EXECUTIVE OVERVIEW FROM EUROPEAN TRIP

Visits to CERN, DESY and RAL were made during Dec 1993 to continue discussions with colleagues on current activities and future plans for support of computing and networking. In particular we focussed on issues raised by the SLAC Computer Advisory Committee concerning the distributed environment and client server computing and how to migrate from today's supported systems to those we will need towards the end of this century. While in England, Cottrell took the opportunity to discuss the European Networking scene with ex SLACers at DANTE in England.

VM

All three Labs are downsizing their IBM mainframe operations. CERN hopes to phase out the service by end 1996, DESY is moving production work to Unix, RAL have downsized from an IBM 3090-600 to a 3090-400 and expect to continue this downsize as users move elsewhere. CERN has assigned one FTE in the CN division to coordinate the move from VM. L3 at CERN is looking at developing some tools to facilitate moving REXX execs to Unix.

VMS

CERN and DESY are "rightsizing" their central VMS clusters allowing for an increase or decrease in VMS CPU power in future. This is being accomplished by converting to scalable ALPHA based clusters and removing their VAX 9000 and associated HSC based peripherals. Both CERN & DESY expect to continue to support VMS in the medium term and have no timetable for terminating VMS support. Both labs are converting to SCSI based disks, and to providing VMS tape access via a Unix staging system based on code from CERN. All labs appear to have very advantageous software licensing relationships with DEC, akin to that of Stanford University and much better than that enjoyed by SLAC.

UNIX

We received strong confirmation that the move from mainframe computing to distributed computing is correct. In particular, the move to Unix is clearly happening elsewhere.

There are no winners yet on sophisticated batch systems for workstation farms. Other labs are looking to see if SLAC will provide leadership and repeat the success of the SLAC Batch Monitor.

CERN and DESY in particular have done some collaborative work on a consistent UNIX environment and interface. They recognize the need for site-wide coordinated user accounts and are working towards that goal. The direction at all labs is towards dedicated servers running a single service.

The management of software costs in the distributed environment is becoming a critical issue at all labs, requiring considerable effort to negotiate reasonable site licenses and investing in license server technology. With the rapid advances in technology, all labs recognize the need for more agility and flexibility in procurements. Vendor partnerships are being pursued at all labs in order to leverage skills.

Both CERN and DESY have are doing some work with MPP technology (CERN's Meiko, DESY's APE, Zeuthen's IBM SP-1) for their theoreticians.

Data Management

There is no clear winner or even contender for a commercial product for staging software. It is worthwhile to examine CERN's SHIFT software mounted on a high speed network (FDDI speeds or better) as a staging service for compute servers at SLAC.

The future of high density tape storage does not have a clear winner yet but SLAC is uniquely positioned with partnerships and/or proximity with Storage Tek, IBM, and Ampex.

Further confirmation was received that 8mm tape backup is not reliable for other than streaming I/O (such as backups). However it is important for importing and exporting data with other sites. RAL have had considerable success with DAT technology and strongly prefers it to 8mm.

Physics Applications

PAW is the analysis tool of choice by CERN and DESY experiments. There are many complaints about the reliability and speed of the COMIS macro facility.

GEANT users can expect to see a 2x improvement in speed in version 3.17 from geometry optimization.

Bulk processing is run on UNIX farms. At CERN, analysis is run on VMS and VM where access to tapes is easier. DESY does not have much of a VMS presence in the offline analysis environments.

CERN and DESY have become accustomed to data volumes and CPU needs far in excess of SLD's.

PCs & Macs

CERN is actively developing PC, Mac and Unix environments to enable them to manage large numbers of users in a cost effective fashion and provide services that are more attractive than VM. The CERN PC management environment appears to be the most advanced and is worthy of serious consideration at SLAC.

All Labs recognize the importance of central support for Macs and PCs and the necessity to provide coordinate sensible software licensing, recommend configurations, coordinate maintenance and training, provide backup services etc. Without this central role, they feel entropy rapidly increases and the costs of running the environment are hidden and become unmanageable.

Both CERN & RAL appear to have large support efforts for PCs and Macs, with over a dozen people at CERN and 6 at RAL. RAL also put together a task force to address the needs for Office Systems. They have made considerable progress and their reports are well worth reviewing for applicability at SLAC.

Networking

All Labs are actively increasing the capacities of their LANs. At the moment they are using FDDI to meet the backbone and major volume server needs. Both CERN & DESY have invested in DEC Gigaswitches for non blocking FDDI access. CERN, who got in on the early Gigaswitch deployment have found it to be very complex. CERN and RAL are also moving away from a bridged network to a routed network.

CERN and RAL are working with pilot ATM experiments to understand the next generation technology that they feel will most likely be successful in the second half of the 90's.

European wide area networking is less centrally coordinated than the US, and efforts to coordinate the networks European wide are in their early stages.

TABLE OF CONTENTS OF EUROPEAN HEP COMPUTING TRIP REPORT, Dec93

These visits were arranged to continue discussions with colleagues at CERN, DESY and RAL on current activities and future plans for support of computing and networking. In particular we focussed on issues raised by the SLAC Computer Advisory Committee concerning the distributed environment and client server computing and how to migrate from today's supported systems to those we will need towards the end of this century. While in England, Cottrell took the opportunity to discuss the European Networking scene with ex SLACers at DANTE in England.

  - DESY Hamburg Unix Services
  - DESY Zeuthen Unix Services
  - Unix for Zeus at DESY
  - DESY Unix Miscellanea
  - RAL Unix Services
- Data Storage Issues
  - CERN Storgae Systems, Data management, FATMEN, TMS, Unitree
  - CERN Robotics, Exabytes, and General Operations
  - DESY Data Management
  - RAL Storage Management
- Physics Applications
  - L3 MC and Data Processing
  - ALEPH Event Display Graphics
  - CERN CN Applications
  - PC & Mac Support
    * PC Support in CN Division at CERN
    * CERN PC/MAC Support Outside CN Division
    * DESY PC/MAC Support
    * RAL Office/INformation Systems Plans
  - Other Applications (Mail & WWW)
    * Electronic Mail at CERN, RAL, DESY
    * WWW Support at CERN
  - Local Area Networks
    * CERN LAN Support
    * DESY LAN Support
    * RAL LAN Support
  - European Networking
  - Appendix A: Full Itinerary
  - Appendix B: List of Persons Met During Trip
    * CERN People Met
    * DESY People Met
    * RAL People Met
  - Appendix C: Documents Obtained on Trip

TRIP REPORT SUMMARY

- Travelers:
  * Roger. L. Cottrell Assistant Director, SLAC Computer Services
  * Richard Dubois Physicist, SLD Experiment
  * Tony Johnson Physicist, SLD Experiment
  * Randy Melen Systems Specialist, SLAC Computer Services
- Dates of Trip:
  * December 4 - December 15, 1993
- Purpose:
  * To Visit CERN, DESY and RAL and learn about their activities and plans in computing and networking
- Estimated cost of travel:
  * $5000, funded by DoE

These visits were arranged to continue discussions with colleagues at CERN, DESY and RAL on current activities and future plans for support of computing and networking. In particular we focussed on issues raised by the SLAC Computer Advisory Committee concerning the distributed environment and client server computing and how to migrate from today's supported systems to those we will need towards the end of this century. While in England, Cottrell took the opportunity to discuss the European Networking scene with ex SLACers at DANTE in England.

GENERAL INFORMATION ON LABS

CERN Computing Organizations

CERN has about 3000 CERN employees, 500 contractors 5000-6000 registered users and about 6000 people on site any given day. They have about 2000 Macs, 2000 PCs, 1200 UNIX workstation and 250 Xterms. There are three divisions at CERN providing computer services. Two of these divisions (ECP & AS) have some functions which have nothing to do with computing.

The CN division, with about 150 people, provides infrastructure support and is divided into 7 groups described in the CERN computer Newsletter 214, Sept - Dec 1993 p1. These groups are vertically oriented. Each group has developers and front line services/operations sections. The operations group is left with only 10 shift operators after the recent reorganization.

The Administrative Services Division provides document handling services, database applications (focusing on Oracle applications) and desktop computing support for Macs and PCs (they have about 13-14 people supporting Macs & PCs). We did not gather information on their ADP activities, this should be a focus of a future visit.

The ECP division has 130 people providing computing support for data readout, acquisition, experiment controls and event simulation, reconstruction and analysis, within the 130 there are about 30 who are focussing in "Progressive Techniques" to improve software production. This includes 10 people in software technology and 5 people in information systems (in particular WWW).

DESY Generalities

DESY appears to be very similar to SLAC in size, though we did not get estimates of the numbers of central support people involved. They appear to be following a policy like SLAC's viz a viz the interrelation between the central Computing Services and the experiments. That is, their CS department supports the central computing and networking. It is up to the experiments to develop and maintain their software environments.

It is clear to DESY that the vast majority of Physics computing is moving to UNIX because the vast majority of computing cycles are there (though the vasy majority of Physics users may not be there yet!) DESY is not yet saturating their SGI CPUs. For example, Zeus now has 18 CPUs but us usually using about 15 CPUs.

At the other end are the Macs, PCs, and personal UNIX systems. The question here is what can central services provide?

Where does VMS fit in? A DEC Alpha cpu is approximately equal to an SGI CPU. The Alpha VMS boxes with 4 processors are about 20 CERN units each. VMS is used by the Synchroton Radiation group, a mostly-DEC shop. And Zeus has its own VMS cluster for online work. DESY expects the casual IBM MVS users to go to VMS while the serious cycle IBM MVS users will go to UNIX. This probably differs from SLAC where we expect those casual users (who just need email, printing, etc.) to go to PCs and/or Macs.

DESY notes that their mainframe use has gone from 100% down to about 80%. I/O is still a big problem. They are using SHIFT from CERN, locally modified, with Ultranet to do tpread/tpwrite via MVS with MVS-ported daemons, although they feel that is still not fast enough. The IBM ES9000 is a bottleneck as the robot drive controller.

DESY runs about 500 X-terminals. Many are in offices and are configured to get service from the same machine that is also an NFS server for the home directory of the X-terminal user, this cutting back on NFS file activity and network overhead.

They also have 40-50 Suns used for CAD with IDEAS, technovision, etc. As well, they are using an Italian highly parallel SIMD system, the APE Quadrix Q16 with 2 banks of 128 nodes.

VM & IBM MAINFRAME INFORMATION

CERN VM

CN Division CERN plans to phase out the CERN VM service at the end of 1996. The existing IBM ES9000-600 mainframe hardware lease runs out at the end of 1994 and the target is to move LEP batch production to the CORE (UNIX servers) by then. Then the mainframe with be downsized to reduce the mainframe expenditures by a factor of 2 to 3.

The remaining interactive load (mainly mail & editing) will be absorbed by workgroup servers supporting the desktop. These will be based on the existing Novell based NICE environment for PC and Macs and a similar environment being develop for UNIX.

They are setting up a task force to faciltate the VM migration and 1 FTE from CN has been assigned to it. They are currently defining the areas in VM that have to be moved and are monitoring VM to understand the current use. Email is recognized as a major component to be moved. For UNIX they are looking at Z-mail. They will be moving the Oracle servers on VM to UNIX. There appear to be no plans to support FORTRAN 90 from Oracle. CERN makes little use of the ProREXX interface so there are no plans to port it to UNIX. On the other hand Pierantoio Marquesini of L3 did not feel it would be hard to port ProREXX to Regina (see section on L3 plans for VM) on UNIX.

L3 Plans for VM

L3 has considerable political pressure to get rid of VM. They have about 40 simultaneous user who are mainly physicists. They feel it will be fairly easy to move these people as soon as they see the grass is greener (access to data is better) on the other side. Some educational effort will be required.

L3's main concern is the many service machines and EXECs which control their production system and provide the master database system. They estimate they have over 100K lines of VM REXX code. The code is stable and often untouched for several years. A largely automated port would be preferable to a buggy re-write. They are actively investigating moving this code to UNIX by using the public domain Regina 0.5H (REXX implementation on UNIX). To facilitate this, they have: developed an RXSOCKET interface; implemented about 60% of the HEP REXX functions for Regina; and are developing automatic conversions for EXECIO and GLOBALV calls. They have found it easy to develop external functions for Regina. They are investigating how to adress IOS3270 (block mode full-screen 3270 interactive access), SPOOL, CMS PIPES, & XEDIT functions, possibly utilizing curses for the full screen emulation. They estimate that it will take item 2-3 man years to move, on top of the additional FTEs providing normal support and development in their Unix environment. They aim to accomplish the move by mid 1995.

They do not expect end users on UNIX to use REXX (Regina), rather they expect them move & move to use UNIX tools commonly used in the Unix community outside HEP (such as Perl).

There is a SENDFILE emulator out there that can simulate spool files from a Rick Trott. L3 uses CURSES to emulate XEDIT macro usage.

DESY Mainframe Down Sizing

In 1994 DESY will downsize their IBM 9000 mainframe running MVS with the production work moving to UNIX in 1995 and 1996. They have invested heavily in SGI Challenge machines with a total of 84 processors. They expect the "serious" users to move to UNIX and the casual users to Macs/PCs, though there is a faction at DESY that feels the casual users should move to Xterminals of which they already have 600, rather than to PCs or Macs.

VMS SERVICES

CERN VMS Services

CERN is 'rightsizing' their VMS cluster to eliminate their VAX 9000 with its HSC and RA90 disks. They are being replaced with 72 GB of SCSI disk in a DEC Storageworks cabinet and three Alpha workstations. The use of individual Alpha workstations will allow them to scale the cluster size up or down in future as appropriate. Three FTE's support the two VMS clusters, VXCERN and VXENG. CERN is happy with their Alpha 7000, finding it very reliable and easy to manage.

They found that Alphas behave well with 128 MB of memory, allowing them to support at least 75-100 interactive users per alpha with no noticeable loss of response.

Access to tapes is via network copying from the SHIFT (UNIX) farm: tape requests go via the TPREAD utility to the SHIFT machines which copy the tape to VMS disk using CERN's RFIO package and an FDDI network. Access to tapes are provided throught the channel-connected rs6000's already in use with the SHIFT farm. This system appears to be portable and could be used at SLAC.

CERN plans to continue to operate their central VMS cluster until a year after LEP experiments are terminated (appproximately the year 2000). They do not believe thee DEC support for VMS will be a problem on this time scale.

CERN has written a utility, called SPACE, to aid in the management of disk space within a group. It provides an easy way to transfer disk quotas between users within the group without having to provide system privileges to the space administrators. This should be investigated for applicability at SLD.

DSNLINK is used to get DEC code fixes online.

A VMS users' guide has been written.

Other HEP sites are saving large amounts of DEC software maintenance fees since DEC makes them eligible for their campus software license agreement. SLAC spends on the order of $240,000 annually on software updates for Ultrix and primarily VMS. The Stanford Educational Software License fee could be shared (like the Sun VolumePac, Transarc, and Oracle site license agreements) for probably $40,000/year or, worst case, SLAC could have its own ESL (like the IBM HESC) for about $80,000/year.

DESY VMS Services

DESY are in the final stages of adding 4 Alpha AXP3000-500's workstations to their VAX cluster. Once these are fully commissioned they will remove their VAX 9000 and the associated HSC's and disks. They will use a GigaSwitch to ensuring good connectivity between the workstations in the VAX cluster. DESY has no timetable for phasing out VMS support. (DESY currently supports up to 250 simultaneous users on their VAX 9000, and have not seem problems).

DESY will use CERN's RFIO and SHIFT tape staging techniques to get tape access to the VMS cluster. DESY has a well layed out network arrangement that provides access from the VMS clutser's Gigaswitch to tape through a Gigarouter (from a company called NetStar). This GigaRouter also forms the hub of their Unix cluster, connecting to their SGI challenges using HIPPI.

They warned us VMS 6.1 will be the last version of the O/S to support DECNET Phase IV.

RAL VMS Services

RAL are upgrading their VMS services by introducing two DEC Alpha 7000's one with 3 cpus and 756MB and one with 1 cpu and 512MB. These machines will be upgraded to higher speeds (180MHz to 200MHz in Jan-94 and to 275MHz later in 1994). They will include access to 100GB of disk via a DEC StorageWorks system. RAL considers the multiprocessor Alpha 7000 to be a very efficient, easy to administer way to couple cpus together.

RAL will keep at least one VAX running VMS for legacy VMS applications that may not be ported to Open VMS.

DEC won the hardware and software maintenance contract for RAL. RAL is eligible for the very attractive DEC campus agreement for software.

UNIX SERVICES

CERN UNIX CORE (batch farms)

With regards to workstation farm service, CERN notes that it depends on the data rate required. With a Gigaswitch and a FDDI ring, you might get 8 or 10 MBps. The advantage of keeping CPU and disk servers separate is to optimize each. Some architectures are better for CPU-intensive work, others for I/O. DEC Alphas are just now becoming reasonable as disk servers. Doing 8 streams together , you can possibly get 5-6 MBps. Using remote file I/O for tape (rfio's tpread), you can get 800 KBps for a single stream.

CERN requires 5-10 MBps to the net for disk service, CPU service and eventually for tape service, with peak aggregates of 25 MBps. Four years ago this required Ultranet but now you can almost do that with segmented FDDI and a Gigaswitch. Ultranet as a company is still shaky but probably still 1-1/2 to 2 years ahead of competitors. There have been management problems in the company plus FDDI muddied the market for them, they spent too much time on the Cray interface and were slow to implement workstation interfaces, and finally they charged too little for their major effort, their software.

CERN is waiting for ATM but looking at FCS. The good news about FCS is that vendors other than IBM are participating but the bad news is that most see this simply as a way to connect disks. HIPPI is suddenly becoming interesting and cheap serial HIPPI cards are coming. It appears to be a good short term alternative to Ultranet.

Five architectures are currently supported. They have at least 1 person per architecture plus a backup person. They have found it important to keep systems configured similarly. In the area of accounting Gordon Lee has been talking via email with Lois White here at SCS. CERN uses standard UNIX accounting and massage the data into Excel spreadsheets for color plots.

They found linking was very slow with NFS so they use rdist to keep local copies of libraries. This is managed by each experiment.

The architectures and their support are as follows:

1. SGI – This is a good multiprocessor architecture, scalable up to 36 CPUs per system and able to handle TBs of data. Graphics is strong, good I/O, and a good SMP design. Probably the best choice. However there is flaky local SGI support (no spares, export license problems, etc.) and SGI was late in delivering the 150 MHz CPUs. They are currently weak on the midrange.
2. H-P – These machines have excellent price/performance. The I/O is expensive though and they have no multiple CPU design. HP-UX is rather old at the base level but has good tools. The hardware and software are reliable but there is a long response time on bug fixes.
3. Sun – Until recently, they had no big machine. They do have very cheap desktops and are widely used as a reference machine by software developers. However Solaris is a "big pain".

4. DEC – DEC really had nothing until to offer until the Alpha arrived. Their disk systems now appear to be good products with good prices. CERN does not have enough experience with DEC-OSF/1 yet but have been hearing about software immaturity.
5. IBM – IBM has had a good relationship with CERN and has done joint projects. So far they have only been used as tape servers but soon will be doing disk serving too. They seem to have very good TCP/IP support.

RAID disk technology at CERN is/will be only used for home directory reliability.

CERN also has a Meiko MPP system that uses 32 Sparc processors. The application can write into local memory to communicate with remote nodes. It seems to have a better network than the IBM SP-x products. It runs Solaris 2.1 on local nodes.

With respect to AFS, CERN expects to move home directories from NFS to AFS. Right now, all "Physics data" is in NFS but they expect to have it all AFS-accessible. They do not use automounter in CORE; rather they have some hard and soft NFS mounts instead. At present, the public home directories are on Sun systems and others on SGI. It is not clear how AFS will work with large Physics data sets and its performance compared to rfio. The AFS token lifetime for batch is a problem to consider.

CERN uses NQS. This is historical because of their Cray and Ultrix systems. They ported NQS to Ultrix and acquired NQS internal expertise. They added some enhancements (e.g.., cluster, portable interface that also runs on VMS and VM, limits on the number of executing jobs per user). Christiane Boissat is their developer.

CERN UNIX (interactive systems)

CERN's customer base is about 1400 UNIX workstations across 7 supported architectures (Sun Sunos & Solaris, Apollo, HP, SGI, IBM, DEC Ultrix, DEC OSF/1). They also support 250 X-terminals. They have one specialist for each of 7 architectures plus one for AFS and one for printers. Staffing is 6 staff plus 2 contractors.

Software is centralized except for Ultrix and OSF/1. A DEC "campus contract" is being negotiated. They expect to pay 50K SFr to buy in plus 400K SFr annual maintenance for 1000 workstations.

General services offered to clients include:

- security advice, CRACK service
- backup service and advice (they use ADSM, had a joint project with IBM on WDSM)
- AFS service, gradually being expanded and experimenting with ADSM
- DCE/DFS tests (trying IBM, promised from HP)
- management tool testing with FullSail, Patrol, Tivoli (though expensive)
- COSE desktop study (VUE evaluation)

Their current challenges are:

- not enough staff as they try to build up IBM RS6000 support, SGI support; they are trying to buy contract support for relatively mechanical jobs
- too much time answering mail/phones and not enough producing guides and doing development work
- developing a standard UNIX environment in collaboration with DESY

A new challenge is that AFS is becoming strategic to them. For example, they developed the ASIS system to do software distribution. It started as a "pull server" and then became an NFS mount server. The next phase is to move to AFS and not pull but point.

They have no site-wide NIS domain. Therefore they will go through some conversion moving to AFS if they want a single user name space (as presumably they will want for a CERN-wide cell).

They have bought reference machines for each platform. These are unmodified entry level machines for standard installation testing and to generate standard binaries for AFS serving. They can be used to check out problems in a standard vendor-supplied environment. They are entry level and not for general use; these machines have their own NIS domain.

Applications support is done from other groups plus some individuals in CN.

For disk quotas, they have no quotas on home directories except for one specific system and for AFS. (The Novell systems have a quota of 50 MB.) They are considering a 100 MB quota for AFS but have not thought it through yet. They feel very strongly that disk quotas should not be an issue for researchers.

Obviously VM migration will have some effect on interactive UNIX systems. They have about 4200 VM users/week and 300-400 VXCERN users/week. Physicists think they'll go to UNIX based on what the collaborations are saying. CERN is driven by what each national lab does and the influence those labs exert on specific experiments.

CN offers the NICE environment, a "Novell Netware club" with Novell file servers for PCs, and CUTE, the Common UNIX and X-Terminal Environment. CUTE will offer AFS home directory servers with backups and hierarchical storage management, AFS ASIS readonly binary servers based on Reference Machines, and an AFS General Staged Data Pool.

DESY Hamburg UNIX Services

UNIX began at DESY with graphics workstations, originally storage tubes and IBM systems, then about 25 Apollo systems, then HP 9000/700 systems and then SGI systems. There are 8 HP model 735 systems for public use and about 11 model 730s for specific use. They are moving from 730s to 735s. The SGI systems started with MIPS 3000-based systems, originally 6 systems and a 7th for public UNIX service, each with 6 CPUs. These became SGI Challenge systems. There are a 7 systems with a combined total of 84 MIPS 4000 CPUs, most using the faster 150 MHz chip. There are 2 public systems (called x4u), 1 system for Hermes, and 2 systems for H1 (called dice), all in the R2 YP domain, and 2 systems (called Zarah) plus 50 DEC Ultrix system for Zeus in the Zeus YP domain. They are merging the R2 and Zeus domains. In total there is about 400 GB of disk. There are also 2 Ampex robotic tape systems each attached to one experiment and with 3 drives each. The Ampex equipment had only been in production a few months as of early December, 1993. DESY also has perhaps as many as 600 Xterminals deployed. They started with Tektronix Xterminals, had problems, and now have mostly NCD terminals. Most of 15" monochrome, perhaps 10% are 17" color, and just a few are 19" monochrome.

They are testing amd at the Hamburg site, using automounters from the vendors. They are also running some font servers for X. The R2 group has about 12 people doing support – 4 or 5 for special purpose, 3 or 4 on networking, 3 for SGIs, 2 for HPs, and 1 for Apollos. They have a standard lockout screen that lets you choose a maximum of 20 minutes for lockout. After that someone can use the console. When logging in from an X terminal, they try to xhost to the same machine that has the home directory of the user to reduce network traffic and latency. They are not using AFS but seem to be interested for the future. They are not doing any file sharing between UNIX and MVS and noted that MVS NFS performed poorly. Like SLAC, they don't have the staff to support all of the public domain packages and enlist the volunteer help from their users for such things.

DESY Zeuthen Unix Services

At the Zeuthen location near Berlin, they have about 150 people with about 70 of those as scientists and 10-20 in the Computer Center. The location is about 300 km. from the Hamburg location, 5 hours of travel no matter which mode is chosen. The Zeuthen location is a complete IBM shop now, having already moved from a VM environment on an IBM-compatible mainframe in 1991. It took about 6 months as that was considered "sudden". They also have HP and SGI systems with some Sun, IBM, and Convex machines. As well, they have about 10 Macs in the H1 experiment and another 50 PCs on the network. Their net is Ethernet with TCP/IP plus some AppleTalk and DECnet for a VMS cluster. They use a FDDI ring between the SGI file server and the computer servers (HP, SGI, and a Convex C3210). They have about 100 GB of SCSI disk for data. They consider it to be cheaper that robotic storage right now though they expect to have cheaper prices from robotics for the TB range in the next 1-2 years. For now, they move data with 8mm tapes though these experience some problems some times. They have about 70 X-terminals.

Zeuthen expected to install an IBM SP-1 with 10 nodes in December 1993. Why SP-1? They needed parallel computing for the theoreticians and they wanted to investigate parallelizing Monte Carlo simulations. They really needed a large mainframe such as a Cray T3D or Convex. They were using part of an external Cray Y-MP and needed their own machine.

UNIX for Zeus at DESY

A view from the Zeus project was given by Till Poser. Zeus sees central services are needed to provide processing and reconstruction of data. They use 18 SGI processors for batch and 18 SGI processors for reconstruction plus about 150 GB – 60 GB for data storage, 25 GB for

Monte Carlo, and the rest for staging, spooling, etc. They are not using Fatmen. Zeus is using rdist right now to keep libraries in synch. They use CMZ for code management and it takes most of an FTE to incorporate code from collaborators. Monte Carlo production should officially be done at outside institutions. However, they have developed a tool to scavenge workstation cycles over the Ethernet though they need to make sure those workstations have adequate memory. The tool is called "funnel", is runs from an SGI "funnel server", and they could probably handle 250,000 events/week though right now they are actually doing about 2500 events/week with just a few workstations. They will use Oracle on UNIX to record run information.

DESY UNIX Miscellanea

IP addresses are assigned within a few minutes and name servers are updated each night. At the moment they do not keep a central database of IP address information.

DESY has done a lot of work to develop a consistent UNIX environment. Details made be found in WWW from the HEPix conference held in Pisa in 10/93.

They have been using Legato Networker for backups but are planning on moving to IBM's ADSM (though not at the Zeuthen location).

They are acquiring LoadLeveler, presumably for the SP-1 at Zeuthen, but would like the HP version as well. They have used an NQS-like system from Convex called Cxbatch for their Convex system. They have found a German supercomputing company called Genias with a package called Codine. It seems to be similar to LoadLeveler, possibly propagated by DEC, and has ports for HP, Sun, IBM, DEC, and SGI. Volkswagen is using it. A major advantage might be that it seems to not be priced based on number of clients. Their cost is about 20K DM the first year and about 8K DM maintenance each year thereafter.

X-terminal support on the SGIs seems to need about *MB of memory for the first window and then 1.8-2 MB of memory for each additional window. DESY will be adding 1/2 GB of memory to both of the x4u public SGI machines, bringing them up to 1GB each. DESY uses the SGI Varsity software licensing program and also an academic DEC software licensing program (equivalent to Stanford University's ESL). Each SGI Challenge machine has 3 Ethernet interfaces (maximum of 4) with 3 IOPs. Each SGI Challenge has 8 Fast/Wide SCSI disk drives. They are negotiating a maintenance contract right now.

RAL UNIX Services

RAL has built a copy of the CERN Computer Simulation Facility (CSF) from 6 HP 735s. This is a low I/O, high cpu farm application used and paid for mainly by the HEP people. In the farm is a file server (on one of the 6 HP 735s) which holds the home directories. Each machine has 48 MB of main memory and 1.4 GB of disk for the OS and file and swap space. Job scheduling is via the CERN/NQS system. There is also a 5 CPU Alpha 3000 OSF/1 cluster each node having 64 MB of main memory and 2 GB of disk and one of the nodes acts as a file server with 24 GB. This cluster is mainly used for batch. There is no support for memory limits, at the moment, but this is considered to be important. LoadLeveler is not available for Alpha OSF/1 but DEC has a similar product called the Load Sharing Facility which is also Condor based. Accounting does not work well for the OSF/1 cluster, at the moment. There is no policy based batch scheduling or reporting at the moment. Upgrades to the OSF/1 cluster will be based on user demand.

RAL's current plans are to skip AFS and go directly to DFS. They have DFS for the IBM RS6000 and hope to have it for OSF/1 next year. They may become an AFS client of ASIS at CERN. On the farms they hard mount NFS filesystems and soft mount NFS elsewhere. They do not use NFS quotas they are considered to be too easy to bypass, it makes disk space usage less efficient and users are not demanding it. There are 2 GB of home directory space for users. The biggest users are moved out to other partitions as they are identified.

DATA STORAGE ISSUES

CERN Storage Systems, Data Management, FATMEN, TMS, Unitree

It was noted that remote tape is troublesome because recovery by tape repositioning is hard to do, while disk is cheap and scalable so copying a while to disk seems much better than reading/writing tapes from an application. The idea is to have tape servers that respond over the net to any tape request. So far, such function's performance has been benchmarked at CERN at about 1.0-1.5 MBps, while about 800 KBps is actually seen using the "parallel channel" interface, probably because of the high level of interrupts required for such an interface.. The ESCON interface is faster though CERN had not yet measured it. They expect actual performance with ESCON to be 1.5-2.0 MBps rather than the current 800 KBps. CERN uses 32 KB blocks with FDDI on IBM RS6000s. They note that the ESCON interface requires a 5xx deskside model, more money that a 3xx desktop, and that the ESCON interface itself is quite expensive. They have lots of parallel channel tape drives and expect to move to the IBM NTP tape drives with Wide SCSI when they are available. They are rated at 9MBps. Their testing with a DEC Alpha doing disk to net gives about 6 MBps and expect tape to be a bit better.

The SHIFT software is currently available from anonymous FTP and is installed at places like DESY, FNL, and IN2P3.

TMS has proven to be useful. It is written as a SQL-DS application originally but has now been entirely rewritten in IBM C. It runs on a separate IBM 4381 but has been ported to Solaris 2.x and oracle 7 on a Sun and is awaiting testing and tuning. TMS will let you have write-protected tape volumes and will let FATMEN do dynamic volume allocation. Both DEC and IBM will have competing products; DEC's will be MLM and IBM's will probably be based on ADSM. Concerns were also expressed about the usefulness of NSL Unitree. It was pointed out that it cannot backup the metadata while the system is up, that a default maximum of 8000 tapes is in effect without changing the source, and that a recent review of such system on the net showed Unitree as a poor choice. CERN is working on new stager code. It will start a daemon to start the copy process and watch it. If the staging disk fills up before copying is done, it will suspend the copy, do disk space garbage collection, and then resume the copying. The staging software at CERN knows how to deal with SL tape.

IBM will be demoing the NTP technology at the IEEE MSS Symposium and probably at CHEP '94 in San Francisco. CERN expects it to be available in mid-1994. NTP has logical volume support. The pricing and connectivity of NTP will determine the market size. Competitors to IBM's technology are be STK and DEC with something called DLT. CERN is buying DLTs to replace 8mm tape backup systems. They're rated at 1.25 MBps and have been measured at 1.0 MBps.

CERN is curious about our STK experience and our future with D3 technology. They have one experiment that is going to collect data on a Sony DL21000 D1 system.

CERN Robotics, Exabytes, and General Operations

The PDP group now encompasses the robotics and tape vault operations with 10 shift operators plus 1 operator for consumables plus 25 contract operators to handle manual tape drives. CERN's involvement began with an IBM Joint Study and an acquisition in 1988. They expect to upgrade to the IBM 10 GB SCSI linear serpentine cartridge technology that is rated at 7-8 MBps. Right now about 50% of the tape mounts are manual. Tape data is staged to disk and erase later when disk space is needed. Right now the robotics are controlled by VM but they have just received the RS6000 software. The current bottleneck is in the control unit, capable of a total of 6 MBps throughput even though each drive can do 6 MBps.

CERN notes that using 8mm Exabyte tapes is taking a step backwards in reliability and performance. It is really only useful for backups on standalone systems where the tapes will be read again on the same tape drive or for transporting data to other places that have inexpensive tape drives. They have a user-operated station for copying 3480 cartridges to/from 8mm tape. The use an 8500 drive, because it is more reliable than the 8200, without compression. They use two DECstation for self-service. For robotics, they also have 8mm tape drives attached to an IBM channel via a SCSI-channel converter. They find that 8mm tape drive heads wear out about every 2 month and so they expect to replace tape drives on a regular basis. The had tried the Summus tape carousel and found it was a poor choice; the Exabyte carousel worked well and it was easier to replace a tape drive yourself when it became necessary.

DESY Data Management

DESY is interested in the Lachmann software, rather than Unitree. They want a central data repository that conforms to the IEEE Mass Storage Model. DESY expects to use a Gigarouter from NetStar to connect various media together (HIPPI, FDDI, etc.). Phase 1 now has the IBM ES9000 as the silo controller. Phase 2 will disconnect one silo from the IBM ES9000. They are shopping for a sophisticated Hierarchical Storage Manager. This is an HSM that need to be able to intelligently choose between STK and Ampex. They feel Ampex will be good for sequentially accessing large data sets with a faster search, though they only have 6 drives. STK will be better for access to smaller amounts, with 36 drives.

DESY has been very happy with their SGI machines. The SGI Challenge has a 1.2 GBps bus, has IO processors that can do 320 MBps, and can have 32 SCSI busses off the IOPs without going through the VME bus. If striped, they can get a read rate > 11 MBps and a write rate > 7 MBps.

Martin Gasthuber discussed hierarchical storage management. DESY has chosen OSM (Open Storage Manager) from Lachman. The concept is that data is produced with intelligent controllers and a central "bitfile" server. The HSM discovers the most recent copy of your data, talks to the Storage Server, and then a direct communication occurs between the Storage Server and the client making the request. The client would have a Migration Filesystem on top of a standard filesystem to make secondary storage appear primary. The Migration Filesystem is the typical client of the Storage Server. OSM clients can be NFS, AFS database access, or even Fatmen. DESY has a license and the package has arrived. They noted that IBM Adstar has also licensed OSM.

Michael Ernst then discussed their Ampex tape system. DESY is quite happy with the system and probably will not buy any further STK systems. With new software, they expect to do better than 14 MBps on reads and writes. One question that always comes up is tape wear, tape re-readability, and head wear and replacement. Michael said that a head appears to be good for about 1000 tape/head contact hours. This is equivalent to reading or writing about 20-30 TB of data. They have also test readability of tapes and found that they could reread tape more than 1500 times with no problems. When the bit error rate begins to climb, they clean the heads. If error rates are still a problem, then they change the head assembly (8 heads). This turns out to take about 15 minutes and is self-service – no Ampex technician needed. Such a head assembly costs about $2500.

RAL Storage Management

RAL has developed a data storage with migration facility. Currently it is based on an IBM 3090-400 running VM which incorporates STK Silos with over 100 GB of disk staging space. On VM they modified the LINK command to load the data from tape to disk space if necessary. Much of the impetus was to minimize the number of drives required in the STK Silos. No user tapes are contained in the Silos. Data is moved from the Silos to DAT tapes when space is required in the Silos. Network access to the Silos is via 3 RS6000's with channel interfaces connected to the VM system.

RAL looked at various tape technologies for archive purposes. They had a Metrum SVHS 1/2" 6 TB Robot on site for evaluation but were concerned about reliability, expense single vendor support and rejected it. They compared DAT and 8MM. The costs were about equal (though the media costs were greater than for SVHS). They Chose DAT since it turned out to be much more reliable and bought 12 DAT drives. These are operator serviced with one mount/drive every 3 hours (based on data rates). They looked at stackers to possibly reduce operator intervention however it was not attractive costwise compared to simply increasing the number of drives. RAL has one 8 mm drive for compatibility purposes. The DAT data rates are (slower than 8 mm) 183 KB/sec today, will increase to 366 KB sec and are soon expected to go to 510 KB/sec. The DAT tapes are 90 meters long and hold about 2 GB/tape. RAL is copying about 30K 3420 type tapes to DAT tapes. The reason for this is to get rid of the older disintegrating tapes, to enable the data to be accessable to current devices, to reduce the floor space for storage needed by probably two orders of magnitude, and to record the data in a well defined DAT format to make future retrieval easier.

RAL are looking at the new IBM Digital Linear tapes, which look attractive when compared to STK. Particular STK concerns are maintenance costs, uncertainty about their ability to deliver the helical scan drives, and the reliability of the helical scan drives.

PHYSICS APPLICATIONS

L3 MC and Data Processing

L3 uses GEANT. They maintain their own copy, which they keep frozen for up to a year at a time. They try to not get more than a year behind the current CERN version.

It takes about 3 months after taking the data for them to derive the constants. Reconstruction is done on Apollo DN10000 workstations connected to their IBM 3090 by a channel-VME interface. The recon output is streamed three ways RECON: DST:miniDST at 200:20:3 kB/evt.

Data quality monitoring is handled by a single person at a time during the run. The feedback from reconstruction takes about 2 days to get back to the experiment.

About 1/2 of the group uses PAW for analysis. The COMIS facility is thought to be slow and too buggy to depend on for final analyses, which are performed with compiled code.

They do not have a very good handle on how they use their disk space (cf D0 in our FNAL trip report).

ALEPH Event Display Graphics - Hans Drevermann

We discussed his 2d event display package, DALI. Notable features are the 'v-plot' for showing 3d effects of tracking information in a 2d plot and a novel method for mapping 3d structures like projective towers of a luminosity monitor onto 2d. He has also studied optimal use of colours and colour combinations, and compression of scale dimensions. These are all reported in lectures he gave at a recent CERN summer school on computing.

He suggested we try using a plot of phi vs theta for looking at CRID data. A ring from the liquid would appear as a vertical line.

CERN CN Applications

GEANT version 3.17 is expected to be 2x faster than its predecessor due to improvements in the geometry optimization, which will become automatic in that version.

They are looking into shareables for UNIX plus dynamic loading. This is intended for people, once they are happy with their COMIS macro file, to compile them, link them in interactively and run the faster, compiled code on their ntuples. This is being developed for HP so far.

Tony Osborne wanted to make it clear to us that CERN CN is not an infinite source of manpower and, that if we wanted features, etc from them, we would be expected to contribute by collaborating on projects.

PC & MAC SUPPORT

PC Support in CN Division at CERN

CERN has put into place the NICE (Novell Integration Cooperation and Evolution) environment for supporting PC's This is supported by 4 people from the CN division. The goals of NICE are that: each PC is based on an operating system (currently Windows 3.1) that is identical; each PC has access to all facilities at CERN; file services are provided by highly reliable, easily managed servers.

NICE provides centrally organized home directory servers (at the moment there are about 20) plus 5 binary applications servers. The Windows 3.1 operating system is also stored on the applications servers, not on everyone's PC. This allows the whole environment to be maintained centrally. The applications servers are replicated each night. The servers incorporate SCSI disks with mirroring for availability and central backup is provided. There are about 1000 users. Every user runs in the same consistent environment so they can easily move from PC to PC and debugging is simplified. Configuration files which contain hardware specific information are constructed automatically.

They have site licenses for many MicroSoft applications and are looking at license servers in particular a product from Funk Software called AppMeter. They actually don't want to control licenses, but rather keep a count of usage (eg. # copies simultaneously in use) so they can buy more licenses as necessary. They have looked at the next version of Windows, Chicago, at the Microsoft developers conference and doubt they will consider OS/2 any further.

Dave Foster stated that the consistent environment and limited choices of NICE could be used for Unix support. It is not a technology issue but rather a cultural issue that has lead to the bizarre (in Foster's view) requirement to support 6 Unix platforms at CERN. Dave proposed a radical vision of desktop PC's for all people with a number of symmetric multiprocessor high performance Unix file servers and some Unix applications servers.

CERN PC/MAC Support Outside CN Division

The Administration Support Division (ASD) has 13-14 people supporting Macs and PCs, including external service contract staff and administrative and clerical staff. Of these there are about twice as many providing Mac support as PC support. THis is due to the lack of a NICE environment for the Macs. They estimate, however, that with there about 2000 PCs and 2000 Macs at CERN and across the CERN site that including AS-DC support staff there are about 24 full-time people dealing with PC support (divisional Novell managers plus local support staff and some service contract personnel). ASD use the CN provided NICE environment for standard PC application distribution.

Software distribution is a problem for Macs. In order to introduce a centralized NICE type structure, an Ethernet is needed on every Mac (using them with a FastPath is too slow). However the CN network group are concerned about adding an extra 1000+ Macs to the CERN Ethernet.

The ASD Mac and PC support group look after software applications procurement,licensing and installing on the binary servers (consulting with the CN NICE section) for the whole site. They have MicroSoft Select for standard PC and Mac applications and are working with Apple to get site-wide agreements for MacTCP, MacOS and MacX.

There are 2 FTE's in ASD who look after purchasing hardware. They keep about 100KSFr of spare parts plus a few current machines which they sell with a 5% markup. The ASD people receive, checkout and install standard hardware and software. PC installations are done over the network. Macs are done locally. Cabling, etc. is looked after by the CN network group.

For mail on the Macs they have about 1200 QuickMail licenses supported by about 25% of an FTE. They use the StarNine gateway product to provide access to and from QuickMail recipients. For the PCs they chose MicroSoft Mail over ccMail. They are running an AppleShare file server on a Sun using EtherShare/Helios software which includes shared printing support.

DESY PC/Mac Support

DESY has about 160 Macs most which are Ethernetted, and 600 PCs of which about 150 are Ethernetted.

For Macs they have about 25 AppleTalk II zones extending to DESY-Zeuthen in Berlin. They have several AppleShare servers, one is on the central VMS machine using Pathworks for central distribution of software and documentation. Macs are also used from home via local modems and AppleTalk Remote Access to provide connectivity to a networked MAC on-site.

They have a sitewide license for MacTCP and are working on a site license for Microsoft products. Training is via outside commercial parties. Backup is provided for the central servers using Retrospect and DAT drives. They are looking at the IBM ADSM product. There is roughly 40% of an FTE in Mac central support plus about 35 people around the site providing Mac and PC support. Mac and PC activities of these 35 people are not coordinated centrally. The main requests from users are for: software distribution, printer support, license management and training.

DESY wants to improve PC support. First they need to find out what people need, then provide advice on what to buy, reduce the diversity and set up network services (e.g. mail., print) to be available everywhere. They are interested in CERN's NICE environment. They have not decided on a recommended email package though many people in the machine group are using Microsoft Mail.

RAL Office/Information Systems Plans

RAL put together an Office Systems Strategy Working Party (OSSWP) in 1992. In 1984 RAL began a pilot office system (0S) based on PROFS now called office Vision/VM (OV/VM), the OSSWP reports the following about RAL's OV/VM experiences:

"The more successful aspects have been the diary function, the document repository, the electronic communication of notes and documents, and the several MIS functions accessible from OV/VM. The most significant failing has been in satisfying in full the specialist needs of the scientific staff resulting in a large portion of them eschewing OV/VM altogether ... The key lesson learned from this experience are the importance of a critical mass of users, and of integrating office functions into the normal computing environment of scientific staff. Almost all the potential benefits available from diaries, email, document preparation and information dissemination are lost if an insufficient proportion of the staff actually makes use of the facilities on a regular basis".

The OSSWP came up with a set of goals (see "RAL Administrative Computing and office systems: A strategy for the Nineties" T. Daniels 27 May 1992.) The working Group agreed that the IBM mainframe would need to be replaced, the provision of office/information services would be based on a client/server approach. They also recommended setting up a technical working group to select and define the suggested components of an alternative office system and to oversee their introduction.

This technical working group of about 16 people provided a progress report (see "Office Systems Technical Working Group: Interim Peport to DL and RAL Management Boards" T. Daniels & A. A. Taylor 6-Dec-93). The report provides extensive appendices including user requirements, current uses at RAL and the evaluation of many possible products.

RAL does not want to lock into one vendor for all office products. They regard OS/2 as a solid system that is the best buy (for PCs) today but likely to be supplanted by NT in 4-5 years time, so they do not wish to lock into OS/2. Even today drivers are more common for Windows than OS/2 which can lead to problems. People who have moved to OS/2 seem to be happy running Windows applications under OS/2. It appears to be the more sophisticated (less than 10) users (similar to Unix types) who are moving to OS/2. As such RAL regards OS/2 as a specialist niche for clients. It may have relevance for servers. There is no interest at RAL in Nextstep, it is too limited a market. There also appears to be little central investigation of NT in progress. They appear to have considerable hopes for Windows 4 with it's 32 bit support and NT like features but without the high overhead required for NT. NT is expected to be mainly needed for servers.

The administration group is using a Novell Netware server to support shared disks. Such a solution is not seen as suitable for the entire site, since it is proprietary. There is a central Unix NFS file server with site-wide access. It provides shared backed-up data. They have not looked at WDSF/ADSM for backup.

Central PC support consists of about 6 FTEs. They charge back PC support at a rate of 1 FTE per 100 PCs. Repair requests are made to the help desk. The PC people provide spare monitors & PC machines for loan in case of problems. Most of the hardware problems are associated with monitors & mice and a few particular models of disks. PCs are purchased from a single selected vendor (Viglund) which is small enough that they respond to RALs concerns. The choice of vendors is reviewed based on reliability,cost concerns. They have a database to keep track of what is supported & the warranty information. They find it impractical to try & keep track of the hardware & software on machines or where the PC is located.

RAL has about 100+ Macs and 1000 PCs About 1/3 of the PCs are Ethernetted. They are buying about 200/year. They also have a few hundred UNIX workstations. The central support group specifies what to buy, makes up the requisition, receives the equipment, and sets it up. The vendor deliver it with DOS and windows 3.1 preinstalled. For applications they recommend Word, Lotus 1-2-3 or Excel (see section on Electronic Mail for Mail recommendation), PC/NFS, FTP Software Inc, some DECnet, & Vista eXceed for Xterminal emulation. The applications and networking software are installed in the central area before the machines are delivered to the user. Most new PCs are Etherneted installations facilitated by using CDROMs and Network installs and can be done in parallel. Accepting and preparing the machines centrally enables RAL to spot infant problems and catch problem trends early. It is observed that once users have a working system they seldom want to upgrade. This is fortunate since upgrading in the field without something like NICE is labor intensive, especially since the machines are usually customized by the users.

OTHER APPLICATIONS (MAIL & WWW)

Electronic Mail at CERN, RAL, & DESY

CERN uses MicroSoft Mail on PC with the MicroSoft Mail Post Office under Novell. They report that it is not MIME compliant and MicroSoft seems not to have plans to make it compliant. On Macs they use QuickMail with the StarNine Gateway. For UNIX they use elm & pine, the Sun mail tool, & zmail which is bundled on SGI. They looked at Oracle Mail but rejected it since it was not RFC822 compliant, had problems with return addresses and the user interface was not easy to use.

DESY is testing email packages. They feel MicroSoft Mail still needs MIME capability. They are interested in POP/IMAP/Eudora. They would like the clients to be consistent across platforms. They are not very happy with zmail on the SGIs.

RAL are looking at MicroSoft Mail and ccMail for PC and Eudora for Macs. At the moment most mail is mainly read and sent on the VM System.

The UK is converting to "world order" email addressing (see Flagship 28 p. 10) so mail will now be in the form Fred@vax.oxford.ac.uk instead of Fred@uk.ac.oxford.vax.

WWW Support at CERN

The ECP division at CERN has 5 people providing Information Systems services. In particular they are providing help to the experiments to set up their own WWW servers. They are also looking at what they call meta documentation to cover the complete lifecylce of a project from planning documents, specifications, software listings etc. all linked together with hypertext links. They have a FrameMaker to WWW converter and also a word to HTML converter.

LOCAL AREA NETWORKS

CERN LAN Support

CERN has several strategic LAN problems to solve including: providing 100 times more integrated bandwidth over 10 years; providing high bandwidth, highly reliable paths especially between CORE and collaborative work groups; and accommodating the need to support multiple protocols, supporting multicast, synchronous applications such as video conferencing.

The strategies to address these problems include: converting the CERN wiring to a structured fiber and copper cabling scheme and select one or more suppliers of network hubs/switches; isolate selected zones behind routers; select an industrial network management system; deploy FDDI as appropriate; deploy FDDI and Ethernet switches as appropriate; prepare for ATM deployment; prepare for the next generation of IP.

They are investing in DEC FDDI Gigaswitches to be used in tandem for semi-hot failover. CERN have been evaluating the Gigaswitch and have found it to be very complex containing over 250,000 lines of code. Management of the Gigaswitch was based on the DECmcc network management station which was de-emphasized in DEC's future plans late in '93, which came as a shock to CERN. They have an FDDI Sniffer and Tekelec for FDDI troubleshooting.

CERN have also been looking at Alantec & Kalpana Ethernet switches to increase aggregate Ethernet bandwidth with the intent to upgrade them to ATM when available. They also have an ATM pilot project with a 4-port ATM switch from Netcom plus an HP workstation with a prototype ATM interface and router.

For an industrial network management system they are looking at OpenView, Net-View/6000, SunNeT Manager and Cabletron's Spectrum. They appear to be particularly interested in Spectrum. Since they have a large legacy of DECnet this must weigh heavily on any selection. In the meantime they continue to use the CERN developed network management system. They generate statistical reports on utilization and errors/pkt from bridges (DEC & HP). They use pings every 5 minutes to establish reachability and display maps with lines between rodes scaled in width depending on the error rates. They do not make ping timing measurements on a regular basis. Trouble-ticketing uses email to notify responsible people. They will look at paging people via phone in the future.

CERN is moving from a bridged network to a subnetted network. Since IP addresses are randomly set with respect to geography subnetting will be difficult. IP addresses can be requested by sending email to tcpip@cern.ch.

DESY LAN Support

DESY have invested in Giga Routers from Net Star in Minneapolis. The Giga router is like the DEC Gigaswitch in that it provides non-blocking FDDI access between 16, 32, or 64 ports (the

Giga Switch supports 32 port maximum). It also supports IP routing, unlike the Gigaswitch, and will be adding MAC layer bridging in the near future. They also plan to add OC-12 ATM access later in 1994 Each port can sustain 1 Gbps and the aggregate bandwidth is 70Gbps. The cost is about $110K including 7 HIPPI ports and 4 FDDI ports. It supports SNMP. DESY have only had the Giga Router for a week so it is too early to relate experiences. John Morrison at Los Alamos was identified as someone who had been testing Giga Routers: DESY is using a DEC Gigaswitch to build an Alpha cluster for open VMS support.

RAL LAN Support

RAL is divided into "villages" each of which is connected to the backbone, & each of which has an identified autonomous village network manager (VNM). The VNM has a bank of IP addresses that she/he can allocate. An email list is maintained of VNMs so they can be notified of changes that may affect the network. News is not used for this purpose since they feel it would add to the information overload, be irrelevant to most users and cause unnecessary concern. They also log such changes by keeping copies of all the notifications. The departments choose the VNMs. The quality of VNMs varies widely. Central support is considering producing some documents to help VNMs. There is a LANs management committee that is run by operations and meets about every three months. For a fee the central support group will manage a village's network. Central support does most of the network installations.

There are about 9 FTEs in Network Operations. Network Development is split across several groups including an Advanced Communications Unit (5 FTEs). IBM Networking Group (3FTEs) and Communications and Small systems (9 FTEs).

They are moving from a bridged to a routed network with the villages being placed behind routers. They are installing UTP5 in 1 wing of the computer building as a test. It is easier to manage and can support FDDI and ATM later. They have an FDDI backbone with DEC concentraters and 3Com and NSC routers. The 3Com routers were chosen after evaluating bids from 3Com, Cisco and Wellfleet. They are highly reliable and easier to configure than Ciscos. They route IP, and IPX and are starting to route AppleTalk. Next they will route Decnet. There were FDDI performance problems related with AIX/6000 release 3.2.2 which are fixed in more recent releases. They have a W&G FDDI analyzer that they selected over an HP and Sniffer. To manage the 3Com routers they bought the 3Com application for SunNet Manager. They also have 4 Ethermeters, but do not appear to be making heavy use of them. They have not taken network management seriously yet since the network is running smoothly and problems are only at the irritant level. They would like to get some baseline measurements. RAL is looking at ATM: they have a 4 port ATM switch from Fore that will be used with Alphas (AIX is not ready yet). They want to investigate the speed of connection for high volume servers and farms. They also have a Netcom ATM switch and will connect the two switches. They are hoping for 3Com to provide an ATM router interface.

For home use they have 19.2 kbps access into a Decserver (using V.Fast modems). ISDN is reasonably deployed in the UK. Initial install costs are about $600.00. Rental costs are quarterly. They are looking to use ISDN to support SLIP for work workstation and PC connections at home.

EUROPEAN NETWORKING

European Networking

European wide area networking is different from the US due in part to their stronger national ties and the continuing telecom monopolies and fragmented telecoms infrastructure (see "Knit Your Own Superhighway" The Economist, Oct 16, 1993 p 121). One result of this is that transatlantic line costs are similar to inter European country line costs.

To first order each European country runs its own national network and are in turn members of RARE (which coordinates technical developments). Members of RARE qualify to be shareholders in DANTE (Delivering Advanced Network Technology to Europe Ltd). DANTE runs the operational side of the EMPB (European Multi Protocol Backhone) setting up and operating international connections, plus providing intercontinental links and network applications such as directory services and X-400 mail.

Separate & competing with DANTE/EMPB IS THE EBONE which links the French, Scandinavian and Austrian national networks. The interchanges between EBONE & EMPB are not fully clear at the moment. There is a gateway agreement which runs until 30 Jun 94; what happens after that is undefined. EBONE appears to be less formally organized (eg. it is not a legal entity, national nets pay for inter country links individually). EMPB has more in the way of performance guarantees. There appear to be culture conflicts between DANTE and EBONE. Before Xmas NORDUnet decided in priciple to take an EMPB connection and have cancelled their EBONE connection from 7/1/94; Austria with some reluctance, hace also submitted a formal cancellation and are talking to DANTE which just leaves France.

Partly due to this competition there are many transatlantic links including: the ULCC fat pipe (1.5 Mbps) which is funded by NSF, DARPA & NASA; ESnet has a 1.5 Mbps Dusseldorf - PPNL link and the Bolonga - FNAL links; the EBONE Stockholm/Paris to U.S. link at 1.5 Mbps funded 40% by the NSF; the CERN - GIX (Washington DC) link which will be taken over by DANTE on Jan 1, 94; plus a new link from Amsterdam to GIX to be installed Jan 15 94. There is also a 64Kbps line from Canada thru the London Canadian Embassy to ULCC. Besides the transaltantic links there is a 64 Kbps link direct from London to Korea which is about to be ordered, and a link to Japan via the U.S. at 512 Kbps.

The future funding for these transatlantic links is in some cases unclear. CERN has decided to contract for 1 Mbps of EuropaNet (th marketing name for EMPB/DANTE) access and will collect money to support this from the HEP community. CERN also will approach the U.S. HEP community to collect money to pay for the transaltantic link.

EMPB is also being extended into Eastern Europe. Links to Hungary, Prague, and Bulgaria are already in place and a link to Roumania is being worked on. DANTE has requested the European Commission to change these links from X.25 to IP and expect this to happen soon. Slovenia is already connected to the EMPB and will bought into the EC funding program in 1994. Next year they also hope to extend connections to the Baltic Republics, and Albania.

DESY is in the process of pulling in an experimental link to Moscow via a 256 Kbps channel satellite link that comes down at Moscow State University and is distributed by microwave to

ITEP, the Lebedev Institute and Troitsk. ESnet is interested in sharing this link and already have the OK to carry Russian traffic. NASA is trying to get fiber links from the west to get to Petersberg & then to Moscow via microwave.

All German Universities are on a common network run by the German PTT. It is an X.25 network with 9.6 Kbps, 64 Kbps, & 2 Mbps links. The 300 km 2 Mbps. DESY - Hamburg to DESY - Zeuthen (Berlin) costs about $250K/year. The majority of the traffic on the Germsn academic network is IP encapsulated in X.25. Unfortunately the performance of the routers in this mode is not good (since it is only a niche market). We observed this via the recommended path to Dusseldorf when trying to telnet to SLAC from DESY, the response time was very poor at best (> 500msec) and very variable. We eventually gave up on this and used the CERN link instead logging (using the CERN-DESY 768 kbps IP link) on to CERN, & telnetting to SLAC from there via the CERN-GIX X link. This avoids the X-25 links & provided a much more usable link.

In the U.K. the academic network is being upgraded to ATM at 34 Mbps. This is discussed in "Super JANET, Special Edition" Nov 1993 published by the Joint Network Team. It is called Super JANET and includes an 18 M pound sterling 4 year contract awarded to BT. It will include 16 sites with SDH connectivity and 50 sites with SMDS connectivity.

APPENDIX A: FULL ITINERARY

December 4: Leave Stanford

December 5: Arrive Geneva

December 6-8: Visit CERN

December 8: Arrive Hamburg

December 9-10: Visit DESY

December 10: Arrive London (Dubois & Melen returned to Stanford Dec 11) (Johnson returned to Stanford Dec 13)

December 13-14: Visit RAL

December 15: Visit DANTE

December 20: Returned to Stanford

APPENDIX B: LIST OF PERSONS MET DURING TRIP

CERN People Met

CERN is organized into 14 divisions, as follows. We met with people from the CN, AS, PPE, and ECP divisions.

CERN Divisions

Div.	Function	Div. Leader
AS	Administrative Support	J Ferguson
AT	Accelerator Technologies	J P Gourber
CN	Computing and Networks	D O Williams
DG	Director - General	C Rubbia
ECP	Electronics and Computing for Physics	P G Innocenti
FI	Finance	A J Naudi
MT	Mechanical Technologies	G Bachy
PE	Personnel	W Middelkoop
PPE	Particle Physics Experiments	J V Allaby
PS	Proton Synchrotron accelerator	K Hubner
SL	SPS and LEP accelerators	L R Evans
ST	Technical Services	F A Ferger
TH	Theory	J Ellis
TIS	Technical Inspection and Safety	B de Raad

CN - Computing and Networks Division The objectives of the Computing and Networks (CN) Division are briefly summarised below:

Provision of computing infrastructure for the laboratory. This includes batch, timesharing, workstation server and mass storage (magnetic media) services; distributed computing support, including onsite and offsite networking, central and remote printing, and workstation services; CERN Program Library and related software. Provision of applications-related hardware, and software services, in collaboration with the divisions concerned. This includes applications oriented towards experiments, the accelerator and technical sectors, engineering, and management information.

The CN Division has undergone a recent reorganization and the current reporting structure was unavailable at the time. The following is a "best guess" taken at the time of the visit. (Any errors are my reponsibility - Randy Melen.)

```
    Division Leader: D.O. Williams
    Deputy to the Division Leader:     D. Jacobs
CN-DI  Division Leader's Office:  J.C. Juvet
CN-AS  Application Software:  A. Osborne
CN-CE  Computing for Engineering:  C. Eck
CN-DCI Desktop Computing Infrastructure:  C. Jones
CN-CS  Communications Systems:  B. Carpenter
CN-OM  Operations Management:  D. Underhill
CN-PDP Physics Data Processing Services:  L. Robertson
Name                    Title (if known)                Email Address
----                    -----                           -------------
                        (Interests)
```

David Williams	CN Division Leader	davidw@cernvm.cern.ch
Julian Bunn		julian@vxcern.cern.ch
Chris Jones		chris@cernvm.cern.ch
Alan Silverman		alan@vxcern.cern.ch
Judy Richards		
Ignacio Reguero		
	(Sun & SGI support)	
Rainer Tobbicke		
	(AIX, AFS, printers)	
Tony Cass		
	(work group servers, TMS)	
Lionel Cons		
	(HP support, security)	
Les Robertson	(leader of Physics Data Processing)	
Jean-Philippe Baud		
	(stager and tape software for SHIFT)	
Gordon Lee		gordon@dxcern.cern.ch
Frederick Hemmer		hemmer@sun2.cern.ch
Ben Segal		ben@dxcern.cern.ch
Brian Carpenter		brian@dxcoms.cern.ch
Mike Gerard		jmg@dxcoms.cern.chr
John Gamble		gamble@vxcern.cern.ch
Dave Underhill		djuct@cernvm.cern.ch
Charles Curran		cscct@cernvm.cern.ch
Jean-Claude Juvet		jean-claude-juvet@macmail.cern.ch
Tony Osborne	ASD Group Leader	tonyo@cernvm
Olivia Couet	PAW Section Leader	couet@hphigz.cern.ch
Miguel Marquina	ACT Section Leader	marquina@cern.ch
	(UCO, accounting, UNIX apps)	
Jamie Shiers	CERN Program Librarian	jamie@zfatal.cern.ch
	(FATMEN)	
Rene Brun		brun@cernvm
	(evolution of software)	
Simone Giani		sgiani@cernvm [did not meet]
	(Geant)	
Sergio Santiago		santiago@cernvm
	(data bases)	

AS - Administrative Support Division The AS Division is responsible for administrative computer applications and much of the Macintsoh and PC-compatible support. The Division Director, John Ferguson, was unavailable for our meeting but we met with his staff.

```
Name                   Title (if known)                Email Address
----                   -----                           -------------
                       (Interests)
                       -----------
Achille Petrilli       Deputy AS Division Leader       achille@dxcern.cern.ch
Frank Ovett            PC & Mac Group Leader           ove@cernvm.cern.ch
Mats Moller            Sys. & Ops. Deputy Group Leader mats.moller@macmail.cern.ch
```

PPE - Particle Physics Experiments We met with researchers from the LEP experiment.

```
Richard Mount          Physicist                       mount@cernvm.cern.ch
Flavio Sticozzi                                        sticozz@dxcern.cern.ch
John Swain                                             swain@cernapo.cern.ch
Rizwan Khan                                            khan@dxcern.cern.ch
                       (UNIX & Perl)
Pierantonio Marchesini                                 marchesi@lepics.cern.ch
                       (REXX conversion)
Robert Clare                                           clare@hpl3.cern.ch
Vincento Innocente                                     innocent@cernvm.cern.ch
```

ECP - Electronics and Computing for Physics Division The Electronics and Computing for Physics Division is made up of staff from the fields of electronics and computing, working on projects of direct relevance to the experimental program of CERN. It is divided into two major units: Electronics and Computing.

```
Division Leader: P G Innocenti
Electronics Unit - Deputy: F Bourgeois
Computing Unit - Deputy: G Kellner
```

The Computing unit contains 5 groups, organised according to related technical activities. Group members are either directly attached to experiments or work on projects, mainly in collaboration with other groups, divisions or outside users.

```
PT   PROGRAMMING TECHNIQUES                        P Palazzi
RA   READ OUT ARCHITECTURE                         S Cittolin
DS   DATA ACQUISITION SYSTEMS                      A Vascotto
SA   SIMULATION, RECONSTRUCTION & ANALYSIS         C Onions
```

At the Zeuthen location near Berlin, they have about 150 people with about 70 of those as scientists and 10-20 in the Computer Center. The location is about 300 km. from the Hamburg location, 5 hours of travel no matter which mode is chosen. The Zeuthen location is a complete IBM shop now, having already moved from a VM environment on an IBM-compatible mainframe in 1991. It took about 6 months as that was considered "sudden".

They also have HP and SGI systems with some Sun, IBM, and Convex machines. As well, they have about 10 Macs in the H1 experiment and another 50 PCs on the network. Their net is Ethernet with TCP/IP plus some AppleTalk and DECnet for a VMS cluster. They use a FDDI ring between the SGI file server and the computer servers (HP, SGI, and a Convex C3210). They have about 100 GB of SCSI disk for data. They consider it to be cheaper that robotic storage right now though they expect to have cheaper prices from robotics for the TB range in the next 1-2 years. For now, they move data with 8mm tapes though these experience some problems some times. They have about 70 X-terminals.

Our discussions were with the PT group.

```
Name                    Title (if known)                 Email Address
----                    -----                            -------------
                        (Interests)
                        -----------
Pier Giorgio Innocenti  ECP Division Leader              innp@cernvm.cern.ch
Paolo Palazzi           Pgmming Techniques Group Leader  palazzi@vxcern.cern.ch
Gottfried Kellner       ECP Dep.Comp. Unit Div. Leader   kellner@vxcern.cern.ch
Jurgen Knobloch         Computing Coordinator, Aleph     knobloch@cern.ch
Robert Cailliau                                          cailliau@www1.cern.ch
                        (information systems for experiments -- WWW)
Bertrand Rousseau                                        rousseau@ptsun01.cern.ch
```

DESY People Met

```
Name                    Title (if known)                 Email Address
----                    -----                            -------------
                        (Interests)
                        -----------
Dietrich Moenkemeyer                                     r01moe@dsyibm.desy.de
                        (IBM mainframe, tape operations)
Hans Kammerlocher                                        hans@vxdesy.desy.de
                        (automation of operations)
Peter-Klaus Schilling   Head of R2 Group                 r02sch@dsyibm.desy.de
                        (general UNIX workstation clusters)
Erwin Deffur                                             f22def@dsyibm.desy.de
                        (general PC and Mac support)
Hans Frese                                               frese@desy.de
                        (external networking)
Michael Ernst                                            r02ern@dsyibm.desy.de
                        (internal networking, high perf. UNIX, MSS technology)
Michael Behrens                                          r01beh@dsyibm.desy.de
                        (user support)
Jan Hendrik Peters                                       r01net@dsyibm.desy.de
```

	(user support)	
Wolfgang Krechlok		krechlok@vxdesy.desy.de
	(VMS, DEC Alphas)	
Wolfgang Friebel		friebel@ifh.de
Till Poser		
Martin Gasthuber		
Otto Hell		

RAL People Met

John Barlow	Manager	Computer Services
Paul Bryant	Group Leader	Communications & Small Systems
Bob Cooper	Manager	SuperJanet
Trevor Daniels	Manager	Software & Development, CCD RAL
Dave Drummond	Systems Analyst	IBM PCs
David Gaunt	Systems Analyst	Unix
Tim Kidd	Systems Analyst	IBM Networking
Chris Osland	Group Leader	Video Conferencing
David Rigby	Systems Programmer	Sorage Management
Geoff Robinson	Systems Analyst	IBM PCs

DANTE People Met

Howard Davies	Director	European Networking
Tim Streater	Systems Specialist	European Networking

APPENDIX C: DOCUMENTS OBTAINED ON TRIP

These documents are available for viewing and copying (please do not remove) from the SCS Technical Library.

Documents Obtained from CERN

C01: Interview with Garber from Epoch about the DMIG (p.56, 11/11/93, OST)

C02: DRAFT Computing and Networks Division (931104)

C03: Proposal to Rightsize VXCERN w/Scaleable Mixed Architecture VMS Cluster

C04: 13'th IEEE Symposium on MSS announcement

C05: From Mainframes to Distributed Computing

C06: not used

C07: Second UniTree User's Group Meeting & Symposia – trip report

C08: A strategy for a structured internal network
C09: The CERN Internal Network Brochure
C10: Dimou's mail re: specific weaknesses of Oracle*Mail 7/20/90
C11: Mail to Dimou from Oracle questioning the validity of RFC822 10/8/91
C12: Proposal for the Evolution of the CORE Services
C13: A New Model for CERN Computing Services in the Post-Mainframe Era
C14: A Strategy for the Interactive Services at CERN over the next 3 years
C15: CORE Highlights Week 47
C16: L3 Regina User's Guide
C17: Report of the High Perf. Computing & Networking Advisory Committee V1
C18: not used
C19: Cartrobo Configuration/Weekday Average Hourly Mounts report
C20: Software Development Tools from ECP-PT
C21: Organisation and Software Development in LHC
C22: Methods of Software Development for the ALEPH Experiment
C23: Aleph Offline chart
C24: Aleph offline workstation cluster chart
C25: WWW Project description flier
C26: ADAMO description flier
C27: LHC Era Computing lecture notes
C28: A Sketch of CORE's Networking Requirements
C29: CERN Computer Newsletter No. 214 October-December 1993
C30: X Terminal Administrator's Guide
C31: Recommended Standard for Unix Workstation Environment Setup
C32: Ethernet. A guide to technical installation.
C33: A Guide to Personal Computer Networks at CERN
C34: IBM Workstation Support at CERN
C35: CERN Support Services for DEC Workstations Running ULTRIX or OSF/1
C36: SUN Workstation Support at CERN
C37: The CORE Services at CERN
C38: Mainframe Svcs. w/RISC-based Workstations: The first 2 yrs' experience
C39: CERN Security Handbook
C40: CERN Computer Newsletter No. 211/212 March-June 1993

C41: CERN Computer Newsletter No. 208 July-September 1992

C42: CERN Computer Newsletter No. 207 May-June 1992

C43: CERN Computer Newsletter No. 205 January-February 1992

C44: CERN Computer Newsletter No. 204 October-December 1991

C45: CERN Computer Newsletter No. 203 July-September 1991

Documents Obtained from RAL

D01: How to use the Previewer "xdvi" and "ghostview"

D02: X-Terminal "Choosers"

D03: Introduction to "NCSA Mosaic"

D04: Remote Login Without a DESY Computer Account

D05: "Emacs" Mini-Reference Card

D06: Introduction to "Unix" at DESY

D07: Update on the DESY UNIX Environment

D08: Sample letter used to justify DESY's academic status for software

D09: GNU Emacs Reference Card

D10: Reference Card for the "vi" editor on Falcos and X-terminals

D11: pico - a very simple Editor on Unix

D12: Connect to a Remote Host Using an X-Terminal

D13: How to Send a FAX Using Email from any Computer at DESY

D14: The First Three Days: DESY Computer Center Primer

D15: How to use TeX at DESY

D16: Using LaTeX at DESY

Documents Obtained from RAL

R01: Flagship 27

R02: Flagship 28

R03: SuperJANET ATM Switch Request For Information, Sep 92

R04: Network News, Nov 1993 (SuperJANET Special Edition)

R05: RAL ATM Plans

R06: A Strategy for the Nineties, Office Systems Strategy Working Party, May 92

R07: Information Systems at RAL, Jun 93

R08: Interim Report to Daresbury Lab and Rutherford Appleton Lab Management Board from Office Systems Technical Working Group, Dec93