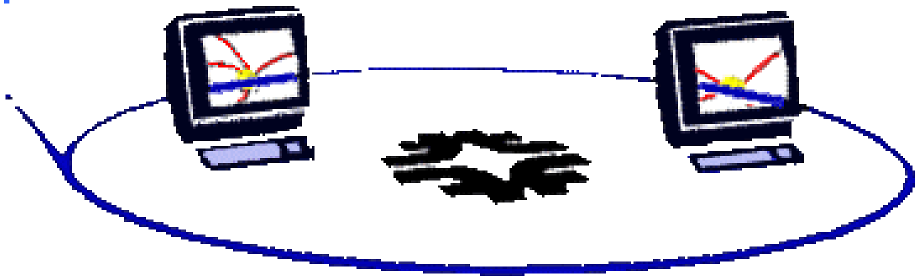
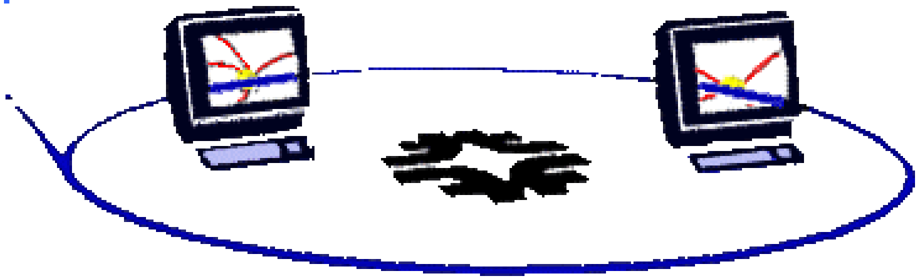


SITE REPORT: FERMI

LISA GIACCHETTI

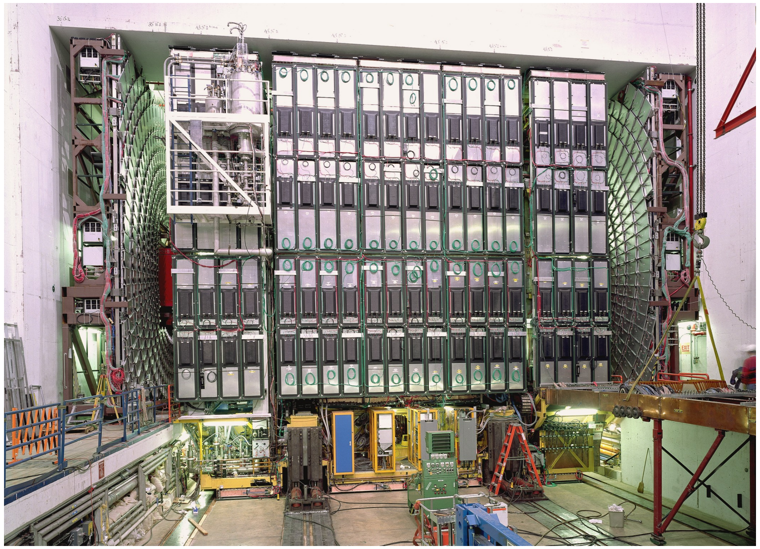
Operating Systems Support



GENERALComputing

- OS Stats
  - 21 Tru64
    - Last Tru64 system in FNALU to be decommissioned Dec ‘02
  - 271 Sun
  - 2159 Linux
  - 84 IRIX
  - 6531 Windows
  - FNALD (CDF Vax cluster) powered off March 2002

Status of Computing for the DØ Experiment at Fermilab



- Total detector data expected ~ 150 TB / yr for ~ 6 yrs
- Total data store, incl simulated & derived data ~ 0.5 PB /yr for 6 yrs
- Total user community > 500 scientists in > 60 institutions on 4 continents

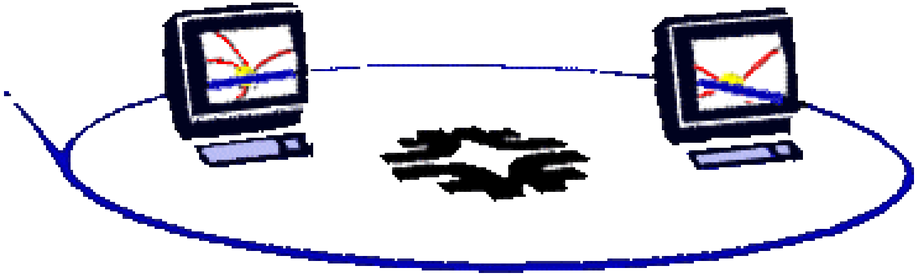
- Logging data at 10 MB/sec
- Moving data around at 90-150 MB/sec
- Have stored 100 TB data in one year
- Creating simulated data at up to 3 MB/sec
- Central Computing Facility
  - 192-processor SGI O2000 w/ 30 TB disk
  - ~6000 SpecINT95 Linux production farm
  - Linux cluster for building code releases
  - Linux analysis server w/ 2 TB disk
  - Sun 4500 & Sun 3500 database servers
-

- Desktop computing at Fermilab
  ~ 200 Linux machines
  ~ 200 WinNT / Win2000 machines
- Remote computing
  ~ 600 Linux machines & 1 192-proc O2000 at 8 remote sites, including 4 in Europe
  currently used for created simulated data, can be used for data analysis

  and reconstruction as well

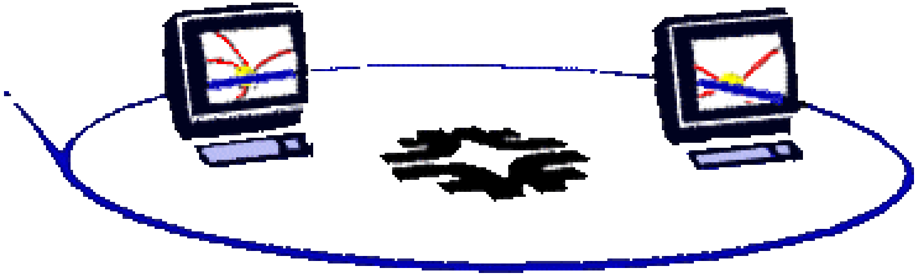
Status of Software for the DØ Experiment

- Using C++ for reconstruction and analysis code
- Reconstruction code working at approximately design data rate
- Release system based on cvs, supported in house
- Weekly test releases, 3-month schedule for production releases
- User analysis based on ROOT, supported in house and externally
- Supporting all code on IRIX and Linux (RH 7.1 w/ some systems still at RH 6.2), some online code on Tru64
- Mass storage handled by ENSTORE software, created in house
  - ~ 300 TB STK robot, 9 STK 9940 drives
  - ~ 750 TB ADIC AML/2 robot, 6 IBM LTO drives
- Data handling system is SAM software, created in house
  - Its bookkeeping is based on ORACLE database to track:
    - file locations and status
    - event locations in files
    - processing information
    - user-defined datasets
  - SAM also manages computing resources: disk (directly), CPU (through batch system interface), tape drives (through ENSTORE)
  - Working with GRID projects (Particle Physics Data Grid) to expand usefulness of remote resources
- Calibration and other databases also in ORACLE



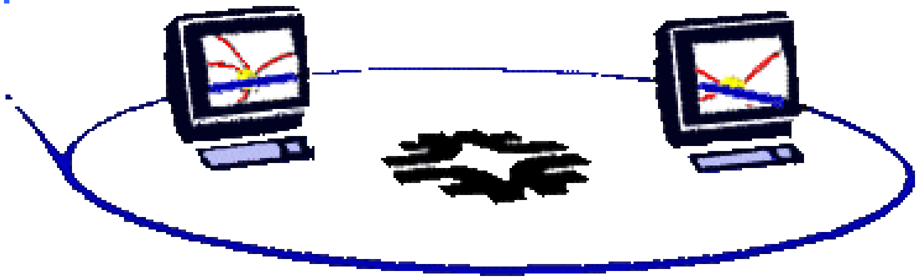
Experiment Updates: CDF

- System Count
  - 128 CPU SGI with 30TB of disk
  - 8-way Linux box for code builds
  - Couple of Sun database servers
  - 4 CPU SGI for data handling
  - ~280 Linux desktops; ~40 SGI desktops
  - Miscellaneous servers
- Data Statistics
  - Take raw data at 20MB/s
  - 100TB of data on tape so far (half raw data, half reconstructed)
  - Will soon use reconstruction farms for Monte Carlo production
    - Means more Monte Carlo data on robot as well
  - Expect 250 TB or raw data/year and 500TB of total data store per year
- More than 500 users from 55 institutions



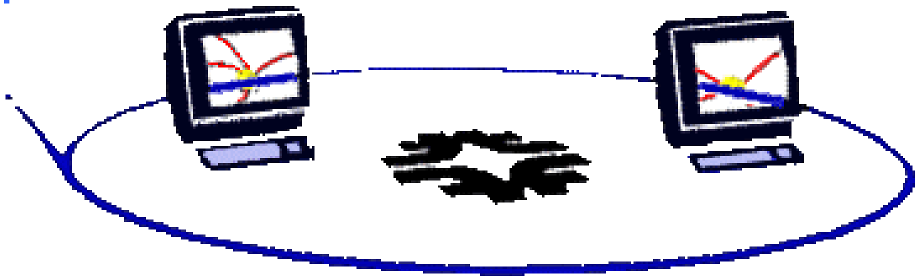
Mass Storage: Tape and Data movement Technology

- http://hppc.fnal.gov/enstore/
- ~200TB in Storage systems
- 8mm tapes have been deprecated
  - IBM LTO under light production in ADIC libraries
- 4 STK silos, 26 9940 drives, ~1.2PB capacity
- 6 -> 12 IBM LTO drives in ADIC AML/2
- 2 AML/2 libraries, 7 Quadra Towers ~3.5PB capacity
- 6 Exabyte M2 and 2 Exabyte M1 for reading old 8mm
  - NO 8mm taped being written in storage system context
- CDF nearly done copying data from AIT-2 to 9940
- All data movement continues via standard MTU ethernet
- All data storage hardware is Linux
  - Lancewoods moving to “succor of STL” mainboard



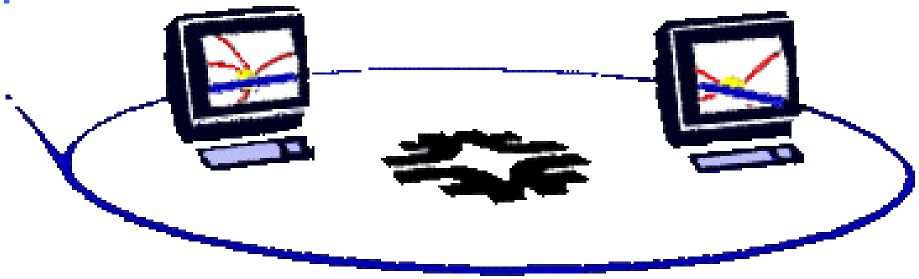
Mass Storage: User Community

- CDF joins D0 as making Enstore its main direction
- Enstore now writes all tapes
  - Don P. says, “responsible for all data loss at FNAL”
- Three instances of Enstore system at FNAL
  - D0en
  - STKen
  - CDFen
- Some work with Lancaster in U.K.
- Data movement statistics
  - http://www-d0en.fnal.gov/enstore/enplot_total_bpd.jgp



Mass Storage: Directions -> dCache

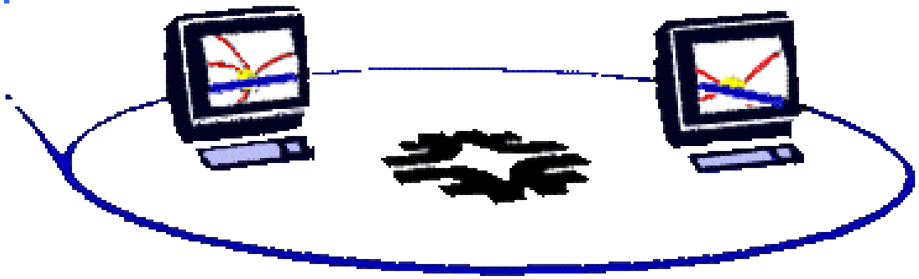
- dCache: disk caching system written collaboratively with DESY
- Enstore's WAN interface
  - Kerberos FTP door for writing, reading
  - Weakly authenticated FTP for reading if experiment wants it
  - Grid FTP server under development
    - Comments about Grid FTP made at the GGf
    - We'd like changes with writes under mode E
  - In production, used by Minos
- Under investigation by CDF, CMS for their data handling system
  - Interested in dccp protocol which provided file system like access to storage system
  - Root interface
  - CDF are planning to expand to have ~15Tb manage by dCache
  - FNAL security compliance under investigation
- SRM door for Grid development



Scientific Computing:

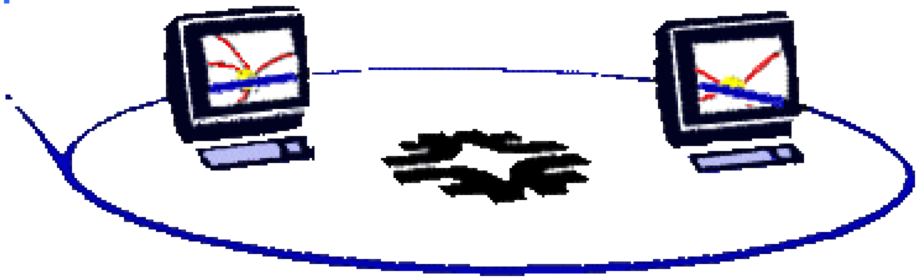
Reconstruction FARMS

- 3 farms in various stages of production
  - CDF, D0, FT
  - Migrating to FRHL 7.1 with 2.4.9-31 kernel on worker nodes
  - Eval in progress for new worker node hardware
  - All farms utilize FBSNG as the batch system
    http://www-isd.fnal.gov/fcs
    - CMS farms are also using fbsng
  - fcp is used to limit need for NFS access to disk
  - dfarm is uses as a tool to manage temporary storage on work node disks
- Fixed Target farms
  - 50 x 500MHz, 40 x 1GHz Linux nodes, 2 SGI IO nodes in production
  - 16 x 1.2GHz more Linux in burn-in process
  - 10 groups currently utilizing this farm
    - Ktev, E871, E781, SDSS, BTeV, Minos, …
  - Most groups now using Enstore for data IO



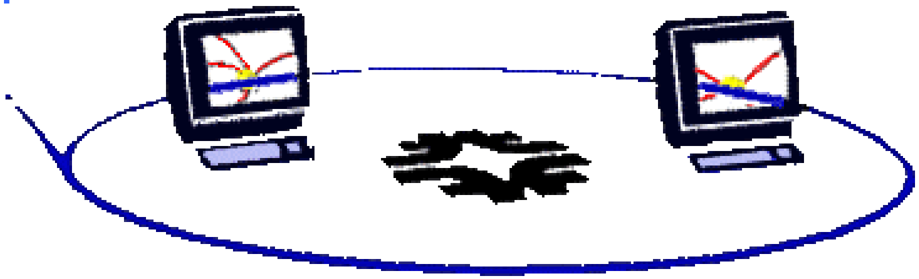
Scientific Computing: CDF Reconstruction FARMS

- Hardware
  - 50 x 500MHz, 23 x 800MHz, 64 x 1GHz; all 1GB memory
  - 32 x 1.2 GHz Linux nodes in burn-in
  - The above totals 13,700 SpecInt95
  - Single 4 x 400MHz CPU O2000 (used primarily job submission host)
- Data Processing stats
  - Processing data as it arrives with small delay to wait for calibrations
  - ~ 200 million events reconstructed so far
  - ~ 26 million events taken in the “final” CDF detector and trigger configuration (since early Feb ’02)
    - 21 million of these reconstructed on farm
- Mass Storage Info
  - Migrating from AIT2 to STK 9940a/Enstore for IO
  - Due to be completed in April/May
- Future
  - Enstore only IO
  - Purchase ~50 more workers in FY02
  - Starting FY ’03 will replace ~50 nodes/yr with newer hardware



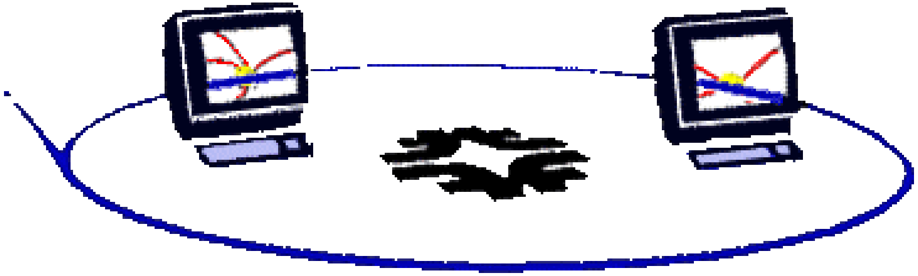
Scientific Computing: D0 Reconstruction FARMS

- Hardware
  - 40 x 500MHz, 50 x 800MHz, 32 x 1GHz; all duals w/ 1GB memory
  - Single 8 CPU O2000 as IO node
  - All Linux systems are now running 7.1 w/ 2.4.9-31 kernel
- Data Processing Stats
  - ~37 million events taken with ~45 million events reconstructed
    - Reconstructed can be > than taken due to reprocessing of some data
- Mass Storage Info
  - Utilizing Enstore to access mass storage
  - Migrated from ADIC robot to STK w/ 9940 tape drives
- Future
  - Additional IO system – probably PC hardware running Linux
  - More workers nodes to be purchased



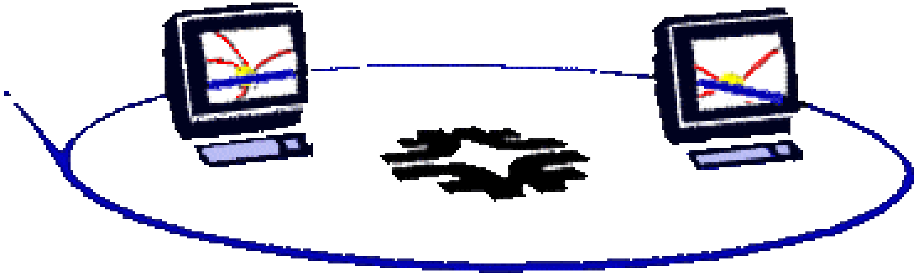
Linux Clusters Update

- CAF:CDF Prototype user analysis farm
  - Current config: 16 duals and 1.4TB file server connected via NFS
  - Stage 1 system (May):
    - 2 8-way interactive Linux servers to submit jobs from
    - 43 dual 1.26 GHz workers (in burn-in)
    - 20TB of NFS attached disk on order
    - Utilizing fbsng as batch system
  - Plan is to scale up to 600 workers and 200TB of network attached disk over next 2 years
- DAQ
  - D0 utilizing embedded Linux in their DAQ



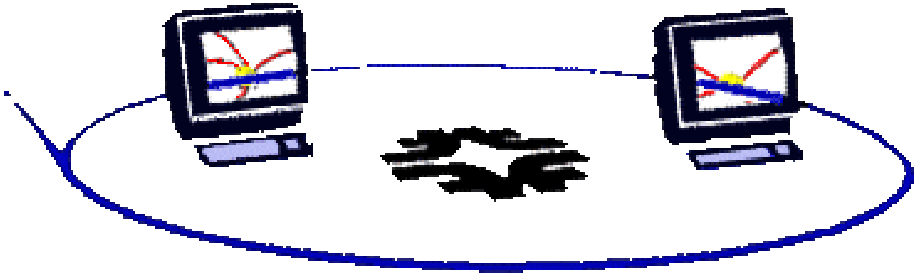
Linux Clusters Update

- pcQCD
  - Lattice gauge Theory cluster
  - 80 node cluster with Myrinet 2000 network interface cards, 700Mhz CPU
  - Cost approaching $1/Mega flop ($100/MF for ACPMAPS of early '90s)
  - http://qcdhome.fnal.gov/
- CDF/MIT Mosix Cluster
  - 10 dual processor PC's
  - Concerns about expansion
  - Extensions needed: batch system, virtual server, single file system
- CMS UAF Mosix Cluster
  - Prototype cluster
  - Similar to above with lower level of usage



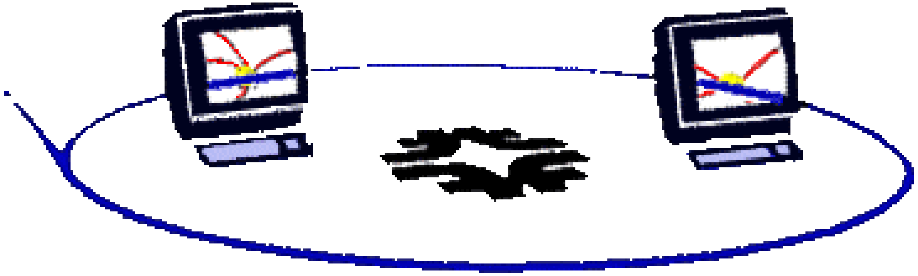
Central Services: AFS

- 11 servers, 1 AFS/NFS translator
  - 3 Ultra-10s (Fileserver, NFS/NIS/DB server, Backup server)
  - 1 Ultra-1 (Translator)
  - 2 Sparc20s (Fileservers; to be replaced)
  - 6 Ultra 60s (Fileservers)
- Majority run Solaris 2.6 to be upgraded to 2.8 3Q02.
- Use Transarc AFS v3.6 server w/ patch 3
- Fileservers serve ~2.5TB of disk (1TB scsi RAID, 1.5 TB FC RAID)
- Backup ~800-900GB of disk with AFS backup tool
- Plans for this year:
  - More disk; 730GB – 3TB
  - Get OS upgraded on servers
  - Retire old hardware
- ~3500 accounts in AFS Cell and growing
  - Unix/Linux and Windows desktop client base growing



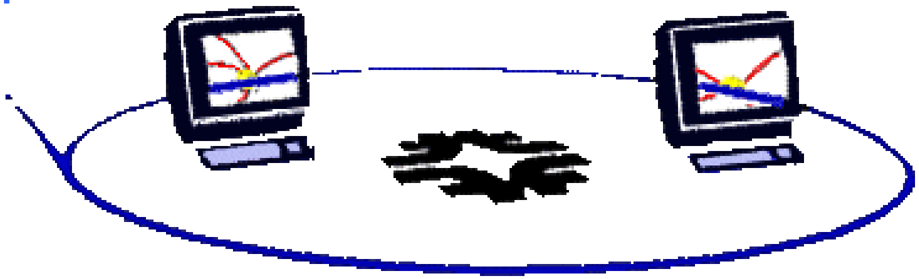
Central Services: E-Mail

- Anti-Virus - 2 stages of detection
  - Sophos AntiVirus on email gateways: scans all mail from offsite
  - Symantec AntiVirus on IMAP and POP servers: all mail received on these machines is scanned
  - This has caught 19,378 viruses since Aug ’01 (avg. 538/week)
- Mail gateway stats:
  - 23,102, 383 messages and 550,106,053 Kbytes of data in last year (444276 messages/week and 10,578,962 KB/week)
- IMAP
  - 3 servers in production: 2300 active users, over 55GB of mail stored online
  - Number of users has grown 32% in last year and mail store doubled
  - Future plans to move to Solaris based servers with a mail store on a SAN



Central Services: Web Servers

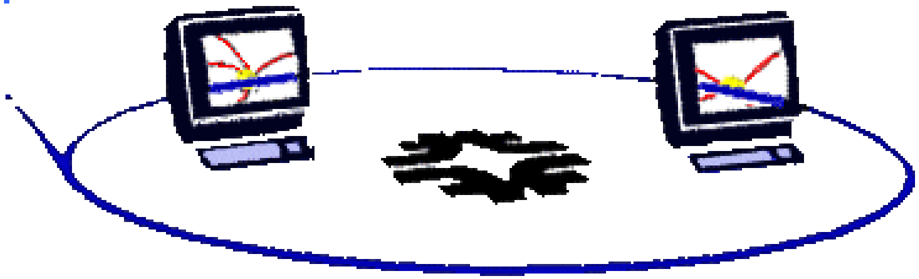
- 52 Web servers managed centrally
  - Some standalone, some virtual
  - Will be converting all to virtual
  - Currently run from 2 Sun Ultra60's
  - Scans of site show ~500 web servers running on site
- Plans for this year
  - Upgrade and increase hardware: 4 Sun Netra's
    - 3 as virtual web servers, 1 as CGI server
  - Put these 4 systems behind an Alteon switch to provide load balancing and fail over capabilities



Central Services: Web Servers

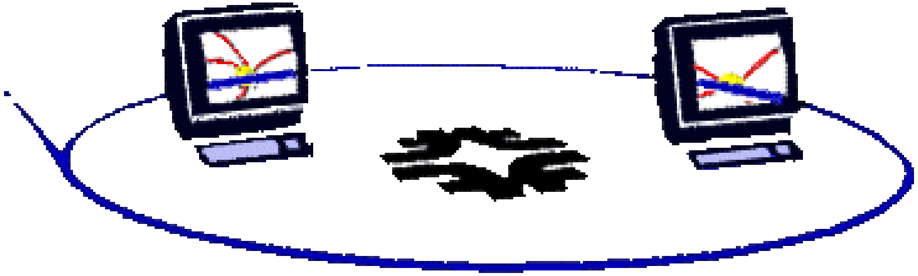
- Statistics Sep 30, 2001 – Dec 31, 2001

Successful server requests	21,706,799 Requests
Successful requests for pages	5,502,740 Requests for pages
Failed requests	439,555 Requests
Redirected requests	290,534 Requests
Distinct hosts served	302,579 Hosts
Total data transferred	179.10 GBytes



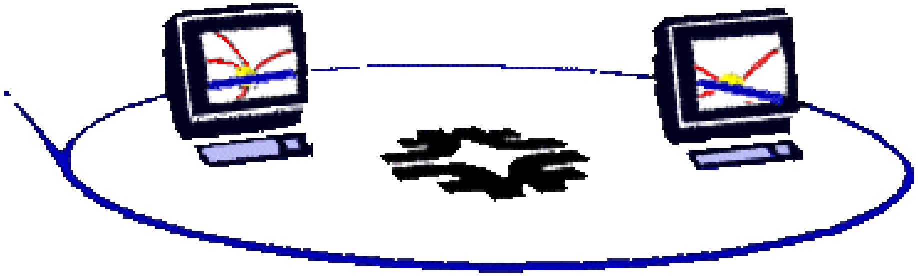
Security:

- Met (for the most part) the Dec 31, 2001 deadline for Strong Authentication rollout on Unix systems
- Running regular site scans for compliance with the policy
- Security problems continue to include:
  - Viruses
  - OS holes



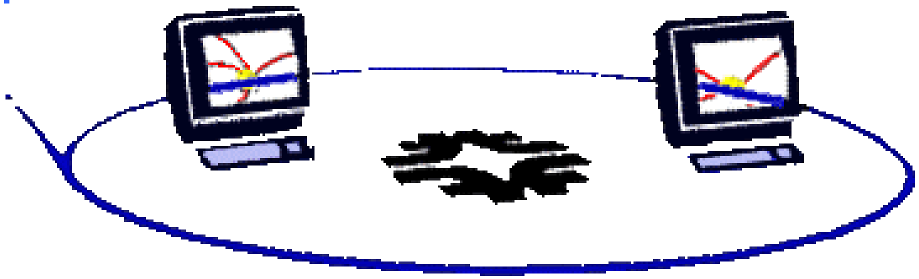
Networking:

- On-Site Networking
  - Core network consists of Cisco Catalyst 6509s with routing (MSFC2) modules
    - Have started upgrade of these to use cross-bar switch fabric, increasing back-plane capacity from 32Gb/s to 256Gb/s
  - Network Connections between core network devices and to work group switches now all gigabit Ethernet
    - Expect to start running multiple gige links in parallel for trunks and uplinks that need more than 1Gb/s
    - 10Gb/s Ethernet is still on the horizon; too expensive right now
  - The number of host systems with gigabit connections >60 and increasing rapidly
    - Gigabit connected systems only within computer rooms; no deployment or immediate plan to support this on desktop
    - Evaluating 1000B-TX (gigabit over copper cables) but have not deployed it yet



Networking:

- Off-Site Networking
  - Connectivity split between 2 OC3s (155Mb/s), ESnet & MREN
  - Traffic levels regularly exceeding 100Mb/s for sustained (>1hr) periods
  - ESnet link to be upgraded to OC12 (622Mb/s) 3Q02,
    - After upgrade, all off-site traffic will be sent over this link
    - MREN link will then revert to redundancy and test network roles
  - The DMZ (off-site…) LAN upgraded to support gigabit host systems
    - Facilitate useful network monitoring at data rates exceeding 100Mb/s
- Monitoring system deployed (6 pc's and growing)
  - Provides MRTG utilization graphs for all 4000+ switch ports on site
  - Graphs available to onsite admins and users
  - Web-based node locator tool finds the local LAN segments graphs
- Remote support of MINOS-Soudan experiment (400mi North of FNAL)
  - Mine LAN (800M deep) supported, monitored and maintained as if it were part of the FNAL campus LAN
  - GRE tunnel used to establish VPN link between mine & FNAL



MISC

- NGOP
  - Monitors all nodes which are the responsibility of CD (~680)
    - Agents (more detailed monitoring) running on farms and Enstore nodes
  - Migrating to NGOP for monitoring on several other clusters including FNALU
  - Presentation on Thursday 4/18