413

Report on the fifteenth meeting of Computing Representatives, 19-20 May 2003

P. Prior (Compiler)

Operations Department

September 2003

Contents

Preface

The fifteenth meeting of Computing Representatives took place on 19-20 May 2003 at ECMWF. Eighteen Member States and Co-operating States, plus EUMETSAT, were represented. The list of attendees is given in Annex 1.

The Head of the Computer Division (Walter Zwieflhofer) opened the meeting and welcomed representatives. He gave a presentation on the current status of ECMWF's computer service and plans for its development. Each Computing Representative then gave a short presentation on their service and the use their staff make of ECMWF's computer facilities. There were also presentations from ECMWF staff members on various specific developments in the ECMWF systems. The full programme is given in Annex 2.

This report summarises each presentation, with additional detail on those topics expected to be of particular interest to Member States and Co-operating States. Part I contains ECMWF's contributions and general discussions. Part II contains Member States'and Co-operating States' contributions; all the reports were provided by the representatives themselves.

Part I

ECMWF Staff contributions

ECMWF Computing Service: Status and plans - *Walter Zwieflhofer*

The configuration of ECMWF's computer system current at May 2003 is given in Figure 1 below.

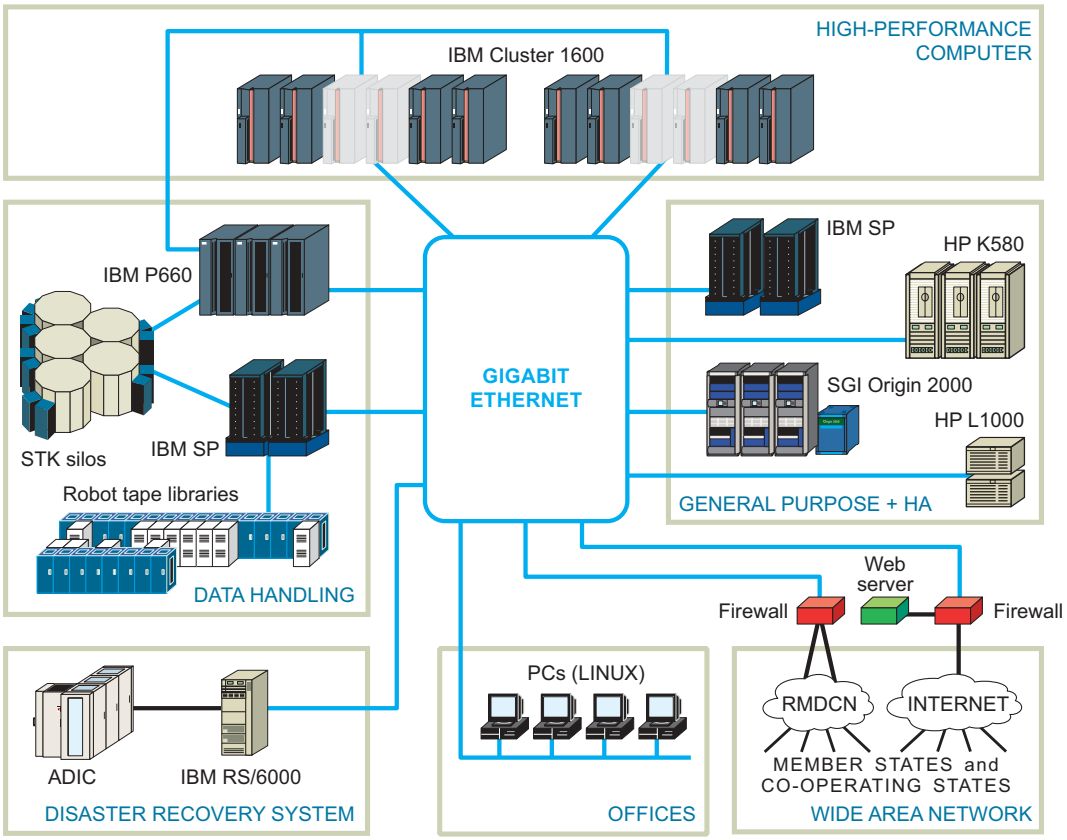


Figure 1

The IBM HPCF - Phase 1

The system comprises two clusters of thirty p690 servers. Each server is logically partitioned into four 8-way nodes, thus each cluster has a total of 120 nodes. All the nodes have 8 GB memory, apart from twelve in each cluster, which each have 32 GB. The processors run at 1.3 GHz (5.2 gigaflops peak). There are 960 processors available for user work in each cluster. Dual plane Colony Switches with PCI adaptors are used for node to node communication. Each cluster has 4.2 Terabytes of disk space available to users. Figure 2 shows the current IBM HPCF configuration. The migration of the operational suite went smoothly. Operational running began on 4 March 2003. The system is now very stable and the reliability high.

The Fujitsu service was terminated on 31st March 2003.

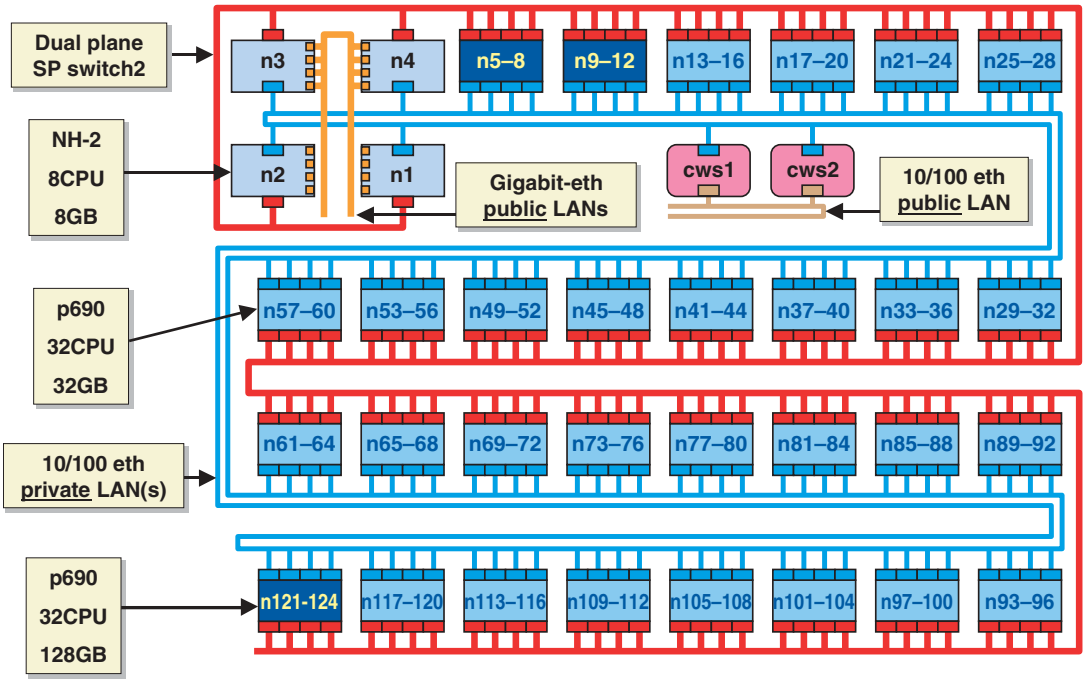


Fig. 2 HPCF Phase1 (cluster A/B)

Access to the HPCF for Member States' users.

The first training course on use of the IBM cluster was held in June 2002. This was primarily for those users who wanted to take advantage of the possibility of restricted early access to the machines from August 2002. The training course was repeated twice more in 2002 and once in Spring 2003. In total 41 users from Member States' organisations attended. A trial service for external users began on 6 November and a full user service was announced on 17 January 2003. The twenty-four man/months of application support provided by IBM under the service level agreement is also available to Member States' users.

Beyond Phase 1 of the HPCF

Originally, a second phase of the HPCF was planned to increase the number of servers and replace the switch adaptors. Various other alternatives, such as replacing all the 1.3 GHz Power4 processors by 1.7 GHz Power4 processors, were investigated but all involved major work on the operational system. Given that the contribution of the Phase 2 upgrade to the performance integral over the contract period is rather small, ECMWF has opted for corresponding performance compensation at the Phase 3 stage in lieu of a potentially risky upgrade of the Phase 1 system.

Phase 3 will entail a complete replacement of the currently installed hardware, apart from the disk space. The system will be based on the Federation switch and Power4+ processors and will use CSM (Cluster System Management) instead of PSSP, the current clustering software.

IBM plans to pre-build the system in Poughkeepsie and give it a three-month trial; ECMWF will be involved in this test phase. It is likely that a small Federation based test system will be available at the end of this year (2003). The main delivery of Phase 3 equipment will be in 2004.

DHS

The new HPSS based system is performing very well. Some of the Phase 2 equipment has been installed and is already in use: two IBM pSeries 650 6M2 servers, each with four processors and one FastT700 RAID with 2.25 TB storage capacity. The new, high capacity tape drives will be delivered later this year.

Migration to the new DHS is progressing well. The MARS-RD service is running exclusively on the new system and MARS-ER and MARS-OD data are being copied from the old TSM based system. The porting of the ECFS service is well underway and the introduction of the new service will be as transparent as possible to users; initially, the new ECFS will handle data stored on both the old and new systems. Full scale testing with selected users and applications will begin in early autumn.

The current data volumes are:

ECFS ~ 170 TB

MARS ~ 590 TB

Servers and desktops

All users' SGI workstations have now been replaced by Linux PCs. The Linux systems continue to be very stable. The migration of internal work from the SGI servers to the IBM Nighthawk II servers is nearly complete. It is planned to shut down the internal SGI servers soon.

The CPU load on ecgate1 is high, so some of the SGI servers which have been replaced by IBM Servers for internal work will be used to upgrade ecgate1. Ecgate1 will receive an additional 8 CPUs, 6 GB of memory and 180 GB of disk, pending its replacement by IBM equipment, which will be acquired under the call-off contract for AIX servers approved by Council in July 2000. Currently, the most suitable system configuration is being investigated and it is expected that the replacement system will be installed later in 2003. The SGI Origin will not be removed before its replacement has been in service for 6 months.

User registration system

Over recent years the number of different classifications of users (Web-only, archive-only etc) has increased. The existing NIS based system was not designed to cope with so many different types of users. An Entity Management System (EMS) is now being implemented to replace and improve upon the existing user registration and authentication system.

The EMS will be able to handle the whole range of different user types and organisation categories which ECMWF deals with. It will also be used to authorise access to the various services offered to various categories of user/site (e.g. login, MARS data, Web pages etc).

By automating many of the tasks required, it is intended that EMS will simplify and speed up the registration process and reduce paperwork to a minimum.

The main system will be operational by summer 2003 and used initially by the Call Desk. Once this basic system is well established it will be extended to enable Computer Representatives to perform certain registration tasks via a browser interface. This interface should be available by the end of 2003.

Web Services

The migration to Linux based web servers is complete and some of the hardware has been upgraded. ECMWF internal users have migrated from ZMAIL to Mozilla for browsing and email services Web services are being continually expanded and improved. The Demeter and ERA-15 data, which are freely available to the research community, can be downloaded.

A "data finder" module and on-line costing and ordering have been implemented to facilitate Data Services.

RMDCN

The RMDCN contract with Equant has been amended to allow the enlargement of the RMDCN beyond WMO RA VI. The United Arab Emirates joined the RMDCN in August 2002 and China in October 2002.

A second price review took place in July 2002 and in December 2002 a contract amendment reducing the overall charges by 27% and waiving ECMWF's right to terminate the contract after the first five years of service was signed. The contract will now run for a total of eight years.

The basic RMDCN package funded for Member States by the ECMWF budget will be increased in October 2003.

ECaccess

The ECaccess portal has been enhanced to provide access to the Centre's computing and archiving facilities via the RMDCN. A second version of ECaccess was released recently and includes many improvements and new features.

Computer security

In accordance with the ECMWF Security Policy an external security audit was performed in the first quarter of 2003. In particular, this audit concentrated on the perimeter systems. The security of ECMWF's Internet access was assessed as high and no major vulnerabilities were identified.

A feasibility study of IPsec based VPN carried out by ECMWF and some of its Member States has been completed and its results published as an ECMWF Technical Note.

Infrastructure review

In summer 2002 ECMWF carried out a detailed review of its electrical and mechanical services and of its computer hall space requirements.

The review covered the eleven kV electricity supply, the UPS (Uninterrupted Power Supply) system, the standby power supply system, the centralised chilled water system, the floor space and the air handling capacity. There were two areas in which the need for improvement was deemed critical: the UPS system and the computer hall floor space. Plans to effect these improvements are underway.

Experience with the new IBM HPCF: Hardware and software - *Neil Storer*

Neil Storer began by explaining the cluster configuration, including internal and external networking, in greater detail. He noted that the system was already exceeding the performance of the VPP system for most applications.

The new system has a much simpler job class structure than the previous NQS queue system on the VPPs which had over one hundred different queues. There are currently four Loadleveler queues: two for operations and two for users' work: one for serial jobs and one for parallel. A job submission filter has been set up to trap the most obvious user errors in submitted jobs. A version of TotalView that runs on Power4 architecture has also been installed.

A simple job scheduler, tailored to ECMWF's specific requirements has been developed that utilises IBM's LoadLeveler features. For instance, LoadLeveler's "backfill" is used to optimise node use: if job n at the top of the queue needs more nodes than are available, backfill searches for jobs lower down the queue than job n, which can use the available nodes and finish in time for all the nodes necessary for job n to be available without delaying the start time of job n. ECMWF has requested from IBM various relatively minor enhancements to LoadLeveler, which would greatly improve and simplify the scheduling of the Centre's workload; discussions are ongoing.

Investigations into using the dual clusters to provide resilience in the most effective way in the event of a serious failure on one cluster are continuing. Currently the operational suite runs on one cluster HPCB with selected, short running research experiments, while cluster HPCA provides a general user service, including that to the Member States. Efficient, reliable mechanisms to enable data to be transferred between the two clusters are being investigated; a heterogeneous shared file system accessible to all the major systems at ECMWF would be ideal.

Multi Chip Module (MCM) problems.

Most of the initial problems related to L3 cache were caused by a bug in the firmware, which was eventually fixed by IBM. Most of these problems caused L3 cache to be disabled at a boot and, since each separate node has its own error log, the error report would be hidden in one of the 240 errlogs and the first indications of problems were when benchmarks ran slowly. Ways of centralising error reporting are being considered. In the meantime, a suite of tests has been developed to run after a boot but before real work begins. The suite includes checks that the nodes and all the switch adaptors are running at the correct performance level and that GPFS file systems are available.

Switch adaptor and Bulk Power Regulator (BPR) problems.

Switch adaptors failed at boot time in several different ways; occasionally, a single redundant power supply failure resulted in the failure of a complete switch plane, which, in turn, caused GPFS problems. IBM spent a good deal of time and effort in trying to solve these problems; eventually, PTF level 20, containing various software fixes, and earthing the frames with dual straps has alleviated them to a large extent.

Software

An alternative, separate set of system disks is used to upgrade the operating system and products. It allows the new system to be built without disrupting machine running and if, by any chance, there are problems, it is simple to revert to the old system. If all goes well, using this means, it takes less than one hour to perform an upgrade.

Floating-point error trapping

It was soon discovered that floating point errors did not trigger a job abort by default. Processes would run to completion producing INFINITE and Not a Number (NaN) results. Unfortunately, it was discovered that switching on the compiler option to trap these errors resulted in very large increases in run times.

A routine (signal_trap) was therefore developed in-house to configure and trap SIGFPE (floating point error signal) caused by overflow, zero division and operations performed on invalid values. Signal_trap has negligible overhead and produces a trace back and an optional core dump. The source code has been made available to several Member States and is available to any other Member States which require it.

Support

As well as the routine support services available to maintain the IBM system, the Service Level Agreement also provides 24 man months of on-site IBM application analyst service. So far, this service has been used to gain expertise in LoadLeveler/Workload Manager, Fortran/MPI/OpenMP, GPFS and Fibre Channel Storage Area Network. An ECMWF consultant is available to assist Member States' users in migrating to the new system.

Locally developed utilities

Since some standard IBM utilities, such as SP Perspectives, were not well suited to the large number of nodes in each cluster at ECMWF, several utilities have been developed in-house, mainly to help in operating the systems:

- node_health - checks for switch adaptor problems and nodes not responding
- llscan - displays the status of LoadLeveler jobs
- eoj - displays information about a specific job
- ll_cpu - displays the usage of the nodes by the LoadLeveler jobs

TotalView debugger

ECMWF installed a beta release of TotalView. There were quite a few problems but its developers, Etnus, responded well and produced fixes for most of them. Version 6.1.0-3 is now installed and supports large applications such as the IFS.

Questions

M. Pithon (France) asked why some of their parallel jobs run more quickly if they use only seven CPUs within a node, rather than all eight. This is also the case for experiments run with IFS, if MPI only is used. However, when a hybrid MPI-OpenMP model is used, it is more efficient and takes less time to use all eight CPUs within a node. A large part of the slowdown is caused by system tasks having to interrupt the model in order to gain access to a CPU. This has a "knock-on" effect during times when all MPI tasks are trying to synchronise in a barrier. The hybrid model means that there are fewer MPI tasks and in general there is usually less contention with the system for CPUs.

Migration to the IBM HPCF: Performance improvements - *Deborah Salmond*

It was stressed that the IFS had been designed to be highly portable, with the same code for all systems, except for a very few exceptions, when the only way to achieve efficient performance was to provide alternative code for different architectures - e.g. the adjoint of semi-Lagrangian interpolation has alternative vector and scalar versions.

With the large number of processors on the HPCF, the 2-D decomposition - full pos post-processing - in grid point and spectral space works well. The messages to be passed in the IFS are long, so the critical factor in IFS message passing is bandwidth, rather than latency.

NPROMA

Within the high level control routines, where subroutines such as physics or other grid point calculations are called, there is a high level blocking scheme which allows the block length to be specified (nproma), so that the vector length can be chosen according to the machine architecture e.g. block lengths of 12 or 24 have been found to be most efficient on the IBM; vector machines perform well with nproma = 1000 or more.

Open MP

The NPROMA loops, large blocks of independent code, are an ideal location for parallelisation with OpenMP. The OpenMP parallel regions are inside the MPI parallelisation. It is very easy to parallelise using OpenMP but there are hidden dangers. The code within OpenMP is assumed, by the system, to be thread-safe, so the coder must ensure that none of the threads has any dependencies. IBM's auto-parallel compiler has not been used. In order to obtain worthwhile advantage from using OpenMP a high percentage of the loops must be parallelised. Scalability is affected by various factors, such as memory bandwidth - one processor can use 3.5 GB/s but an 8 processor LPAR has only 10 GB/s available; thread despatching causes a 20 microsecond overhead, so loops running for less than one millisecond are not worth parallelising; data being in the wrong cache and cache line interference also cause delays.

The advantages of using OpenMP in addition to MPI:

performance is improved;

the number of MPI tasks is reduced and, therefore, MPI communications (both semi-Lagrangian and global communications) are reduced;

memory use per node is reduced;

some automatic load balancing is carried out for each of the parallelised loops.

There is, however, a disadvantage in the maintainability of the code. Once high-level parallelisation has been implemented, it is essential that no thread unsafe code is inserted at a later stage. This requires strict discipline and at ECMWF this is achieved by limiting the number of people who are authorised to implement code changes.

ECMWF requirements

ECMWF has various particular requirements, which affect optimisation:

Error trapping must always be on, to ensure that problems are detected early; it is also essential to know at least in which subroutine the error has occurred and, preferably, in which line of code.

The code must continue to be able to run efficiently on a vector machine.

Bit-wise reproducibility on different numbers of processors, or at different NPROMA lengths is essential, so optimisation must be limited to -03 -qstrict, which is the highest level which strictly maintains the order in which the computations are performed. 64 bit arithmetic for reals and 64 bit addressing are used.

Hints for optimisation for the IBM p690

1. Add timing
   - use IBM Xprofiler (call graph and source statement profiling)
   - or ECMWF's GSTATS
2. Optimise data access
   - try to avoid copying or zeroing out arrays
   - combine multiple passes through the data

3. Be aware of the implications of bit reproducibility
   - use -qstrict
   - remove divides (by hand) and either replace them by reciprocals or put them outside the loop and identify common expressions
4. Use IBM MASS library (for EXP and LOG etc) and ESSL library (mainly for matrix multiplies)
5. Insert more OpenMP into the code.
6. Do not allow overlapping of communications with CPU. Buffer creation should be separated from message passing.
7. Use MPI global communications as much as possible.
8. Use low overhead Floating Point Trapping (developed at ECMWF from the IBM functions).

HPCF Communication Optimisation - *Peter Towers*

The p690 switch is much slower than the VPP5000 crossbar for a number of reasons:

- it has one eighth the bandwidth of the VPP5000 crossbar;
- it has twice the latency;
- it uses CPU cycles to move data into switch buffers;
- it uses a 'rendez-vous' protocol (not required by the VPP5000)
  which means that large 'sends' block until a 'receive' signal has arrived.

All these factors may have an impact upon the scalability of parallel applications. The following steps can be taken, to ensure that communications are fully optimised:

- check environment variables
  if message passing is used inside a node, set
  MP_SHARED_MEMORY=yes
  which will ensure that shared memory is used, otherwise the switch would be used.

Very short messages are sent via intermediate memory, avoiding the need to wait for a rendez-vous signal; the default maximum message length for using intermediate memory is 1kB; this should be increased to, say 64kB setting the environment variable:

MP_EAGER_LIMIT=64k

- Check LoadLeveler directive

This must be set correctly to ensure that the full switch capacity is used:
Network_MPI= csss, shared, US

- Run MPI trace tools

These tools were developed by IBM and have been installed at ECMWF in /usr/local/lib/trace. They use the standard MPI profiling interface and report on how much time is spent in sends, in receives, in barriers, the number of messages, size of messages and time taken to perform transfers, so that communication intensive 'hot spots' can be easily and quickly identified.

The Call Graph section analyses individual user applications and summarises the types of MPI routine called, number of times called, volume transferred and time taken in transfers, so that hot spots can similarly be identified for intensive investigation.

- Use non-blocking routines such as MPI_ISEND (Immediate Send). This avoids rendez-vous delays by initiating the send, then performing other tasks while the send takes place, returning in time for send completion.
- Use collective routines to reduce latency e.g.
  MPI_ALLGATHER
  MPI_BROADCAST
- Remove MPI_Barriers

MPI_Barriers are often inserted to aid in timing and statistics gathering but, if not removed before operational running, they can increase the effects of operating system interference and exacerbate any effects of load imbalance.

- Recode inefficient parts of applications

Applications originally developed to run on only a few processors may contain global communications operations which create bottlenecks once scaled up to run on hundreds or thousands of processors. Investigation may establish that the communications can be removed completely from the code and inserted elsewhere.

- Consider mixed mode MPI/OpenMP
  As described in the previous presentation.

In response to a query, D. Salmond confirmed that addressing and floating point numbers were 64 bits and integers 32 bits long.

Dissemination update - *John Hennessy*

At last year's meeting ECMWF reported that 575,000 products and 6.9 Gigabytes of data were disseminated daily. This year, that number has more than doubled to 1,200,000 products (11.5 Gigabytes) disseminated on the RMDCN with 85,000 products (4.4 Gigabytes) additionally disseminated via the Internet.

Many additional parameters have been introduced into the dissemination both for the deterministic forecast and the EPS. Products disseminated from the 00 UTC forecasts are now almost exactly the same as those from the 12 UTC run. Many new wave products have been introduced from the global and European shelf forecast from the Wave EPS. Representatives were reminded that detailed information on new products is available on the 'News' page of the web Dissemination pages. A unified catalogue for MARS and dissemination has been established, which means that any products available from MARS are also available from the dissemination. (The reverse, however, does not apply; for instance model level data products from the EPS are available as dissemination products but are not archived in MARS because of their enormous volume).

The operational migration to the IBM HPCF on 4 March went very smoothly, apart from analysis files being sent earlier than usual. New dissemination management tools on the web have been introduced. The tools are Java based and provide improved management and monitoring. A Help facility has been included to assist users and navigation has been simplified.

The option of dissemination via the Internet has been introduced in addition to the RMDCN dissemination. Should the Internet fail for any reason, the dissemination automatically reverts to RMDCN. This option has proved particularly useful for Boundary Condition project members, as twice per day the BC products are generated and ready for dissemination simultaneously with the main forecast products and would otherwise be competing for RMDCN bandwidth.

Plans

Currently, the products are disseminated in a particular sequence, according to a defined dissemination schedule. Some of these products are, in practice, available for some time before they are actually disseminated, so in the future it is planned to work to a production schedule, according to which, products will be available on the telecommunications system from a specified time. Member States will be provided with utilities to manage the sequence in which they receive their products and specify priorities for individual dissemination files. The overall intention is to allow Member States to receive dissemination products as early as possible.

It is planned to begin migration to GRIB Edition 2 this year. Work on encoding/decoding software for GRIB2, BUFR and CREX will begin in the near future. This software will be made generally available to any country in the world and to manufacturers of automatic observing equipment. The software will be Fortran and C-callable. It will run on Linux and on as many UNIX versions as possible.

In the past the atmospheric, global wave and European Shelf wave models were assigned different version numbers, and some users hard coded checks to distinguish between them. This caused problems when model version numbers were changed, so each model was allocated a range of version nos viz;

Atmospheric model	121-203
Global ware model	104-120
European Shelf ware model	204-220

Representatives were reminded that the Atmospheric Model was currently at version 203 and that the next change would revert to version 121. The other models are not yet close to the end of their version ranges.

Questions

K. Holmlund (EUMETSAT) asked whether ECMWF had any plans to archive BC data. ECMWF replied that at the inception of the BC project archiving of the products was specifically excluded. Currently, the cost of a short-term archive is being investigated.

R. Rudsar (Norway) enquired about automatic back-up via the RMDCN, if there were problems on the Internet: would it be possible to reduce the amount of data to be transmitted via the RMDCN? ECMWF replied that this would not be possible, as the dissemination files would already have been created. However, if smaller RMDCN files were created and assigned a lower priority, then they would not interfere with the routine Internet dissemination. It would be possible to routinely generate two streams: a large one for the Internet and a smaller one for RMDCN, created in delayed mode. The RMDCN stream would be created daily but would never be sent unless manually initiated.

MARS status - *Baudouin Raoult*

Perhaps the most important development in MARS for the Member States is the ability for MARS stand-alone client and the MARS Metview module to access the archive via ECaccess.

MARS now also supports Member States' archiving on a MARS server dedicated to the Member States. The system is operational for DWD, Germany, and in the final stages of testing for Italy. A system for the COSMO-LEPS project is currently being designed.

In conjunction with the installation of the new Data Handling System all data archives must be migrated to the HPSS. So far, all research data, all ERA-15, Demeter, PROVOST and other such projects' data have been transferred. The migration of ERA-40 data is in progress. Once this is finished, operational test suites and other operational data will be transferred. The migration is approximately 60% complete.

A public data server has been made available. It is a small Linux MARS server running outside the ECMWF firewall. It provides public access to some of the ECMWF data, which can be retrieved in GRIB, NetCDF or as plots. The server hosts the Demeter dataset and the 2.5° ERA-15 dataset. In the future, the ERA-40 dataset at 2.5° will be included.

Archive volumes

- Operations 131 TB
- E-Suites 83 TB
- Research 299TB
- Member States (IFS) 6 TB
- Member States 6 TB
- ERA (15 & 40) 59 TB
- Other 2 TB

In response to an earlier query on the most efficient way of using MARS, representatives were referred to:

http://www.ecmwf.int/publications/manuals/mars/guide/Request_scheduling.html

A general 'rule of thumb' was also given:

1 month at a time

All levels

All parameters

All time steps

All ensemble members

For a very large request of ensemble data this might have to be broken down into two fortnights of data or four weeks of data, but on the whole, the larger the volume retrieved, the better the scheduling.

It is also possible to retrieve observations via the web.

The Data Finder system has been created to assist users in searching for any operational or re-analysis data. (ERA-15 is already available, ERA-40 is currently being added).

MARS information sources

The web catalogue is automatically, dynamically updated from the MARS metadata. This means that it reflects the contents of the archive exactly but, being based on MARS metadata, is not in the most convenient format.

The "data finder" is updated daily and includes all data types such as monthly forecasts, seasonal forecasts etc. The "last archive" page is generated in the same process as the data finder, so reflects the same information.

Some users have asked why MARS does no checking for valid requests. ECMWF had considered this but were concerned that it would be difficult to avoid rejecting unusual, yet valid, requests. Moreover, error monitoring has established that very few mistakes are made.

Users have asked why they sometimes have to wait for a MARS request for some time. There are four broad reasons for delays:

- MARS is heavily used, dealing, on average, with 100,000 requests per day (40,000 on the DHS). Approximately 500 GB are archived daily and 555 GB retrieved from the DHS, 350 GB from the FDB. In spring this year the old and new DHSs were running in parallel, archiving everything twice.

Users can view the status of the MARS queues at:

http://www.ecmwf.int/services/archive/d/activity

- Inefficient retrievals

Many users understand how to use the system efficiently and, in one day for instance, 100 GB is retrieved in 200 requests. On the other hand, 0.5 GB may be requested on the same day in 10,000 requests.

- Migration to the new DHS

The new DHS has completely new hardware, including the disks. Much time has been and continues to be spent in tuning the software to the new equipment. The whole of the back archive has had to be copied from the old TSM storage to the new HPSS. Four tape drives have been used continuously to achieve this, and so are not available for users' requests. The ERA-40 data comprises data from many different sources, with many different version numbers. The whole ERA-40 dataset is being systematically retrieved, so that all GRIB headers can be made to match and all version numbers are being set to 1.

- Cache issues

The MARS archive comprises tapes and disks. Various datastreams have differing levels of cache hits. For the operational analysis and forecasts most requests are for the most recent data, so there is an 82% or 63% hit rate respectively. The operational EPS is not so popular, so has only a 20% cache hit rate. For ERA-15 and ERA-40 the hit rates are even lower (15% and 6% respectively) firstly because the archives are huge (33 Terabytes for ERA 40) compared to the size of the cache (1.2 TB) and secondly because users' requests are spread over the whole reanalysis period of 40-45 years.

Graphics update - *Jens Daabeck*

New features introduced into Magics include:

- New coastlines based on the NOAA GSHHS database which has now become the default for both Magics and Metview.
- Support for data in Polar Stereographic Grib (NESDIS).
- The Magics test and installation procedures have been improved.
- Use of shared Magics libraries under the IBM AIX operating system is being considered but so far, this does not appear to give any great performance advantage.

Currently, there is limited support for NetCDF in Metview. It is planned to expand this and also provide some support in Magics.

The new level of export Magics (6.8) will be made available in June.

There are versions for the following UNIX platforms:

Linux	SuSE 7.3 & 8.0 (Portland Fortran compiler)
IBM	AIX 4.3
SGI	IRIX 6.5
HP	HP-UX B.11
HP/Alpha	OSF1 V5.1
Sun	SunOS 5.9

User Guide in HTML, PDF and PostScript format.

New-Magics

Work on the phased migration of the Magics Library from Fortran to C++ continues. The internal structure is being analysed and optimised but it is intended that existing Magics user programs will need minimal changes to use New-Magics.

Internally, Magics comprises various independent modules, for contouring, observations, wind plotting etc, so it is possible to introduce the changes gradually, module by module. The design of New-Magics is based on the style of Metview, which will make it easier to integrate into Metview, of which it is an important component, in the future. Figure 3 represents the New-Magics design.

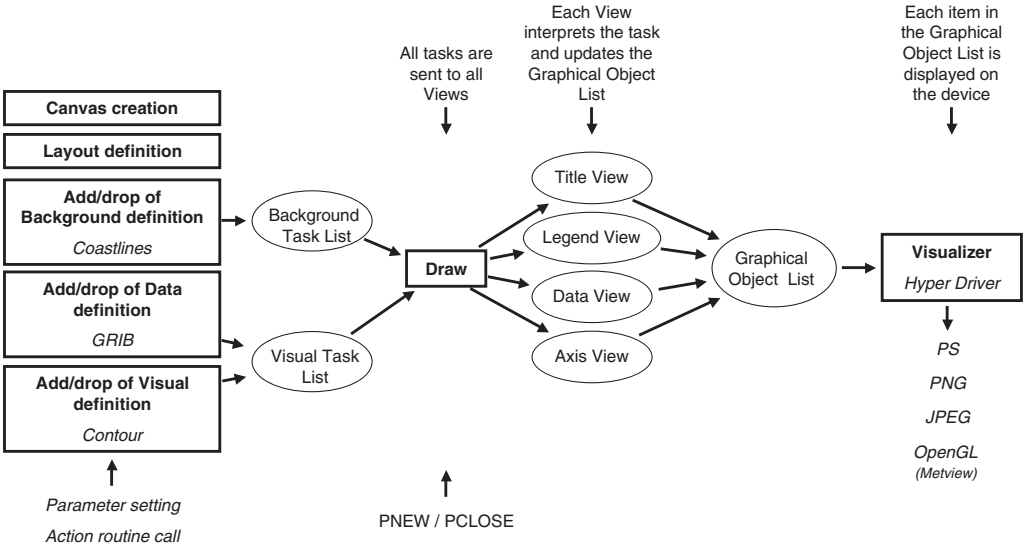


Fig. 3 New-Magics Design

The Akima method has been chosen for contouring in New-Magics. It differs fundamentally from the Conicon contouring method. Akima is basically an interpolation technique, coupled with simple contouring.

It is hoped to demonstrate New-Magics at the 9th Meteorological Operational Systems Workshop in November this year. It is planned to have a prototype running in the first quarter of 2004 and a pre-operational release in the second quarter of 2004.

Metview Update

New-Mars access scheme

ECaccess is replacing ECbatch. If Member States' Metview is built with MARS_remote access to ECMWF Mars archives, then the Metview MARS modules and the MARS configuration file to which etc/MARS_remote is pointing will need to be updated. (If Metview is built with MARS_none, then no changes to the Metview MARS module are needed).

A distribution package containing the updated Metview MARS module will be available from Data Services from this June. It is suitable for Metview 2 and 3.

If an external MARS client is used as a common cache holder for all Metview sessions, then this software, too, will need updating. A fully updated Metview 3.5 export distribution package is planned to be available for the third quarter of 2003.

New features in Metview

- Support for handling vector type point data format (GeoVectors)
- ODB support
- Examination of geopoint values and satellite pixel values
- Running of several Metview sessions simultaneously
- Support for missing data values in Grib fields has been added in macros
- Support for missing data values in BUFR messages has been added to geopoints and the ObsFilter editor.
- Control over plotting grid point values has been improved
- A Metview fix log file has been introduced for better feedback to users reporting problems
- A trajectory model has been integrated as part of Metview, providing Metview users with a fast and easy facility to compute and visualise trajectory data.

Plans include:

- the development of support for various ODB applications
- support for high volume satellite data
- support for New-Magics
- support for the new MARS access scheme

Metview 3.5 export will be available to the Member State in the third quarter of 2003.

There will be versions for the following UNIX platforms:

Linux	SuSE 7.3 & 8.0 (Portland Fortran compiled)
IBM	AIX 4.3
SGI	IRIX 6.5
HP	HP-UX B.11
HP/Alpha	OXF1 V5.1
Sun	SunOS 5.9

H. Baldursdottir (Iceland) asked whether there were any plans to port Metview to Windows. ECMWF replied that there were not.

She also observed that Metview was complicated to install, were there any plans to make installation easier? ECMWF replied that Metview was inevitably complex because of the complex functions it fulfilled. Binary distribution had been tried in the Netherlands but was unsuccessful. If Member States have useful hints on installation or have made a Metview version for a new platform, they are requested to pass the details to the Graphics section, who will include the information in the Metview web pages.

J. Daabeck noted that support for the trajectory model could be included in the next Metview upgrade, if there were demand from the Member States but they would need access authorisation to the required data.

ECaccess - project status – *Laurent Gougeon*

Thirteen ECaccess gateways are now installed on servers in Member States and two Member States use ECaccess facilities via the two Gateways installed at ECMWF (one dedicated to Internet access and one to the RMDCN). The latest release of ECaccess includes an embedded database engine which gathers local statistics.

ECtrans has been enhanced to allow the reception of files from the Member State in addition to the sending. Files can be received on the system running the Gateway, on a system running an FTP server or into a local Member State proprietary archive.

The web interface now allows users to specify the host name, login username and password of the system in which they want to send and receive files. The passwords are encrypted in the database of the remote Gateway.

New ECaccess Plugins

- SSH Plugin
  Provides users with a Secure Shell (SSH) access
  Implements SSH protocol version 2
  Supports OpenSSH and SSH.COM clients
  Allows secure X11 connections
- VNC Plugin (Virtual Network Computing)
  Allows X11 applications to start using a VNC viewer (a thin-client application)
- The VNC viewer is about 150K in size and can be run directly from a floppy
- The VNC viewer exists for Windows, UNIX and Java platforms (e.g. within a Web browser)
- It is free (http://www.uk.research.att.com/vnc)
- The VNC plugin uses the SSH plugin to secure connections

ECaccess Certificates

The validity duration of ECaccess certificates has been refined, so that each service is classified by category and has a specific duration.

File Management	copyFile, getFile, putFile, readFile writeFile, moveFile, deleteFile, chmod, getFileSize, getFileList, getTempFile, getFileLastModified, mkdir, rmdir.	7 days
Job Management	submitJob, deleteJob, getJobList, getJobSize, getJobResult.	7 days
Queue Management:	getQueue, getQueueList	7 days
Transfer Management:	getTransferList, frlrtrTransfer, getTransferResult, retryTransfer, spoolTransfer	7 days
Other:	useECmars	1 month

Durations can be specified per category and/or action and/or user; they are maintained at ECMWF in the ECaccess database of the ECcmd daemon.

ECaccess gateway administration has been facilitated. There is now a web interface (JMX) and each plugin can be managed and monitored independently.

An HSQL database manager provides ECtrans statistics and management and monitoring tools.

A log manager (log4j) can be used to monitor the operation of the Gateway.

ECaccess Plans

- Integrate LoadLeveler to ECaccess for submission of jobs to the HPCF
- Introduce service routing to allow the administrator of the Gateway to specify a network (Internet or RMDCN) per service and according to the time of day
- ECaccess Registration Authority (ECRA) to create a distinct ECRA for each Member State and allow mapping between local and ECMWF users
- ECaccess symmetry to provide Member State users with ECtools (ecget. ecput etc) running either on their Member State workstation or the ECMWF servers

Questions

M. Pithon (France) asked when it was anticipated that job submission facilities to the HPCF would be available. ECMWF replied that its development was a high priority and it was hoped to be available in 1-2 months.

P. Halton (Ireland) asked the specification for the Gateway server in a Member State with 20-30 users. ECMWF replied that the system would need to run on a system with Java Virtual Machine version 1.3 . For 20-30 users at least 256 MB of memory would need to be reserved for the Gateway application.

RMDCN Update - *Tony Bakker*

T. Bakker reported that the RMDCN contract had been amended in June 2002 to allow the inclusion of non-WMO RA VI members.

The United Arab Emirates was the first to take advantage of this enlargement, signing its Accession Agreement in August 2002 for a 64 kbps PVC to Germany, 16/16 kbps, connected in November 2002; accepted in February 2003.

China signed its Accession Agreement in September 2002 for a 64 kbps PVC to Germany, 16/16 kbps, accepted in April 2003. Japan is currently in discussion with EQUANT concerning a connection to China.

An overview of the other most recent connections and upgrades was given, along with a summary of core network outages and access line problems.

The second contracted price review took place in July 2002. It was established that only PVC prices had reduced (by approx. 15%), other cost items had remained unchanged. The overall agreed reduction was, therefore, 9%.

There were significant amendments to the EQUANT contract on 1 January 2003:

- ECMWF waived its right to a 5 year breakpoint; the contract will run until March 2008
- Annual price reviews will take place
- A price review, in addition to the review carried out in July 2002, led to an approx. 27.5% reduction in overall prices; individual countries may receive slightly greater or smaller reductions in their prices
- A cost neutral bandwidth upgrade option was introduced. Rather than receive the calculated price reductions, RMDCN sites can opt to continue to pay 1Q 2002 prices and receive service upgrades.

ECMWF has decided to use this option to upgrade the basic RMDCN telecommunications package which is supplied to ECMWF Member States as part of their membership benefits.

The table below compares the old and new standard packages.

	Current (Since March 2000)	New (Planned for October 2003)
Access Line/Port	128 kbps	384 kbps
ISDN Backup	128 kbps	384 kbps
PVC ECMWF (in/out)	96/96 kbps	256/128 kbps
PCV to other ECMWF Member State	2* 16/16 kbps	2* 32/32 kbps

Virtual Private Network (VPN) feasibility study

In order to ascertain the feasibility of using the Internet as backup to the RMDCN, a series of IPSEC based VPN tests were carried out over the Internet in 2002 with France, Germany, Greece and the Netherlands. A variety of equipment was used including Cisco router, PIX Firewall, Checkpoint Firewall and FreeS/Wan (Linux) which enabled its interoperability to be tested.

As a result of the tests, the following recommendations were made:

- X509 certificates should be used for device authentication, as they are the most secure and most scalable option
- Use ESP HMAC for data integrity authentication
- Use ESP NULL or ESP 3DES for data encryption
- Ensure that IPSec equipment supports: IKE key exchange protocol and X509 certificates;
  ESP_NULL and ESP 3 DES;
  preferably hardware encryption acceleration cards available.

The recommendations and results of the feasibility study have been published in an ECMWF Technical Note, May 2003: IPSec Feasibility Study (in co-operation with DWD, Météo France, HNMS and KNMI): Summary and Recommendations.

The next stage of the trial will be to establish operational site-to-site VPN connections for dissemination.

Questions

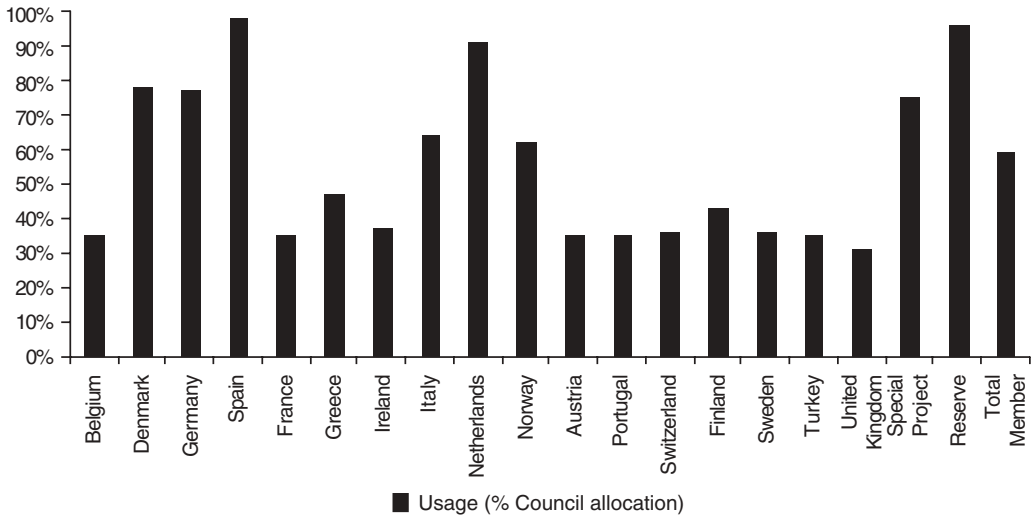
ECMWF confirmed that those Member States which are already paying for services equal to or greater than the increase in standard funded package will receive a reduction in charges, as a greater proportion of their charges will be covered by the standard fees paid as part of the ECMWF budget.

User Services: status and plans - *Umberto Modigliani*

U. Modigliani summarised developments over the past year, noting that much effort continues to be invested into putting material onto the web. He invited comments and suggestions for improvements to the website.

The migration to the new HPCF went well. Four training courses were organised and attended by forty Member State users. Over ninety users executed jobs on HPCA in the first four months of 2003.

The table below gives an overall summary of Member States' usage of the supercomputer in 2002.



It was noted that Member States' use of ECMWF computing facilities had increased significantly. There are now over 1,200 registered users and approximately 65% of them connected to ECMWF in the last year. In the last 3 years, the number of users of ecgate1 has doubled.

Representatives were reminded of some general advice:

- Before retrieving large amounts of data from MARS, please ask for advice
- Consider using the compiler optimisation flags for more computing intensive programs, especially if run frequently
- Consider moving very computing intensive codes to the HPCF
- Please contact ECMWF if you would like to allocate bandwidth to certain traffic on your RMDCN line (dissemination versus user related data transfer)

Discontinuation of telnet, ftp, X11 gateways

The Centre will discontinue the telnet, ftp, and X11 gateways for access to ECMWF via the Internet at the end of November 2003. The following table compares the old and new services.

telnet	tn-gw.ecmwf.int	telnet	ecaccess.ecmwf.int
ftp	ftp-gw.ecmwf.int	ftp	ecaccess.ecmwf.int
xhost	+x-gw.ecmwf.int	xhost	+ecaccess.ecmwf.int
telnet	x-gw.ecmwf.int	telnet	ecaccess.ecmwf.int
x-gw	<your host IP address>	ecxterm	

Access via RMDCN will not change

Discontinuation of ecbatch/eccopy services

- The Centre will discontinue the ecbatch/eccopy services at the end of November 2003
- ECaccess provides all the services provided by ecbatch/eccopy including support for standalone MARS client "ecmars"
- If you are using ecmars you will need to install the new ecmars software and at least the ECaccess tools
- If you have access to MARS via Metview you will need to install the new MARS module which will be available in June 2003

Changes to the accounting formula

A new formula has been created for all parallel jobs. It is designed to reflect actual usage, whilst encouraging optimal usage but remaining fair.

New formula for all parallel jobs to be implemented in June 2003:

SBU = 88600* (Tot Elapsed Time/952* 86400).

There are no accounting changes for serial/single task jobs.

Entity Management System - *Petra Kogel*

For some time the need to simplify and rationalise various user permission and registration systems has been recognised. The Entity Management System (EMS) has been devised as the unified solution. It is based on a distributed database with a core on a high availability cluster. It will be kept small, simple and flexible with an expected data volume of less than 200 MB. Figure 5 shows the Entity Management System architecture.

PostgreSQL, a Public Domain version of INGRES, has been chosen for its speed and comparative sophistication, though any developments will be kept as generic as possible, to allow migration, if ever another solution becomes more suitable than PostgreSQL.

Security will be strict: authentication will be by means of a Java based ticketing system for access via the web and by an identification system for access from Unix.

Authorisation to access various services - the primary purpose of this system - is based on a codified access policy system, which will also control access to the EMS itself. There will be no direct user access to the central database.

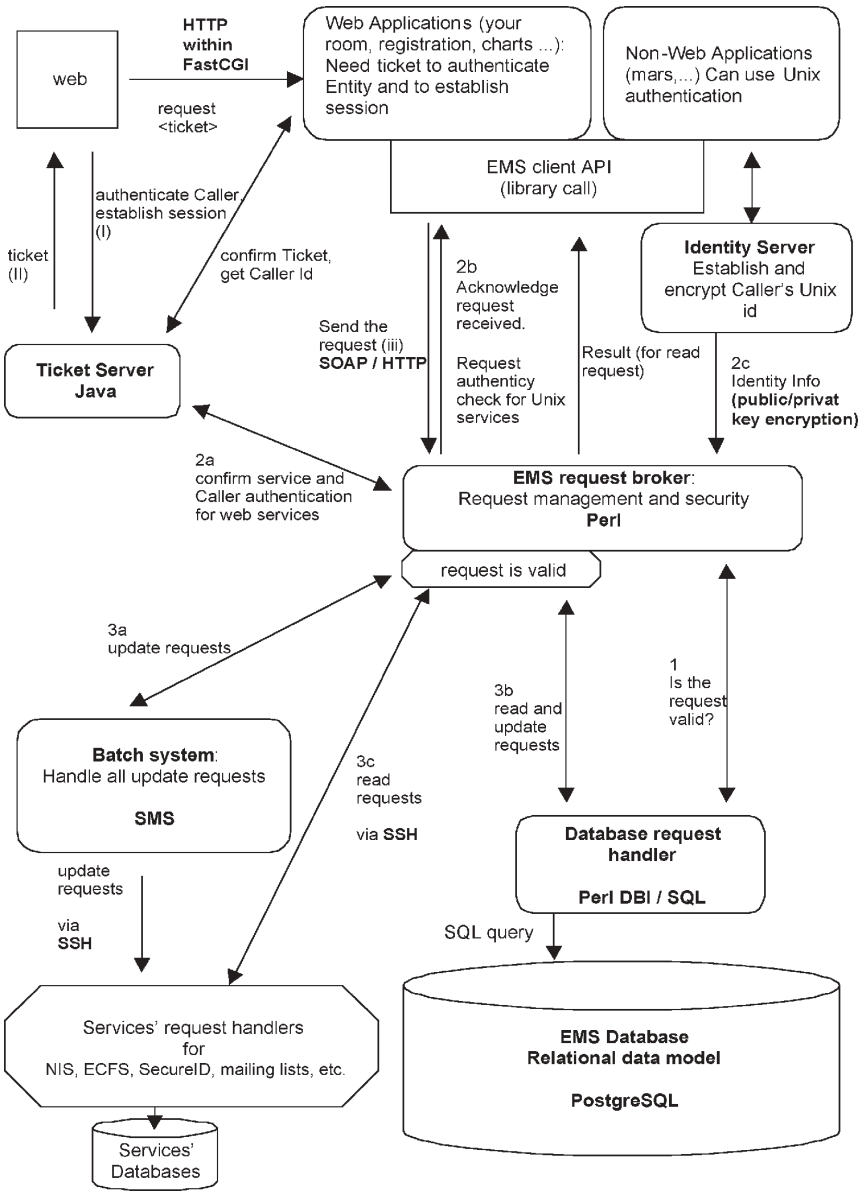


Fig. 5 Entity Management System architecture

Access to data will be strictly controlled by the request broker. All communications are secure. Write requests will be handled by the SMS-based batch system for managing the distributed database. Read requests will be handled directly.

Communication with the EMS server will be via SOAP/HTTP. Communication with services such as NIS, ECFS etc will be via SSH. Identification information will be secured by encryption, using a private/public key technology, which is, as yet, not selected.

The system is designed to be highly reliable. The central components: ticketing system request broker, EMS database and batch system will form a service package on a high availability cluster.

It is planned to have the kernel working by the summer for use by the ECMWF Call Desk. A web-based interface will then be developed for Computer Representatives' use.

DISCUSSION - TIME OF NEXT MEETING

It was unanimously agreed that the next meeting of Computer Representatives should take place in 12 months' time (May 2004). It was the general preference that the meeting run from Monday, 10 am to Tuesday lunchtime. The dates of the meeting will be confirmed later.

PART II

Member States' and Cooperating States' Presentations

AUSTRIA AUSTRIA

Gunter Wihl, Central Institute of Meteorology and Geodynamics, Vienna

Computer equipment

a) Production Server:
SUN Server 420, 2 CPUs/450 MHz, 2GB Memory, Disk 2 GB, CD-ROM
SUN Server 420, 2 CPUs/450 MHz, 2GB Memory, Disk 2 GB, CD-ROM

b) Development Server:
SUN Server 420, 4 CPUs/450 MHz, 2GB Memory, Disk 2*18 GB Raid1, CD-ROM

c) Fileserver:
NET APPLIANCE Network Attached Storage, Disk 500 GB proprietary Raid (~Raid 4)

d) Short-Range_Database Server:
SUN Ultra Enterprise 450 Server, 2 CPUs/300MHz, 2 GB Memory, Disk 4*9.1 GB, CD-ROM, Floppy 3.5″
SUN Ultra Enterprise 450 Server, 2 CPUs/300MHz, 2 GB Memory, Disk 4*9.1 GB, CD-ROM, Floppy 3.5″

e) Long-Range_Database Server:
SUN Enterprise E3500 Server, 4 CPUs/336 MHz, 2GB Memory, Disk 4*9.1 GB, CD-ROM
SUN StorEdge A3500 Disk Array, Disk 2 x 51*9.1 GB
SUN Enterprise E3500 Server, 4 CPUs/336 MHz, 2GB Memory, Disk 4*9.1 GB, CD-ROM
SUN StorEdge A3500 Disk Array, Disk 2 x 51*9.1 GB

f) ECMWF-Server:
SUN Ultra-10, 1 CPU/440 MHz, 524 MB Memory, Disk 2*19 GB, CD-ROM
SUN Ultra-10, 1 CPU/440 MHz, 524 MB Memory, Disk 2*19 GB, CD-ROM

g) GTS-Server:
SUN Ultra-10, 1 CPU/440 MHz, 524 MB Memory, Disk 2*19 GB, CD-ROM
SUN Ultra-10, 1 CPU/440 MHz, 524 MB Memory, Disk 2*19 GB, CD-ROM

h) Internet- and Product Server:
SUN LX50 Server, 2 CPUs/1.4GHz, 1 GB Memory,, Disk 2*72 GB, CD-ROM
SUN LX50 Server, 2 CPUs/1.4GHz, 1 GB Memory,, Disk 2*72 GB, CD-ROM
SUN LX50 Server, 1 CPU/1.4GHz, 512 MB Memory,, Disk 2*36 GB, CD-ROM

i) Intranet-Server:
SUN Ultra-1, 1 CPU, 65 MB Memory, Disk 10.5 GB, CD-ROM

j) Domainname-, Administration- and Operating Server:
SUN Ultra 5_10, 1 CPU, 132 MB Memory, Disk 5.2 GB, CD-ROM
*SUN Ultra-1, 1 CPU, 65 MB Memory, Disk 4.2 GB, CD-ROM

k) Mail-Server:
SUN Netra T1, 1 CPU/500 MHz, 512 MB Memory, Disk 18 GB, CD-ROM
SUN Netra T1, 1 CPU/500 MHz, 512 MB Memory, Disk 18 GB, CD-ROM
SUN Netra st D130, Disk 2*36 GB

l) Backup- / Archive-Server:
SUN Enterprise 250 Server, 2 CPUs, 128 MB Memory, Disk 26.4 GB
Single Equipment with double Access:
DLT Cartridge Roboter (3.5 TB, 4 drives)
Single Equipment:
Tape 0.5″, 9-track, (6250/3000/1600/800 bpi)
Optical Disk Roboter (4 Drives, 144 Slots re-writeable Magneto-Optical-Disk, 650 MB Cartridge)

m) RC-LACE Model Group:
Digital Personal Workstation 600 AU, 1 CPU, 1 GB Memory, Disk 8.6 GB, CD-ROM, Tape 4 mm DAT
SGI Origin 3400, 20 x R14000 CPUs/500MHz, 20 GB Memory, Disk 2*18 GB, 8*73 GB, Tape 4 mm DAT

n) FIREWALL:
XXXXXXX, Confidential

and more than 60 other Servers and Clients depending on special needs at the several Departments and Regional Services of ZAMG, and a flock of nearly 300 PCs, some of them used for routine work, e.g. for forecasters and to supply special Media (Broadcast and Television, Newspapers).

AUSTRIA AUSTRIA

Software

SUN-Systems

Operating System:	Solaris (UNIX)
Compiler:	Fortran 77, 90, 95, C, ANSI C, C++
Script language:	Perl
Graphics:	Xelion GKS, MAGICS, PV-Wave, OpenGL
Libraries:	IMSL, NAg
Database:	SYBASE
GIS:	ARC/INFO
Backup SW:	Veritas Netbackup
e-mail:	Z-mail

LX50

Operating System:	Sun Linux

Digital Workstation

Operating System:	Digital UNIX
Compiler:	Fortran 90, C++
Graphics:	NCAR Graphics

SGI-System

Operating System:	IRIX64
Compiler:	Fortran 77, 90, C, C++
Graphics:	NCAR Graphics

Personal Computer

Operating System:	Windows NT, Windows 2000, Linux (SuSe, REDHAT),MacOS
Compiler:	Fortran, Visual Basic, C
Graphics:	Xelion GKS, MAGICS
Applications:	MS Office, ABC Flowcharter, ARC/VIEW, CorelDraw,Exchange, Exceed, PageMaker, PhotoShop, SYBASE ODBC,OnNet interdrive
Internet/e-mail:	Netscape, Internet Explorer, Outlook / Outlook Express

The Usage of ECMWF-Data

1) Processing the ECMWF-Data

Most products are transmitted to Austria on a daily basis by using the ECMWF-Dissemination System, but the EPS-Meteograms and the Sesonal Forecasts are transmitted by a user job running on ecgate1.

The ECMWF-Data are copied to the Servers zaaecm1 and zaaecm2 in Vienna (zaaecm2 is used when the zaaecm1 has broken down).

On the zaaecm1 and -2 the data are checked and sent to the multi-user server zaamus1p (coupled with zaamus2p) for public use, when they are available and correct.

On the zaamus1p and -2p the derived products are generated and disseminated (e.g. QFA's for selected locations and weather charts using Magics).

The software consists of Korn-Shell Scripts (Solaris Unix-System) for operational control of the data and Fortran77-Programs for deriving special products. The gribdata are processed by using the GRIBEX-Software.

The special products are stored on ASCII-Files and Gribfiles with a local table for Vienna, which has been established in Vienna.

Corresponding grib data are also available from the DWD (German Weather Service), based on the obervations at 00, 12 and 18 UTC, but only up to 72 hours after the base time and with fewer elements and no Ensemble products.

The Programmes written for ECMWF-Data are applied also to the DWD-Data.

Additionally also the Output of the ALADIN-Modell is used (fine meshedmodel, but only for middle europe and two forecast days).

AUSTRIA AUSTRIA

2) Users of ECMWF Products

ECMWF-Data are used by the Austria Weather Services internally (operational and scientific purposes) and to derive products for private customers :

2 a) Internal use within the Austrian Met Services for operational products

- Central Institute and Regional Departments in Innsbruck, Salzburg, Klagenfurt and Graz ("ZAMG")
- Flight Meteorology ("ACG" - Austro Control)
- Military Weather Service ("Bundesheer")

Maybe the firtst two Services will be combined to one (private) Weather Service "Met Austria" in the near future.

2 b) Scientific purposes (special projects)

- internal use in the Central Institute of Meteorology (e.g. the Model Working Group and the Satellite Working group)
- Institutes University of Vienn, University of Agriculture in Vienna, University Innbruck, University Graz (Steinacker, Haimberger; Kromp-Kolb; Mair; Kirchengast and advanced students)

2 c) Private and public customers (derived products), e.g.

- ORF - Austrian Broadcasting Corporation
- local authorities
- some newspapers
- organizers of sport and culture events
- tourist traffic offices
- street services
- environmental purposes

ECMWF-Products in Austria

The following data streams are currently used by Austrian Weather Services:

1) A0D-Data : Main Production Stream (T511 gribdata), described in detail

The A0D-Data contain data from 12 UTC run at ECMWF from the T511 model

- deterministic forecasts on pressure levels 1000-200 HPA

Temperature, geopotential, u/v/w-wind, relative humidity

- surface parameters

cloud amount (total, low, medium, high), precipitation amount (large scale, convective), 2M temperature/dewpoint, 10M u/v-Wind, surface temperature, sea level pressure, snowfall, snow depth)

- Area 90W - 90E, 90N - 18N with 1.5*1.5 deg grid point distance (121*49 gridpoints in longitude and latitude)
- Analyses 18 00 06 12 UTC and forecasts from +6 to +240 hours after base time in 6-hr intervals
- a smaller set of global data is also available (forecasts from +12 to +96 hr) with 8 elements (500 HPA geopotential, 10M u/v wind, cloud cover (general), 2M temperature/dewpoint, sea level pressure, total precipitation)

A number of special products are derived from the A0D-Data

- several additional elements derived from the A0D-Gribdata (e.g. the Baroclinic Zone Indicator, a specific Humididy Index, Temperature Advection, Showalter-Index, level with 0 Celsius temperature and others)
- QFA-forecasts for 500 different locations in the whole world (Qfa's are ECMWF-forecasts interpolated from the gridpoints to the locations) (e.g. the newest application: the Austrian Television shows weather forcast charts for Europe derived from the ECMWF QFA's daily in the evening news since May 2002)

- graphical products (weather charts for Europe and Austria by using MAGICS) as hardcopies and visualized in the Intranet Server. Mainly used are charts for Europe/North Atlantic for sea level pressure, geopotential 850 and 500 HPA, equivalent relative topography, humidity index, precipitation, temperature and vorticity advection, baroclinic zone indicator, total cloud cover, temperature 2M and 850 HPA.
- a new application of A0D-Date comprises special forecasts of temperatur and precipitation amount for selected regions (are used by the main electric supply company for managing hydroelectric power plants)
- the Satellite Working Group uses the A0D-Data to combine them with the satellite weather images producing a special analysis, called "satrep".
- the Model Working Group uses the A0D-Data for verification and compares them with the data from other forecast models
- A0D-Data are used also as input for AUSTROMOS (Model Output Statistics for Austria) forecasting local weather for 112 weather stations in Austria and 150 locations (main towns) outside of Austria

(temperature, relative humidity, 10M wind, probability of catagories of precipitation and cloud amount, probability of fog and thunderstorm, ceiling and visibility)

- selected products (QFA's and MOS-Data) are stored in a data bank used internally for deriving products for specific customers

The other gribdata streams are used only for special purposes and are only shortly described:

2) A4D-Data: Main Production Stream (T511 gribdata) for 00Z run

The A4D-Data Stream contains the same data as A0D-Data, but refer to the 00 UTC model run, the A4D-Data are processed in the same way as the A0D-data.

3) A3D-Data : High Resolution data for Austria from 12Z run

The A3D-Data contains the same parameters as the A0D/A4D-data stream, but only for the Area of Middle Europe (6-21 deg longitude, 53-44 deg latitude) and a grid point distance of 0.5 deg instead of 1.5 deg.

The same products are derived from A3D-Data as from the A0D-Data.

4) A6D-Data : High Resolution data for Austria from 00Z run

The A6D-Data are equivalent to the A3D-Data, but use the model output from the 00 UTC run of T511-model.

5) A1D-Data : Gribdata on Model Levels

The A1D-Data contain deterministic forecasts for the same region as A0D but on model levels instead of pressure levels and with a higher resolution (1*1 deg) from +6 to +84 hours in 6-hr interval. They are used in the Environmental Department of the Central Institute to compute trajectories.

6) A0E-Data : Ensemble forecasts on grid points

The A0E-Data Stream contains cluster means, ensemble means and standard deviation and probability forecasts for the same region as the A0D-Data for the day 1 to 7 at 12 UTC (precipitation, temperature 850 HPA, geopotential 1000/500 HPA).

The A0E-data are processed with MAGICS and shown as graphical weather charts (only internal use).

7) AYA/AYB-Data : Weather Parameters (the only product in bufrcode)

The AYA and AYE Data are deterministic (AYA) and EPS-Forecasts (AYB, perturbed and control forecasts) of temperature for Vienna used by the Vienna Main Heating plant ("Fernwärme").

8) A0M-Data : Deterministic Wave Produkts for Mediterranean Sea

The A0M-Data contain forecasts of significant wave height and mean wave direction for the days 1-3 at 12 UTC in the mediterranean area. They are used, together with the A0D-Data, to generate sea weather forecasts for private (sailing) boats and are published in the Austrian HF-Radio.

AUSTRIA AUSTRIA

9) A2D/A2E-Data : Forecasts for Italy

The A2D and A2E-Data contain deterministic (A2D) and EPS Forecasts (A2E, Ensembles 1-50) for a grid point in Northern Italy and is sent to ARPA (Regional Weather Services in South Tyrol / Alto Adige and Friul Julish Venetia / Friuli Venetia Julia) per e-Mail (12 to 240 hours after base time in 12-hr intervals).

10) APE-Data : precipitation forecasts for AUSTRIA

The APE-Data contain precipitation forecasts for the area of Austria from the EPS-System.

11) A3E-Data : Ensemble Forecasts for Austria

The APE-Data contain Ensemble Forecasts for the area of Austria (ensembles 1-50, ensemble means and standard deviations, cluster means and standard deviations, forecast probabilities). Additionally, graphical products als Postscript-Files are transmitted to Austria by user jobs :

12) EPS-Meteograms (Postscript Files)

EPS-Meteograms are drawn for Vienna and the capitals of the 8 districts in Austria by a user job (ecgate1) and are transmitted to Austria by FTP printed out and visualized in the Intranet Server. In the last months the generating of EPS-Meteograms has failed for several times due to new data structures, missing access rights and installation of new Versions of Metview unknown to the users.

13) Seasonal forecasts

At the middle of a month the actual graphical Seasonal Forecasts are sent to Austria with FTP as Postskript Files and are visualized in the Intranet-Server (only restricted internal use for special users and scientific purposes) :Precipitation, 2m-temperature, surface temperatures (Sea) and mean sea level pressure (global or for individual continents; probabilities and ensemble means).

Plans for the next year(s)

1) Extension of the usage of EPS-Products

2) Implementation of the data from the 00Z run in operational weather forecasts in Austria

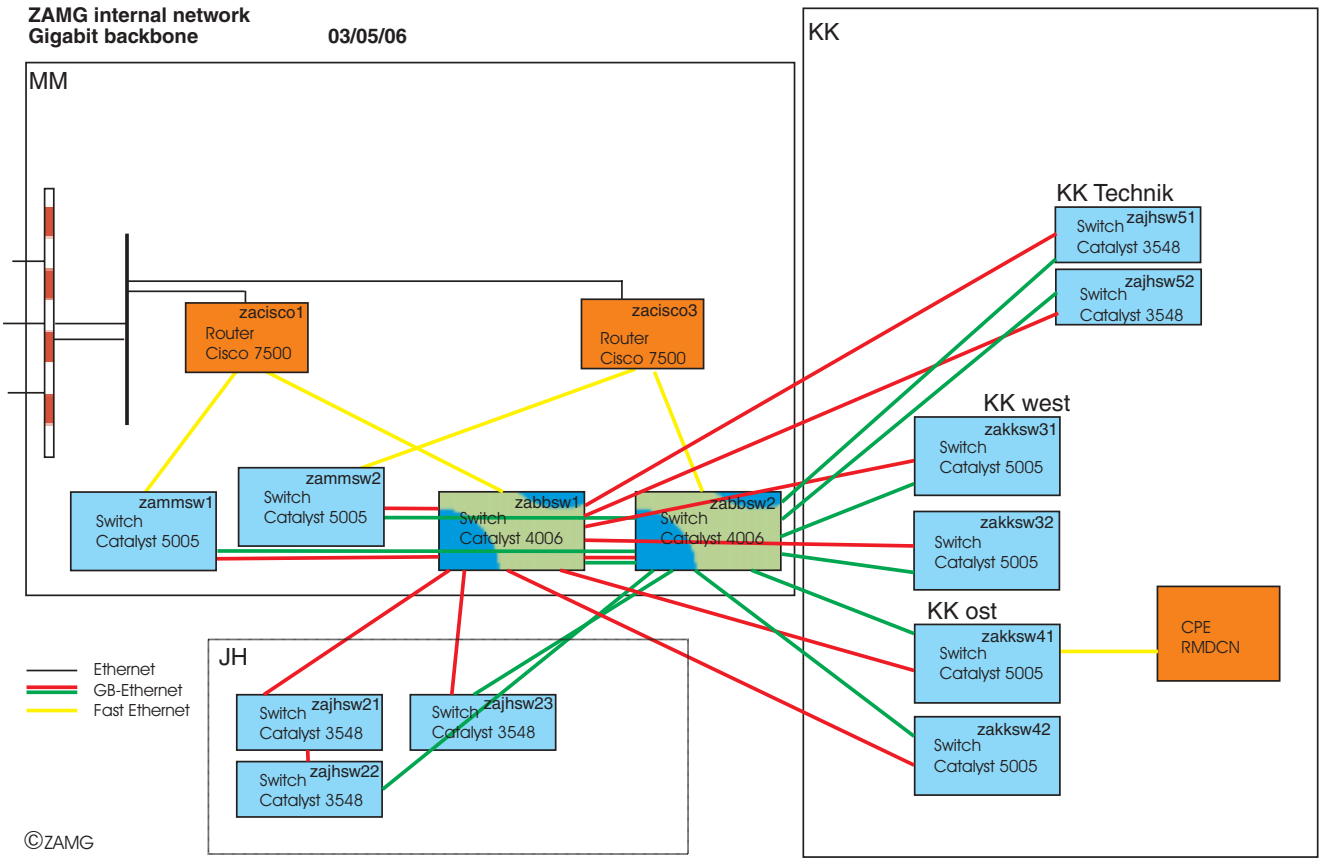
3) Generation of new products (esp. forecasts for individual locations)

4) Derivation of new version of statistical forecast AUSTROMOS

5) The newest Magics-Version will be installed and tested and metview will be installed.

AUSTRIA

AUSTRIA



BELGIUM BELGIUM

Roger Vanlierde, Royal Meteorological Institute, Brussels

Computer infrastructure at the RMI

Parts of the network and the computing resources that are used in the three institutes sharing the same site (the Belgian Institute for Space Aeronomy, the Royal Observatory of Belgium and the RMI) are continuously upgraded.

1. Shared infrastructure

Network

The three institutes' networks are connected by a Gigabit Ethernet backbone. A Cisco Catalyst 4006 routing switch is located at the core of the backbone. This switch is connected with a gigabit Ethernet connection over fibre to one HP Procurve 8000 switch at each institute. Each institute has its own switched Ethernet network connected to these Procurves.

The computer server and file server cluster are directly connected to the backbone routing switch.

The previous shared servers and each institute's most important servers are connected to a dual FDDI ring. This will be disabled, as the last shared server connected to it will be withdrawn from use in January 2004.

Computing resources

The servers infrastructure consist of

- 2 HP 9000/L2000 (2 processors, 3 GB memory, 2x18 GB internal disks): file server cluster. Both nodes are active (load balancing).
- 1 HP FC-60 Fibre Channel disk arrays with 2 TB of storage, to be complemented by an additional FC-60 with 2 TB in two years.
- SGI Origin 3400 (24 processors, 24 GB memory, 584 GB internal disks): compute server.
- 2 HP 9000/A500 (1 processor, 1 GB memory): mail, DNS, Internet server ...

In the near future

- 1 HP 9000/A400: split horizon DNS

2. Computer infrastructure specific to RMI

In addition to the shared resources listed above, the RMI itself has the following servers in use:

- HP 9000/L2000 (4 processors, 2 GB memory, 72 GB internal disks): login and application server.
- HP 9000/L1000 (2 processors, 1 GB memory, 34 GB internal disks): oracle database server. The database data are stored on a SAN (Storage Area Network) consisting of a HP VA 7400 Virtual Array with 15 disks each of 73 GB. A HP Superstore Autoloader insures the backup of the data on tapes.

These 2 servers form a high availability cluster.

- HP 9000/D390 (2 processors, 128 MB memory, 9 GB internal disks): Met Office server.

This server and the following two form a high availability cluster where each node is capable of taking over the functions of the other nodes.

- HP 9000/D370 (2 processors, 128 MB memory, 2x2 GB internal disks): Telecommunications server.
- HP 9000/K100 (1 processor, 256 MB memory, 2x3 GB internal disks): login server management and pilot database server.
- HP 12H AutoRAID disk array with 200 GB disk space: storage for the cluster detailed above.
- HP 9000/R380 (1 processor, 512 MB memory, 2x9 GB internal disks): nearline storage server and intranet server.
- DLT jukebox with 2 drives and 30 tape slot: nearline storage.
- HP 9000/A500 (1 processor, 512 MB memory, 2x18GB internal disks): web server.

The servers in the high availability array (with the Met Office server) are marked for gradual replacement over the following years.

BELGIUM BELGIUM

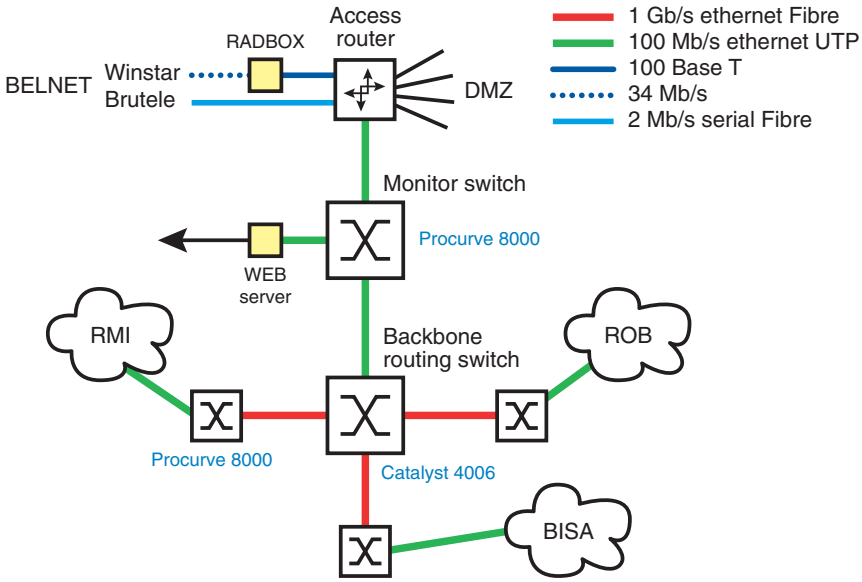
External communication links at the RMI

The institute has the following external communication links:

- Internet connection: over a directional wireless link with Winstar (34 MB/s maximum), and from there to Belnet, the Belgian leg of the European research network. A backup to Belnet at 2 Mb/s via a television cable (Brutele) was added in 2001.
- RMDCN: leased line to SITA POP Brussels (256 Mb/s) with dual ISDN backup.

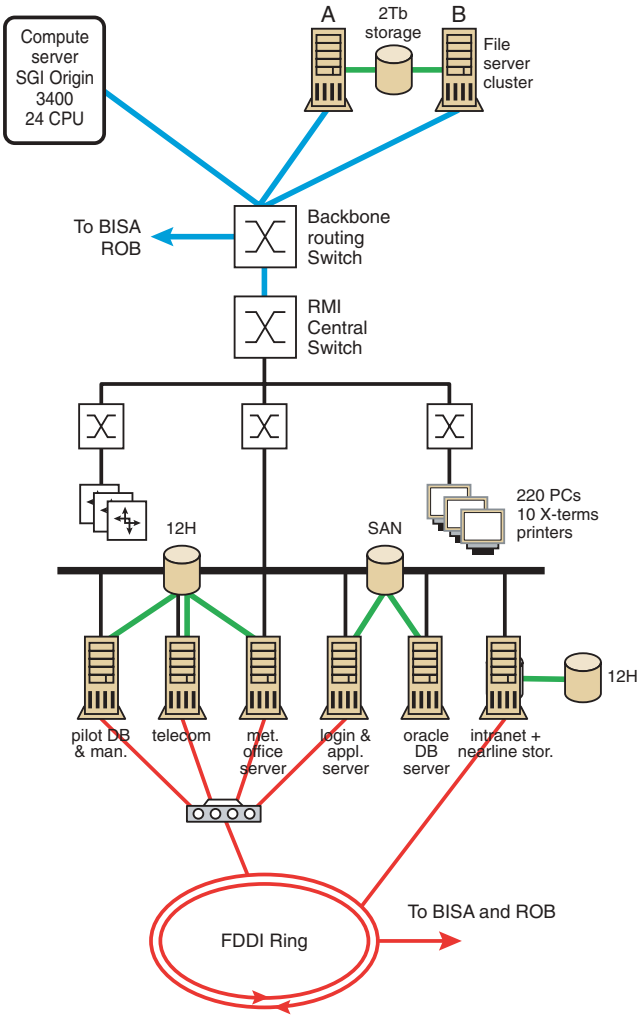
The following CIRs are used:

- ECMWF: 96 kb/s
- Brussels to Toulouse: 32 kb/s
- Toulouse to Brussels: 96 kb/s
- Bracknell: 16 kb/s
- Belgocontrol (civil aviation service): 64 kb/s leased line through Cisco router
- MeteoWing (military aviation): 64 kb/s leased line through Cisco router
- branch office in Zeebrugge (AWZ): 128 kb/s ISDN through Cisco Router
- branch office in Dourbes: 96 kb/s leased line through Cisco Router, with binary radar data multiplexed on the line
- leased line to VRT (national television): 256 kb/s
- several modems for data collection and distribution
- a modem pool for users (dialback)
- to the lightning detection system SAFIR
- to Météo-France RETIM system (via satellite)



BELGIUM

BELGIUM



CROATIA CROATIA

Vladimir Malovic, Meteorological and Hydrological Service of the Republic of Croatia

Dr_avni hidrometeorolo_ki zavod (DHMZ) is the governmental meteorological and hydrological service of the Republic of Croatia. It is situated in Zagreb and has two branch-offices at the coast for a maritime services (weather reports and forecasts for Adriatic Sea), one in the town of Split and the other in Rijeka.

Computer network

The main computer resources are located in DHMZ headquarter in Zagreb. The branch-offices on the Adriatic coast, Zagreb airport, GTS network, automatic station network, Ministry of defence, national TV and the Internet are connected by WAN via ISDN and modems.

We are connected through RMDCN to Reading and Vienna.

LAN consists of several machines under the UNIX operational system, 3 VAX VMS machines, terminals and PCs under Windows OS and few Linux PC computers. Some of the computing tasks are performed on a old VAX machines.

We expected that the migration of all software from old computers will be done at the end of this year. Applications are currently performed on both platforms.

Computer resources

The two main servers are SGI Origin 200 computers called Cirus and Lahor. They are responsible for the processing of GTS and climatological data. The Lahor computer is our gateway to RMDCN connections to ECMWF.

SGI Indy is the graphical workstation used for post-processing of ECMWF dissemination data and their graphical representation.

2001. We have bought a 16-processor SGI Origin 3400 that now runs ALADIN-Croatia twice a day.

EUMETSAT images are downloaded on a Linux PC (rapid scan images) and transferred to the Intranet. Data from DWD (FAXE) are received on a dedicated HP workstation.

Observation station and automatic observation station data are collected on two PCs under WIN NT and processed on SGI servers.

The Marine centre in Split and the Marine office in Rijeka are equipped with PC servers. They are connected to Zagreb via ISDN.

Linux PC computers with Trivis software are used for visualisation of meteorological data for Intranet, TV and newspapers.

Radar centres, situated in the northern part of Croatia are equipped with PC servers, mainly under WIN NT and they are connected to Zagreb through ISDN. The product from Puntijarka Radar Center is sent to Graz, and then processed as part of the Central European Radar Composite, available trough GTS products exchange.

Operational service

The majority of our operational service is done by our own application related to GTS, climatological and hydrological data processing. Products are prepared for hardcopy and Intranet. Some products are also available on the Internet.

We just start with developing software based on Mysql database and PHP services.

Weather forecasts rely on products of several numerical models (ECMWF models, ALADIN-Croatia, ARPEGE and DWD models).

Use of ECMWF products

- Operational forecast model products (12Z based products)
- Ensemble products
- Mediterranean wave model products

Received data are in GRIB and BUFR form. They are used mainly for weather forecasts.

- Occasional retrieval of MARS data.

CROATIA

CROATIA

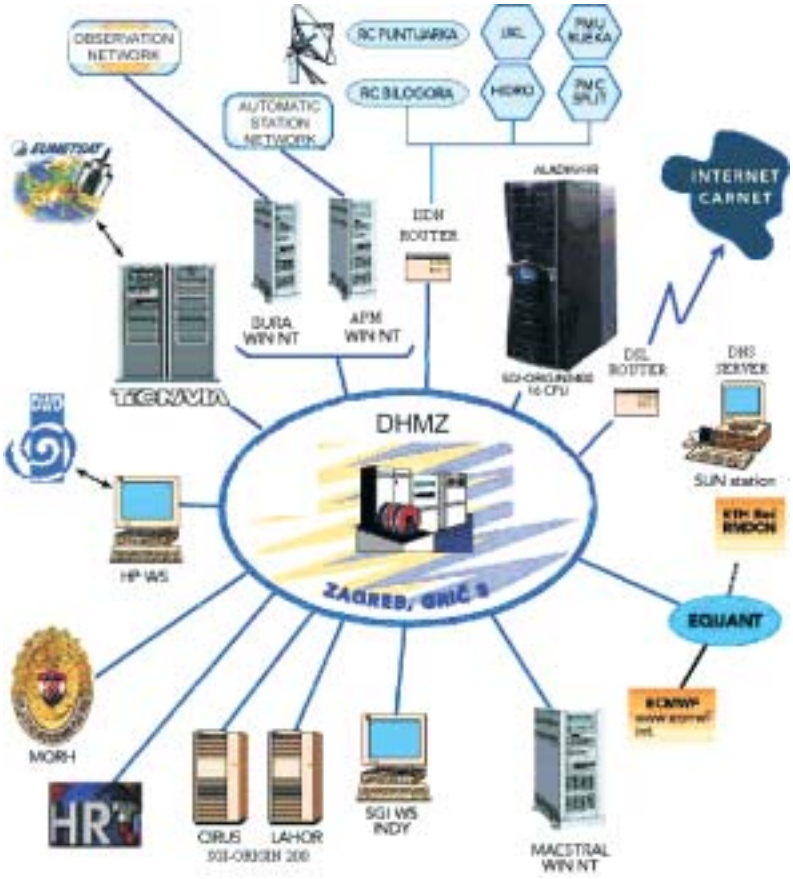


Fig. 1: Computer configuration at DHMZ

DENMARK DENMARK

Niels Olsen, Danish Meteorological Institute, Copenhagen

The computer system at DMI

The software that manages the GTS communication is installed on to sun Sparc stations and has functioned for several years without any problems.

The received data is transferred to the SGI Origin 200 computers for decoding and to two sun Sparc stations for AFTN use.

The decoding programs for SYNOP and Sealevel have been modified in order to make some extra quality control of the observations from Denmark, Greenland and Faeroe Islands. The new quality control will be taking the different climatological limits in account when checking the observations. The implemented controls consist of max/min check, step check and check for internal conflicts in the observation. All decoded observation is stored in the GTS-database in BUFR-format.

Data from 4 Danish weather radars are handled by micro-Vax 3100 computers and transferred from these to DMI via ISDN using ftp.

Data from ECMWF are received on a sun ultra Sparc and stored in the GRIB database for use in the operational suite.

Meteosat and NOAA data are received on PCs and transferred to two sun ultra Sparc stations for processing and generation of products.

The main computing system for running different weather prediction models consists of a 8 processor NEC SX-6 with 32 Gb MMU, 8Gb XMU, 1 Tb disc storage. The HIRLAM weather prediction model is run for an area around Greenland and an area around Denmark. The run consists of nested models, where the first model run uses ECMWF as boundary fields, while the nested models uses the HIRLAM as boundary conditions. The data generated is stored in the mass storage system, that holds up to 50 Tb.

The NEC SX-6 computer was installed at DMI last year with 2 nodes. In March the system was upgraded to 8 nodes.

ECMWF Products

Denmark receives approx. 1.25 Gbytes per day from the operational forecast model, the wave models and the Ensemble Predictions.

12Z based products:	450 Mbytes
00Z based products:	10 Mbytes
Ensemble Products:	3 Mbytes
Wave model products:	20 Mbytes
Boundary products:	800 Mbytes

The data received from ECMWF is transferred to the computers holding the GRIB database. The data is the accessible by the graphics package, which is based on Metview that is originally obtained from ECMWF.

Projects run at ECMWF

The remote jobs run at ECMWF are to a large extent for data retrieval from the MARS archives.

The research department is running some experiments on the HIRLAM 3D-Var code. Some jobs have been run in order to compare the new 3D-Var code with the reference HIRLAM system.

They have also been used in calculating trajectories for the stratosphere and there have also been run jobs connected to a research project on reduction of the ozone layer.

DMI is currently running two special projects.

The two projects covers the areas ‘Heavy rain in Europe’ that is connected to ‘A European Flood Forecasting system’ and ‘detection of Changing Radiative Forcing over Recent Decades’.

DENMARK DENMARK

UNIX servers

2 SGI Origin-200

- 1 server with 4 R10000 processors, 2 Gb main memory, 300 Gb disk storage
- 1 server with 4 R10000 processors, 2 Gb main memory, 150 Gb disk storage

These systems handle the data pre-processing, message switching, the generation of the GRIB database, the post-processing and most of the production of graphical products.

SGI Origin-2000

- 4 R10000 processors, 1 Gb main memory, 64 Gb disk storage.
- File server for the research department.

SUN Ultra 10 Servers

- 1 processor, 128 Mb main memory, 9 Gb disk storage.
- Serves as file servers for the different departments.
- 2 Servers handling satellite date
- 2 servers handling GTS-communication
- 6 servers handling data to be send to external users

SUN Ultra enterprise Server

- 2 processors, 256 Mb main memory, 68 Gb disk storage of which 60 Gb is raided.
- Climatological database based on Ingres DBMS

FTP Servers

- 2 SUN Ultra 10 with 1 processor, 128 Mb main memory, 9 Gb disk storage.
- Different versions of Intel based PCs running Linux.

Firewalls

- Different versions of Intel based PCs running Linux.

Other computers

Data from the 3 of the 4 Danish weather radars are handled by micro-Vax 3100 computers and transferred from these to DMI via ftp, and the last weather radar is handled by a Linux-box.

Later this year the micro-Vax computers will be replaced with computers running Linux.

UNIX Workstations

There are about 50 UNIX Workstations, of which most are PCs that has installed Solaris. Most of these PCs are equipped with two screens.

Linux Workstations

The last 6 months we have started replacing the SGI and Sun workstations with Intel based computers running Linux.

Network

On the Local Area Network we link the network by use both routers and switches.

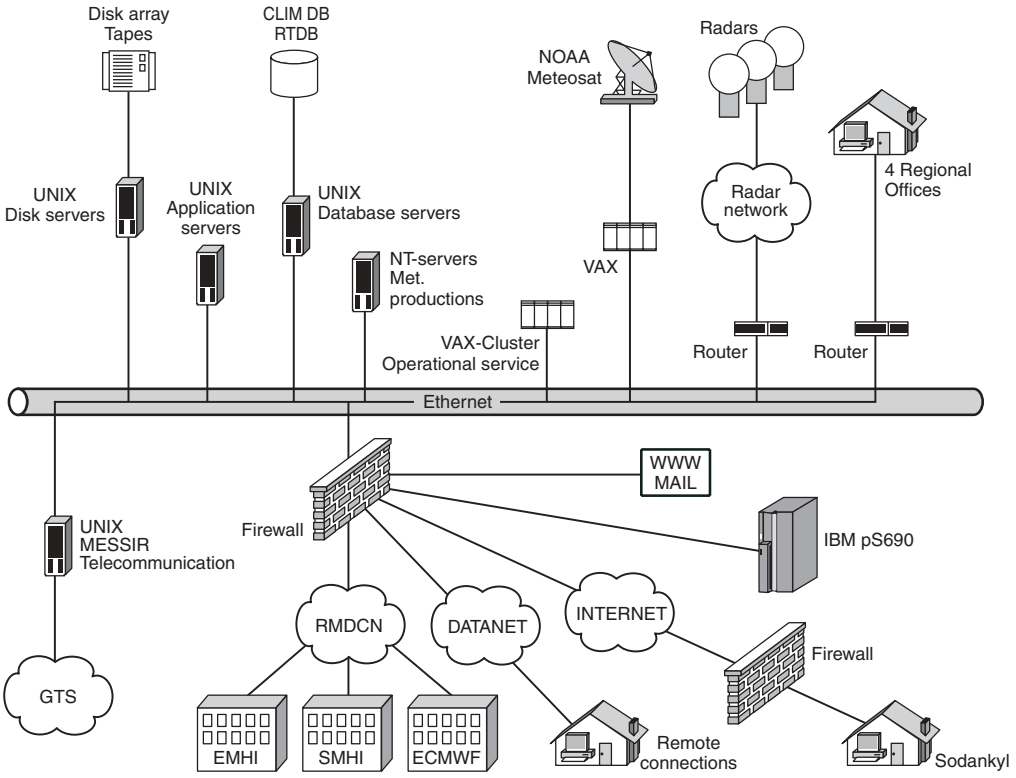
The networks at the regional centres are linked by bridges via leased lines, using ISDN as backup.

FINLAND FINLAND

Kari Niemelä, Finnish Meteorological Institute, Helsinki

Use of ECMWF-resources in Finland

1. Computing environment at FMI



– There are no major changes in our computing configuration since last year. Some servers have been updated into more powerful ones.

– The new database servers are in full use running both climate and real time databases. The disk server provides all systems with centralised disk space.

– The Finnish supercomputer has been in full use at CSC (Centre for Scientific Computing) 10 km away from our institute. The line speed to CSC is 100 Mbps. Our agreement with CSC is such that one of the 16 IBM pSeries 690 nodes is dedicated solely to our own use. The Hirlam model is run 4 times a day and the rest of the computer time is available to other use. The fail-over system is still under consideration. One possibility that has been thought of is to run the model here at ECMWF. The software is fully interchangeable between these two machines.

– Our RMDCN-line has a speed of 192 kbps. On times when there is dissemination traffic from ECMWF it is fully loaded but at other times it seems to be empty. This is due to use of Internet in file transfers.

2. Experiences from the use of ECMWF systems

– The ECaccess-facility has proved to be very handy. I have guided our users to use it instead of RMDCN and since then the RMDCN-line has not suffered from overloading. Now the developers here say you can also use ECaccess through RMDCN, but this is then question of the line speed. To us ECaccess means Internet. ECaccess has gained positive judgements from everyone I have talked to.

– There are a couple of projects running in our institute involving copying of considerable amounts of data from ECMWF. Acquiring this data through RMDCN would be impossible.

– You don't have to log into an ECMWF-machine to run requests on MARS.

- What I would wish from ECaccess-developers is to have a bit lower level instructions.
- ECaccess has also made it easier to make use of HPCF. Our Hirlam-developers run test runs at ECMWF. Now that the power of the HPCF has increased again it will be harder than before to use the same percentage of the allocation. It seems that very few countries have reached the 50 % line of allocation usage even before IBM era.
- The ECMWF web pages are used very actively by many people at out institute and even routinely by the forecasters. The pages contain an enormous amount of information. It is a good idea to have the training material in the web. It diminishes the need to attend courses personally and surely diminishes the number of contacts to user support.
- There's one problem I have encountered using the list of meteorological parameters for MARS and dissemination. It is sometimes difficult to tell which the corresponding parameter in the WMO-list would be. This problem counts especially for radiation parameters. I would be grateful if ECMWF could put a link pointing to an explanation page or just another column where there is the corresponding number of WMO parameter list where applicable.
- To get some feedback from the users I launched an email inquiry. Obviously the users are very satisfied with the service, because I only received one answer, and it was all positive. The only one of the users who answered told me to give special thanks for the improvement of the user information and for it that whether the problem is big or small you always get a nice and friendly answer from the user support. And this in spite of that the problem is mostly caused by yourself.

3. Future plans

We have no specific plans on institute level concerning the use of ECMWF facilities. As to the HPCF we have a fair share of IBM resources in Finland, but having fully compatible equipment here at ECMWF enables the users to easily switch from one centre to another. Here you have considerably more power to run, and that fact attracts the Hirlam testers. There also seems to be some interest to use trajectory facilities too.

One of our projects is planning to copy a huge amount of ERA-40 data to Finland to make some calculations on it. I have tried to persuade them to make the calculations here and then transfer the results, but I have not been very successful. It seems they have a lot of tested software, which they are not willing to port to another platform.

4. New building

The rental contract of the new building for our institute has been signed 2 months ago. This means that we are starting to plan the removal into a new address. The new building will be situated 5 km towards the North among the faculties of natural sciences of the University of Helsinki. All the units in Helsinki will be united under the same roof. The new building should be ready in two years time. This removal and its planning will require substantial extra effort from especially the systems and networking personnel of our data services division.

FRANCE FRANCE

Marion Pithon, Météo France

1. Computing environment and Network Configuration at Météo France - Future plans

1.1 Central computing system

The main enhancements regarding the computing environment since the previous meeting are related to the supercomputer, the backup system and the backbone of the local network :

- Vector computing: VPP5000 (124 Pes : one 64 Pes computer and one 60 Pes computer)
- File server and archiving system: O2000 (8 procs) , 2 STK 9310 silos and a STK 9710 silo.
- Production system (pre and post processing, databases ...): HP servers
- Scalar systems: HP servers
- Visualization system: SUN servers
- Climatological data base: a SUN server
- Backup system: a SUN serverV880 and 2 STKL700E silos with LTO drivers
- Telecommunication system: SUN servers
- Internet Web server: 2 SUN servers running Solaris and 4 PC running LINUX
- Office servers: 11 Windows NT4 servers for office applications and development purpose.
- Desktops: personal computers running Windows NT or Linux and Sun workstations.

Supercomputer

The upgrade of the VPP5000 supercomputer was decided in the beginning of 2002 : the choice of an "open" Invitation To Tender seemed to be risky, at that time, in terms of performance increases, migration work load and cost. Météo France decided to keep a vector system for 3 years (until 2006) and the availability of ECMWF VPP5000 was a good opportunity to seize on.

The official decision and financial agreement were signed in the early 2003 and the installation is in progress now.

The final configuration will be 2 machines:

– One VPP for production (60 Pes- 280 Gbytes memory- 2.7 Tbytes disks)

– One VPP for development, climate research and backup for production (64 Pes- 300 Gbytes memory- 3.4 Tbytes disks)

The 64 Pes machine is expected to be set up by the end of May, the 60 Pes machine at the end of July. The performance increase is 4 times the actual system and there will be no migration effort for research teams.

File server and backup system

The hierarchical storage system is based on DMF (Data Migration Facility) from SGI with the following hardware:

– 1 server SGI O2000 (8 procs- 4 Gbytes memory- 1.8 Tbytes of RAID3 disk cache)

– 2 STK 9310 silos (Powderhorn) with 42 drives

– 1 STK 9710 silo (Timberwolf) in a remote location for double copies.

The data handling system manages currently 125 TB of data for a maximum capacity of 150 Tbytes and a monthly growth of 5 Tbytes!

As the actual system is overloaded and as the compute server's upgrade will not improve the situation, the DHS replacement is essential.

A procurement process took place at the end of 2002 and the choice should be made in July.

The backup service, which is a different system from the file server, was upgraded:

– To face the increasing number of servers to backup (170 Unix or Linux servers, 20 Windows NT servers, 5 Oracle data bases and the supercomputer system)

– To provide a high level of data security (backup data and original data located in different buildings)

– To provide the same system for central service and regional centres.

FRANCE FRANCE

The new system is in production since February 2003. The system can cope with a total storage capacity of 45 TB. It consists of a 4 processors SUN server, 2 Storage Tek L700E silos with IBM LTO drives and 2 Fibre Channel Switches (Storage Tek 4116) to connect the silos to the server. The software used is "Time Navigator" from Atempo.

HP servers

A set of HP servers is in charge of the operational tasks:

- 2 N4000 handle pre-processing and post processing of meteorological data and the databases (Oracle) for observations and forecast products.
- 2 L1000 servers host products database.
- 2 D370 are used to start the models on the supercomputer, to monitor and supervise all operational tasks.
- 1 L2000 is for tests and integration.
- 5 HP servers (1 L2000 with 4 procs, 2 K570 with 3 procs and 2 D and J class servers with 2 procs) are available to users for development purpose, use of graphical tools, code management, front end server for the VPP (job submission to the VPP, use of graphical libraries for visualization...)

SUN servers

- 2 E3500, 2 E450, 9 Ultra 60 (2 procs), 9 Ultra 10 are used for the interactive visualization system for forecasters (production and backup, development, visualization).
- 1 SUN Fire V880 (4 procs, 8 GB of RAM, 600 GB of Fibre Channel disks) houses the backup system.
- 1 SUN Enterprise 5000 (8 procs) and 1 SUN E3500 are used for a climatological data base (based on ORACLE).
- 4 SUN E3000 (2 processors) handle the telecommunication system (locally developed software TRANSMET) that are in charge of real time reception and broadcast of meteorological data.
- 2 SUN E450 for the production part of the WEB server.
- 2 SUN E450 for the lightning database.
- 2 SUN E450 for the Radar data production server.

LINUX servers

A total of thirty-two LINUX (Red Hat 7.3) servers (18 in Toulouse and 14 in regional centres) are used at Météo France for different tasks.

- 4 HP LP1000r for the WEB server (client part)
- 4 NEC express 1200f for http servers
- 8 DELL P2600 servers for development servers and some climatology production servers
- 4 Fujitsu-Siemens servers for climatology production (java-Corba applications).

Currently, at Météo France, production and development Unix servers are divided into HPUX, Solaris or LINUX operating systems. These heterogeneous platforms are expensive in terms of hardware upgrade, software availability, system administration, user support and supervision. A study is in process to try to standardize the Unix operating systems. LINUX should be a good candidate because the cost/performance ratio is good.

Desktop systems

Currently, in Toulouse, there are many (>800) PCs running Windows NT4 or LINUX, (mainly Red-Hat or Mandrake distribution), some (~50) X-Terminals and some (~50) workstations.

- 11 Windows NT4 servers are used for office applications and development purpose.

FRANCE FRANCE

1.2 Network

Local area network

The backbone of our LAN was replaced in March this year to provide Gigabit Ethernet. It is based on CISCO equipments (Catalyst 4006 with 6 slots, one for the central department and one for the research department). This provides a 32 Gbits/s Full duplex bandwidth.

HIPPI links enable fast transfers between the 2 VPP5000 and the O2000 (file server). These links are backed up by Gigabit Ethernet connections between the three computers.

Wide area network

The connections between central services, the 7 regional centres and almost all other Météo centres (about 90) are covered by a Frame Relay network. The Committed information rate (CIR) between Toulouse and the regional centres is now 1Mb/s, and 64 kb/s between Toulouse and the local centres. Some leased lines remain for connection with outside services or X25 applications.

Connections to Internet (access to RENATER, the French academic and research network) are made through "GALACTIS", a high bandwidth ATM network in Toulouse, with a 10 Mb/s access since March. These connections are protected by a firewall implementation. The software used is Firewall1 from Checkpoint on a dedicated equipment (NOKIA).

1.3 Connection to ECMWF

Our connection to the GTS and to ECMWF is made through the RMDCN network. The delivered bandwidth is asymmetric with a CIR of 512 kbps in and 256 kbps out and is dedicated to dissemination.

Users are highly encouraged to use Internet (10 Mbits/s) to connect to ECMWF and for file transfers to avoid concurrent traffic with dissemination. The convenient transfer tools (part of ECaccess) have enabled to promote the use of Internet among users. A packet shaper, at Météo France, has been set up to prioritise traffic on the port numbers used by Ecaccess over general Internet upcoming traffic.

2. Use of ECMWF resources

Operational products

From mid May, the volume of data received at Météo France through the dissemination of ECMWF products will double because of the 00 UTC IFS results availability. So, the total volume of data concerned by the dissemination at Météo France will be more than 2 GB compressed per day (30% to 50% savings). The products transferred are: IFS 12 UTC and now 00 UTC results, some raw products (divergence, vorticity, some surface fields...), Wave models and EPS results. All these products are used for operational meteorological activity.

The projects

42 Météo France projects and 6 special projects are registered for this year with a total of 185 users; the main activity is data retrieval from MARS on ecgate1 (almost 75% of Météo France users).

The more active projects are:

- Climate simulations.
- Mesoscale model experiments.
- Participation in DEMETER project (ended in March 2003)
- Atmospheric predictability studies.
- Statistical works on ECMWF data.
- Control and monitoring of data.
- Studies on EPS.
- Seasonal forecast.

The following are registered as special projects in 2003:

- MERCATOR project: build a global ocean circulation simulation tool based on a high-resolution model assimilating real time data (SPFRMERC)
- Universal software for data assimilation: variational method for global ocean (SPFRPALM)

FRANCE FRANCE

- Chemistry cloud and radiation interactions in a meteorological model (SPFRPECH)
- Seasonal to interannual predictability of a coupled ocean-atmosphere model (SPFROASP)
- Variational data assimilation (SPFRVODA)
- Forecasting optical turbulence of Astronomy applications with the MesoNH mesoscale model coupled with ECMWF products. (SPFRSEE)

The majority of users are MARS users and work on ecgate1. They are mainly concerned with MARS data access, local storage on ECFS (or on disks on ecgate1), efficient file transfers from ECMWF to Météo France (their workstations or Météo France file server) and Web services.

Some projects, however, need to work on the supercomputer. In 2000, the CPU usage represented 5% of the total allocation, in 2001 it increased to 8%, in 2002 to 35% and in May 2003 it's already 15% of the allocation, so we notice a steady increase in the use of supercomputer resources. These projects are concerned by the migration to the IBM system.

Migration to the IBM system

A survey has been made among the 5 or 6 projects, which started to use "hpca" about their migration experience.

Migration support

- About 10 French users have attended the IBM course at ECMWF and they found it was a good introduction to IBM system.
- The migration support from ECMWF provided by user support is experienced as very efficient (as usual!) and very useful (good advices).

Performance issues

- gprof and mpi_trace are useful tools to get information about a code performance.
- Up to now, the users benefit of the very good throughput (almost no waiting in batch queues, for small jobs - 4 nodes or less-). This good throughput is likely due to the increase of total performance (compared to VPP systems) and to the moderate use of this new system at the moment (all projects have perhaps not achieved yet their migration?).
- The level of performance compared to VPP system observed on pure MPI codes are the following:
  - Climate models (ARPEGE with a low resolution) use 4 nodes with 7 processors each and get the equivalent performance observed on 4 VPP Pes.
  - MESONH code : more than one node (from 10 to 14 processors) are needed to get the equivalent performance of one VPP processor. MPI communications take time. Users need some advices to analyse MPI_TRACE outputs to see if they could improve things or if they are limited by the interconnect.
- Users, from both climate model and MESONH teams notice bad performance when they use full nodes (8 procs/node). Performances are much better when they run on more nodes and less procs by node. To a certain degree, a test case (of MESONH code) on 16 4-procs nodes takes less elapsed time than the same test on 16 8-procs nodes. Besides, using half full nodes is less expensive in terms of billing units than using full nodes. However, if you use more than one node and don't fill up the nodes you are noticed by the operation team and warned by mail.

Compiler and debugging issues

- Are there suitable tools for "openmp" debugging? Sometimes code failures have this only error message "segmentation fault" and it's very difficult to find where the problem is.
- Some users ask to be warned when there is a change in the compiler level to maintain consistency in their codes.
- Some MESONH routines don't compile with "-O3" standard optimisation option.
- The "-q save" option doesn't work well with recursive routines, so it cannot be used for MESONH code.
- Sometimes, the compilation of some routines never come to an end, so users have to limit the process cpu time limit at loadleveler level to be able to reach a time limit and an end to the job.
- The "-q extname" is set by default (/etc/xlf.cfg), which is not coherent with other IBM configuration in other French labs (CINES, IDRIS, CNES, SHOM...). That causes some trouble to MERCATOR project.

FRANCE FRANCE

– Users experienced cases where the compiler error message was nothing but "call service representative". It was difficult in these cases to see what the problems were!

Load leveler problems

– At job submission, loadleveler checks for the non ASCII characters in the jobs and that cause problems with comments inside programs. In that case, the only solution is to change the jobs by hand to force it to restart.

Specific requirements

– HOME file systems on hpca are really too small. Some projects really need some permanent disk space (visible from ecgate1 also) whose sizes are respectively:
  • 1 GB (at a minimum) for MERCATOR
  • 3.5 GB for MESONH
– MERCATOR team whishes to use CVS (code management) from hpca (a CVS client which could access the CVS database on their CVS server). They wonder, if it would be possible to have a sort a "CVS" tunnel through ECaccess or another way to achieve this?
- MERCATOR team wonder if it could be possible to modify the generic user ".profile" file for the use of rsh.

ECaccess

At Météo France, the gateway has been running, since last year, without major problems. It is running on a HP server (HPUX11 with a Java virtual machine 1.3). The current version of the gateway is 2.0.1 and it appears to be very stable. As soon as a problem was reported it is solved immediately and new requested features have been developed very quickly by ECMWF. Another gateway is also running at Cerfacs for their own users (about 20).

As the "ectrans" facility to transfer files from ECMWF to local servers is much more convenient than the former "eccopy" command, the migration was easy and is over, now.

The "ftp" and "telnet" connections through the gateway should replace very soon the use of "tn-gw" and "telnet-gw" gateways and should also replace the direct telnet or ftp connections through RMDCN.

Job submissions to "hpca" with "ecjput" are not yet available. ecqsub was used by the MESONH community to submit jobs to the vpp's.

Up to now, 40 users and about 10 different destinations are registered in Météo France gateway database. 10 of them use "ectrans" for operational tasks or at a regular basis.

The average bandwidth obtained for a transfer from ECMWF to a local destination is about 300 kB/sec.

GERMANY GERMANY

Elisabeth Krenzien – Deutscher Wetterdienst

1. Computer equipment at DWD and connections to ECMWF

The major changes in DWD computing centre environment since the last meeting refer to the replacement of the telecommunication system (AFSV) and the installation of a GBit-Ethernet network. The current configuration is illustrated in Figure 1.

	SP RS/6000	Routine Servers		O2000	Data Servers
Processor	POWER3 SMP	MIPS R 10 000 +			
CPU /processors	1280	8	14	24	24
Main Memory (GB)	792	6	8	24	24
Disk Space (GB)	3790	94	680	7853	7853
Tape Drives	-	-	-	26	-
Networks	HIPPI Fast Ethernet	Fast Ethernet HIPPI / ATM			
Operating System	AIX 4.3.3	IRIX 6.5			

Table 1: Current specifications of the central servers

The DWD compute server consists of 80 16 way NightHawk II compute nodes equipped with 375 MHz Power3 processors with a peak performance of 24 Gflops per node, with 8 GB of memory per compute and 16 GB memory per special purpose node, adding to an aggregated performance of 1.92 Tflops and 792 GB of main memory. The nodes are interconnected with an intermediate single plane SP Switch2 (Colony switch) which allows for an upgrade of up to 128 nodes.

The system has RAID-5 SSA disk management for the IBM parallel high performance file system GPFS. The communication network for MPI and the GPFS is the single plane Colony switch which serves all nodes. The system has been reconfigured into a single system although operation and developers use their own GPFS filesystems and -servers and are served by different loadleveler queues on distinct nodes. Three nodes are dedicated to login and interactive usage. Two control workstations in high availability mode serve as central management points and together with the login nodes are the only visible parts in the LAN.

The system software installed consists of the AIX 4.3.3 operating system, the SP system management software PSSP and the LoadLeveler as batch system. Currently the user environment includes the Fortran, C and C++ compilers, MPI, the ESSL library and third party software packages like TotalView, Vampire and the NAG library.

Currently the production SP system and one of the routine servers are the main platforms for the operational model chain, Lokal Modell (LM)/Global Modell (GME), that will stay in production until 2004 approximately. The second routine server hosts non-time-critical operational tasks and user activities that originate on X-terminals, SGI workstations and Windows NT4.0 systems within the DWD LAN. An upgrade of the system is inevitable to meet the new DWD initiative to respond to the great public pressure for better forecasts of severe weather events and resulting disasters.

An upgrade of the compute server is planned to provide the required performance of the original SGI contract signed in 1997. A purchase order to IBM will increase the system by additional 44 NightHawk II nodes in autumn this year. Performancewise an increase by a factor of 1.55 will result.

The data handling system at DWD comprises two O2000 servers in high availability SGI Failsafe mode, running about twenty Oracle database instances (Oracle 8.1.7), the Adic AMASS / DataMgr HSM software package and a three component StorageTek Powderhorn silo with 26 fibre channel tape drives (both 9840 and 9940). The total amount of data stored is 360 TB in about 6 million files and the monthly growth exceeds 10 TB; because of budget shortcuts users had to abandon the automatic backup-copy of their data, this restriction drastically reduced the volume of stored data. A replacement of the O2000 system by an IBM system is ongoing. Based on the user changing requirements new functionality is needed which is already tested on a small SAN based Enterprise Storage Server architecture attached to an IBM p690 server with 8 nodes. An upgrade of this system is in an early planning phase, but nevertheless the migration from the current data server to this new system has to be completed early next year.

GERMANY GERMANY

The DMRZ (Deutsches Meteorologisches Rechenzentrum), the distributed computing centre based on HPC resources of the national (Offenbach, OF) and military (Traben Trarbach, TT) weather service connected by a 34 Mbit/s direct data link is in full operation and now, guarantees the backup functionality for the operational GME forecasts. To attain binary compatibility within the DMRZ the required replacement of the HPC systems (Fujitsu VPP300, SGI O2000) is planned by the installation of a 27 node IBM p655 cluster (8x 1.1 GHz Power4, 32 GB memory per node) with a SP switch2 and a RAID-5 SSA disk management for a GPFS in late summer 2003. A new data server, an additional node, connected to a 14 TB IBM ESS storage system will also be installed, as well as two one node systems with 16 GB memory each to cover the post processing and user testing requirements.

Users from external governmental organisations successfully use the UNICORE software to access the DWD compute server.

Almost completely transparent for the partners on the GTS the replacement of the telecommunication system based on the Tandem, and therefore hardware failure resilient system (AFSV) by a high availibility two cluster system (Meteorological Telecommunication System Offenbach, MTSO) built by Siemens, took place in October 2002. An overview of the MTSO integration is given in Figure 3. The use of network address translation provided the time necessary for all partners to implement MTSO in their environments which is nearly completed by now.

DWD uses RMDCN as its main telecommunication link to Reading for exchange of operational data and products, while users are highly advised to use the Internet access to the Centre. DWD aims to finally set-up the backup management centre for RMDCN. The data communication to ECMWF is schematically shown in Figure 2.

2. Projects and experience using ECMWF resources

The ECMWF ecfs software package is in production on DWD platforms since late autumn 2002 and allows the users to access their archival data more directly. The retirement of the nfs protocol for the central servers resulted in a more reliable service for the users.

The co-operation contract signed in late 1999 by DWD and ECMWF to provide DMRZ users an interface to both computing sites in a seamless environment based on UNICORE has been renewed for another two years giving the partners the opportunity to operationally introduce the CSOMARS interface which is currently in an user acceptance test at DWD as part of the new ecaccess service. It offers the opportunity of a uniform access to both MARS and DWD databases via csobank.

First preparations for the delayed implementation of UNICORE software at ECMWF are on the way, while more planning of the security and authentication issue is necessary.

The testbed for the next generation of the GME model on the VPP5000 system, finally installed in July 2003, offered the GME model output for three different mesh sizes (60, 40, 30 km), 35 layers and a forecast range of 174 h which provided valuable information on the role of the assimilation data and will stay in use for reference purposes for the rest of the year.

The LM implementation is DWD's contribution to the COSMO LEPS Project that runs a limited area ensemble prediction system based on selected EPS members from ECMWF's EPS forecasts for severe weather prediction. This work is carried out by Italy for the COSMO Consortium (Switzerland, Poland, Italy, Greece and Germany).

In 2002, nine projects with 60 users and 13 Special Projects with 55 users have been registered, the majority of them retrieves data from MARS on ecgate1 via the Internet. About 78% of the allocated Fujitsu units have been used by mainly two DWD projects: the LLM studies that have come to a preliminary end, sensitivity and benchmark studies with the LM and GME models and the GME test bed.

The Special Projects of the Max-Planck-Institute of Meteorology run their climate and coupled ocean/atmospheric model studies on the VPP5000 with great success.

The majority of new users now comes from outside DWD and requests access to MARS data with growing interest in the operational forecast data to support short term international field campaigns. An easy procedure to handle the resulting amount of administration work would be helpful.

In principle, there are no difficulties in using the ECMWF computer equipment because of the portable code development strategy of DWD. The professional support provided by the ECMWF user support certainly is part of this success, the visit of Carsten Maass to DWD in November 2002 has been very useful to introduce the ECMWF web services.

GERMANY GERMANY

3. Plans

The Research Department of DWD still considers implementing the Observational Data Base (ODB) developed at ECMWF to meet NWP requirements and plans to test it in a systematic comparison of observational data sets at DWD and ECMWF but wants to know about the status of the project.

The delayed implementation of UNICORE (V4.1) at ECMWF is in progress and will be used for testing and research purposes. The CSOMARS software is expected to be optimised and expected to see more functionality resulting from active user requirements.

As a partner in the ELDAS Project (EU Project 'European Land Data Assimilation System') DWD intends to use ECMWF resources to develop an assimilation infrastructure for estimating soil moisture for better prediction of floods and droughts. The requirements for the experimental runs are available while the reliability of the transfers back to DWD is still an issue to be solved.

With respect to the homogeneous IBM platforms in both centres DWD expects to use the new resources as soon as they are fully available.

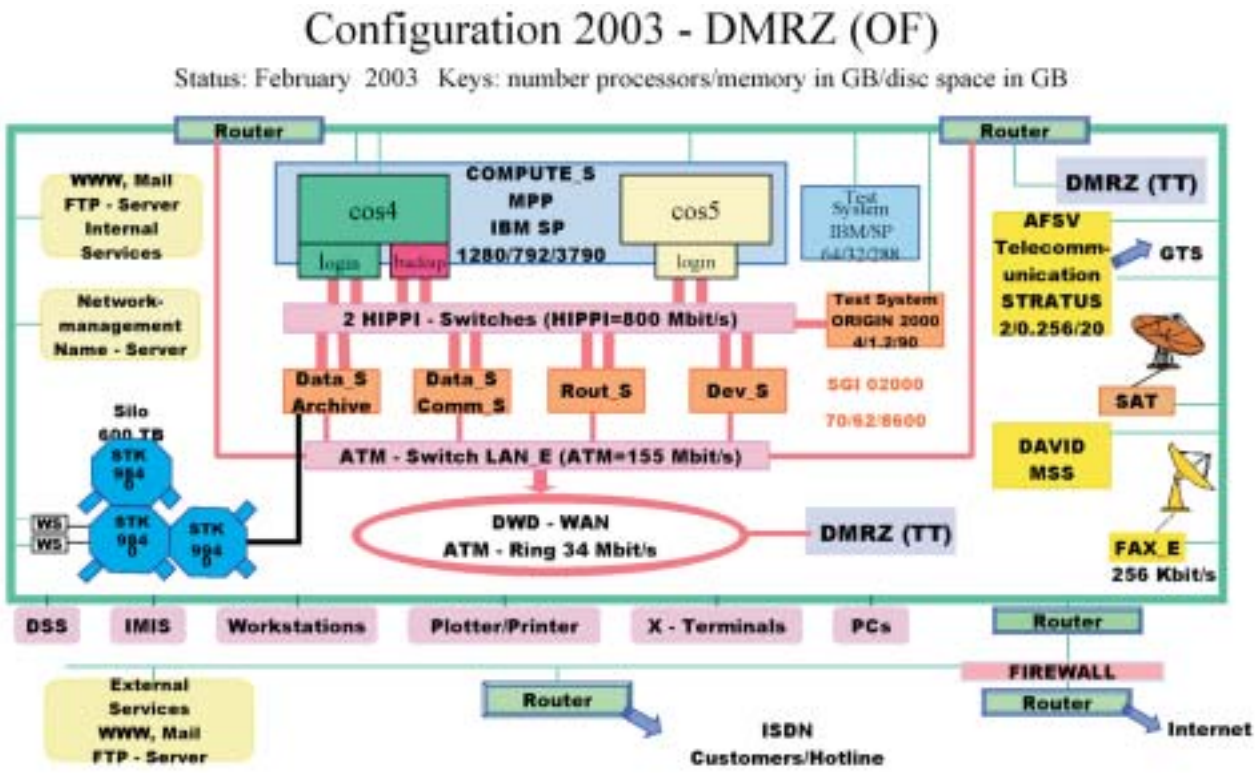


Fig. 1:

GERMANY GERMANY

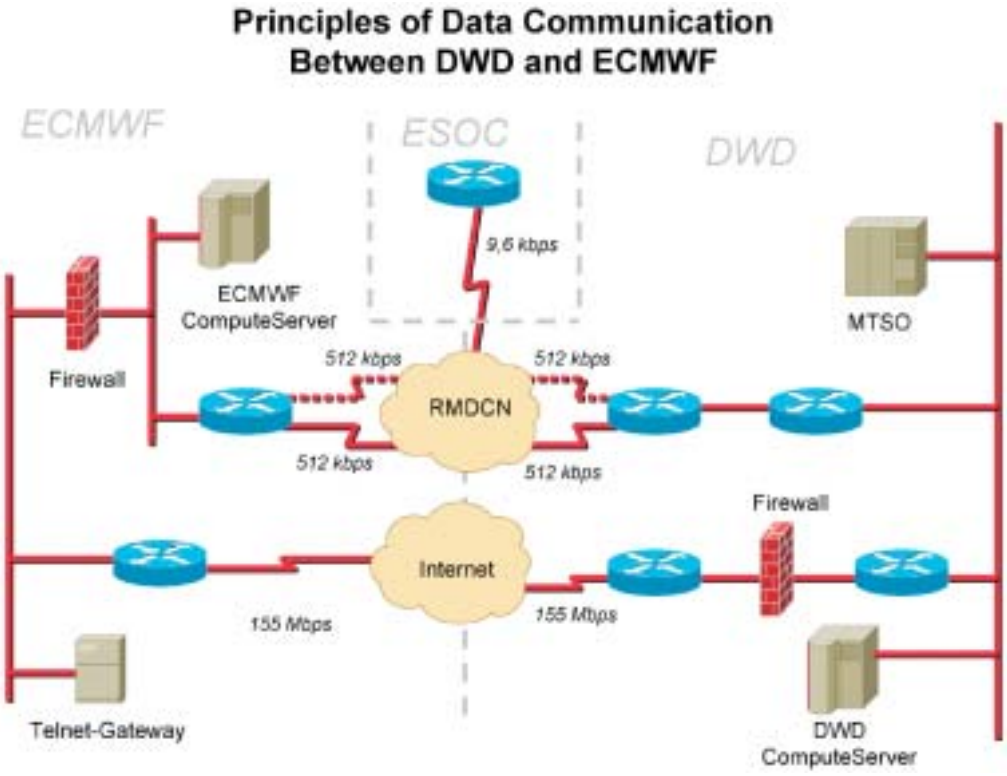


Fig. 2:

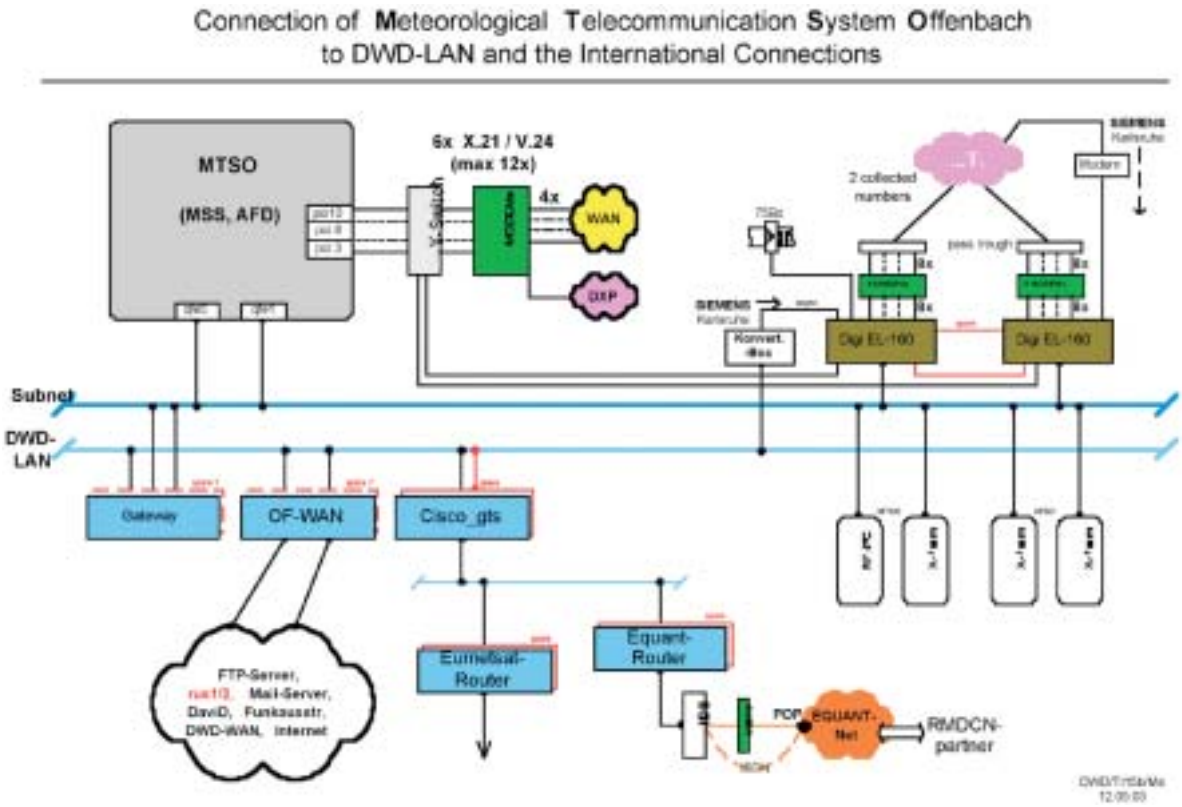


Fig. 3:

GREECE GREECE

Nikos Kamperakis – Hellenic National Met. Service (HNMS)

- Modernization plan of HNMS
- Use of ECMWF facilities

Modernization plan of HNMS

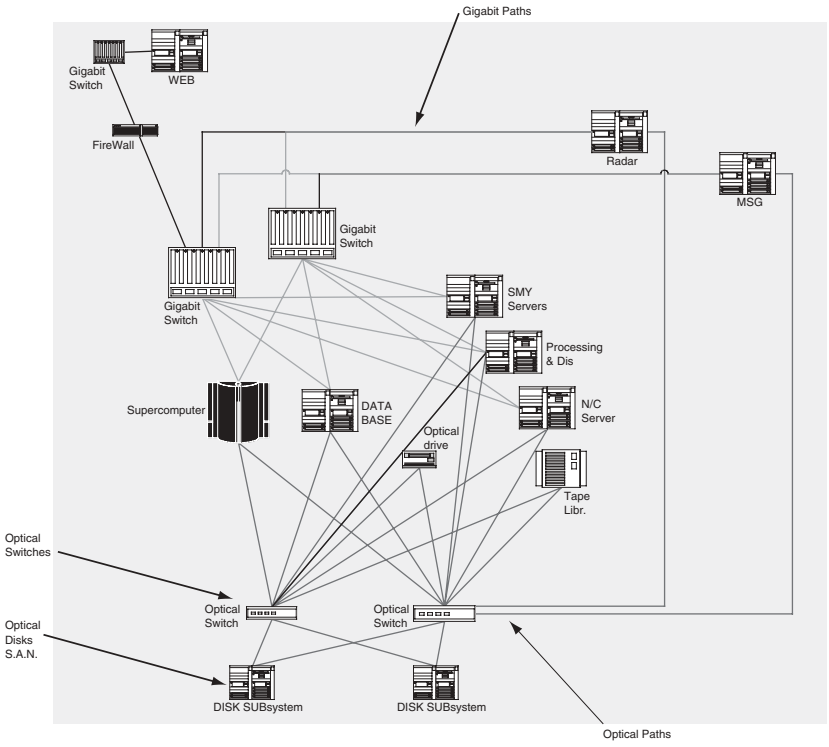
Computer and System Infrastructure

- Main Computer for the run of HNMS NWP Models
- Meteorological Visualization W/S System.
- Nowcasting System.
- Data Base H/W Upgrade.
- Radar Network.
- Lightning Detection Network.
- Wind Profiler Radar.
- MSG Satellite Station.
- Automatic - Manned Weather Station Network Upgrade.

Goals

- Meteorological support of the Athens 2004 Olympic Games.
- To upgrade the observation network and processing systems in order to provide more accurate weather products orientated to customer and internal end-user needs.
- Improve HNMS support to Hellenic early warning system for hazardous weather phenomena.

Planned Main Computer Facilities



GREECE GREECE

NWP Super Computer

64 bit OS. UNIX

Mem 80GB

Link Interconnection >= 1GB/sec

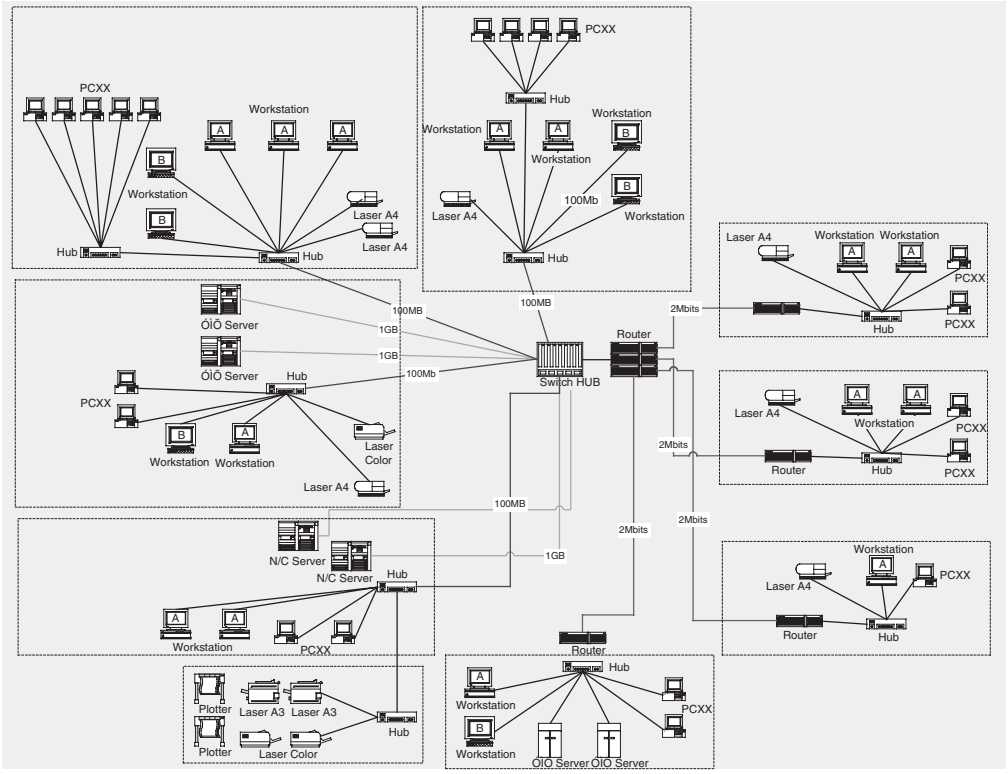
DISK SubSystem

- SAN Connectivity
- RAID Support
- 100MB/sec
- 2 Tbytes / 600 GB

MPI & OpenMP

RUNs LM (in an HOUR about 100Gflops)

Meteorological Support System



NowCasting System

Update Interval 15min

Forecast UP to 3 Hours Ahead

Use:

1. Weather Radars (10)
2. NWP products
3. OBS system
4. Satellite MTP/MSG/HRPT
5. Quality Control
6. Orographic Enh.

Forecast :

1. Temperature
2. Visibility
3. Dew Point
4. Synoptic Scale Precipitation
5. Convective Scale Precipitation
6. Type of Precipitation

GREECE GREECE

Data Base Technical Aspects

OS. UNIX

Mem 16GB

65.000 tpm TPC-C

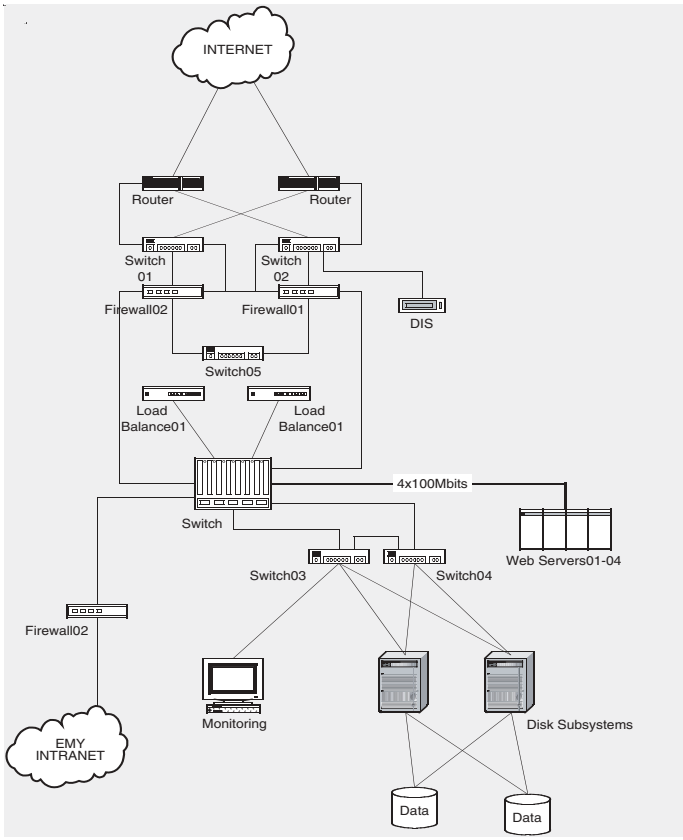
OPTICAL DISK SubSystem

- SAN
- RAID Support
- 100MB/sec
- 10 TBytes

Robotic Library

- 100TBytes

Web Server



Use of ECMWF Facilities

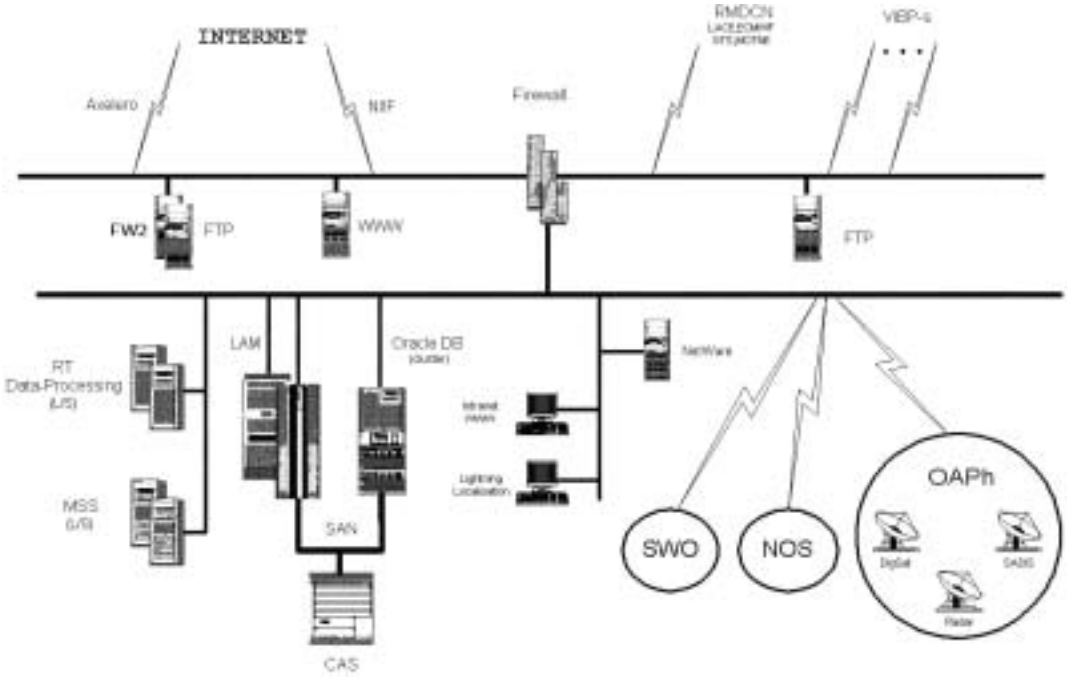
- Deterministic Model
- EPS
- Wave Model
- Boundary Condition
- Daily run of LM
- Trajectory Model
- MARS Retrieval

We are very SATISFIED

HUNGARY HUNGARY

Istvan Ihasz – Hungarian Meteorological Service. Report is presented by László Tölgyesi.

HMS logical network diagram

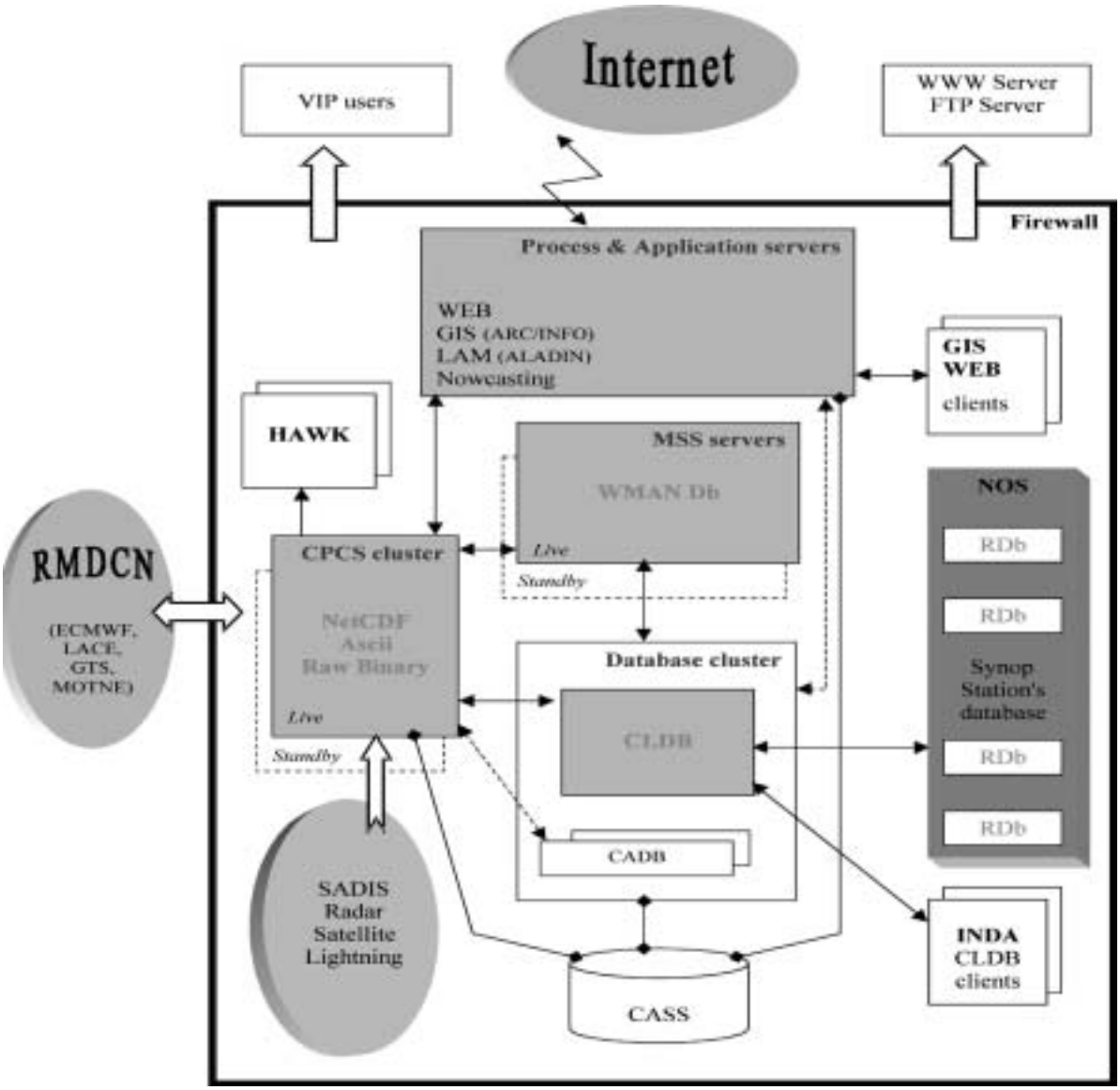


Computer resources

- Message Switching System (2 PC-s; Linux): life-standby WeatherMan
- Central Processing & Controlling System (HP L2000 cluster-PKG2; 4 CPU, 3 GB RAM, SAN-CSS, GigaLAN + HP D280, K250): scripts, programs, CDS-CASS
- Database server (HP L2000 cluster-PKG1; 4 CPU, 3 GB RAM, GigaLAN, SAN-CSS): ORACLE (8.1.6), CLDB, CADB
- Process & Application Server I. (IBM pSeries 690 Regatta; 32 CPU, 64 GB RAM, 850 GB disk, SAN, GigaLAN): LAM Aladin-HU and other research studies
- Process & Application Server II. (SGI ORIGIN 2000; 16 CPU, 8 GB RAM, 88 GB disk, SAN, GigaLAN): GIS, WEB, modelling (nowcasting, MM5, ...)
- Central Storage System: (CLARiiON FC4700; ~1.9 TB, with HP SureStore Ultrium 2/20 for system backup and HP DLT 1/8 for data backup)
- Other (firewall, mail, news, file, printer) servers: Unix, Linux, Netware
- DEC, SUN, HP and Linux WS's for visualisation and development
- about 300 PC (Windows NT and 2000)

HUNGARY HUNGARY

HMS information system



Network

- LAN: 10/100 Mb/s, structured UTP cabling system, CISCO switches
- GigaLAN: 1 Gb/s, UTP cabling system, among main servers, CISCO sw.
- SAN: 2Gb/s, redundant FC-switches, CLARiiON - IBM,HPcluster,SGI
- WAN: 64 kb/s and 128 kb/s leased lines, 2 Mb/s and 4*2 Mb/s microwave channels
- Internet: 512 kb/s and 128 kb/s
- RMDCN (256 kb/s with ISDN backup, upgraded from 192 kb/s):
  - ECMWF: 128 in / 16 out kb/s;
  - Austria (AC RTH): 32 in / 8 out kb/s
  - Slovakia: 16 in / 16 out kb/s
- Management: OpenView NNM, WEB-Netsaint, Netenforcer (QOS)
- Firewall: Internet segment and DMZ

HUNGARY HUNGARY

Changes related to ECMWF

- MARS, EMOS, ECLIB, MAGICS, METVIEW, EPS plume software latest versions are installed on HP-UX, SGI and Linux platforms
- Our RMDCN connection was upgraded from 56+8 (in/out) kb/s to 128+16 (in/out) due to cheaper prices.
- Ten registered users (since March 2003)
- Operational use of 0.5 x 0.5 degrees deterministic forecast
- Several operational applications (flextra, monthly and seasonal forecast) run on ecgate 1.
- Domain and resolution of ALADIN/HU LAM model was significantly increased in the begining of November 2002 (model runs on IBM p690)
- Installation of the newest version of ecaccess software is planned this May.
- No projects run at ECMWF

ECMWF data coming to HMS

Data type	number of files	MB/day	arriving time [UTC]
Data coming via dissemination			
H1D - GRIB DET 12 UTC	48	98	10.00 pm – 2.00 am
H7D - GRIB DET 12 UTC	53	56	"
H1E - GRIB EPS 12 UTC	10	3	2.00 am – 3.00 am
H4E - GRIB EPS 12 UTC	41	32	"
H3A - BUFR DET WORLD	1	5	2.00 am – 3.00 am
H5A - BUFR DET HUNGARY	1	1	"
H5B - BUFR EPS HUNGARY	1	1	"
H6B - BUFR EPS WORLD	1	1	"
Data downloading from MARS			
EPS 00 UTC Control Forecasts	30	2	4.00 pm – 5.00 pm
EPS 00 UTC Perturbed Forecast	15	2	"
EPS 00 UTC Forecast Probability	30	1	"
DET 12 UTC Model level	29	63	5.00 am – 6.20 am
Seasonal Forecast for Hungary	5		

Future plans

- Complete the ORACLE meteorological database (CLDB) with special (air-pollution, forecast, ...) data
- Establish the ORACLE database for registry of Central (file) Archives and Central IT Data Bank
- Automatic product generation and dissemination system (FTP, e-mail, SMS, FAX, ...)
- Meteorological Web portal for users registered
- New WEB based meteorological visualization system
- New WEB based application management system (NWP monitor, ...)
- Further development of WEB based visualization for ECMWF forecast and verification (Intraweb)
- Operational use of EPS clusters for Central European area (based on Metview macro)

Thanks to User Support and staff of ECMWF for fruitful co-operation.

ICELAND ICELAND

Halla Björg Baldursdóttir – The Icelandic Meteorological Office

The computer system at IMO

- Servers:
  - Sun/Solaris (File and printer servers, special projects)
  - IBM/AIX (Database server, DB2)
  - Linux (Special projects)
  - Windows 2000/NT (File and printer servers, application servers)
  - Cisco (Firewall and VPN)
  - VAX/VMS (hopefully phased out 2004) (Special projects)
  - Alpha/Digital Unix (phased out 2003)
- Clients:
  - PC's running Windows 95/98, NT, 2000, XP
  - PC's running Intel Solaris or Linux
  - Sun workstations running Solaris
- Communication:
  - Telephone lines (Internet, direct lines, ISDN, X25, RMDCN)
  - Satellite receiver, radar receiver...

New Database Server

- Start of project: April 2002
- Tools
  - Database DB2 on IBM/AIX P620
  - WebSphere Studio for development
  - Java programming language, JDBC and XML
  - ClearCase for code and documentation
- Main projects
  - Move all older data from Ingres to DB2
  - Develop a data warehouse including observations, Kalman filtered parameters and selected gridpoints and parameters from the direct model output

The Microsoft Dilemma

- Office 97 was standard on all Windows workstations
- Upgrading all workstations to Office 2000 too expensive
- Only advanced users of MS Office were upgrated, 10-15 licences (of 60)
- Other users must use Open Office
  - not without problems

Model output received at IMO

- From ECMWF - Operational forecast model
  - UTC 12, 42 MB/day:
    - 1.5° resolution up to +240 hrs
    - 0.5° resolution up to +240 hrs
  - UTC 00, 42 MB/day:
    - 1.5° resolution up to +240 hrs
    - 0.5° resolution up to +240 hrs
  - Global wave model, 1 Mb/day
- From DMI – Hirlam, 83 Mb/day
- From UKMO – UKG and UKL, 6 Mb/day

ICELAND ICELAND

Usage of ECMWF products

- MARS archive
  - For research projects
- Deterministic model
  - Forecasting data for meteorologists
  - Kalman filtering of ECM 12 (T2m,W10m) for more than 100 local weather stations
  - Adaptive Kalman filtering, gives reliable prediction intervals, pre-operational, will soon replace the older KF
  - Post processed T2m maps using a Digital Elevation Model
  - Precipitation forecast maps based on a statistical model
  - Probability of precipitation based on a statistical model
  - Ice accretion on ships
- EPS
  - Use of EPS on ECMWF website
- Wave
  - Used by the Icelandic Maritime Administration to calculate “dangerous waves”

Experience

- Very satisfied with the staff of ECMWF
- Very satisfied with the web site, especially electronic documents related to various events
- Problems updating ECWMF certificate on workstation
- Problems accessing ECMWF Dissemination System
- ECaccess system, information on policy, use and structure would be of great help
  - Better flow of information to contact persons
- Install Metview 3 on Linux
  - Not an easy task!
  - Are there any plans of implementing Metview on Windows platforms?

IRELAND IRELAND

The Irish Meteorological Service

(Department of Environment, Heritage and Local Government)

Paul Halton – Met Éireann

1. Computer systems at Met Éireann

The main computing environment at Met Éireann is shown in Figure 1. The Linux Operating System (Red Hat or SuSe) is the preferred platform for the development of in-house applications. A number of dual processor Linux PCs, with the PGI Fortran compilers, are used to run NWP experiments.

NWP Servers

- The IBM RS/6000 SP with 10 WinterHawk II nodes is used to run the operational 3D-VAR Analysis, the HIRLAM model, the WAM model and the hourly analysis.
- This server has 9 x Power3 SMP Nodes (each with 4 x 375MHz CPUs). It has one I/O node. Each node has a 9.1GB disk pair. It has a total of 19GB of distributed memory and RAID-5 SSA disk management. The disks are connected in a loop configuration and there are two paths to each disk. The system has a Control Workstation that has 256MB memory and 3 x 9.1GB disks.
- In March 2003, IBM updated all the system software including the AIX operating system, Loadleveller, PSSP and FORTRAN and C compilers. After the upgrade, all applications were recompiled and tested.
- Backup versions of HIRLAM and WAM models are run on a Dell dual processor Precision 530 MT, Xeon 2GHz PC running Red Hat Linux version 7.1.
- A dual-processor Linux PC is used to run a Foot-and-Mouth Disease (FMD) model.

Graphics Servers

- Two SGI Origin 200 servers continue to decode and process GRIB bulletins received from the GTS and ECMWF over the RMDCN circuit. Configuration: 1 x MIPS R10000 processor (180 MHz each), 128MB Memory, IRIX 6.5, and 12.7GB disk capacity each.

  The graphics servers support X-based applications to display NWP products from HIRLAM, UKMO, DWD and ECMWF. The graphics servers also support Plotting applications; ADE database (BUFR), DNS services and file and print services for UNIX systems.
- An SGI O2 system is used as a FAX server to store and forward Analysis, Forecast and WAFS charts. Aviation weather charts are automatically faxed to aviation forecast offices. This server also supports the delivery of weather forecast products for the Met Éireann web site (www.met.ie).

UNIX Workstations

- An SGI O2 workstation is used to display NWP products in the General Forecast Office. All Research and Applications development staff use high spec PCs with Red Hat or SuSe Linux.

Climatology Servers

- Sun Ultra 170 running Solaris (2.5); Ultrasparc 170 MHz processor; 18.5 Gbytes storage. Runs OpenIngres RDBMS from Computer Associates (CA). This is the Climatological database server used for long term storage and on-line retrieval of historical weather data.
- Sun Enterprise 250 Server running Solaris (2.6); Ultrasparc 300 MHz processor; 13 GB storage. Runs OpenIngres RDBMS from Computer Associates. This is an applications server. It handles climate data processing and automated transmissions of climate data to customers. It also hosts climate user logins and hosts the Climate INTRANET web site.
- Sun Fire V880 server with 2 x 900 Mhz UltraSPARC(tm) III Processors, 6 x 73GB FCAL Disks, 3310 Array (12 x 36GB Disks), DVD, SDLT220. Operating System Solaris 8, Fortran 77 and C/C++ compilers. Supplied by Fujitsu in December 2002. Used for Climatological database & OpenIngres applications.

IRELAND IRELAND

Weather Radar Systems

- EWIS (Ericsson) Radar data from **Dublin Airport** is received over a 2Mbps Microwave link and stored on a VaxStation-3300 linked to a MicroVax-3100 through a Local Area VAX Cluster. Runs VMS V5.5-2 and Multinet 4.1. On receipt, these radar data are forwarded to the Rainbow applications server.

  RAINBOW (Gematronik) Radar data from **Shannon Airport** is distributed on a network of 5 x DEC AlphaStation 200 4/100, running Compaq Tru64 UNIX V4.0E. Two servers (Shannon & HQ) have 128MB RAM and 10GB disk capacity each. Three clients have 64MB RAM and 6GB of disk capacity each. RainBow V3.3 is installed. High Resolution 1km x 1km Irish Radar Composites are generated and sent to the TV graphics presentation systems.

- **UKMO RadarNet Composites** (5km x 5km) and Sferic data are received on the RMDCN link from UKMO and stored on a new dedicated WRADS2 server. The Java-based WRADS2 server and clients have been developed in-house and are in operational use since Q1, 2003.

  Since June 2002, radar data from 4 beams are sent to UKMO in place of Pseudo-Cappi products.

  – Four BUFR files are sent every 15 minutes instead of one from each radar
  – GTS headers were changed from PAIE40 EIDB to PAAL40 EIDB
  – Radar images are downloaded to the INTRANET site and TV Graphics workstations. Radar data is archived in, and can be retrieved from, the Climatological database.

Communications Computers

- The Vax-4200 cluster handles the real time reception and pre-processing of data from the AFTN, IAA and Irish observing stations. A project to replace the VAX-cluster is currently in progress.

RMDCN Servers

- 2 x PC's running Red Hat Linux are used to handle the processing of all incoming RMDCN data, including data from ECMWF and the GTS. They also transmit outgoing GTS data.

Web Site, FTP Site and Intranet Site

- Hosting and support of the Met Éireann web site (www.met.ie) is outsourced to Fusio Ltd. Forecast and recent weather data & static content are prepared at HQ and routinely uploaded to the main web site.
- An SGI O2 system at HQ supports the delivery of weather forecast products for the web site and it hosts the main Met Éireann Intranet Site. Many users have developed Intranet pages.

Internet Firewall

- Gauntlet Firewall on a Windows NT server protects the permanent Internet 256kbit link to a local ISP, Eircom.net. The firewall will be upgraded in 2003 with another Proxy Server. (Sidewinder)

Data Collection Systems (DCS)

- 2 x Linux Servers support the DCS system for the collection of weather reports from Automatic Weather Stations. Reports collected by the DCS are sent to the VAX-4200 cluster for bulletin distribution.
- The Linux-based PC application OBS, is installed at each manual synoptic station. The OBS system applies local quality control to all Synoptic reports before they are sent by dial-up to the DCS at HQ.

Satellite Systems

- **PDUS**: METEOSAT and GOES data are received, stored and processed on a DecAlpha Station 255, with OpenVMS v6.2H. VCS, Germany, supplied the PDUS software. Special images are produced for TV Graphics Systems, SATREP, Intranet and Web Sites. From March 2003 PDUS data is used to test the to new MSG reception and product production system.
- **EARS**: In March 2003, a new server and DVB reception equipment was purchased from a German Company, T-Systems for the reception of EARS data for input the NWP data assimilation system.
- **MSG**: In 2002, Met Éireann, following EU procurement procedures, awarded VCS the contract for the supply of MSG data reception, processing and display equipment. In March 2003 VCS installed the computer equipment and software and set up data access links to the PDUS and DVB reception servers. VCS also set up a remote access facility to enable them to logon to the MSG system over the Internet by prior arrangement. VCS returns to Dublin to install the MSG antenna at the end of May 2003. The entire system will be site accepted when MSG data is received and processed and displayed correctly.

IRELAND IRELAND

Road Ice Prediction System (RIPS)

- A Linux Server supports the RIPS system which collects reports from 50 Road Side Automatic Weather Stations. 2 x Windows 98 client PCs are used to view the observation data and to input NWP forecast data which is transferred to the National Roads Authority.
- NWP products from the fine-mesh version of the HIRLAM model are uploaded to the RIPS.

Office Systems

- A Dell PowerEdge 4600 server was procured in 2002 to enhance desktop file & print facilities
- Two Windows NT Servers provide office services, including e-mail, file and print facilities.
- MS Exchange on the Windows NT servers supports MS-Outlook on c. 130 user PC's.
- An SMTP Gateway Server on BSD UNIX supports Internet mail and DNS services.
- Dr. Solomon's Anti-Virus and MAILsweeper with Sophos anti-virus toolkit are used

Graphics Output Devices

- 2 x HP DesignJet 1050C plotters with Postscript are used for plotting charts at CAFO & Shannon.
- Laser printers in use include: HP 4000N, 4050N and 8000N; 1 x Tektronix Colour A4 LaserJet. A Mita photo-copier with multiple printing options is connected to the HQ LAN.

TV Graphics Systems

- 4 x SGI O2 workstations, running IRIX v6.5, 512MB RAM are installed - two at RTÉ and two at HQ. All systems run Metacast Ultra from Weather One AB, Norway. At HQ two O2 workstations are used for TV forecast preparation. At the TV stations, the O2's are set up as clients.
- The 'Borealis' TV Graphics package from Weather One is currently being evaluated on Linux workstations.

Networks

- The HQ LAN operates at 100Base-T (Fast Ethernet).
- The LAN equipment consists of 5 x Cisco 3548-XL switches, 5 x GBIC 50cm cables. The installation also includes CiscoWorks and "What's up Gold", for Windows to enable IT Operations staff to monitor the status of the nodes on the network.
- LAN equipment was installed at Valentia Observatory and the instrument PCs and desktop PCs systems have been networked and are managed from a central point. A new Windows-2000 File and Print Server was installed there in 2002.

Future Plans

New NWP Projects at Met Éireann

- Project for development of new SAF Nowcasting and Climate products from MSG satellite data
- Derma Project in collaboration with DMI. Funding dependent on Environment Protection Agency
- A small Linux Cluster will be procured in 2003 for the development of NWP experiments

Special Projects at ECMWF

In 2003, two special projects commenced in Ireland. These are as follows:

- **An IRISH Project, based at Met Éireann, Dublin**: The Community Climate Change Consortium for Ireland (C4I) will establish a capability to address the climate issues of vital national concern in an effective, coordinated way. At the core of C4I will be a Regional Climate Analysis, Modelling and Prediction Centre (RCAMPC) to simulate the detailed climate of Ireland and model climate change.
- **Joint DANISH and IRISH Project, based at UCC, Cork**: Verification of Ensemble Prediction Systems for a new market: Wind Energy.

RMDCN Link

- The RMDCN base bandwidth is 128kbps (96 to ECMWF; 32 to UKMO). The C4I project could significantly increase demands for additional bandwidth on the link from mid 2003 onwards

IRELAND IRELAND

ECaccess

- ECaccess not fully installed at Met Éireann so far. In May 2003, initial steps were taken and some difficulties were encountered. These have yet to be resolved. ECcopy is still in use in some user scripts.

Product Dissemination - Current Requirements

- ECMWF Product dissemination – Currently 3 data streams operational (3 spare)

IR Stream (Main dissemination run)			
Boundary files for Hirlam Swift	(SC),	13508 Products, Size	658.0Mb
Deterministic data	(DA),	2158 Products, Size	19.2Mb
Wave data (Global)	(WV),	506 Products, Size	2.5Mb
Wave data (EUR Shelf)	(WV),	231 Products, SIze	4.3Mb
Ensemble	(EF),	72 Products, Size	0.2Mb
IW Stream			
Boundary files for Hirlam Humid	(SC),	13504 Products, Size	91.9Mb
Boundary files for WAM	(WV),	11520 Products, Size	2.67Mb
IH Stream			
Data files for Marathon	(WV),	14400 Products, Size	3.45Mb
			Total per day 782.00Mb

Future Dissemination Requirements

If dissemination of the 0000z data is restored, that would add an extra 25 to 29Mb per day, depending on what we request. We can use the spare data streams to assess the impact on the link if required.

The Deterministic schedule includes:

- data for every 3 hours up to 48hrs
- data for every 6 hours from 48hrs out to 240hrs

SMS Batch queues

A number of jobs currently run under SMS batch at ECMWF

- Trajectory data for UCG under bc00h072 smsmars/ Traj Job launched at end of smsmars
- Probability charts for CAFO under bc00h072 probpltmag
- Sea Temps Mars request under bc00h072 sstreq
- Wave Plumes for CAFO under ef12h240 plumes_ireland

We are unable to run no more than three jobs under bc00h072. Is there a limit of 3 jobs?

IRELAND

IRELAND

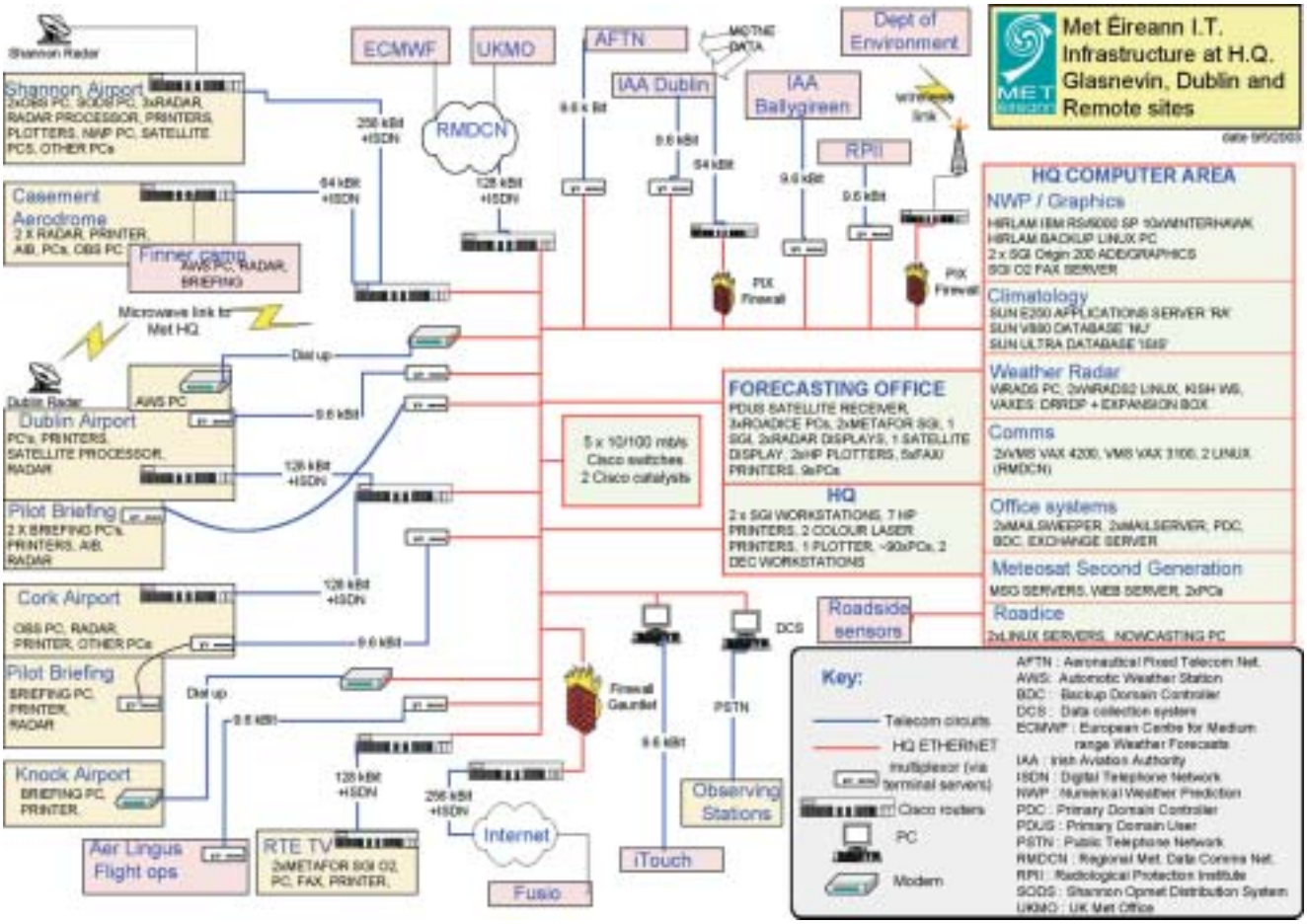


Figure 1

ITALY ITALY

Col. Giuseppe Tarantino – Italian Meteorological Service

1. Computer equipment and ECMWF connection.

During last year IMS has managed to complete the upgrade of the Computing equipments improving CPU and Disk Capacity of the Servers and installing the new equipments for the RADAR Area, the DVB satellite broadcast system and the new Computing facility system.

a. Radar Area

We are removing the old MOTOROLA Computers both at the RADAR sites and at the collected data Center with the new ones:

- Server COMPAQ DS20 (Two CPU) at every RADAR Sites,
- Server COMPAQ ES40 (Four CPU),at the collect data center
- WS COMPAQ XP1000 or Compaq DS10 for graphical presentation.

At present we have finished the installation at two RADAR sites and the collect data Center and are concluding the acceptance text.

b. Broadcast area

We have installed a WS COMPAQ XP1000 as server for a DVB satellite Broadcast system to distribute data to users according to definite schedule. The system is operational.

c. NSED (New Computing Facility System)

The acceptance tests of the new Computing System have been concluded and it will be operational in a short time.

This area is based on a Cluster memory channel connection of five COMPAQ Alpha Servers, each of them employed to carry out different jobs:

- Number crunching Area - One COMPAQ GS60 Server with 4 processors and 2GB RAM memory. This server is used for the run of Italian High Resolution Model;
- Operational Meteorological Area - One COMPAQ GS60 Server with 4 processors and 2GB RAM memory. This server is used to produce messages and products for operational duties;
- Meteorological Data Base Area - One COMPAQ AS 4100 with 2 processors and 1 GB RAM memory. This server is used to get the access to the online meteorological data base;
- Climatological Data Base Area - One COMPAQ AS 4100 with 2 processors and 1 GB RAM memory. This server is used for climatological purposes.
- WS Areas - WS DS10 with 512 MB RAM memory and 40 GB HD;
- System and Application Supervisor Area - One COMPAQ AS 1200 with 1 processors and 512 MB RAM memory. This server is used for the monitoring of activities, processes, LAN, WAN, Data Bases, ecc.

The total amount of Storage for the cluster is about 100 GB on a RAID Storage Cabinet connected via memory channel with each member of the cluster.

d. DMZ

We have upgraded our DMZ according the the following scheme:

- 2 Server COMPAQ DS20E (Two CPU) in a cluster memory channel configuration, used as Internet/Intranet Web Server;
- 1 Storage Device COMPAQ RA3000 with 50 GB of disk storage total amount in a RAID 0+1 configuration.
- 1 Server Compaq ProLiant ML350 Generation 2 Hot Plug (1 CPU - 1 GB RAM) used as Internet/Intranet Mail Server;
- 1 Server Compaq ProLiant ML350 Generation 2 Hot Plug (1 CPU - 1 GB RAM) used as Antivirus Gateway

A new INTRANET SYSTEM, named "PROMETEO", has been developed in order to allow the access to Meteorological Data and Products by INTERNET.

ITALY ITALY

e. DataBase Area

We have upgraded our DataBase according the the following scheme:

- 2 Server COMPAQ DS 20 (One CPU) and COMPAQ AS4100 (Two CPU) in a cluster memory channel configuration used, as DataBase Server;
- 2 Storage Device COMPAQ RA3000 with 200 GB of disk storage total amount in a RAID 0+1 configuration.

f. Message Switching Area including RMDCN(GTS & ECMWF)

We have changed one of our Message Switching Server (DEC AS2100) with a COMPAQ AS1200 (Two Cpu - 512 MB RAM Memory) used as Operational MS Server (the stand-by one is still a DEC AS2100).

2. Project, experience and plans

We have about 110 users using ECMWF services and most of them use INTERNET access. These users are mainly from:

- University
- Regional Meteorological Services
- Research Agencies
- Environmental Agencies
- Armed Forces
- Environmental Hazard Department

The main usage of ECMWF services is the retrieval of MARS data associated with the decoding software to run either models or MAGICS and METVIEW applications.

ECMWF GRIB data are routed in real / delayed time to users for their operational activities (environmental hazard, agriculture, pollution etc.).

At the Operational Center ECMWF data are also used:

- as support for the operational meteorological activities
- as boundary condition for the forecast models
- as input for post processing programs
- as input to produce information and maps usefull for aeronautical activities

A subset of ECMWF GRIB data are still distributed to MDD users by Rome MDD Uplink Station.

Projects:

- A test suite of the Lokal Model daily runs at ECMWF to verify the impact of Boundary Condition (originated from IFS) versus the operational version running with BC from GME. Moreover the model is tested by introducing additional parameters like cloud ice content. A selection of these products are going to archive on the Member State MARS area.
- COSMO LEPS (Limited area Ensemble Prediction System) project.
  The project is to produce routinely probabilities and description of severe weather events over a geographic region enclosing Germany, Italy, Greece and Switzerland. Multiple runs of the COSMO limited area high-resolution model (set at 10km resolution) will be initialised by appropriate ECMWF EPS members selected via a specific clustering technique in order to maximize possible meteorological perturbations. The SW to check the output has been carried out and the first results of this project will be available next autumn.

The current special projects are:

- Evaluation of the performance of the ECMWF meteorological model at high resolution.
- Non linear aspects of the systematic error of the ECMWF coupled model
- Limited Area Model Targeted Ensemble Prediction System (LAM-TEPS).

ITALY

ITALY


NSED
ECMWF
Bracknell
Offenbach
Toulouse
Athens
Ankara
Beirut
RMDCN
Numerical Products
WMO-ICAO Messages
UMS
Satellite Data & Products
SATELLITE
RADAR Data & Products
RADA
DMZ
Military & Civilian USERs
Data Base Area
Data
Data Subset
End User products
Climatological Data
End User Products
DVB USERs
Number Crunching Area
Numerical Products
Data Subset
End User Products
Operational Area
Archives & Climatological Area
Remote WS Area
Graphic Data Handling
WS Area

NETHERLANDS NETHERLANDS

Hans de Vries – Royal Netherlands Meteorological Institute, De Bilt

KNMI's Computer Infrastructure and ECMWF

KNMI is the Dutch national data and knowledge centre for weather, climate and seismology. It performs scientific research, maintains the extensive infrastructure for observations and information processing and provides data and services, general weather information and severe weather warnings.

In 1999 all market related ('commercial') activities have been transferred to a newly established company HWS. Since then the only products which go directly to the public are the severe weather warnings and weather reports and forecasts on the Internet.

KNMI operates from its main office in De Bilt. In 2001 also the operational activities which were formerly performed at branch offices at Schiphol Airport (aviation) and Hoek van Holland (maritime meteorology) have been transferred there.

1 Infrastructure

Figure 1 gives a schematic view of KNMI's current computer infrastructure. The LAN is centered around a Gbit/s backbone with connections of at least 100 Mbit/s to all endpoints.

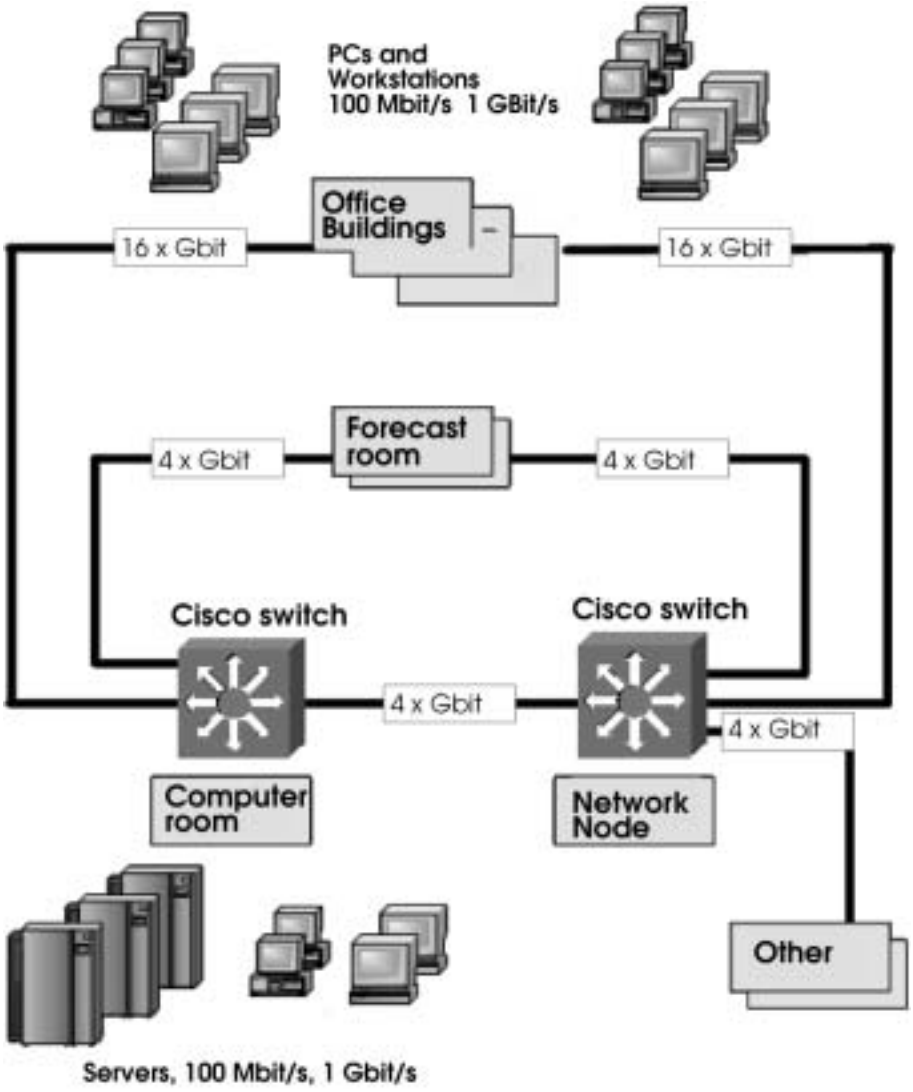


Figure 1: KNMI's computer Infrastructure

NETHERLANDS NETHERLANDS

Computer facilities

- A StorageTek PowderHorn 9310 mass storage system, capacity 30 TB.
- A 2 CPU SGI Origin 2000 (backup 2 CPU Origin 200) server for the mass storage system (DMF software).
- A Sun Fire 15000 server (StarCat) with 6o UltraSPARC 111 900 MHz CPUs (12 as compensation for bad OS performance), 1 GB memory per CPU, for (operational) HIRLAM, Climate research and applied research. The number of CPUs will be increased to 84 later this year.
- A cluster Of 3 DEC-Alpha workstations for operational applications.
- Many (≥ 20) DEC-Alpha meteorological workstations.
- Several (~ 4) DEC-Alpha workstations for development of operational applications.
- 66 SGI-Linux workstations.
- 33 Linux workstations on Compaq hardware.
- ~ 15 Linux workstations on HP hardware, waiting for installation (and a Linux upgrade).
- Still several SGI-Irix (O2) personal workstations.
- Two identical SGI-Linux NFS servers with general software and site licenses for workstations.
- Many (> 500) Windows-XP PCs.
- 5 SGI Origin 200 Web servers.
- A WNT exchange server which handles all email.
- WNT Windows Terminal Servers to give access to WNT applications in a UNIX environment (Citrix client).
- Several WNT application servers.
- A dual Windows 2000 server for message switching.

Connections

Connections to the Internet and other telecommunication services are separated from the LAN by a firewall. KNMI's Public web server (www.knmi.nl) is located outside this firewall

Table 1 lists the main outside connections. Other sites, service providers and institutions also have dedicated connections to KNMI.

Connection		Capacity [bit/s]
Internet		1 G
	Firewalls	4 × 100 M
RMDCN	Access line	128 k
	ECMWF	96 k
	Met Office (GTS)	24 k

Table 1: Main KNMI connections

Developments in 2002

- LAN upgrade toward a Gbit/s backbone and ≥ 100 Mbit/s everywhere.
- Breaking up the high-availability SGI-Linux NFS server into two mirrored separate servers: the high-availability software was unsatisfactory.
- Installation of a new observations network, in production since April this year. This makes observations with a time resolution of 10 minutes available in real-time inside KNMI and for external service providers.

NETHERLANDS NETHERLANDS

Challenges for 2003

- Implement the ECaccess gateway (finished).
- Implement the Trajectory model on the new HPCF platform (finished).
- A new infrastructure for image data, in production this summer. This makes extensive use of HDF-5.
- Upgrade the tape technology for the mass storage system to cope with the present 1 TB/month growth.
- Utilize the wasted power of workstations outside office hours.
- Find a solution for continuing performance and stability problems of the WTS servers.
- Upgrade firewalls.
- Protection and integrity of information.

2. ECMWF use

Operations

For its weather forecasts, KNMI uses various ECMWF products. To give an impression:

- The deterministic model for medium-range (up to day 5) forecasts.
- The Ensemble Prediction System for forecasts up to 9 days ahead (also available on the Internet).
- Trajectory computations.
- Boundaries from the deterministic model for HIRLAM.
- Forecasters like to use products available on the ECMWF web.

Software packages

The following software packages from ECMWF have been installed for general use:

- emoslib
- eccopy
- Magics
- Metview
- ECaccess gateway

Users

On May 7, 2003, there were 87 Dutch users (92 on May 13, 2002) with an account on ecgate 1. Of these 45 had logged in to ecgate 1 earlier this month or in April; 15 had logged in earlier in 2003; 14 had not logged in since May 28, 2002. Three new users had never logged in yet.

Twelve users are from outside KNMI: 8 from Netherlands' universities and 4 from research institutes.

Projects

Table 2 lists the Dutch projects for access to the HPCF services with their total use of resources from 1999. Of the allocated 356 kSBU for 2002, 312 kSBU or 88% have been used. Special projects, of which some were extended to deal with the high demand of computing power this year, however, used another 51 kSBU or 71% of the allocated 72 kSBU.

HIRLAM users also used the HIRLAM Special project, which is administered by Sweden.

62 users have access to the HPCF in one of the regular projects. Furthermore, 13 users have access to at least one of the special projects. 25 users have no access to HPCF.

For 2003 the total allocation for regular projects is 579 kSBU. Until May 7 45 kSBU or 7% had been used. Three remaining Special projects have a total allocation of 95 kSBU, of which 200 SBU had been used.

NETHERLANDS NETHERLANDS

		1999		2000		2001		2002	
		Used [kSBU]	Alloc [kSBU]	Used [kSBU]	Alloc [kSBU]	Used [kSBU]	Alloc [kSBU]	Used [kSBU]	Alloc [kSBU]
Operations									
nlcwdtra	Trajectory computations for local emergencies	0.2		0.2		0.2		0.2	
Climate research									
nlglomac	Atmospheric Composition	1.4		3.1		63.2		97.6	
nlhirkli	Climate research with HIRLAM	0.2		3.5		1.5		0.7	
nloocean	Oceanography	7.1		6.7		21.4		87.4	
nldmwgac	Other climate research	39.4		21.3		26.0		28.9	
Applied research									
nldmlam	HIRLAM maintenance	14.3		25.0		60.0		96.8	
nlsatnwp	Satellite data research	14.1		0.0		0.0		0.0	
nlwmcalc	Other applied research	0.3		0.5		0.3		0.4	
Total		77.0	142.0	60.4	348.0	172.6	336.0	312.1	356.0
Special Projects									
spnlagul	Agulhas ocean model	1.7	15.0	0.3	20.0	6.5	20.0	9.1	10.0
spnlctmk	Chemistry and transport	1.8	3.0	0.3	3.5	0.0	4.0		
spnlflux	Validation of A/S fluxes	17.1	20.0	0.0	20.0	0.1	20.0		
spnlles	Large Eddy cloud simulations	0.7	10.0	2.5	15.0	4.3	25.0	2.8	17.5
spnlmix	OGCM modules and assimilation	2.5	15.0	0.0	15.0	2.9	15.0	15.3	20.5
spnlozas	Assimilation chemical species	0.02	1.7	0.0	1.7	0.0	1.7	0.0	0.0
spnlprob	Short-term regional prob fc	5.4	5.0	8.7	10.0	0.0	10.0		
spnlsiep	Stratosphere - troposphere exchange	1.2	2.0	2.5	5.0	0.8	4.0	24.0	24.0
Total		30.3	71.7	14.3	90.2	14.6	99.7	51.2	72.0

Table 2: Netherlands' projects in 2002

3 Experience and suggestions

New HPCF

The new IBM HPCF computers have not yet been used by many people. Those who have are generally pleased by the increased speed. As most of them are using HIRLAM or IFS, which have been ported by others, there is still not much conversion experience.

A reason given for not using ECMWF anymore is that present-day PC processors are just as fast for non-parallelized code and that it is easier to run jobs on your own workstation then.

ECaccess

The ECaccess gateway has been up-and-running since mid March. Users who have started to try out the different possibilities are enthusiastic. However, there is a threshold to overcome before people actually start And not enough people have the experience to show their colleagues yet.

One user is using *ecxterm* to start PrepIFS and is very happy with that.

Logging in through ECaccess to ecgate 1 and the web interface for file transfer have already found widespread use.

NETHERLANDS NETHERLANDS

Suggestions for ECaccess

- Something to handle transfers with ectrans to the same user on different machines.
- Include (recently) completed ectrans transfers in the ectls listing,
- How to make a simple switch for ECDOMAIN?
- Include subdirectories in ECDOMAIN?
- Could ecxterm use the same certificate as e.g. ecget, ecput?
- The web interface of ecaccess.ecmwf.int and ecaccess.knmi.nl look exactly the same, which can sometimes be confusing if people e.g. like to change their MS user association.
- An early warning when the gateway certificate expires would prevent many problems.

General comments

The help of User Support, especially John Greenaway, is very much appreciated. Any questions are always answered promptly and adequate.

Laurent Gougeon has been a great help in getting the ECaccess gateway running.

Aware of the large amounts of data which are requested and what John Hennessy commented last year on this issue, users are still sometimes disappointed by the turn-around times for MARS tape retrievals.

What is the status of user (de)registration over the Web?

NORWAY NORWAY

Rebecca Rudsar – met.no, Oslo

The computing environment at met.no

The communication for data acquisition and routing of observations is managed by the NORCOM software. This software now runs under Linux OS.

Observation data is transferred to the SGI Origin 2000 for decoding. The decoded observations are stored in an internal format and temporarily in BUFR format for input to the 3D-VAR analysis program. Since May 2002, ATOVS (AMSU-A) radiances have been assimilated in the 3D-Var analysis of the main model. The plans are to also introduce QuikScat winds later this year, together with a software upgrade of the HIRVDA assimilation system.

VAX4000-200 is used for running the Nordic Radar system, NORDRAD, which is a communication system for the distribution and collection of radar data between Nordic countries.

In addition to the radar unit (Hægebostad) covering the very south of Norway there are two new radar units. One at Bømlo, just north of Stavanger and one at Rissa, near Trondheim.Two Sun Ultrasparc10 computers and two Hewlett Packard workstations are used for acquisition, for product generation and monitoring and for display.

Processing of data from NOAA satellites is done by the MEOS system (Multi Mission Earth Observation System). The MAPP (MEOS Advanced Product Presentation Facility) is installed on five SGI workstations.

Alpha-200 is used for processing the data obtained from the geo-stationary satellite Meteosat. The data consists of satellite pictures, DCP (Data Collection Platform) data and MDD (Meteorological Data Distribution) data.

Oracle software is used for the Climate database and a Verification database. It is also used for the Climate data warehouse.

A new system (KVALOBS) for quality control of observations is under test. The system contains automatic control routines which satisfy specific criteria for real time transmission, forecasting and climatological use.

The total number of drives in the StorageTek tape robot is three. In addition to being used for archiving observations and model results, the tape robot is heavily used by several research projects. A DLT tape station, HP Surestore DLT80, connected to a Linux PC is used for the remote backup of specific disk areas. Backup copies are taken approximately every 3 months and the tapes are stored at another site.

A new Linux-based file server was installed in December. This is being used as a 'Short term Archive' i.e. a selection of files from the operational routine are stored on this server for a period ranging between 6 months to one year. The data archived on the Tape Robot is a subset of these files. The file server is used mainly by the Research Department but also by the meteorologists for easier access to the most recent data.

The SGI Origin 200 receives the disseminated data from ECMWF and transfers it to the operational suite. Data from jobs run regularly at ECMWF and transferred using ECaccess (eccopy) is also processed on this computer. The RMDCN connection is 256 kbps.

Met.no is connected to the University network, Uninett, via a 10 Mbps connection for Internet and a 100 Mbps connection for access to the supercomputer in Trondheim. Supernet (ATM) is a part of Uninett and at present the theoretical transmission rate between the University of Oslo and the University of Trondheim (NTNU) is 123Mbps.The Supercomputing Project (NOTUR) at NTNU operates and supports High Performance Computing equipment for use by Norwegian universities and four other partners including the Norwegian Meteorological Institute.

The computer configuration at present is a SGI Origin 3800 with 384 500 MHz processors, 304 GB memory and a SGI Origin 3800 with 512 600 MHz processors, 512GB memory. Both computers are connected to a 7 TB fast disksystem through a common Clustered XFS. The operating system is TRIX - 6.5 Release. A third SGI computer is also connected to the common disk system.

On March 20, 2003, all HIRLAM models were upgraded to version 6.0.0. At the same time, the 50 km model (with 31 levels) was replaced by a 20 km model with 40 levels covering the same area. The 10 km model, that used to be nested inside the 50 km model, now get boundary frames directly from the ECMWF LBC project, except for the analysis, which is interpolated from the analysis of the main 20 km model. The 10 km model is also upgraded to use 40 vertical levels.

The main computers running the operational suite are still two SGI computers. The Supervisor Monitor Scheduler (SMS) and the X Command and Display program (XCDP), developed at ECMWF, are used to control and monitor the operational suite.

NORWAY NORWAY

We are in the process of migrating the operational suite to computers running Linux. Contracts for new servers and disk system are being negotiated at this moment. As a part of the process of porting jobs to run under Linux two Linux based computers have been defined as part of the operational computer configuration. It is possible under SMS to specify on which computer a job shall execute thus enabling the testing of individual jobs.

The tele-communication links to the forecasting centres at Bergen and Tromsø are 2 Mbps. Operational products are distributed to two fileservers at each centre using the TCP/IP protocol. ISDN links (128 Kbps) are used as backup.

The satellite distribution system (Infosat) will be fased out in June. The customer database for the distribution of products via a WEB interface (Værbutikk) is now running under Linux OS. The system for processing data from automatic observing stations (AutoObs) also runs under Linux OS. The newer version of the distribution system (Transfer-Products) is operational and a web interface is being developed. The automatic fax distribution system (Autofax) is still running under AIX OS

The interactive DIgital ANAlysis application (DIANA) has been upgraded several times and is now used as the main graphical viewing system. The forecasting meteorologists run this application on 2-screen Linux PC's.

As well as running the MM5 model (1km resolution) in connection with an Air Pollution project (Bedre Byluft) the Linux Cluster is used as a backup computer for the supercomputer in the case of a failure of the network to Trondheim. Some of the backup models have a lower resolution but the results are then interpolated in order that products can be generated.

ECMWF Products

Disseminated data from the operational forecast model, the global wave model, the Ensemble Prediction System and Boundary Condition data are received from ECMWF. This data amounts to approx. 475 Mbyte per day.

Deterministic model:	43	Mbytes
EPS:	18	Mbytes
WAVE_Models:	3	Mbytes
Boundary Condition 1	350	Mbytes
Boundary Condition 2	61	Mbytes

Dissemination data received from ECMWF is converted from GRIB format and placed in our present fields database. The data is then accessible by the graphics packages which have been developed at met.no.

The data is also used

1. for general forecasting by the forecasting department.
2. as boundary values for the Norwegian limited area models.
3. as backup for the Norwegian limited area models.
4. as input to the maritime and air pollution models.
5. as input to a ship routing program for the Pacific.
6. the Norwegian Institute for Air Research still receives ECMWF data on a regular basis. The data is utilized in the European Arctic Stratospheric Ozone Experiment.
7. by a commercial weather forecasting company.
8. by the commercial department of Norwegian Meteorological Institute.

Data retrieval from MARS is used for research projects.

Projects at ECMWF

The following are registered as Special Projects in 2002-2004:

- Parametrization of clouds in general circulation models.
- Ozone as a climate gas.
- Targeted ensembles providing boundary values for limited area models.
- REGCLIM: Regional Cimate modelling.
- The HIRLAM Project.

NORWAY NORWAY

Experiments include :

- The generation of climatological fields.
- LAMEPS - limited area model ensemble prediction system. Using ECMWF EPS to perturb the HIRLAM model. Producing re-runs of EPS for a target area covering Northern Europe (on model levels).
- REGCLIM:Producing singular vectors and forcing singular vectors for all four seasons.

Feedback from the users

Total number of users = 32

Computer Rep., Operations = 2, met.no users = 24, external users = 6

The total usage in 2002 was 43% of our quota. The last account overview from April 2003 showed that we have so far used 8%.

In response to a request for feedback for this talk 13 users replied.

Seven of these replies were from users who had used/were using MARS. Two used PrepIfs. Three users had no immediate plans for using ECMWF.

Comments from users include:

"The Web pages contain a lot of information but it can be a bit difficult to find want you want. I have not found any information about how the MARS archive can be used most effectively, i.e. how the data is organised on the tapes. With this information one could, when fetching a lot of data, organise the jobs to use the least number of tapes."

"I would like a description of how to fetch observations from ECMWF"

"The information on the web- pages is not updated in accordance with the data which is available from MARS."

"It can sometimes be difficult to find out from which analysis/experiment a particular parameter is available and there should be more information about the differences between the datasets. The different types of analysis data available in ERA40 is especially confusing."

"Suggest a start-date/stop-date check in the Mars interface. When I, by mistake, defined 'date=20020101/to/200101001' then MARS chewed until MEMORY OVERFLOW or some such error message appeared!"

"When running eps_fc (PrepIFS) from my own singular vectors, I believe it is not possible to run for several days at once. This is because the variable that set the path to the place where the SV's are is set to default each time a run starts. Is it possible to change this so the right path to the SV's can be set for each date in the date-list?"

"It might be useful if the calculation of the PV (potential vorticity) structure of SVs (singular vectors) existed as an option in PrepIfs (SVs are now available as vorticity, divergence, temperature, stream function, geopotential height)"

"Information from WMO states that ECMWF will offer standard software for BUFR-decoding/encoding. What is the status and time plan here?"

"Will ECMWF also offer standard software for GRIB-decoding/encoding?"

"met.no has for many years used the observation decoding programs written in Fortran at ECMWF. Are these programs still being used at ECMWF? Have they been compiled and run under Linux OS?"

The users are enthusiastic about User Support and mention fast response and very good service.

All users use Internet for data transfer.

Plans include:

- many of the Special Projects continue into 2004.
- use ERA40 data in connection with Climate modelling.

The IT detpartment at met.no is responsible for a national project. The aim of the Norwegian Service Centre for Climate Modelling (NoSerC) is to help climate researchers and it offers technical assistance for managing data, data analysis and organisation and optimisation of climate model code.

- a global pollution model driven by meteorological data from ECMWF.

SLOVENIA SLOVENIA

Miha Razinger – EARS, Met. Office, Ljubljana. (Slovenia was not represented at the meeting but submitted this report).

1 Computers and network in environmental agengy of Slovenia, Meteorological Office

In Environmental Agency of Republic of Slovenia (EARS), Office of Meteorology we got a new cluster computer system in October 2002. It will provide computing facilities for operational computation of ALADIN local area model and research and development work of NWP team in Meteorological Office. The computer system has 14 nodes, each with two 2.4 GHz Intel Xeon processor, 2 GB RAM and 0.5 TB of disk space. Processors are connected via gigabit fiber ethernet. We are using open source (LGPL) SCore 5.4 global operating system maintained by PC Cluster Consortium (www.pccluster.org) together with Linux. SCore operating system offers optimized PM library for efficient MPI and PVM communication. Other significant features are customized MPICH with more bandwidth and reduced latencies, single system image, flexible scheduling in time and processor space domain trough gang and batch scheduling, checkpointing and pre-empting. It comes with set of tools to monitor and administrate the cluster. The cluster is in testing phase and will come into operational use in the middle of the year. More information in Jure Jerman's presentation:

http://www.ecmwf.int/publications/library/ecpublications/proceedings/high_performance_computing_2002/jerman_jure.pdf.

In parallel we are using our old cluster with five 533 MHz EV5 Alphas processors with 128 MB RAM, FastEthernet connections, MPICH message passing library and Linux.

We are working on a new server for central meteorological archive, where all measurements and observations from synoptic, automatic, climatological, hydrological, precipitation, airport, phenological and other measuring stations will be stored.

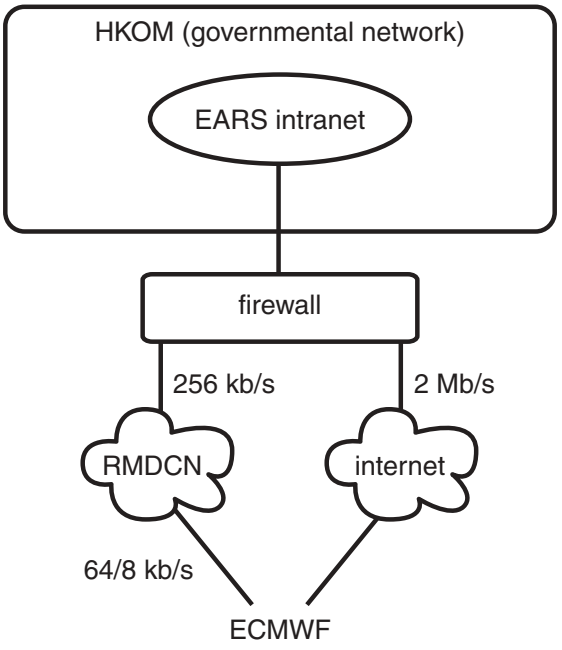


Fig 1 EARS connection to ECMWF

For Office's workservers and workstations we use PCs and Alpha boxes with Linux and some PC with Windows 2000/XP operating system. All computers are connected to local area network via 10 and 100 Mb/s ethernet. For access we use ssh, nfs, ftp and samba protocols.

In March 2003 we became a part of the HKOM, Slovenian governmental network. Because of a strong security policy, biggest problem now is the access to our LAN from the outside world. Some employees have gotten SecureID card for secure direct dial-up access, but for people outside the Agency or employees currently outside the country the access is virtually impossible.

SLOVENIA SLOVENIA

2 Use of ECMWF services

We are still a co-operating state, so we don't use ECMWF's high-performance computing services.

Two times daily we are transferring model output data of deterministic model, ensemble prediction results, multi-analysis data and results from the Mediterranean wave model through RMDCN line. There is limited use of new ECAccess facility, which users find elegant.

Recently we installed Supervisor Monitor Scheduler (SMS) software package for better control over our operational computation, visualization and dissemination jobs.

3 Future plans ECMWF services

We plan to use extensively ERA-40 products in the future. Due to relatively low spatial resolution of ERA fields it was decided that a regional dynamical adaptation (most probably using operational limited area model) should be performed. Suitable tests have already been made and gave satisfactory results.

Since a lot of information is needed to successfully nest a limited area model with available tools into ERA fields (global model level data) a considerable amount of ERA-40 data should be and is planned to be transferred from MARS gradually this year (about 2 G per month; planned 10 years = 240 G).

SPAIN SPAIN

E. Monreal – Instituto Nacional de Meteorología, Madrid

1. Main computer systems

The main computing environment at INM is shown in figure 1. The major changes since previous meeting have been the replacement of both the CRAY C94A and the message switching system and the move to a new computer hall - still in progress - with new LAN equipment.

Main computer:

The contract to replace the CRAY C94A awarded to CRAY Inc. included, as interim system, an SV1 configured with 16 processors, 16 Gbyte of memory and 160 Gbytes of disk storage. Installed in November 2002, it is operational since mid February running the HIRLAM NWP model as it was done on the C94A, that is 4 times a day with .5° and .2° resolution in a nested configuration.

CRAY's ultimate solution is, however, based on its new X1 system that designed as the successor to both the Cray T3E and SV1 product lines, combines powerful SMP vector nodes in a scalable MPP architecture. The first phase of the solution, with a 100 times increase over the C94A performance capability, is to be installed early next July. The second phase, to be accepted by the first quarter of 2005, will install additional X1e nodes to raise the total capability to 300 times C94A.

The specification of the initial configuration, is as follows:

> 10 computing nodes plus 1 system node. Each node has four vector Multi Stream Processors (MSP) sharing 16 Gbytes of high bandwidth (20.5 Gbytes/s) memory. An MSP is made of 4 vector chips with 2 dual pipelines each, running at 800 Mhz, 4x 400 Mhz scalar processors & 2 Mbytes of cache. Peak performance is 12.8 Gflops per MSP and 512 Gflops in total. Nodes are connected in a 51.2 Gbytes/s full-duplex 2D modified torus. I/O is distributed along nodes: 4x 1.2 Gbytes/s System Port Channels (SPC) per node. Our X1 will be configured with 1.8 Tbytes of directly attached disk storage and 2 Gigabit Ethernet connections.
>
> The solution also includes a Storage Area Network which connects the X1 with 3 Tbytes of additional disk and an ADIC scalar 100 robotic system with 4 LTO-2 drives and a capacity of 14.4 Tbytes of uncompressed data. Other SAN equipment provided by CRAY are the FC fabric (switches & hubs), two data servers and StorNext Management Suite from ADIC to provide a heterogeneous shared SAN filesystem, HSM and SAN management capabilities.

Final system configuration:

> 4 additional X1e nodes. According to CRAY's roadmap each X1e node will have 8 MSP running a 50% faster, i.e. 19.2 gflops peak performance and 32 Gbytes of memory. Thus, the final configuration will have a total peak performance of 1,126 Gflops.

Data archiving and retrieval system

It comprises two subsystems, a data handling system and a data management system:

> Data handling system: Acts as the central repository for user files and all operational data. It is based on AMASS software running on a 2 processor HP 9000 K570 server configured with 2 Gbytes of memory and 300 Gbytes of disk cache, which controls an IBM3494 tape library with ten 3590 drives and a total capacity of 17 Tbytes. At present, the system is storing about 9 Tbytes in 1,200,000 files.
>
> Data management system: Designed to provide access at the atomic level (fields) to model and observational data with an user access interface rather similar to MARS. There are four Relational Data Bases: observations, NWP model fields, satellite images and products, that were developed using ORACLE 8. The system comprises a cluster of two HP 9000 K460, configured with 2 processors, 1 Gbytes of memory each and 220 Gbytes of shared disk (in RAID5) running Service Guard High Availability software.

Main UNIX servers:

Sun Enterprise 10000:

Recently updated to:

- 6 system boards with 18 Sparc II processors running at 466 MHz, 18 Gbytes of memory and 720 Gbytes of disk.

SPAIN SPAIN

It is divided in four domains, the main are:

- New message switching & graphics dissemination. Operational since October 200, it is based on an in-house development that replaced UMS from Global Weather Dynamics.. It runs in a master/slave configuration using two domains of the Sun 10K that share 504 Gbytes of disk. All lines, including low speed circuits are managed by routers.
- Primary: 2 System boards with 6 processors, 8 Gbytes of memory, 36 Gbytes of disk.
- Secondary: 1 System board with 4 processors, 4 Gbytes of memory, 36 Gbytes of disk.
- NIS master server and file server for SUN workstations
- 1 System board with 4 processors, 4 Gbytes of memory, 144 Gbytes of disk.

2 Sun ULTRA Enterprise250:

- 2 processor, 512 Mbytes of memory, 36 Gbytes of disk each.
- This two systems handle the data pre-processing and Report DB as well as the reception of ECMWF dissemination, most of graphics production and post-processing.

Sun Enterprise3000:

- 2 processors, 512 Mbytes of memory, 12 Gbytes of disk storage.
- Applications development and testing.

Sun SPACRserver 1000E:

- 2 processors, 512 Mbytes of memory, 24 Gbytes of disk storage.
- Climate database based on ADABAS DBMS.

Sun Ultra10:

- 1 processor, 128 Mbytes of memory, 12 Gbytes of disk.
- Intranet Webserver and anonymous ftp server, secondary NIS server, Mars Cache server.

Sun Ultra1:

- 1 processor, 128 Mbytes of memory, 6 Gbytes of disk.
- NQE server for job submission to the CRAY C94A. It is also used as a gateway to ECMWF for non-operational usage.

Sun Enterprise-450:

- 1 processor, 256 Mbytes of Memory, 12 Gbytes of disk storage.
- Internet Webserver, anonymous ftp server

McIdas servers:

The McIdas production system which deals with satellite image processing and serves as operational WS for forecasters runs on a distributed environment. The main McIdas servers, handling data ingestion, are:

2 Sun ULTRA60:

- 2 processor, 1 Gbyte of memory each. and 18 Gbytes, 36 GBytes of disk storage each.
- national radar centre (radar & satellite image composition) and model grids data ingestion.

2 Sun ULTRA1 170E:

- 1 processor, 256 Mbytes of memory each. and 16G Bytes, 13 GBytes of disk storage each.
- GOES images and observation data ingestion.

2 Sun SPARCstation 20/712:

- 2 processors, 256 Mbytes of memory and 20 Gbytes, 8 Gbytes of disk storage each.
- TIROS and METEOSAT images ingestion.

SPAIN SPAIN

Other computers

Radar Network: A VAX 4500, running open VMS, handles radar images reception from the 14 meteorological radar currently in operation.

Network

All the computers in the central office are connected through a local area network (see figure 1). The LAN is basically Switched Ethernet, with a Gigabit Ethernet backbone between main switches, although an ATM backbone at 622 Mbps also links the data archiving and retrieval systems. A number of Cisco Catalyst switches manage the Ethernet giving 100 Mbps to each of the main UNIX servers and 10 Mbps to the rest of the systems.

All the subsidiary offices, 15 regional centres (CMTs) and 60 airport's Meteorological Offices (OMAs), are connected via Frame Relay to the central LAN in a wide area network. The current configuration is as follows:

	Access lines number x speed	**PVCs to Madrid (central office): number x C.I.R. (In/Out)**
Central office	8x 2 mbps (channels)	N.A.
11 CMTs (forecasting activities)	1x 512 kbps	2x 192/32 1x 32/32 (non meteorological data)
4 CMTs (no forecasting activities)	1x 256 kbps	2x 96/32 1x 32/32 (non meteorological data)
60 OMAs	1x 128 kbps	2x 48/16

The central office is linked to two different Points of Presence (PoP) through two diversely routed lines, whereas subsidiary offices have single access lines and ISDN as backup.

Connection to RMDCN: Backup ISDN connection and the access line are diversely routed. There are PVC to ECMWF, Météo-France, DWD, UKMO and Portugal.

The two connections to the Internet (a 2 mbps leased line to REDIRIS, the Spanish academic and research network, and a 64 kbps link to IBERNET, a commercial provider) are protected by a firewall implementation based on Firewall-1 software.

2. Connection to ECMWF

ECMWF dissemination is received in GRIB code on a Sun ULTRA-250 server via ftp, where is decoded and plotted as well as sent to the C94A to be used as boundary conditions for the HIRLAM model, to the data management system to be archived and to the McIdas grid server where is again decoded and ingested. The other Sun ULTRA-250 server is used as backup.

A Sun ULTRA 1 workstation is used as gateway for the non-operational usage; interactive work, file transfers -ftp and eccopy-, etc. Version 2.0.1 of Ecaccess gateway was recently installed, and at the moment, only a few tests have been made, specially regarding file transfers.

Submission of operational jobs to ECMWF computers is done, for the most part, in ecgate1 through SMS. Ecbatch software runs on all the users workstations and on the two Sun ULTRA-250 servers for operational jobs. Use of MARS remote client software, with the Sun Enterprise-10 acting as cache server, is allowed to any user Workstation within the LAN.

Although access via the Internet is possible now without authentication in our firewall through Ecaccess gateway, the majority of our users still access ECMWF through RMDCN.

3. Experience using ECMWF computers

10 INM projects and 1 special project are registered for 2003, with 58 users in total, of which 26 are active users -logged on ecgate1 several times this year-. The majority of them access the Centre's computing resources to extract data from MARS for both operational and research work. We also run Metview in batch mode on ecgate1 to produce derived products from EPS.

SPAIN SPAIN

Concerning the usage of the Centre's High Performance Computing Facilities, during past year nineteen users accessed regularly the FUJITSU's VPPs. They basically worked in the following areas:

- HIRLAM model runs using the reference system
- Large-Eddy Simulations (LES) of the stable atmospheric boundary layer (SBL).
- Trajectory computations.
- Studies on variability
- Statistical downscaling

On 2002 our allocation units were actually exhausted. 58% of the spent resources were used by HIRLAM -including preparation and testing of benchmark versions for the replacement of the CRAY C94A-, a 30% by LES experiments, another 9% by trajectory computation and the remaining 3% by the statistical downscaling and the variability studies. This year the use of our allocation dropped dramatically -we have so far used less than a 6% compared with a 45% on May 2002

In response to a request for feedback, less than half of active users replied. Here is a summary of their comments:

In general, they are very satisfied with ECMWF computer services and the assistance and help they get from User Support is very much appreciated.

Some users would like more permanent disk space ($HOME).

Users of HIRLAM reference system are very much satisfied with job turnaround on the new IBM, in particular those who used before the VPP 700. MARS, however seems to them not to cope well with the new IBM capability, and currently it is seen -from these users point of view- as a bottleneck.

Mini SMS for HIRLAM reference system is a distributed application and therefore it is sometime difficult to find out where problems occurred.

ECFS/MARS suffer from frequent system sessions.

Move to IBM batch environment was smooth and easy for the half a dozen of users that tried so far the new HPCF, but they would like to have a command like vppqsub to remotely submit jobs from ecgate1.

Users of HPCF who run programs that are either difficult or not worth to parallelize find not easy to get a good performance on the IBM compared with what they get on the VPP 5000 without much care on optimization.

4. Future plans

Our use of the Centre's High Performance Facility is expected to decrease once the CRAY X1 is available after next summer -most users will also be involved at migration effort -. In fact, as I mention above, this drop already happened, although main reason is that very heavily CPU consuming LES experiments are now part of special project SPESTURB and preparation of benchmark codes is no longer needed.

For the rest of the use of ECMWF resources no changes are envisaged in the immediate future.

SPAIN **SPAIN**

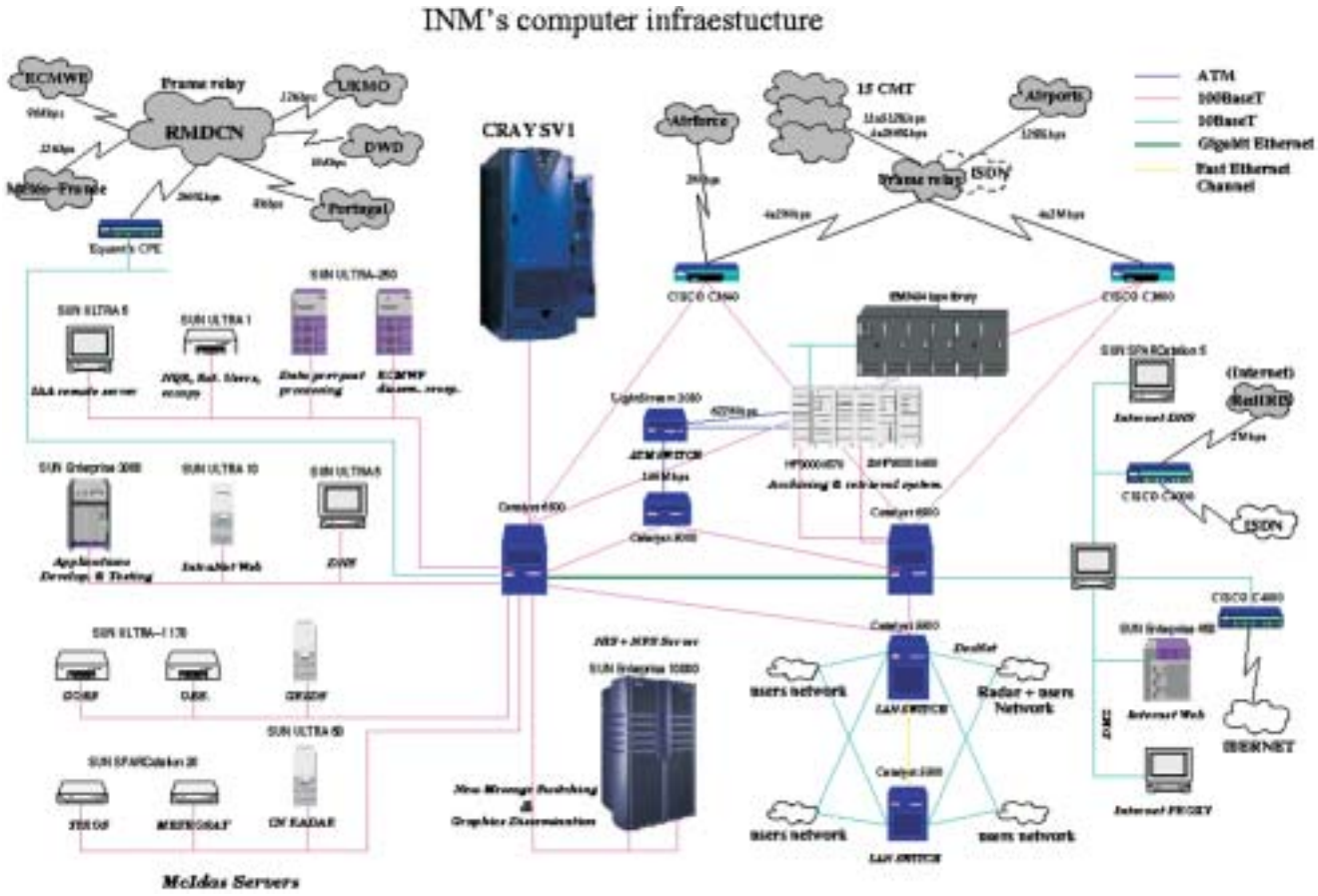


Fig. 1

SWEDEN SWEDEN

Rafael Urrutia – SMHI

The computer equipment at SMHI

Our Internet platform is based on Sparc/Solaris and Compaq Intel/Windows 2000 and Linux. We are using software like Apache, Jakarta-Tomcat, Ldap, ProFTPd, MS IIS, ht://Dig. Our Firewalls are today based on Cisco Pix

The main computer system is based on Unix environment at SMHI. You find here Solaris, Tru64 and Linux as a growing server platform. Important systems are also based on OpenVMS and Windows NT/2000. We have plans to gradually terminate OpenVMS under a period of three years.

We have a Unix Alpha TruCluster working as Fileserver. It provide FS via NFS for Unix and CIFS for Windows (via Compaq's ASU).

The office environment is mainly provided by MS Office for Windows users and Star Office for Linux users. Metaframe and Windows Terminal server is also used both from Unix and Windows clients. The mail services are done by MS Exchange.

The satellite receiving system comprises two AlphaStation 200 4/166 running OpenVMS, one system for Meteosat data and one system for NOAA data.

The system for satellite image processing and for production of various image products is based on two AlphaServers DS10 466MHz, 512 Mb memory with a total of 52Gb disk space in an Available Server environment (TruCluster).

We try to provide the same development environment as we have in our production one. This means Unix, OpenVMS and Windows platforms. The Unix systems are of type Sparc Solaris, Intel Linux, Alpha Tru64 and Alpha OpenVMS.

SMHI has the responsibility for running the Swedish system connecting to the GTS. SMHI is now using a new MSS application, MESSIR-COMM. This is a High Availability system (with software-managed automatic failover) on two Compaq Proliant DL380 servers, with Windows 2000.

Many Internet tools and freeware are used. Example Phyton, Apache, Perl, Gimp, Ghostscript, ghostview, ImageMagick, Jakarta-Tomcat, and more...

The National Supercomputer Centre in Sweden (NSC) is used to run operational models such as the local implementation of the HIRLAM model and HIROMB (High Resolution Operational Model of the Baltic Sea Area), and also climate models.

There are presently five Vax systems supporting weather-radar, one in Norrköping, one in Stockholm, one in Göteborg, one on the island of Gotland and one in Leksand. There are also connections between Norrköping and the military weather radar.

The radar in Sweden, Norway and Finland are connected using concentrator nodes in the three countries, which in turn will be connected to the computers of each radar system. In the future it will be supported by Intel system running Windows and Linux.

Experience using ECMWF computers

In Sweden we have more than 60 ECMWF users. 19 of them are coming from outside the SMHI office. They are from the Swedish University and the Swedish armed forces

The assistance, help and other relations we have with ECMWF are very good.

SWEDEN SWEDEN

What are the users doing

The projects registered at the ECMWF are related to the areas:

- High-resolution limited area model development.
- Aerodynamics and air-pollution modelling.
- Extraction of data for operational usage.
- Extraction of data from era40 and MARS.
- Extraction of data MATCH.
- Hydrodynamic models.
- Research on regional transport.
- Trajectory studies.
- Atmospheric chemistry.
- Daily archive, analysis and forecast.

Question from the users

- Too long time to get ERA-40 data via MARS, too many queue in the MARS request. Can ECMWF here get more computer resource?
- If SMHI wants to get 5-10 TeraByte data from ERA-40 how can we do that?

SWITZERLAND SWITZERLAND

Peter Roth – Meteo Swiss, Zurich

MeteoSwiss

Actual Computer Environment

Since the last meeting, much progross were made in the project 'Server, Storage & Client' (SSC). The actual state is:

- at Zurich
  - The main computer is a SunFire 6800 with 4 domains. One domain is for the fileserver controlling the NAS (Network Attached Storage) and the backup/recovery tapes. A second domain is for the databaseserver of our climatological database (Oracle). The other 2 domains - one for the forecasters, one for the other users - are for producing meteorological and climatological products. The system is actually not operational because the migration from the old systems has not been completed yet.
  - The servers for the message handling system (MHS) had been replaced by SunFire V880 maschines and the radar servers had been integrated in theses maschines. Others systems will follow (e.g. the data aquisition system).
  - The servers (IBM 345 Series) for the 'MS-World' (office applications, mails) are installed and operational, running an AD server, an Exchange server and a MetaFrame server. Users with a workstation (SUN / Solaris 8) use Outlook from the MetaFrame server while PC users (DELL, W2k) have a local installation of the MS software. For special 'MS'-tasks in the Unix-environment, we still use PC boards installed in the workstations.
  - No changes in the DMZ.
  - The system should be fully operational by the end of June 2003.

The drawing below gives further details.

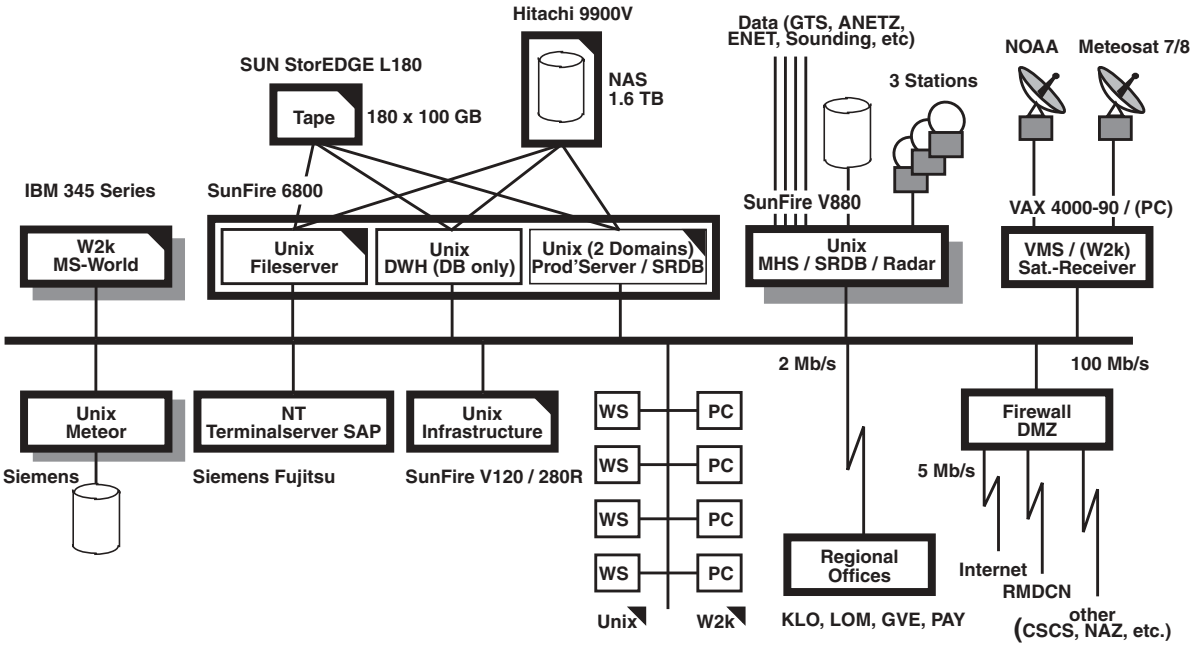


Fig 1: Configuration at Regional Offices (smaller Equipment)

SWITZERLAND SWITZERLAND

- at the Regional Offices (Zurich-Airport, Geneva, Locarno, Payerne)
  - The configuration is the same as at Zurich, but with smaller maschines (SunFire V880 instead of SunFire 6800) and less storage capacity.
  - Actually, only the Sun environment at Zurich-Airport is operatinal.
  - All systems at all offices should be fully operational by the end of this year (that also is the end of the project).

Network

- LAN: 100 Mb/s
- WAN: 2 Mb/s (backup-lines are part of SSC / not yet available)
- ETH/CSCS: 10 Mb/s
- Internet: 5 Mb/s
- RMDCN:
  - ECMWF: 96 kb/s
  - DWD: 128 kb/s
  - MeteoFrance: 16 kb/s

Experience using ECMWF Computer Services

MeteoSwiss makes use of:

- the dissemination system (different data sets)
- MARS (incl. ERA40 data)
- several services form 'ecaccess' (the migration from 'eccopy' to 'ectrans' should be completed in a few months)
- MAGICS applications runninng at ECMWF
- using MetView
- SLEPS-calculations on HPCF (COSMO-LEPS project)

The users in Switzerland are very satisfied of the user support and the services from ECMWF.

UNITED KINGDOM UNITED KINGDOM

Paul Burton – Met Office, Bracknell

1. Main Computer Systems

The current computer infrastructure at the Met Office in Bracknell remains much as described in the previous meetings. However, by the end of 2003, the Met Office will have completed its relocation to Exeter. Once relocation is completed, the main systems and their interconnections will be as illustrated in Figure 1.

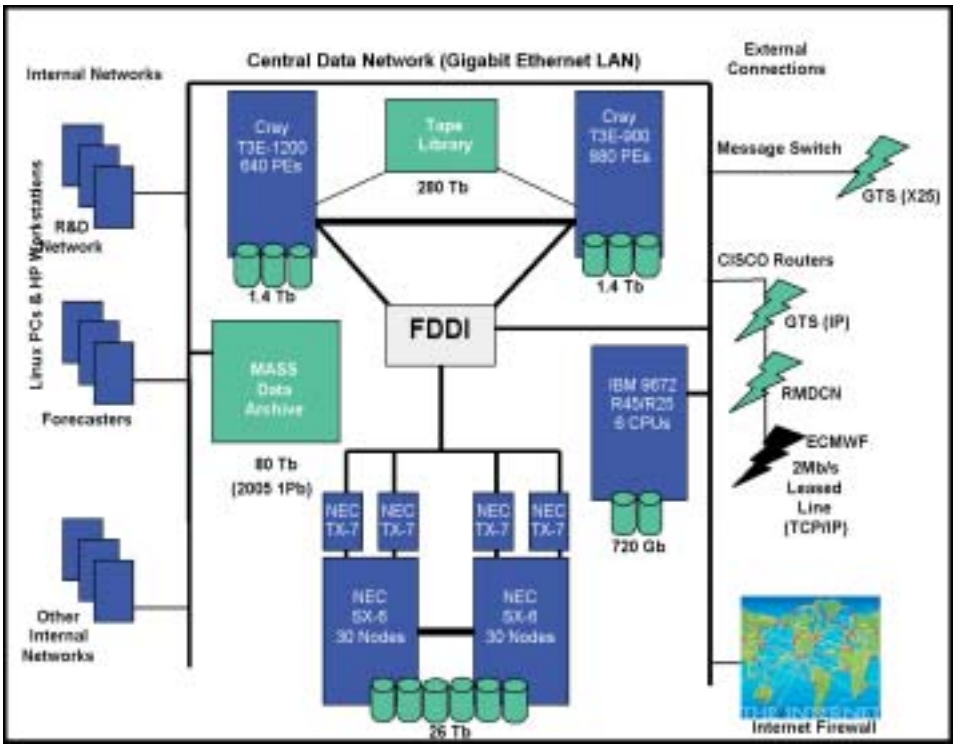


Fig. 1: Main Met Office Computer Systems

At present, the main supercomputer system consists of an 880 processor T3E-900 (installed 1997) and a 640 processor T3E-1200 (installed 1999). About 60% of the CPU capacity of these machines is dedicated to climate prediction, and the remaining 40% to Numerical Weather Prediction.

The IBM 9672 acts as the database server for observations and forecast products and controls the flow of data between the T3Es and the GTS message switch. The R45 module has four CPUs and is split into MVS and UNIX partitions. The R25 module has two CPUs and is dedicated to the pre and post-processing of operational forecast information.

Our data handling system, known as MASS, comprises a SUN E6500 multi processor computer running StorHouse proprietary software, and a StorageTek Powederhorn tape silo with up to 32 StorageTek 9840 tape drives. MASS uses relational database technologies to provide access to model and observational data at the atomic (field, observation group) level. It can also act as the central repository for user files and other data. The amount of data being managed by MASS is expected to grow from the initially installed 80Tb to around 1Pb by 2005.

We are now taking delivery of our new supercomputer. The new system being installed in Exeter consists of 30 NEC SX-6 nodes accessed through a total of four NEC TX-7 front ends. Each SX-6 node has 8 CPUs. The system will be divided between the two new computer halls, with 15 SX-6 nodes and two TX-7 front ends in each hall.

2. Connections to ECMWF

All of our prime intra-regional links are connected via RMDCN, however we continue to use the existing 2Mb/s leased line to ECMWF, which is utilised for the exchange of all data and products with the Centre as well as for user access to the Centre's computing facilities.

UNITED KINGDOM UNITED KINGDOM

Data is sent to and from an Alpha Server 1200 on the Central Data Network (CDN) at Bracknell, with a second Alpha Server 1200 acting as a backup. The connection supports the receipt of observational data and the dissemination of ECMWF products. Along with Offenbach, the UK provides one of the connections between ECMWF and the GTS.

Telnet, FTP, eccopy and ecqsub are available from the HP workstation network and from both T3Es. The ECaccess gateway service has now been installed on a new HP server in Bracknell as well as on a staging post server in Exeter. Initial tests of the new access portal have been successful and initial problems with the speed of the X access have been overcome thanks to much gratefully received help from ECMWF experts. Special thanks are also due to John Greenaway and Dominique Lucas of ECMWF User Support who visited the Met Office on 7 May 2003 to provide some well-received on-site user training on the use of ECaccess.

3. Changes to the system

By the end of 2003 the Met Office will have completed its relocation from its premises in Bracknell to a new purpose built facility in Exeter. This major move is also being used as an opportunity to upgrade and replace many parts of the IT infrastructure we currently use.

The most important change will be the replacement of our supercomputer systems. Although we have begun to install our new NEC SX-6 supercomputer system in the new computer halls, both T3Es will be moved to Exeter to allow the smooth transition of our business. The move of the T3Es to Exeter will begin when the Operational Forecast is switched to T3EA on 15/06/03. T3EB will be powered off, dismantled and moved to Exeter on 17/06/03. It is anticipated that T3EB will be back online by 24/06/03 and, after some parallel operational testing, will switch to operational running on 20/07/03.

The second stage of the move is planned for 02/09/03 when T3EA will be powered off, dismantled and shipped to Exeter. By 06/10/03 both T3Es will be installed and running in operational mode in Exeter.

The first phase of the NEC SX-6 supercomputer system will be accepted in early 2004, and will provide the Met Office with a 6 times increase over our current performance capability. For a maximum period of 3 months after final acceptance, the T3Es and SX-6 will run in parallel. This will allow the Met Office to choose the most appropriate time to switch systems.

The second phase, to be accepted by spring 2005, will install additional SX-6 hardware to raise the total capability to 12.5 times current.

We will also be replacing the desktop systems. The current scientific desktop, consisting of HP servers and X terminals will be replaced with flat-screen PC hardware running the Linux operating system. Some of the more powerful HP servers will be kept as a resource for running more computationally intensive work, removing its impact from the desktop systems.

The Central Data Network will be upgraded to Gigabit Ethernet, with local connections to this generally running at 100Mb/s.

The Tandem Message Switch is being replaced by a Unix based system running proprietary message switching software. This new system will be sized to cope with an expected doubling of traffic over the next 3-5 years.

4. Use of ECMWF Computer Systems

As ever, there continues to be a relentless rise in the number of registered users - numbers have increased from 114 users at the last meeting, to 129 at the last count - and a further increase is anticipated over the next months with the availability of ERA-40 data. The majority of these users have fairly simple data retrieval requirements. Some of these users really only require a one-off access to ECMWF data. Those with relatively simple data access requirements, continue to report a high level of satisfaction with the documentation and ease of use of the system.

There are a small number of users with more complex requirements that involve some post processing of retrieved data. To date, this post processing has, in general, been undertaken on ECgate1 but there now is a move to perform some of this on the IBM. To aid this transition, it has been suggested that it would be useful for Metview (or some of its functionality, e.g., the facility for sorting large field sets) to be made available on the IBM.

We continue to use Metview on our HP workstation system, and users benefit from a transparent access to ECMWF data via the MARS client software. The current version of Metview used on the HP workstations is version 3, the port having been undertaken with much gratefully received assistance from ECMWF experts. We look forward to our new Linux desktop, which as well as making the installation of new Metview versions much smoother, offers a

far better performance and user experience than our current HP hardware. It is hoped that the move to Linux will also allow us to include the ECaccess MARS module.

In preparation for the change in the ECMWF supercomputer system, work was undertaken during the past year to port the Unified Model system to the IBM hardware, running in MPP mode over 8 processors (IBM single node). The port was achieved with relative ease although some difficulty was encountered in extracting useful information for optimisation purposes from the available profiling tools.

5. Current Use and Projects Run at ECMWF

Our use of HPCF units continues to grow - with most of our time being taken running multi-model ensembles for our long range and seasonal forecasting groups. We run atmosphere only and coupled ocean-atmosphere versions of our Unified Model system. As development of the system continues we will soon require a greater number of runs.

The long running work on the DEMETER project, which was being undertaken on the old Fujitsu hardware, is currently being completed on the IBM. This work will finish in the near future and will cover a total of 42 years.

Last year, we also started to use the ECMWF computing facilities for performing large ensemble predictions of long term anthropogenic climate change in order to quantify the uncertainty in the predictions.

Our Ensemble Research team makes considerable use of EPS data. The PREVIN (PREdictability and VIsualisation) project provides visualisation of processed EPS data as a forecaster's tool. EPS data is also used to generate warnings to aid forecasters in support of the National Severe Weather Warning Service as part of the First Guess Early Warning project. We are also using ECMWF forecasts together with a two member ensemble run as part of the Poor Man's Ensemble Prediction System (PEPS). Finally, experimental use has been made of EPS data to drive the Nuclear Accident Model (NAME).

We are also running 1DVar retrievals of GPS based radio occultation measurements using ECMWF short range forecasts as a first guess. The forward model is based upon ECMWF's model level structure. Help from Mike Fisher with the background error covariances is gratefully acknowledged. It is hoped that collaboration will be intensified if ECMWF starts its own work on radio occultation data assimilation later this year.

6. Future Plans

Although the DEMETER work will soon come to an end, we will continue to run our long range and seasonal forecasts on the IBM, with the seasonal forecast using 40 member ensemble runs supported by hindcasts. Some further experimentation and tuning of the ocean analysis, together with an investigation into the physics sensitivity of a simple ocean model will be undertaken with the aim of improving the seasonal forecast.

Our large ensemble predictions of climate change will also continue for the foreseeable future. From September 2003 onwards, there are also plans to use time on the ECMWF computer to obtain predictions of the climate on a decadal time scale using coupled models initialised from observations.

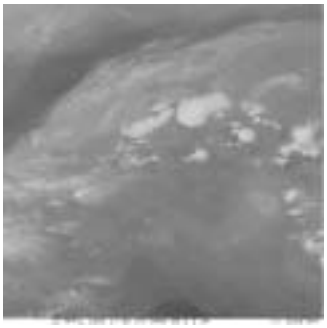
In addition to the continuing Ensemble Research work, there are plans to use the recent port of the Unified Model to the IBM for ensemble experiments. An investigation into possible methods of generating model perturbations for use in limited area ensembles will also be undertaken and an experimental use of EPS perturbations for generating initial perturbations is planned.

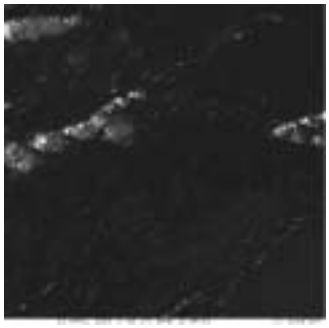
EUMETSAT

EUMETSAT

Ken Holmlund – Eumetsat

MSG loops





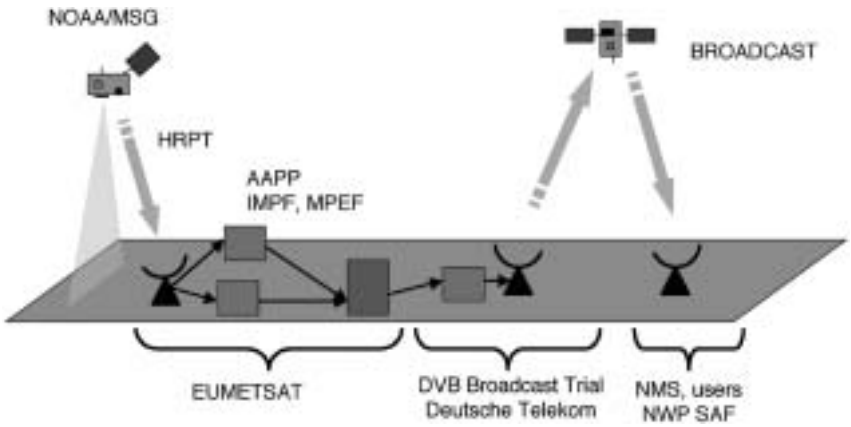
Introduction

- MSG status
- Storage update
- Relocation
- Reprocessing
- Network
- ECMWF data

Meteosat Second Generation (MSG)

- Launched August 2003
- Anomalies
  - "Wobble" (Solved)
  - RTU (Currently on redundant unit)
  - SSPA ==> Impact on dissemination
- Alternative Dissemination Mission (ADM)
  - MSG-1 to be re-located to 3°W
  - Dissemination using DVB
  - Phase-1 Europe
  - Phase-2 Full coverage
  - Trial Dissemination On-going

MSG ADM and EARS Dissemination Chain



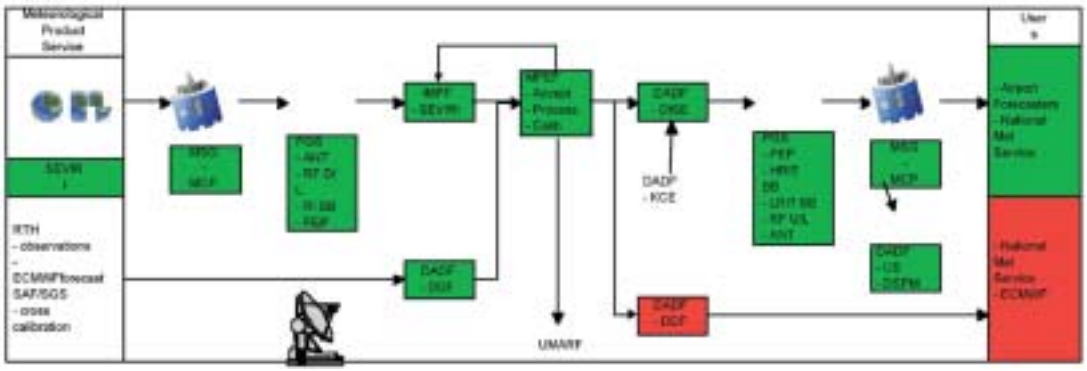
EUMETSAT

EUMETSAT

Meteosat Second Generation (MSG)

- Schedule
  - Pre-operational dissemination September 2003
  - End of commissioning December 2003
  - 1 year parallel operations with Meteosat-7
  - MSG-2 launch 2005

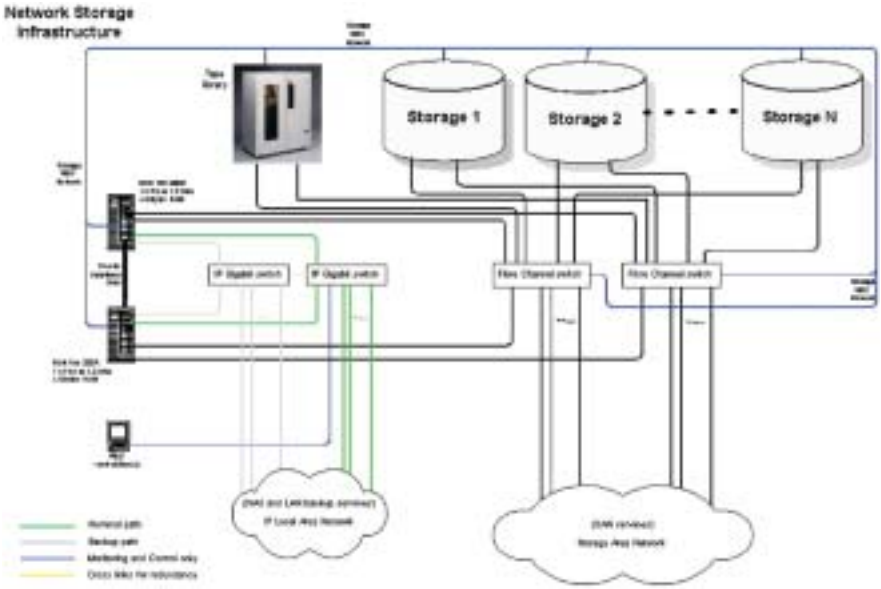
MSG Meteorological Products Service Chain Flow



Data-storage upgrade

- SAN-based solution
- Initially ~7 TB
- MSG I + Meteosat image archive inititially 3 TB
  - Current EPS support activities (3 TB)
  - Other, 1 TB
- Further storage next year
  - 10 Tb
  - Full Meteosat archive (rectified, compressed ca 7 Tb)
  - EPS support activities

Data-storage Infrastructure



EUMETSAT EUMETSAT

Relocation of MTP GS

- Enable installation for EPS CGS
- Ensure operational service levels
- Compress GS equipment
- Clean-up infrastructure
- Current deadline 9/2003

Affected sub-systems

- CF
- MDRS
  - GTS, DCP upgrades
- MPEF
  - h/w upgrades, new rapid-scan MPEF
- MARF
  - new connectivity
- MSG
  - new monitoring consoles
- Additional
  - Routers, cables, links, etc

Re-processing Meteorological Product Extraction Facility

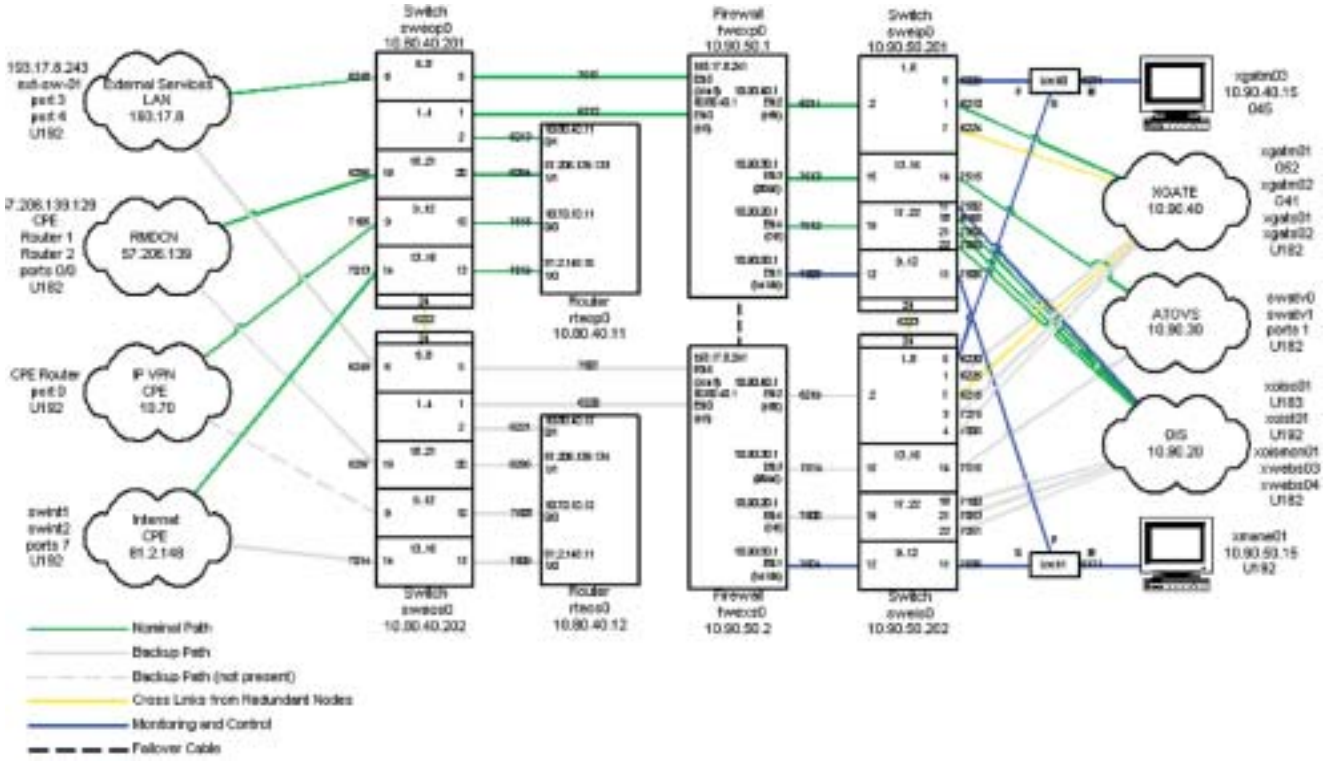
- ERA-40 (ECMWF Re-Analysis Project)
- SSCC (SEVIRI Solar Channel Calibration)
- MSA (Meteosat Surface Albedo)
- MPE (Multi-sensor Precipitation Estimate

RMPEF

- Initially a simple project to support ERA-40
  - Based on redundant equipment and services
  - High performance access to image data
  - Currently 7 days/day ==> 2-3 times faster by the end of the year
- New projects ==> New requirement
  - h/w upgrades
  - Development environment
  - Configuration control
- Formalisation of activities

EUMETSAT **EUMETSAT**

EFTS, RMDCN, ATOVS and OIS network overview



RMDCN Status

- MTP MPEF implementation
  - Operational
- MSG MPEF implementation
  - Delayed due to other activities (DADF upgrades)
  - End of 2003/Early 2004
- EPS implementation
  - Part of CGS delivery
  - CGS integration starts III/2003

ECMWF data use

- MTP
  - RMDCN, EPS-data
  - 31 levels, 1.25°, T, u, v, Rh, p
- MSG
  - GTS, EPS-data
  - 31 levels, 1.0°, T, u, v, Rh, p (Oz)
- EPS
  - RMDCN
  - All levels (60, TBC), 0.5°, T, u, v, Rh, p, Oz
- Re-processing
  - ftp
  - 31 levels, 1.25°, T, u, v, Rh, p
  - SSM/I
- Scientific studies
  - ECaccess
  - variable resolution and data

ANNEX 1

ANNEX 1

Fifteenth Meeting of Computing Representatives

ECMWF, Shinfield Park, Reading, U.K., 19-20 May 2003

Participants

Austria	Gunter Wihl
Belgium	Roger Vanlierde
Croatia	Vladimir Malovic
Denmark	Niels Olsen
Finland	Kari Niemelä
France	Marion Pithon
Germany	Elisabeth Krenzien
Greece	Nikoloas Kamperakis
Hungary	László Tölgyesi
Iceland	Halla Bjorg Baldursdottir
Ireland	Paul Halton
Italy	Giuseppe Tarantino
Netherlands	Hans De Vries
Norway	Rebecca Rudsar
Spain	Eduardo Monreal
Sweden	Rafael Urrutia
Switzerland	Peter Roth
United Kingdom	Paul Dando
	Kelvyn Robertson
Eumetsat	Ken Holmlund
ECMWF:	Tony Bakker
	Sylvia Baylis
	Petra Berendsen
	Jens Daabeck
	Matteo Dell'Acqua
	Richard Fisker
	Helene Garçon
	Laurent Gougeon
	John Greenaway
	John Hennessy
	Norbert Kreitz
	Dominique Lucas
	Carsten Maass
	Umberto Modigliani
	Pam Prior
	Baudouin Raoult
	Deborah Salmond
	Neil Storer
	Peter Towers
	Walter Zwieflhofer

ANNEX 2 ANNEX 2

Programme

Monday, 19 May

09.30 *Coffee*

10.00 Welcome

ECMWF's computer status and plans . . . . . .W. Zwieflhofer

Member States' and Co-operating States' presentations

Each representative is asked to speak for a maximum of 10 minutes, outlining their Member State's or Co-operating State's involvement (actual or planned) in the computer service at ECMWF. This should include:

- diagram of own computer equipment, and of connection to ECMWF
- projects run at ECMWF (if applicable)
- experience using ECMWF computers/services, including suggestions

and queries regarding the present service

- plans (involving ECMWF usage over next couple of years)

12.30 *Lunch**

14.00 Experiences with the new IBM HPCF: Hardware and software . . . . . .N. Storer

Migration to the new IBM HPCF: Performance improvements . . . . . .D. Salmond

Migration to the new IBM HPCF: Communication optimisation . . . . . .P. Towers

15.30 *Coffee*

16.00 MARS and dissemination update . . . . . .J. Hennessy/B. Raoult

Graphics update . . . . . .J. Daabeck

Ecaccess status and plans . . . . . .L. Gougeon

RMDCN update . . . . . .T. Bakker

17.40 *Cocktails, followed by an informal dinner*

* During this lunchbreak machine room tours will be arranged if required. Interested participants should contact their User Support Contact Point

Tuesday, 20 May

09.00 Member States and Co-operating States presentations (continued)

10.30 *Coffee*

11.00 User services: Status and plans . . . . . .U. Modigliani

Discussion

12.30 **End of meeting**