Bits & Bytes

No. 174 RZG Computer Bulletin June 2003

Computer Center of the Max Planck Society and the Institute for Plasma Physics*

Trends in Videoconferencing in IPP & MPG

Videoconferencing is used regularly for connections in- and outside IPPs sites in Garching and Greifswald. This comprises between 10–30 independent point-to-point conferences per week between rooms on both sites. Some of the rooms, namely the lecture hall in building D2, the seminar rooms 1 & 2 in Greifswald and the smaller rooms of the "Wissenschaftliche Leitung", RZG, and the three meeting rooms in Greifswald are therefore used rather frequently.

New are the two rooms operated by ASDEX Upgrade – the big seminar room is equipped like the other big lecture halls, the smaller one for division E4 has been set up like the existing meeting rooms. In addition about 20 ViaVideos, small desktop units, are used on both sites by technicians and scientists.

A talk by U. Schwenn, given recently at CEA/DFRC in Cadarache, France is available in pdf format under the following two links:
www.vc.dfn.de/doku/berichte/schwenn-may2003.pdf
users.jet.efda.org/pages/rem-part/-documents/schwenn-may2003.pdf

IPP Scenarios: The RZG video group supports three scenarios - desktop, seminar rooms and lecture hall like rooms. The video conference clients recommended and used throughout IPP are ViaVideo desktop systems, Tandberg 500 and 880 settop systems and Tandberg 6000 for big rooms. These allow conferences between all scenarios. The user interfaces for the smaller two are easy and can be operated by anyone, the big rooms need a dedicated technician to assist the speakers during their presentations.

Gatekeeper, Firewall, Global Dialing Scheme (GDS): All IPP VC systems and several (≈ 20) external are registered on RZGs gatekeeper, an openH.323 Gnu software run on a Linux PC. Using the Global Dialing Scheme based on the E.164 standard of ViDeNet throughout IPP allows to keep all VC systems inside the RZG firewall. They are reachable only by their E.164 number by control of the gatekeeper. The structure of the numbering is as follows: [<EZ>] <worldgk> <cc.> <org.> <client no> [<suffix>] e.g. 00 49 140 1000 (Schwenn's ViaVideo at IPP). The numbers allow standardized international and national access together with individual choice of local extensions (<client no>). See also www.wvn.ac.uk/support/h323address.htm

DFN Videoconference (DFNVC) Service: As with April 1, 2003 the former pilot project of the DFN Verein for a Multipoint infrastructure for the users of the GwiN became operational. This implies for members of MPG, that they will have to expect a billing. Details are under discussion between DFN and the MPG general administration (GV). Until then the standard service and a beta version of the special MPG premium service can be used by anyone inside the MPG. It is obligatory to be registered on one of DFNs gatekeepers, the one of RZG or any other local gatekeeper which must be registered by DFN on the German country gatekeeper (see www.vc.dfn.de/). For questions concerning choice of systems, gatekeepers etc. contact schwenn@rzg.mpg.de, the respective web pages of RZG will be replaced by an updated English version this summer.

Ulrich Schwenn

IBM p690 Supercomputer

System

The IBM "Regatta" system was moved into two steps to the new machine hall of the RZG in February. First 8 nodes were connected to a second "Colony" SP Switch in the new room. After successful test operation this new cluster configuration, the remaining "Regatta" nodes and the interactive node "psi" were also moved. All "Regatta" nodes were connected to the "Colony" SP Switch in a symmetric configuration and are now members of the LoadLeveler batch cluster.

The global parallel file systems /u and /ptmp had also to be reconfigured since the overall disk capacity was increased to 25 TBytes. New Fibre disks have been attached to the system. Both capacities of /u and /ptmp have been significantly increased. The size of the file system /u is now about 4 TBytes, the size of the file system /ptmp is 9 TBytes. For security reasons, all data are mirrored till the file system will be filled up.

Additionally 2 nodes with 256 GBytes memory each have been put into operation with full 64 bit kernel support.

Ingeborg Weidl

*Max-Planck-Institut für Plasmaphysik, Boltzmannstraße 2, D-85748 Garching bei München, tel.: +49(89) 3299-01, e-mail: benutzerberatung@rzg.mpg.de, URL: http://www.rzg.mpg.de/
Editorial: Dr. Roman Hatzky, Tel. -1707

Optimizing for Performance

The IBM pSeries model 690 ("Regatta") uses the POWER4 processor. A tutorial which describes the performance optimization facilities available in the C, C++ and Fortran compilers and the most effective tactics for leveraging them can be found under
www.spscicomp.org/ScicomP7/abstracts.html
Detailed information about the C, C++ and Fortran compilers for AIX, e.g. language references, compiler reference and user's guide are presented as pdf-files under the following two links:
www-3.ibm.com/software/awdtools/caix/library/
www-3.ibm.com/software/awdtools/fortran/-xlfortran/library/

Roman Hatzky

GPFS on the Linux Clusters

For a distributed Linux compute cluster the need for a performant parallel file system is obvious. As standard global file system AFS is in use at RZG. The Blade Center and the Rack optimized Linux Cluster at RZG were additionally equipped with IBM's General Parallel File System (GPFS). This system is already known from our IBM "Regatta" system (/ptmp and /u). In a test configuration we have set up several GPFS systems.

In the contrary to the "Regatta" system where we have two dedicated I/O nodes and all compute nodes as clients participating in the same file system, we employ a different strategy with the Linux clusters. Here, in principle, all compute nodes are both clients and servers for a GPFS file system. In one Blade Center Cluster, e.g., each compute node contributes a 40 GByte partition of its internal hard drive to the GPFS mounted on /gpfs4. Without mirroring data and meta data we would create a single point of failure. Just rebooting one machine would require a shutdown of the whole file system. By mirroring we are able to reboot up to $n/2 - 1$ nodes of a n-nodes cluster without loosing a single Bit or disturbing the file system usage (except performance). Problems occur, if nodes of different failure groups become unavailable. GPFS has the advantage over other parallel file systems such as PVFS that it is able to recover by itself to a certain degree. But if mount points become unavailable, e.g., due to hanging jobs, even GPFS is unable to remount and most often only a reboot helps. Luckily, such situations are in general rare.

In our test configuration we achieve a sustained transfer speed up to more than 100 MBytes/s for a single host I/O. This is almost at the edge of the networks bandwidth. Read/write speed can even be faster, if the data are processed in parallel. This requires to make use of special POSIX compliant I/O libraries. MPI-IO which is going to be able to make use of GPFS' parallel I/O in the near future, and already works with AIX, is another solution. More about this will be available from July on under www.rzg.mpg.de/docs/linux

Naturally, GPFS needs bandwidth, and, at the moment, it has to compete with communication intensive jobs for that bandwidth. We are aware of this bottleneck and will provide a dedicated network for GPFS on the Blade Centers.

Thomas Soddemann

Using the NAG library from C and C++ programs

Although the NAG libraries have been produced for Fortran environments, it is possible to make use of them also in C/C++ programs. Of course, the C/C++ application has to conform to the Fortran conventions (i.e. pass parameters by reference, store arrays in "column major order", and add extra parameters for string lengths). This is described in a NAG technical document
www.rzg.mpg.de/from_external/NAG_mk20.doc/-C-headers/techdoc.html
and the section about interlanguage calls in the IBM compiler user's guide
ibmdoc5.rzg.mpg.de/doc_link/en_US/vac/vacen/-compiler/tasks/tusublnk.htm

The problem remains of producing a working executable, because both the C/C++ and the Fortran environments are required. This is a bit tricky, since it is highly operating system and compiler dependent, and several different NAG libraries produced with different Fortran compilers may exist. How to do it under the RZG (AIX, Linux and Solaris) environments can be found at www.rzg.mpg.de/docs/libraries/nag+c.html We also provide a "makefile" that you may just copy and modify as you see fit. It is hoped that even if you cannot use that makefile directly, the web page provides sufficient information for you to to discover which adjustments to make.

Werner Nagel

New Color Printer at RZG dispatcher room

The Canon CLC 1180 is a color printer with integrated copying facility. It is the substitute for the old Ricoh. The supported paper sizes range from A5 to A3 (80–200g/m^2). Transparencies are loaded and duplex-printing is possible. There is also a 20-fold-sorting unit and an integrated tacker. The printer-name is rzl4c. For more information please see:
www.rzg.mpg.de/visualisation/hardware/printers/

Silvia Groß