

# Highlights of Spring 2004 HEPiX

BNL Technology Meeting

14 June 2004

Tom Throwe

# Outline

- Summary of Site Reports
- Highlights from Site Reports
- RedHat Presentation and Discussion
- Other Topics
- Mass Storage Workshop
- Presentations:  
<http://hepwww.rl.ac.uk/hepix/nesc/agenda.htm>

# Site Summary

Site		Linux		Solaris	AFS	Batch
	Version	Satellite	Update			
SLAC	RHEL	No	YUM	2.9 (2.10)	Open	LSF
DESY	DESY L5 SuSE 8.2	No		2.8		SGE
FNAL	FL LTS (2 FTE)	No	YUM		Open	
JLAB	RHEL	Yes	up2date	2.8 (2.10)		
CASPUR					Open	
DAPNIA					Open	
CERN	CERN (3 FTE)		APT			
Manchester	Mac				Open	

# Highlights

- DESY
  - Starting to do 64bit testing and found that ROOT benefits
- FNAL
  - Facing space, heat and power problems (many sites echoed this problem)
  - Renovating old experimental hall
  - Cabling overhead and only A/C under floor

# Highlights (cont.)

- JLAB
  - Problems with SATA box, poor soldering and broken components on backplane
  - Moving everything to VLANs at recommendation of audit team
- CASPUR
  - Using AFS aware SSH 3.8p1 with GSSAPI or K5
  - Abandoning Heimdal for MIT

# Highlights (cont.)

- NERSC
  - BRO is Grid enabled and have DOE grant to produce BRO Lite for distribution
  - Looked at Panasas, Lustre and other new file systems and found them lacking. They are using GUPFS
- GSI
  - Many problems with IDE RAID, getting SATA system

# Highlights (cont.)

- TRIUMF
  - Westgrid (IBM blades) having many problems and still not stable. Order 1 failed node per day
- Manchester
  - Plan to expand Mac cluster to 1500 nodes and 200 TB of local disk

# Highlights (cont.)

- Oxford
  - Recently obtained an Eonstor IDE RAID system, said last IDE RAID system failed

# RedHat Presentation

- Only new information is that there is now a “Desktop” version of RHEL below “Workstation”
- SCO lawsuit
- Cluster definition (recent DOE phone conference)

# RedHat Discussion

- CERN
  - Discussion with RH took too long, use their version through 2005
  - Add and change RPMs (OK with RH)
  - Run RHEL on small # machines with commercial software (Oracle)
  - Evaluating RHEL

# RedHat Discussion (cont.)

- FNAL
  - Will run Fermi Linux LTS 3.0.1
  - Have RHEL on small # machines running commercial software (Oracle)
  - To continue with Scientific Linux requires community support

# RedHat Discussion (cont.)

- SLAC
  - Bought entitlements for on-site and off-site machines
  - No satellite server
  - Share a TAM with Livermore

# RedHat Discussion (cont.)

- Discussion
  - General agreement that a free HEP wide version of Linux based on RHEL SRPMs would be nice
  - Most agreed that Scientific Linux could work as the base with the “site” option being used for customization
  - FNAL said that Fermi Linux LTS could become a “site”
  - CERN said they could probably do the same

# Other Topics

- Grid
  - Update on LCG-2
  - LCG testing suites
    - Tools developed to verify LCG software installation and to detect failures
  - GridPP: a, mainly UK, Grid connecting Universities and Labs and CERN
  - LCG user registration and VO management

# Other Topics (cont.)

- CERN installation and monitoring software
  - Quattor for node configuration
  - LEMON for monitoring: ~1GB data/day for 2100 nodes and plan to integrate (Monalisa)
  - LEAF for hardware and state management (lifecycle and reboot, software install, etc.)
  - Widely used at CERN and some other European sites

# Other Topics (cont.)

- PDSF
  - CHOS or chroot OS allows running multiple versions of the Linux OS on the same machine (provided all happy with a single kernel)
  - Tested Lustre 1.0.2 and found to be better than NFS, but it had very poor recoverability
  - Plan to do a survey of Linux auditing and accounting packages, and to produce a package for distribution

# Other Topics (cont.)

- Disk
  - Major RAID failure at Glasgow
    - 2 months downtime
  - CERN
    - 370 EIDE Linux servers, 544TB on 6700 spindles
    - Mirrored: had 3% broken, exchanged WD disks
    - Want to go to RAID 5 + striping + XFS
    - Getting 60 TB SATA system
    - Project to look at USB2 or firewire attached disk to exploit # devices supported. Failed on SMP

# Other Topics (cont.)

- Security View from SLAC
  - Windows:
    - Viruses more sophisticated: selling as SPAM engines
    - IE exposures: do not make default browser
  - Linux/UNIX
    - Local exploit = remote exploit
    - CDE dt
    - One-time-passwords in your future
  - HTML mail format is bad

# Other Topics (cont.)

- AFS
  - All sites reporting said they are running OpenAFS on Solaris or Linux
  - Report on “AFS Best Practices Meeting” at SLAC
    - Windows version broken, scrapping and rewrite
    - Sine Nomine Association – support
  - Administration tools Arc 2 and K5cron

# Mass Storage Workshop

- IBM SANFS
  - Global, shared namespace consisting of all storage attached to the SAN
  - Manage by policy
  - Robust to adding and subtracting storage
  - Metaserver manages data
  - Access control is mapped to the local file system semantics of the client
  - Future – go to an OBS model

# Mass Storage (cont.)

- STK
  - Information Life Cycle management
  - New robot to replace Powderhorn
    - 1500 – 6000 cartridges (750 – 3000 TB)
    - Up to 64 drives (4 section, 4 robots)
    - Next version will have 2 robots per section
    - Pass through ports and handoff between robots

# Mass Storage (cont.)

- CASPUR Storage Lab
  - Well established with vendors and have lots of loaner equipment (SANFS)
  - Horizontal disk orientation preferred
  - Linux server with Qlogic or Emulex FC HBAs
  - Qlogic SANbox 5200 stackable switch
  - AFS speedup: replace RxRPC or use Reuter AFS. 30-45 MB/sec went to 425 MB/sec
  - Lustre needs further investigation

# Mass Storage (cont.)

- Discussion of LCG Tier 0 to Tier 1 data transfer test
- Presentations on SRM and SRB

# Conclusion

- Next HEPiX at BNL in October (18-22)
  - Workshop on “Alternatives to 32 bit X86”
    - 64 bits
    - Opteron
    - Itanium
    - Mac