Annual Report of the Computing and Networking Advisory Panel 1998/9

June 1999

1. Introduction

Much of CNAP activity during 1998/99 has been the consolidation and rationalisation of centrally provided services, the overseeing of better network connections and the impact of the start of a new era of HEP computing using object oriented programming methods and database facilities with the imminent coming online of the BaBar experiment. The main focus of CNAP in the future will be the provision of computing for the LHC experiments and the management of resources from other sources of funding such as JREI and GIF.

Section 2 of this report deals with centrally provided computing and data store services. Network issues are covered in section 3, JREI/JIF bids in section 4, CNAP sub-group reports in section 5, training in section 6 and user representation in section 7. A brief summary and list of recommendations are given in section 8.

2. Central Facilities

2.1 Unix Farms

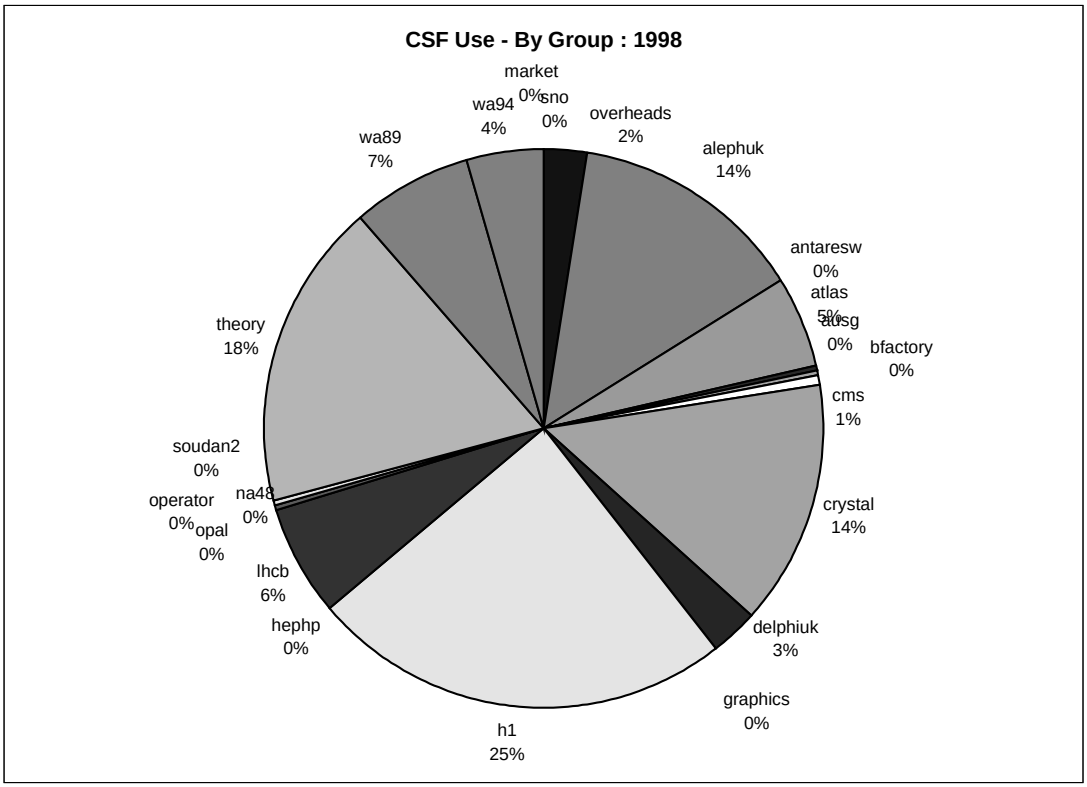
2.1.1 HP

The CSF HP upgrade of five C200 workstations, ordered towards the end of 1997, was delivered early in 1998 and installation went smoothly, as did the integration into the farm. The extra capacity was measured at 43% by averaging the performance of the most used experiments' codes, and was fully used right away.

The performance of the new HP C200 was disappointing. For a couple of years the SPECint95 benchmark was a good indicator of performance on HP codes but recently processor developments have improved the benchmark figures without corresponding increases in typical HEP codes. Compilation with optimisation specifically for the PA8200 cpu produces some improvement but the binaries then become unsuitable for the rest of the farm. The upgrade also included a substantial upgrade to the front-end machine. The extra memory and faster cpu have greatly improved this part of the service.

At the same time as the upgrade there was also a major networking reorganisation: a Fast Ethernet switch (100Mbit/sec) was installed; the new C200s came with Fast Ethernet and the older C110s were converted; the 20 HP710 farm nodes were moved from two FDDI-ethernet switches to a single ethernet switch with a Fast Ethernet uplink. Many of the HEP central systems have now been moved behind a router and insulated from general DCI traffic and this work will continue. The performance observed to the disk servers behind the switch was much improved and this new configuration also allows the possibility of implementing extra security in the router.

The service was heavily used throughout the year by a wide range of experiments from almost all universities (see figure 1). Notably Crystal Barrel received a pulse of extra resources in early 1998 to help them meet a deadline and H1 were consistently heavy users throughout the year resulting in a total use of 86,000 HP712 cpu hours.



2.1.2 Linux

There were several reasons for investigating Linux in 1998. The relatively poor price-performance of HP hardware seemed due to a marketing decision by HP not to pursue the education sector. Comparisons with Sun, Digital (now Compaq), IBM, and even SGI were poor. The price-performance differential with Intel hardware was even more marked. There were also indications that with future HP hardware and operating systems it might no longer be possible to run the same level of operating system across the whole farm. These possible problems came to light just after CERN had announced that they did not intend to move to HP-UX 11 for several years. At the same time several groups were making enquiries about support for Linux on Intel hardware as their experiments were heading that way. These, coupled with the increasing demand for resources on CSF lead CNAP to suggest that future capacity be provided by Linux/Intel. A development system was implemented on a redundant PC and made available to several experiments for tests. These were successful and ten dual 450MHz Pentium II cpu boxes were purchased at the end of the year. These will be integrated into the CSF farm as closely as possible, sharing a batch system.

2.1.3 Digital Unix

Although the general Digital Unix service for particle physics closed down in 1997, a temporary system was approved by CNAP in response to a problem experienced by the BaBar experiment. The Objectivity software on which their code is based was not supported on HP-UX for most of 1998. CNAP asked DCI to provide a solution for their program development, which they did by redeploying a Digital disk-server, increasing its memory and adding AFS software. This provided useful work for most of 1998 until it was superseded by the Sun development service described below.

2.1.4 Sun

After the experiment with a dedicated program development platform BaBar made a longer-term case for a substantial development platform for UK-based development based on the operating system with a guaranteed long-term future for their code - Solaris. CNAP agreed to this with the caveat that other experiments could make the case to use it too. A development service was brought up by enhancing a desktop system and this provided a bridge before the installation at the end of the year of a bigger E3500 4 cpu system.

2.2 NT

The NT Farm of 11 Pentium Pro cpus (one front-end and ten batch cpus) became a production system at Easter 1998 and was full for most of the rest of the year. Apart from tests by a number of people most of the work has been from ALEPH; the expected production simulation from ATLAS did not materialise. Discussions are taking place with CMS, ATLAS, and LHCb over future work.

The NT farm was converted to Fast Ethernet as part of the networking changes to the CSF farm described above.

A new Platform Manager, David Salmon, started work in July and has done a lot of work on the environment, resilience, monitoring, and installation procedures of the farm as well as planning a new front-end based on Microsoft NT Terminal Server Edition.

2.3 Data Storage

The central data storage facilites at RAL continued to evolve with enhancements made by DCI through 1998.

A number of improvements were made to the DataStore to improve the speed of data access and reduce wait times. These included commissioning two new servers and disk drives to provide a faster interface to the IBM 3494 tape robot.

A tape (3590) archiving system went into production in March 1998 as part of a particle physics initiative to release pressure on the tape robot and to provide increased capacity at low cost. This has already been used by Soudan and H1 currently have 50 volumes stored offline. Nonetheless, tape storage continues to grow steadily with particle physics using 1095 robot slots (as of March 1999) out of the 1200 slot allocation (see figure 2). This offline system is regarded as a temporary measure until the future direction, particularly for the LHC experiments, becomes clearer.

A DLT7000 system with a 15 slot stacker was also installed during 1998 to improve the import/export facility which was previously based around DLT2000 media.

Emphasis has continued to be placed on improving disk storage with the Disk Farm reaching 0.5TB in 1998 using a mix of 4GB and 9GB disks. A recent upgrade using 18GB disks will result in a capacity of 1TB. With the advent of a distributed OO database in the UK for BaBar (part of a JREI project described below), new central disk systems will need to be implemented during 1999.

°

The data storage group formally met once during 1998 in June. At this time, it was hoped that a bid made by RAL to HPC97 would provide a basis for new large-scale robotic facilities for particle physics. However, the HPC tender was eventually awarded to Manchester and this option was removed. An additional meeting to a wider audience took place in October. As part of this, the role played by data storage in potential bids to the Joint Infrastructure Fund (JIF) was extensively discussed. Although, one individual UK group (CDF) has so far bid for storage hardware for their specific experiment, it remains to be established how the large scale upgrade necessary for the LHC era can be taken forward. However, during 1999 is hoped to build a closer liaison with CERN and learn from their experience with Hierarchical Storage Managers, particularly the HPSS and Eurostore projects.

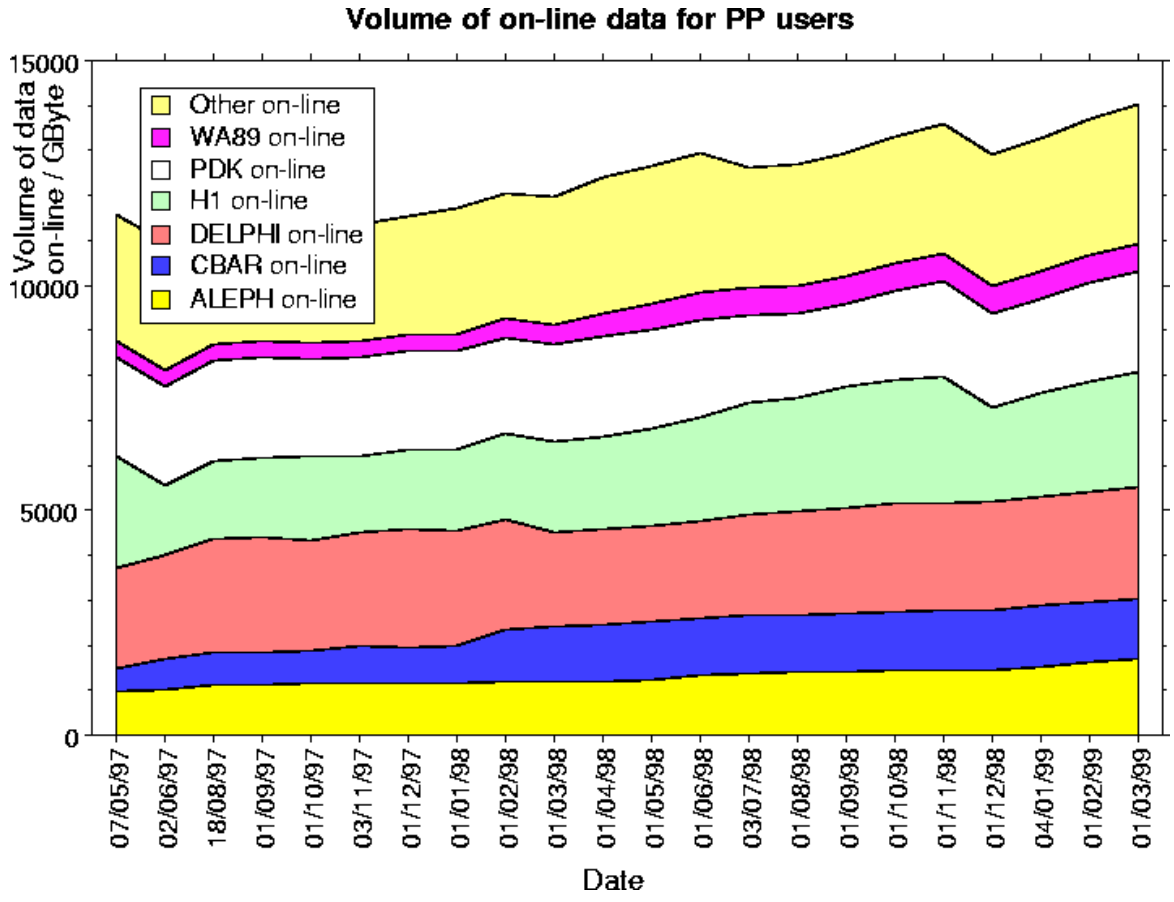


Figure 2

Data storage is going to be an extremely important issue as we head towards the LHC era. The data storage group is currently conducting a survey of data storage requirements which will be reported to the PPC in due course.

2.4 Central Computing Resources

2.4.1 FY98-99

For the Financial Year April 98 to March 99 the central computing budget was spent as shown in Table 1 and the spend on new equipment in Table 2.

	£K
Maintenance	39
Travel	9
Other	25
OO Training	21
New Equipment	161
Total	255

Table 1 Central Computing Spend 98-99

	£K
Unix Farm	15.5
Linux Farm	38.5
BaBar	38.6
NT Farm	18.4
Atlas Datastore	3.0
Disk Farm	47.0
Total	161.0

Table 2 New Equipment

There were three major elements to the new equipment spend:

- Extra cpu capacity to meet increasing demand from experiments. The extra capacity installed in 97-98 was filled as soon as it was installed. Providing this by Linux was a new direction for this year as described in 2.1.2.
- A program development platform for BaBar UK. BaBar made convincing cases for a separate program development machine with a lot of memory and for the platform to be Sun.
- An increase of the Disk Farm to 1TB to meet increasing demand.

All of these were debated and agreed by CNAP. The demand for the Atlas Datastore was flattened by the provision of offline storage so little was spent this year.

2.4.2 FY1999-2000

For 1999-2000 the priorities are:

- To meet the increasing demand for cpu on the platforms demanded. It is proposed to buy Intel cpus and implement them as Linux or NT as required and to integrate all the various architectures as closely as possible through a shared infrastructure.
- To meet the target of 20TB of datastore by the millennium that was set by the CNAP review of 1997 as advised by CNAP's Datastorage Subgroup.

It is foreseen that the resources should support approximately 9 staff years of effort and that there should be £250K for capital and recurrent spending and £20K for OO training of UK physicists. This represents an overall increase from £657K to £700K. The foreseen breakdown of staff effort is shown in table 3.

Service	Staff
Unix Farms	2.0
NT Platform	1.5
Datastore	3.0
Network Monitoring	1.0
Video Conferencing	0.5
Development	0.5
Indirect	0.5
Total	9.0

Table 3: Staff Effort for ITD Services

The increases in staff effort are required for the continued provision of Solaris services for BaBar and the support of extra services obtained through JREI and JIF bids (see Section 4).

The capital expenditure will be split between the provision of extra cpu resources for the farms and raising the capacity of the Atlas Datastore to the 20TB that the CNAP review of 1997 set as a target for 2000. There may be demand for extra disk space as well but the large JREI award to BaBar should relieve pressure on existing disk.

3. Networking

The PPNCG is a subgroup of CNAP which is concerned with general networking requirements of the Particle Physics community. It also has a member of the Astronomy community to represent their needs. The main provision of European and Transatlantic networks are discussed below. Network monitoring, a strategic activity of the PPNCG, and Video-conferencing are discussed in Appendices A and B.

3.1 SuperJANET-III and European TEN-155 networks

SuperJANET III has been in production service since April 1998. A core 155 Mbits/s ATM network connects to a set of 'backbone nodes' and metropolitan area networks (MANs) are connected to SJ-III via the appropriate backbone nodes. SJ-III has generally worked very well. Both PPNCG monitoring and user experience indicate that this provision is adequate for intra-UK use by the HEP community (except possibly packet based video-conferencing discussed in Appendix B).

Until 10th December 1998 European traffic was carried by TEN-34, the Trans European Network. On that date it was upgraded to 155 Mbits/s (TEN-155). The HEP community has been reliant on TEN for some time, and has not maintained an independent high capacity private leased line. The TEN-155 bandwidth provision has proved to be just adequate so far; there are no major bottlenecks and it has performed reliably. From the point of view of PPNCG monitoring we see slightly higher round trip packet times and packet loss than purely over SJ-III, but this is marginal. From the user experience point of view interactive connections to CERN are acceptable at most times. The connections into DESY are not quite as good, this being due mainly to the TEN-155 to DESY networking within Germany.

3.2 Transatlantic Links

Until 22nd June 1997 the traffic to the USA was carried on a single 45 Mbits/s link (T3 link) following the award of a contract to Teleglobe. For a short while the performance was good, but soon packet loss reached rates which rendered it unusable for interactive HEP use. On 28th May 1998 an extra 45 Mbits/s line was purchased. The same pattern was repeated in that within a short time it became essentially useless for interactive HEP needs, although average file transfer rates are probably acceptable provided one can wait. Hopes that the introduction of charging (currently set at 2p/Mbyte) would curb demand were disappointed, traffic volume instead continuing to increase. Thus in the last reporting period it is fair to say that the needs of the US

experiments have essentially not been satisfied.

On 6th April 1999, somewhat overdue owing to technical difficulties, a single 155Mbits/sec line replaced these aggregated links, with a second 155Mbits/sec link due in May/June. While it is true that since the 6th April interactive response has been very good, and that within a couple of months the total transatlantic bandwidth will double to 310Mbits/sec, two developments are then almost certain:

1. That the demand will, as before, grow to saturate it very quickly. This will inevitably lead to the same packet loss syndrome rendering interactive use impossible. It should be stressed that that truly interactive use is not the prime purpose of the links, and so this problem does not affect most users. TAU estimates that it will be overloaded within 9 months.

2. The US provision will further unbalance traffic flows on SJIII, already perturbed by the upgrade of TEN-35 to TEN-155. This could well lead to bottlenecks within the UK until SJ-IV.

The PPNCG has recently held discussions with representatives from CDF, D0, Minos and BaBar where it was unanimously confirmed that the transatlantic provision is wholly inadequate. The seriousness of the situation cannot be understated. The UK now has collaborators in several high profile US experiments, with significantly more UK physicists needing to connect interactively to the US. This is however essentially impossible at present, leading to an unacceptable loss of efficiency and limiting the UK impact. Unless this situation can be ameliorated quickly the UK will not be able to make full use of its investment in the US experiments.

The PPNCG is currently in discussions with UKERNA to see if a short-term technical solution is possible. We have now held two meetings, and two possible solutions are being investigated (i) a limited form of prioritised routing (ii) extension of the managed bandwidth service which UKERNA are about to offer within the UK and Europe. Both of these methods would incur costs, and should we obtain a reasonable technical solution we will present details to the PPC for a decision on whether and how to go ahead.

3.3 Interactive connections: Quality of Service and Differentiated Service Piloting

We wish to draw the attention of the PPC to this specific item which will have important ramifications in the years to come. In the last item we pointed out the severe difficulties caused to the US experiments by the transatlantic link. The problem ironically has nothing to do with bandwidth per se. It is important to understand that an interactive connection requires very little bandwidth (100 Kbits/s), but requires special treatment of its packets so that a reliable and short round trip time is assured. It is known that low latency requirements on the transatlantic line can be implemented on very modest amounts of b/w. Therefore, for example, if only a small portion of the US link were reserved for such use it would have negligible impact on the main traffic, whilst satisfying our immediate needs. We have repeatedly stressed this point to the relevant bodies.

There is much discussion at present in the wider networking community concerning the related, but longer term issue of "quality of service" (QoS) of which "differentiated services" (DiffServ) is one proposed implementation. Briefly, the concept of QoS is the management of a set of network characteristics, i.e. loss, latency, congestion, jitter etc., so that a network service of known quality may be delivered. DiffServ does not guarantee bandwidth; it refers to the technique of aggregating traffic, i.e. packets, to receive particular forwarding behaviour at routers throughout the network, each such forwarding behaviour being based upon an associated resource allocation. At its simplest this may result in one aggregation receiving a higher priority than another, although a wide variety of services can be implemented using these building blocks. It is envisaged that the different aggregations would attract differing levels of cost to the end user.

The awareness of the need for this has been raised in the relevant bodies, and at least part of

this is due to both PPNCG lobbying and to Dr. P.W. Jeffreys, who represents the RCs on the ACN. As a result of this and other pressure from the RCs there is now a plan to pilot the techniques needed to offer QoS/DiffServ as a production service. Our best guess is that this may be available in 3 years time. The PPNCG is following developments.

3.4 Transatlantic Charging

On 1st August 1998 a system of charging all HE institutes for their transatlantic link usage was introduced to raise revenue to fund additional bandwidth and to moderate demand. While some concessions (e.g. overnight usage) were included, under the scheme all institutes are now charged annually for all their incoming traffic from the US, on a volume basis. It should be noted that in England and Wales the HEFCE subsidised these costs by £1M (of an expected £2M total) in the first year. So far most institutes have chosen to pay by top slicing. Consequently, even in the cases where bills have been passed down within institutes hitherto the typical liability for an HEP Group appears to be very modest. However, the balance of revenue raised against transatlantic link provision costs are subject to adjustment, because demand has not been significantly curtailed by these measures and, as indicated above, purchase of yet more bandwidth is planned.

3.5 The CERN leased line and HEPTEL

The HEPTEL service continues to offer a private telnet connection between the UK HEP sites and CERN and thence to DESY. It runs on a 64 Kbits/s RAL-CERN leased line and may be used as fallback circuit when interactive response over TEN-155 is poor. HEPTEL is also configured to route transatlantic calls across a private CERN/US link, of which we have negotiated a small share. This route to the USA, via CERN, has been essential during long periods of poor network performance over the normal transatlantic link.

In late 1998 some of the HEPTEL hardware situated at RAL was upgraded. A very old DEC Ultrix machine had earlier been adjudged unmaintainable and was replaced by a Viglen Genie2+ PentiumII PC with 256Mbytes memory and running Red Hat Linux. It is currently providing an identical service, although with enhanced security using TCP Wrappers, but work is underway to strengthen this even more through the use of the secure shell, ssh.

4. JREI/JIF

It is clear that there is a genuine need for networking and computer equipment in the future well beyond that which can be funded from current budgets. There now several new sources of possible funding such as JREI and GIF as discussed below. If these future JREI and JIF bids are not successful then it is not clear where the funding will come from. If the bids are successful there are serious implications on what happens afterwards and how this equipment is managed and supported in the most cost effective way.

4.1 JREI

We note that there have recently been two successful JREI bids for computer equipment:

1. The BaBar Collaboration was awarded £800K for data servers which is currently out for tender. The specification is for a 4GFLOP of cpu at RAL with 4GB memory serving 6TB of disk with 4 ‘large’ institutes with 2GFLOP, 2GB serving 1TB disk and 5 ‘small’ institutes with 1GFLOP, 1GB serving 0.5TB.
2. Liverpool were awarded £275 for a 300 processor Linux farm for LHCb Monte Carlo production.

More JREI computer bids are expected in the coming round.

4.2 JIF

One computer related JIF bid was submitted to the first round. This was tape, storage, cpu

power and networking to Fermilab in the USA for the CDF and MINOS collaborations. Some of this equipment would be sited at RAL and some in the institutes.

In the second JIF round there are expected to be bids from BaBar for cpu power to supplement their JREI bid and for networking to SLAC in the USA and from LHCb for disk and analysis cpu to supplement their JREI bid.

In the third JIF round there are expected to be bids from D0 and for a UK LHC Regional/Analysis centre.

The PPNCG has provided technical formulations for two possible contributions to JIF bids.

1. We have suggested provision of good quality ISDN video conferencing equipment for all HEP institutes. We have provided the technical and costing information required for this. This idea may be put forward in a (i) Manchester HEP group JREI bid and (ii) an LHC wide computing bid mentioned above.

2. We have published a roughly costed draft technical proposal to install at every University-based HEP site within the UK appropriate fibre and copper cable to provide it with direct access to the WAN at its point of Campus entry. This infrastructure would be designed to enable Groups to take early advantage of any QoS developments and would facilitate the configuration of virtual network paths between groups and Accelerator Laboratories and Regional Centres respectively. This has not been adopted as yet.

5. Other Group Reports

5.1 Computer User Forum (CUF)

The Computer User Forum (CUF) was set up at the end of 1997 by CNAP to improve information flow between the computing service providers, policy makers, users and experiments.

As a consequence, three meetings were held in the CUF series during 1998 and were chaired by Glenn Patrick (RAL) with representatives from most of the experiments with UK involvement. These meetings covered many of the current issues concerning institute computing, central computing and networking. Typically, they attracted an attendance of about 30 people from the particle physics community, split equally between the universities and RAL. The general feedback was very positive and meetings with the user community were felt to be an important element in developing computing strategies within the UK. In October, a special meeting was also arranged by Glenn Patrick and Peter Renton to assist in the formulation of computing bids for the Joint Infrastructure Fund (JIF), particularly in the area of data storage. This provided a mechanism whereby a number of potential bids were presented and some of the common elements and problems discussed.

All CUF meetings are publicised in advance using Hi-Phi, as well as a dedicated LISTSERV mailing list. A Web site has also been established to provide details of meetings and to store documents and presentations.

As from April 1999, Tim Adye (RAL) has taken over the role of co-ordinating the central computing facilities for particle physics and this will include organising future CUF meetings for the community. Future developments could include holding some meetings away from RAL to reach a wider audience. As we enter the LHC era, it may also be important to focus on specific themes at each meeting to reach a co-ordinated viewpoint on the requirements of the new generation of experiments.

5.2 System Managers Group (HEPSYSMAN)

It was reported last year that CNAP had mandated a small group to reconsider how to address the twin issues of receiving "information on future trends" and maintaining "[vital] contact with institutes on computing issues". This group took the opportunity to discuss options with UK

system managers at the 2-day meeting organised by Dave Kelsey on 19/20-May 1998 at RAL. The conclusion was that a system managers' group (UK HEPSYSMAN) should be formed, with direct representation on CNAP.

The May meeting was deemed to be the first meeting of this group, and a second meeting was held at UCL on 30-Sep-1998. Terms of reference were agreed at the UCL meeting and proposed at the Oct-1998 meeting of CNAP. The basic hope is that this organisation will provide a bi-directional flow of information between institutes, central laboratory sites and CNAP. The focus would be on local institute-based computing facilities, leaving experiment-related issues to be addressed by CUF.

A key product of the HEPSYSMAN meetings was the setting up of a mailing list and a Web page, complete with contact list. The mailing list in particular has already proved extremely useful.

At the May meeting institutes reported on their configurations and plans and highlighted their chief concerns. It was interesting to note the variety of problem situations and solutions, but a major common concern was clearly lack of manpower for a system management task that shows no sign of reducing. Most sites felt they had no option but to continue with at least two platforms and, in practice in many cases, three or four. Other pressing issues were seen to be the future of Linux in HEP, the use of AFS, and the cost of complying with collaboration choices of operating systems and/or software.

The two issues that have dominated discussions over the past year, however, have been video-conferencing and security. Video-conferencing is being largely addressed by PPNCG, some of whose members have given very useful presentations to the HEPSYSMAN meetings, whilst members of the HEPSYSMAN Group have been actively concerned in several aspects of computer security.

In the first place it was important to ensure that system managers themselves were aware of the new level of risk arising from the explosive growth of the Internet. Case reports of serious intrusions at UK HEP sites were accompanied by detailing sources of information on how best to guard against attack. It was also felt important to agree on the organisation of incident reporting and security policy for remote access. For the former, Andy Sansum (RAL) has agreed to act as the primary contact point, but it has not proved so easy to agree a set of remote access procedures which would be acceptable to all institutes and collaborations. The problem is, of course, that any security measures have associated inconveniences and a balance has to be struck between risk and hassle.

The main concern of security provision at system management level is to ensure adequate local support, a) in allowing for the system effort and other resources involved in implementing the necessary protective measures, and b) in obtaining the requisite cooperation of both individual and collaboration users. To these ends CNAP is distributing two documents: a letter to group leaders highlighting the new priority that must be given to local computer security, and a set of guidelines for users, explaining the new importance of computer security and the part that they must play.

6. Training

In previous years, Steve Fisher (RAL) has organised training in Object Oriented techniques for the UK HEP Community with ad hoc funding arrangements. This year CNAP assigned pounds 25K for him to spend on appropriate courses. In the Autumn two sets of introductory courses were given, each with a four day Introduction to Object Oriented Analysis and Design and later a 3 day Introduction to OO Programming using C++. Those who wanted to go further with C++ were offered a more advanced four day course. Courses were also run on Java (3 days), C++ templates and the STL(1 day) and a four day course provided on program development with Objectivity (an Object Oriented Data Base). The courses were all provided by an independent consultant except that the Objectivity company provided training on their own product.

The feedback provided by those attending the courses remains generally very encouraging. The introductory courses were especially popular and were always full, but the other courses have only had about half of the available seats taken. However those attending do then benefit from more of the lecturer's time.

We would like to request continued funding for this IT training but at a slightly decreased level. It turns out that experiments are generally hiding Objectivity so that only a few people need to be aware of it and the course this year was poorly attended in spite of broadcasting warnings that it would not be repeated. Otherwise it is hoped to offer courses similar to those offered this year.

A previous application to the PPARC Education and Training committee for support was unsuccessful, but a request will be made this year for funds in FY00/01.

The IT training budget requested by CNAP for FY99/00 is £20K.

7. Representation

User representation was reorganised last year with the Computer Users Forum having representatives from all experiment and the System Managers sub-group having representatives from all Institutes.

CNAP membership has also been reviewed to reflect the changing situation whereby BaBar is now a running experiment and the main future concern is the provision of computing in the LHC era.

We have recently appointed one more member to CNAP itself to be the identified point-of-contact for physics computing within the UK on behalf of the PPC, PPESP, CNAP and Director of Physics (RAL). The co-ordinator is responsible for monitoring developments with, advising on the formation of bids against, and aiding tendering and procurements resulting from various funds available to support computing initiatives (including JIF, JREI, PPARC Opportunities as well as awards made through the PPGC). The co-ordinator will report to each CNAP meeting and will give a status summary of recent developments, advice on how developments interface with the central computing provision made available through the SLA. The co-ordinator will be available to participate in experiment-based committees which are formed to plan bids or oversee the expenditure of funds resulting from successful bids.

8. Summary & Recommendations

The last year has seen a move away from commercial Unix systems towards more commodity computing based on PCs as well as a move from traditional programming methods to object oriented programming involving more commercial software such as large object oriented databases. Many more UK physicists are involved in experiments in the USA than has been the case for several years. We note the following:

- That the cost of providing central services is increasing.
- That the demand for data storage in the future is likely to far exceed what is currently provided or that can be provided through current budgets.
- That a number of additional sources of funds are now available but optimising the use of these fund over which we sometimes have little direct control will require careful management. An additional member of CNAP has been appointed to coordinate this very important area.
- That the UK and European bandwidth provision has been by and large adequate over the last reporting period.
- That both UK and international bandwidth provision is increasing rapidly, particularly in the UK following the governments priority to provide Internet access in all areas of education.

We may in fact have 2.6 Gbits/s SuperJANET provision within 2 years, with 10Gbits/s likely to follow. As a result the pure HEP requirements appear even smaller than before when judged purely as a fraction of annual integrated bandwidth usage. However pure bandwidth alone is not the only factor, and some of our users require more specialised services. We still remain largely at the forefront of such needs and accordingly concentrate our efforts on lobbying involvement in piloting where this is useful.

- That the transatlantic link remains essentially unusable for the main needs of our US experiment involvement. This is a serious problem and we strongly feel that this has to be ameliorated in some way if we are to make sensible use of our investment in money and effort in these experiments. The main requirement is for small amounts of protected bandwidth. We are investigating possible solutions and will report in due course.

We recommend:

- That the RAL DCI budget to support the central computing facilities and datastore be increased to 9SY and £250K for capital and recurrent spending plus £20K for OO training of UK Physicists.
- That the PCC takes note of the likely necessity to buy protected bandwidth to the USA.
- That wherever possible, CNAP be mandated to co-ordinate or comment on potential funding bids so that resources from all sources can be managed in the most cost effective way.

Appendix A Network Monitoring

Network monitoring, a strategic activity of the PPNCG, has been established for some time now for key routes to the US and Europe as well as sample connections within the UK. The normal state of the network from the UK site to the remote node is recorded and any changes, beneficial or otherwise, can be detected with ease. The objectives of this effort are threefold: to identify bottlenecks or other problems, enabling us to respond with some authority to HEPGRIPE queries; to enable us to deliver accurate evidence to the relevant reporting channels, in making representations to them; and to enable us to plan future requirements. The effectiveness of this work has gained the respect of the TAU and of UKERNA who consult our material in their planning and troubleshooting respectively.

The RAL monitoring suite generates data on packet loss and round-trip times which are collected at RAL, CERN and DESY. FTP throughput is also recorded. The use of the HEPTEL facility is monitored and gives a convenient indicator of our users' views on the performance of the networks. All of this information is presented graphically, via WWW, in a user-friendly format. Historical data are available for comparison and analysis of long-term trends. The "traceping" tool (developed by PPNCG) is used to monitor the intervening route. Again the most useful aspect of this tool is that changes in the running network are readily detected. Traceping data are collected and stored at Oxford but it also runs at RAL, SLAC, CERN, DESY, FNAL and Soudan. The traceping utility is portable and can be run from both ends of a link at a moment's notice, aiding a rapid analysis of network faults or routing anomalies. At the moment this tool is being re-written in Perl for greater portability and easier user access via WWW.

Currently monitoring is operated from RAL as a part of the ICFA Network Monitoring Initiative which seeks to characterize network performance between HEP sites around the world. A PPNCG development is currently underway to concentrate the various elements of this monitoring into a single machine which will also be used to host the PPNCG network monitoring Web pages. When complete this will allow PPNCG network monitoring to be based on software which is in general use across the world rather than on in-house solutions that are more costly to maintain. We are offering traceping for worldwide use.

Appendix B Video-conferencing

Video-conferencing is now an essential requirement of the HEP community. Currently two distinct transmission technologies are in use by HEP groups:

1. Dial-up ISDN circuits: Traditionally the equipment for these has been installed in a dedicated room within the institute. For more than two sites a central switch is required, dial-up connection being initiated at either end according to billing arrangements and operational policy. The bandwidth is therefore guaranteed, the quality is good and the conference is generally reliable. Within the UK tariffs have been roughly double those of equivalent voice calls, but BT, one ISDN service provider, has just announced significant reductions.

   Not all institutes with an HEP Group have such facilities and a major disadvantage is that rooms and the switch need to be booked in advance, frustrating spontaneity. Moreover for conferences involving US–based collaborations, time-zone differences may cause scheduling and operational difficulties in the use of institute VC facilities outside UK working hours.

   With careful selection it is now possible to purchase a set of videoconferencing components for between £2K and £3K which would be scalable from desktop (one conferee) through rollabout (several conferees) to many (room-based). The PPNCG proposes that each group should be provided with at least a basic set of such equipment, comprising a minimum set of

hardware components and an ISDN connection. This would facilitate participation of 2-3 persons per site in the regular conference meetings now typical of the new generation of experiments. A larger investment of between £10K and £15K would be sufficient to either scale up the basic kit or purchase additional equipment suitable for studio operation, excluding room alteration costs.

2. Packet switched connections: These use the Internet and may use the MBONE (a worldwide superstructure on the Internet) to distribute traffic. Bandwidth on the MBONE is severely limited and quality, both internationally and within the AC.UK domain, is barely adequate. Work is in progress to improve this infrastructure within the UK with the aim of offering an MBONE that can sustain business quality conferences using standards compliant hardware. In the meantime IETF-developed tools suitable for use in the existing harsh environment of the Internet (with or without using the MBONE) are freely available and have been adopted widely in other academic arenas. This method of VC has the great advantage of requiring almost no booking, and being free at the point of use. The equipment costs are minimal (£200-300). Several groups within UK HEP have successfully tested this technique for various small meetings, although it is limited to one conferee per site, and in each case much time has been invested in initial session setup. Such hardware may also support NetMeeting, a freely downloadable videoconferencing application from Microsoft which has proved to be very popular because of its accessible user interface and its support for sharing Microsoft documents (e.g. Word, PowerPoint). At present it is perhaps more widely used in HEP than the IETF tools, but only in two-site sessions. We have recommended that groups try these facilities, even if only for simple one-to-one or few-to-few ad-hoc communications (e.g.: student supervision).

Members of the PPNCG have researched and tested extensively both ISDN-based equipment and kit suitable for use with the IETF tools. As a result of this we have published and distributed a document recommending to HEP groups specific products, concentrating on low-end entry-level IP-based conferencing using the IETF tools. A complementary second document, reviewing current entry-level ISDN-based options and making specific product recommendations, is already drafted and will be published for HEP Groups very soon. A third document, likewise surveying standards-compliant IP-based products, is planned as this market matures.

UKERNA have instigated a large-scale project aimed at identifying service requirements to operate within the UK an MBONE capable of sustaining business quality videoconferencing. The PPNCG believes that internationally this multicast structure will be of rapidly growing importance to the HEP community and accordingly it is monitoring this project, which began in December 1998. It is already clear that MBONE performance is unpredictable, unreliable and, in the UK at least, under resourced. Within the context of the present project, the PPNCG is measuring parameters critical to good MBONE performance and is developing tools that address the unpredictability element. It has already discovered a number of quantitative performance anomalies, the data for which are being analysed.

Appendix C . CNAP Committees and Membership

C.1 CNAP

Name	Institute	Function	Rotate Off
S.L.Lloyd	QMW	Chairman	2001
T.Adye	RAL	Central Facilities/CUF	2002
P.Clarke	UCL	Chair PPNCG	2000
R.Cranfield	UCL	Chair System Managers	2000
D.Newbold	Bristol	Coopted (CMS)	2002
P.Renton	Oxford	Chair Data Store	2000
L.Thompson	Sheffield	Institutes Representative	2002
A.Halley	Glasgow	Computing Coordinator	2002

In attendance:

Name	Institute	Function
A.Coates	PPARC	PPARC
S.Fisher	RAL	Training
A.Flavell	Glasgow	UK HTASC Rep
J.Gordon	RAL	ITD
P.Jeffreys	RAL	PPARC SLA
D.Kelsey	RAL	Chair HTASC
R.Middleton	RAL	Secretary

C2 Data Storage Sub-group

P.Renton (Chairman), T.Adye, T.Folkes, N.Geddes, J.Gordon J.Hart, P.Jeffreys, G.Patrick, D.Sankey, H.Wittig, S.Wotton.

C3 PPNCG

P.Clarke (Chairman), R.Hughes-Jones (Secretary), G.Fayers, A.Flavell, J.Hart, P.Jeffreys, D.Kelsey, J.Macallister, D.Sankey, D.Terrett (Astronomy representative), F.Wickens

C.4 Computer User Forum (CUF)

ALEPH	R. Edgecock(RAL)	
ALICE	I. Bloodworth (Birmingham)	
ANTARES	L. Thompson (Sheffield)	
ATLAS	J. Baines (RAL)	S.L Lloyd (QMW)
BaBar	T. Adye (RAL) (Chairman)	A.J Martin (QMW)
CMS	G. Heath (Bristol)	G.N. Patrick (RAL)
DELPHI	J. Guy (RAL)	
H1	D.P.C. Sankey (RAL)	J.V. Morris (RAL)
LHC-B	N. Brook (Glasgow)	C.P. Ward (Cambridge)
MINOS	R. Edgecock (RAL)	G.F. Pearce (RAL)
NA48	B. Hay (Cambridge)	
OPAL	D. Charlton (Birmingham)	G.N. Patrick (RAL)
SNO	N. West (Oxford)	
SOUDAN	G.F. Pearce (RAL)	
UKDMC	N.J.T. Smith (RAL)	
WA94	O. Villalobos-Baillie (Birmingham)	
ZEUS	J.C. Hart (RAL)	

C5 System Managers Group (HEPSYSMAN)

Birmingham	L Lowe
Bristol	J-P Melot
Brunel	P.Hobson
Cambridge	J.Hill
Edinburgh	S.Gowdy
Glasgow	D.Martin
Imperial	G.Fayers
Lancaster	R.Henderson
Liverpool	A.Moreton
Manchester	A.McNab
Oxford	I.McArther
QMW	M.Landon
RAL	D.Kelsey
RHUL	S.George
Sheffield	L.Thompson
UCL	R.Cranfield (Chairman)

Appendix D References & Contacts

D1 CNAP web sites

CNAP	http://hepwww.rl.ac.uk/CNAP/
Data Storage Sub-Group	http://hepwww.rl.ac.uk/CNAPDS/cnapdata.htm
PPNCG	http://hepwww.rl.ac.uk/ppncg/
CUF	http://hepwww.rl.ac.uk/CUF/
SYSMAN	http://hepwww.rl.ac.uk/SYSMAN/

D.2 DCI web sites

CSF Farm	http://www.itd.clrc.ac.uk/Activity/CSF
WNT Farm	http://www.cc.rl.ac.uk/dci/services/hepnt/
DataStore	http://www.itd.clrc.ac.uk/Activity/DataStore

D.3 Other web sites

HEPCCC	http://nicewww.cern.ch/~Djacobs/Hepcccw3/HEPhome.htm
HTASC	http://wwwcn.cern.ch/~eauge/htasc/public/
HEPiX	http://wwwinfo.cern.ch/hepix
HEPNT	http://hepnts1.rl.ac.uk/HEPNT/
HEPpc	http://hepwww.ph.qmw.ac.uk/HEPpc/