

Report on the eighteenth  
meeting of Computing Repre-  
sentatives  
8 - 9 June 2006

P. Prior (Compiler)

Operations Department

November 2006

*This paper has not been published and should be regarded as an Internal Report from ECMWF.*

*Permission to quote from it should be obtained from the ECMWF.*



**Series: Technical Memoranda**

A full list of ECMWF Publications can be found on our web site under:

<http://www.ecmwf.int/publications/>

Contact: [library@ecmwf.int](mailto:library@ecmwf.int)

© **Copyright 2006**

European Centre for Medium Range Weather Forecasts  
Shinfield Park, Reading, RG2 9AX, England

Literary and scientific copyrights belong to ECMWF and are reserved in all countries. This publication is not to be reprinted or translated in whole or in part without the written permission of the Director. Appropriate non-commercial use will normally be granted under the condition that reference is made to ECMWF.

The information within this publication is given in good faith and considered to be true, but ECMWF accepts no liability for error, omission and for loss or damage arising from its use.

## Contents

Preface .....	2
<b>Part I: ECMWF Staff contributions</b>	
ECMWF Computing Service: Status and plans – <i>Isabella Weger</i> .....	4
HPCF & DHS update – <i>Neil Storer</i> .....	13
Early experience on the IBM Phase 4 system – <i>Deborah Salmond</i> .....	18
Overview of ECMWF electrical and mechanical infrastructure – <i>Sylvia Baylis</i> .....	22
RMDCN status and migration plans – <i>Tony Bakker</i> .....	25
ECaccess status and plans – <i>Laurent Gougeon</i> .....	31
New MS job submission via SMS – <i>Dominique Lucas</i> .....	35
User registration update – <i>Paul Dando</i> .....	41
Survey of external users: status of following actions – <i>Umberto Modigliani</i> .....	45
ECMWF web search revisited – <i>Daniel Varela</i> .....	47
Graphics update – <i>Jens Daabeck</i> .....	50
Technical overview of Magics++ – <i>Stephan Siemen</i> .....	53
Introduction of the new GRIB API – <i>Enrico Fucile</i> .....	58
<b>Part II: Member State and Co-operating State contributions</b>	
Austria .....	64
Belgium .....	68
Croatia .....	70
Denmark .....	71
Finland .....	76
France .....	79
Germany .....	84
Hungary .....	89
Iceland .....	93
Ireland .....	95
Italy .....	102
Netherlands .....	110
Norway .....	115
Romania .....	120
Serbia .....	124
Spain .....	131
Sweden .....	135
United Kingdom .....	139
CTBTO .....	142
EUMETSAT .....	144
Annex 1 – Participants .....	150
Annex 2 – Programme .....	151

## **Preface**

The eighteenth meeting of Computing Representatives took place on 8–9 June 2006 at ECMWF. Nineteen Member States and Co-operating States, plus the CTBTO and EUMETSAT, were represented. The list of attendees is given in Annex 1.

The Head of the Computer Division (Isabella Weger) opened the meeting and welcomed representatives. She 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. 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 and general discussions**

## **ECMWF Computing Service: Status and Plans – Isabella Weger, Head of Computer Division**

### **Major activities over the past 12 months**

#### *High Performance Computing Facility*

The results of the Multi-cluster GPFS testing are very promising. A pilot study has been completed and the testing of MC-GPFS over a WAN is in progress. There was a spate of MCM failures which were eventually solved by slightly increasing the voltage. There have been more improvements in Job Scheduling, achieved by combined use of Advance Reservation and Floating resources and better understanding of the backfill-scheduling mechanism. Preparations for the installation of the HPCF Phase 4 are complete.

A market survey of HPC vendors has been carried out in readiness for the forthcoming HPCF ITT.

#### *Data Handling System*

The 2005 Phase of the DHS has been installed.

An evaluation of HPSS 6.2 is underway. (Migration is planned in the next few months.)

A market survey of Automated Tape Library and Tape Drive technology was carried out in 2005. ITT 189 to obtain a replacement for the ATL in the DRS building is underway.

#### *Linux Cluster*

The Linux cluster is in full production as a general purpose server.

#### *LAN*

The replacement of the general purpose LAN is complete.

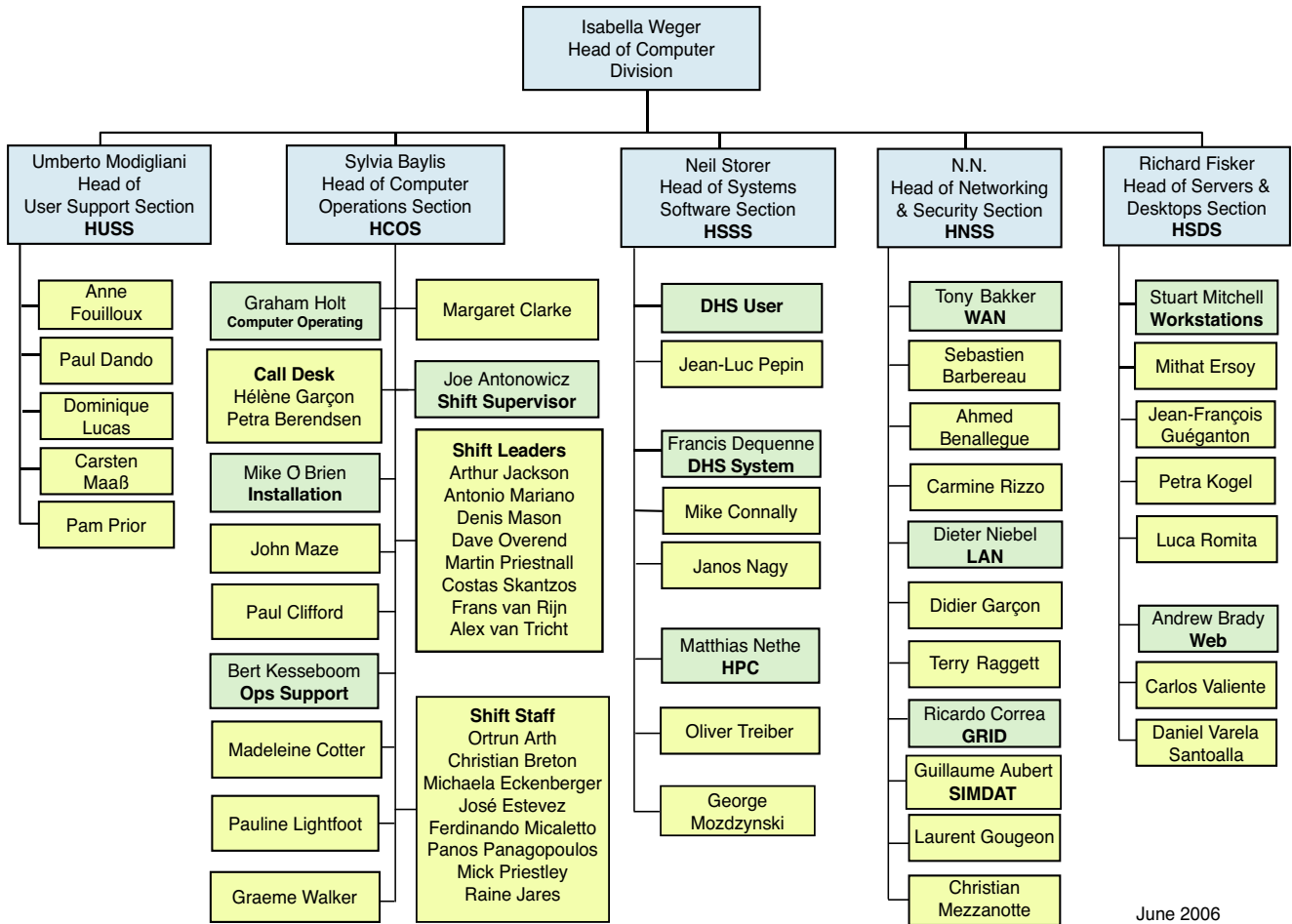
#### *RMDCN*

The contract has been amended, reflecting the new SLA for IP/VPN and detailing the migration to MPLS technology.

#### *Infrastructure*

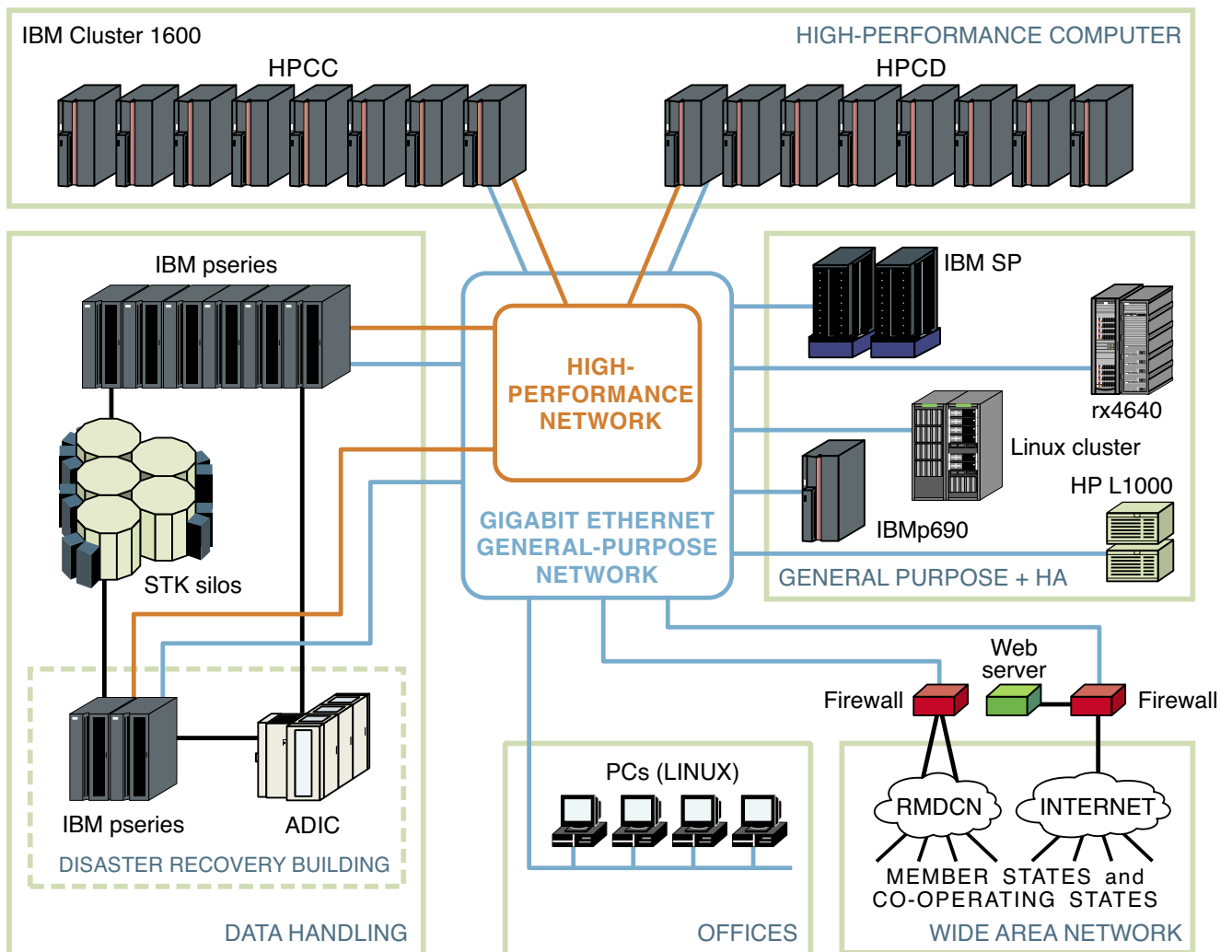
Various upgrades to the mechanical and electrical infrastructure have been made. The fourth Uninterruptible Power Supply machine has been installed and integrated into the UPS system. An additional chiller has been installed to provide more chilled water capacity. The Computer Building extension has been completed and the fire detection and suppression systems have been extended into the new computer hall.

### Computer Division Organigram



June 2006

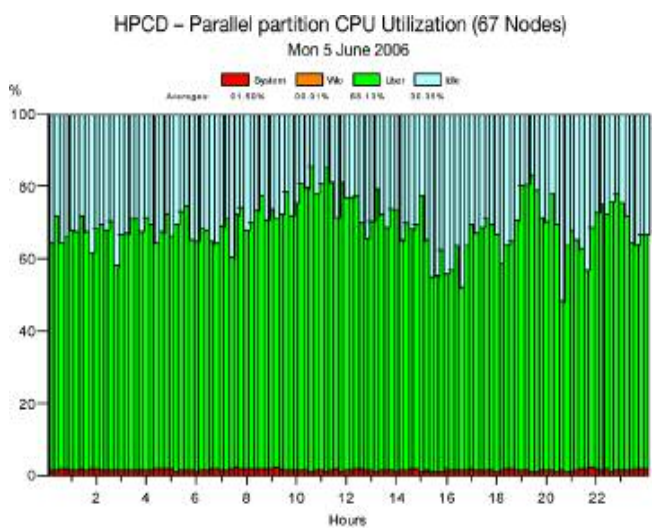
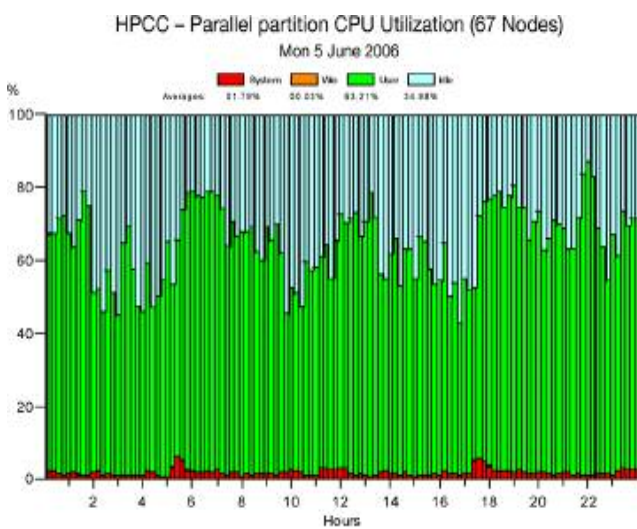
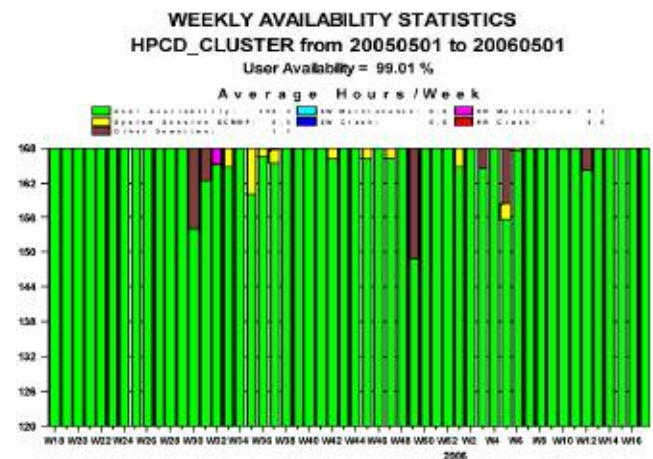
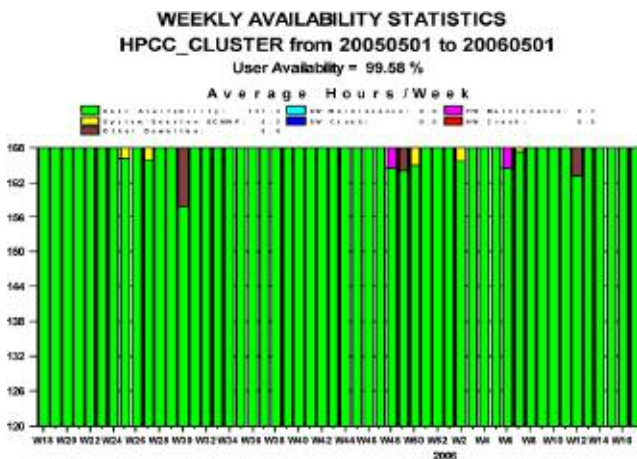
ECMWF Computer Environment



High Performance Computing Facility

- The IBM Phase 3 continues to deliver a very good service. User availability was generally very high. Several service interruptions were caused by unscheduled power-downs, which were related to UPS issues.
- Member States' usage of supercomputing resources has increased significantly, particularly since the end of 2005. In 2005 Member States used more than 50% of their combined HPCF allocation. Several Member States (Denmark, Finland, Italy, Netherlands and Sweden) used more than 70% of their allocation and some of them more than 90%. Resource usage has remained quite high in 2006, with Member States using more than 80% of their allocation of HPCF Phase 3 resources (daily average).





**Framework for MS time-critical applications**

- Current MS activities

The NORLAMEPS system has been implemented as “option 3” and has been running at ECMWF since February 2005. The COSMO-LEPS suite has been implemented as “option 2” and was declared “time-critical” in November 2005. Both the NORLAMEPS and the COSMO-LEPS suites were updated when the high resolution EPS (TL399) was implemented at ECMWF.

- New MS activities

A SMS suite to run the IFS-EuroHRM-EuroLM activity as “option 2” is running daily and is being finalised.

A request to support the UKMO THORPEX ensemble forecast suite as “option 2” has just been received: the suite is already running on a daily basis.

- A solution to handle option 1 “Simple job submission monitored by the Centre” has been implemented.

**HPC system – Phase 4 (2006 - Q1/2009)**

Phase 4 is the last phase of the IBM service contract. The first cluster is already installed and will begin formal acceptance soon. The second cluster will start arriving in August.

There are two identical IBM clusters 1600 - p5-575 servers Each cluster has ~ 140 to150 nodes, all connected by the pSeries High Performance Switch. Each node has:

- 16 Power5+ @ 1.9 GHz SMT processors (8 dual-core chips)
- 32 GB physical memory per node (a few will have 128 GB)
- 10 nodes per cluster are dedicated to I/O and networking
- 50 TB of FC disk storage per cluster.

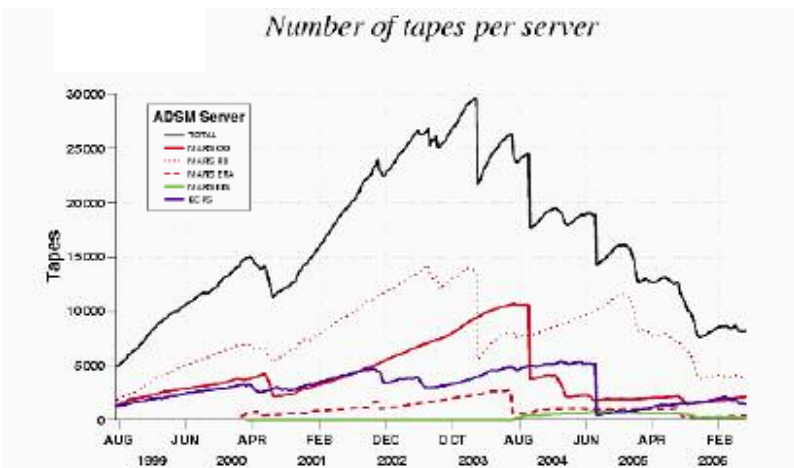
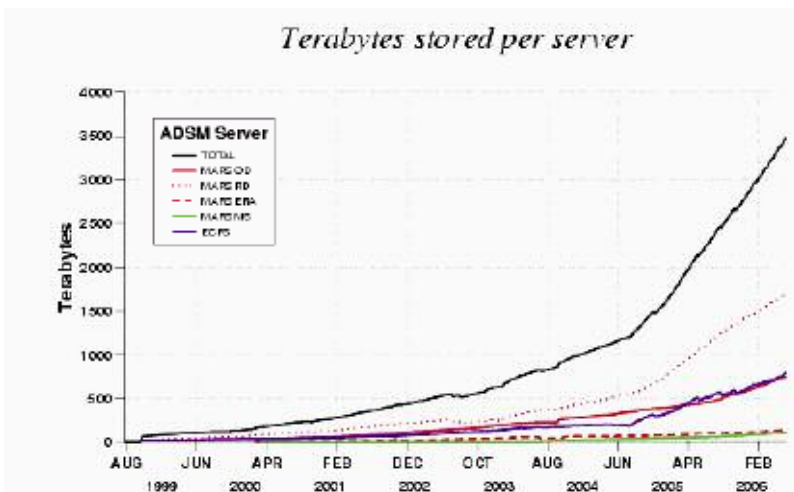
A MC-GPFS (Multi-cluster global parallel file system) will enable each cluster to have concurrent access to data.

**Data Handling System**

The HPSS-based system continues to perform very well. The 2005 Phase of the DHS has been installed. New tape technology (IBM 3595-J2) and new SATA disk technology has been evaluated and installed. In mid-April this year the volume of data in the DHS reached the 4 PByte threshold. The system now contains approx. 3.25 PB of primary data in the DHS (excluding backup copies).

An ITT for the replacement of the Automated Tape Library in the Disaster Recovery Building, which has reached its end-of-life, has been issued.

**DHS archived data (excluding backup copies)**



### Servers and Desktops

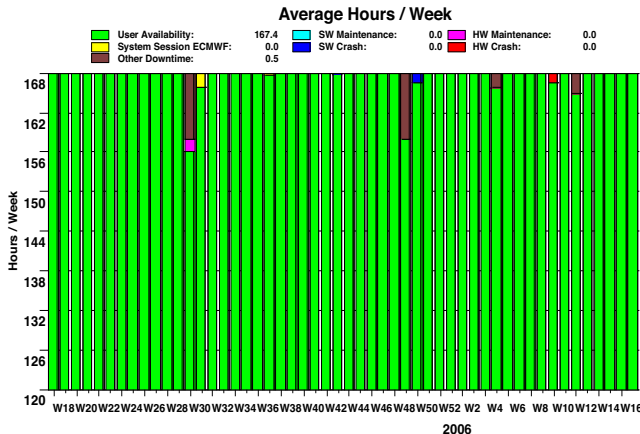
The internal IBM AIX servers leda & metis (which are Nighthawk2 nodes) are being decommissioned.

The Linux Cluster is now in full production as a general purpose server. This took longer than originally anticipated, mainly because of incompatibility problems with the Korn Shell.

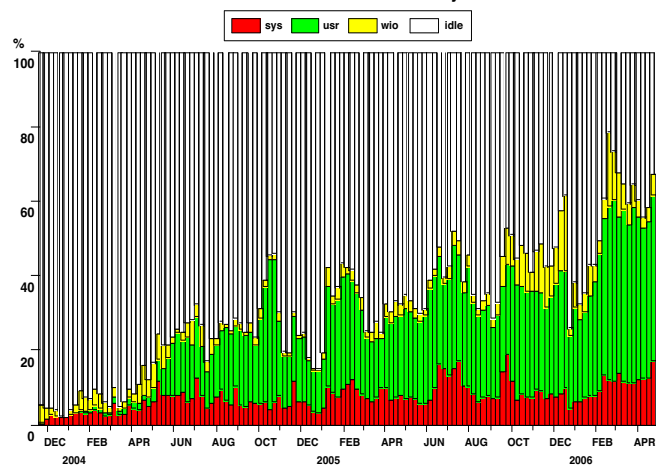
The Itanium-based Highly Available System for data acquisition, pre-processing and dissemination has proved very stable and copes very well with the workload.

ecgate remains very stable and is providing an availability of 99.6%. ecgate is nearing its capacity - a substantial increase in workload has been noticed over the last 12 months. A successor system is being planned. It will be based on a Linux Cluster, but no details have been decided yet. The current service on ecgate will continue until at least mid-2007.

**WEEKLY AVAILABILITY STATISTICS**  
**ECGATE from 20050501 to 20060501**  
 User Availability = 99.63%



**ECGATE weekly CPU Utilization**  
 From 24 November 2003 to 1 May 2006



### Web Service

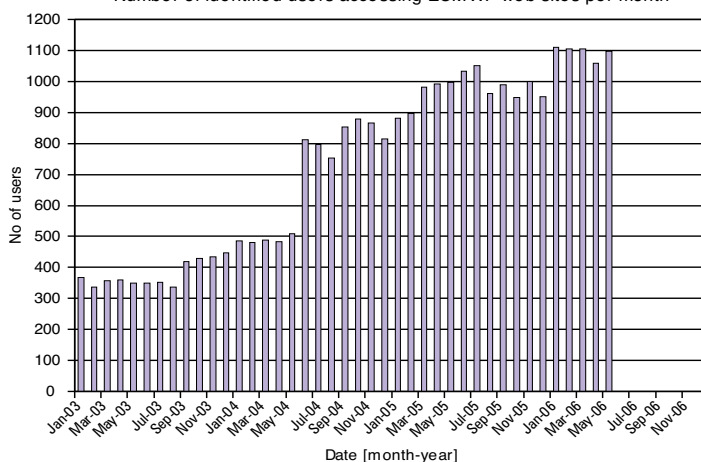
The ECMWF web servers continue to provide a stable and reliable service. The response time of the web site remains excellent. At the end of April this year bug fixes to the web login ticket expiry closed some access holes in the web site.

The web search is being improved.

The growth in use of the ECMWF web site continues. Not only are existing users increasing their usage of the site but there is also a growth in new users. Overall there was a 50% increase in access by registered users between 2004 and 2005. The early 2006 figures also suggest further strong growth.

### Web Services – No. of identified users

Number of identified users accessing ECMWF web sites per month

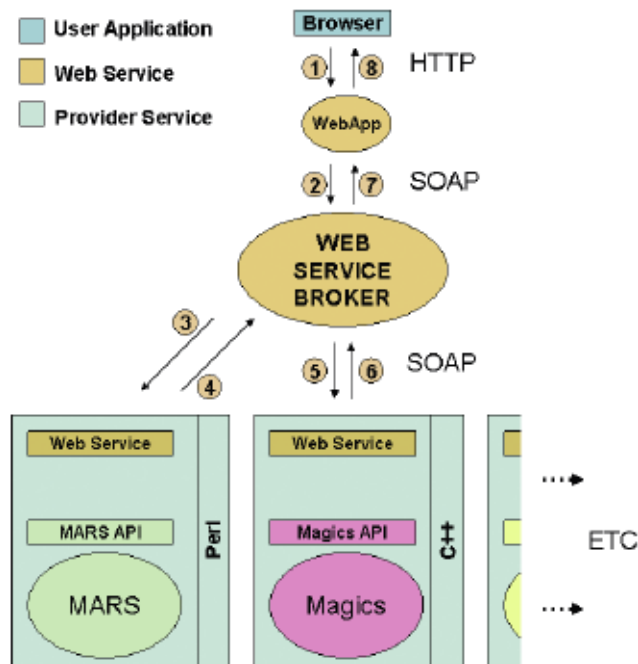


**Web Service – Statistics**

	2002	2003	2004	2005	2006*
Total number of page accesses by all users (millions of pages/year)	8.09	10.9	13.6	17.6	23.1
Change compared with previous year	+98.0%	+35.0%	+25.2%	+29.4%	+31.2%
Total number of page accesses by identified users (millions of pages/year)	0.95	1.56	2.02	3.05	4.31
Change compared with previous year	+64.2%	+68.7%	+26.5%	+51%	+41%
Average time between page accesses (seconds)	3.9	2.89	2.31	1.79	1.37
Ratio of total users to identified users	8.5	6.8	6.8	5.7	5.4

\* Based on first five months

**Web Services – Plots on demand**



**2005**

A strategic project to develop web service interfaces to main ECMWF tools was started under the "Plots-on-Demand" project. This will expose MARS, ODB, Verification and Magics via a common Web Service API and enable the development of a new application for delivering plots on demand.

**2006**

- Initial development is complete.
- Features of individual web services are very much dependent on the API of underlying provider services.
- Examples of provider APIs have been written in Java, Perl, C++ and Python.
- A simple application was demonstrated at the MetOp workshop in November 2005.
- A production environment is planned for later this year and future web application developments are then planned to use this framework.

**LAN**

Replacement of the General Purpose Network was completed. Following an ITT in Feb. 2005, a bid from Matrix was selected and the Riverstone/Enterasys equipment was replaced by Foundry equipment. Phase 1 of the migration to Foundry Equipment (with Big Iron MG8 in Core) was completed in October 2005. Phase 2 of the migration, to replace MG8 with RX16 and to relocate half of the Core into the second TCC Area, was completed in April/May 2006.

Wireless LAN is now deployed throughout the office building.

HPN was expanded to accommodate Phase 4 of the HPC.

## WAN

44 sites are now connected to the RMDCN. Saudi Arabia's connection to the RMDCN passed acceptance on 4 July 2005. Various upgrades have been carried out, e.g. MeteoFrance, Spain (in progress).

The migration of the transport technology from Frame Relay to MPLS (Multi-Protocol Label Switching) is underway. Supplement No 4 of the RMDCN contract has been signed and site configurations are being finalised.

The VPN Project Phase 2 is finished, though some additional tests are to be concluded.

ECMWF's Internet link was upgraded from 70 Mbps to 100 Mbps (Dec 2005) and to 250 Mbps (March 2006).

## Security

The development of the new Certificate Authority is complete. The various existing ECMWF Certificate Authorities are being migrated to the new system.

A security audit was carried out in spring 2006. The final report will be available later in June. The audit focused on the perimeter systems and on the acquisition and dissemination of data. The audit concluded that:

- ECMWF's internet access is well secured and no major vulnerabilities were identified.
- Some issues regarding documentation and procedures will have to be addressed.

## Infrastructure work

A new 2 MVA UPS machine was installed and integrated with the other three UPS machines. This provides increased UPS capacity and restores N+1 resilience in the UPS system. The second standby generator, which had reached the end of its life, can now be replaced.

The Computer Building extension was handed over to ECMWF. Staff moved into the offices in January. One cluster of HPC Phase 4 has been installed in the extension computer hall. The Gaseous Fire Suppression system has been extended into the Computer Building extension. The water mist fire suppression system has been completed to provide protection to the offices in the Computer Building and Link Building and the new Loading Bay.

## GRID activities – DEISA

ECMWF continues to actively participate in DEISA, thus obtaining a better understanding of GRID middleware, Multi-Cluster GPFS and Multi-Cluster Loadleveler. In the critical area of security, results are still disappointing. The former Phase 3 test cluster, HPCU (6 nodes), has been redeployed as a DEISA host for the purpose of evaluating a file system shared with the partners. This ensures that ECMWF's operational computers can be kept isolated from other DEISA systems.

The DEISA Common Production Environment has been installed on HPCU and HPCD

## GRID activities – SIMDAT

ECMWF continues to co-ordinate the meteorological activity of the SIMDAT project.

A demonstrator of the V-GISC (Virtual Global Information System Centre) has been developed.

SIMDAT collaborations with Japan, China, and Australia are going well. Japan is already connected to the infrastructure publishing real data sets and China is in the process of installing a node.

The development of a new version of the software is nearly complete. It is expected to be finished by the end of June.

## Major ongoing/planned activities

- The HPCF Phase 4 installation and commissioning will be carried out in the next few months.
- HPSS will be upgraded to version 6.2 later this year.
- ECFS continues to be enhanced to provide additional features for users.
- The Automated Tape Library in the DRS building will be replaced later this year.
- The migration of data to new tape technologies will be carried out throughout the year and will be transparent to users.
- Improvement of Web Search

- Replacement of the Intranet web site with a new version utilising a web content management system. Benefits include better organisation of existing information, a simpler editing environment and managed navigation (not under user control).
- Implementation of a scalable architecture for web services, including a framework for chart generation (plots-on-demand).
- Consolidation of Linux/Intel based server systems and upgrades of existing systems requiring increased capacity.
- The ECaccess architecture will be reviewed, to benefit from new technologies and move towards standard protocols (Web Services).
- Co-ordination of the migration of the RMDCN from frame relay technology to MPLS (Multi Protocol Label Switching).
- The resilience of the network infrastructure (both LAN and WAN) will be increased by using two TCC areas.
- ECMWF's token-based Strong Authentication System will be reviewed.
- Options for the replacement of the ECMWF telephone system will be investigated.
- Another chiller is being installed to provide some resilience in the chilled water system during the parallel run of HPC systems this summer. As the new chiller is more efficient than the oldest chillers, one of the old chillers will be switched off, after the HPCF parallel run, and the two new, efficient chillers will be relied upon more heavily.
- A survey of energy use is almost complete. Assuming that we cannot reduce the electrical load from the computing and associated cooling, indications are that a saving of about 1% from the remaining load might be possible.
- A review of the electrical infrastructure will be conducted to determine how best to develop the electrical infrastructure in the coming years.

R.Urrutia asked whether the new Certificate Authority which was being implemented would be open to Member State users. *After the meeting R. Correa replied to R. Urrutia in detail that, although the ECMWF Certificate Authority could be used for some purposes, it was not to be recommended for the purpose for which it was required for SHMI.*

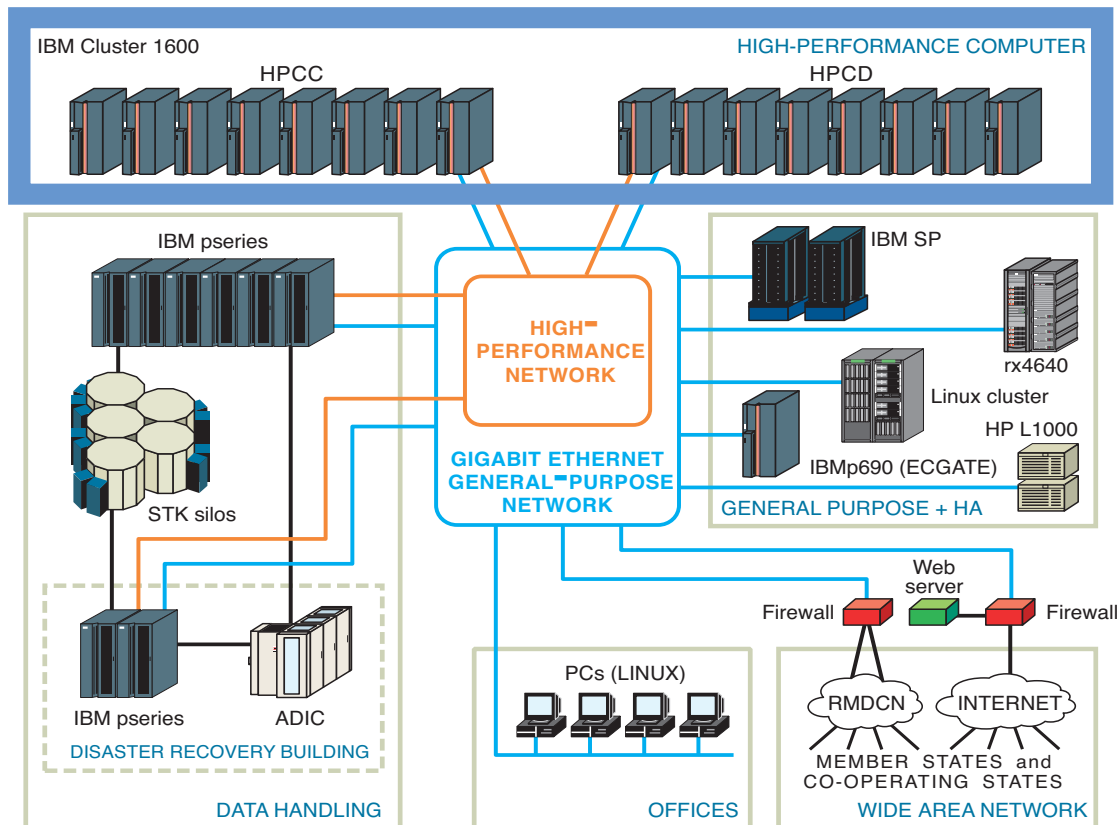
M. Pithon asked the purpose of ECMWF's intended testing of Multi-Cluster GPFS over the WAN. *I. Weger replied that MC-GPFS will be used for the file sharing of some files between ECMWF's two HPC clusters. The intended testing of MC-GPFS over the WAN is in conjunction with ECMWF's participation in the DEISA project, to allow file sharing between GRID members. Testing is particularly aimed at the security aspects and operational implications of file-sharing and will be using ECMWF's test cluster, which is not connected to the ECMWF LAN.*

T. Lorenzen asked whether the improvements which had been made to LoadLeveler at ECMWF had been incorporated generally into LoadLeveler worldwide. *I. Weger replied that ECMWF had worked in close co-operation with IBM and IBM had taken up some of the ideas for general implementation. Furthermore, ECMWF had built tools on top of Loadleveler and made sophisticated use of some Loadleveler features*

E. Krenzien asked what new functionalities had been incorporated into ECFS. *I. Weger replied that the possibility to move files and directories within ECFS was one example. See also N. Storer's presentation.*

## HPCF Phase 4 update & DHS update – Neil Storer

### Current computer configuration



### Phase 3 (HPCC & HPCD)

Phase 3 (POWER4+-based system), continues to provide a very good service. It has been very stable (once the problem of excessive MCM (Multi-Chip-Module) failures was resolved). Over the last year several problems that were caused by unscheduled power-downs, related to UPS issues, arose. Also power restrictions curtailed some of the Phase 4 testing on the familiarisation machine.

The problems associated with paging worsened with the ML06 upgrade of AIX 5.2. Several times (twice last week) it affected a whole cluster, when it caused GPPS to hang. We have tried several workaround solutions and are liaising closely with IBM to come up with a better solution. On the Phase 4 system any job that attempts to exceed the consumable memory it requests will be killed (with an e-mail notification).

### Member State file systems (on HPCD)

- ms\_home (34 GB ~30% full today)
  - backed-up (by ECMWF) to protect against loss of the file system
  - usage quota-controlled (80 MB per user), shares with ECMWF “home”
- ms\_temp (6 TB ~80% full today)
  - **not** backed-up (by ECMWF)
  - usage controlled by select-delete
- ms\_perm (300 GB ~20% full today)
  - **not** backed-up (by ECMWF)
  - **not** controlled by select-delete
  - usage quota-controlled, “administered” by User Support

- ms\_crit (850 GB ~20% full today)
  - not backed-up (by ECMWF)
  - not controlled by select-delete
  - usage quota-controlled, “administered” by User Support

#### **Phase 4 (HPCE & HPCF) – hardware**

2 clusters based on POWER5+ that will replace the two Phase 3 clusters. Each of the Phase 4 clusters has:

- 140 to 150 p5-575+ nodes
  - 16 1.9 GHz SMT CPUs per node (4 results per clock cycle giving a peak of 7.6 GFLOPs per CPU)
  - 32 GB physical memory per node (~25 GB usable memory)
  - 8 nodes per cabinet
- 8 16-way POWER5 I/O nodes (VSD servers)
- 2 16-way POWER5+ network nodes, each with dual 10 Gbps ethernet adapters

~100 TB of usable disk space, installed within a Storage Area Network (SAN), based on IBM fibrechannel SAN directors (intelligent, highly-reliable switches)

It is possible to configure the SAN such that all I/O nodes on both clusters see all the disks, thus enabling us to use Multi-Cluster GPFS effectively.

#### **Phase 4 – software**

- Operating system:
  - AIX 5.3 ML04 (Phase 3 runs AIX 5.2 ML06)
- Cluster services:
  - CSM 1.5.1
- MPI environment:
  - Poe 4.2.2.3
- File system:
  - GPFS 3.1
- Batch subsystem:
  - LoadLeveler 3.3.2
- Default compilers:

xlf 9.1 and xlc 7.0 with the 1.5 runtime environment. This is the same as on the Phase 3 system, since testing has shown that the newer xlf 10.1 compiler with rte 1.6 gives problems with the IFS code.

#### **Phase 4 migration timetable**

- Cluster-4E is already installed. Some users have had early access for the last 4 weeks or so and several Member State applications have been run on it by User Support (eg LM, OPA, HIRLAM). It will go into formal acceptance within the next week or two.
- Cluster-4F will start to be delivered in August (once HPCC has been decommissioned). It will go into formal acceptance in October.
- Full access to Cluster-4E will be given to MS users towards the end of July but note that the Operational Suite will also be running on this system.
- A full service on HPCE will be available to MS users from early November, when the O-Suite will move to Cluster-4F.
- HPCD will be decommissioned at the end of November.



## POWER5 features

- Upwards binary compatible with POWER4
- 90nm technology (POWER4+ was 130nm)
- 1.92 MB of L2 cache (128-byte, 10-way LRU), increased from 1.44 MB (8-way LRU)
- 36 MB of L3 cache (2x128-byte, 12-way LRU), increased from 32MB (4x128-byte, 8-way LRU)
- L3 cache has lower latency and higher bandwidth and acts as a “victim cache extension” of L2 cache (rather than an in-line cache for data from memory)
- The memory controller is now on-chip, giving faster memory access and fewer chips on an MCM (8 - was 16)
- The intra-MCM bus is twice as fast, the inter-MCM bus 50% faster
- Dynamic power management and reduced power leakage
- Faster barrier synchronization (special hardware)
- Increased rename-register resources (120 FPRs – previously 72)
- New instructions:
  - enhanced data pre-fetch
  - a faster SQRT and DIVIDE that does not conform strictly to the IEEE standard regarding rounding and exception handling
  - reciprocal approximation in 64-bit [FPRE()]
  - reciprocal SQRT approximation in 32-bit [FRSQRTES()]
  - byte-wise POPCOUNT [POPCNTB()].
- Simultaneous multi-threading (SMT) which produces ~30% gain in performance on IFS

All these features help the Phase 4 system to sustain a higher performance as a percentage of peak.

See: <http://www.research.ibm.com/journal/rd49-45.html>

## Different page sizes

- Small – 4 KB (default on all AIX systems)
- Mid-size – 64 KB (available only on AIX 5.3):
  - fully supported by the operating system - no O/S configuration needed
  - gain on limited testing of OPA 9.0 (0.25° resolution) at ECMWF ~10%
  - no evident drawbacks
- Large - 16 MB (available on all AIX 5 releases):
  - not well supported by the operating system - O/S configuration needed
  - previously Large Page pool had to be allocated statically at boot time

The new dynamic allocation feature (root-callable) is poorly implemented, with several problems, e.g. increasing the number of Large Pages can take several minutes and often aborts part way through. There are also fragmentation issues.

- scheduling and efficiency issues (Large Page applications can use Small Pages, but not vice versa)
  - gain on IFS code over small pages is ~1%
  - no benefit over Mid-size Pages seen during (very limited) ECMWF testing.
- Huge - 16 GB (for very special applications only)

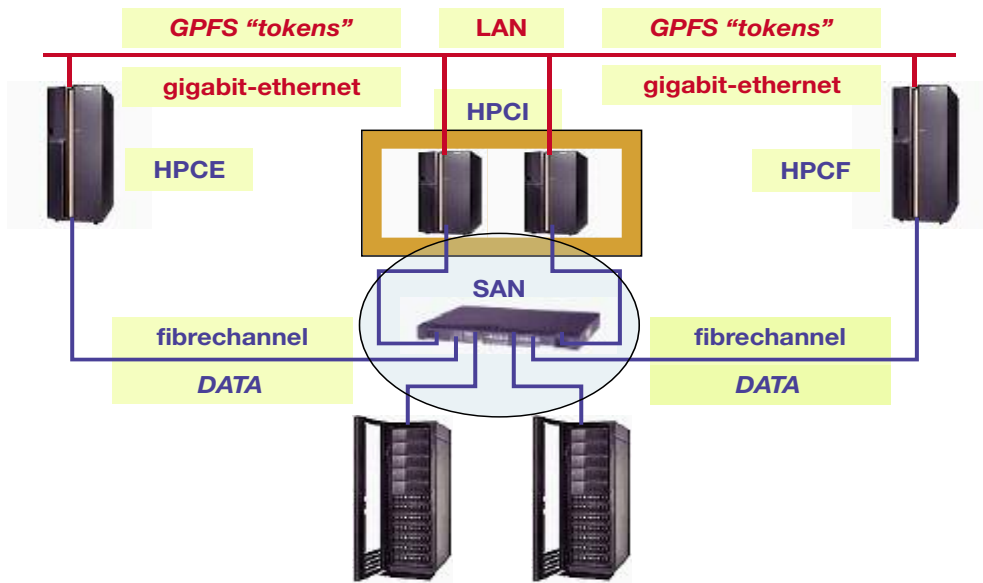
**Multi-cluster GPFS for Phase 4**

Data are transferred between the Phase 3 clusters over the LAN, either using FTP-like applications or via ECFS. MC-GPFS will enable each Phase 4 cluster to have concurrent access to the data, directly over a fibre-channel storage area network, at much higher data rates than are possible using NFS.

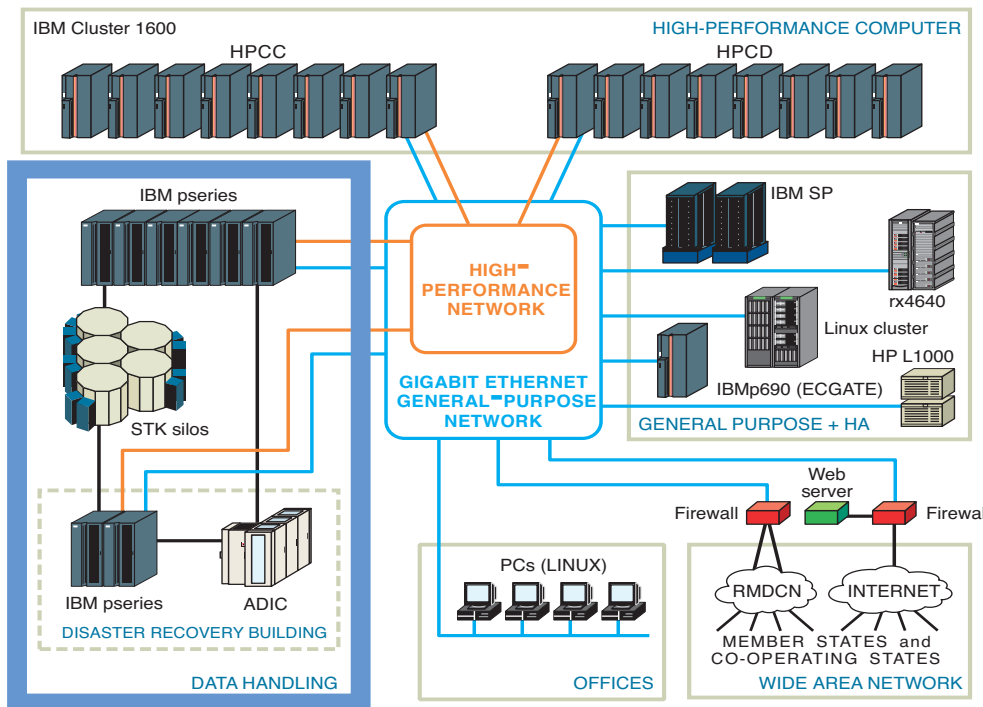
Various data are replicated on both Phase 3 clusters, effectively reducing the amount of usable disk space. MC-GPFS removes the need to replicate the data. The plan is to have ~50% of the disk space shared by both the Phase 4 clusters. The SAN has been set up with sufficient flexibility to enable this ratio to be changed easily.

MC-GPFS removes synchronisation problems (e.g. out-of-date copies), since there need only be one version of the data, and also helps with data management.

**Multi-cluster GPFS configuration**



**DHS**



## ECFS

- The “emove” command is available to rename files and directories in ECFS. This only works if the source and target files are in the same family within HPSS (all MS users and “normal” ECMWF users are in the same family).
- The recursive option (“-R”) for “els” should be fully implemented before the end of the year.
- A scheduler has been developed to control better the way that ECFS responds to requests. This helps to ensure that HPSS does not become overloaded and also enables us to prioritise access to the data in ECFS.
- HPSS version 6.2 is currently being tested. Changes must be made to the way ECFS and HPSS validate users by making use of the Entity Management System (EMS), but this will be transparent to end-users.

## ECFS back-up copies

- Please take note again that (unlike in the old TSM-based ECFS system) in the HPSS-based ECFS system, **by default: no secondary (backup) copy is made of ECFS data**
- The user has to specify the “-b” option on the “ecp” command to request that a secondary copy be made of any data that cannot easily be reconstructed, should the primary copy be destroyed.

## DHS plans

We plan to upgrade to version 6.2 of HPSS later this year.

Various hardware will be purchased later this year to cope with the increased load that will be generated by Phase 4 of the HPCF.

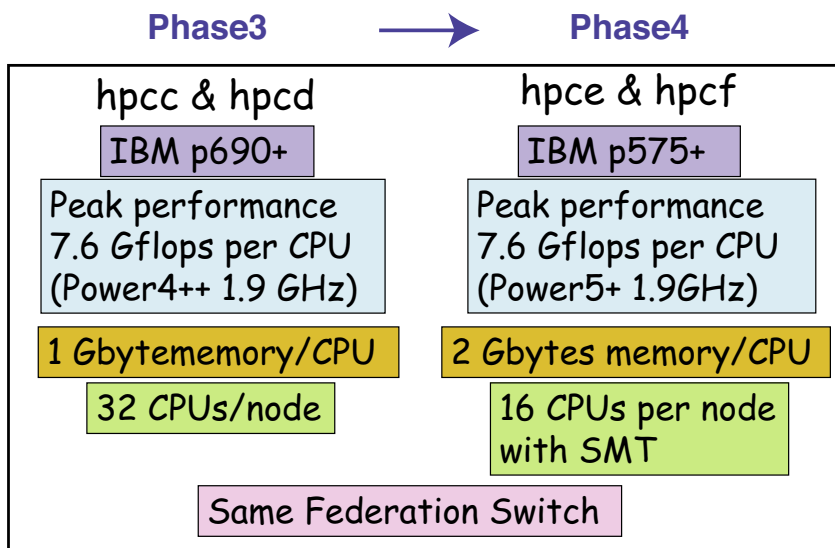
The current automated tape library (ATL) in the DRS building, an ADIC AML/J, reached its end-of-life almost 2 years ago. We are currently running an ITT to replace the ATL and its tape drives.

The migration plan calls for the LTO tape cassettes in the old library to be physically moved to the new one and for the data that exist currently on AIT cassettes to be copied.

The STK silos in the main computer hall are nearing their end-of-life, however we plan to keep them in service for the next few years at least.

T. Lorenzen noted that N. Storer had stated that performance gains of 30% on the IFS application had been experienced with SMT (Simultaneous Multi-Threading), was this what had been expected.? *N. Storer replied that less improvement had in fact been expected. He explained that SMT does not give full duplication of all the functional units, only more registers and some other shared units which are duplicated. This means that if the functional units are heavily used most of the time, it is very difficult to achieve a gain with SMT. (See D. Salmond's presentation.)*

Early experience on the IBM Phase 4 system – Deborah Salmond



**SMT = Simultaneous Multi-Threading**

- Node has physical 16 CPUs = 8 dual-core chips  
AIX thinks that the node has 32 CPUs  
2 'logical CPUs' are allocated to each 'physical CPU'  
These CPUs can be used with MPI or OpenMP
- Programs benefit from SMP if there is a mix of memory / FP ops  
e.g. IFS
- Some programs don't benefit from SMT  
e.g. If they have a lot of memory traffic per FP op  
or if they are 'FP bound' like SGEMM  
or the program doesn't scale well - uses 2\*CPUs
- User can choose not to use SMT inside loadleveler  
# @ tasks\_per\_node = 16

**Power4++ to Power5 with SMT for IFS and UM**

*UM (288x217x38) & IFS (T255L40) 1-day forecast*

	CPUs	IFS % of peak	UM % of peak
Power4++ 1.9GHz p690+	32	7.7%	4.0%
Power5 1.5GHz p575	32 no SMT	9.3%	7.4%
Power5 1.5GHz p575	32 with SMT	11.7%	8.0%

### IFS - T799L91 10-day forecast from RAPS9

	CPUs MPI x OMP	WALL (secs)	% of peak
Power4++ 1.9GHz p690+ hpcd	768 192 x 4	3848	7.6%
Power5+ 1.9GHz p575+ hpce	768 SMT 192 x 8	2457	11.8%

### IFS - T799L91 10-day forecast from RAPS9

	CPUs MPI x OMP	WALL (secs)	%Comms	Gflops	% of peak
Power4++ 1.9GHz p690+ hpcd	768 192 x 4	3848	12.6%	444	7.6%
Power5+ 1.9GHz p575+ hpce	768 SMT 192 x 8	2457	14.0%	696	11.8%

### Binding CPUs to chips and memory

- export MP\_TASK\_AFFINITY=MCM

keeps MPI task & its OpenMP threads on same resource

where resource is 'dual-core chip' on hpce = 2CPUs (was 'MCM' on hpcd = 8CPUs)

- On hpce this is OK for up to 2 OpenMP threads ( or 4 threads with SMT)
- If using more than 2 OpenMP threads, user must explicitly bind CPUs to tasks for best performance

IFS speeds up by ~5% with binding

- export MEMORY\_AFFINITY=MCM

allocates memory on nearest memory to task, as on hpcd

### How to use binding

- call `jfh_bind()` after `mpi_init` in your Fortran source
- compile and link with `jfh_bind.F90` & `jbind.c`
- Set the following environment variables in the run script:

```
unset MP_TASK_AFFINITY
```

```
export JFH_BIND=map
```

```
export JFH_SMTFILE="/tmp/smt.proc"
```

*If NOT using SMT:*

```
export JFH_BMAP="16 0 2 4 6 8 10 12 14 16 18 20 22 24 26 28 30"
```

*If using SMT:*

```
export JFH_BMAP="32 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31"
```

**IFS - T799L91 12 hour 4D-Var – Cycle 31**  
**2nd minimisation step**

	CPUs MPI x OMP	WALL (secs)
Power4++ 1.9GHz p690+	640 80 x 8	2379
Power5+ 1.9GHz p575+ No binding	640 80 x 8	2094
Power5+ 1.9GHz p575+ Binding & no SMT	640 80 x 8	1768
Power5+ 1.9GHz p575+ Binding & SMT	640 160 x 8	1604

**Other differences hpcd → hpce**

- xlf90\_r -qarch=pwr5  
targets compiler at the POWER5 processor  
e.g. can use software divide
  - more pwr5 features are expected in the version 10 compiler
- Memory limit set for loadleveler with e.g.  
# @ resources = ConsumableMemory(780mb)  
is now enforced – with MEMKILL
- Medium pages = 64k bytes (compared with 4k bytes)  
gives ~3% improvement on IFS  
Users only need to 'ldedit' their executable.  
ldedit -btextpsize=64K -bdatapsize=64K -bstackpsize=64K a.out

**Early results for UM, LM and HIRLAM**

- LM (Umberto)
  - 306 x 258 x 40
  - 12hr forecast
  - Ver 3.17

	CPUs	WALL
hpcd	128	196
hpce	128	129
hpce + SMT	(256)	122

- UM (Paul)
  - 288 x 217 x 38
  - 15 day forecast
  - Vn 6.0

	CPUs	WALL
hpcd	64	1950
hpce	64	1185
hpce + SMT	(128)	1075

- HIRLAM (Dominique)

- 438 x 336 x 40
- 48hr forecast

	CPUs	WALL
hpcd	16	2274
hpce + SMT	16 (32)	3434
hpce + SMT	15 (30)	1617

**Summary hpcd → hpce**

- New for hpce
  - SMT
  - Better aggregate memory bandwidth for a node
  - Dual core chips
- Same as hpcd
  - Clock speed is 1.9GHz
  - Federation switch
  - ~2000 physical processors per cluster

## Overview of ECMWF electrical and mechanical infrastructure - *Sylvia Baylis*

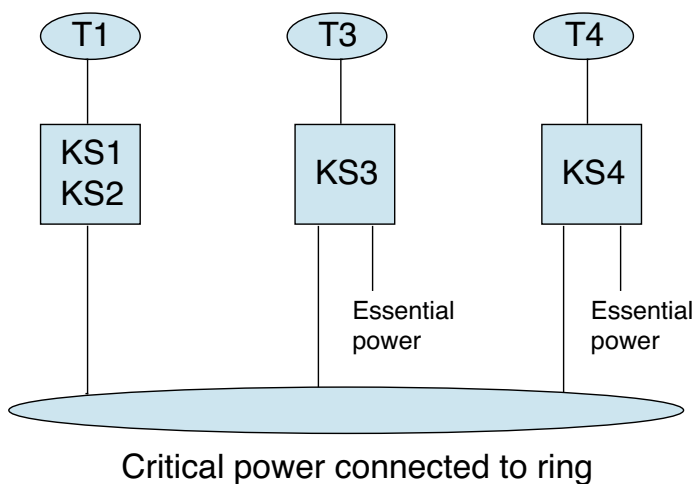
### Uninterruptible Power Supply (UPS)

The Uninterruptible Power Supply is based on rotary UPS machines.

- 1994 two machines installed – 800kW each
- 2004 third machine installed – 800kW / 800kW
- 2006 fourth machine installed – 800kW / 800kW

The two new machines provide a mixture of UPS and Essential (short break) power. The critical outputs from the UPS machines are connected in a ring to provide power to the site.

### Current power configuration



### Critical Loads

- Computer equipment in the main hall and tape library
- Computer equipment in the extension
- Denco air handling equipment
- All networking equipment
- Chilled water pumps
- All desktop systems except those in Terrapin Towers

### Essential Loads

- Chillers
- Lights and small power in computer building
- Desktops in Terrapin Towers
- All equipment other than compute / networking equipment throughout the rest of the site

### UPS resilience

The UPS system is designed as a resilient system for the critical loads. There is 4800kW of total on-site generation capacity, split between UPS and Essential supplies.

- UPS
  - a total of 3200kW critical power (all 4 machines running)
  - 400kW of resilient critical power (i.e. one UPS machine can be out of service and we can maintain this power level).



- Essential
  - 1600kW of essential (short break) power. There is no resiliency for the essential loads.

### **Current power loads**

We are currently powering 3 HPC clusters (during the parallel run) and the site power load is ~ 3,000kW (3 MW). Once HPCD is removed in 4Q2006, we will only have two clusters, so this will reduce to ~2,500kW. The total load of 2,500kW is split approximately:

- 1,500kW critical loads
- 1,000kW essential loads.

### **Mains problems – what happens**

The critical load is always supported by the UPS system, without a break, during mains problems. The way the essential load is supported depends on the total power load on the UPS system:

- if the total load is less than the available UPS capacity, essential supplies can be supported without a break;
- if the total load is greater than the available UPS capacity, the essential supplies will suffer a short break when the UPS goes into emergency mode.

### **UPS bearing problems**

Each UPS machine has 5 main bearings with an anticipated life of 10 years. As part of the installation of KS3 in 2004, all the bearings on the original UPS machines were changed. The bearing replacement was not completed to a sufficiently high standard and there have consequently been bearing failures.

With one machine out of service, UPS resilience is lost. The supplier has acknowledged the problem and is now working to resolve the underlying difficulty of below standard refurbishment.

### **Service interruptions in the last year**

There have been 5 power interruptions in the last year for a variety of reasons:

- 28 July 2005 – After a mains failure, a failure in the UPS control system resulted in total power outage.
- 31 July 2005 – After a mains failure, a failure in the UPS control system resulted in total power outage.
- 5 December 2005 – A switchgear failure during a normal switching operation resulted in loss of supply to the computer building.
- 30 January 2006 – Failure of a fuel level sensor caused no fuel to be available for UPS machines during a period of repeated mains fluctuations, leading to loss of supply to the computer building.
- 22 March 2006 – Failure of a temporary cable during KS4 installation resulted in loss of supply to the computer building.

### **Latest UPS machine failure**

The most recent UPS machine failure (KS2) occurred on 10 May 2006; there was no operational impact. The failure was the result of a balancing screw working loose and eventually becoming free within the accurotor. The screw flying around the machine caused damage to the electrical windings within the machine. The machine is away being repaired, so we are currently running without resilience. The machine should be reinstalled and operational later this month.

### **Electrical distribution in the future**

Two of the main 415V switchboards are now obsolete and will need to be replaced in the near future, as spare parts are not available.

The 11kV switchboard is now 30 years old. Whilst there are still spares available at the moment, consideration must be given to a timely replacement. A feasibility study is currently being carried out to determine:

- the options for replacement of these switchboards with minimal disruption to service. Installation of new 415V boards must be planned around any HPC installation.
- how to reduce single points of failure in the main electrical distribution system, when switchboards are replaced.

## Chilled water system

The chilled water system provides chilled water to:

- cool the computer equipment within the computer building via air handling units;
- condition fresh air before it is passed into the computer building;
- supply chilled beams in the offices;
- supply cooling for the air conditioning systems in the conference building.

In 1995 3 Carrier chillers were installed, this was increased to 4 in 2003. In 2005/06 a fifth chiller, using new technology, was installed to cope with the increased capacity required during the summer months. In 2006 a sixth chiller is being installed to provide increased capacity and resilience during the HPC parallel run in the summer (3 HPC clusters are being powered). Once we are back to two HPC clusters, we plan to switch off one of the older, less efficient chillers.

## Choice of refrigerant

Before purchasing a new chiller in 2005, a study was undertaken to determine the best refrigerant to be used in any new chillers at ECMWF.

The report covered the various options available:

- Ammonia
- Hydrofluorocarbon (R134a, R407, R410 etc)
- Hydrocarbons (Propane, Butane, Isobutane and Propylene)

All of these have zero Ozone Depletion Potential (ODP) and varying Global Warming Potential (GWP).

It was decided that ammonia and hydrocarbons would not be selected for a variety of reasons, including explosion potential, chemical handling and storage problems.

The refrigerant chosen was HFC R134a (0 ODP / 1300 GWP / 14yr life). In selecting R134a great consideration was given to obtaining:

- the most efficient Chiller possible to reduce energy costs, reduce electricity consumption and hence reduce CO<sub>2</sub> emissions from electricity generation;
- the most robust containment construction, to reduce the probability of releasing refrigerant to the atmosphere.

## Chiller efficiency

Chiller efficiency is referred to as the Coefficient of System Performance (COSP) and is the ratio of heat load rejected to electrical power input.

At full load the Star chillers are approximately 35% more efficient than the original Carrier chillers, therefore reducing the electrical load of a chiller by at least 75kW in 215kW averaged over the year. This will produce annual electrical savings of circa £35,000 / Star chiller operating at 100% load. ECMWF does not operate the chillers fully loaded and, unlike other chillers, the efficiency of the Star chiller increases at part load - thus even greater savings can be achieved.

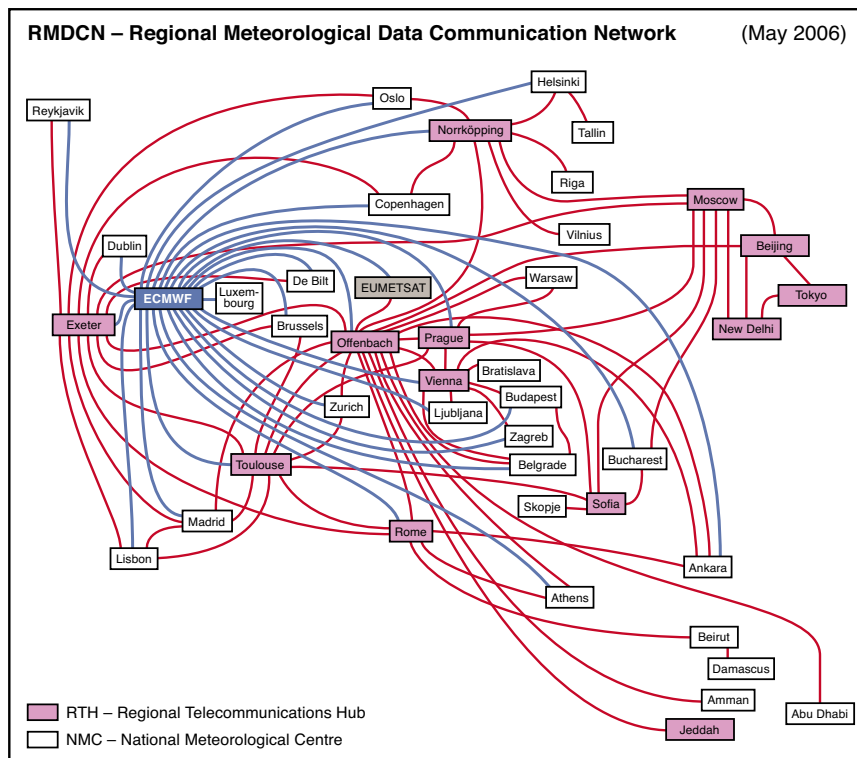
## Service interruptions in the last year

There has been one service interruption for a problem with the centralised chilled water system. On 17 May 2006 the failure of a main valve on the centralised chilled water system resulted in a planned outage for replacement of the valve.

## Chillers in the future

The three 11 year old Carrier chillers at ECMWF have an anticipated life of circa 15 years and use R22 as the refrigerant. In 2010 the manufacture of R22 will be banned; only reclaimed gas will be available for use. In 2015 the use of R22 will be banned completely. Starting within the next 2 – 3 years, it will be necessary to begin replacement of these three chillers. If the performance of the Star chiller is as predicted, consideration should be given to replacing the 11 year old Carrier chillers at the earliest opportunity, to take advantage of electricity cost savings.

## RMDCN status and migration plans - Tony Bakker



### Current Frame Relay Network

44 sites (41 Countries+EUMETSAT+ECMWF+Germany Disaster Recovery)

– Saudi Arabia accepted on 7 July 2005

Spain: Access Line 2 Mbps, Access Port 1 Mbps, ECMWF PVC 768/128 kbps, ISDN Backup 512 kbps

Austria: ECMWF PVC 384/128 kbps

EUMETSAT: Access Line/Port 512 kbps, PVC Germany 128/128 kbps

Latvia: Access Line/Port 128 kbps, Sweden PVC 64/16 kbps

China-Japan PVC 96/48 kbps

Russia: Access Line/Port 256 kbps, Sweden PVC 16/16 kbps

Russia-Romania PVC 16/8 kbps

Finland move HQ

### Major Outages

Luxembourg: 1 day + 4 hours : modem failure

Norway: 5 hours 12/5/06: PTT equipment failure

FYR Macedonia: 1 day 15-16/12/05: PTT strike

Romania: 3 days 30/11/05 - 3/12/05: Access Line problem

Italy: prolonged service degradation lasting 2 weeks in November 2005: Access Line problems

Norway: 18 hours 7/1/05: EQUANT node switch port failure

ECMWF: 2.5 hours 31/7/05: site power problems

ECMWF: 5 hours 28/7/05: site power problems

FYR Macedonia: 8 hours 8/7/05: flooding at site

Sweden: 5 hours during 6-21 May 05: EQUANT node switch port failure

## Other Issues

Annual Price Review, 1/4/06, resulted in a 10% discount on all elements including access lines.

1/6/06 – EQUANT, Orange and France Telecom merged to form Orange Business Services.

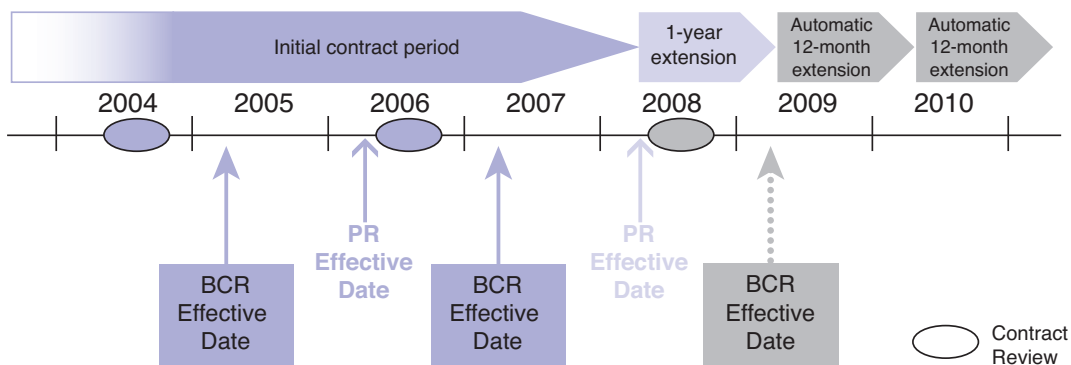
## RMDCN Migration to IPVPN MPLS

The Biennial Contract Review in 2004 resulted in an offer from EQUANT for migration from the current Frame Relay technology, used in the RMDCN, to IPVPN MPLS technology. A proposal was made to ECMWF TAC and Council in 2004. After adoption of the proposal by the Council, all MS User Sites and WMO were informed. Once all User Sites had agreed to the proposal, contract negotiations were started. Supplement no. 4 to the RMDCN contract was signed on 8 May 2006.

## Extension of the Term

Equant had requested a 1-year extension to the contract Initial Term.

ECMWF asked Equant to include a continuation option with automatic extension at the end of the extended term, but with ECMWF having the option to terminate the contract on 6 months' notice. This provides ECMWF with both a strong negotiating position for the 2007 biennial contract review and the necessary flexibility to issue an ITT, if necessary.



## Service Level Agreement

### Scheduled Maintenance

– No single outage must exceed 60 minutes; in practice, only 15 minute slots are envisaged by EQUANT.

### Site availability

Site availability depends on the level of resiliency and on the location

– Mission Critical : 100%

- requires diversity of Access Lines

– Enhanced resilience (Dual router, Access Line with automatic NAS backup): 99.9% for most of the sites, 99.5% for India, FYR Macedonia

- Equant and User Sites must test the backup on a quarterly basis

Service Rebates are due when the Site availability is not achieved:

- 1/25th of the Relevant Monthly Charge for each hour of Outage in excess of the committed Site availability
- Cumulative but cannot exceed 100% of the monthly charge

### Guaranteed Time To Repair (GTR)

The baseline is a 2-hour restore time for any Severity 1 fault (Severity 1 implies that the service is unavailable).

For Mission Critical sites EQUANT agree to commence the resolution of a severity 1 fault within 10 minutes of the start of the Service Interruption.

GTTR of less than 24 hours from the start of Service Interruption for a Severity 2 fault (Service is being provided by a backup).

GTTR of less than 10 Working Days for a Severity 3 Fault (a problem that does not seriously affect the Service).

GTTR of less than 20 Working Days for Severity 4 Fault (a minor issue, such as a change request).

#### **Service Rebates**

Service rebates are due, when the GTTR for a Severity 1 fault is not achieved.

Rebate rates:

0.5% Relevant Monthly Charge, if the problem is not resolved within 2 hours. 1%, if the problem is not resolved within the next 2 hours. Thereafter 1% for each additional 4-hour period that the problem resolution takes.

#### **Packet Loss Ratio (PLR)**

- guaranteed on a PE-to-PE path only
- measured on a monthly basis
- baseline is a PLR between 0.1% to 0.4%, depending on source and destination
- Service Rebates are due when the PLR is greater than 10% of the agreed level.

#### **Round-Trip Delay (RTD)**

- The round-trip time of a 128 bytes packet from CE router to CE router is measured. The baseline is a RTD between 60ms to 580ms, depending on source and destination
- Service Rebates are due when the RTD exceeds 10% of the agreed level.

#### **Service Degradation**

- Measured throughput must not fall below 80% of the IP bandwidth.
- Measured throughput for the COS must not fall below 80% of the IP bandwidth for each COS.
- A backup service must be available for at least 90% of the time in any one calendar month.

#### **Global Service Degradation**

- When 30% or more of all the User Sites fail to meet the specified SLA

Service Rebates are due when Service Degradation has been demonstrated.

#### **Annual Price Review**

The 'Head of Agreement' envisaged a first IPVPN Price Review on 1 April 2006.

The first Price Review Date for the IPVPN Charges will be 1 April 2007, however, ECMWF requested Devoteam to undertake a financial analysis of Equant IP VPN pricing, which resulted in a 10% reduction of the original EQUANT offer (2004) for the IPVPN MPLS migration.

From 1 April 2007 the Price Review will be based on Equant's commercially competitive prices.

- Equant will provide ECMWF with a breakdown of all charges on a per country basis, reflecting Equant's pricing structure.
- Only reductions are acceptable.
- ECMWF has the option to request an independent 3rd party to assess the RMDCN charges.

#### **Charges (Schedule 4)**

There will be no overlap of the charges during the migration from Frame Relay to IP VPN.

ISDN installation (which includes ordering the circuit) and rental charges at the User Site are the User Site's responsibility.

Silver Service Type will be available for User Sites that do not require the support of differentiated COS, at a discount of 8% on Gold Service Type charges.

For User Sites with a dual connection (Mission Critical Sites) load balancing is available at an additional cost of 5%

of the relevant monthly Charges. Monthly charges for IP bandwidth backup are 33.3% of the primary IP bandwidth charges.

### **Cost redistribution**

Equant had agreed to reassess the costs per site, i.e. to reduce the IP VPN charges for the sites where MPLS charges would be higher than 1 April 2004 Frame Relay charges, provided that the total cost of the proposed RMDCN configuration remains unchanged.

Cost balancing will be re-assessed at each Price Review. The number of sites needing assistance should decrease, as MPLS prices will continue to go down because of competition between providers.

Pricing for upgraded configurations will be calculated by adding the discounted cost of the additional service to the "balanced" cost, e.g.:

Initial: 100 kbps	discounted cost: £100	balanced cost: £80
Upgrade: 200 kbps	discounted cost: £150	balanced cost: £80 + £50

### **Service Particulars**

#### *Service Particulars include:*

- Access line speed
- Type of router(s)
- Port speed
- COS
- Extended service management and fault management
- Backup topology
- User Site contact details
- Equant help desk details

### **ECMWF Basic Package**

ECMWF will fund the following basic package (this is an Enhanced Resilience package with GOLD Service) for Member States:

- 768 kbps Access Line
- 768 kbps IP Bandwidth
- 384 kbps NAS Backup
- primary router and backup router
- maintenance charges
- funding of the redistribution charges (See Table 4 Schedule 4)
- one-time installation charges according to Table 1 Schedule 4

**Class of Service**

D 1 = 75 %	D 2 = 20 %	D 3 = 5 %
Traffic coming from local Message switch or system generating GTS traffic Traffic going to ECMWF dissemination system Environmental warnings ECMWF dissemination to MS	Traffic coming from radar system (Sweden, Norway, Finland)* Traffic coming from lightning system (Sweden, Finland)* Product transmission from Sweden to Lithuania and Latvia LACE coupling data from ECMWF to LACE countries Uploading LACE data from MF to ECMWF MF- Czech ALADIN input Time-critical dissemination to Member States (Italy, MF?) Lightning data Italy Spain EUMETNET and EU project pre-operation satellite data (radiance, ...) Access to ECMWF via EAccess Monitoring and routing traffic	EUMETNET and EU project during development phase

Note: ECMWF will limit its traffic to 80% (or other agreed value) of the MS access speed

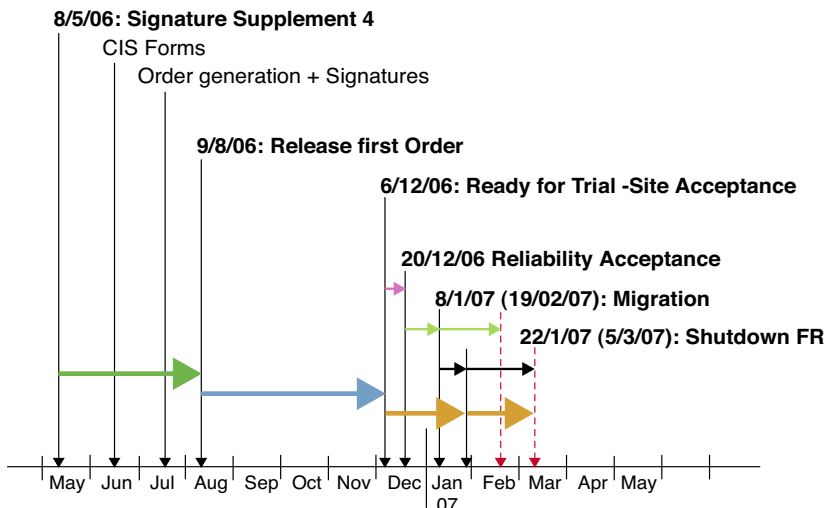
**Implementation plan (Schedule 5)**

*Big Bang approach*

The implementation of the IP VPN will have no effect on the current Frame Relay network. There will be regular project meetings to review the implementation progress. Once the order forms have been returned to Equant, there will be no changes to the requested configuration until the final migration of the network. Modifications to the FR network will not be implemented after the IP VPN Ready For Trial Date.

All User Sites must be accepted before the start of the Reliability Acceptance test. Failure of a User Site to perform the Site Acceptance test should not lead to delay in the commencement of the Reliability Acceptance test. The Reliability Acceptance test will commence on the day following the end of the Site Acceptance test.

Equant will arrange for the IP VPN network to go into operation on the 11th day following the completion of the Reliability Acceptance test.



## Acceptance (Schedule 6)

### *Site Acceptance*

- To demonstrate functionality / installation at every site
- Verify the general workmanship, access to help desk, etc.
- Test TCP/IP protocols and measure throughputs

### *Reliability Acceptance*

- 20 consecutive day availability test
- 90% of the sites must meet the SLA
- The rest have to achieve 99.5% of the committed Service Level

### *Migration*

- 10 working days to migrate the operation to the IP VPN

Site acceptance or migration may be suspended, if there is a severe weather event.

P Halton asked whether the rigorous Terms & Conditions negotiated with Equant were protected after Equant's takeover by Orange. T Bakker replied that the initial contract was signed with Equant UK Ltd and at that time ECMWF specifically requested a parent company guarantee from Equant BV in the Netherlands. At the time when France Telecom became a major shareholder in Equant, the situation was reviewed. Since the merger between France Telecom, Equant and Orange the situation is being re-investigated.

In regard to the Service Level Agreement and Guaranteed Time to Repair: Severity Level 3 Fault ("Problem that does not seriously affect the service") T. Lorenzen asked who determined whether a fault did not seriously affect the service. T. Bakker replied that a more precise definition of fault severity was included in the contract.

In regard to Packet Loss Ratio and Round Trip Delay, for which Service Rebates are due, if PLR or RTD exceed certain defined levels, he asked who carried out the tests to ascertain the failure to achieve the required levels. T. Bakker replied that ECMWF runs the procedures which form part of the User Site Acceptance Test, if other problems, such as throughput delays, are experienced on the network.

C. Gambuzza asked whether MSs would be offered the opportunity to use the load balancing facilities. T. Bakker replied that this was one of the options in the site definition schedule. MSs who have dual-type connectivity can select this facility at an additional 5% cost on their monthly charges. Configuration details for all user sites are currently being finalised for the new MPLS system. Order forms reflecting the selected options will then be created by Equant for signature by the MSs. Configurations can be altered at any time up to signature. Once an order has been signed, no changes will be accepted until after the rollout of the new MPLS based network is complete.



## ECaccess status and plans - Laurent Gougeon

### ECaccess in a nutshell

ECaccess is a secure portal to provide access to ECMWF archiving and computing facilities either through the Internet or RMDCN Network.

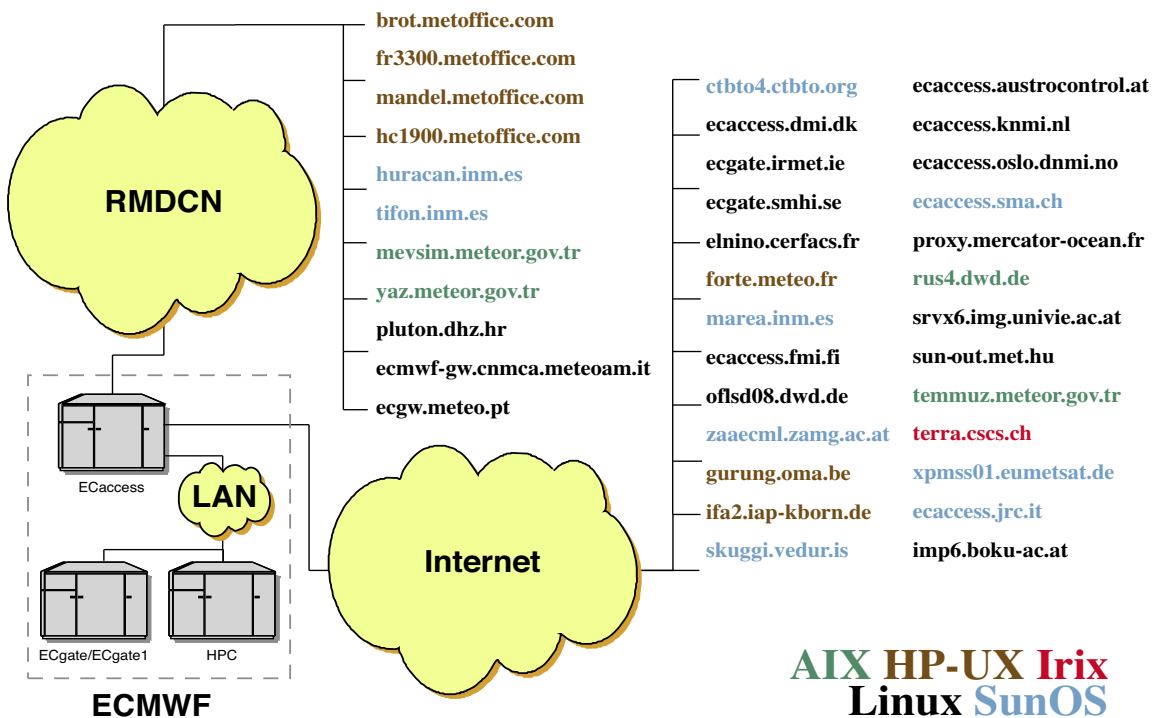
It provides file and job management in batch or interactive mode through an extended FTP server or Web server.

It allows secure file transfers between ECMWF and remote sites running an ECaccess Gateway or a FTP/SFTP/GFTP daemon. It has a retry mechanism for network connectivity issues and dissemination is through remote ECaccess Gateways.

It supports secure interactive sessions to ECMWF platforms:

- UNIX command prompt via a Telnet or SSH client
- Graphical access via an X11 server or a VNC client
- Graphical access via a Web browser.

### ECaccess network



### ECaccess v3.1.2

#### New features

- Graphical access via a Web browser (already available at ECMWF)
- Ectrans enhancement (priority and retry mechanism) - 2 simultaneous transfers allowed per user
- Job scheduler enhancement (MS time-critical applications)
- Access to ECFS projects (e.g. ec:/hirlam)
- Support for the dissemination

The ECaccess Server at ECMWF is scheduled to be upgraded in mid June

A release of the new ECaccess Gateway, first deployed at ECMWF for the local Gateways, should be available for download at the end of June.

## Gateway versions

- 2.0.1\_031  
proxy.mercator-ocean.fr
- 2.0.2\_003  
ctbto4.ctbto.org
- 2.0.3\_000  
ecaccess.austrocontrol.at, ecgate.irmet.ie,  
srvx6.img.univie.ac.at, sun-out.met.hu,  
terra.cscs.ch
- 2.1.0\_006  
brot.metoffice.com, fr3300.metoffice.com,  
hc1900.metoffice.com, mandel.metoffice.com,  
oflso8.dwd.de, xpmss01.eumetsat.de
- 2.1.0\_008  
ecaccess.knmi.nl, ecaccess.oslo.dnmi.no,  
ecaccess.sma.ch, elnino.cerfacs.fr,  
gurung.oma.be, ifa2.iap-kborn.de,  
pluton.dhz.hr, skuggi.vedur.is
- 2.2.0\_2004060701  
ecgw.meteo.pt, huracan.inm.es, marea.inm.es,  
mevsim.meteor.gov.tr, temmuz.meteor.gov.tr,  
tifon.inm.es, yaz.meteor.gov.tr
- 2.2.0\_2004061501  
rus4.dwd.de, zaaecm1.zamg.ac.at
- 3.0.0\_2005060601  
ecaccess.dmi.dk, ecaccess.fmi.fi,  
ecaccess.jrc.it, ecgate.smhi.se,  
ecmwf-gw.cnmca.meteoam.it, forte.meteo.fr,  
imp6.boku.ac.at

**! Systems in red are recommended to upgrade (see section below)**

## Upgrade to ECaccess v3.1.2

### *Automatic Gateway upgrade via a script*

**Those sites listed in red in the above list, whose Gateway is at a level < v2.2.0, are highly recommended to use the automatic upgrade, which:**

- upgrades any Gateway to v3.1.0
- imports the configuration from the initial installation
- is safe, because:

the initial ECaccess installation is not altered and the start-up script is not modified (to be done manually)

- If required by a simultaneous Java upgrade, database upgrade (converts ECtrans associations) and management of the Java key store conversion are performed

### *Manual Java upgrade*

- **An upgrade is highly recommended for those sites listed in red in the above list, whose Java is at < 1.4.1:**
- To achieve support for Java1.5

## Java versions

- 1.3.1  
proxy.mercator-ocean.fr,  
ecaccess.oslo.dnmi.no, ecgate.irmet.ie,  
sun-out.met.hu, srvx6.img.univie.ac.at,  
hc1900.metoffice.com, fr3300.metoffice.com,  
brot.metoffice.com, mandel.metoffice.com
- 1.4.0  
ecaccess.sma.ch, ifa2.iap-kborn.de,  
ctbto4.ctbto.org
- 1.4.1  
ecaccess.knmi.nl, ecaccess.austrocontrol.at,  
rus4.dwd.de, xpmss01.eumetsat.de,  
ecaccess.jrc.it, temmuz.meteor.gov.tr,  
terra.cscs.ch, zaaecm1.zamg.ac.at,  
mevsim.meteor.gov.tr, yaz.meteor.gov.tr
- 1.4.2  
ecaccess.fmi.fi, oflso8.dwd.de,  
ecaccess.dmi.dk, skuggi.vedur.is,  
forte.meteo.fr, gurung.oma.be, ecgate.smhi.se,  
elnino.cerfacs.fr, imp6.boku.ac.at,  
marea.inm.es, ecgw.meteo.pt, tifon.inm.es,  
huracan.inm.es,  
ecmwf-gw.cnmca.meteoam.it, pluton.dhz.hr

## What next?

### *Maintenance releases of ECaccess*

- Support for the new hpcf system
- Support for the new Linux cluster
  - Integrate ECaccess with the Sun Grid Engine

### *ECaccess new generation (2 year project)*

The architecture will be reviewed to benefit from new technologies and there will be a move toward standard protocols (Web Services)

- Impact on Member States:
  - A new Gateway application will be required
  - It will initially run simultaneously with ECaccess
- The project is currently in the requirements phase, collecting past user comments and wishes. **Any suggestions are welcome at this stage.**

## ECaccess new generation

### *Review the user interface*

- Web interface (add / remove / update / filtration)
- New layout

### *ECaccess domains*

- Host based vs. file system based
  - ECGATE, HPCD, ECFS
- Special project with ECFS (easier to use)
- Remote domains?

### *Load balancing / backup*

- RMDCN vs. Internet
  - Simultaneous connections
  - Priorities per service
- ECaccess nodes (use GRID technologies)?

### *ECaccess tools*

- Binary version of the tools will be made available
  - Support for meta-characters
- Sftp vs. ftp based (e.g. ecpur or ecget)
- Resource files
- ECaccess directives- parameters to be added
- Harmonisation (too many options)
- Map UNIX commands on remote servers
  - Investigate GRID shell (ECaccess Shell)

### *The symmetry*

- File transfers between remote Gateways
- Job synchronization between remote Gateways
- All existing services under ECaccess?

### *User transfers and dissemination*

- Transfers monitored by the operators
- Association management from the command line
  - Push a local association to a remote Gateway

- Priorities across different users

### *Deployment and administration of remote Gateways*

- Documentation (FAQ and release notes)
- Controlled upgrade (from the admin interface)
  - Bundle Java with the Gateway distribution
- Enhance administration tools
  - Web based database / log manager
- Deal with Firewall and traffic issues (e.g. timeouts and priorities)

### *Any other suggestions?*

### **ECaccess status**

- To learn the latest status of the project:

**<http://www.ecmwf.int/services/ecaccess/>**

- Download now!
- User's manual (on-line & PDF)
- Administrator's manual (PDF)
- Registration centre (SecurID card required)

- To test the new facilities:

**<http://ecaccess.ecmwf.int/sshvnc/>**

To upgrade please contact: **[ecaccess@ecmwf.int](mailto:ecaccess@ecmwf.int)**

P. Halton asked whether the Linux operating systems would need to be upgraded for the new implementation of ECaccess. L. Gougeon replied that it would depend what level the Linux was. The necessary level is listed on the ECaccess website: **<http://www.ecmwf.int/services/ecaccess/>**.

## **New MS job submission via SMS – Dominique Lucas**

### *Framework for MS time-critical applications:*

- Simple job submission monitored by the Centre
- Member State SMS suites monitored by the Centre
- Member State SMS suites managed by the Centre

### *Data access from real-time archive*

- ECMWF data distribution - dissemination
  - ~100 GB/day
- Real-time MARS access by MS users on ecgate
  - ~250 GB/day
  - MS jobs under SMS
  - Crontab entries

## **Current system**

This system has been in place since about 1991.

Users put jobs in specific directories, e.g. .../fc12h240.

The ECMWF operational task "fc12h240" takes all jobs in matching directories and submits them.

Some environmental variables are passed to the user jobs to give the date and time of the current operational run.

About 500 jobs are submitted in this way every day.

No monitoring is available.

Direct job submission to the supercomputer was lost, when the IBM HPCF system was installed.

## **New system**

### *Enhanced ECaccess batch system*

The scheduled running of jobs and a retry mechanism (in ectrans) were already available in ECaccess.

The new concept of "events", also known as notifications, has been added to ECaccess. Events are defined by one user; they can be made publicly available. E.g.

Event name: "fc12h240"

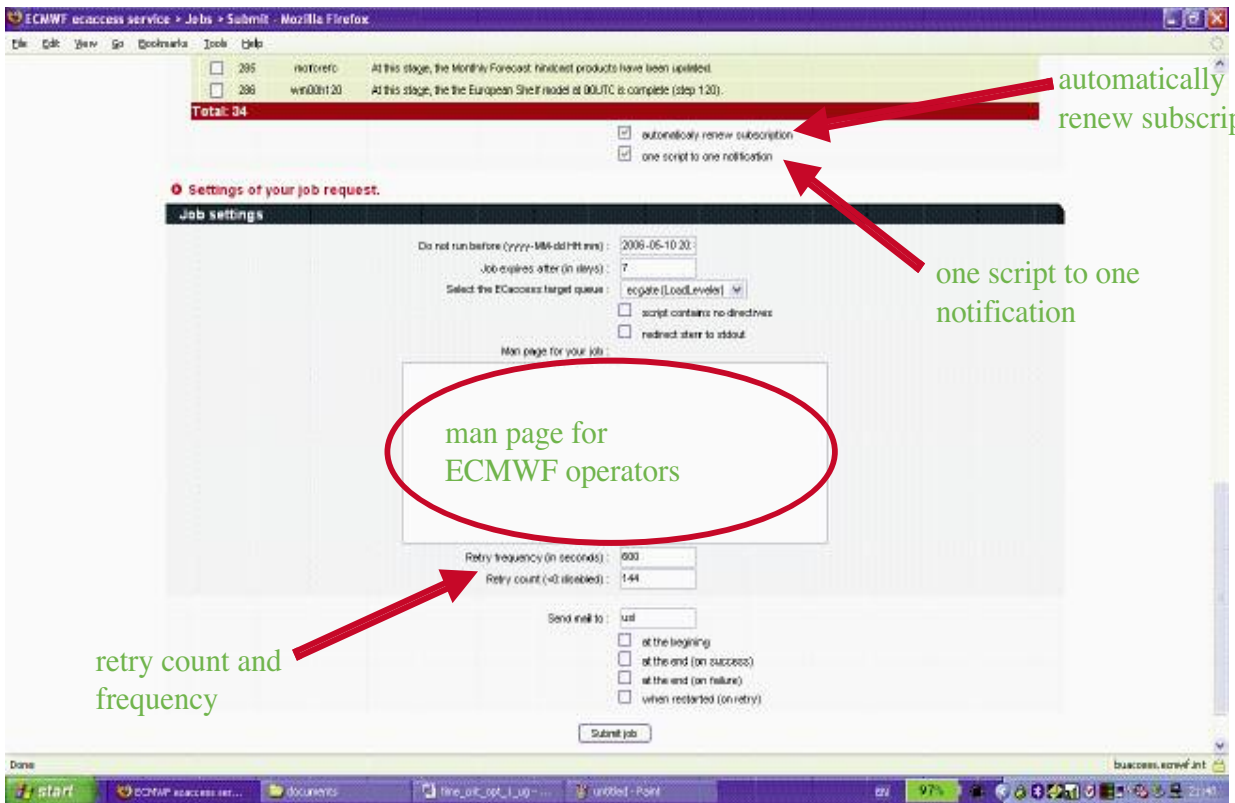
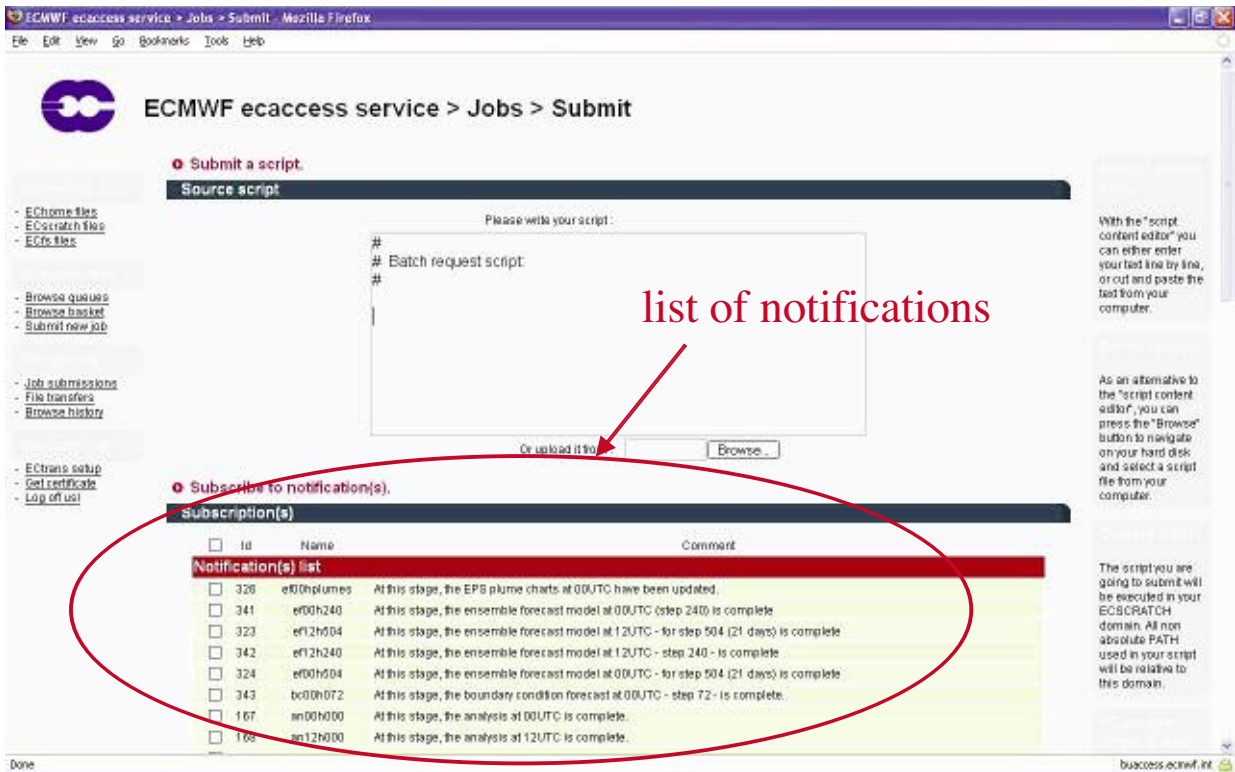
Event description: "at this stage, the 10 day forecast data from the high resolution 12UTC run is available".

Users can subscribe their own jobs to "public events"; these jobs will remain in standby mode, until the event owner sends a notification to an event; ECaccess will then submit the jobs subscribing to that event.

The special environmental variables are passed to the jobs, when the notification is given to the event.

Lastly, soon after the notification of an event, ECaccess will schedule a new version of the jobs subscribing to the event, ready to be submitted at the next notification.

User interface: web submission



### User interface: web job monitoring

Use this interface to track jobs you have submitted to ECMWF.

Jobs submitted by us2

Jobid	ECaccess queue	Notification(s)	Schedule	Try number	Status
35925	hpcd (LoadLeveler)	bc00h072 (344)	Jun 02 05:03	0/1	INIT
35926	hpcd (LoadLeveler)	bc00h072 (344)	Jun 02 05:03	0/1	INIT
35954	hpcd (LoadLeveler)	fc12h000 (185)	Jun 01 19:43	0/1	STDBY
35983	ecgate (LoadLeveler)	fc12h000 (185)	Jun 01 19:43	0/1	STDBY
35982	hpcd (LoadLeveler)	fc12h240 (182)	Jun 01 19:43	0/1	STDBY
35981	hpcd (LoadLeveler)	wg12h240 (180)	Jun 01 19:43	0/1	STDBY
35950	hpcd (LoadLeveler)	fc12h144 (181)	Jun 01 19:43	0/1	STDBY
35959	ecgate (LoadLeveler)	fc12h240 (182)	Jun 01 19:42	0/1	STDBY
35958	ecgate (LoadLeveler)	fc12hmetgram (184)	Jun 01 19:42	0/1	STDBY
35957	hpcd (LoadLeveler)	fc12hmetgram (184)	Jun 01 19:42	0/1	STDBY
35956	ecgate (LoadLeveler)	wg12h240 (180)	Jun 01 19:42	0/1	STDBY
35955	ecgate (LoadLeveler)	fc12h144 (181)	Jun 01 19:42	0/1	STDBY
35954	ecgate (LoadLeveler)	bc12h072 (345)	Jun 01 19:42	0/2	STDBY
35952	ecgate (LoadLeveler)	bc12h072 (345)	Jun 01 19:42	0/1	STDBY

All jobs are kept track of by their job identity number. The job status can be either "STDBY", "INIT", "WAIT", "RETR", "EXEC", "DONE" or "STOP".

To get more details about a job, select it with your mouse in the list by clicking the "Track job" icon.

You can delete one or more than one job by ticking the job(s) in the list and clicking the "Delete selected" icon.

### User interface: ETools: list events

*ecels - list events available to user:*

```

> ecels
326 ef00hplumes At this stage, the EPS plume charts at 00UTC have been updated.
341 ef00h240 At this stage, the ensemble forecast model at 00UTC (step 240) is complete
323 ef12h504 At this stage, the ensemble forecast model at 12UTC - for step 504 (21 days) is complete
342 ef12h240 At this stage, the ensemble forecast model at 12UTC - step 240 - is complete
324 ef00h504 At this stage, the ensemble forecast model at 00UTC - for step 504 (21 days) is complete
343 bc00h072 At this stage, the boundary condition forecast at 00UTC - step 72 - is complete.
167 an00h000 At this stage, the analysis at 00UTC is complete.
168 an12h000 At this stage, the analysis at 12UTC is complete.
172 ef00hmetgram At this stage, the EPS metgram database at 00UTC has been updated.
...
> ecels 343
Notification id: 343
Name: bc00h072
Public: true
Owner: usl
Title: /od/o/msjobs/00bc/ms072
Comment: At this stage, the boundary condition forecast at 00UTC - step 72 - is complete.
    
```

## User interface: ETools: job submission

*ecjput - ecjls*

```

> ecjput -help
Syntax: JPUT ECaccess-queue local-script [args ...]
...
-ni - notifications ids (list separated by ';' or ',')
-ro - renew subscription off (default is on)
-oo - one script to one notification off (default is on)
-mr - send mail when the execution/transfer retries
-mp - man page content (comment for the operators)
-rc - define the number of job retries (default is 0)
-rf - define the frequency of job retries in seconds (default is 600)
...
> ecjput ecgate sms.cmd -nd -ni 343 -mp "nothing to be done" -rc 1 -rf 300
35853
> ecjls 35853
Jobid: 35853
Location: ecgate@ecgate.ecmwf.int
Notification(s): bc00h072 (343)
Schedule: May 31 20:06
Try number: 0/2
Status: STDBY
    
```

## Operators' interface: monitoring

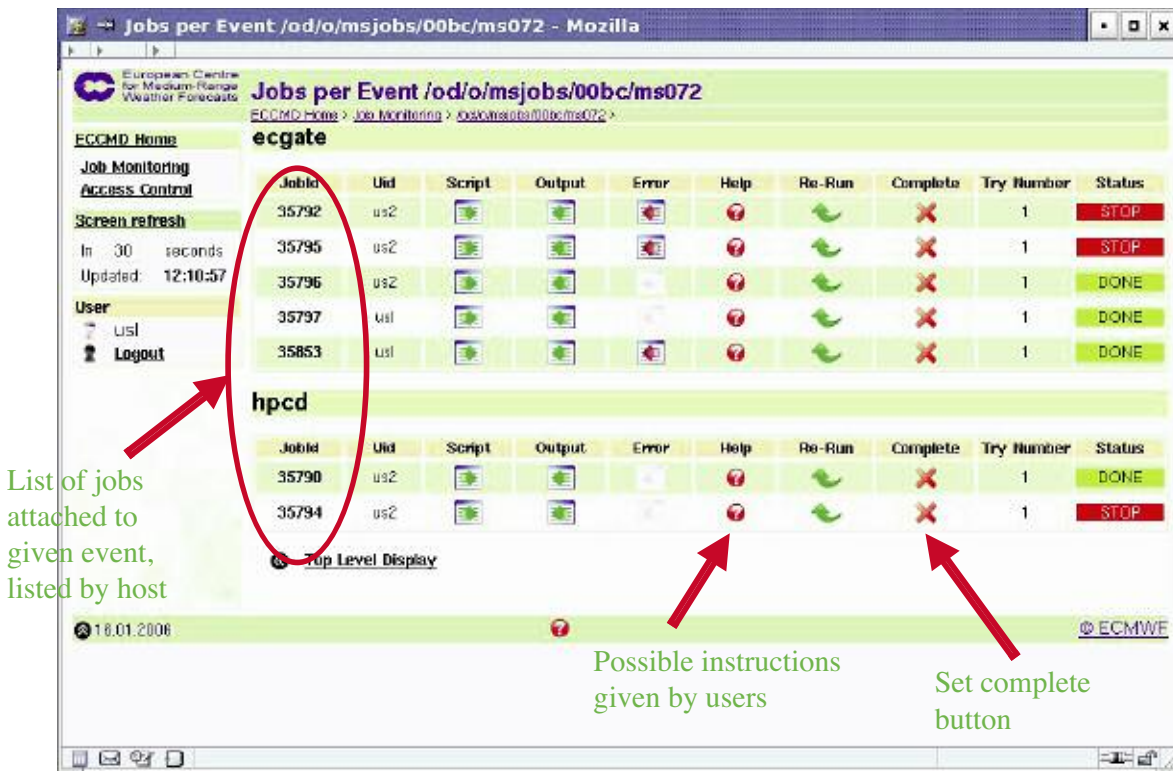
The screenshot shows a web browser window titled "Top Level Display - Mozilla" displaying a monitoring interface for the European Centre for Medium-Range Weather Forecasts (ECMWF). The interface is titled "Top Level Display" and includes a navigation menu on the left with options like "ECMWF Home", "Job Monitoring", "Access Control", "Screen refresh", "Filter", and "User". The main content area displays several tables of job status information, organized by directory paths. Each table has columns for "Day and Time", "Title", and "Jobs Status".

A red arrow points to the "Day and Time" column of the first table under the path "/od/law/lag/12/euroshelf", with the text "Date and time when last notification received for this event" written next to it.

Path	Day and Time	Title	Jobs Status
/od/law/lag/00/euroshelf	01 07:51:26	msjobs	2 DONE
	01 07:23:48	ms240	2 DONE
/od/law/lag/12/euroshelf	31 19:51:28	msjobs	2 INIT
	01 07:16:52	msmetgram	2 DONE
/od/ma/msjobs/12	01 05:31:17	ms144	2 DONE
	01 05:59:53	ms000	3 DONE
/od/mc/msjobs/00	01 08:26:07	ms072	7 STOP
	01 08:25:49	msplume	2 DONE
/od/o/msjobs/00bc	01 09:06:07	ms504	2 DONE
	01 08:49:34	msmetgram	2 DONE
/od/o/msjobs/06bc	01 08:48:52	ms240	2 DONE
	31 12:20:54	ms072	7 INIT
/od/o/msjobs/12	31 19:27:40	ms240	2 INIT
	31 19:17:17	msmetgram	2 INIT



**Operators' interface: monitoring (2)**



**Differences from existing system**

**ECaccess jobs:**

There are unique entities for each run. They are kept in the ECaccess pool.

Users can remove an "SMS" job by removing the job entity in STDBY mode.

Users can modify an SMS job by removing the existing one, then submitting the modified version.

Access to the HPCF system is available.

**Monitoring by operators:**

Errors will only be reported, if ECaccess is notified about errors in a job.

- Users need to notify loadleveler about errors in jobs.
- "set -e" should be used or correct exit codes should be returned, e.g. with "trap".
- Loadleveler will notify ECaccess about errors in jobs.

**Implementation schedule**

Mid June: upgrade of ECaccess at ECMWF (v3.1)

End of June: new ECaccess gateway and ETools (v3.1) will be available for distribution.

Early July: announcement of new MS job submission via SMS

End of 2006: closure of existing "MS jobs under SMS" service

ECtools are available on ecgate.

Requirements for remote usage (with remote gateways):

- New version of ETools
- New version of ECaccess gateway

H. de Vries asked whether computer operators would have time to comply with users' special instructions. *D. Lucas replied that it is assumed that most users will not give special instructions. By default the operators will set failing jobs to 'complete', unless there appears to be a general problem affecting more than one job.*

G. Wotawa asked whether the existing service would be terminated. *D. Lucas replied that the existing SMS submission would be replaced by this service. There were no plans to withdraw crontab. Users are, however, encouraged to switch to the new service as soon as possible.*

## User Registration Update – Paul Dando

### EMS – Entity Management System

EMS is a user registration system with a web-based interface.

The core system has been operational since December 2003 and has been used for all user registrations since then. A demonstration of the interface was presented at last year's meeting. Computing Representatives were given access to the system soon after it became operational.

It is now being used regularly by 13 Computing Representatives to register and manage their users. Any Comp. Rep. can use the system but must first check with User Support that access has been enabled.

### Actions possible with EMS

Registering new users

Modifying user info and access rights

– E.g., access to MARS current forecast data, hpcd, etc

On-line query of user info and access rights: up-to-date information is obtainable directly from the EMS database

Comp Reps CANNOT (yet !) use the system for

- Deregistering users
- Registering or deleting Special Projects or project accounts
- Registering new Section IDs
- Changing user quotas

Please contact User Support ([advisory@ecmwf.int](mailto:advisory@ecmwf.int)) for these operations

### Number of users registered

From 1 June 2005 to 31 May 2006:

Total of 512 'entities' registered

- 160 by Calldesk
- 173 by User Support
- 128 by Computing Representatives
- 51 by others

275 new, fully registered Member State users

27 new internal ECMWF users

Remainder are web-only or other users

413 self-registered web-only users

135 fully registered users were deregistered

### Roaming web access

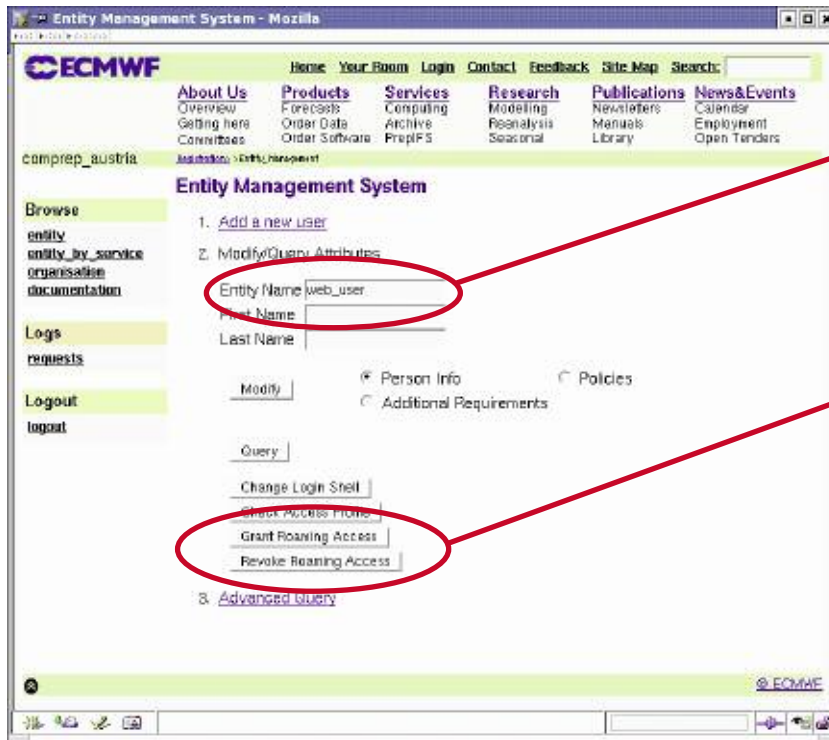
Some users only need access to the ECMWF web site, which wastes a SecurID card. These users will be registered for web-only access by means of a "roaming" password. No SecurID card is needed.

Password access to the web site is the same as with SecurID, i.e. users are able to access the same pages from work or home PC.

This service was announced last year as "almost ready" but technical problems delayed its release. These problems have now been resolved and the system is ready for implementation.

**Users need a SecurID card only for login access to hpcd or ecgate**

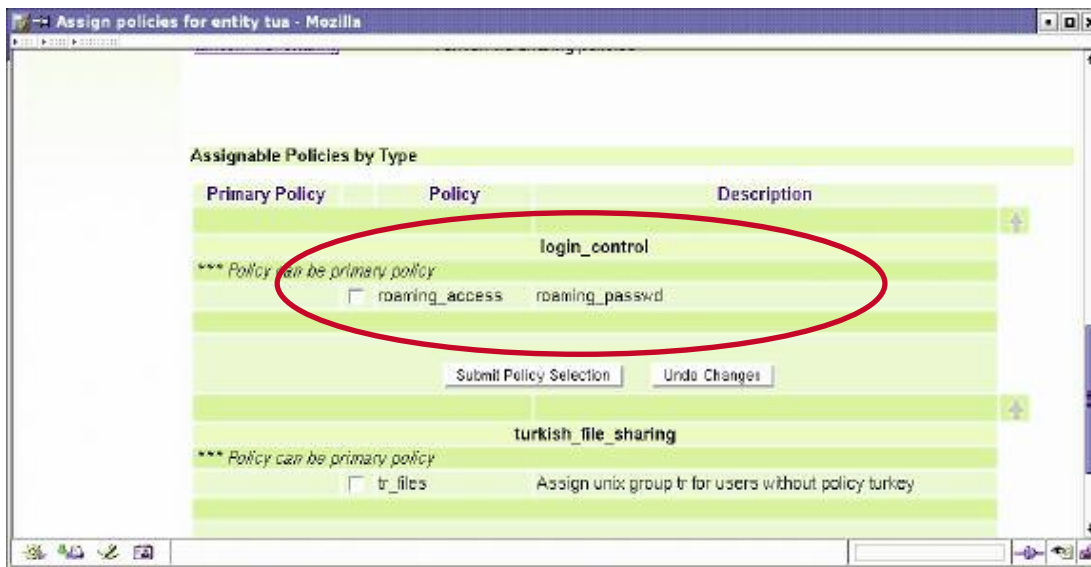
Giving an existing user roaming access



- Enter user name – can be a web-only or self-registered user
- Use “Grant Roaming Access” to give a roaming password
- Use “Revoke Roaming access” to remove

Registering a user with roaming access

Tick box for the `roaming_access` policy on the “Assign policies” page in EMS



## Paperwork

User registration forms can still be used in parallel. If a Comp. Rep. uses EMS to register users, there is no need to use a form. Forms should be sent to User Support, only if you want ECMWF to register users.

Current registration forms will be changed to reflect the new system.

Users will still need to contact the authorising organisation- access will be authorised by Comp. Reps. (as at present).

Instead of signing the SecurID declaration, users will have to acknowledge acceptance of the ECMWF Terms and Conditions on- line.

New User Packs – all information will be sent by e-mail or made available electronically.

## Who sends SecurID cards ?



- “Use a SecurID token from the local pool” – Comp Rep supplies SecurID card from their “SPARES”
- “Use a SecurID token from the ECMWF token pool” – ECMWF Calldesk supplies the SecurID card

## Future of SecurID cards

RSA SecurID has been used since May 1995

900 SecurID tokens will expire at end of April 2007 and would cost £36,000 (€54,000) to replace. Use of the RSA SecurID solution is currently being reviewed. Other solutions are being investigated:

- Tokens that last up to 10 years
- No “expiration” date
- Cost not based on token - e.g., licence per user

A decision will be made in October 2006, with implementation from October 2006 - February 2007. 900 spare tokens will be deployed before 28 February 2007.

## Reducing the number of SecurID cards

Currently we have

- 1,625 enabled SecurID cards
- 240 disabled SecurID cards
- 120 expired SecurID cards

We want to try to reduce the number of cards. We propose to remove expired cards but keep users registered in EMS (and any files on ecgate/hpcd). Comp. Reps. would have to use EMS to see these users - eclist would **not** show them.

Many users have not used their cards to access ecgate for some time. – **Are some of these users candidates for roaming passwords ?**

## Possible future developments

### *SecurID card maintenance*

We aim to make all user information available via the EMS web interface and phase out old SecurID admin commands. We will implement an EMS “eclist” in the near future. Other admin commands will be included in EMS when the new solution is implemented.

There will be a web-based form for the registration of new accounting projects (as requested by M. Pithon last year) and on-line acceptance of “ECMWF Terms and Conditions”.

## Summary

The web-based user registration system (EMS) is available for all Computing Representatives to use:

**<http://www.ecmwf.int/services/ems/d/registration/>**

Remember: You must first login with your SecurID card!

You can still send registration forms to ECMWF.

You must be given access, if you want to use EMS – check with your User Support contact point.

Roaming access is now ready for implementation – you will need to use EMS instead of eclist to see all your users.

M. Pithon asked whether EMS could be used to check other users’ quotas. *U. Modigliani replied no; the ecquota command could be used for this, although it is not a documented feature.*

H. de Vries mentioned that he would notify User Support of users to be deregistered more promptly, if he were able to signal this via the EMS system, rather than sending a fax. *I Weger stated that such a facility could be added to the ‘wish list’.*

H. de Vries asked whether ‘web only’ access included WebMars. *P. Dando replied that it did. Currently, users are not obliged to acknowledge the ECMWF Terms & Conditions but this may become obligatory, on-line, in the future.*

R. Sharp asked whether EMS could show a list of all disabled cards. *P. Dando replied it could not, at present, but a facility to do so will be introduced into EMS. In the meantime, ‘eclist’ could be used to list disabled users.*

## Survey of external users: status of follow-up actions – *Umberto Modigliani*

### Update on survey/report

Following the problem reported by France, all users at Meteo-France were sent another invitation to complete the survey.

38 replies (31%) were received from 120 invitations. The inclusion of this additional information did not change the original statistical results presented last year.

The report of the survey was published on the ECMWF website and all users were informed. The results of the user survey were also presented at the Technical Advisory Committee in October 2005.

We invited further feedback from users during visits to NMSs.

### Improving the information flow

Some users, in particular those outside the NMSs, wished to be kept better informed, especially regarding meteorological changes.

A specific effort was made to inform all users about the planned implementation of the higher resolution model upgrade:

- A web page was created and frequently updated;
- All users were informed of the plan and of the available information;
- Specific users were directly informed about issues of particular concern to them.

We continued to use ad-hoc mailing lists for specific issues.

Sympa mailing lists, open to all users, were recently created for GRIB, BUFR, CREX, and Interpolation software. A single registration is required for all subscribed mailing lists.

### Improving the documentation

Work to address the comments received about the quality of certain documentation has started.

#### *The ECaccess User's Guide*

has been rewritten in Latex, which allows much easier updating of both the HTML and PDF versions. It also facilitates a better document structure, which should improve its usability and searchability. The guide was also updated to include recent changes.

#### *The MARS User's Guide*

has been rewritten in Latex. It now contains a full index of MARS keywords. A first draft is available.

#### *The new web search tool*

should also help in finding the available documentation more easily.

We are planning to review the external web site, after we have completed the restructuring of the internal web site.

### Making more disk space available

- Since more disk space (~6 TB) was made available on /hpcd/tmp/ms/ no major shortages of space have been experienced. The select/delete process runs every day to remove files older than 3 weeks.
- More users have asked for and obtained specific space in /hpcd/ms\_perm.
- The quota for \$SCRATCH on ecgate has been increased for all users requiring more space:
  - The select/delete process is running less frequently than once a month.
  - We will soon increase /scratch/ms to 1 TB.
  - We will also increase the default quota of all users to 6 GB.
- The default quota on \$HOME will be increased to 100 MB
- Separate disk space can be made available for specific purposes.

### Improving the shell environment

- The user's login shell on ecgate can be changed via the EMS.
- Several users have been advised on how to set up working arrow keys, when using the Korn shell.
- Work has started on a new "Modules" based environment which will simplify the definition and usage of an application programming environment (compilers, libraries, tools) and make it easier to know which software packages are available. It is already used by other supercomputing centres and in DEISA and it is planned to be used on all ECMWF systems.

### Other actions regarding access to ECMWF

#### *SecurID card*

There is a project to investigate other means of supporting strong authentication.

#### *Bandwidth*

The Internet access line has been increased to 250 Mbit/s.

The RMDCN migration to MPLS will increase the available bandwidth.

The ECaccess review will address the possibility of differentiating and prioritising usage of the connection, allowing bandwidth reservation.

### ECFS enhancements

- The new **emove** command, which allows the renaming or moving of files/directories within the same ECFS domain, is available:

```
emove [-o | t | n | e] [ecdomain:]<src_file> [ecdomain:]<trgt_file>
```

It is not possible to move files between ectmp: and ec: domains

- **ecfs\_audit** and **ecfs\_audit.tmp** files are created once a month. They show the complete list of users' files in ec: or ectmp:

They also show if a secondary copy of the file has been created and allow detailed monitoring of the usage of ECFS.

Both these enhancements were described at the last Computer User Training.

T. Lorenzen asked how to sign up to the new BUFR and GRIB mailing lists. *U Modigliani replied that this could be done via the website. Go to: <http://www.ecmwf.int/products/data/software/index.html>, where all the ECMWF software packages are listed. Click on the package(s) you are interested in and you can join their individual mailing list(s).*



## ECMWF Web Search Revisited – Daniel Varela

### Current Search: HtDig

*The Good* It works, more or less (see User Survey)

Simple to set up: CGI process and configuration files

Extensive configuration options (what to index, URL substitution, etc)

*The Bad* The CGI process loads the indexes for every search; not thought to be very scalable

Very simple query language

Only indexes static content, not databases or any other source

*The Ugly* Monolithic design; difficult to modify

Project seems rather dead

### Finding something else

#### Google Search Appliance

– The top player in Internet search; a closed box, ready to use

– but is Internet search technology applicable to single web sites without customization?

#### Another Open Source search engine

– Many new engines are based on Lucene.

– Some are J2EE WebApps, which could be deployed in our current Tomcat installations:

– The following applications were assessed:

- Regain
- Nutch
- Compass

### Lucene and Nutch

**Lucene** is just an API. It provides:

– An index creation API. Indexes are files with searchable fields.

– A search API that uses the index files.

word "some phrase" -butnotthis field:onlyinthisfield thisismoreimportant^4

**Nutch** is a full Search System, based on Lucene. It provides:

– A Crawler: a program that explores web sites and gets content.

– A Content Extraction Framework.

– A Search Interface; the actual page with the search box and the results listing.

– Tools to analyze results: it is very important to understand WHY a result appears in a given position.

### Other features of Nutch

Nutch can scale from Intranet/Single Domains to the whole internet. It can be clustered.

It has a Link Graph information gatherer. This is similar to Google's PageRank technology.

It has a modular design. It is possible easily to add support for new file formats or change the way search terms are interpreted, without touching the main source code.

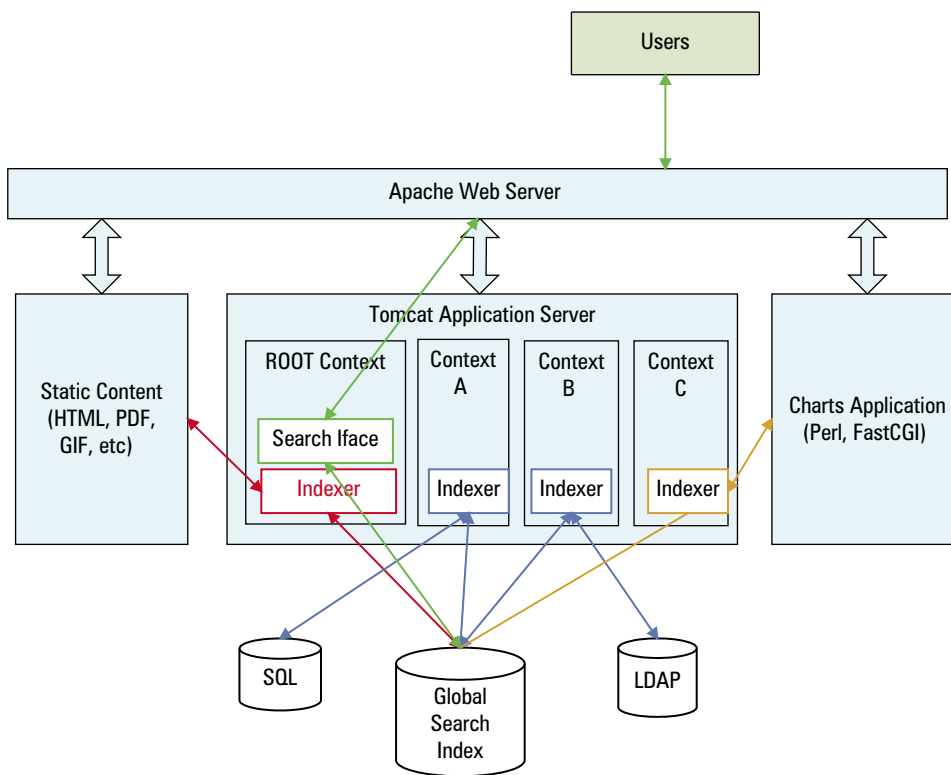
It is supported by Yahoo Research Labs. They currently employ Nutch's lead architect.

**Building on top of Nutch**

The ECMWF implementation:

- assigns special weight to the URL field. At ECMWF, it is a very good indicator of the page content;
- has an improved Search Page, with better pagination features;
- has integrated database indexing (SQL, LDAP);
- has crawlers integrated into the web applications, so that they are completely self-contained; it just needs a Tomcat installation and the .war file.
- It has integrated Lucene SpellChecker. This is similar to Google’s “Do you mean?” and suggests another query when the original one is not successful.
- It has Query filters: Authorization, sub-area search, preset results for matched keywords.

**Runtime structure**



**TO DO**

- Improve indexer to try to cope with the most common user editing errors (e.g. lack of title in a page)
- Avoid indexing common areas (header, footer, menus) in the page
- Search graphic charts catalogue
- Search “tuning” as an iterative process:
  - Gather data from users (logs, etc)
  - Analyze it (Results for the searches? Are most adequate?)
  - Apply information to indexation/search.

### For more information

Lucene: <http://lucene.apache.org/>

Nutch: <http://lucene.apache.org/nutch/>

Introduction to Text Indexing with Lucene: <http://www.onjava.com/pub/a/onjava/2003/01/15/lucene.html>

Advanced Text Indexing with Lucene: <http://www.onjava.com/pub/a/onjava/2003/03/05/lucene.html>

Introduction to Nutch, Part 1: Crawling: <http://today.java.net/pub/a/today/2006/01/10/introduction-to-nutch-1.html>

Introduction to Nutch, Part 2: Searching: <http://today.java.net/pub/a/today/2006/02/16/introduction-to-nutch-2.html>

Search and You May Find (Jakob Nielsen's July 1997 Alertbox, on Search Usability):

<http://www.useit.com/alertbox/9707b.html>

Applying Lucene to the Web: <http://www.nutch.org/twiki/Main/Presentations/serverside2.pdf>

Intranet Search With Nutch: <http://www.nutch.org/twiki/Main/Presentations/serverside1.pdf>

Did You Mean: Lucene? <http://today.java.net/pub/a/today/2005/08/09/didyoumean.html>

### Conclusions

- It is possible to find high quality Open Source search products.
- Good quality search in Intranets or single domains/sites depends on customization and "tuning".
- Achieving consistently high quality is an iterative process.
- Planned deployment: 5th July. It is available for testing beforehand.

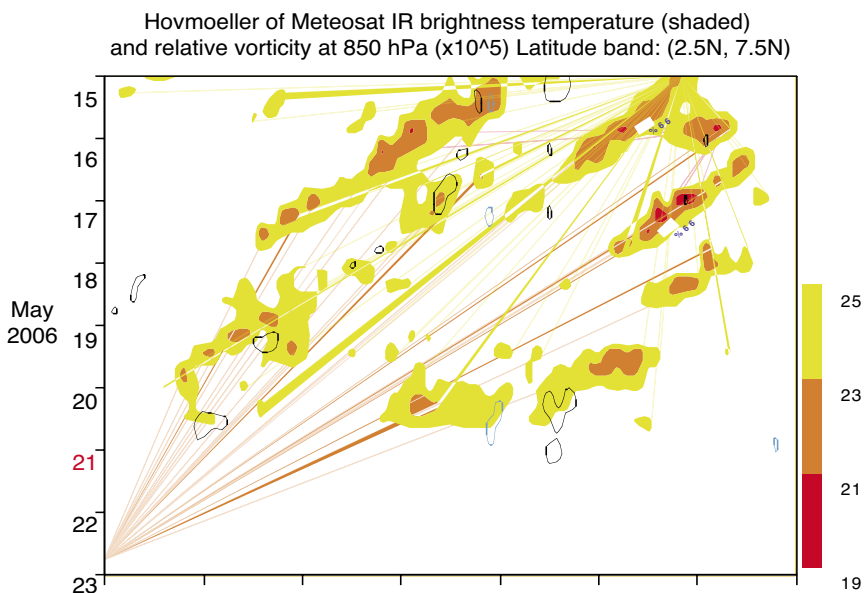
R. Urrutia asked whether a content management system was planned and, if so, what kind. D. Varela replied that it had been developed and customised internally and will be indexed by the 'search'.

## Graphics Update - Jens Daabeck

### Metview Development

- There is continuing routine effort to incorporate new facilities into Metview. This includes constant improvements to
  - handle new data types, e.g. Metview Macro operations on fields stored in polar stereographic and Lambert grids
  - respond to new requirements from the meteorological community, such as support for "harmonized access times"
  - support new BUFR input into the TimeSeries application to plot time series from dissemination weather parameter products
  - provide extended satellite image support
  - handle Embedded Fortran routines in Metview Macro
  - provide classic Metgram (Deterministic Model) support for RD experiments.

### Hovmoeller Diagram of Meteosat IR Brightness Temperature



### Embedded Fortran Routines in Metview Macro

```
File Edit Search Preferences Shell Macro Windows Help
/metview/macro_tutorial/prop/gradient_ex 325 bytes L: 17 O: 0
extern gradienth(f fieldset) "fortran" inline
PROGRAM GRADIENTH
PARAMETER (ISIZE=50000)
C GET FIRST ARGUMENT AS A FIELDSET
CALL MOETO(IORIS1, ICNT)
end inline
-----
q = extractv(param "q")
q = gradienth(q)
```

- New software facilities have been added to Metview (such as use of Perforce for source code management, title information from a common database, debugger choice).
- New features have been added to Metview to meet user requirements, including several Metview Macro functions:
  - subsample, geosort, univertint, addmonths, mgetv, msetv, float, indices, sortindices, offset
- Routine effort also includes training courses, software maintenance and support for both internal and external Metview users.

### **Metview Releases**

The latest internal Metview version is 3.8.5, which is based on Magics 6.10 and Emoslib cycle 000300.

Metview 3.8-export, released 30.8.2005, supports the High Resolution Forecasting System.

Platforms:

- Linux: SuSE 9.1      – IBM: AIX 5.1
- SGI: IRIX 6.5        – HP: HP-UX B.11.00
- Sun: SunOS 5.7

### **Plans**

Work on using the `grib_api` within Metview has started.

A new application in Metview, called ToolPlus, has been developed as a first attempt to integrate Metview with Magics++. This new facility explores Metview's current features and combines them with new ones provided by Magics++, building a framework for the execution of interactive user tasks.

A new visualisation module is currently under development to take advantage of all the benefits provided by Magics++, including:

- improvements in user interactivity, such as legend/text handling
- the ability to dynamically toggle certain graphical objects, such as contour labels and maxima/minima
- querying the properties of graphical objects displayed on screen.

The first prototype aims to execute current operational Metview Macros

### **MAGICS**

The maintenance of the current operational MAGICS has continued. The changes for the switch to the High Resolution Forecasting System were successful.

The latest internal Test Version is 6.11

The currently available Export Version is 6.10, released 5.10.2005.

Platforms

- Linux: SuSE 9.1      – IBM: AIX 5.1
- SGI: IRIX 6.5        – HP: HP-UX B.11.00
- Alpha: OSF1 V5.1    – Sun: SunOS 5.7

### **Summary**

#### ***Metview***

- Metview 3.8 Export Version
- Next Export Version 3.9 planned for 4Q2006

#### ***MAGICS***

- MAGICS 6.10 Export Version
- Next Export Version 6.11 planned for 3Q2006

In regard to embedded Fortran routines, T. Lorenzen assumed that Metview would not contain a complete Fortran compiler and asked whether there was a particular compiler which users were recommended to use. Also, was it limited to Fortran? J. Daabeck replied that currently Portland Group Fortran 5.2-4 is used, so this version is guaranteed to work. It seems that some Fortran system library names have changed in later versions and this requires some editing of the config/config.linux.2 file. (Metview modules are linked using a C++ compiler which doesn't know about Fortran libraries needed and thus the required libraries must be explicitly stated). Currently, there is only a Fortran version, as there was no user demand for other versions. Should users' needs change, then other languages could be supported. H. de Vries commented that if GRIB and BUFR decoders became available in C, then his users would very much appreciate support for C in Metview. This will be investigated by Graphics section. (A detailed reply was sent to T. Lorenzen after the meeting.)

## Technical Overview of Magics++ – Stephan Siemen

### MAGICS 6.x

Meteorological Applications Graphics Integrated Colour System has been in operational use since 1984. It is used by Member States and many other weather services. It is also used in Metview and Synergie as the graphics engine. Hundreds of plots are produced daily for Met Ops and thousands of plots for the Web, with demand growing constantly.

### Why redesign MAGICS?

20 years of code history / legacy

To take advantage of recent developments in software engineering

To adapt to changing work practice : from paper to web usage (e.g. formatting text)

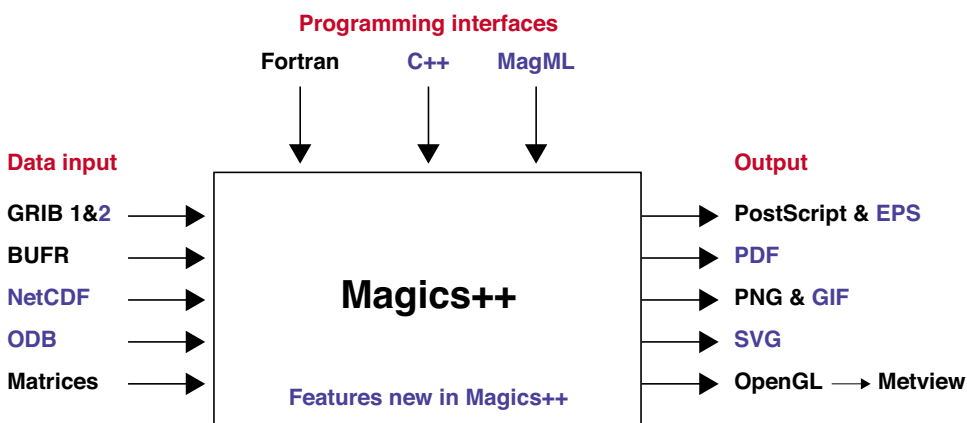
To create new architecture which allows more interactivity for users with Metview and formats such as SVG

To achieve independence from CONICON (licences needed)

MAGICS 6.x way of handling pointers is not working on IA64/AMD64 bit platforms

**but backwards compatibility will be assured!**

### Overview



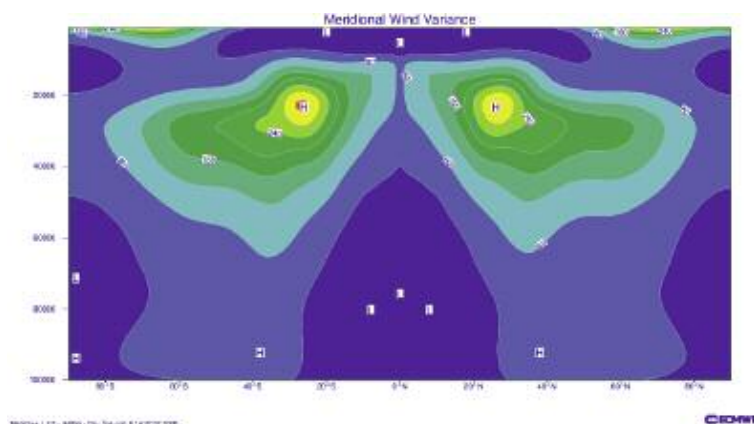
### New decoders

Thanks to ECMWF Grib\_API, there will be better support of GRIB 1 and new support of GRIB 2

NetCDF is supported in Magics++

Magics++ now uses the same code as Metview to decode BUFR

New decoders can easily be added

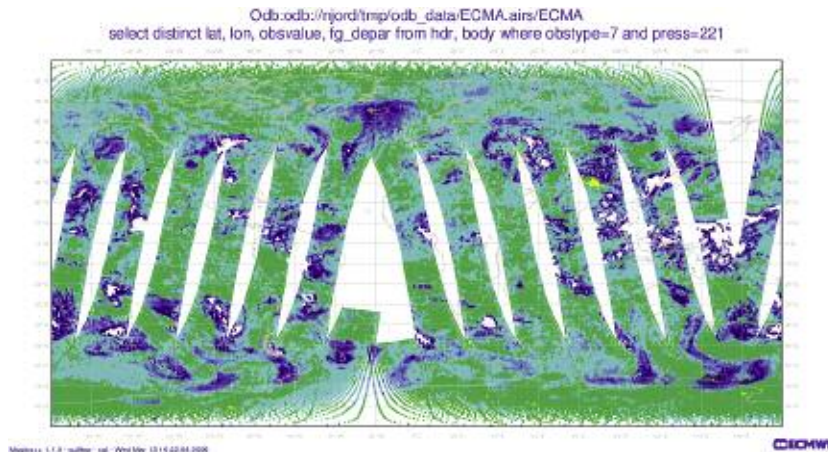


## Bindings to databases

ODB support

SPOT database is used to produce EPSgrams

Architecture allows easy extension to read data from other (database-like) sources



## Programming interfaces

Definitions of Magics++ plots can be written in Fortran and MagML (XML).

Software currently using MAGICS (e.g. Metview, Synergie) can use the new C++ interface.

Procedural (Fortran/C) and object-oriented (C++, MagML)

The simple API for Fortran with its parameter concept stays, with only a few default values changed.

The aim is to plot meteorological data as simply as possible with meaningful automatic scaling and titles.

## MagML

MagML is an XML based format to describe Magics++ plots. It is descriptive, not procedural. There is no need for (re-)compilation.

A *MagML\_Interpreter* program processes a MagML file.

An Interpreter can easily be called in user code.

The description is very different from Fortran API - more closely reflecting the internal structure of Magics++. It can be integrated into more complex XML request descriptions. It is also used in the *Plot-On-Demand* project.



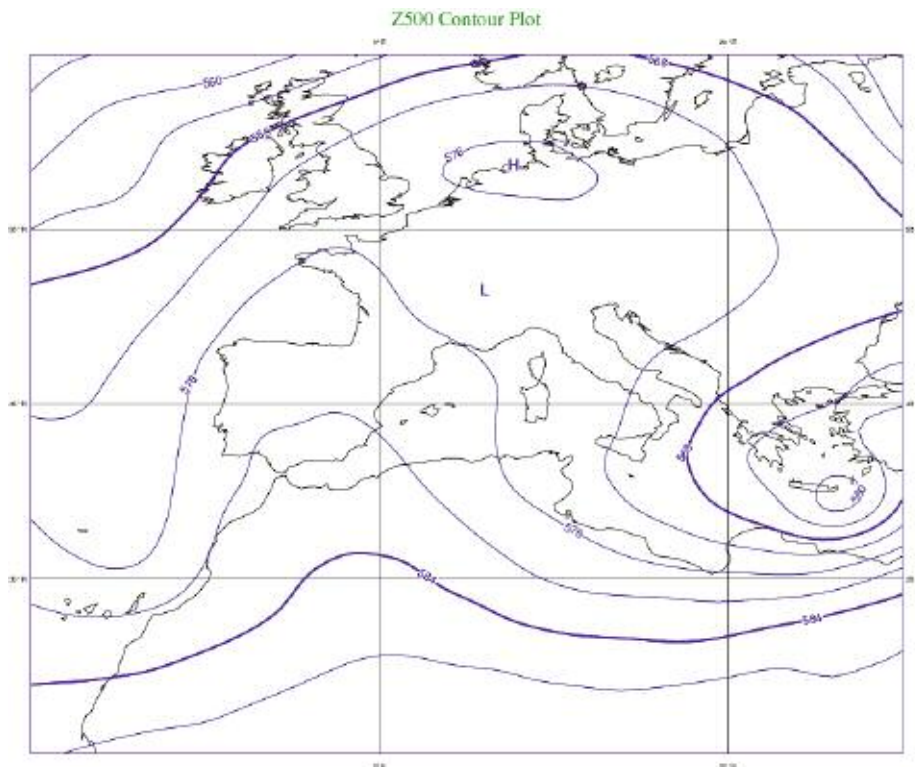
## MagML example- the code

```

<magics>
  <page format ='a4' orientation = 'landscape' >
  <subpage>
    <mapview>
      <cylindrical>
        <corners min_longitude = '-20'
          min_latitude = '20'
          max_longitude = '30'
          max_latitude = '60' />
      </cylindrical>
    </mapview>
    <coastlines>
      <coast colour = 'black' />
    </coastlines>
    <layer>
      <grib path = 'data/z500.grb' />
      <geocontour />
    </layer>
    <text font = 'Times-Roman' colour = 'rgb(0.1,0.6,0.2)'>Z500 Contour Plot</text>
  </subpage>
</page>
</magics>

```

## MagML example- the output



**Output formats**

The new object-oriented architecture allows easy implementation of the new output drivers. In a single run Magics++ can produce multiple output formats to save computing time (calculations are only performed once).

The Postscript driver is much improved and also supports EPS and PDF.

The Raster Driver (based on GD) now also supports GIF, animated GIF and transparent backgrounds in GIF/PNG (and is much faster than with ImageMagick's convert).

OpenGL output has been improved to enable interactive work in Metview.

SVG support has been added.

There is better support of fonts.

User defined logos (in GIF and PNG) can be added.

**Future benefits for Metview**

Magics++ and Metview will both use C++

More input and output formats

More interactivity for users

- toggle contour labels, shading, HiLo
- change quality of contours/coastlines

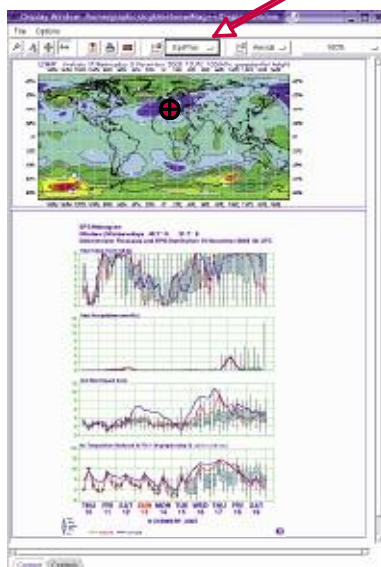
Better display through the new OpenGL driver

- higher quality text
- improved import of graphics

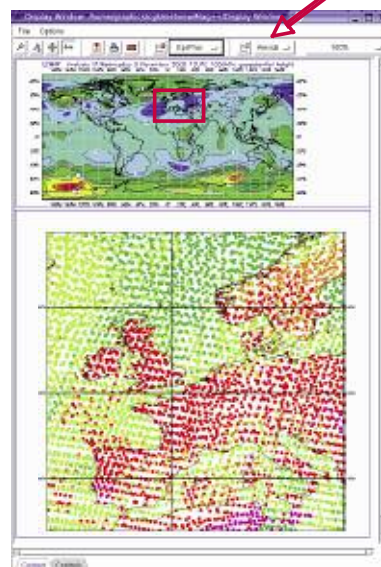
Generate MagML for plot-on-demand (web)

**Magics++ in Metview - ToolPlus**

Point selection



Area selection



## Installation

Now based on autotools (configure)

It checks for dependencies and flags errors

Flags to disable and enable features (e.g. -enable-odb)

There is the possibility of installing static and shared library versions

It requires a recent C++ compiler (stdC++98)

It can be installed in parallel to Magics 6.x

## Availability

It is supported on the following platforms at ECMWF:

- Linux on x86, AMD64 (32 and 64bit!) using g++ 3.3 & 4.0
- AIX 5.x using g++ 3.3

Other platforms will be considered on request

Magics++ 2.0 will be released to Member States at the end of 2006

## Conclusion

Magics++ is used operationally for EPSgrams

Testing has started for Plot-on-Demand, verification plots and Metview Macro

New documentation will highlight new features

For updates please visit our external webpage at <http://www.ecmwf.int/publications/manuals/magics>

H. de Vries asked whether it was straightforward to make changes, such as adding data formats. S. Siemen replied that it was hoped to document the processes, so that users could also make their own changes, but he didn't yet know when the documentation would be complete.

T Lorenzen asked whether MAGICS++ was available on ecgate. S. Siemen replied that a test version was currently available on ecgate and users were encouraged to test it and pass comments to the MAGICS group.

T. Lorenzen asked whether MAGICS++ produced any performance gains compared to MAGICS. S. Siemen replied that performance increases had been seen, for instance with GRIB API, but features were still being added and optimisation would take place once the code was fairly complete.

P. Halton asked that a module on MAGICS++ be added to the training programme next spring. S. Siemen replied that in the last two years 1/2 a day of the 2 day MAGICS module had been set aside to introduce new features in MAGICS++. Next year, it was planned to increase the coverage.

**Introduction to the new GRIB API – Enrico Fucile**

**GRIB edition 1 vs. 2: Different Structure**

GRIB 1	GRIB 2
<b>SECTION 0</b> Indicator	<b>SECTION 0</b> Indicator
<b>SECTION 1</b> Product definition	<b>SECTION 1</b> Identification
<b>SECTION 2</b> Grid Description	<b>SECTION 2</b> Local Use
<b>SECTION 3</b> Bitmap	<b>SECTION 3</b> Grid Definition
<b>SECTION 4</b> Binary Data	<b>SECTION 4</b> Product Definition
<b>SECTION 5</b> End	<b>SECTION 5</b> Data Representation
	<b>SECTION 6</b> Bitmap
	<b>SECTION 7</b> Binary Data
	<b>SECTION 8</b> End

**GRIB edition 1 vs. 2: Different units**

*GRIB 1*

LatitudeOfFirstGridPoint is in **milli**-degrees

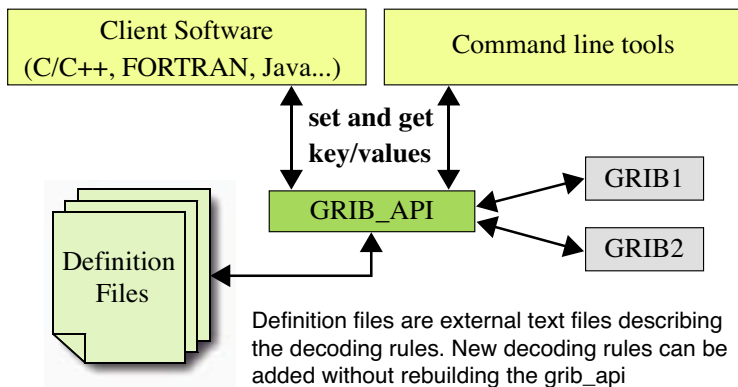
*GRIB 2*

LatitudeOfFirstGridPoint is in **micro**-degrees

or in an arbitrary unit, defined by the fraction of two parameters:

`basicAngleOfTheInitialProductionDomain/`  
`subdivisionsOfBasicAngle`

**GRIB\_API – Design**



**GRIB\_API - Keys**

*Required (set and get access)*

- Complete set of keys describing a GRIB 1 or 2 message
- Same names in both editions
- Not always mapped directly to octets
- Higher abstraction level provided for end-user access

**Optional (set and get access)**

- Closer to binary coding rules
- Usually mapped directly to octets
- Requires higher level of expertise

**Read only (get access only)****GRIB\_API - Key examples****Required (set and get access)**

```
latitudeOfFirstGridPointInDegrees
jDirectionIncrementInDegrees
editionNumber
```

**Optional (set and get access)**

```
latitudeOfFirstGridPoint (edition dependent)
numberOfPointsAlongAMeridian (calculated)
```

**Read only (get access only)**

```
identifier = "GRIB"
end = "7777"
```

**GRIB\_API - Definition Files**

octets

Keys → decoding algorithm  
encoding algorithm

Flow control ( if (condition) {...} else {...})

Aliases, include, templates, transient

Text files. **No link or build needed**

**GRIB\_API - Basic features****Create a new grib from template:**

```
h=grib_handle_new_from_template(c,"LatLonGrib1")
```

**Load a grib from file:**

```
h=grib_handle_new_from_file(c,inFileHandle,&err)
```

**Dump content to a text file**

```
grib_dump_content(h,fout,"serialize",flags,NULL)
```

**Create a new grib from dump file:**

```
h=grib_handle_new_from_dump(c,"myGrib2.txt")
```

**Set values through keys:**

```
grib_set_double(h,"latitudeOfFirstGridPointInDegrees",10)
```

**Get values through keys:**

```
grib_get_long(h,"numberOfPointsAlongAParallel"
&numberOfPointsAlongAParallel)
```

**Iterate through gridded data values:**

```
i=grib_iterator_new(grib_handle);
while (grib_iterator_next(i, &latitude, &longitude, &value))
{
    do something with latitude, longitude, value
}
```

**GRIB\_API – Interfaces / Platforms****Interfaces**                      **Supported Platforms**

Compiled	Linux
C/C++	AIX
FORTRAN 77 / 90	HP-UX
JAVA (JNI)	Cygwin

**Scripting**

Perl  
Python  
PHP

**GRIB\_API – Command line tools*****grib\_dump***

- Dump grib file content to a text file as key / value pairs
- Options: `—read_only`, `—optional`, `—values`

***grib\_gen***

- Generate a grib file from a dump file and a data values file
- After a dump is produced, it can be edited and modified, before a new grib is generated using `grib_gen`

***grib\_set***

- Set values through keys

***grib\_get***

- Get values through keys

**Conclusions**

GRIB\_API will provide transparent read and write access to GRIB 1 and 2 for expert and novice users.

Great effort is being taken to provide a library which always produces well formed messages.

Using external text definition files to describe the decoding workflow gives the flexibility to change decoding rules without rebuilding or linking. This is particularly useful for adding new templates.

Alpha version 0.8.2 can be downloaded from [www.ecmwf.int](http://www.ecmwf.int). Version 0.8.3 will be available within a few weeks

A release version, including command line tools and C/C++, FORTRAN interfaces, is planned for September.

H. de Vries asked whether there was an interface to the old emolib GRIB handling routines (GRIBEX). E. Fucile replied that the philosophy and interfaces were completely different, so it would be impossible to map the new to the old.

T. Lorenzen asked whether users of the old emolib routines who wanted to use GRIB2 would be supported. E. Fucile replied that it would be impossible to modify emolib. B. Raoult stated that it was planned to split emolib into 3 levels; the first level would be GRIB API. GRIBEX will be maintained for as long as it is needed but it will not be migrated to GRIB2. The Data and Services section may be able to provide a GRIB2 to GRIB1 converter.

R. Rudsar asked whether ECMWF could add an 'ec' prefix to environmental variables which it added, to show that they were part of ECMWF libraries. U. Modigliani replied that it could inconvenience existing users to change current names. R. Rudsar suggested that the rule apply to new environmental variables only.

## Responses to comments made during the course of the meeting

H. de Vries' users would like MARS support for NetCDF. B. Raoult replied that this would be quite complex. His section would maintain (on a best efforts basis) a post-processing tool within MARS to convert GRIB retrievals from MARS into NetCDF.

In regard to the reports of slow ECFS transfers via ecgate in January, N. Storer encouraged Computer Representatives to send comments, complaints or queries as soon as problems arose, rather than waiting until this meeting to report them.

## Organisation of the meeting

Several representatives expressed a preference for holding Comp. Rep. meetings back-to-back with Security Rep. meetings. This would facilitate attendance for some representatives whose NMSs had travel restrictions. Considering that recent meetings had always been rushed, the majority of Representatives present agreed that a 2 day meeting over 3 days would be preferable, starting at 14.00 hrs on day 1 and ending at midday on day 3. This would also allow Representatives plenty of time for informal discussions with fellow Representatives or ECMWF staff.

## NEXT MEETING

It was agreed that the next meeting take place in spring 2007, back to back with the Security Reps' meeting.

## **PART II**

### **Member States' and Co-operating States' Presentations**



AUSTRIA

AUSTRIA

Cornelia Hammerschmid – Zentralanstalt für Meteorologie und Geodynamik, Austria

Operational ECMWF - data in Austria

The operational ECMWF-Data for Austria from the 00Z and 12Z model run are transmitted to Austria by using:

- ECMWF Dissemination System (most products) : 746 Grib files and 4 Bufr files
- MSJ-Jobs : 9 EPS-Meteograms and Spaghetti files as Postscript files
- simulated satellite images from MARS
- monthly forecasts from MARS (only every Thursday evening)
- seasonal forecasts from MARS (day 15 of the month)
- special products are downloaded from the internet (e.g. EFI-Forecasts)

Zentralanstalt für Meteorologie und Geodynamik

Overview: Operational DataStream's in the ECMWF Dissemination System

1. Data from deterministic model T799
  - Grib data from area Europe/North Atlantic with 1.5 deg grid distance
  - High-resolution grib data from T799 model, area Central Europe
  - Global data from T799 model (restricted parameters)
  - Grib data for a grid point in Northern Italy
  - Grib data on Model Levels, area Europe/North Atlantic

Zentralanstalt für Meteorologie und Geodynamik

Overview: Operational DataStream's in the ECMWF Dissemination System cont.

2. Data from Ensemble Forecast System (EPS)
  - EPS-products for the area Europe and North Atlantic
  - Ensembles for the area Central Europe
  - precipitation forecasts from EPS for Austria
  - Ensemble forecasts for Regional Weather Service Northern Italy
3. Data Wave Forecast System
  - wave forecasts in the mediterranean area
  - extended wave data (additional North/East Sea and Biskaya)

Zentralanstalt für Meteorologie und Geodynamik

## AUSTRIA

## AUSTRIA

Survey of data processing (ECMWF Products) ComptRep  
8-9 June 2006

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

4. Operational use by the Austrian public weather services

- ZAMG : Central Institute for Meteorology in Vienna and the 4 Regional Departments in Innsbruck, Salzburg, Klagenfurt and Graz
- MWD : Military Weather Service of the Austrian army and the air force
- ACG : Civil Aviation Weather Service (Austro Control)

Zentralanstalt für Meteorologie und Geodynamik

Survey of data processing (ECMWF Products) ComptRep  
8-9 June 2006 **cont.**

5. Scientific purposes (special projects) - actual data from dissemination and archived MARS-Data

- Internal use at ZAMG:
  - + Model Working Group (e.g. for the project "verification")
  - + Remote Sensing Group (e.g. combining with satellite data)
  - + Environmental Department (e.g. for computing trajectories)
- University Institutes in Vienna, Innsbruck, Graz

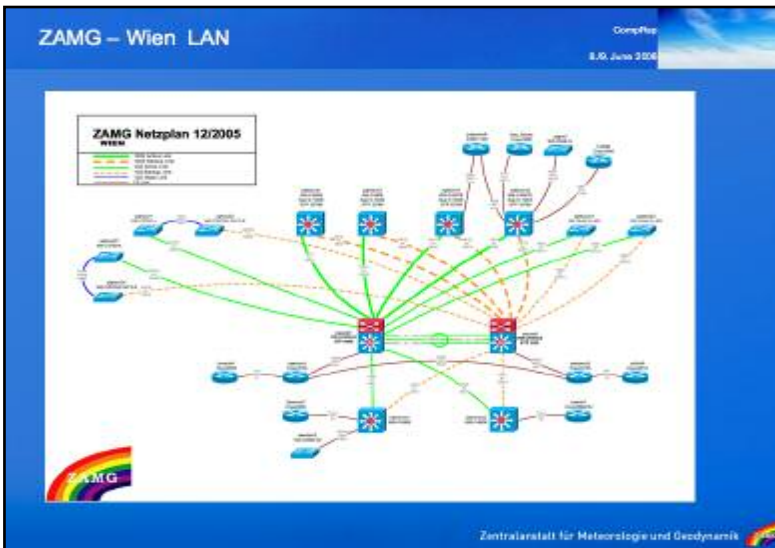
Zentralanstalt für Meteorologie und Geodynamik

Survey of data processing (ECMWF Products) ComptRep  
8-9 June 2006 **cont.**

6. Private and public customers (only derived products), e.g.

- ORF - Austrian Broadcasting Corporation
- local authorities
- newspapers
- organizers of sport and culture events
- tourist traffic offices
- street services (esp. snowfall and freezing rain)
- environmental purposes
- electric supply companies (forecasts of precipitation and temperature)
- warning of extreme weather situations (placed in Internet) such as strong wind, extreme precipitation amounts, thunderstorms, icing conditions

Zentralanstalt für Meteorologie und Geodynamik

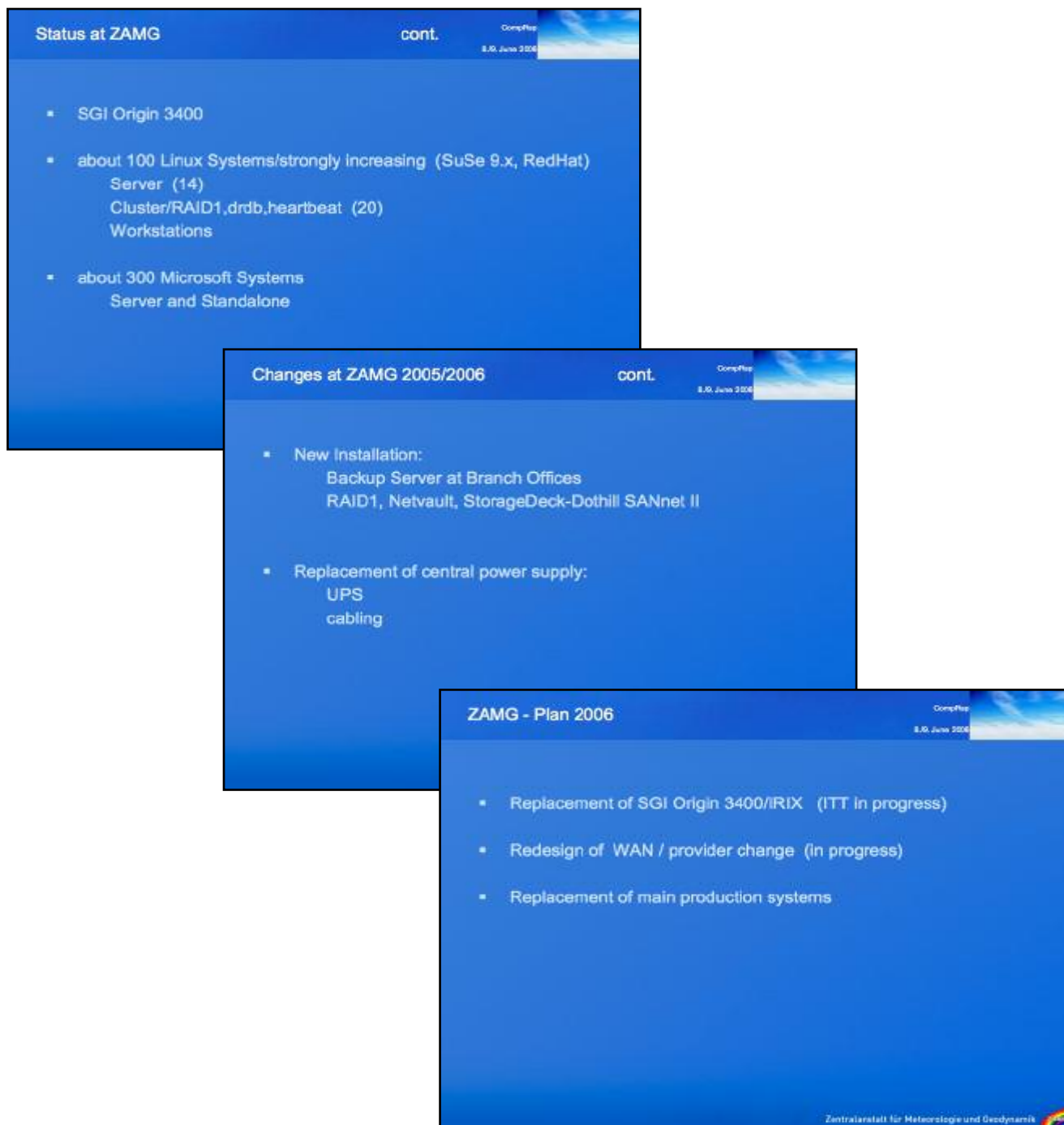


- Changes at ZAMG 2005/2006
- Comptel  
8.9. June 2006
- Couple of SUN/Solaris systems replaced by LINUX/Suse 9.x
  - New Installation: LINUX/Suse 9.x Systems  
LINUX/Suse 9.x Systems - Cluster/RAID1  
LINUX RedHat 9
  - Replacement of central Backup Server:  
SUN V250/Solaris 9, Veritas Backup 5.1  
Quantum DX100 Virtual Tape Library (4 TB)  
Quantum M2500 LTO Tapes (200/400GB)
- Zentralanstalt für Meteorologie und Geodynamik

- Status at ZAMG
- Comptel  
8.9. June 2006
- about 60 SUN Systems  
Enterprise 420r, SunFire V65x  
Ultra 10, Blade 1500, Netra T1  
StorEdge 6300, SunRack 900
  - Sun V250  
Quantum DX100  
Quantum M2500
  - NetApp Filer (Cluster)
- Zentralanstalt für Meteorologie und Geodynamik

## AUSTRIA

## AUSTRIA



**Status at ZAMG** cont. 8.10. June 2006

- SGI Origin 3400
- about 100 Linux Systems/strongly increasing (SuSe 9.x, RedHat)
  - Server (14)
  - Cluster/RAID1, drdb, heartbeat (20)
  - Workstations
- about 300 Microsoft Systems
  - Server and Standalone

**Changes at ZAMG 2005/2006** cont. 8.10. June 2006

- New Installation:
  - Backup Server at Branch Offices
  - RAID1, Netvault, StorageDeck-Dothill SANnet II
- Replacement of central power supply:
  - UPS
  - cabling

**ZAMG - Plan 2006** 8.10. June 2006

- Replacement of SGI Origin 3400/IRIX (ITT In progress)
- Redesign of WAN / provider change (In progress)
- Replacement of main production systems

Zentralanstalt für Meteorologie und Geodynamik

*In reply to C. Hammerschmid's comments on ECFS downtime, N. Storer explained that HPSS 5 uses DCE as its authentication service. Certain tables in DCE become filled up and do not get cleared automatically. HPSS has to be taken down in a weekly system session to allow DCE to be cleared, thus ECFS is not available during this time. HPSS 6.2 does not use DCE for authentication, so the weekly ECFS downtimes will not be needed.*

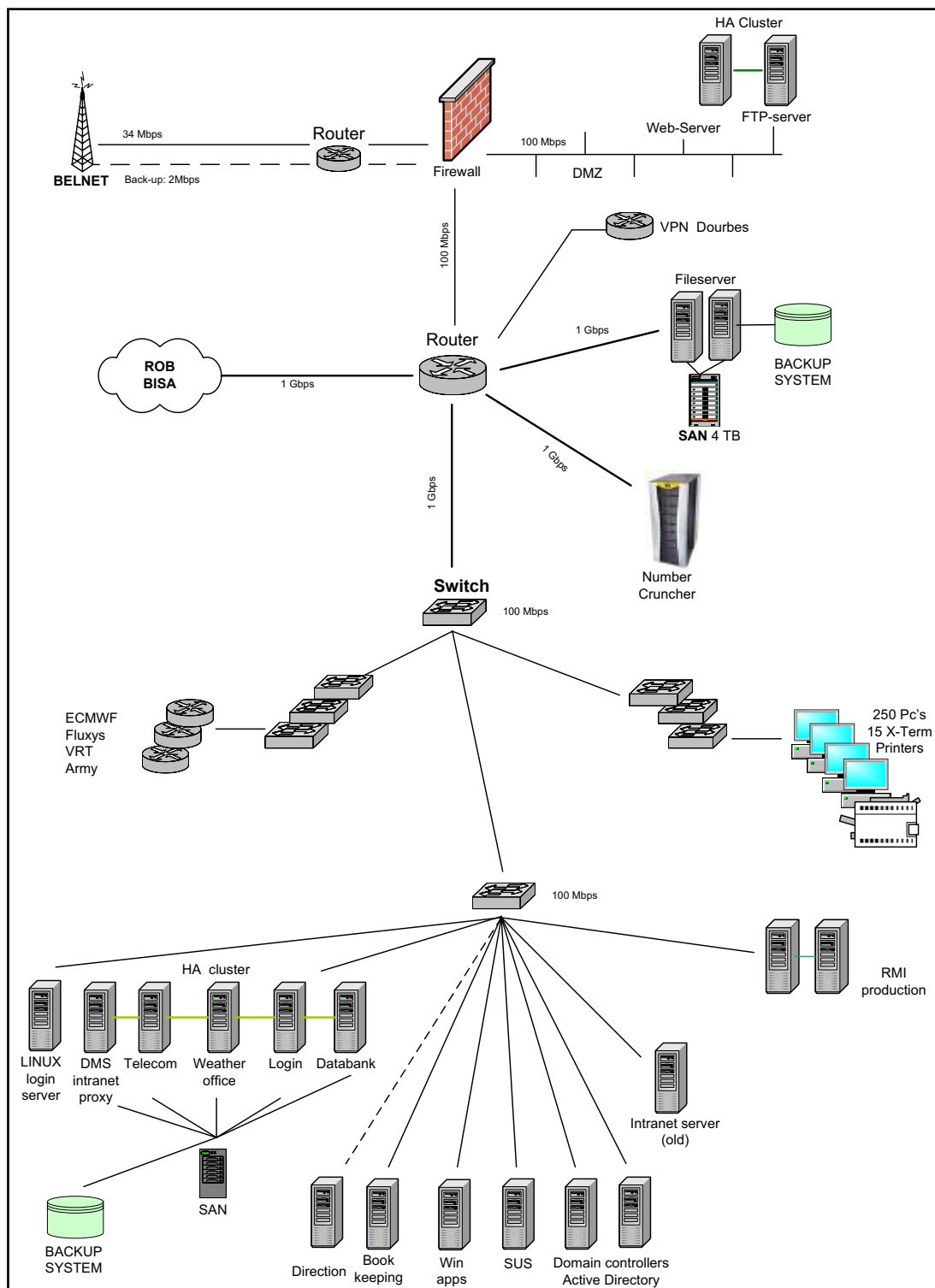
*In regard to the need to resubmit some SMS jobs from Austria recently, U. Modigliani noted that since the recent implementation of the higher resolution model many SMS-triggered jobs are taking longer than they used to, so users need to adjust the time limits specified in the jobs.*

*Also the load on Ecgate has increased significantly, leading to variable run-times. Wall clock limits specified by users should allow a buffer to take account of run time variations.*

Liliane Frappez – Royal Meteorological Institute, Brussels

**Computer infrastructure evolution at the RMI of Belgium**

Three institutes located on the site of Plateau d’Uccle (the Belgian Institute for Space Aeronomy (BISA), the Royal Observatory of Belgium (ROB) and the Royal Meteorological Institute RMI) share some common servers: file server, intensive computing server, external communication server (FTP, Web). Each institute also has a local network with specific infrastructure.



## BELGIUM

## BELGIUM

### In the shared infrastructure:

- the common compute server has been replaced by a SGI Altix 3700 BX2 (56 Itanium 2 processors, 2 GB per cpu, ~2 TB disk space). It is running Suse Linux Enterprise Server 9 (SLES 9) with PBS Pro 7 as its queuing system. The compilers are Intel. The operational ALADIN forecast is run since recently 4 times per day on this server.
- Another “cluster” of 3 AMD-Opteron machines is installed at BISA: each node has 4 cpus with 2 GB/cpu and works under Linux. These machines are intended for big or long mono cpu jobs such those running IDL, matlab,...
- The file server is an HP FC-60 with 2 TB of storage, with 1 VA7400 with 38 disks, each of 73GB, giving a total storage of 4.9 TB. Its contract comes to an end this year. The invitation to tender will soon be issued.
- In the DMZ, the Web server has been replaced by a Dell Power Edge server running Linux and it now forms a high availability cluster (heartbeat) with the FTP server.

### The computer infrastructure specific to RMI is mainly based on HP servers with the progressive introduction of Linux servers:

- HP 9000/L1000 (2 processors, 1 GB memory, 34 GB internal disks): oracle database server and direction administration data server (samba). The data are stored on a SAN (Storage Area Network) consisting of an HP VA7400 Virtual Array (2.4 TB).

There is a project to move the direction administration data service to a new server that will also be used as exchange server.

- HP 9000/L2000 (4 processors, 2 GB memory, 72 GB internal disks): login and application server. It should be replaced next year.
- HP 9000/A500 (1 processor, 512 MB memory and 34 GB internal disks): telecommunications server. All our meteorological products (RMDCN) are sent and received on this server.
- HP 9000/L1500 (2 processors, 1 GB memory, 34 GB internal disks): forecast office server.
- HP rp3440 (PA8800 processor): used as web applications server (Tomcat), proxy server (Squid). It also runs several software packages such as a DMS (Apache Jakarta Slide), workflow (OSWorkflow), access to Oracle data base (Hibernate), server monitoring (Bigbrother), problem logging (Elog, freeware using a ticket system).

These servers are tied in a high availability cluster.

- 2 Dell Power Edge (Xeon) servers running SLES 9 are assembled in a cluster (Heartbeat) for the creation of products delivered by the RMI.
- a Compaq Proliant DL 350 (2 cpus pentium 3, 15 GB disk) has been refurbished and converted into a Linux login server. It runs RHES 4 and is in test phase.

For the **external communication links to the RMI**, the backup link to BELNET (gateway to Internet) will be replaced by a 100 Mb connection.

The RMDCN link is a leased line to SITA POP Brussels (384 Mb/s) with 3 ISDN backup lines.

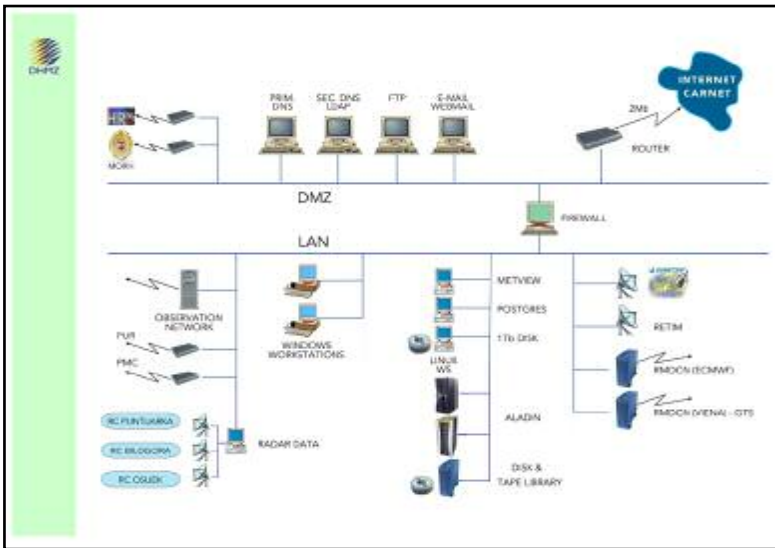
### ECMWF facilities usage

All Belgian projects (20) but two are using only access to MARS archive data. One project deals with a model of the North Sea. The second is handling ECMWF ensemble forecasts to calculate probabilistic precipitation forecasts over the area of Belgium operationally.

There is also the project to implement the ECMWF's SMS software on our SGI compute server to monitor the forecasts and other operational runs. The project will start probably in 2007.

The users expressed their satisfaction with the Support service and the manuals.

*Oleg Perčinić – Croatian Meteorological and Hydrological Service*



### Computing resources

- **16 processors SGI Altix 3000, 32 GB RAM, 2x150 GB disks**  
- LAM Aladin, **Regional Climate Model (RegCM)** - Norway donation
- **16 processors SGI 3400, 12 GB RAM, 120 GB disk**  
- LAM Aladin, RegCM
- **SGI Origin 200**  
- GTS, e-mail, radio sending
- **SGI Origin 200**  
- climatological db, radio sending db (file system)
- **SGI Indy**  
- postprocessing ECMWF dissemination data and graphical presentation
- **SGI Origin 350**  
- Qualstar tape library (5 TB online)  
- 1,5 TB disks space (raid 5 configuration)

### Computing resources

- **LINUX servers**  
- MetView (v. 3.6)  
- PostgreSQL database (climatological data)  
- 1 TB disks space (raid 5 configuration, Intel)  
- TrVis (visualisation for TV)  
- EUMETSAT data  
- RETIM (Meteo France satellite dissemination, GTS)
- **VAX VMS**  
- climatology, accounting
- **MS Windows computers**  
- DVD data  
- Observation network and automatic station data

### Network

- **LAN (100/1000 Mb UTP)**  
- Observation station network  
    - over 45 main meteorological stations  
    - over 30 automatic met. stations  
    - over 30 investors' met. stations  
- Maritime centre Split (PMC - ISDN)  
- Maritime office Rijeka (PUR - ISDN)  
- Radar stations network (8 x ISDN)  
- RMDCN (ECMWF/64 kbps, RTH Vienna/ 64 kbps)  
- RETIM  
- EUMETSAT  
- DWD
- **WAN**  
- Internet: 2 Mbps link to CARNET  
- Governmental organisations: 640 kb  
- Croatian national TV: 2 Mbps radio link

### Future plans

- **3 radar stations for nowcasting at the Adriatic coast**
- **RMDCN - 512 kbps**

DENMARK

DENMARK

Thomas Lorenzen – Danish Meteorological Institute, Copenhagen

Current weather radar positions and coverages

The "triple intersection" part of the radar coverages in the eastern part of Jutland is actually an area with bad precipitation monitoring due to the radar beam being elevated to much.

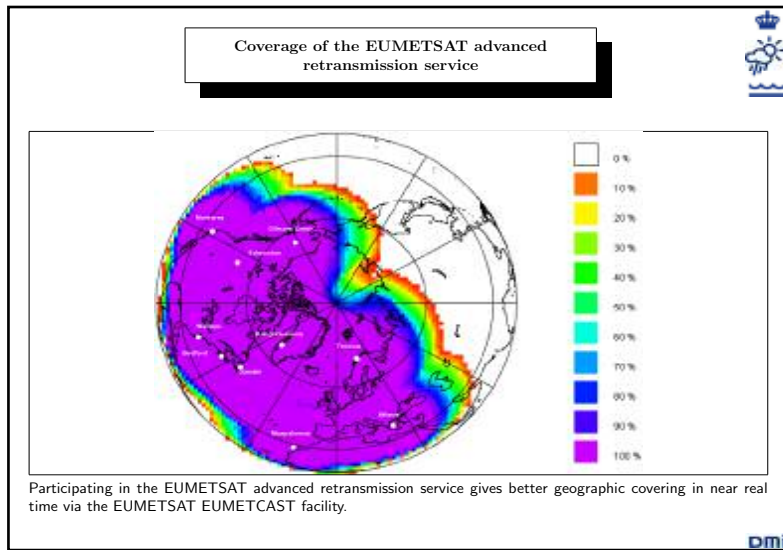
Possibilities for new weather radar location

The blue bullets on the map are currently targets of investigation for a new weather radar.

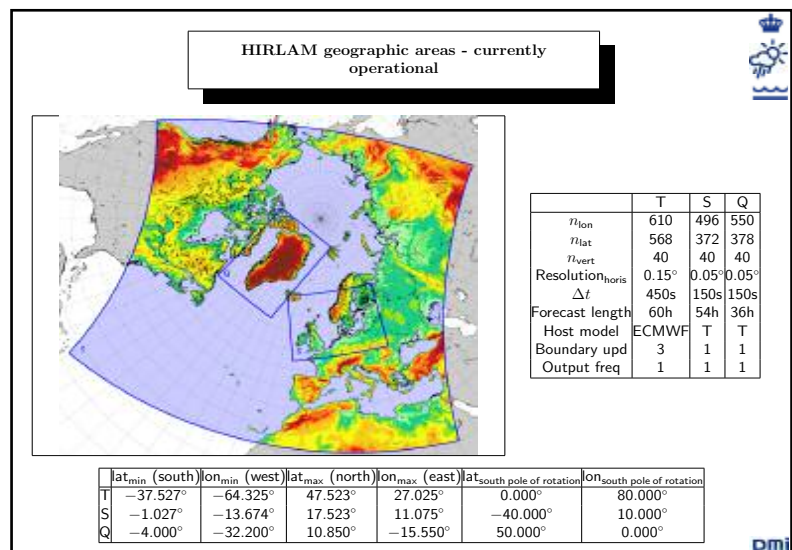
Coverage of the new EUMETSAT metop polar orbiting satellite

The EUMETSAT metop satellite to be launched on 20060717 is the first European polar orbiting satellite equipped with meteorological instruments primarily for operational meteorological purposes. DMI will receive data from two satellite receiving stations located in Denmark and Greenland. Data from the metop satellite (wind profiles, atmospheric soundings, etc etc) will be used as data assimilation input for the numerical weather prediction.





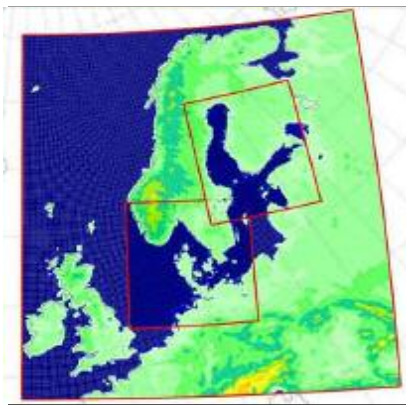
- Operational machinepark**
- NEC SX6 64M8 (8 vector nodes with 8 CPU's each)
    - 64 \* 8 Gflops
    - 32 \* 8 Gbyte RAM
  - NEC TX7 16M2 (2 scalar nodes with 8 CPU's each)
    - 16 \* 1.30 GHz Intel ItaniumII
    - 16 \* 2 Gbyte RAM
  - SX6 is used for running the HIRLAM forecast model four times a day. TX7 takes care of scalar pre-processing and post-processing.
  - Other operational workloads, among which also GTS receiving and archiving from RMDCN, are done on quite a big park of LINUX and SUN machines. No commodity type clusters are used at present, but may become relevant.
- DMI




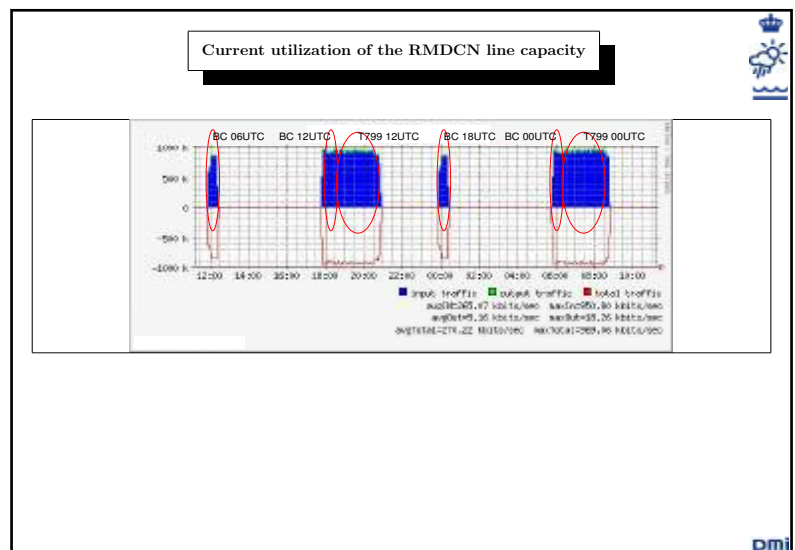
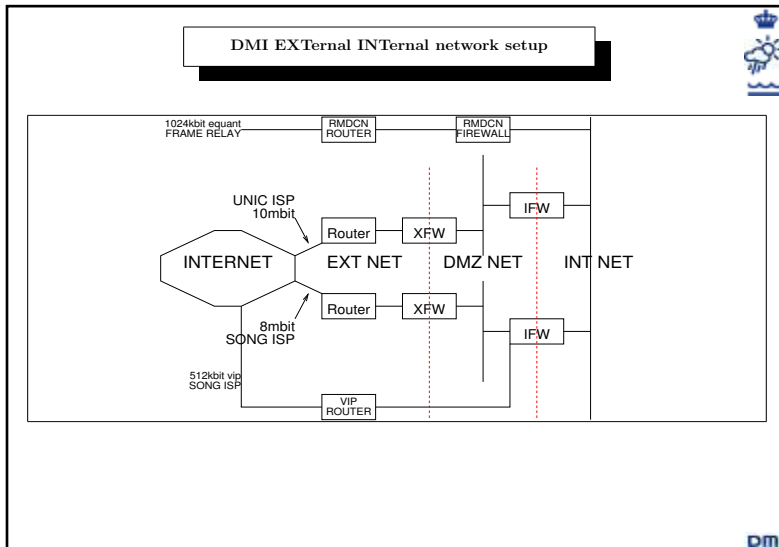
DENMARK

DENMARK

**ALADIN geographic areas - an ongoing research project**



	Scandinavian Hydrostatic	Danish Non-hydrostatic
$n_{lon}$	256	384
$n_{lat}$	288	400
$n_{vert}$	40	40
$Re_{shoris}$	0.11°	0.025°
$\Delta t$	300s	60s
FC length	24h	24h
Host	DMI-HIRLAM-T	Scandinavian
Bdy upd	1	1
Output freq	1	1
lat <sub>center</sub>	59.2°	56.3°
lon <sub>center</sub>	9.9°	14.0°

### Operational data flow from ECMWF to DMI



- DMI receives from ECMWF via RMDCN data amounting to close to 2.5GByte per day.
- Frames for hirlam account for  $\frac{1}{4}$  of the data volume. Global gauss surface fields for web and other visualization purposes account for  $\frac{1}{2}$  of the data volume. The remaining fraction contains a mixture of products for other purposes.
- Around dissemination hours the primary 1Mbit RMDCN line is fully saturated.
- The backup ISDN line only holds one third of the bandwidth of the primary line, so a failure of the primary line will cause operational delays at DMI. DMI may opt for a true mission critical setup, when RMDCN turns to use MPLS.
- For testing and quasi time critical activities, DMI receives data via internet amounting to about 1.0GByte per day.
- Data is transferred via internet via plain ftp as well as via our ecaccess gateway. Using our ecaccess gateway for this purpose has been ongoing for a year with no problems.
- Suggestion for improvement : *Could the MS ecaccess gateway be instructed to distribute data to more than one local host.*
- Suggestion for improvement : *Could the MS ecaccess gateway act as a buffer for received data, so data retransmission is not needed, in case of transient destination host problems.*



### DMI use of the ECMWF facilities



- Presently 32 registered users on ecgate and hpcd.
- Of these users, 24 seem to be using the system regularly.
- Two special projects are presently active.
  - EUCOS about data impact study concerning either observation data or model parameters and the model sensitivity of certain data types.
  - Potential of local area model ensembles for wind power prediction.
- The mars data archive and especially its era40 data have been used in several projects.
  - NordRisk (Nordic countries) about nuclear safety and probabilistic safety assessment.
  - EnviroRisk (EU funded) also dealing with conventional pollution.
- Of the DMI share of the compute resources, most activities are billed to hirlam related activities.
- About twenty percent of this years share has been used.



### Hirlam use of the ECMWF facilities



- Presently, the hirlam reference system is run on a daily basis at FMI. No quasi operational hirlam runs are presently conducted at the ECMWF system.
- There are a few practical obstacles in running the hirlam reference system on a regular basis on hpcd.
- The realtime ECMWF observation data are not available in due time, they have a two to three days delay.
- It may take too much compute resources which need to be shared among members.
- Staff to do monitoring is needed as well, but the member state job submission via SMS may prove helpful.
- Thus, the main hirlam user on ecgate just maintains the hirlam source code and updates the standard hirlam system installation from the main "members only" subversion repository <https://hirlam.org/trac/browser> located beneath the project homepage <https://hirlam.org>.



## DENMARK

## DENMARK


L. Gougeon commented that, if the Danish Met Service wished to implement a local dissemination system, they could consider installing the ECMWF product dissemination application, ECPDS, at their site or include part of it in their ECaccess gateway.

U. Modigliani asked for more information on how Hirlam was run. He explained that general access to ECMWF Reports Data Base was not allowed for 2 days to avoid conflict with ECMWF operational requirements. A set of observations used for back-up purposes is available on a separate filesystem, so if the lack of access to observations was the main reason preventing regular Hirlam runs at ECMWF, then the Hirlam community might like to consider use of this set with the introduction of the new framework for time-critical applications.

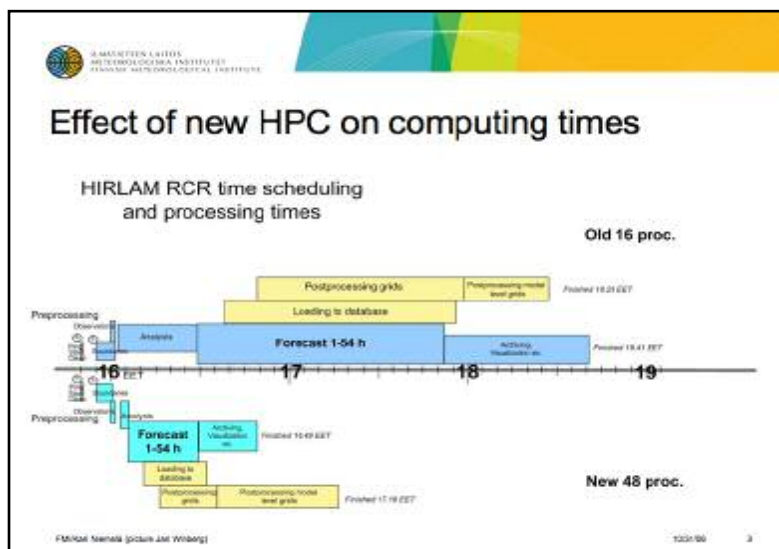
*Kari Niemelä – Finnish Meteorological Institute*

**New HPCs**

- SGI Altix-3700 / 256 CPU / 256 GB memory
- SGI Altix-3700 / 48 CPU / 48 GB memory
- The larger is used for research purposes
- The smaller is dedicated to model runs



FMI/Kari Niemelä 8/12/06 2



**Models run at FMI**

- Hirlam RCR version 7.00, ~ 20 km resolution, 40 levels, 438 x 336 points, 54 hours. Later this year update to ~ 16 km / 60 levels
- Hirlam MBE, ~ 8 km resolution, 406 x 306 points, 54 hours
- Both produce 31500 fields
- Arome under testing, 2,5 km resolution
- Silam air quality model

FMI/Kari Niemelä 8/12/06 4

FINLAND

FINLAND

**Removal to new premises**

**Target: to duplicate the systems so that all routine production could be run also during the removal either in the old place or in the new place.**

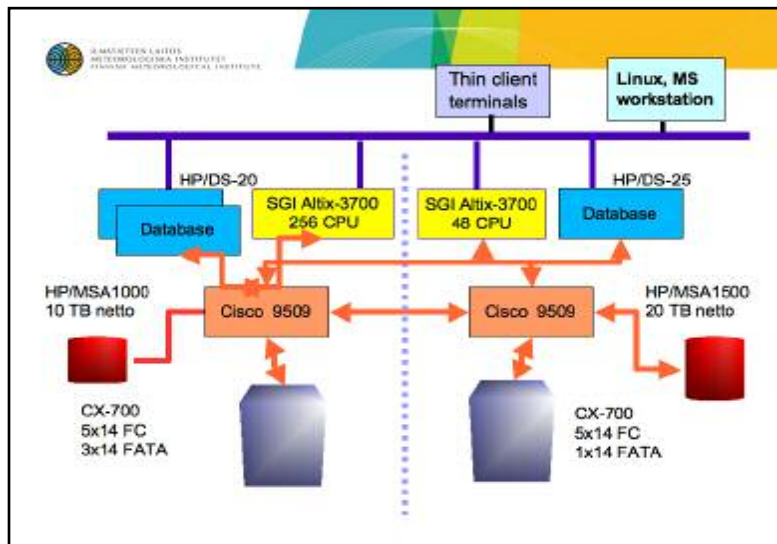
**Result: the target was accomplished.**

**However: some breaks caused by network problems**

**Very soon after the removal →**



FINMIG Meeting 8/12/06 5




**Use of ECMWF resources**

- **About 70 SecurID users. Positive: number of passive users has reduced during last years**
- **About 50 self-registered users**
- **7 active HPC users, so most users benefit from MARS archive**
- **In 2005 percentage of 70 reached in use of HPC allocation (record for a number of years)**
- **Most of HPC use is Hirlam test runs**

FINMIG Meeting 8/12/06 7

FINLAND

FINLAND



**Influence of new ECMWF model on dissemination requests**

- **Two basic areas: global 1.0 deg, Europe 0.25 deg**
- **Extensive use of internet stream for dissemination**
- **Only basic data through RMDCN**
- **Priority to RMDCN stream**
- **Volumes: 18 500 fields / 316 MB through RMDCN**  
**48 700 fields / 2.65 GB through internet**  
**(daily volumes)**

**much of the volume is due to model level data**

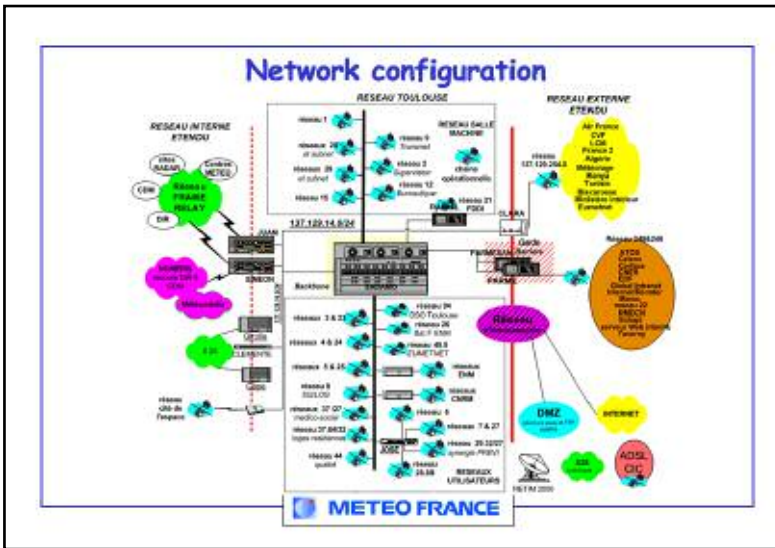
- **Before February the volume was 350 MB**

FINMIG/Netinfo 01/2006 4

FRANCE

FRANCE

Marion Pithon – Météofrance



**Network configuration**

- **Wide Area Network (between Toulouse, the 7 regional centres and the 90 local centres)**
- Migration from Frame Relay to MPLS in progress
- **Bandwidth:** 4 Mb/s for regional centres (8Mb/s for Paris and Lannion) 512 kb/s in general for local centres  
60 Mb/s for Toulouse (2 links of 30 Mb/s)
- All equipment is redundant for reliability.
- End of migration is planned for end of July 2006.
- **Internet connections.**
- "commercial" link : 16 Mb/s for MF Web server, backup of our satellite broadcast system, ...
- "research" link : RENATER : 80 Mb/s upgraded in May 2006:
  - 40 Mb/s reserved for ECMWF access through ecaccess
  - 10 Mb/s reserved for ftp
  - 16 Mb/s reserved for http
- 2 Firewalls for each link (CISCO and NOKIA)

08/06/06

**Network configuration**


- **Local Area network.**
- **General purpose network:** Gigabit Ethernet backbone 32 Gbit/s Full duplex
- **High performance network:**
  - HIPPI network between the 2 compute systems and the DHS
  - Replacement at the end of 2006 by redundant GigaEthernet switches to connect aggregate GE links on the compute system to 10GE interface on the DHS

08/06/06




**Compute system replacement**

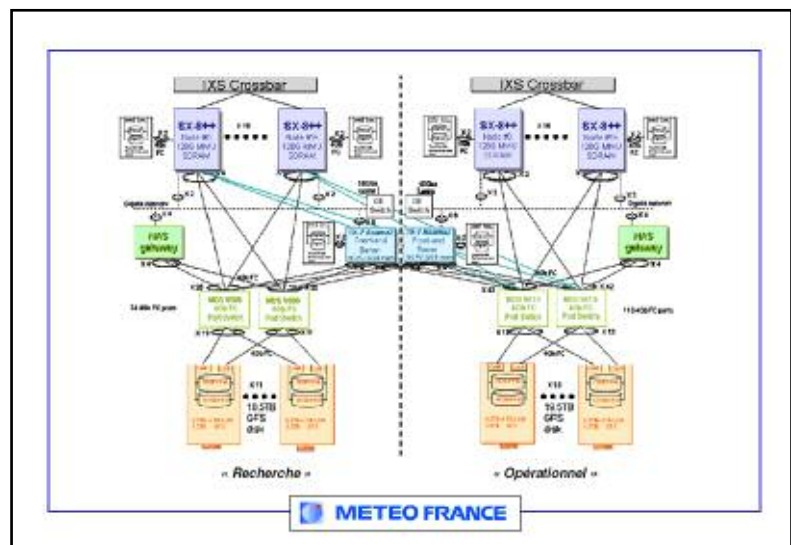
- **Requirements**
- Upgrade needed to match plans in NWP (new mesoscale model AROME, new resolution of MERCATOR model, improvements of ARPEGE-ALADIN) or research activities
- Upgrade in two phases (end of 2006/end of 2008)
- Performance required: x4 (x5 expected) for phase 1  
x8 (x16 expected) for phase 2
- End of Fujitsu contract on 08/2007 → beginning of installation at the end of 2006
- Iterative process of the procurement procedure (3 stages of discussion with tenderers)

08/06/06 

**Compute system replacement**

- **Procurement schedule**
- Call to tender in December 2004
- 7 replies (Bull, Cray, Fujitsu, IBM, LNXI, NEC, SGI)
- Only 3 left (IBM, NEC, SGI) for the first results on May 2005
- Second set of results in Sept. 2005 after interviews with candidates and site visits
- Final specification of requirements in Nov. 2005
- Last offers in Jan. 2006
- NEC selected in April 2006
- Performance commitment of x5 (phase 1) x21 (phase2)
- First installation planned for mid September 2006

08/06/06 



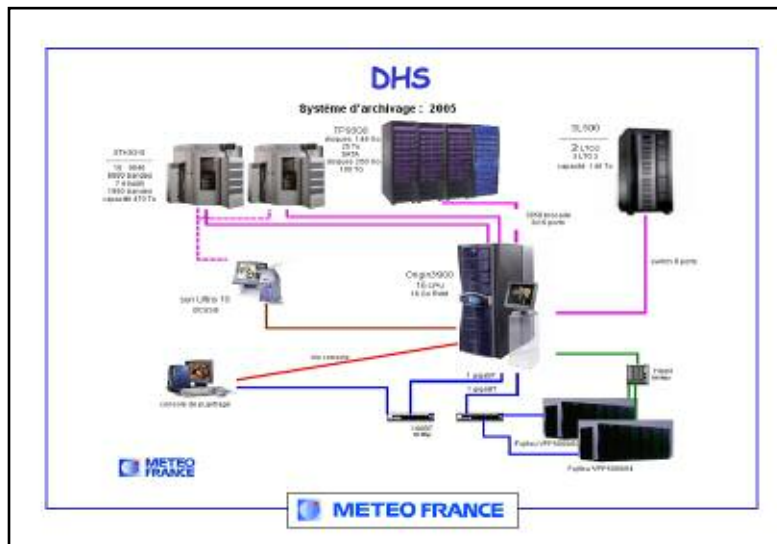
FRANCE

FRANCE

### Data Handling System

- Installed in March 2004; last phase in May 2005
- Software: DMF from SGI
- 3 different storage levels:
  - fast Fibre Channel disks: 25 TB for cache
  - Serial ATA disks: 100 TB in June 05
  - "Fast" tapes (9840) in STK 9310 silo
  - "Slower" tapes (9940) in STK 9310 silo
  - Only 5% of data have a "backup" copy in a different building
- Server SGI O3900 (16 procs - 16 GB mem)
- Total capacity 570 TB
- Actual use: 370 TB (+10 TB/month) 12Mfiles
- Upgrade needed for 2007 (negotiations with SGI)
- Replacement planned for 2009

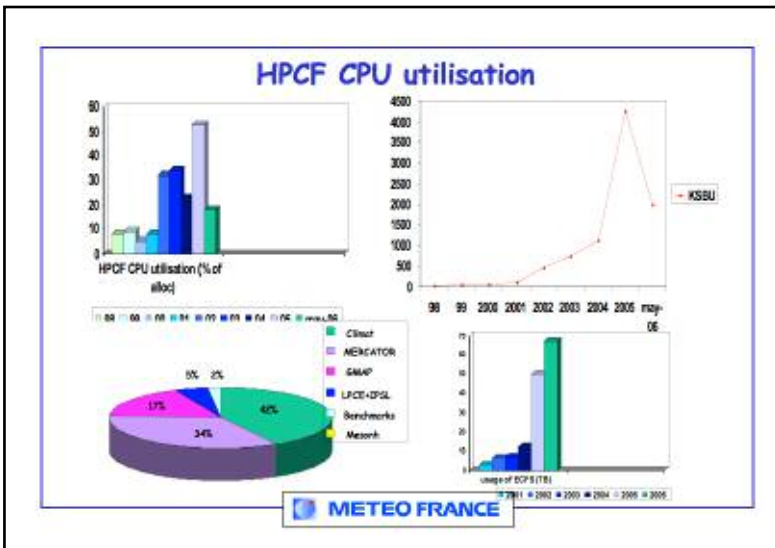
08/06/06



### Use of ECMWF facilities

- 51 M.F. projects + 4 Special projects:
  - ✓ 262 users
  - ✓ 140 from Meteo France - 120 from universities or other organisations
  - ✓ A lot of work for user support team ...
- 7 projects used HPCF resources in 2005.
  - ✓ Climate, MERCATOR, research in Modelling...
- Increase of HPC use → need for data storage → use of ECFS will grow
  - ✓ 87% of the total amount of storage allocated is used (68 TB).
  - ✓ Some warnings and advice were given (using ectmp)

08/06/06



- ECMWF/LAM Special Project (spfrcoup)**
- ✓ 11 countries (4 MSs 5 Co-op States + Morocco and Slovakia)  
23 registered accounts (some users were already registered)
  - **Scientific and technical goals of the project**
    - ✓ Investigation of coupling LAM ALADIN and AROME to other large scale model data (IFS and ERA archives)
    - ✓ Technical issues when coupling Aladin/Arome to ECMWF model data (file preparation, data compression, telecom issues,...).
    - ✓ Relative merits of coupling Aladin/Arome to IFS vs Arpege models for NWP (various geographical domains)
    - ✓ Impact of large scale model resolution on the quality of LA modelling
    - ✓ Coupling issues linked to the representation of the surfaces
    - ✓ Coupling issues linked to differences in physical packages
- 08/06/06 METEO FRANCE

- ECMWF/LAM Special Project (spfrcoup)**
- **How will resources be used ?**  
Scientists of different states will:
    - ✓ Access HPCF to perform boundary condition files
    - ✓ Transfer them to their own sites
    - ✓ Run the LAM on their own sites
    - ✓ Compare Arpege/Aladin and IFS/Aladin forecasts on case studies
  - **Activities planned / Resources needed**
    - ✓ Offer a unified software environment for preparing Aladin and Arome boundary condition files
    - ✓ Storage on ECFS of Aladin/Arome coupling files for chosen test periods and geographical domains
    - ✓ Some LAM runs on HPCF for testing purposes → Neither regular nor long forecast experiments at ECMWF
- 08/06/06 METEO FRANCE

FRANCE

FRANCE

### Other projects

- **Hirald group**
  - ✓ Scientific collaboration with HIRLAM/ALADIN communities
  - ✓ Offer a shared software environment on HPCD (scripts, libraries, object files,...) to run Arpege or Aladin configurations (ms\_perm file system)
  - ✓ Used through OLIVE by MF users to run Arpege models
- **OLIVE**
  - ✓ Web-like interface to run ARPEGE or ALADIN experiments at ECMWF
  - ✓ Use SMS (local) and ECaccess (ecjput, ectrans, genericExec plugin)
  - ✓ Use of "hirald" software environment

08/06/06

### ECMWF services

- **User Support**
  - All users are very satisfied with support: help of D.Lucas very efficient
  - Excellent organisation (call desk/user support)
  - Web site useful for support
- **EMS**
  - Easy to use; Helpful to comp.representative (speed up the process of user registration)
  - Sometimes it blocks (when add user or organisation) - Problem of communication ("Web error")? Refresh solves the problem
- **ECACCESS**
  - All file transfers between ECMWF and MF use ECaccess tools or ectrans
  - Used by OLIVE mechanism
  - Very reliable

08/06/06

### Plans

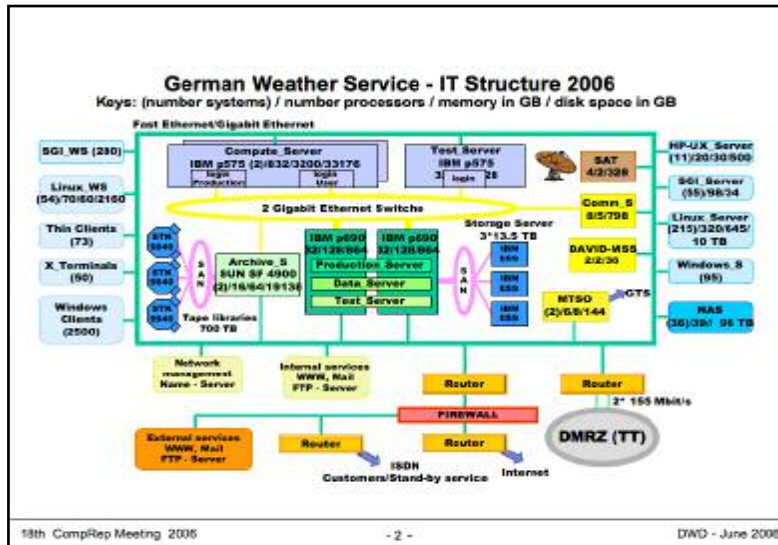
- **New gateways for ECaccess:**
  - On 2 LINUX PCs in DMZ
  - To replace our current HP-UX gateway on MF network
  - Installation next month (Jul 06)
- **Continue using HPCF resources at the same rate**
  - More ARPEGE or ALADIN runs (thanks to OLIVE tool and upgrade of Internet link)
  - Climate team: seasonal forecasts
  - Mercator: tests on MNATL12 (regional models)

08/06/06

T Lorenzen commented that MeteoFrance’s procurement by the iterative process of ‘competitive dialogue’ had been very time consuming: had the time and effort involved been worthwhile? M. Python replied that it had been well worthwhile, as the many technical discussions between MeteoFrance and manufacturers meant that their technical requirements were very clearly defined by the end of the process.

In response to M. Python’s comment about occasional EMS communications problems U. Modigliani noted that communications problems with no immediately obvious cause were also occasionally experienced by ECMWF staff and would be investigated. Meanwhile, he agreed that a ‘refresh’ usually cleared the problem.

Elisabeth Krenzien – Deutscher Wetterdienst



### Compute server Upgrade

5/05 Installation of Power5 replacement System

6/05 Acceptance test passed

7-8/05 Migration of applications (operation, user)

15.08. NWP suite operational on the system

31.08. Shutdown of Power3 cluster

06.09. Power3 cluster removed completely

11/05 Installation of second Power5 cluster

4 add. compute nodes, 2.2 TB GPFS disks

AIX 5.3, operational load only

1/06 Acceptance test passed

2/06 Migration of NWP operations (15.03.)

Current production system since

3/06 Upgrade of development system hardware (4 nodes, disk storage..) software (AIX 5.3)

4/06 Acceptance test passed

nodes: 52 p575 nodes, 8 processors Power5, 1,9 GHz

number: 44 compute and 4 login nodes 4 I/O nodes

memory: 32 GB, I/O nodes 16 GB

peak: 3.161 Tflops/s

GPFS FS: 6x DS 4300 Turbo (14.4 TB) FC SAN

networks: dual plane HPS Switch (2x 2.5 GB/s bi-directional)

2 CSM, 6 HMC Servers, test system (4 nodes)

18th CompRep Meeting 2006 - 3 - DWD - June 2006

### Compute server

Performance-neutral replacement of hardware

better communication and I/O performance

LM\_RAPS\_3.0 on the LME domain:

- Only half of the machine user
- I/O intensive: read 3.5 GByte, write 6.1 Gbyte
- Lower Flops rates than former tests with LM\_RAPS 2.0, because of worse relation: computation / (commun. + IO) in this test
- efficiency increase from 7% to 9%
- SMT activated

### Performance benefit

	P3	P5/SMT
# Processors	900	160
Total Time (s)	1172.5	969.6/776.4
Comp. Time (s)	549.1	782.5/586.9
Comm. Time (s)	208.7	118.3/96.5
I/O Time (s)	300.0	58.3/106.3
# Operations (10 <sup>12</sup> )	75.1	93.9
GFlop / seconds	64.1	96.8

18th CompRep Meeting 2006 - 4 - DWD - June 2006

GERMANY

GERMANY

### Compute server

OS level: AIX 5.3  
 Cluster Managem.: CSM 1.4  
 Parallel Environ.: PE 4.2  
 LoadLeveler: 3.3.3  
 GPFS File system: 2.3  
 Compiler: xlf 9.1; xlc 7.1  
 Debugger: TV 7.3  
 Libraries: ESSL 4.2  
 Service Pack: SP 6

**Twin System:**  
 Two independent Compute Servers for production and development and back-up  
 99.8 % availability per month for operational NWP system  
 regular update of data files and check pointing on application level  
 manual switch of NWP suite from production to development system in 15 min  
 DMRZ is centre of HPC competence for the Ministry of Transport, Housing and Urban Development since May 2006

### Utilisation

18th CompRep Meeting 2006 - 5 - DWD - June 2006

### Archiving System

05.2005 Start of installation  
 28.6.- 27.7 Acceptance test passed with a few problems (missing functionality)  
 Data server -> Archive server: installation of special file transfer software „dmSCP2“  
 13.09.2005 Migration of ECFS user archive no problems  
 19.09.2005 Start of migration of AMASS archive into SAM-QFS archive (meta data) several major problems Functional down-time (3 weeks)  
 11.2005 Start of bulk migration  
 10.03.2006 End of migration 650 mill. files migrated, 4 PB data handled, (one corrupted data block, 7 files lost during a system crash, one „mystery file“)

### Specification

Fallover Cluster 2 x SunFire 4900 server  
 Test cluster (SunFire 490 server)  
 Per Cluster Element 8 Proc. / 32 GB Mem.  
 18.7 TB disk space (SAM-QFS File system)  
 30 Archives (currently 750 TB)  
 SAN 2FC Gb switch 16 port connections to 3 StorageTek Silos with 46 STK FC-drives : STK 9940 (16), STK 9940A (10), STK 9940B (20)

Solaris 9 (SUN OS 5.9)  
 Fallover Software SUN-Cluster 3.1  
 HSM-Software SAM-QFS V 4.4.3  
 ECFS user archive (based on ECMWF version)

Current Usage:  
 SAM Files: 5 836 825  
 Data volume 660 890 TB  
 Tapes: 9 479

18th CompRep Meeting 2006 - 6 - DWD - June 2006

### Plans

**New DWD building in Offenbach**  
 cornerstone: 23.03. 2006  
 moving in: spring 2008  
 computer hall moving in: end 2007

**NinJo** parallel operation since 03/2006  
 version 1.2 in 07/2006

**Replacement of DMRZ**  
 Procurement schedule 2006/2007

**ITT Compute server and Data Handling system:**  
 Performance enhancement factor 10-15

**ITT Storage System:**  
 Replacement of STK Powderhorn silos

**Installation Schedule** 2007/2008

**New computer hall**  
 area: 1100 m2  
 false floor: 2 m  
 UPS: 3000 kW  
 2 diesel-engines with alternator

18th CompRep Meeting 2006 - 7 - DWD - June 2006

**ECMWF resources: DMRZ Operations**

**ECFS enhancements:** AIX 5.3 client  
early-access tests of ECFS\_FTP Portal (ECcmd based)

**ECaccess v2.2:** Internet: Linux (DMZ v2.0), AIX (LAN)  
RMDCN gateway: planned  
Csomars version 1.2 in full operation, interactive client

**SMS v4.4.5:** single instance, solely used for operations

**Operational Products:** via dissemination 657 MB; (transfer rate ~43 kB/s)  
(per target date) via Csomars 717 MB; (transfer rates 115 - 1540 kB/s)

18th CompRep Meeting 2006 - 8 - DWD - June 2006

**Deutscher Wetterdienst**

**ECMWF resources: Research Modelling**

**Global Model (GME)** 2 reference runs (daily) for DWD production

**GME-EPS** testing assimilation schemes for new model

**Local Model (LME)** test bed for model improvements

**COSMO LEPS** LM Local area Ensemble Prediction System

**EPS studies** MUSTOK Project 'Storm surges in the Baltic Sea'

	GME	GME-EPS	LME	MUSTOK	COSMO-LEPS
2006	40 km/L40 20 km/L50	220 km/L40 M32 to M64	7 km/L40 3TL, 2TL	T <sub>1</sub> 399L62 M51	10 km/L40 M10
KSRU	4200	1000	3300	2000	1900
2007	40 km/L40 20 km/L50	220 km/L40 M32 to M64	7 km/L40 2TL, 2TL	T <sub>1</sub> 399L62 M51	10 km/L40 M10

18th CompRep Meeting 2006 - 9 - DWD - June 2006

**Deutscher Wetterdienst**

**Storm Reanalysis with GME, LMQ, LMK**

**GME Sensitivity study**

18th CompRep Meeting 2006 - 10 - DWD - June 2006

GERMANY

GERMANY

**Deutscher Wetterdienst**

**Simulated weather conditions during real storm surges with EPS**

storm surge year	residual surge meas. / sim. potential [cm]	storm surge category meas. / sim. potential	Number of EPS
2002	158 / 142	severe / severe	21
1995	202 / 217	extreme severe / extreme severe	21
1987	169 / 161	extreme severe / extreme severe	11
1979	157 / 185	severe / extreme severe	19
1978	120 / 171	weak / extreme severe	39
1968	155 / 231	severe / extreme severe	17

18th CompRep Meeting 2006 - 11 - DWD - June 2006

**HPC Allocations**

Storage used in 2006 (April)  
**ECFS: 8.4 TB**  
**MARS: 44.3 TB (available 89.4 TB)**

DWD user	2006	2005	2004
	72	69	64
Last login	57	12	-

**Annual usage in 2005 / 2006**

18th CompRep Meeting 2006 - 12 - DWD - June 2006

**Special Projects (SP)**

Currently 24 SP are registered for German universities (11) and different research facilities such as Max-Planck-Institutes (MPI, 4), DLR (5) and other research centres (4) and 5 new applications for 2007 have been handed in.

number of users	2006 ( April )	2005	2004
	87	81	66
last login	62	16	-

**Main activities:**  
 HPCD Resources are used by 9 SP mostly for large scale climate modelling  
 Retrieval of MARS data  
 Current forecast data for flight guidance during field campaigns (several DLR institutes)  
 ERA-40 data access

18th CompRep Meeting 2006 - 13 - DWD - June 2006



<p><b>Experience</b></p> <p>All users appreciate the professional support from ECMWF User Support, Call Desk and staff members.</p> <p>many thanks for:      MS visit by Carsten Maass in November 2005 and support on using <i>ectrans</i> and <i>EMS</i> (<i>learning by doing</i>) running large scale jobs: 30 y of T213L31 with ECham (across several Special Projects) excellent throughput on HPCD</p> <p>concerns raised are:    shortage of permanent on-line disk space on HPCD system appreciated would be: 2 - 6 TB ECaccess transfer rates drop drastically for large transfers to about 1/4 of ftp</p> <p>questions refer to:     accounting information for new Special Projects</p>		
18th CompRep Meeting 2006	- 14 -	DWD - June 2006

In regard to the *ectrans* problem L. Gougeon replied that a German user had tried to transfer files of more than 6 GB. The file size was encoded in 32 bit and the application to submit files on *ecgate* had failed for this. The problem was fixed on the day of discovery and files of 11 GB now transfer successfully. However, this particular user had so much data (~ 2.5 TB) to transfer at once that it was decided to send it on hard disks.

U. Modigliani commented that there was planned to be more file space for Member States on *hpce* but exact allocations had not yet been decided.

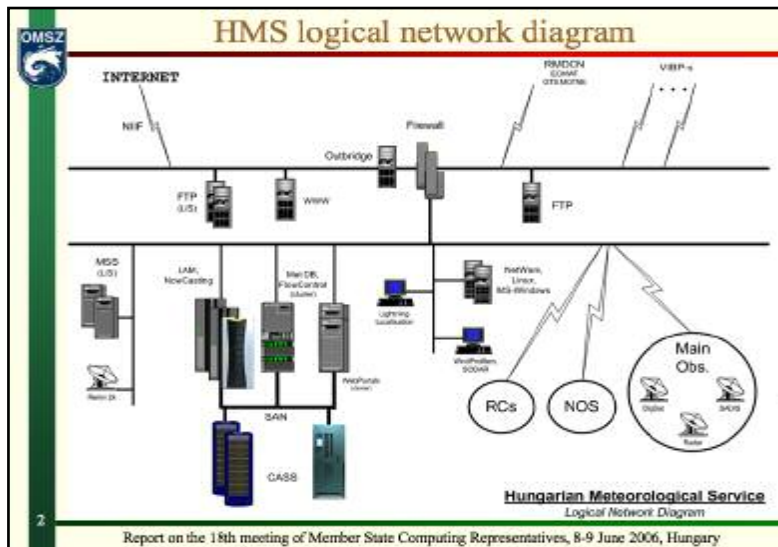
N. Storer noted that there was already 6 TB for the Member State temp file system on *hpcd*. *hpce* will have twice the disk space of *hpcd* and use of Multi-Cluster GPFS will obviate the need to replicate as much data as with the old system, so the MS file space will certainly be increased, doubled if possible.

N. Storer took the opportunity to reiterate the need for users to backup their own essential data: default backups were too inefficient to be performed.

HUNGARY

HUNGARY

István Ihász – Hungarian Meteorological Service, Budapest



Report on the 18th meeting of Member State Computing Representatives, 8-9 June 2006, Hungary


### Network

- LAN: 100/1000 Mb/s, structured UTP cabling system, CISCO switches
- SAN: 2Gb/s, redundant FC-switches, CLARiON disk unit with backup and IBM LTO3 Ultrium tape librare\* connected to IBM, HPcluster, SGI, Linux servers
- WAN: from 64 kb/s to 256 kb/s leased lines, 4\*2 Mb/s microwave channels
- Internet: 128 kb/s via ISP for commercial users (FTP)  
1 Gb/s via University Network for general Internet usage
- RMDCN (256 kb/s with ISDN backup):
  - ECMWF: 128 in / 16 out kb/
  - Austria (AC RTH): 32 in / 8 out kb/s
  - Slovakia: 16 in / 16 out kb/s
  - Serbia: 8 in / 16 out kb/s
- Firewall (ZORP): Internet segment and DMZ
- Satellite broadcast: SADIS 1G, RETIM 2000, MSG
- Management: OpenView NNM, WEB-Nagios

Report on the 18th meeting of Member State Computing Representatives, 8-9 June 2006, Hungary

### Computer resources I.

- IBM pSeries 690 (2002); 32 POWER 4 CPUs, 64 GB RAM, 850 GB disk, SAN: operational LAM (Aladin/HU); regional climate modelling and other research studies
- IBM pSeries 655 (2004); 32 POWER 4+ CPUs, 128 GB RAM, 4x36 GB disk, SAN: LAMEPS Aladin, ultrashort range NWP modelling (AROME) and other research studies
- SGI ORIGIN 2000 (1999); 16 CPU, 8 GB RAM, 88 GB disk, SAN: operational nowcasting (MM5) and GIS
- SGI Altix 3700 (2006); 144 Itanium 2 CPUs, 288 GB RAM, Numalink 26 Gbps, 900 Gflop/s 2x140 GB mirroredisk, SAN: nowcasting, ultrashort and short range forecast, LAMEPS, climate modelling



Report on the 18th meeting of Member State Computing Representatives, 8-9 June 2006, Hungary

**Computer resources II.**

**Central Processing & Controlling System**  
 HP L3000 cluster-PKG2; 4 PA8700 CPUs, 6 GB RAM, SAN: scripts, programs; CDS-CASS storage systems


**Database server**  
 HP L3000 cluster-PKG1; 4 PA8700 CPUs, 6 GB RAM, SAN: CLDB, CADB; ORACLE (8.1.6)

**Central Archive and Storage System**

- CLARiiON CX700 and FC4700 disk unit: ~13 TB native capacity, with tape backup
- IBM 3584 LTO3 Ultrium 2/60 tape librare: ~30 TB capacity

**System, database and data backup**

- HP SureStore Ultrium 2/20 tape librare
- HP DLT 1/8 tape librare



5 Report on the 18th meeting of Member State Computing Representatives, 8-9 June 2006, Hungary

**Computer resources III.**


**WEB servers**

- SGI Altix-350; 6 Itanium 2 CPUs, 6 GB RAM, 36 GB HD; SAN: Public portal [www.met.hu](http://www.met.hu)
- PC Linux cluster; 4 nodes, Pentium IV. CPU, 0.5 GB RAM, 40 GB HD; SAN: Intra portals

**Message Switching System**  
 2 PC-s; Linux: life-standby WeatherMan

**Other (Linux, Unix, Netware) servers**

- firewall, mail, printer, WAP, FTP, etc.
- ECaccess (Internet) and MSaccess (RMDCN) gateway
- DEC, SUN, HP and Linux workstations for visualisation and development
- about 300 PC (Windows, Linux)



Report on the 18th meeting of Member State Computing Representatives, 8-9 June 2006, Hungary

**Changes related to ECMWF**

- Twenty-seven registered users (10 in 2003, 16 in 2004, 21 in 2005)
- ECaccess facility via Internet (ECaccess gateway) and RMDCN (MSaccess gateway)
- Local questionnaire on use of ECMWF resources (March 2006)
- Generate 100 member EPS plumes operationally on *ecgate* server since September 2005
- Special project: Investigation of coupling the ALADIN and AROME models to boundary conditions from ECMWF and ERA model data co-operation with Meteo France

7 Report on the 18th meeting of Member State Computing Representatives, 8-9 June 2006, Hungary

**Summary of questionnaire on use of ECMWF resources (cont)**

**Q.1: computer usage**  
 75 % work on both ecgate and local computer  
 25 % work on only local computer

**Q.2: type of work on ecgate**  
 25 % operational and research & development (R&D)  
 60 % only R&D

**Q.3: data source (more answers)**  
 60 % deterministic model  
 35 % ensemble model  
 15 % monthly forecast  
 25 % seasonal forecast  
 10 % DEMETER (multi model EPS seasonal forecast)  
 10 % ERA-15 (re-analysis 1979-1993)  
 40 % ERA-40 (re-analysis 1957-2001)  
 20 % Special project  
 5 % others /EUCOS

Report on the 18th meeting of Member State Computing Representatives, 8-9 June 2006, Hungary

**Summary of questionnaire on use of ECMWF resources (cont)**

**Q.4: Trouble shooting (more answers)**  
 75 % use ECMWF web  
 85 % ask Computing Representative  
 85 % ask colleagues  
 75 % read printed documents  
 85 % ask colleagues  
 25 % occasionally ask User Support

**Q.5: Quality of printed documents**  
 50 % said: good, clear and well organised  
 50 % said: suitable

**Q.6: Quality of ECMWF web information**  
 55 % said: good, clear and well organised  
 45 % said: suitable

**Q.7: Assistance of Computing Representative**  
 All of them is satisfied

**Q.8: Assistance of User Support**  
 All of them is satisfied

Report on the 18th meeting of Member State Computing Representatives, 8-9 June 2006, Hungary

**Summary of questionnaire on use of ECMWF resources**

**Q.9: Need of additional information and/or training (more answers)**  
 10 % local training courses  
 25 % more information on ECMWF web /significantly decreased/  
 55 % don't know the future needs,  
 30 % have no additional needs

**Q.10: Subject of local training course on METVIEW (June 2005)**  
 65 % were fully satisfied,  
 10 % said: training was good and it was just enough,  
 20 % do not participate on it  
 5 % it was a little bit to much

**Q.11: Experiences of local training course**  
 80 % said: it was easy to follow  
 20 % do not participate on it

Report on the 18th meeting of Member State Computing Representatives, 8-9 June 2006, Hungary

ECMWF data by ECPDS and MARS			
Data type	number of files	MB/day	arriving time (UTC)
<b>European area (70N,15W, 34N,40E; DET: 0.5x0.5, EPS: 1x1 degrees)</b>			
H2D - GRIB DET 00&12 UTC	53	2*216	6.45 pm/pm - 09.05 am/pm
H2E - GRIB EPS 00&12 UTC	49	2*299	8.45 am/pm - 00.30 am/pm
H9E - EFI GRIB EPS 00&12 UTC	34	2*0,7	9.25 am/pm - 10.50 am/pm
<b>North Atlantic area (90N,90W, 18N,90E; DET: 1x1 degrees)</b>			
H9D - GRIB DET 00&12 UTC	42	2*2,1	6.10 pm/pm - 8.10 am/pm
<b>Northern hemisphere (90N,0E, 18N,0W; DET: 1x1, EPS: 1.5x1.5 degrees)</b>			
H8D - GRIB DET 00&12 UTC	25	2*88	6.05 pm/pm - 9.45 am/pm
<b>Northern hemisphere (90N,0E, 18N,0W; DET: 1x1)</b>			
H1D - GRIB DET 00&12 UTC	48	2*104	5.52 pm/pm - 9.42 am/pm
<b>Weather parameter BUFR files:</b>			
H3A - BUFR DET WORLD	1	2*5	9.40 am/pm - 9.50 am/pm
H5A - BUFR DET HUNGARY	1	2*1	7.18 am/pm
H5B - BUFR EPS HUNGARY	1	2*1	,8.28 am/pm-9.13 am/pm
H6B - BUFR EPS WORLD	1	2*1	11.53 pm
-			
Monthly EPS Forecast for Hungary	8 files/week by MARS retrieval		
Seasonal EPS Forecast for Hungary	5 files/month by MARS retrieval		
21 days Control Forecast for Hungary	84 files/day by MARS retrieval		

Report on the 18th meeting of Member State Computing Representatives, 8-9 June 2006, Hungary


Future plans	
<ul style="list-style-type: none"> <li>• Establish the possibility of dissemination via Internet (for backup, test and operational use)</li> <li>• Further development of WEB based visualization for ECMWF forecast and verification (Intraweb)</li> </ul>	

Report on the 18th meeting of Member State Computing Representatives, 8-9 June 2006, Hungary

ICELAND


ICELAND

*Vigfús Gíslason – Icelandic Meteorological Office*




### Current systems

- Servers
  - VAX (still being decommissioned)
  - Sun/Solaris (file/web & products)
  - Windows server (terminal/file/web)
  - IBM/AIX for the DB2 database
  - Linux (file/web/MSS/MSG)
- Clients
  - Workstations (Windows, Linux & Solaris)



### Current networks

- Main external connections
  - 2 Mbit main internet connection
  - AFTN to Icelandic Flight Control and Coastal Guard
  - RMDCN (64 kbit)
  - EUMETCast reception
- Internal network is mainly 10 Mbit for users with a few 100 Mbit for forecasters
- Fiberconnections
  - interconnecting SAN, BladeCenter and DB2 database



### Changes since last meeting

- Migration to a new message switch and visualisation tool for the forecasters
  - MovingWeather (MW) and VisualWeather (VW) from IBL Software Engineering
- Joined EUMETSAT and are now receiving via EUMETCast
  - 1.8 metre satellite dish for stable reception
  - IBL Software Engineering SW/HW and connection to MW/VW
- New resources for production/research
  - BladeCenter and extension of SAN (>4 Terabyte)
  - Fiber interconnection of BladeCenter, main DB and SAN



## Current Activities

- BladeCenter – production and development centre
  - decommissioning of VAX/Sun machines
  - migration of all products
- External/Internal Web
- Network upgrades
  - Internal fibre backbone and 100 Mbit ethernet to users
  - 1Gbit national connection with 315 Mbit international
- Re-evaluating security systems like UPS's, cooling systems and user access control. To guarantee a stable environment.



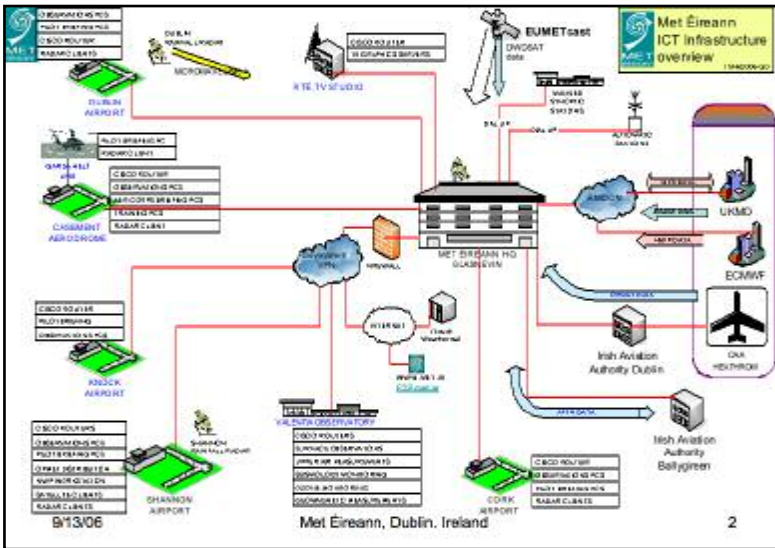
## Future plans and activities

- Pursuing the possibility of outsourcing services like backup, external web and e-mail
- Stabilizing production and research runs after many changes
- Setting up a more extensive monitoring system (nagios)

IRELAND

IRELAND

Paul Halton – Met Éireann



### SecureID Card Users

- 21 Active Users (logged in within past 4 months)
- 4 Inactive Users (didn't login in 7 months)
- 2 Recent new applicants awaiting new cards

- 3 Cards Disabled (returned to Computer Rep.)
- 4 Spare Cards (2 expire on 30 April 2007)
- 2 Expired Cards

*SecureID Cards: current versions have awkward key pads - hit and miss!*

24 Cards are due for replacement before the end of **April 2007**


9/13/06 Met Éireann, Dublin, Ireland 3

### User Feedback

- Number of registered users has increased
  - Higher proportion of allocated HPC resources being used recently
  - Special Projects BC, SPIEC4I and SPIESERG
- User Support
  - Users express thanks to Paul Dando & his predecessor John Greenaway
  - Special thanks to Vesa Karhila, for on-site visit in June 2005 to install MetView on a central Linux server and for the user training provided.
  - Thanks to Anton Beljaars who helped on a scientific/technical enquiry
  - Users have found User Support staff to be extremely helpful and they are impressed with the speed of response to queries.
  - During the switchover to T799 one user had several contacts with User Support. In several instances, he was contacted by User Support pointing out potential problems with some of his SMS jobs.

9/13/06 Met Éireann, Dublin, Ireland 4





## Other User Comments


**ECaccess gateway**

- ECaccess gateway is working very well -
- Since May 2005, extra deterministic parameter data are requested routinely from MARS via ECaccess gateway
- Ectrans is much more flexible than the old ECcopy

**EPS Plumes**

- 3 EPS plumes received by fax daily
- An electronic version of these data was required
- User Support investigated and found a suitable FORTRAN program compatible with SMS
- User Support contacted a user following a recent change in MARS and advised him to move his program to a different EPS directory

9/13/06 Met Eireann, Dublin, Ireland 5



## Special Projects

**SPIE C4I Project**

C4I experience has been very positive - User Support very helpful at all times.  
 Only one thing which caused some problems (Mainly during office hours!!): The transfer from ECFS to ECgate was very slow during the period from Dec 2005 until now (May 06). The transfer from ECFS to HPCD was not a problem.


**SPIE SERG Project**

Project team had no comments or special experience reports from last year.  
 Project work could be carried out without much interaction from User support.  
 Project is continuing with experiments to get data for a longer period to enable them to start working on the combination of wind power and precipitation for "hydro power" forecasting.

**Boundary Conditions (BC) Optional Project**

At present we only receive 31 of the available operational BC levels in the "frames"  
 From July 2006, new servers at Met Eireann will be used to receive and store all BC vertical levels in the range surface level up to the 10hpa level [i.e. levels 18-90]  
 The extra levels will be added to the routine dissemination of the BC data for use with the operational HIRLAM model.

9/13/06 Met Eireann, Dublin, Ireland 6



## RMDCN reception Servers

- Early in 2005 the old RMDCN servers crashed on a few occasions. It was clear that they were not fast enough to process the higher volumes of data.
- In August 2005 two new Linux (RedHat) servers were installed to replace the two old and much slower low spec servers.
- The new servers were set up identically to the previous servers and they use the same IP addresses.
- A 'Dlink' 4 port switch is used to connect the two servers.
- The new servers are faster, more reliable and have much more disk capacity.
- The transition to the new servers was transparent to RMDCN systems at UKMO and ECMWF.

9/13/06 Met Eireann, Dublin, Ireland 7

## ECMWF High Res. Model Changes

- Preparations began in Sept 2005 for the High Res. Model upgrade.
- The postponement of the changeover from October 2005 to Feb 2006 allowed sufficient time for users to plan and test NWP systems and applications
- Using the e-suite test data provided by ECMWF, the NWP systems and applications were updated on time to receive the new Model output
- Operational reception of ECMWF High Resolution data [T799 (deterministic) and T399 (EPS)] started on 01 Feb 2006
- After a 10-day overlap period, the dissemination schedule was changed to use the Gaussian area N400 (replacing the old N256 area)
- Operational use of the N400 data began on 10 February 2006.
- The various NWP parameters requested by the forecast offices are now available at 3 hourly time steps out to 72 hours and 6 hourly time steps thereafter for surface parameters.
- Most U/A parameters have been changed to 6 hourly timesteps.
- On 16 March 2006, the ECMWF dissemination streams for BC data for Ireland were re-organised in preparation for the reception of additional vertical levels.

9/13/06 Met Eireann, Dublin, Ireland 8


## WAFS Charts for Aviation Users

- GIS-Meteo from MapMakers Limited is used to generate WAFS charts operationally since 1 July 2005.
- GIS Meteo package can produce Wind/Temperature charts for any part of the globe.
- GRIB data received from UKMO is used every 6 hours to produce 12, 18 and 24 hour Wind/Temperature Charts for NAT, EUR, MID and EURAFI areas in the 'PNG' format.
- The system has also been set up to decode BUFR SigWx data
- It is ready to produce SigWx charts but these charts are not operational as yet.
- ICAO deadline for the operational introduction of BUFR SigWx data is 30 November 2006.

9/13/06 Met Eireann, Dublin, Ireland 9

## Sample SigWx Chart for EUR area

(Produced with GIS Meteo)



9/13/06 Met Eireann, Dublin, Ireland 10

**ICT SLA for Aviation Services**

**ICT Service Level Agreement with Aviation Services Division**

An internal SLA for the delivery of ICT Services to Aviation Services Division was agreed and signed off in May 2005.


- The SLA defines & clarifies the role of the IT Division as a provider of services to the Aviation Services Division.
- Aviation Services Division has obtained ISO-9001:2000 certification.
- The first Annual internal audit is taking place presently!

9/13/06 Met Eireann, Dublin, Ireland 11

**Redesigned WEB Site Launched**

**Redesigned WEB Site Launched**

- On 30 May 2006, a newly designed Met Éireann web site ([www.met.ie](http://www.met.ie)), was launched with improved layout & accessibility.
- Content now includes animated half-hourly loops of Irish radar composites
- The web site has a channel for the provision of public service weather information in text format for mobile users (<http://mobile.met.ie/mobile/>)
- A new web-based Self Briefing Unit for airline pilots is currently being tested and will be launched in Q3, 2006.



9/13/06 Met Eireann, Dublin, Ireland 12

**Current Projects**

- **MSG / SAF Projects**
  - EUMETcast data reception system worked well up to recently (segment loss became more frequent)
    - RAM disk successfully set up on the DVB computers to prevent segment loss
    - Since installing RAM disks there have been no operating system generated disk I/O related errors
    - A watchdog function was also set up to restart DVB reception when it stops due to bird interference
  - Meteosat PDUS system will be removed from service on 14 June 2006
  - DWDSAT data reception via EUMETCast was enabled in May 2006
  - Nowcasting SAF products are generated on a Linux server & made available on our Intranet site
- **TUCSON Project**
  - 13 x new AWS stations are now installed around Ireland over past 3 years
  - Pseudo-SYNOP reports in CSV format are produced locally for NWP assimilation
  - Plans to disseminate data on the GTS were postponed to mid 2007
- **BUFR encoding of observation data**
  - BUFR Encoding & Decoding software was downloaded from ECMWF & tested in 2005
  - Project work postponed to mid 2007 due to a shortage of staffing resources

9/18/06 Met Eireann, Dublin, Ireland 13

IRELAND

IRELAND



### UPS Replaced

- At Met Éireann the Uninterruptible Power Supply (UPS) was replaced during the period 16-19 January 2006.
- A temporary generator and power supply was installed to enable electricians to safely rewire the Power Distribution Units (PDUs)
- PDUs are connected to two new 80KVA UPSs supplied by AES
- The UPSs have been connected in parallel for additional power resilience.
- ICT system managers followed a detailed schedule to:
  - Shut down & disconnect each unit from the old UPS
  - Connect to temporary power supply & reboot
  - When the new UPS was ready all systems were shut down again, disconnected from temporary power supply & connected to new UPS
- The UPS replacement project was very successful thanks to good forward planning with stakeholders and attention to detail by all.

9/13/06 Met Éireann, Dublin, Ireland 15

### Computer Room facilities?

A recent site survey concluded that the computer room facilities are inadequate for next generation NWP systems. Specifically:

**Cooling**


Computer Room is cooled by:

- 2 x floor standing Samsung 5Kwatt Units
- 3 x Fujitsu 9Kwatt ceiling cassette cooling units
- These have swirlers fitted to cause air movement.

**Floor**

- Replace raised floor – the existing raised access floor is of a very old design and is obsolete
- Reinforce Load bearing capacity of mass concrete floor
- Space in Computer Room is very limited

9/13/06 Met Éireann, Dublin, Ireland 16




## Future Plans - NWP Platform

**NWP Platforms**

- In April 2006, IBM announced 'end of life' for most of the operational system software including the AIX operating system, Loadleveler, PSSP and FORTRAN and C compilers.
- We decided not to upgrade to the latest versions of the software and the software support and maintenance contract was terminated.
- Since then we have no software support or maintenance on the IBM RS-6000/SP (SWIFT). Annual hardware maintenance contract was renewed.
- From Q3 2006, the 10 node Dell Linux Cluster, will be the operational NWP server
  - 1 x master node: 2 x Xeon 2.8Ghz/512k 533Mhz FSB, 4Gbytes ECC DDR memory
  - 9 x slave nodes: 2 x Xeon 3.2Ghz/1MB 533Mhz FSB, 2Gbytes ECC DDR memory per node with Scali interconnect and Dolphin cards
- Research staff are preparing a benchmark for HIRLAM 6.x before an ITT for the replacement of SWIFT is published.
- In Q1, 2007 SWIFT will be taken out of service.
- SWIFT replacement options include outsourcing to the new ICHEC (Irish Centre for High End Computing).
- An internal Working Group will make recommendations soon.

9/13/06 Met Eireann, Dublin, Ireland 17



## Future Plans - NWP Platform (2)

**NWP Applications planned to run on the proposed new system over next 5 years**


- **HIRLAM** forecast suite:
  - initially, this will be based on **3DVAR and HIRLAM 6.x**
  - over a 5-year period the analysis is likely to move to 4DVAR.
  - Performance requirements are based on the assumption of using 3DVAR.
- **Wave model (WAM)**, requires more modest resources compared with the forecast suite (typically 10%).
- An **ocean model** (typically ~15% of HIRLAM in terms of resources)
- A **nested version of HIRLAM** (on a Mesoscale grid).

There are no plans in the immediate future to run short-range ensemble forecasts.

**Future NWP platform or operating system?**

- A Unix/Linux OS is required
  - With usual resources (Shells: Bash, C, Korn; Compilers: C, Fortran-90, with debuggers, etc.).
  - Preference is for a system that supports MPI for parallel architecture
  - A 64-bit architecture for the OS would be our choice but 32-bit would be acceptable.

9/13/06 Met Eireann, Dublin, Ireland 18



## Future Plans – NWP Platform (3)

**Target NWP system resources?**

- The target is to double the horizontal/vertical resolution of the current operational HIRLAM (~15km)
  - i.e. a performance scaling by a factor ~16. (compared to IBM RS6000)
- Performance will be evaluated by the elapsed time for the throughput of the operational suite on the target system.


**Disk Capacity?**

- HIRLAM needs boundary files received from ECMWF each day.
  - For a single run the files occupy about 35 MB in storage but when processed to suit HIRLAM they consume about 1GB.
  - The output from a model run varies but a typical working directory for the current suite holds about 30GB.
  - With a doubling of the resolution the file sizes will increase by a factor of 4-8 i.e. about 8 GB for boundary files,
  - 150GB working HIRLAM directory (holding forecast products).

9/13/06 Met Eireann, Dublin, Ireland 19

## IRELAND


## IRELAND



### Future Plans – Forecaster Workstation

- In May 2006 two Forecaster Workstations were set up in order to facilitate detailed evaluation over the next four month period.
  - NinJo
    - Static Data
  - IBLsoft – "Visual Weather" & "Aero Weather"
    - Live data streams including RMDCN, DWDSAT, MSG and NWP data from HIRLAM, ECMWF, UKMO and DWD


9/13/06 Met Eireann, Dublin, Ireland 20



### Future Plans - RMDCN Link

- Plans are being made for the Migration of the RMDCN from Frame Relay to MPLS by the end of 2006
- Upgrade of the RMDCN link:
  - from Frame Relay (at **384kbps**)
  - to MPLS (at **1Mbps**) for the Primary Access Line.
  - NAS ISDN Backup speed will be **512kbps**, with a ...
  - Warm Standby Router
- Completed Connection Information Sheet (CIS) returned to ECMWF in May 2006
- EQUANT will be asked about configuration of local ISDN connections for both Frame Relay and NAS Backup to facilitate "Share & Move" during the transition period.

9/13/06 Met Eireann, Dublin, Ireland 21



### Migration to Phase 4 (HPCE, HPCF)

Finally....

- Thanks to ECMWF Call Desk staff for keeping Member State users well informed on a weekly basis regarding planned system sessions and their expected impacts.
- The News page on the ECMWF web site is very informative and keeps Member State users up to date on topics of common interest.
- All users have been advised of ECMWF plans to migrate from HPCD to Phase 4 (HPCE, HPCF) between now and November 2006.

9/13/06 Met Eireann, Dublin, Ireland 22

*Carmelo Gambuzza – CNMCA, Italian Met Service*

**Last year's changes at the Italian Met. Service:**

- New Meteorological Message Switching System (MMS)
- New Gateway for Ecaccess
- FireWall upgrade
- New Proxy Sever
- New Domain Name Server
- New AntiVirus-AntiSpam Server
- New Antivirus Client
- New public FTP Server


**Use of ECMWF facilities**



ECMWF, 8-9 Jun 2006

**New Meteorological Message Switching System**

**Meteorological Message-handling System**



**Why an upgraded MSS?**

- New data codification for the WMO/ICAO meteorological informations (GRIB-BUFR-CREX)
- The increased number of meteorological messages available on the International Networks (GTS-RMDCN, NATO-ACOWEX, ICAO-AFTN/CIDIN)
- New hardware resources
- A better man-machine interface
- HW and SW standard de facto


MMS is an HW and SW evolution of the old UMS - Unified Message Switching running at the IT. Met Service since 1997 - developed by GWDI (Monterey, CA) and tailored by I&SI (Rome, It.)

ECMWF, 8-9 Jun 2006

**New Meteorological Message Switching System**

**Meteorological Message-handling System**

**Hardware Configuration - O. S. - D. B.**



**MMS Servers**

- 3 HP Proliant ML350L G4
  - Primary MMS01
  - Secondary MMS02 } HotStandBy
  - Test and develop MMSDEV ColdStandBy
- 2 CPU Intel® Xeon™ Processor 3.20GHz/800MHz FSB
- 4GB 333 MHz DDR Memory
- controller SCSI Ultra320 RAID 0,1,5,10
  - » n°2 72.8GB Pluggable Ultra320 SCSI 10,000 rpm
  - » n°4 146.8GB Pluggable Ultra320 SCSI 10,000 rpm
- Gigabit Server Adapter (Interconnect)
- DAT 72 Tape Drive
- Redundance of power supply and cooling

ECMWF, 8-9 Jun 2006


ITALY

ITALY

New Meteorological Message Switching System

### Meteorological Message-handling System

Hardware Configuration - O. S. - D. B.



**Network Interface Devices**



- 4 Sync Server Model "2000 RS-232 SS-2000-232" 2 ports to interface X25 circuits:
  - ICAO-AFTN/CIDIN (link with the Italian International Communication Centre ICC - Rome Ciampino Airport)
  - Eumetsat MDD Uplink Station
- 2 Moxa Terminal Server to interface ASCII Asynchronous NATO-ACOWEX circuit

ECMWF, 8-9 Jun 2006

New Meteorological Message Switching System

### Meteorological Message-handling System

MMS Servers & Network Interface Devices


3 Servers    4 Sync Servers    Ultrium    2 Moxa Terminal Server    Metlab

ECMWF, 8-9 Jun 2006

New Meteorological Message Switching System

### Meteorological Message-handling System

Hardware Configuration - O. S. - D. B.



**New Features:**


- Fault-tolerance through hardware redundancy
- 1 Gb Ethernet Servers Interconnection
- Asynchronous channel for monitoring & backup
- All the real time processes are embedded into MMS Core
- Real time DB replication

ECMWF, 8-9 Jun 2006



New Meteorological Message Switching System

### Meteorological Message-handling System




Hardware Configuration - O. S. - D. B.

- O. S. Linux Red Hat Enterprise ES3  
Software licences with support for 3 years
- DBMS MySQL v.4.1.10a-max  
Licences and support for 3 years relating "MySQL Network SILVER"

ECMWF, 8-9 Jun 2006

New Meteorological Message Switching System

### Meteorological Message-handling System



#### Message Switching Software


Tailorization of the MS SW Core allows:

- management of standard RMDCN tcp/ftp connections into the message switching SW core
- management of non WMO standard bulletins and lists of bulletins
- easy database management
- easy update (SW standard de facto: html, php, cgi, c++, applet, javascript)

ECMWF, 8-9 Jun 2006

New Meteorological Message Switching System

### Meteorological Message-handling System



#### Differences UMS (old) – MMS (new)

UMS (from Aug 1997 to Nov 2005)	MMS (from Dec 2005)
<ul style="list-style-type: none"> <li>• X Terminal</li> <li>• Line commands from shell prompt</li> <li>• Difficulties in the management of some meteorological messages</li> <li>• Not upgradable for decoding the new WMO/ICAO message codes (Crex, Bufr)</li> <li>• Use of different protocols for different connection types</li> </ul>	<ul style="list-style-type: none"> <li>• WEB Interface</li> <li>• Windows command input</li> <li>• Management of all meteorological message types</li> <li>• Management of new message codes</li> <li>• Login with username and password</li> <li>• User profile based on the login</li> <li>• Use of FTP protocol for every connection</li> </ul>

ECMWF, 8-9 Jun 2006

ITALY

ITALY

New Meteorological Message Switching System

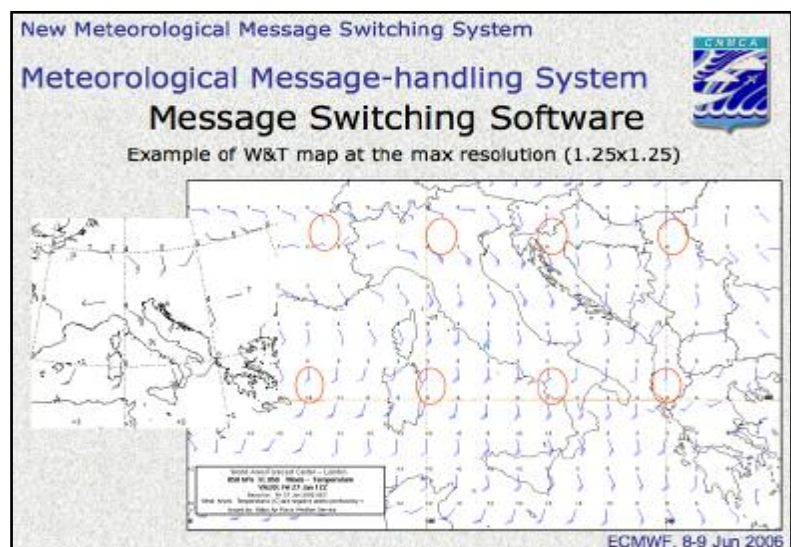
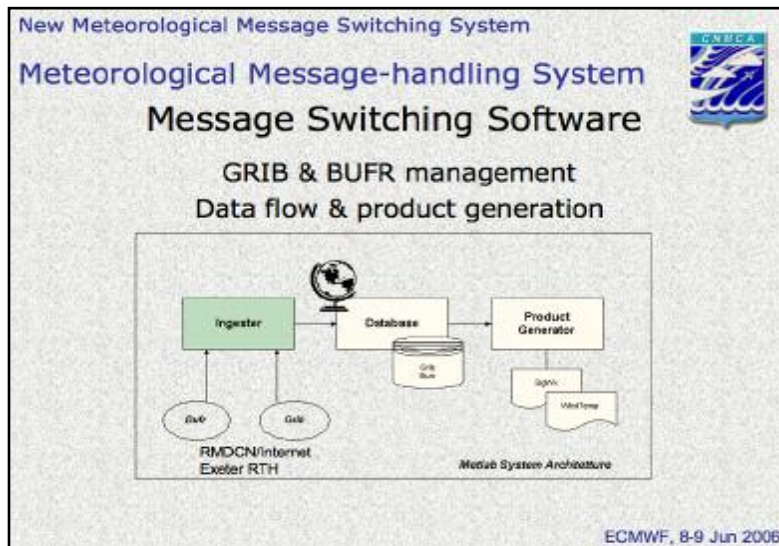
**Meteorological Message-handling System**

**Message Switching Software**


**GRIB & BUFR management**

- Binary messages GRIB & BUFR coming from U.K. Met Office (Exeter RTH) are now decoded and plotted to generate:
  - ✓ New Wind and Temperature aeronautical maps at the standard and non-standard flight levels
  - ✓ New Significant Weather maps at medium flight levels (SWM FL 100-450) and high flight levels (SWH FL250-630)
- 4 daily runs (00,06,12,18)
- New GRIB & BUFR code format in the place of the old T4 fax-simile digital coded

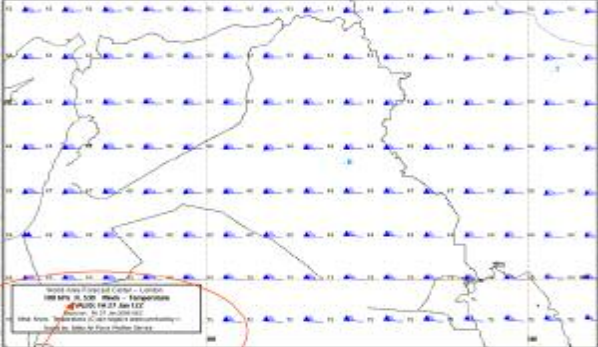
ECMWF, 8-9 Jun 2006



**New Meteorological Message Switching System**  
**Meteorological Message-handling System**  
**Message Switching Software**




Example of W&T map at the max resolution (1.25x1.25)



ECMWF, 8-9 Jun 2006

**New Gateway for Eaccess**




HW: IBM eServer xSeries 226 2xXeon 3.0Ghz

- Raid1
- OS RedHat Linux Enterprise ES3 - update 5

SW: EAccess v.3.0.0

ECMWF, 8-9 Jun 2006

**FireWall Upgrade (Two Cisco PIX 515)**




<p>Old</p> <ul style="list-style-type: none"> <li>• Firmware: 6.0(2)</li> </ul>	<p>New</p> <ul style="list-style-type: none"> <li>• Firmware: 6.3(5)</li> <li>• SW: PDM v.3.0(4) (Pix Device Manager)</li> <li>• New security policy</li> <li>• Crypted VPN password</li> <li>• New DMZ implemented by 2 switch Cisco Catalyst 3550</li> </ul>
---	--

ECMWF, 8-9 Jun 2006

ITALY

ITALY


### New Proxy Server



<p><b>Old</b></p> <ul style="list-style-type: none"> <li>• HW: HP Alpha Server au500</li> <li>• No Raid</li> <li>• OS: RH Linux 7.2 Alpha</li> <li>• Proxy: Squid 2.4.6</li> <li>• Firewall: IP-Tables 1.2.5</li> <li>• Cold Backup</li> </ul>	<p><b>New</b></p> <ul style="list-style-type: none"> <li>• HW: HP Proliant ML350 G2 x86</li> <li>• Raid1 su U320</li> <li>• OS: RH Linux Enterprise ES4 – update 3</li> <li>• Proxy: Squid 2.5.6</li> <li>• Firewall: IP-tables 1.2.11</li> <li>• Cluster compliant</li> </ul>
--	--

ECMWF, 8-9 Jun 2006


### New Domain Name Server



<p><b>Old</b></p> <ul style="list-style-type: none"> <li>• HW: HP Alpha Server au500</li> <li>• OS: Tru64 Unix 5.1A</li> <li>• SW: named included in Tru64 Unix.</li> <li>• No Backup</li> </ul>	<p><b>New</b></p> <ul style="list-style-type: none"> <li>• HW: IBM eServer 325 Type8835 (2xopteron 64bit 2.2 Ghz + 2 Gbyte)</li> <li>• Raid1</li> <li>• OS: RedHat Linux Enterprise ES4 - update 3</li> <li>• SW: Caching Nameserver v.7.3.3</li> <li>• Cluster compliant</li> </ul>
--	--

ECMWF, 8-9 Jun 2006


### New AntiVirus & AntiSpam server



<p><b>Old</b></p> <ul style="list-style-type: none"> <li>• HW: HP Proliant ML350 G2 x86</li> <li>• OS: Windows Server</li> <li>• SW: TrendMicro VirusWall 4 for win.</li> <li>• No antispam</li> <li>• No Backup</li> </ul>	<p><b>New</b></p> <ul style="list-style-type: none"> <li>• HW: IBM eServer 325 Type8835 (2xopteron 64bit 2.2 Ghz + 2 Gbyte)</li> <li>• Raid1</li> <li>• OS: RedHat Linux Enterprise ES3 - update 6</li> <li>• SW: TrendMicro VirusWall 5 for Linux + Antispam feature</li> <li>• Cluster HA: HeartBeat + DRBD (latest stable version)</li> </ul>
---	--

ECMWF, 8-9 Jun 2006


### New AntiVirus Client



<p><b>Old</b></p> <ul style="list-style-type: none"> <li>• HW: Generic Personal Computer</li> <li>• OS: Windows 2000</li> <li>• SW: TrendMicro Office Scan 6.5</li> </ul>	<p><b>New</b></p> <ul style="list-style-type: none"> <li>• HW: IBM Intellistation APro</li> <li>• OS: Windows 2003 Server</li> <li>• SW: TrendMicro Office Scan 7.0</li> </ul>
---	--

ECMWF, 8-9 Jun 2006


### New public FTP Server



<p><b>Old</b></p> <ul style="list-style-type: none"> <li>• HW: Alpha Server 1000A</li> <li>• OS: Tru64 Unix 4.0f</li> <li>• SW: ftpd included in Tru64 Unix.</li> <li>• No Backup</li> </ul>	<p><b>New</b></p> <ul style="list-style-type: none"> <li>• HW: IBM eServer 325 Type8835 (2xopteron 64bit 2.2 Ghz + 2 Gbyte)</li> <li>• Raid1</li> <li>• OS: RedHat Linux Enterprise ES3 - update 6</li> <li>• SW: VSFTD 1.2.1-3E.1</li> <li>• Cluster HA: HeartBeat + DRBD (latest stable version)</li> </ul>
--	---

ECMWF, 8-9 Jun 2006

### Use of ECMWF facilities



**ECMWF Products:**

Total of 1.4 Gbytes of dissemination GRIB data received per day:


- 600 Mbytes for internal purposes
- 800 Mbytes for external users:
  - ✓ Regional Met. Services
  - ✓ Universities
  - ✓ Research Centres
  - ✓ Civil Protection Organization
  - ✓ Antarctic research programme
  - ✓ Italian Space Agency
  - ✓ European Space Agency
  - ✓ TV Networks
  - ✓ Motorways society
  - ✓ Electrical Power company
  - ✓ Other public and private companies
  - ✓ MDD (EumetSat)

ECMWF, 8-9 Jun 2006

ITALY

ITALY

**Use of ECMWF facilities**



**ECMWF Users & Projects**

Total of 140 users:

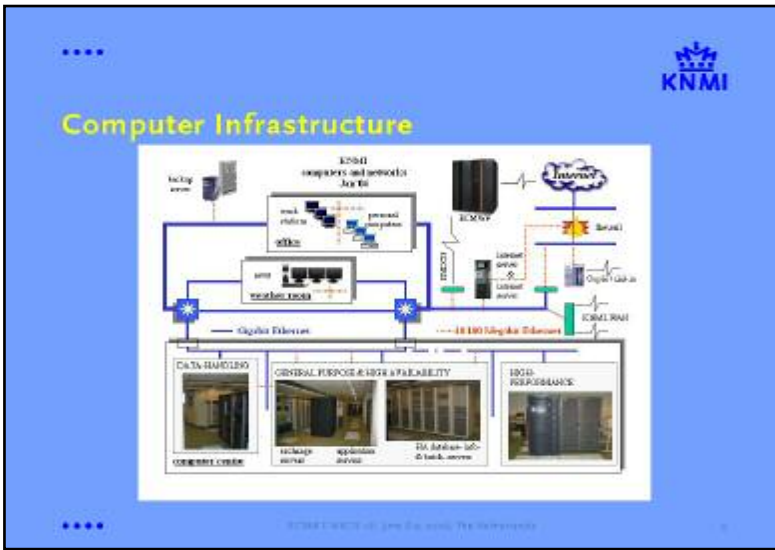
- 50 from National Met Service
- 90 from other institutions

The main usage of ECMWF services is the retrieval of MARS data associated with the decoding software to run local models, or for MAGICs and METVIEW applications.

Total of 10 Special Projects

ECMWF, 8-9 Jun 2006

Hans de Vries – KNMI, The Netherlands



- ### Computer Infrastructure: highlights
- HPC: Sun Fire 15000 (68 CPUs)
  - HA: Dec-Alpha clusters
  - StorageTek PowderHorn 9310 tape silo (100 TB data; 200 TB — 800 TB space)
  - Linux Workstations (Compaq, HP)
  - Citrix servers for Windows applications in a UNIX environment
  - 100 Mbit/s – 1 Gbit/s internal network

### Computer Infrastructure: Main connections

Connection	Capacity [bit/s]
Internet	1 G
Firewalls	4 x 100 M
RMDCN Access line	384 k
ECMWF	> 256 k
Met Office (GTS)	> 64 k

**Remote Access**

**FirePass (F5 Networks)**

- email
- Intranet
- ssh-vpn

NETHERLANDS

NETHERLANDS


\*\*\*\* 

### Computer Infrastructure: developments

- New website (July 2005): <http://www.knmi.nl>
- First steps towards KODAC (KNMI Operational Data Centre), a portal for meteorological data
- New developments in Radar
- New computers for HPC and HA

\*\*\*\*

TCMWF-0020-06 June 8-9, 2006 The Netherlands 1


\*\*\*\* 

### Computer Infrastructure: developments in 2006

**New HPC system** operational Hirlam and research


**SCI ALTIX 3700 bx2**

- 240 Itanium-2 processors
- 3 GB/cpu
- 6 TB storage
- SuSE Enterprise Linux



\*\*\*\*

TCMWF-0020-06 June 8-9, 2006 The Netherlands 1


\*\*\*\* 

### Computer Infrastructure: developments in 2006

**New HA system** operational models and databases

**SUN Fire V440**

- 17 systems
- UltraSPARC III processors
- 8 GB/cpu
- 12 TB storage
- Clustering:
  - 2 clusters: 3 systems with 4 cpu/system
  - 3 clusters: 2 systems with 2 cpu/system
- Solaris 10



\*\*\*\*

TCMWF-0020-06 June 8-9, 2006 The Netherlands 2



\*\*\*\*

KNMI

## Computer Infrastructure: developments in 2006

**New standard Linux distribution**

- Now: SuSE 9.1
- Red Hat Enterprise 4?
- SuSE Enterprise 10?
- SuSE 10.1?

\*\*\*\*

© 2006 ECMWF, all rights reserved, The Netherlands

\*\*\*\*

KNMI

## ECMWF Users

Status on May 17, 2006

Total	92		
Last logged in since 1 March	68	Access to HPCF	54
Last logged in earlier in 2006	6	Access to Special Projects	16
Not logged in more than 1 year	8	Outside KNMI	20

\*\*\*\*

© 2006 ECMWF, all rights reserved, The Netherlands

\*\*\*\*

KNMI

## Projects

The chart displays the fraction of NL allocations (%) for various projects from 1995 to 2006. The Y-axis ranges from -25 to 100. The X-axis shows the Year. The chart is divided into two main categories: Regular NL (top) and SPNL (bottom). The legend includes: Atoz, Comp. use (yellow), Atoz, Comp. prod (blue), HIRLAM decadal (green), Oceanography (red), Predictability (dark blue), HIRLAM (light blue), and Satellite HMF (purple).

\*\*\*\*

© 2006 ECMWF, all rights reserved, The Netherlands

NETHERLANDS

NETHERLANDS

Accounting

GEREEDSCHAPPEN  
KNMI SCHWAF geleend in 2004

Kategorie	Begroed	Geleend	Looftijd (Aanv.)
	2004	2004	2004
Totaal	20000	10000	10000
Schwaf	10000	10000	10000
Schwaf	10000	10000	10000
Schwaf	10000	10000	10000
Schwaf	10000	10000	10000
Schwaf	10000	10000	10000
Schwaf	10000	10000	10000
Schwaf	10000	10000	10000
Schwaf	10000	10000	10000
Schwaf	10000	10000	10000

Accounting

Kategorie	Begroed	Geleend	Looftijd (Aanv.)
	2004	2004	2004
Totaal	20000	10000	10000
Schwaf	10000	10000	10000
Schwaf	10000	10000	10000
Schwaf	10000	10000	10000
Schwaf	10000	10000	10000
Schwaf	10000	10000	10000
Schwaf	10000	10000	10000
Schwaf	10000	10000	10000
Schwaf	10000	10000	10000
Schwaf	10000	10000	10000

Hiram

Cooperation with the Afadin consortium for research and development

Management team  
 Programme Manager: Jeanette Onvlee (KNMI, NL)  
 Project Leader for System and Applications: Xiaohua Yang (DMI, DK)

Usage of ECMWF HPCF

- Hiram Reference System: meso-scale
- Hiram Ensembles
- Use national resources as far as possible
- Special Projects:
  - Reference System: 1250 kSBU in 2007
  - Ensembles: 600 kSBU in 2007

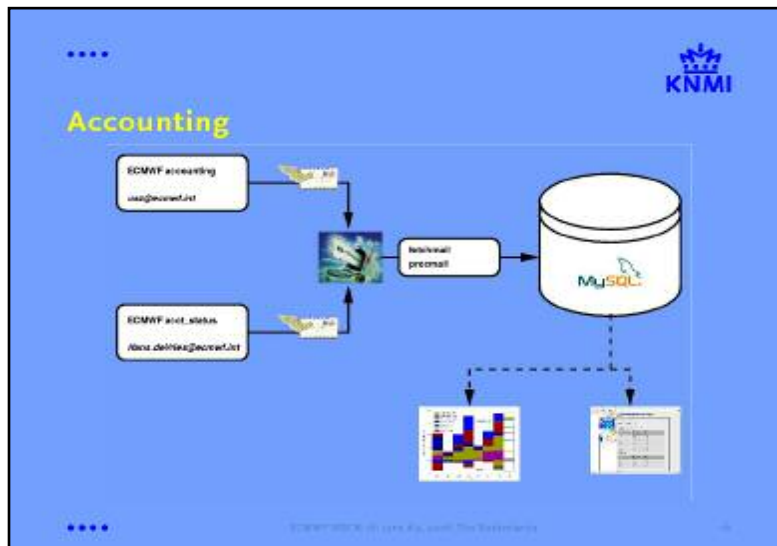
..... KNMI

### Comments

- The support from ECMWF is very much appreciated, e.g. User Support (Dominique Lucas, John Greenaway), Call Desk
- Concern about the many power failures in the last year
- Data required in netCDF by research community
- Dutch Sec Rep works on an overview of computer facilities in the HR NPI's

.....

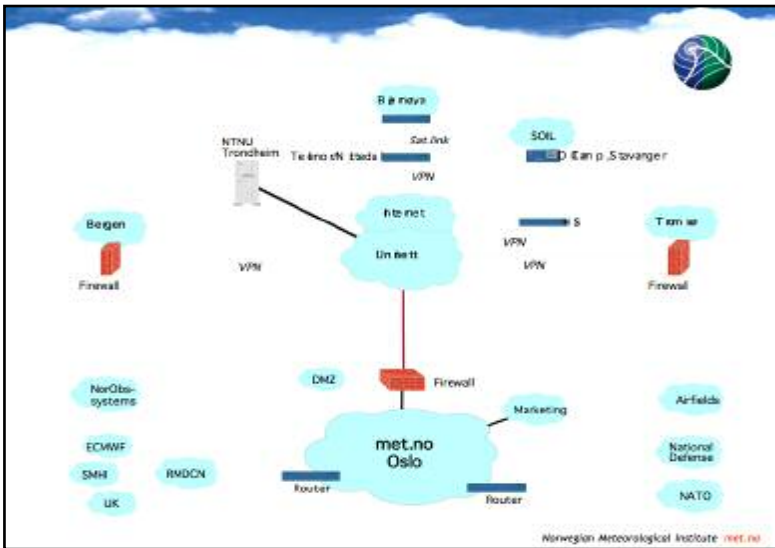
ECMWF 8-9 June 2006, The Netherlands 14



NORWAY

NORWAY

Rebecca Rudsar – Norwegian Meteorological Institute



### New since last time

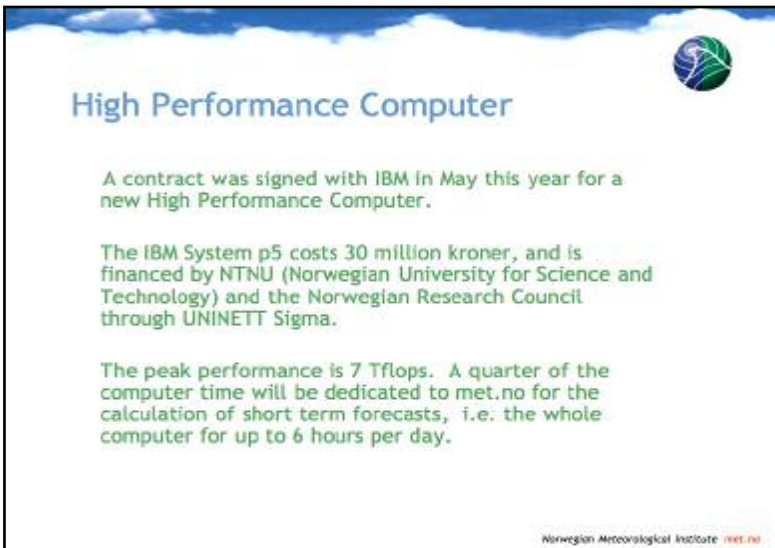
- **FVTKilden.met.no** webservice (UML) for military use.
- **Laptops and VPN** development of a new standard for met.no.
- **Web Map Service** the first server based on bdiana was established.
- **External computer room** further work has been done to establish a reserve computer room with duplicate network and operational computers.
- **Upgrade of the military network** including the communication with the airport offices.

Norwegian Meteorological Institute met.no

### New since last time (continued ....)

- **Text TV** automation of production and development of new products.
- **System for Continuous Presentation** development and installation at 3 different sites.
- **21 new servers** mostly upgrades for existing systems.
- **Approx. 100 new PC/workstations** all Linux workstations have been upgraded to Fedora Core.

Norwegian Meteorological Institute met.no



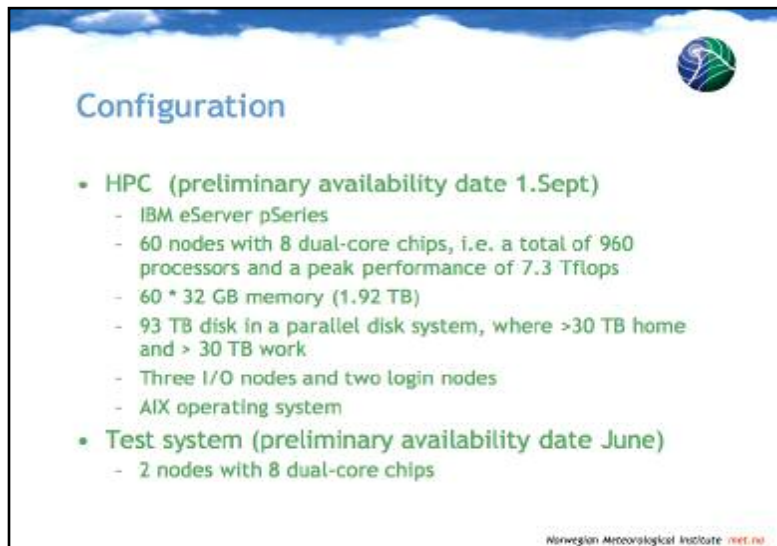
## High Performance Computer

A contract was signed with IBM in May this year for a new High Performance Computer.

The IBM System p5 costs 30 million kroner, and is financed by NTNU (Norwegian University for Science and Technology) and the Norwegian Research Council through UNINETT Sigma.

The peak performance is 7 Tflops. A quarter of the computer time will be dedicated to met.no for the calculation of short term forecasts, i.e. the whole computer for up to 6 hours per day.

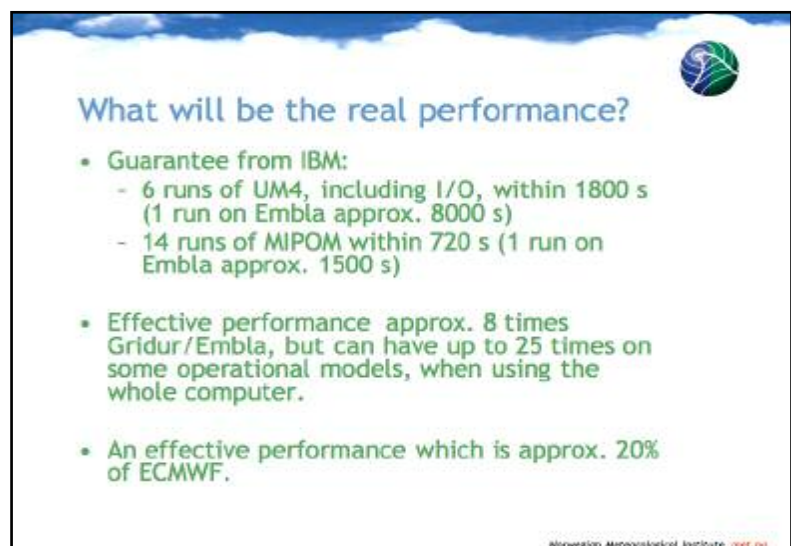
Norwegian Meteorological Institute met.no



## Configuration

- HPC (preliminary availability date 1.Sept)
  - IBM eServer pSeries
  - 60 nodes with 8 dual-core chips, i.e. a total of 960 processors and a peak performance of 7.3 Tflops
  - 60 \* 32 GB memory (1.92 TB)
  - 93 TB disk in a parallel disk system, where >30 TB home and > 30 TB work
  - Three I/O nodes and two login nodes
  - AIX operating system
- Test system (preliminary availability date June)
  - 2 nodes with 8 dual-core chips

Norwegian Meteorological Institute met.no



## What will be the real performance?

- Guarantee from IBM:
  - 6 runs of UM4, including I/O, within 1800 s (1 run on Embla approx. 8000 s)
  - 14 runs of MIPOM within 720 s (1 run on Embla approx. 1500 s)
- Effective performance approx. 8 times Gridur/Embla, but can have up to 25 times on some operational models, when using the whole computer.
- An effective performance which is approx. 20% of ECMWF.

Norwegian Meteorological Institute met.no

NORWAY

NORWAY

### Timeplan

- **Move applications**
  - From now until the autumn
  - Operational models must be ready to run on the new computer in September. Other models must be ready by 31.12, which is when the old HPC is to be turned off.
  - Use test systems, for example, that at ECMWF
- **Redesign and implement the operational suite**
  - Ready for running the acceptance test starting 4.9
  - Co-operation with NTNU and IBM for testing the batch system and scheduler
  - Co-operation with NTNU, IBM and Met Office for the optimisation of UM

Norwegian Meteorological Institute [met.no](http://met.no)

### ECMWF disseminated data

- **via RMDCN : 875 Mbyte**
  - low resolution DA(114), EF(40), Wave(10)
  - boundary conditions BC1(596) and BC2(101)
  - disseminated directly to the operational computer.
- **via Internet : 8.96 Gbyte**
  - high resolution DA(3.89), EF(1.13) and Wave(1.32) (old hi-res data(2.6))
  - disseminated to a computer in the DMZ.

Norwegian Meteorological Institute [met.no](http://met.no)

### ECMWF data

- **via Internet using ECTRANS : 3.7 Gbyte**
  - TEPS data
  - MARS data for UV forecasting
  - MARS data for Norwegian Institute for Air Research
  - seasonal forecast data
  - disseminated to a computer in the DMZ.

Norwegian Meteorological Institute [met.no](http://met.no)




## ECMWF Products

The data disseminated via RMDCN is more or less the same as that disseminated via Internet, the difference being the geographical resolution. The data from both streams are processed similarly.

There are parallel jobs for

- a) extracting data from the grib files and distributing to customers,
- b) converting the grib data to the internal met.no format. The converted data are written on separate files, e.g. 'ec\_atmo\_geo\_00\_r.felt' and 'ec\_atmo\_geo\_00\_i.felt'.
- c) checking which data has arrived and whether an event should be set. The event triggers a set of jobs which further process the data, depending on which data has arrived.

Norwegian Meteorological Institute met.no




## ECMWF Products

The system for switching between the two datasets, which we had just started to design last year, has just been put into the operational suite. The low resolution RMDCN-data, 'ec\_atmo\_geo\_00\_r.felt', is interpolated to the same resolution as the internet-data, 'ec\_atmo\_geo\_00\_r\_interp.felt'.

Normally the link 'ec\_atmo\_geo\_00.felt' would point to the high resolution internet-data, 'ec\_atmo\_geo\_00\_i.felt'. If the internet-data is delayed, the link name is switched to point to the interpolated data.

The criteria for switching will probably have to be adjusted, after we have had some more experience.

Norwegian Meteorological Institute met.no



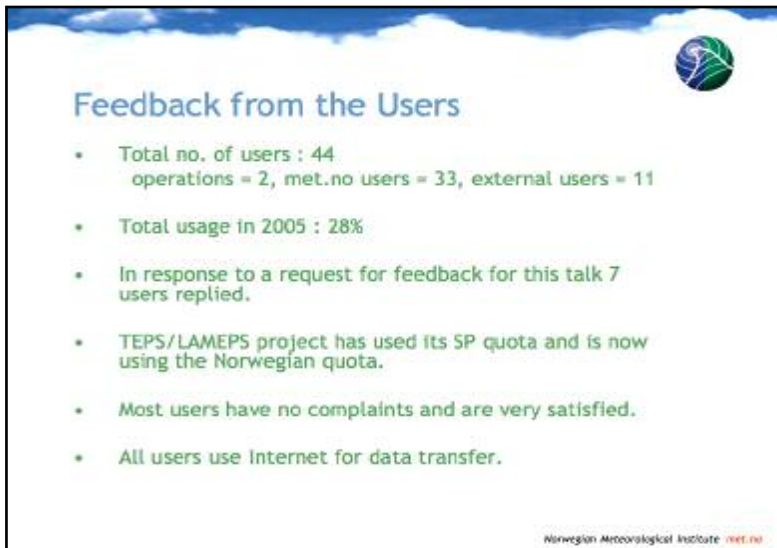
## ECMWF Projects

- Ozone as a climate gas.
- REGCLIM: Regional Climate Modelling.
- HIRLAM project.
- Targeted ensembles providing boundary values for limited area models.

Norwegian Meteorological Institute met.no

NORWAY

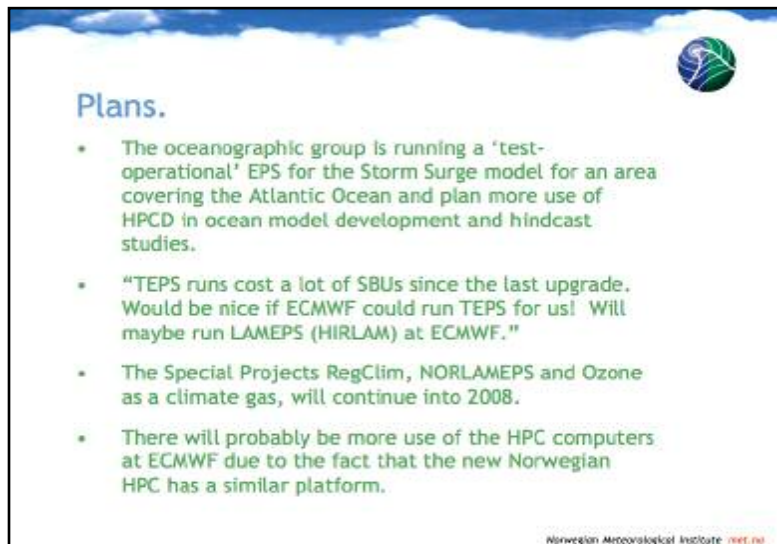
NORWAY



### Feedback from the Users

- Total no. of users : 44  
operations = 2, met.no users = 33, external users = 11
- Total usage in 2005 : 28%
- In response to a request for feedback for this talk 7 users replied.
- TEPS/LAMEPS project has used its SP quota and is now using the Norwegian quota.
- Most users have no complaints and are very satisfied.
- All users use Internet for data transfer.

Norwegian Meteorological Institute [met.no](http://met.no)



### Plans.

- The oceanographic group is running a 'test-operational' EPS for the Storm Surge model for an area covering the Atlantic Ocean and plan more use of HPCD in ocean model development and hindcast studies.
- "TEPS runs cost a lot of SBUs since the last upgrade. Would be nice if ECMWF could run TEPS for us! Will maybe run LAMEPS (HIRLAM) at ECMWF."
- The Special Projects RegClim, NORLAMEPS and Ozone as a climate gas, will continue into 2008.
- There will probably be more use of the HPC computers at ECMWF due to the fact that the new Norwegian HPC has a similar platform.

Norwegian Meteorological Institute [met.no](http://met.no)



Roland Cotariu – Romanian Meteorological Service

**THE MAIN ACTIVITIES**



- **BASIC OPERATIONAL ACTIVITY**
  - Weather forecasting
  - Observation system
  - Telecommunication
- **RESEARCH ACTIVITY**
  - Numerical modelling
  - Climatology and Agrometeorology
  - Remote Sensing and GIS
- **EDUCATION and TRAINING**

2

18th Computing Representatives Meeting, 8-9 June 2006

**Meteorological observations in the surface network**




Network within the National Meteorological Administration is made up of 7 Regional Meteorological Centres:  
North Transilvania (Cluj), South Transilvania (Sibiu), Banat - Crişana (Timişoara),  
Oltenia (Craiova), Muntenia (Bucureşti), Moldova (Iasi) and Dobrogea (Constanţa).

3

18th Computing Representatives Meeting, 8-9 June 2006

**COMPUTING RESOURCES**



- > PC cluster (for both LM and ALADIN models integrations)
  - 14 dual Xeon nodes
  - 2 Gb memory each node
  - 200 Gb SCSI hard disk on master node
  - Linux REDHAT 3 (next month will be changed with a DEBIAN version)
  - OpenMedia
  - MPICH library
  - PGI Fortran
  - INTEL Fortran
- > SUN 4500 Enterprises (for ALADIN model integration)
  - 8 processors
  - SOLARIS 8
  - Fortran
- > SUN Ultra 90 (for visualization at each regional meteorological center)
  - 2 sparc processors
  - 1 Gb memory
  - Solaris 8
  - NexRep software (produced by HARRIS Company in U.S.)

**MAIN PROJECT for 2006-2007 : HPC computer**


4

18th Computing Representatives Meeting, 8-9 June 2006

ROMANIA

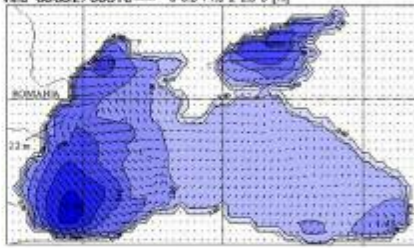
ROMANIA

**ALADIN integration characteristics**  
 on SUN 4500 platform (8 processors)




Hidrotatic version  
 Regular grid in Lambert projection  
 Dynamical adaptation mode  
 Initial and lateral boundary condition from the French global model ARPEGE (6 hours frequency)  
 Digital filter initialisation  
 2 TL semilagrangean scheme

Base: 0000925 00UTG (p=2) VADROW = Significant wave height / direction  
 Wind: 0000527 00UTG (p=2) 0 0.5 1.5 2 2.5 3 [m]




5

18th Computing Representatives Meeting, 8-9 June 2006




**ALADIN DOMAINS**



144x144 points; dx=10km  
dt= 450 s

41 levels




120x90 points; dx= 24km  
dt=900 s

RUN 00 => 78 h ( integration time :- 2 h )  
 RUN 12 => 60 h ( integration time :- 2.5h )

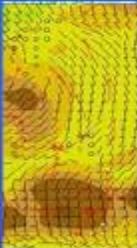
6

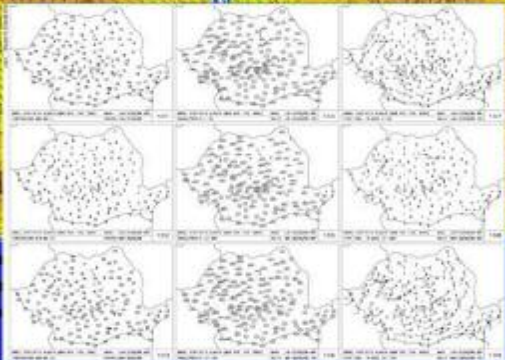
18th Computing Representatives Meeting, 8-9 June 2006

National Meteorological Administration - Romania




**ALADIN Output (2)**





- parameters  
 cloudines, p  
 - for all rom

18th Computing Representatives Meeting, 8-9 June 2006



### LM integration characteristics

**Horizontal resolution:**  
14 km ( 81x73 , grid points)

**Vertical resolution:**  
35 levels

**Time step:** 80 sec.

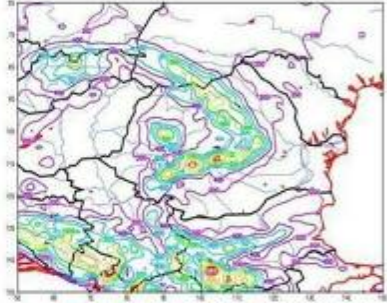
**IC & LBC:**  
GME 00 , every 3h

**Grid-scale pp:**  
2-ice category scheme, prognostic


**Correction scheme:**  
Tiedke

**Data assimilation:** no

**Time of integration:** 54h



18th Computing Representatives Meeting, 8-9 June 2006



### MODEL IMPLEMENTATION

**Period of LM compilation :** 15 min

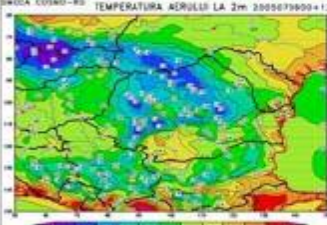
**Period of LM execution (including post-processing) 26 procesors :**

- 81x73x35 ; dt=80s ; normal conditions - 14 km  
3 min / 6h forecasts - 30 min / 54h forecasts
- 161x147x35 ; dt=60s ; normal conditions - 7 km  
22 min / 6h forecasts - 3.20h / 54h forecasts
- 41 x 41 x 50 ; dt=30s ; normal conditions - 2.8 km  
14 min /6h forecasts - 54 min / 24h forecasts

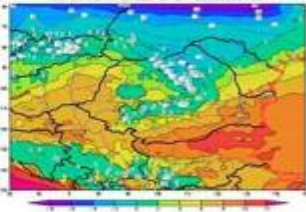
( PGI Fortran compiler)

18th Computing Representatives Meeting, 8-9 June 2006

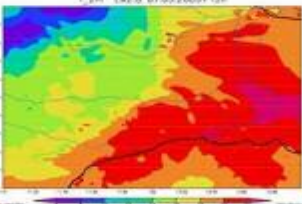
### EXAMPLES IV – 2m TEMPERATURE




**14km**



**7km**




**2.8km**



18th Computing Representatives Meeting, 8-9 June 2006

## ROMANIA

## ROMANIA



**ECMWF Users**

- Currently 15 registered users , all at NMA
- Many users of ECMWF web site

**ECMWF User support**


- All users are very satisfied with support
- Web site useful for support
- Excellent organisation

**ECMWF data used at NMA**

- Products from ECMWF web site
- Data received by dissemination
- Data retrived from MARS

11

18th Computing Representatives Meeting, 8-9 June 2006



**ECMWF software used at NMA**

- EMOSlib
- MAGICS
- METVIEW

= implemented on Linux PC  
= compiled with PGI and INTEL FORTRAN

currently used in the NMA operational suite for both ECMWF and limited area numerical models which are operationally run ( ALADIN, HRM, LM)


- ECMWF Web site

EPS products from ECMWF web site are used in operative weather forecast

12

18th Computing Representatives Meeting, 8-9 June 2006

Vladimir M. Dimitrijevic – Republic Hydro-Meteorological Service of Serbia




### Report of Republic Hydro-Meteorological Service of Serbia

#### ECMWF products in operational use

- ECPDS-ECMWF Product Distribution System
- Products from deterministic forecast (resolution 0.5° and 0.25°) in GRIB and BUFR based on 00Z and 12Z
- Boundary conditions for limited area Eta NWP model based on 00Z and 12Z
- Initial and BC for MM5 NWP model based on 00Z and 12Z-experimentally
- ECMWF software MetView, MAGICs, SMS
- MARS files on request
- Web available daily forecast including EPS
- Monthly and Seasonal forecast products-experimental use


18th meeting of Computing representatives, 8-9 June 2006



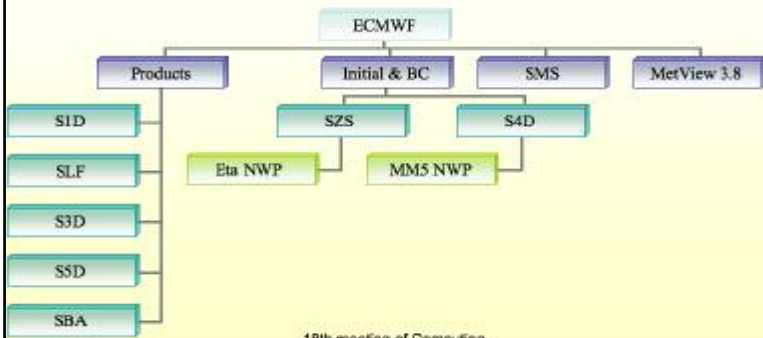
### Report of Republic Hydro-Meteorological Service of Serbia

Name	Data type	No. of products	size
SZS	BC – regional Eta	1868	14.53Mb
S1D	Deterministic forecast 0.5°x0.5°	1372	52.65Mb
S2F	Monthly forecast – raw data	121	489.63Kb
S3D	Experimental (CAPE)	50	2.32Mb
S4D	Initial and BC for MM5 – experimental run	1246	43.59Mb
S5D	Deterministic forecast 0.25°x0.25°	1802	2.46Mb
SBA	BUFR - Weather Parameters	468	506.39Kb

18th meeting of Computing representatives, 8-9 June 2006



### Report of Republic Hydro-Meteorological Service of Serbia



```

graph TD
    ECMWF[ECMWF] --> Products[Products]
    ECMWF --> InitialBC[Initial & BC]
    ECMWF --> SMS[SMS]
    ECMWF --> MetView[MetView 3.8]
    Products --> S1D[S1D]
    Products --> SLF[SLF]
    Products --> S3D[S3D]
    Products --> S5D[S5D]
    Products --> SBA[SBA]
    InitialBC --> SZS[SZS]
    InitialBC --> S4D[S4D]
    SZS --> EtaNWP[Eta NWP]
    SZS --> MM5NWP[MM5 NWP]
    
```

18th meeting of Computing representatives, 8-9 June 2006

SERBIA

SERBIA




### Report of Republic Hydro-Meteorological Service of Serbia

Graphical presentation of ECMWF Products

**192.168.30.20 - /Bitmaps/**



- Tuesday, March 21, 2006 8:49 AM <dir> [agro](#)
- Tuesday, March 21, 2006 8:49 AM <dir> [hydro](#)
- Tuesday, March 21, 2006 8:50 AM <dir> [temperature](#)
- Thursday, June 01, 2006 7:50 AM <dir> [WeatherMaps](#)
- Thursday, June 01, 2006 8:01 AM <dir> [Eatheta](#)
- Thursday, June 01, 2006 7:55 AM <dir> [MaxT](#)
- Thursday, June 01, 2006 7:41 AM <dir> [Metgram](#)
- Thursday, June 01, 2006 7:55 AM <dir> [MinT](#)
- Thursday, June 01, 2006 7:00 AM <dir> [Prec03](#)
- Thursday, June 01, 2006 7:12 AM <dir> [Prec06](#)
- Thursday, June 01, 2006 7:53 AM <dir> [Prec12](#)
- Thursday, June 01, 2006 7:55 AM <dir> [Prec24](#)
- Thursday, June 01, 2006 8:05 AM <dir> [RT](#)

18th meeting of Computing representatives, 8-9 June 2006




### Report of Republic Hydro-Meteorological Service of Serbia

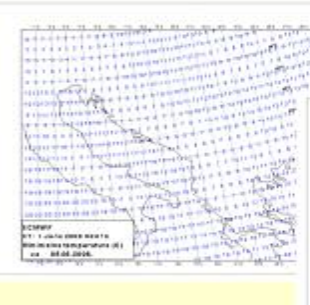
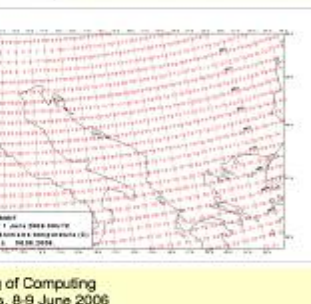
Graphical presentation of ECMWF Products

18th meeting of Computing representatives, 8-9 June 2006



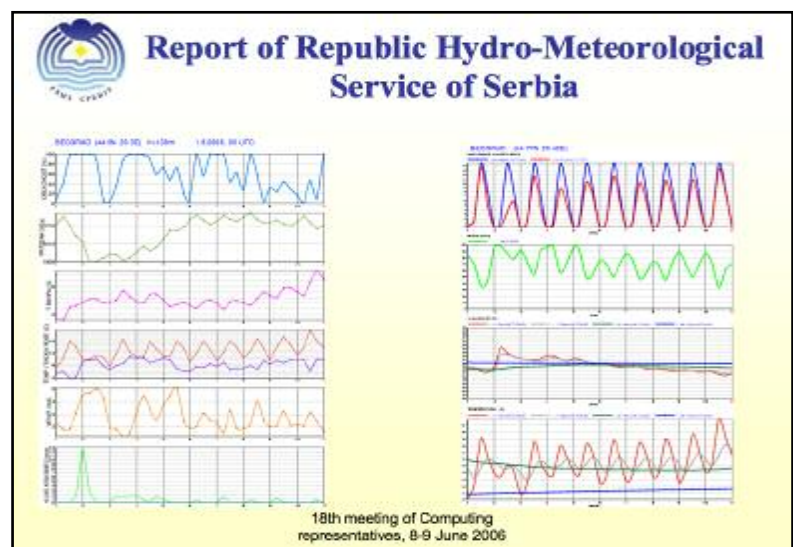
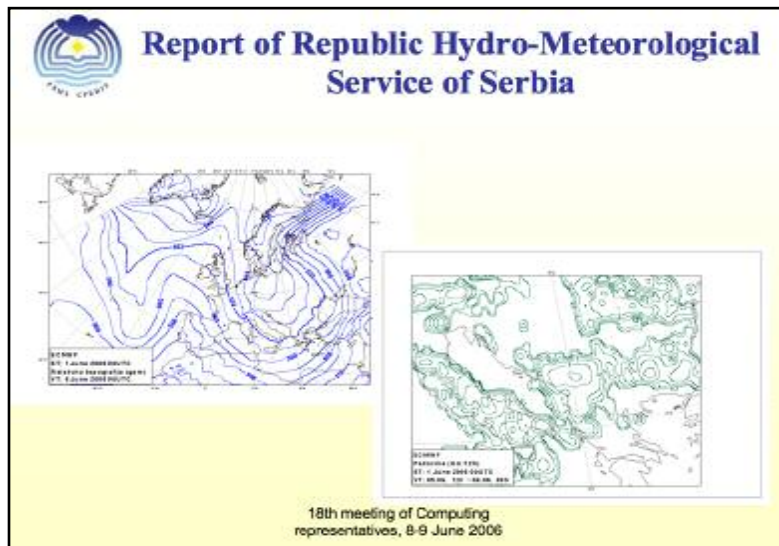
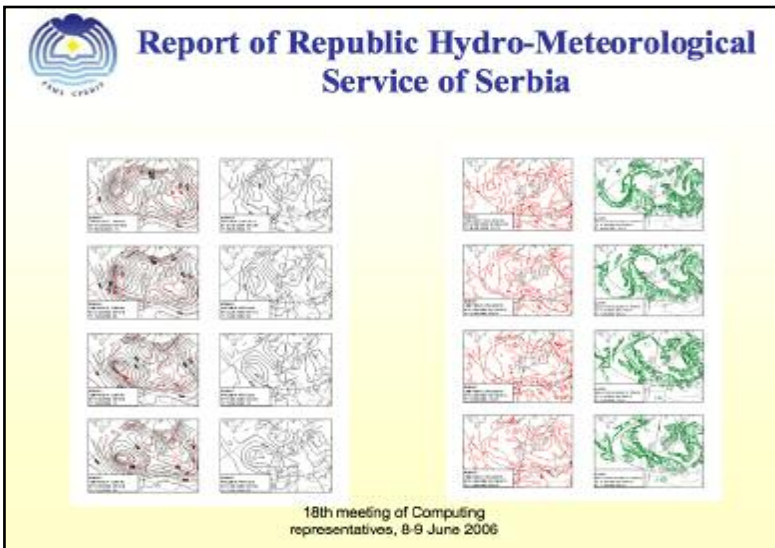
### Report of Republic Hydro-Meteorological Service of Serbia

18th meeting of Computing representatives, 8-9 June 2006

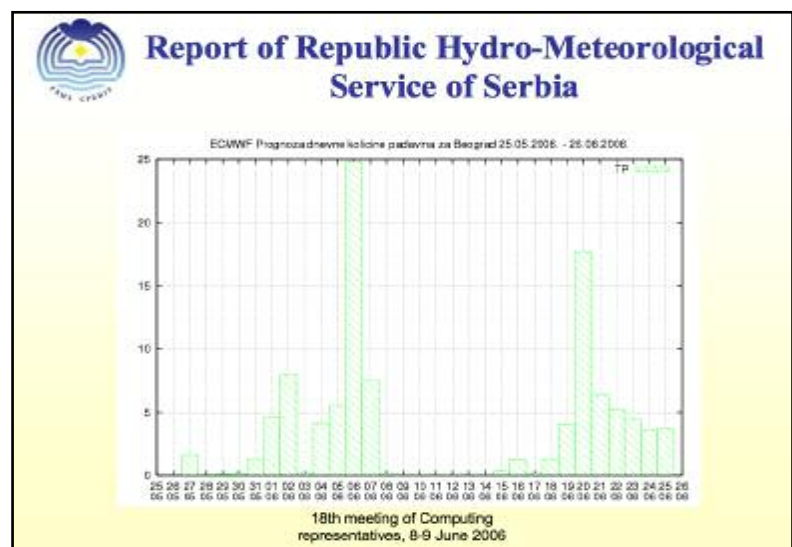
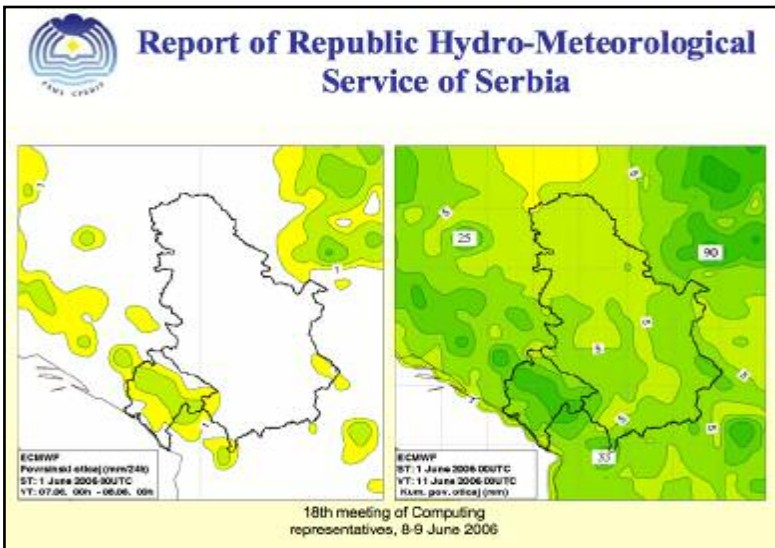
SERBIA

SERBIA



SERBIA

SERBIA





SERBIA

SERBIA

**Report of Republic Hydro-Meteorological Service of Serbia**

SMS

18th meeting of Computing representatives, 8-9 June 2006

**Report of Republic Hydro-Meteorological Service of Serbia**

Products of The Eta NWP model based on 00Z and 12Z ECMWF BC

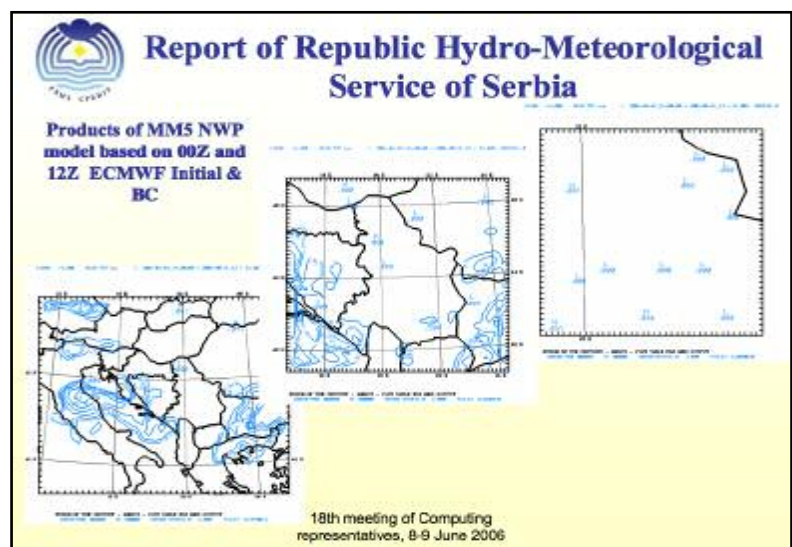
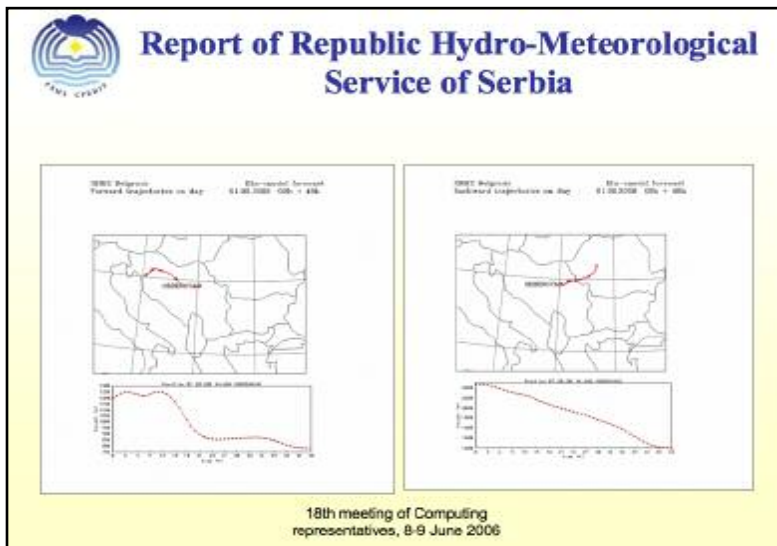
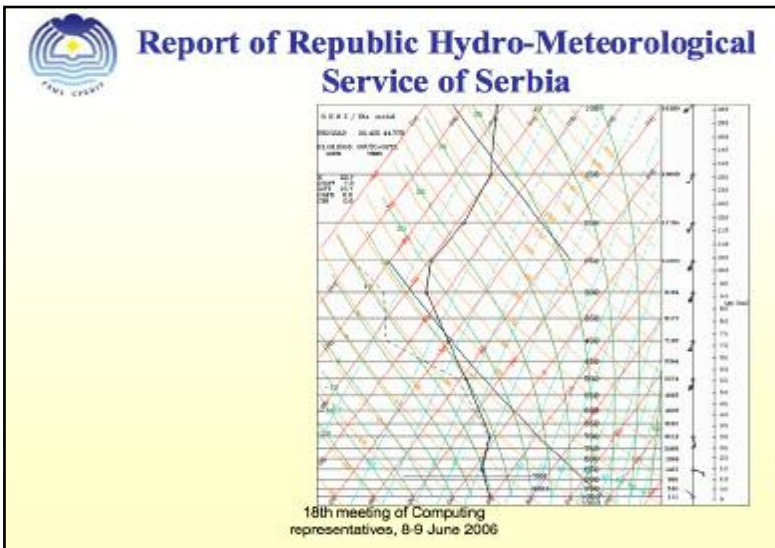
18th meeting of Computing representatives, 8-9 June 2006

**Report of Republic Hydro-Meteorological Service of Serbia**

18th meeting of Computing representatives, 8-9 June 2006

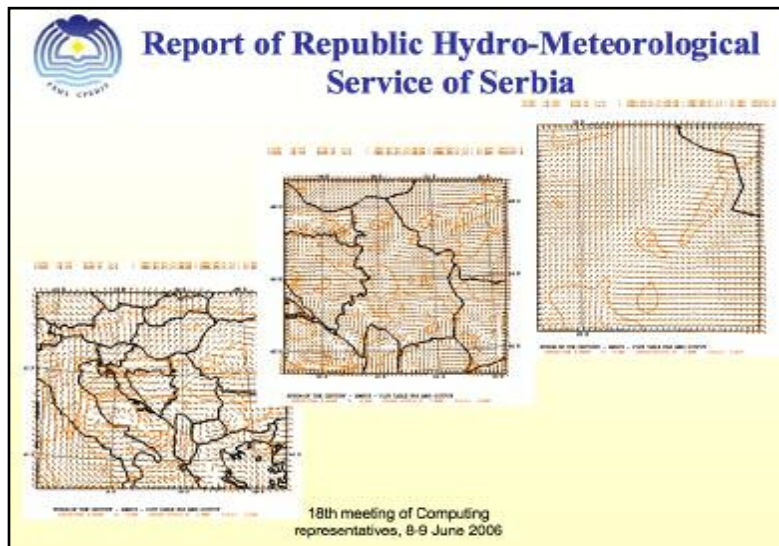
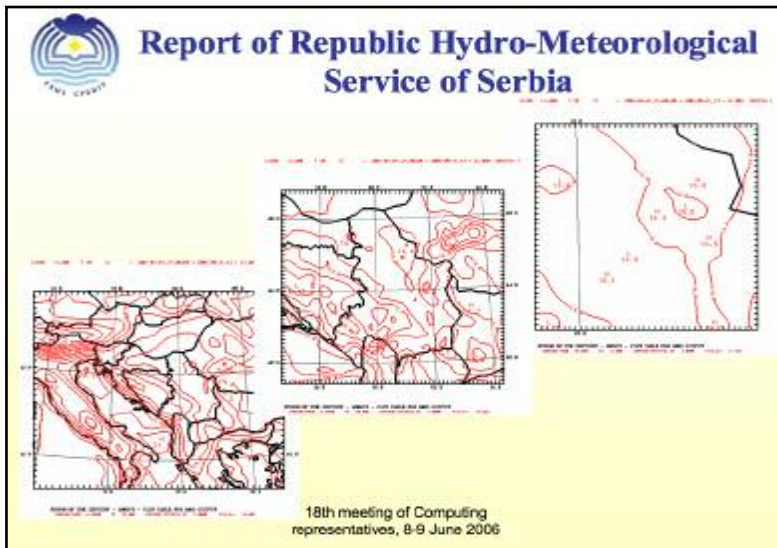
SERBIA

SERBIA



SERBIA

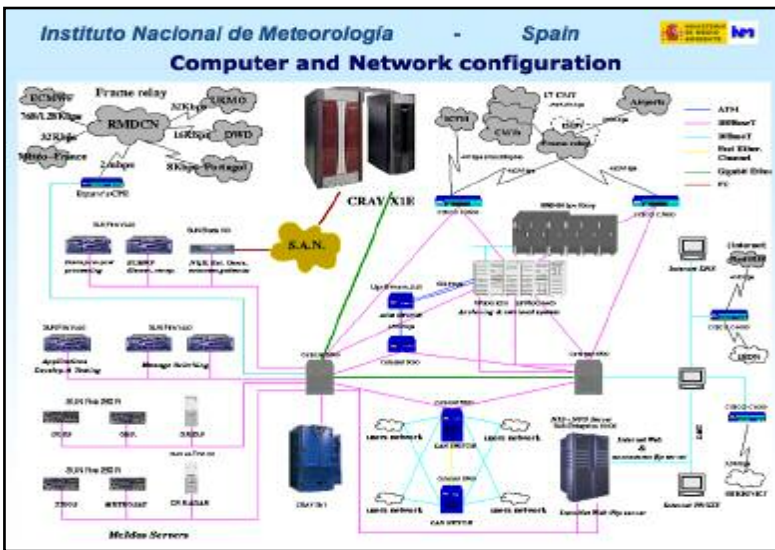
SERBIA



SPAIN

SPAIN

Eduardo Monreal – Instituto Nacional de Meteorología, Spain



**Instituto Nacional de Meteorología - Spain**

### HPC, the CRAY X1E

- **Final phase accepted end of June 2005:**
  - 128 MSPs in 16 modules
  - 2.3 Tflops theoretical peak performance
  - 320 Gbytes of memory
  - 3.6 Tbytes of direct attached disk space
  - A Storage Area Network with:
    - 3 Tbytes of disk space
    - ADIC scalar 100 robotic system: 14 TB
    - ADIC's Stornext Management Suite (Stornext File System & Stornext Storage Manager)
- **Additional 25 TB disk space to be installed 3Q2006:**
  - Direct attached disk to be doubled
  - Usable SAN disk space to be increased up to 18 TB


**Instituto Nacional de Meteorología - Spain**

### The new Data Handling System

- **ITT issued end July 2006**
- **Contract awarded to SGI**
  - 3 storage levels in a 4 Gb/s FC SAN:
    - Fast FC disks: 4.5 TB
    - SATA disks: 55 TB
    - ADIC scalar 2K tape library: 6 LT03 tape drives, 80 TB initial capacity
  - 4 Altix 350 servers:
    - 4 Itanium2 1.5 Ghz processors
    - 16 Gbytes memory
  - Software: CXFS & DFM (HSM)
- **To be installed by July 2006**
- **Expected operational 4Q2006**


SPAIN

SPAIN

*Instituto Nacional de Meteorología* - *Spain* 


### Main UNIX servers

- **New Sun Fire V440 servers**
  - A cluster of 2 Sun Fire V440:
    - 4 x 1.6 Ghz processors, 8 GB of memory
    - Sun StorEdge 3510 FC array: 432 GB
    - These two systems now support the Message Switching (SCM) and Graphics Dissemination (DIAGRAM)
  - 2 Sun Fire V440:
    - 4 x 1.6 Ghz processors, 8 GB of memory
    - 288 GB of disk each
    - Replaced the Sun Ultra 250 that handled data pre-processing and Report DB as well as reception of ECMWF dissemination, most of graphics production and post-processing
  - 1 Sun Fire V440:
    - 4 x 1.6 GHz processors, 8 GB of memory
    - 288 GB of disk space
    - Applications development and testing

*Instituto Nacional de Meteorología* - *Spain* 

### WAN developments

- **Connection to the Internet**
  - upgraded to 40 mbps in March
  - further upgrade to 100 mbps planned for 2007
- **Connection to RMDCN**
  - Upgraded end of May
    - Access line: 2 mpbs
    - CIR to ECMWF: 768/128 kbps (in/out)
- **Connection to Regional Centres (CMT)**
  - Currently on Frame Relay at 512 kbps
  - Upgrade to 2 mbps planned for 2006: 13 sites out of 17
  - Migration to MacroLan planned for 2007:
    - 13 sites running at 10 mbps
    - 4 sites at 2 mbps

*Instituto Nacional de Meteorología* - *Spain* 

### ECMWF users

- **Users Registered**
  - Currently 72 (71 last year)
  - 60 INM and 12 Universities
  - 56 logged in 2006 (45 in April & May)
- **Work done is for the most part MARS data retrievals, particularly access to ERA-40 dataset**
- **Metview used in batch mode to produce derived EPS products**
- **All users use the Internet for data transfer**
  - Data transfer over RMDCN only for dissemination and other operational use

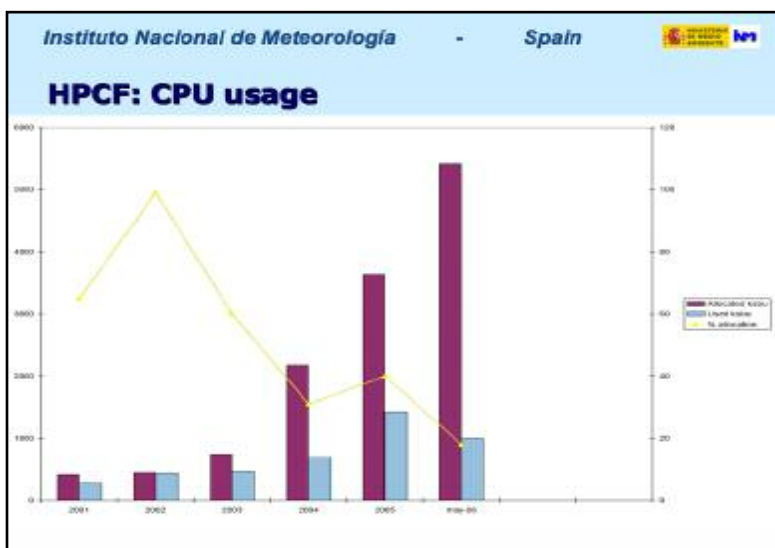
SPAIN

SPAIN

*Instituto Nacional de Meteorología - Spain*

### Use of HPCF

- **Last year 17 users used the HPCF. They mainly worked in the following areas:**
  - HIRLAM model runs using the reference system
  - Trajectory computations
  - Studies on Climate variability
  - Statistical downscaling of seasonal forecast outputs
  - Integration of RCA/HIRLAM within EU ENSEMBLES Project framework
- **HIRLAM accounts for the 99% of 2005 used allocation:**
  - Data assimilation group (4D-Var): 59%
  - Model development: 27%
  - Parallel runs: 13%



*Instituto Nacional de Meteorología - Spain*

### Experience using ECMWF computers

- **Comments & queries from users:**
  - Just a few replied to a request for feedback
  - As a general comment, users are very satisfied with ECMWF computer services. They have no complaints
  - Assistance & help from User Support very much appreciated
  - Only one query: A user asks whether it would be possible to have a sftp plugin on eaccess
- **My own comments:**
  - Last Monday I tried EMS for the first time to register a new user
    - Not difficult procedure, well documented
    - Pleased by its potential, very useful for profiling review
    - ...still need more experience!

SPAIN

SPAIN



Instituto Nacional de Meteorología - Spain

### Future plans

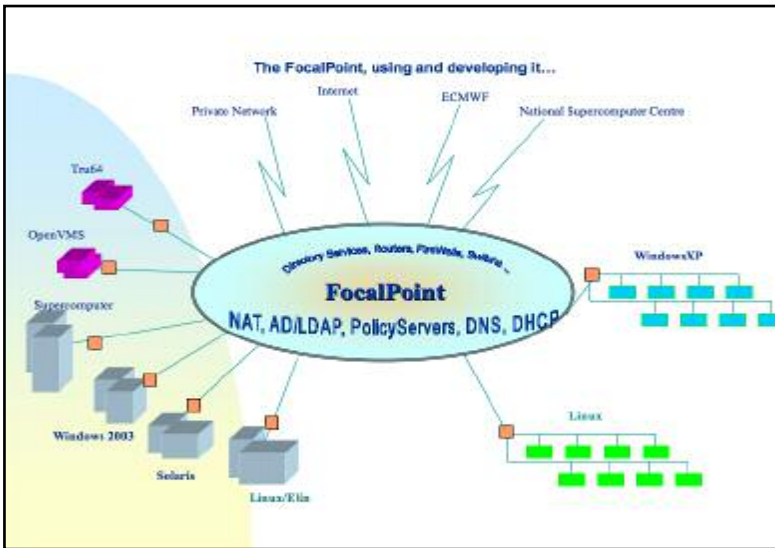
- **New projects at ECMWF:**
  - Integration of RCA/HIRLAM for dynamical downscaling of outputs from Seasonal Forecast System-3:
    - Phase 1: as soon as hindcasts from SFS-3 are available
    - Phase 2: on a regular basis when SFS-3 is operational

*In reply to a query from E. Monreal, L. Gougeon replied that it was hoped to implement sftp in the next implementation of the ECaccess gateway.*

SWEDEN

SWEDEN

Rafael Urrutia – Swedish Meteorological and Hydrological Institute (SMHI)



**IT infrastructure – Some activities in focus**  
SMHI, ITI 2006

- Focal Point and Authentication-
  - Unix/Windows via Goldfinger/AD
  - Internet services, HTTP and FTP via Goldfinger/SevAD
- Linux and Windows Standard-
  - Linux/Elia and Windows2003/Edwin
- Network management-
  - Hamilton Monitor/Statistics, Based on OpenSource Nagios & Cacti
- Printer system for Windows, Unix and VMS-
  - LpdLISST, Based on OpenSource Linux HA, CUPS and Samba
- Backup and Data Management
  - Central resources via SBCSI & Fiber Channel (FC) gives all 'data' is centralized
- Consolidation via Virtualisation
  - We use VMWare, VirtualCenter...
- New web/ftp cluster farm
  - Based on Elia and OpenSource, Apache, Tomcat, JBoss, LVS

**IT infrastructure - Directory Services**  
SMHI, ITI 2006

**Goldfinger**, is an Active Directory based Authentication system used in our FocalPoint.

Goldfinger used for Windows, Linux, Linux systems and application.

Goldfinger implementation are used in both Intranet and Internet zone

Goldfinger/Internet is used for Authentication needs on HTTP and ftp.

Goldfinger/AD supports Ldap v3, SSL v2/3, TLS 1, good Integration via PAM\_LDAP and NSS\_LDAP.



**IT infrastructure – Linux and Windows Server Standard**  
SMHI, ITI 2006

- Linux / Elin, based on Linux Redhat EE and SMHI's ELIN
- SystemImager and PPM
- Windows 2003 environment HP's Rapid Deployment
- Linux systems are installed or recovery by ELIN management system.
- Elin/PPM can deploy a new server or recovery on existing one in a total time of 30 min.
- Windows servers with HP's Rapid Deployment Pack for installation of new servers.
- Estimated time for disaster recovery is less than one hour.

**IT infrastructure – Network management**  
SMHI, ITI 2006

- OP5 delivered Monitor and Statistics
- Applications based on Nagios and Cacti

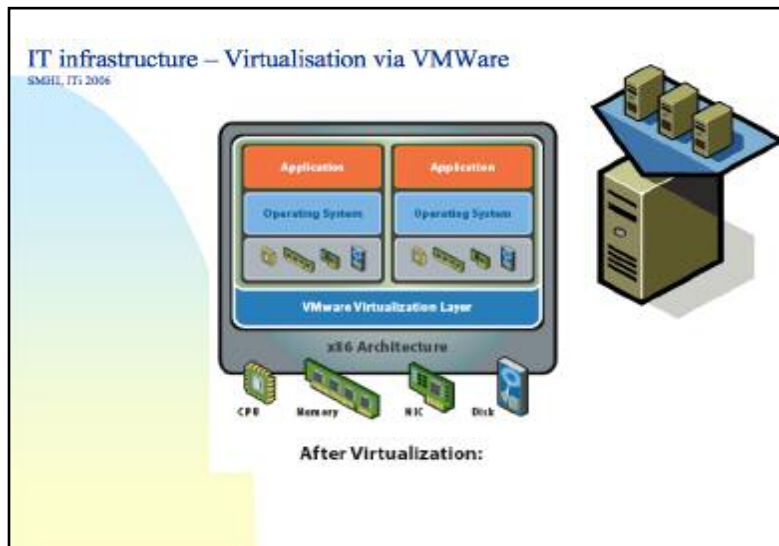
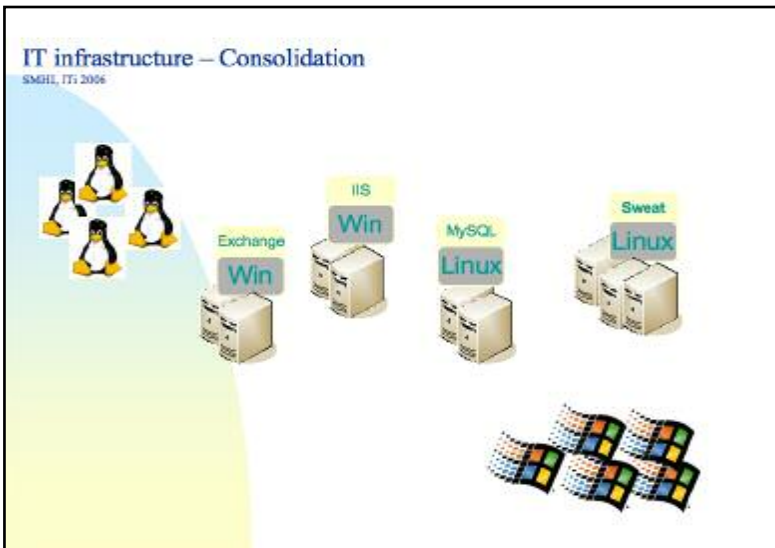
**IT infrastructure – Data Management**  
SMHI, ITI 2006

Actual view of file sharing, 2006

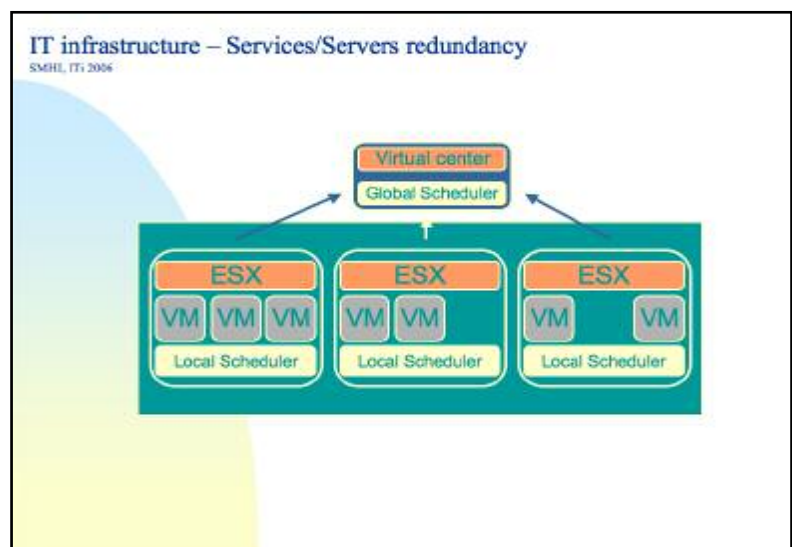
**NAS, CIFS and NFS**  
**NAS, SCSI and Fiber Channel**  
**Multiple servers and volumes**  
One interface for the users  
Different classification of data  
Faster restore and more flexible scalability solutions

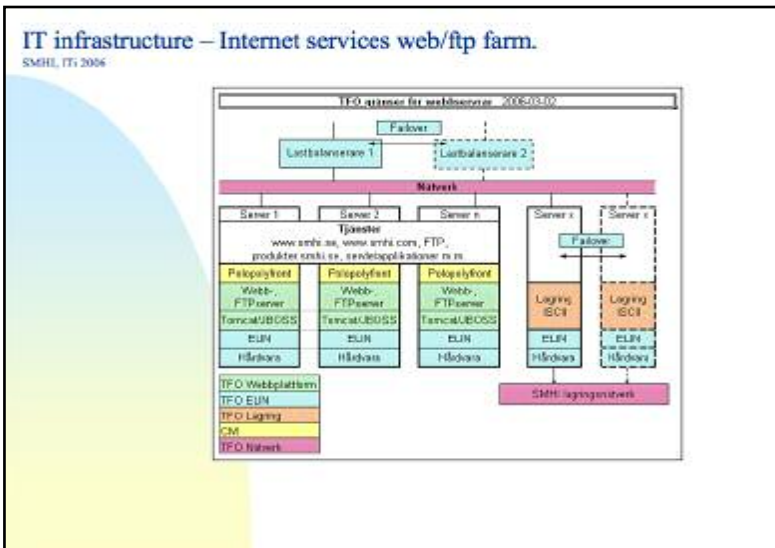
SWEDEN

SWEDEN



After Virtualization:





SMHI, ITI 2006

**More information**

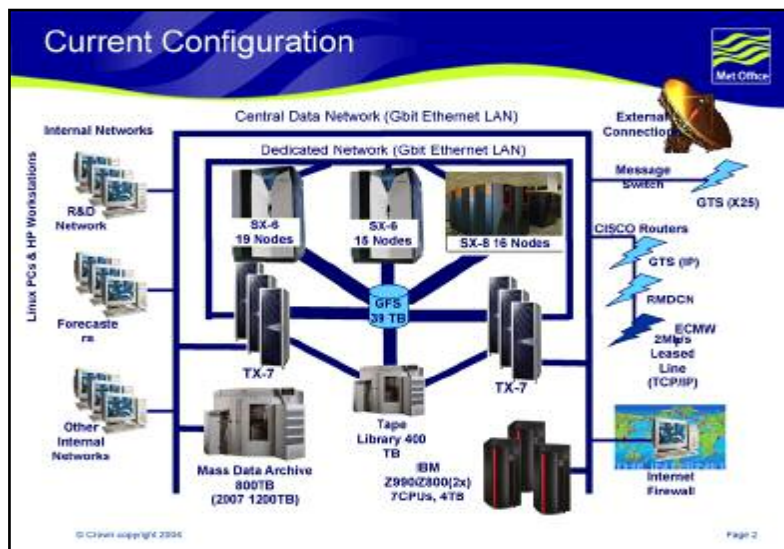
- PAM\_LDAP and NSS\_LDAP  
• <http://www.smhi.se>
- Active Directory  
• <http://www.microsoft.com>
- Network management  
• <http://www.netman.org>  
• <http://www.cacti.net>  
• <http://www.snmp.com>
- Supercomputer  
• <http://www.csc.fi/en/sectors/inf.html>
- Image management  
• <http://www.systems.com>
- Web/Ftp cluter farm  
• <http://www.fusionserver.com>  
• <http://www.sparc.com>  
• <http://www.ibm.com>  
• <http://www.pcsipd.org>

At the end of his presentation R. Urrutia invited Computing Representatives whose service was considering the introduction of Open Source solutions such as those introduced by SMHI to contact him for any advice he could give.

UNITED KINGDOM

UNITED KINGDOM

Roddy Sharp – Met Office, United Kingdom



### NEC SX-6 & SX-8 Supercomputers

- 16 Node (128 processor) NEC SX-8
- 34 node NEC SX-6 system split between 2 halls
  - One 19 Node (152 Processor) Cluster
  - One 15 Node (120 Processor) Cluster
- Accessed via 6 NEC TX-7 front ends
- 39 TB of disk space forming a GFS across 2 halls


Operational forecast Switched to SX8 13<sup>th</sup> April 2005

© Crown copyright 2004 Page 3

### Desktop Environment


- Desktop PCs used across the Met Office
  - Around 300 Linux desktops for scientists
  - Around 600 Windows XP desktops for other staff
- 15 Linux & 21 HP workstations used as servers for compute intensive work
  - HP workstations being rationalised, and gradually replaced by Linux compute servers as machines are retired.
- Network capacity
  - Gigabit Ethernet backbone
  - 100Mb/s to individual desktop

© Crown copyright 2004 Page 4

**ECMWF Users** 


- **Users Registered**
  - Currently 148 registered users (131 last year)
  - Majority at Met Office
- **Many users make simple MARS data retrievals**
  - Find system easy to use
  - Good documentation
- **Few users with large / complex data sets**
  - THORPEX work requires larger transfers to/from Met Office
    - This has further increased the load on leased line

© Crown copyright 2004 Page 5

**Use of ECMWF Systems** 

- **ECaccess**
  - Currently running 4 gateway servers (2 research, 2 operational)
  - All gateways give access via the leased line
- **Metview**
  - Many local macros
  - Automated MARS retrievals
- **Use of HPCF**
  - Unified Model ported to IBM (using MPP code)
  - Used 54% of total SBU allocation in 2005
  - Up from 26% in 2004
  - With several large projects getting underway, likely to use most of allocation this year.

© Crown copyright 2004 Page 6


**Current use & Projects** 

- **Multi-model ensembles, QUMP and FORMOST** : continue as described last year
- **THORPEX work new this year**
  - Production style runs started end of March
  - 24 member ensemble forecast
  - Run using UM vn6.0
  - Forecast to 15 days
  - Contributes to TIGGE
- **10 Special Projects**

© Crown copyright 2004 Page 7

UNITED KINGDOM

UNITED KINGDOM

**Future Plans** 


- Upgrade to Leased line (July 2006)
  - Currently 2Mb/s moving to 10Mb/s
- Continue with long-range and seasonal forecasts
  - ENSEMBLES Project (EU FP6) will start soon
    - Seasonal to decadal predictions of climate with ocean data assimilation
    - Ensembles will be used to sample uncertainty in both initial conditions and model parameters
- Continue THORPEX project contribution.
  - Possible upgrade to UM vn6.1

© Crown copyright 2004 Page 8

CTBTO

CTBTO

Gerhard Wotawa – Preparatory Commission for the Comprehensive Nuclear-Test-Ban Treaty Organization, Provisional Technical Secretariat, Vienna, Austria




## Usage of ECMWF Data Services by the Provisional Technical Secretariat

### Daily data transfer from ECMWF to PTS (1)

- The PTS currently uses the MS-SMS-System to retrieve daily analysis data from ECMWF

- Triggering points are availability of Analysis 12 UTC (around 18 UTC) and Analysis 00 UTC (around 6 UTC)
- Data are retrieved on all 91 model layers on a global 1°x1° grid
- Analyses 06/12/18 are supplemented by 4V-data or FC 3 hours data, as available, to obtain a 3 hour data resolution
- Five 3-D and a number of 2-D fields are extracted

IDC/WM
June 2006
Page 2



### Daily data transfer from ECMWF to PTS (2)

- Data are transferred through the Internet via ECACCESS software

- The MS jobs utilise the ECMWF "ectrans" utility for transfer from the ECMWF server ecgate to the PTS
- The PTS has currently running ECACCESS Version 2.0.2 Cycle 003
- ECACCESS runs on the PTS file transfer server ftp.ctbto.org (Netra T1 Server, Solaris 9) situated in the DMZ outside the PTS internal LAN
- The DMZ hosts are configured with public IP addresses, but protected against the Internet with a firewall

- Connection speed, data volume and speed of data transfer

- The PTS is connected with the internet via FastEthernet to a pair of routers of two different ISPs with a bandwidth of 2 MBit and 2x2 MBit, respectively.
- The transferred data volume from ECMWF is currently 504 MB per day (8 files of 63 MB); before introduction of the ECMWF high resolution model upgrade, it was 340 MB per day
- Typical transfer speed is 5-8 minutes per file; data are typically in-house about one hour after the execution of the retrieval routines has been triggered

IDC/WM
June 2006
Page 3



## Data usage

- The PTS uses ECMWF data as part of its daily Atmospheric Transport Modelling (ATM) Operations

- The data serve as input to the Lagrangian Particle Diffusion Model FLEXPART (Version 5)
- With FLEXPART, Source-Receptor Sensitivity (SRS) information is computed for all Radionuclide Stations that are part of the International Monitoring System
- The SRS information is made available to the States Signatories

- The PTS uses ECMWF data for development purposes

- ECMWF data are fed into a infrasound signal propagation model currently under development; aim is to locate origin of an infrasound signal
- Data are fed into the Lagrangian Particle Diffusion Model HYSPLIT Version 4.7 operated on the Development LAN


- The PTS uses ECMWF data for International backtracking exercises with the WMO

- A near-real-time international backtracking response system is currently build up in cooperation between CTBTO and WMO

IDC/WM
June 2006
Page 4

CTBTO

CTBTO

  
PREPARATORY COMMISSION

### Next steps

- The PTS plans to install ECACCESS Version 3 during 2006 (postponed from 2005)
- The PTS needs to upgrade and maintain the data retrieval software supporting the FLEXPART model. This software does currently not work for 0.5 or 0.25 degree grids and for sub-domains. One solution to be investigated would be the inclusion of the basic functionality into the ECMWF Mars system


IDC/MJune 2006Page 5



Marc Jenner – EUMETSAT


**Introduction to U-MARF**

- U-MARF: Unified Meteorological Archiving & Retrieval Facility
  - Unified
    - Ingests data from all EUMETSAT programmes
  - Archiving
    - archives data in a long-term reliable archive
    - 20 years of operation data already available
    - overall storage capacity required: **2200TB** all **2200TB near-line**
  - Retrieval
    - provides products for internal and external users

8-9 June 2006      ECMWF 18th Meeting of Computing Representatives      


**Introduction to U-MARF cont.**

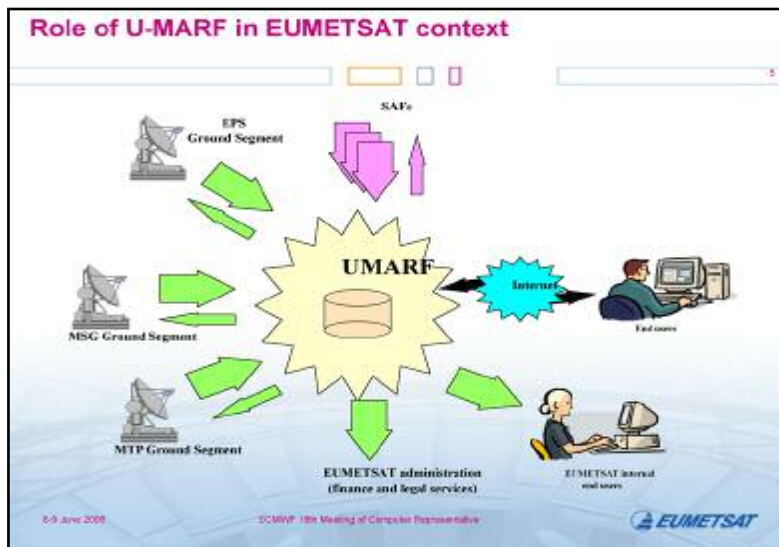
- V1
  - MSG connectivity, data archiving & internal retrievals
  - MTP migration from MTP-MARF
  - MSG & MTP Product Formatting & Delivery
  - MSG & MTP User Services (dedicated to Geostationary satellite)
- V3
  - EPS connectivity, data archiving & internal retrievals
  - V1 functionalities (MSG & MTP)
  - Drastically improved performances and capacity
  - Improved USV (function to process Polar satellite and technology: migration from HTML to Java)

8-9 June 2006      ECMWF 18th Meeting of Computing Representatives      

**U-MARF and ECMWF**

- Current
  - Supporting ERA-40
- Possible future development
  - Access to U-MARF archive data from ECMWF for EPS related experiments (IASI level 0 data)

8-9 June 2006      ECMWF 18th Meeting of Computing Representatives      



- ### Types of Orders
- Orders are submitted by registered users thanks to the USI 'web' interface
  - There are different type of ways to issue orders:
    - **Query & order:** issue a search query and browse the result list to select products to order
    - **Direct order:** issue a search query for which each product from the result list is automatically ordered
    - **standing order:** ask to be periodically delivered all new products corresponding to a given set of criteria. The submitted order has to be authorized by the U-MARF operator before being activated.
    - **Bulk order:** the user wants to retrieve all the products for a limited set of metadata: sensing time range, product type, satellite, algorithms. As the order size will surely exceed his quota, the order has to be authorized by the U-MARF operator before being processed.
- 8-9 June 2006      ECMWF 18th Meeting of Computing Representatives      EUMETSAT

- ### Delivery Methods
- Electronic delivery
    - FTP
    - HTTP
  - Physical delivery
    - DLT tape
    - SDLT 320
    - LTO-2
    - DAT (DDS 2-4)
    - CD-ROM
    - DVD-ROM
- 8-9 June 2006      ECMWF 18th Meeting of Computing Representatives

- ### Delivery Formats
- EPS Native
  - OpenMTP
  - Native
  - HDF-5
  - JPEG
  - PNG
  - BUFR
  - GRIB
- 8-9 June 2006      ECMWF 18th Meeting of Computing Representatives      EUMETSAT

**Operations**

- All operations are automatically handled by the U-MARF system except for:
  - inserting blank tapes into the storage libraries and tape devices
  - handling of offline tapes from the storage library
  - completion of order delivery
  - authorization of orders from users, in special cases

8-9 June 2006      ECMWF 18th Meeting of Computing Representatives      EUMETSAT

**U-MARF Performance**

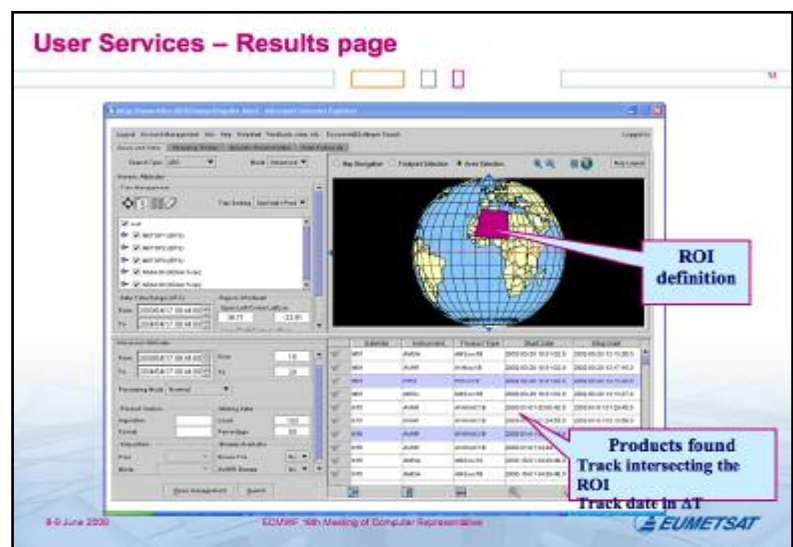
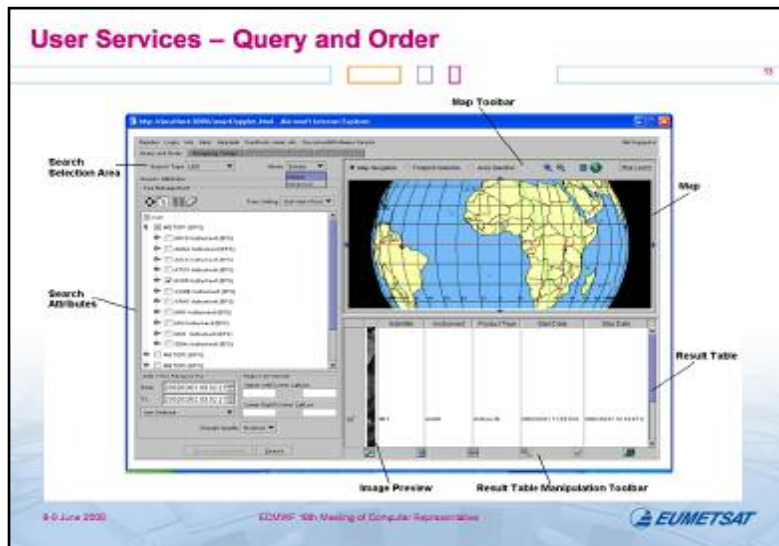
- The U-MARF V3 requirements are challenging regarding both ingestion and retrieval activities
  - Daily ingestion in worst case = 28GB for MTP + 180GB for 3 MSG + 470GB for 2 METOP + 10 GB for 2 NOAA
  - Total = **1288 GB/day** with 540GB in double copy
  - **600GB/day** for Retrieval
  - Actual retrieval performance reaching **1.5 TB/day**
  - Most popular products also available on disk, allowing quicker retrievals

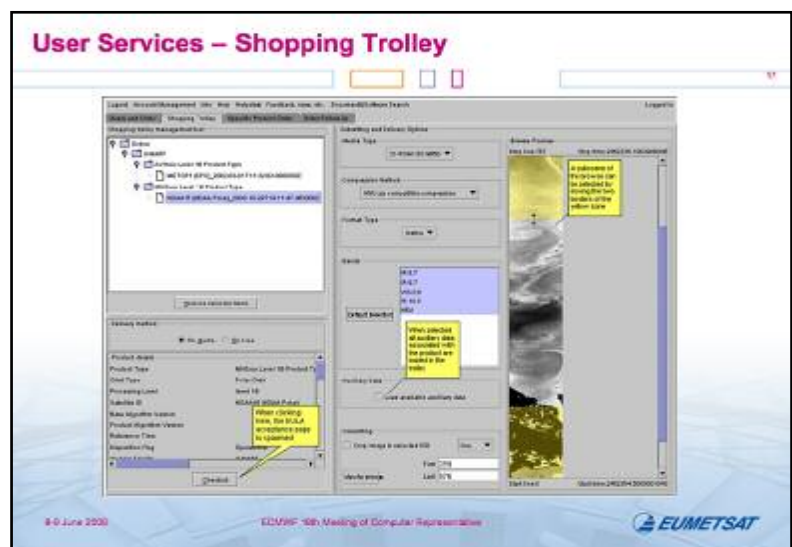
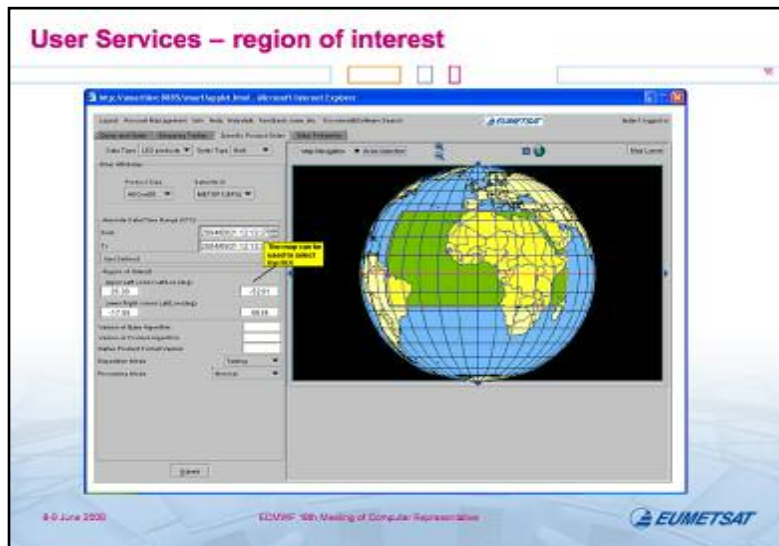
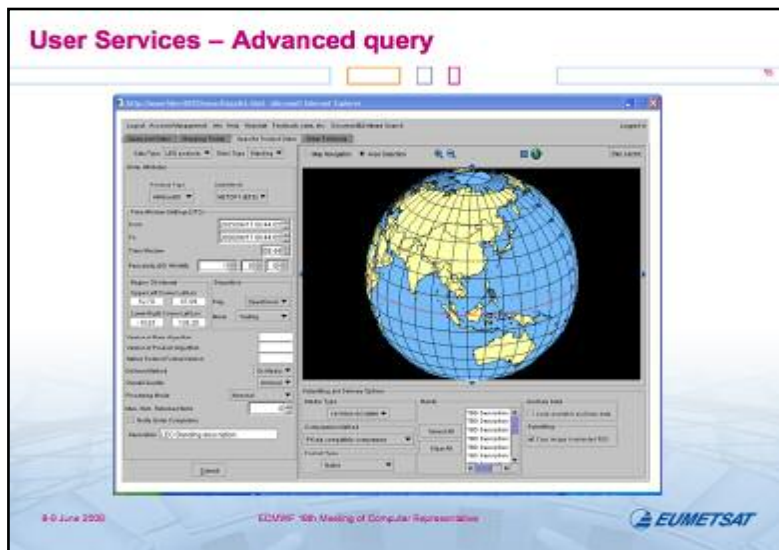
8-9 June 2006      ECMWF 18th Meeting of Computing Representatives      EUMETSAT

**User Services - Online Ordering**

- 2 ways of ordering online
  - HTML Based (Geo data)
  - Java based (all data sets) currently under beta-testing

8-9 June 2006      ECMWF 18th Meeting of Computing Representatives      EUMETSAT







**ANNEX 1****ANNEX 1****Eighteenth Meeting of Computing Representatives  
ECMWF, Shinfield Park, Reading, U.K., 8-9 June 2006****Participants**

Austria	Cornelia Hammerschmid
Belgium	Liliane Frappez
Croatia	Oleg Percinic
Czech Republic	Karel Ostatnicky
Denmark	Thomas Lorenzen
Finland	Kari Niemelä
France	Marion Pithon
Germany	Elisabeth Krenzien
Hungary	Istvan Ihasz
Iceland	Vigfus Gislason
Ireland	Paul Halton
Italy	Carmelo Gambuzza
Netherlands	Hans de Vries
Norway	Rebecca Rudsar
Romania	Roland Cotariu
Serbia	Vladimir Dimitrijevic
Spain	Eduardo Monreal
Sweden	Rafael Urrutia
United Kingdom	Roderick Sharp
CTBO	Gerhard Wotawa
EUMETSAT	Marc Jenner
ECMWF:	Tony Bakker
	Sylvia Baylis
	Petra Berendsen
	Jens Daabeck
	Paul Dando
	Richard Fisker
	Anne Fouilloux
	Enrico Fucile
	Helene Garcon
	Laurent Gougeon
	Dominique Lucas
	Carsten Maaß
	Umberto Modigliani
	Pam Prior
	Sylvia Rozemeijer
	Deborah Salmond
	Stephan Siemen
	Neil Storer
	Daniel Varela
	Isabella Weger

**ANNEX 2**

**ANNEX 2**

**Programme**

**Thursday, 8 June 2006**

- 10.00 Welcome
  - ECMWF's computer status and plans . . . . . I. Weger
  - Member States' and Co-operating States' presentations
- 12.00 Lunch
- 13.00 Visit of Computer Hall (optional)
- 13.30 Member States and Co-operating States presentations (continued)
  - HPCF: Phase 4 update and migration plans . . . . . N. Storer
  - Early experiences on IBM Phase 4 system . . . . . D. Salmond
  - Overview of ECMWF electrical and mechanical infrastructure . . . . . S. Baylis
- 16.00 Coffee
- 16.30 RMDCN status and migration plans . . . . . T. Bakker
  - ECaccess status and plans . . . . . L. Gougeon
  - New MS job submission via SMS . . . . . D. Lucas
  - User's registration update . . . . . P. Dando
- 18.00 Cocktails

**Friday, 9 June 2006**

- 09.00 Member States and Co-operating States presentations (continued)
- 10.30 Coffee
- 11:00 Survey of external users: status of follow-up actions . . . . . U. Modigliani/D. Varela
  - Graphics update . . . . . J. Daabeck
  - Technical overview of MAGICs++ . . . . . S. Siemen
  - Introduction to the new GRIB API . . . . . E. Fucile
  - Discussion
- 13.00 End of meeting