

# HEPiX TRIUMF Meeting Report

Minutes of the Meeting held April 10th - 12th, 1996

---

## 1. Introduction

### o **Logistics**

The meeting took place at the TRIUMF Laboratory in Vancouver, British Columbia, Canada from 10th to 12th April, 1996. It attracted some 36 people, nearly half from Europe. 20 different HEP sites were represented, 7 of them European. Most of the overheads presented are now on the web and pointed to in the web version of these minutes below. The meeting was organised by Corrie Kost for local arrangements and by Alan Silverman for the programme. Large parts of it were broadcast over the MBONE.

### o **TRIUMF Introduction by Prof. Alan Astbury**

Prof. Astbury welcomed HEPiX to TRIUMF and described the work done on the site and its relations to other HEP sites. TRIUMF, Canada's national lab for high energy physics, is run by a consortium of 4 universities. Its principle HEP tool was a 500 MEV Cyclotron for subatomic physics. The lab had a basic science programme including nuclear physics; particle physics which was mostly offsite support for experiments taking place at other labs; and a condensed matter programme. There were also substantial Life Sciences and Medical Science programmes.

There were some 300 staff at the lab plus students and research assistants. There were several technology transfer collaborations with industrial partners which helped in funding some of the programme.

The future goals of the lab were to maintain much of the basic science programme, to build a radioactive beam facility, and to provide Canada's collaboration with LHC "in kind".

## 2. Site Reports

### o **TRIUMF by C.Kost**

Since the Rio HEPiX meeting, the number of VMS nodes on site had dropped and the installed UNIX power had increased as had the number of X terminals. Also, an Alpha 600 server model 5/333 had been installed. Internet connectivity had been improved along with more dial-in capability. The lab had migrated to an FDDI backbone and they will move to a switched desktop network infrastructure.

Their DAT robot had suffered two major failures, one of which had destroyed very many tapes. They would now be investigating DLT devices. There had been an explosion in the PC presence and they were actively investing in integrating UNIX and Windows with the NCD Wincenter tool (see below).

Trends for the future included

- Integrating UNIX and Windows
- C++ replacing Fortran
- Home-based computing, helped by Canadian PTT rates
- A plan to migrate away from VMS over several years and concentrate more on UNIX and PCs.

Nevertheless, some visions were still blurred --

- Should the future be PCs, workstations or X terminals? For the moment they had chosen X terminals.
- Should tapes be 8mm or DLT? DLT looked more promising.
- Was C++ the future? It seemed likely.
- What was the future of wide-area audio and video?
- What was the file system of the future - NFS, AFS, DFS? For the near-term it seemed that AFS was the preferred choice.

### o **Yale University Physics Dept. by R.Lauer**

There were 6 distinct groups in the department, largely funded by DoE Research Grants; they performed or supported experiments at various different laboratories. Although funds were pooled, this led to diverse requirements.

Since the Rio meeting, the network had become based on DEChub devices and included some 65 nodes of different architectures. The main goals for their UNIX cluster remained the same however - a

consistent environment, ease of maintenance and reliability. Usage of the cluster had increased greatly, thus proving the principle of the scheme adopted, as described during the Rio meeting. Future plans for the cluster included making it fault-tolerant using Digital software but this had not yet been achieved.

The support people had still not decided on the preferred file system and there was a short debate with the audience present on the relative merits of AFS or DFS today; AFS was clearly preferred at the present time. Also, the speaker felt that Digital UNIX lacked ports of many of the application found on other platforms and that her users wanted but this was not echoed in the audience apart from one or two specific programs.

She stated that there was a diversity in the user environment found at many HEP sites, which inevitably led to a discussion about the merits or otherwise of at least making the DESY and CERN-developed HEPiX scripts available, not necessarily the default, at all HEP sites. In closing, the speaker made a plea that HEPiX becomes committed to active advocacy on behalf of HEP sites, especially the smaller ones. As an example she felt that Digital should be encouraged to boost the level of support they provide for Digital UNIX.

o **CCIN2P3 by W.Wocjik**

The IN2P3 Computer Centre in Lyon provided computing services to all 17 IN3P3 sites across France; there were more than 3000 IP addresses, all connected by PhyNet. CCIN2P3 itself had an ATM switch interconnecting its various UNIX clusters. VM had recently been downgraded by a factor of 3 but was still used to stage tapes to the Basta farm and was still serving the STK robot.

The speaker then covered the various farms and clusters:

- Basta: 18 HP 735s and 11 IBM RS/6000 390s used for CPU-oriented batch; 90% used.
- Anastasie: 14 HP 735s and 9 IBM 390s used for I/O oriented batch; 55% used.
- Bahia: 2 HP J200s and 4 IBM 390s used for interactive work and job preparation. The 390s were actually nodes in the SP2 in the Centre.
- Sioux: 2 HP 735s and 4 IBM 390s for interactive work; in particular this cluster made use of load sharing for heavy interactive loads.
- Dataserv: some 550GB of disc space (including 213GB of RAID) on 6 IBM servers used for file access; the data was accessed via ATM to the other clusters. It made heavy use of AFS plus some NFS automounted discs.
- Tape server: 4 IBM 370s connected via parallel and ESCON channels to robots and 2 HP 735s connected to DLTs. Tapes were accessed via XTAGE (see previous HEPiX meetings) and the RFIO package.

The SP2 was used for a mixture of interactive work (Sioux and Bahia) and as a PIAF server. It was also used for remote execution of tasks requested by other clusters. Security was implemented by AFS/Kerberos. This scheme was found to have the advantages of saving licence costs on commercial software and offered better loadsharing.

o **BNL by T.Schlagel**

BNL ran a very mixed architecture site. They were gearing up for RHIC which was due to start operation in Spring 1999. There would be 4 major experiments with a total of some 840 physicists from 80 institutes in 15 countries. RHIC planned to have an offline computer centre with an 8 node SMP SGI system and a 16 node IBM SP1. For their batch farm, they would use DNQS (already described at previous HEPiX meetings, see for example HEPiX Fermi in 1994). They were also investigating using Pentium Pro systems running either Solaris or Windows NT.

Of their 3 AFS cells, one was currently not active, one ran AFS 3.3A on AIX 3 servers and the third ran AFS 3.4A on Solaris 2.4 servers. There was a total of some 100GB of file space. Various tape media were planned or already present on site and an ATM test bed was being set up.

BNL had developed SQIRT, an installation tool for installing public domain software. SQIRT stands for Software Query Installation and Removal Tool. They had evaluated WinDD and WABI but had finally settled for WinCenter with which they were very impressed.

Plans for the future included investigating --

- Kerberos ticket services
- Windows NT versus UNIX for physics
- PGP public key services
- AFS to DFS migration

o **RAL by J.Gordon**

The IBM cartridge robot now had 5 3590 drives and with these they were seeing more than 9 MBps tape to memory transfers, 6 MBps to disc. There were about 900 tapes in the robot today and

maximum capacity would be 21TB with 10 GB per tape.

Due to the lack of direct access to tape I/O, they were considering to implement CERN's RFIO package on their (non-UNIX) data servers; other alternatives included stage software or dedicating disc for certain users. Use of the CERN tape stager was likely, at least for some users.

Their CSF installation, used for more than just Monte Carlo jobs, had been upgraded with the addition of 4 HP C110 nodes. It was used by several LEP and LHC collaborations as well by DESY and SLAC (BaBar) experiments. The HEPiX profiles and ASIS were installed. The RAL production AFS cell was built round a single server today with clients spread across JANET.

o **CERN by A.Silverman**

The main areas of expansion in CERN continued to be X stations on desktops and storage capacity, both disc and tape, in the CORE Batch services. There were now some 5TB of disc space and tape cartridge drives and robots of almost every type. Central data recording had become very popular and this year's run of NA48 should set new speed records. Outside of UNIX, a Windows 95 service was about to be announced and released (scheduled for May 1st) and great interest was being shown in Windows NT.

CERNVM (the IBM mainframe) switch-off date was set for June 30th and a major effort was being put into user training, user environments and so on but at the date of the meeting, over 2300 users still remained firmly active on CERNVM.

AFS 3.4 had finally arrived and been installed but many AFS problems resulted which unfortunately coincided with a number of server hardware incidents, one of which the vendor has so far been unable to trace. As a result of these various problems, a second server architecture had been introduced (SUN/Solaris) and some AFS client-based services had been down-graded back to AFS 3.3 while the newer version was fully debugged and understood. CERN had proposed to the HEPiX AFS Working Group that Transarc be told not to work on AFS 3.5 but concentrate on DFS instead but they seemed to be in a minority.

CERN buildings were being steadily recabled with twisted pair cables and the network migrated to being routed instead of bridged. This involved a lot of work by a lot of people, including those in UNIX support since all the vendors' remote installation procedures broke when the protocols (often propriety) they each had chosen to use during the remote installation procedure, usually for remote discless bootstraps, were refused by the routers between individual systems distributed across the site and the central installation servers. Some solutions had been found but not for every platform.

Other areas of work included a new central twin-node mail server (to be covered at the Rome meeting), implementation of the Gnats problem tracking scheme (talk later this meeting) and much work around X11 and the HEPiX environment, both login scripts and X11 look and feel.

o **DAPNIA by P.Micout**

Since some time now, DAPNIA shared the central computing resources of CCIN2P3 and discussions were currently under way to discuss arrangements for the years from 1997 onwards.

The SP2 used by DAPNIA had had an additional 6 nodes installed as well as 2 Magstar units in the IBM 3494 tape cartridge robot. There was also a DLT robot present and RAID arrays from IBM and Digital. AIX 4.1 and Loadleveler were now installed on the SP2. DAPNIA had now established its own AFS cell. There was a "mini" TMS service which would be reported on later in the meeting and a Pine IMAP-based service. The main building had been recabled with twisted pair cable rated at 100Mbps and an FDDI ring was planned. Lastly, there was a new print server and several new printers.

o **CEBAF by S.Philpott**

CEBAF had selected SUN/Solaris and IBM RS/6000 model 43Ps running AIX as their preferred UNIX platforms after a formal procurement procedure as mentioned at HEPiX in Rio. The choice had been made on integer performance, price/performance and support for high-end tapes as principle reasons. Other factors considered as important but not crucial included good floating point, 64 bit support and file, file system size maxima and a few other things. Their existing HP-UX systems would be maintained for now but their ULTRIX systems would be decommissioned by October 1st.

The fact of introducing two new UNIX platforms gave the opportunity for a new environment, a new start. A list of requirements had been drawn up:

- What file system to choose: should they wait for DFS?
- What about clustering?
- How to configure X?
- Work group servers and/or dataless clients
- Whether to put software in /usr/local/bin
- How to access non-standard versions of /bin programs

They had established a [draft documentation](#) on the web at URL <http://www.cebaf.gov/~chambers/unixdraft.html> and invited comments.

o **FNAL by S.Hanson**

Use of FNALU has increased, in both CPU and disc quotas, as VMS migration has advanced. There were now some 330 active users. The speaker showed some interesting graphs to illustrate the usage trends. More AFS server disc space was due to be added plus more CPUs on the SGI Challenges, a four-CPU Alpha 2100, a new IBM PowerPC system and a new HP system. CDF was adding more CPUs to its Challenge and D0 were now running on Challenges for both PIAF and general computing.

FNAL was beginning to use arc for the delegation of system authority. Mail users were switching from VMS to UNIX and MAC for mail and tests were underway of both IMAP and POP servers for UNIX and PCs. Use of AFS was increasing steadily as was the use of licensed software packages and products. Other evaluations in progress were a dedicated WABI server and the WinCenter product to provide X access to Windows applications (see later).

A study group was being setup on the potential uses of Intel CPUs in farms. Software under consideration included Linux, FreeBSD and Solaris as well as Windows. Meanwhile, FNAL intended to perform some pilot tests on DCE/DFS later this summer and work had begun on an update to the Fermi Tape Tools Project.

o **SLAC by R.Melen**

The central SLAC compute farm now consisted of 41 RS/6000 nodes for batch and 12 for interactive work. SLAC could relate many experiences with the latest 43P models. In addition, there were 4 file servers with 60GB of disc each, including 9GB Seagate drives, 4 tape servers connected to 8 STK silos with 3490E drives and a final RS/6000 server driving Exabyte and DAT devices. The software in the farm included LSF and CERN's tape staging package. They expected further expansion of the compute farm, possibly using a different architecture based on suggestions from the BaBar Collaboration. They were adding 3 new servers for NFS and news and looking at STK's Redwood tape product.

The SLAC AFS cell now had 3 servers, 48GB of disc each, over 350 active users, of whom 110 used AFS for the home directories.

Plans were well advanced for VM shutdown - end December 1997 - and migration had started. One large open question concerned what to do about the SPIRES HEP database:

- discontinue it and replace it by a commercial product, perhaps based on ORACLE?
- convert it to ORACLE themselves?
- move to USPIRES, currently under test

SLAC was under severe pressure on staff, their system administration team having been cut from 7 persons to 3 for example. PC and MAC support was staffed by only 2 FTEs.

o **GSI by J.Heilman**

GSI was a heavy RS/6000 user, still running AIX 3.2.5 as the production version. They had a 16 node SP2 and an 11 node RS/6000 cluster; there was a total of 40-50 workstations plus more than 100 X Terminals. File backup was still based on MVS but they were almost ready to switch to using RS/6000s and ADSM.

Other hardware present included 2 HP nodes which ran and would continue to run HP-UX 9.01 for historical reasons plus an ULTRIX-based print server connected to 50 printers around the site.

There were some 900 user accounts of whom about half were currently active and there was about 100GB of disc space, all NFS. Loadleveler was in use.

Near-term plans included moving to AIX 4.1, customising CDE and the backup migration to ADSM on RS/6000 servers.

---

[Alan Silverman](#), 4 June 1996