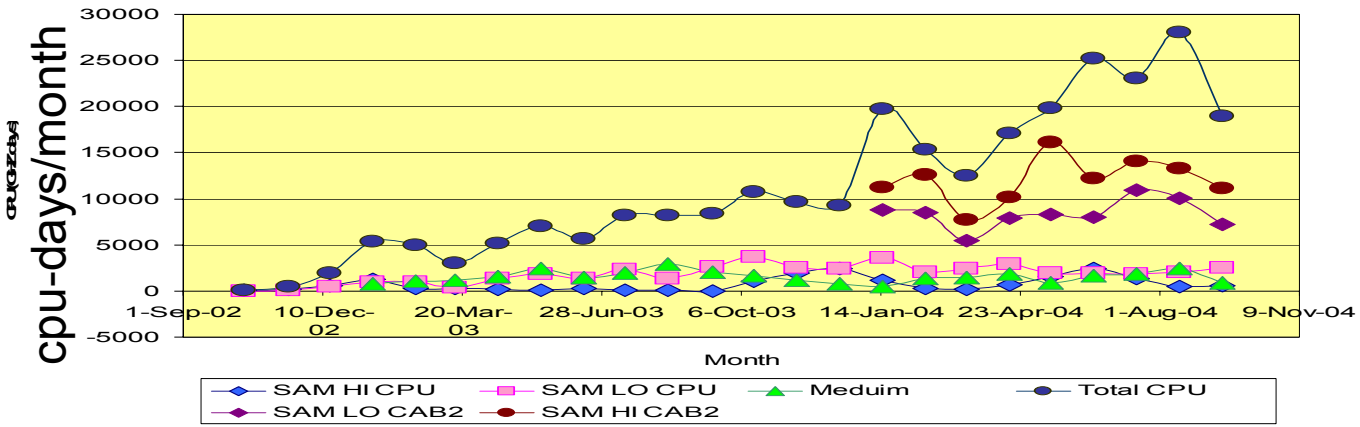


Virtualizes all resources (Including remote resources) behind a single user interface

Amber Boehnlein, FNAL

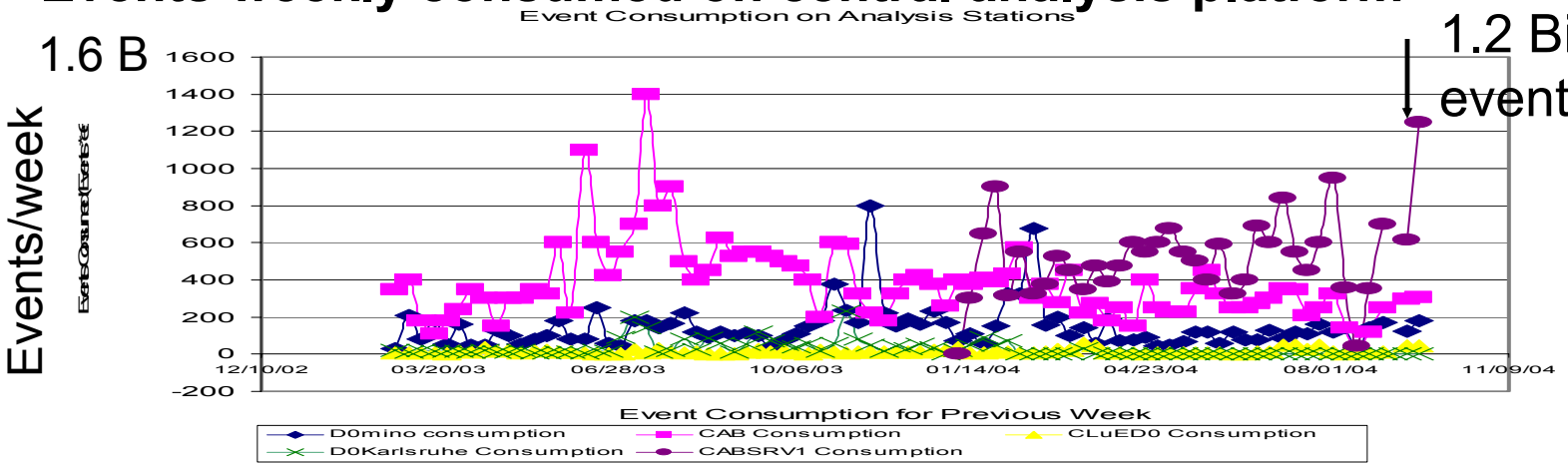
DO Central Analysis Systems

CAB usage in GHz*days



Typically spin through 1 billion events per week at 1 GHz*sec/event <plot normalized To slower cab nodes>

Events weekly consumed on central analysis platform



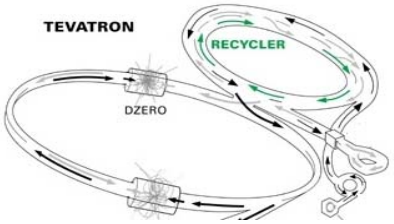
Amber Boehnlein, FNAL

Computing Model

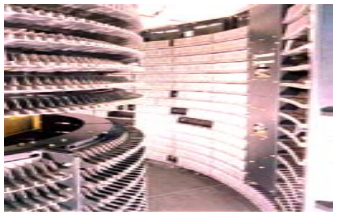
Remote Farms

Central Farms

Raw Data
RECO Data
RECO MC
User Data



Data Handling Services

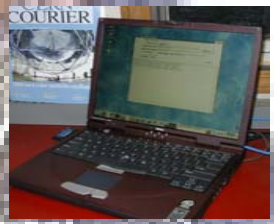


Central Storage

Remote Analysis Systems

Central Analysis Systems

User Desktops



Amber Boehnlein, FNAL

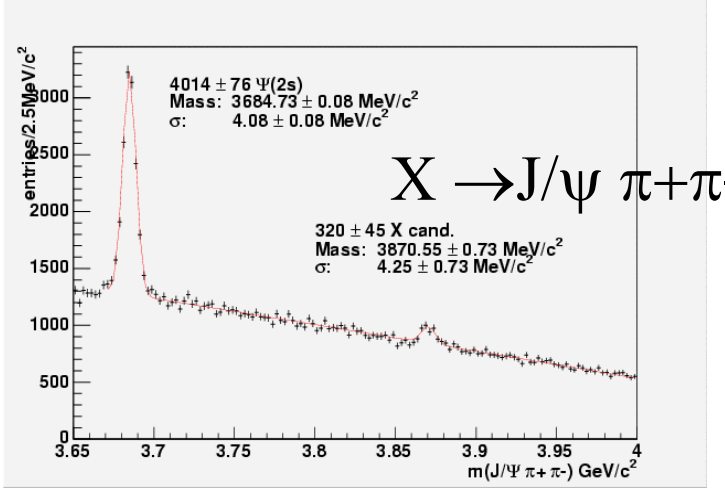
Remote Analysis

- Active SAM stations: 40 DØ (9 @ FNAL)
  26 CDF (2 @ FNAL)

35% of CDF analysis resources are available outside of FNAL

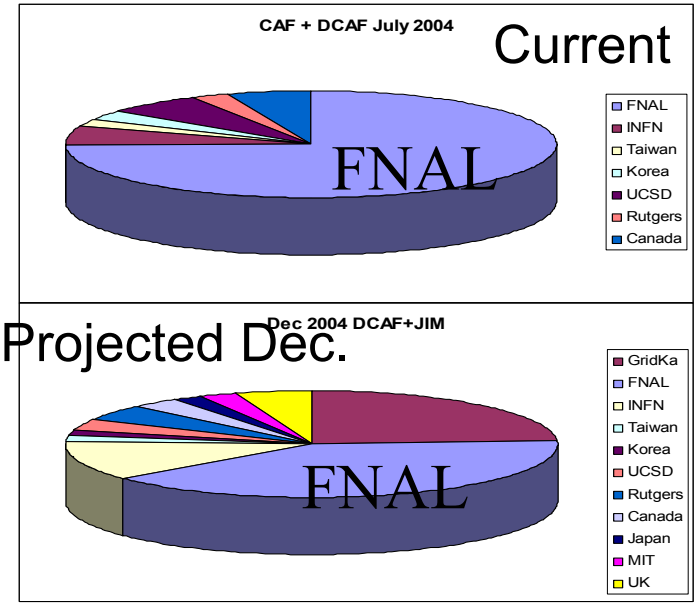
10% of DO analysis projects are run on the remote stations.

CDF--GridKa



60 processes/3000 files: jpmm0c
Analysis performed by a student

CDF—Offsite CPU



Amber Boehnlein, FNAL

Summary and Outlook

- Both experiments have a complete and operationally stable computing model
  - CDF pioneered commodity fileserver usage at 100 TB scale
  - DO pioneered grid data handling
  - Replacing legacy hardware, components, and software
- Both experiments refining computing systems and evaluating scaling issues
  - Planning process to estimate needs
  - DO costing out a virtual center to meet all needs
- CDF deploying SAM on the central systems
- DO defining interfaces, isolate experiment specific exes, deployed OS compatibility product
- DO using Grid tools for MC job submission, extending to data processing this fall
- CDF uses common user interface over virtualized distributed resources
- Experience leading towards second generation in many applications

Both experiments are finding common computing ground and moving towards global and grid computing and continue to provide excellent computing to a diverse user community

Learn More!

By necessity, this talk focused on some areas, ignored many, and did justice to none. For more information, please see posters and talks relating to Run II computing

- Databases <Completely neglected in this talk> and FroNtier [204] [205]
- Networking [359] [369]
- Enstore and dCache [107] [190] [464] [471]
- SAMGrid—performance, monitoring, Metadata
  - [335] [451] [455][462][400] [38][293][481]
- CDF Posters –Monitoring the CAF [390] [484]
- DO Posters

  reprocessing [362], Virtual Center [372]

  Interfacing to other Grids [55] [58]