CN/DPG/93/MIN/5

THE CN DIVISIONAL PLANNING GROUP

Minutes of the Meetings held on 14th. and 15th. June 1993

PRESENT: R.Brun, J.J.Bunn(Secretary), B.Carpenter, P.Defert, F.Fluckiger, F.Gagliardi, F.Hemmer, D.Jacobs, S.Jarp, C.M.S.Jones, J.–C.Juvet, M.Marquina, P.Moroni, A.Nathaniel, A.Osborne, J.Richards, I.Reguero, H.Renshall, L.M.Robertson, J.Shiers, A.Silverman, R.Tobbicke, D.O.Williams(Chairman).

- *MINUTES OF LAST MEETING (CN/DPG/93/MIN/4), AND MATTERS ARISING*

The Minutes were accepted as written.

- *CHAIRMAN'S REPORT*

D.O.Williams said that LEP had had a very good start up, and there would consequently be a lot of data for analysis from the experiments. COLLECT had heard a presentation on the H1 and Zeus DESY experiments. A draft document existed that gave advice to the LHC experiments on how to compose the computing sections of their technical proposals. The CS2 machine from Meiko (or at least part of it), should arrive at CERN sometime in June (1993). DOW expected a status report on GP–MIMD2 to be heard at the DPG in October.

- *TAPE STATISTICS*

J.Richards presented statistics for tape mounting on the central systems. The load was split four to one between the LEP and non–LEP experiments. There had been 500,000 mounts so far this year, 72% of which were on CERNVM, and 13% on CORE. Without staging systems, the load would imply mounting 35,000 tapes per week. In fact, there were 15,000 mounts per week. More than 80% of the tapes had been mounted less than five times so far this year. However, over 50% of the mounts were of the 7% of tapes that had been mounted more than twenty times! She said how nice it would be if it were possible to to identify the tapes that were mounted most often. JR was asked to make a summary of her copious data, and if she could produce the "virtual" mount figures.

A long discussion followed during which several questions arose:

— How is the rackspace being managed?
— How often are experiments removing tapes and putting them in the Vault?
— How could the users be allowed to insert themselves their tapes in the Robots?

- *CORE*

L.M.Robertson presented the current status and plans for expansion of the CORE services. Copies of his transparencies are attached to these Minutes. SHIFT was (or was planned) being used by Opal, Aleph, Delphi, Atlas, CMS, SMC and CPLEAR. L3 were also going to join. Utilization of the 430 available CERN Units was running at around 50%. There were about two incidents per week that affected one or more CPU members, leading to an MTBF of 360 hours. (MTBF for disks was around 140000 hours.) A new component of SHIFT would be introduced later in the year: this was called SAGA and consisted of a number of Alpha OSF workstations clustered across a GIGAswitch. It was a Joint Project with DEC. The main purpose of SAGA would be to provide a parallel platform for running Aleph's analysis program, modified to use PVM, and eventually HPF. Fifty percent of the SAGA machine would be guaranteed for Aleph.

CORE networking strategy was to move towards FDDI, although UltraNet would continue to be used. All tape and disk servers would eventually be equipped with FDDI. The existing Ethernet wiring was undergoing some reconfiguration to cure network stability problems.

LMR emphasised that the long term goal of the SHIFT project was to provide an integrated and flexible set of data handling services for the community. The short term goals included providing a platform on to which CERNVM batch work could be moved. The target was to move away 80% of the CERNVM batch load by 1994. He elaborated by describing the various migration issues like disk space, tape unit availability and, most importantly, the major problem of where to put CERNVM's interactive load. He felt that simply providing Unix versions of the REXX program were only crutches for the die–hard invalid CERNVM user, and by no means a cure.

Expanding on his theme, LMR said that it was important to stem any remaining flow of new users towards CERNVM by providing new experiments with immediate resources on SHIFT. Analyses of the requirements of the experiments should be undertaken to determine their migration requirements. Most importantly, the experiments themselves would simply have to do a lot of work themselves in order to benefit from a move to SHIFT.

There followed a discussion, during which several people expressed their belief that the Division should already be actively and openly selling the concept of a general move to SHIFT–based services for the community. Some felt that there was a clear need for a front–end interactive service in SHIFT, whereas others disagreed, preferring the idea that this should come under the umbrella of the Work Group Servers scheme. It was clear that "interactive" could mean many things: dealing with electronic mail, code development, PAW–like activity, were just a few examples.

- *COMMON STAGING*

P.Moroni introduced the topic of a universal staged data access method by describing the existing and successful system used on CERNVM. This featured good performance, stability, shareability and scalability. Unfortunately the network connectivity was poor, and staged data were not readily available to systems other than CERNVM itself.

He described the FITMEN system, which allowed staged data on CERNVM disks to migrate onto cartridges in the Robot. This meant much faster turnround in the cases where it was required in the future. The FITMEN system was ready for use.

Because the IBM had such good I/O capacity, much of which was still not yet utilized, PM said that the machine could make an excellent data server for the network. This would involve enhancements to VMSTAGE, and implementing a new service called CLIO that allowed fast ftp access to staged data by remote machines.

F.Hemmer took the stage to then describe the CORE data staging model. In this scheme, the CORE tape server machines received requests from across the network for stage in or stage out of data. The handling of the data streams was by a daemon that collaborated (via a socket interface) with a server on the remote target machine. This architecture was very flexible in the sense that the staging could be made from a third party machine (such as a workstation). The software was portable, even to non–Unix systems, so allowing, for example, VMS machines to benefit and participate in the scheme. It was intended to improve the software by rationalising the configuration method, by adding support for remote VMS machines, by adding support for other tape formats (e.g. VBS). (There was a debate as to the necessity and merit of providing VBS support.)

FH continued by proposing that there be work started on a common access to CERNVM and CORE stage areas by system software (not at the user level). This would require a VM server to facilitate network query and access to the IBM Stage disks. An important feature would be dynamic device allocation between CORE and the IBM. The benefits would be homogeneous remote access to staged data, effective use of the available resources, flexibility and scalability.

- *AFS*

  R.Tobbicke presented AFS, a file system comparable in some ways to NFS, but with a much wider scope and much better performance. Copies of his transparencies are attached. For the ben–efit of most of the audience, who were not expert in the subject, he began with a short tutorial on how AFS worked, and its terminology.

  An AFS pilot had been started at CERN, and collaborating platforms included machines running HP–UX, AIX, Sun–OS, Solaris, Ultrix, Next and Aegis(?). There were 7 GBytes of disk in the CERN AFS cell, with around fifty simultaneous client machines. A user's guide was being written which described how to make use of AFS. There was a short digression on the subject of Ker–beros authentication "tickets", which expired after 25 hours, and then required renewal before the user could regain access to the files in AFS. AFS could be made accessible to machines running only NFS by use of an AFS to NFS "exporter". This was already being used for example on VXCERN.

  In the discussion that followed, it was asked what the future of AFS might be when DFS appeared. It was felt that this was, in any case, a very long time off. The problem of migration of AFS cells to DFS "aggregates" was being addressed in the industry. Some were concerned that there seemed to be a limitation to one file backup server per AFS cell, and that this was an unpleasant restriction. RB said that issues like backup policy, registration and so on, had to be settled in a quorum. It was proposed that WGSWG manage the AFS service, and that a discus–sion started in order to arrive at a unified view on AFS, and the other file systems enjoying some currency, to name a few: Netware, Appleshare, DEC's DFS, NFS and DFS.

- *ASIS*

  P.Defert presented the status of the ASIS ftp repository, and the plans for its future. Copies of his transparencies are attached. The ASIS server concept had been approved by the C5 committee in April 1991. Since then it had been providing both ftp and NFS access to CERN and public domain software for users. Software had been written which allowed remote users to install select–ed pieces of software from the ASIS machine. The latest versions of the CERN Program Library were available, as were various parts of the "gnu" distribution, some Frequently Asked Questions (FAQ) lists, and other assorted oddments.

  Requirements from the user community for improvements to ASIS included AFS, and a more reli–able implementation of the installation server. There was also a perceived lack of manpower to operate the repository. In order to move to an AFS–based repository, it was intended to purchase a new machine and some 12 GBytes of extra disk space.

  PD highlighted a particular maintenance problem: it was hard to provide a software distribution for the a particular architecture if a machine of that architecture type was not available on which to test it. Solving the problem clearly involved having the various types of machine available, but also that individuals were named as responsible for testing the various software on each, and as responsible for support. It was agreed that the ASIS server, being the "P–Disk" of the Unix world, had an increasingly important role to play at CERN in the future.

- *PIAF*

  R.Brun outlined the present situation regarding preparations for the PIAF service. Work had been done on implementing a version for a single server (rather than the target five), and this had now been completed as a first pass. He estimated that to parallelize the system would take an extra equivalent amount of work. The system had been "given" to ten Delphi users for trials, and they had experienced no real problems. Staging was presently being done over Ethernet, although the goal was to use FDDI. RB said that he would consider the final system successful if it managed to process at least 200 GBytes of PAW data per week, by the end of 1993. He commented that the "new" PAW n–tuples (which were the sort that had to be used on PIAF) were not enjoying massive popularity: only about 10% of PAW users had taken them up.

  In the discussion that followed, H.Renshall asked for clarification on who the target customers of the PIAF service were. He believed that the majority of physicists would process their PAW data on the desktop, rather than in a central server of this type. RB replied that only time would tell. In reply to a question on the location of the PIAF hardware, RB said that the machines would be moved out of a third floor office to the computer room [1]

- *ANY OTHER BUSINESS*

  None

- *NEXT MEETINGS*

  Members of the DPG will be deemed to have taken note of the following information on future meetings dates:

  1. July 19 (FDDIdeployment,CHORUS)
  2. August 23
  3. September 20
  4. October 18
  5. November 15

[1] The office occupants had recently enjoyed balmy temperatures of forty degrees Celsius, due to the heat generated by the hardware.

Table 1: Action Items

This list is intended to show the items which are currently awaiting presentation at the DPG. The Entry date is the date of the meeting at which the action item arose, and the Target date is the date by which the action item should have been dealt with.

Name	Entry Date	Target Date	Task
F.Fluckiger	Nov 1991	1993	Give Computer Colloquium on Multimedia
A.Osborne	Nov 1992	1993	Session login sources
MPP Steering Group	Nov 1992	1993	Amdahl's Law relevance
M.Marquina	Nov 1992	Dec 1992	64 bit CERNLIB
J.Altaber	Nov 1992	1993	Use of VSAT for poor countries