HEPiX report

Helge Meinhard, Rafal Otto, Matthias Schröder / CERN-IT

Computing Seminar / After-C5

5 November 2004

Outline

- Site reports, Unix related topics (Helge Meinhard)
- Windows topics (Rafal Otto)
- Large Cluster SIG workshop (Matthias Schröder)

HEPiX

- Global organisation of service managers and support staff providing computing facilities for HEP
- Covering all platforms of interest (Unix/Linux, Windows, Grid, …)
- Aim: Present recent work and future plans, share experience
- Meetings ~ 2 / y (spring in Europe, autumn in North America)

HEPiX Autumn 2004 (1)

- Held 18 – 22 October 2004 at Brookhaven National Laboratory, Long Island, NY (USA)
  - Home of RHIC (Star, Phenix, …), US-Atlas Tier 1
- Format:
  - Mon – Wed Site reports, HEPiX talks
  - Thu – Fri Large Cluster SIG on
    - Platforms for physics computing
    - Scientific Linux, RedHat Subscription model
    - Experience with Grid Operations
- Well organised by Tom Throwe
  - Getting access to US and to Tier 2 DoE lab not trivial
- Full details: https://www.bnl.gov/hepix

HEPiX Autumn 2004 (2)

- 70 participants, of which 11 from CERN-IT
  - Bird, Duellmann, Gaspar, Iven, Kleinwort, Meinhard, Otto, Schröder, Schulz, Silverman, van Eldik
  - Other sites: FZK, BNL, DESY Hamburg, SLAC, LAL, LBNL, FNAL, INFN, JLAB, Madrid, DAPNIA, RAL, Prague, TRIUMF, NIKHEF, GSI, IN2P3, CNRS, DESY Zeuthen, Stony Brook,
  - Vendors: Intel
- 58 talks, of which 15 from CERN

Next meetings

- Karlsruhe, 9 – 13 May 2005
  - Large Cluster SIG: Batch schedulers, …
- SLAC, autumn 2005

Site report highlights (1)

- Linux: HEP converging to SL3
  - Mentioned by almost everyone
  - Even DESY committed to migrate
  - Exceptions: GSI (Debian), SLAC (RHEL 3), CERN (small part RHEL)
  - Major success for FNAL (and HEPiX)
- Cluster mgmt: Increasing interest in Quattor
- Disk pool mgmt: Most sites use dCache (DESY/FNAL)
- Serial console (SLAC/CERN) being considered by LAL et al.
- HEPiX has played catalytic role in many of these converging evolutions

Site report highlights (2)

- Batch schedulers: random walk in phase space spanned by PBS (PBSPro, OpenPBS, Torque/MAUI), LSF, SGE, Condor, BQS
  - Only commonality is no-one migrates into $$$ solutions
  - Trend towards integrated farms with fairshare
- Disk storage: No clear tendency
  - Everything being procured – from cheap NAS boxes over full-fledged SAN solutions to specialised HW/SW implementations (e.g. Panasas)
  - Increasing interest in NAS
  - No-one wants IDE disks, much more credit for SATA
  - Parallel file systems not taking off
    - Niche applications running Lustre (LBNL)

Site report highlights (3)

- Platforms: converging to duality Xeon / Opteron
  - Performance comparisons suggest advantage of 64 bit applications, and price/performance advantage of Opteron vs Xeon
  - Some sites now running Opterons in production (few %)
    - SLAC bought 300 dual-Opteron systems
  - MacOS X (on Power) not taking off very rapidly for server applications
    - Only SLAC reported having installed a MacOS X cluster (for non-HEP clients), some particularities
- AFS:
  - Transarc to OpenAFS migration completed almost everywhere
  - Most sites have successfully migrated to krb5
  - Commercial support available for OpenAFS (SLAC subscribed)
  - Perl monitoring and debugging library developed at SLAC

Site report highlights (4)

- Tape storage: Workhorses are STK 9940B everywhere
- Spam fighting mentioned occasionally
  - Legal problems simply deleting spam mail
  - JLAB going for commercial off-site solution
- Security…
  - Biggest problems are self-maintained PCs as well as laptops
  - Viruses, phishing, spyware, … getting ever more sophisticated. Applying patches not always easy
  - Fedora Legacy far too slow providing security fixes
- WAN: catching up with LANs
  - All sites reported at least 2.5 Gbit/s connectivity

Role of HEPiX

- **IHEPCCC request**
  - IHEPCCC, subcommittee of ICFA, somewhat a HEPCCC successor – top-level managers of HEP computing centres
  - Need technical input
    - HEPCCC had standing technical committee HTASC
    - IHEPCCC doesn't want that, suggesting rather that HEPiX provide input on request
    - Would imply that HEPiX cover new areas, e.g. networking, virtual organisations, …
    - First topics suggested: Linux, VOs
  - HEPiX attitude positive, details to be worked out

Interesting points (1)

- Schön / GSI: SATA file servers
  - Many problems with SCSI – IDE disk arrays
  - Bug in ASIC of 3W-8506-8
  - Good experience (including reliability) with Maxtor Maxline Plus 2 and WD Raptor
    - Raptor: 10 k rpm… a lot faster than 10'000/7'200
  - Read-ahead buffer is sensitive parameter
- Petkus / BNL: Panasas experience
  - BNL used NFS-mounted SAN infrastructure, running Veritas VxFS so far
  - Panasas: integrated HW/SW storage solution
  - Difficult start, but now running very well
  - Software issues: constantly changing, requiring Linux kernel patches

Interesting points (2)

- Canon / LBNL: CHOS
  - Developed at LBNL (NERSC)
  - Linux kernel module allowing to run different Linux versions on a host
  - Advocated as a migration tool