ECMWF's NEW DISSEMINATION SYSTEM

Dragan Jokic

European Centre for Medium-Range Weather Forecasts

Summary: In December 1998, ECMWF established a project to develop a new UNIX based Dissemination System and to replace the existing VMS system. As a result, the Dissemination System has been completely redesigned, rewritten and now runs on the Hewlett-Packard cluster. A Web based interface, accessible by Member States has been provided for requirements maintenance and for monitoring and controlling dissemination transmissions. In the process of project development it was decided to develop ECMWF's own queuing system and file distributor for the automatic transmission of files.

The project has been successfully realised and implemented into ECMWF operations in June 1999 when dissemination transmission from the new system to the Member States started. Over a period of four months, all Member States and external users have successfully been migrated to the new system.

1. INTRODUCTION

In December 1998. ECMWF established a project to develop a new UNIX based dissemination system. An expert group was formed and target dates set as June 1999 to develop the new system and the end of September 1999 to migrate all Member States to the new system. As a result, dissemination system has been re-designed and re-written and a new transmission system developed.

The following were the main reasons which led to the decision to develop a new dissemination system:

- Resolution of our products has been constantly increasing over the years resulting in bigger size fields. Still, part of our system was bound by the limits of the VMS logical record structure files, so we had to use WMO BLOCK format at the time to enable large fields in dissemination;
- Number of available forecast time steps has been increasing, resulting in a larger number of files;
- Number of 'data streams' available in ECMWF's dissemination has also been increasing (EPS, Global Wave, European Wave) resulting in the increase in the number of files;

- All the above put heavy I/O pressure on our VMS systems, and the performance of dissemination started to be affected; Also, system support for our VMS started to be more and more limited;

- Monitoring tools were hard to develop and maintain on VMS;

- TCP/IP has become a more preferred way of transferring data then DECNET;

- Dissemination system on VMS could not properly follow other components of our operational system, such as MARS, which has been moved to UNIX and is used as the language handler;

- The main reason for realising the project in 1999 was the fact that our VAX cluster was not Y2K compliant;

2. PRODUCTS AVAILABLE THROUGH ECMWF's DISSEMINATION

The following 'data streams' are available in ECMWF's dissemination:

- 12 UTC based deterministic products,
- Ensemble Prediction System (EPS) products,
- Global Wave Model products,
- European Wave model products,
- 00 UTC based deterministic products (optional project) and
- Wave EPS products.

Data representations available in dissemination are:

- Spherical harmonics,
- Latitude/Longitude grid,
- Gaussian grid (regular and quasi-regular),
- Polar stereographic and
- Rotated grids

Data formats supported in dissemination are:

- GRIB,
- BUFR and
- GRID.

3. BASIC CONCEPT FOR THE NEW DISSEMINATION SYSTEM

Our aim was to build dissemination system which would be open to users as much as possible. In due course the basic concept was drown for the system where:

1) Products are requested/requirements maintained by users themselves;

2) Products are transmitted automatically according to dissemination schedule (or);

3) Products are generated at ECMWF and picked up later by users at their convenience;

4) Transmission priority is supported (initially on dissemination stream level);

5) System is independent of file and data format;

6) UNIX based;

7) To work on ECMWF's HP cluster (also portable);

8) Easy to install;

9) Reliable (as the previous one);

10) Easy to make changes;

11) Modular;

12) Easy to monitor;

4. NEW DISSEMINATION SYSTEM

New ECMWF dissemination system consists of the following components:

a) Dissemination requirements (files, directories);

b) Member States' pre-defined bit-maps (files);

c) Dissemination language;

d) Dissemination requirements checking tools (used by Member States to validate and install requirements);

e) Dissemination requirements combining programs (preparing dissemination requirements for input to operational products generation);

f) Predefined Member States' bit-maps combining programs;

g) Dissemination (products) files;

i) Dissemination process;

j) Transmission System;

k) Repeat process;

l) Cleanup procedures;

Dissemination system with all its components will work on UNIX system (Hewlett-Packard B.10.20).

5. DISSEMINATION REQUIREMENTS

Dissemination requirements are defined by the following parameters:

1) Country name (or the project ECMWF supports);
2) Destination name;
3) Dissemination stream name;

Dissemination requirements are grouped into dissemination streams. Each dissemination stream can contain requirements from any of the available data streams (DA-12 UTC based, EPS, WV-Global, WV-Mediterranean, DA-00 UTC based). Each destination can have more then one dissemination stream and also, each Country can have more then one destination.

Table 1 shows the total number of dissemination requirements (fields) per data stream in ECMWF's operations as it was on 17. November 1999. Dissemination requirements are fully maintained by users themselves through the web provided interface.

DATA STREAMS	TOTAL NUMBER OF PRODUCTS
12 UTC BASED T319 DETERMINISTIC AN/FC	136 066
EPS	72 923
WAVE GLOBAL	32 652
WAVE EUROPEAN	2 700
00 UTC BASED	50 090
TOTAL	297 431

Table 1 Total number of dissemination products

Current, operational and previous dissemination requirements are maintained by the new system. Also, the history logs about requirements changes are kept.

Changes in the number of dissemination requirements over the years are shown in Figure 1.
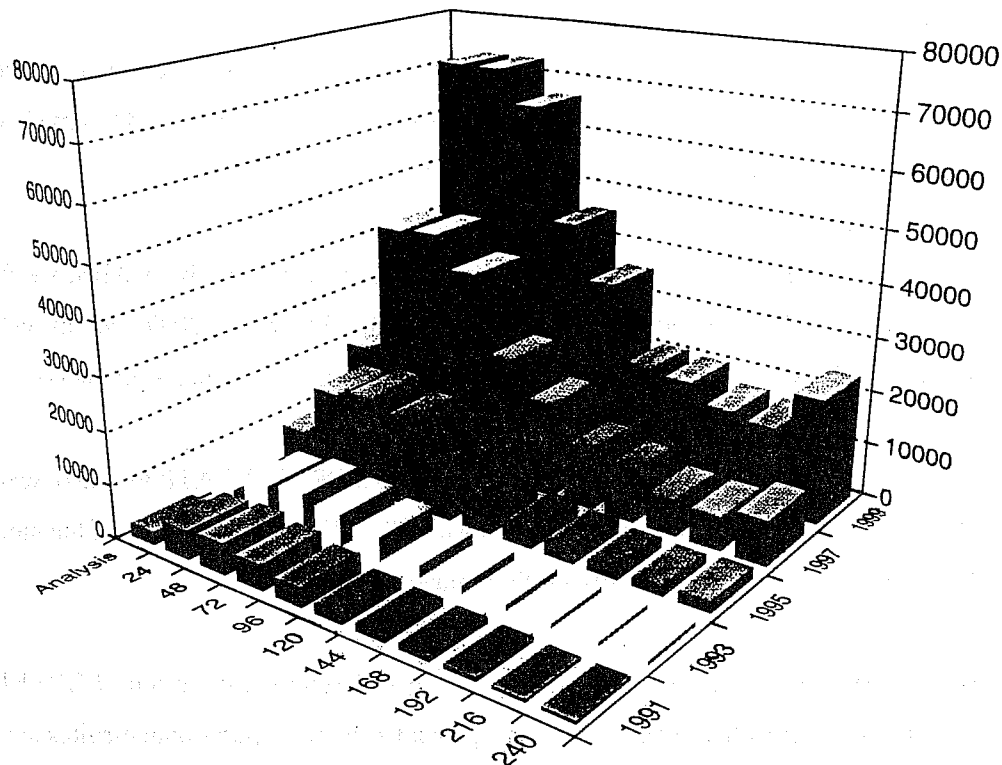
Number of products for dissemination

Figure 1. Number of products for dissemination

What is also new in the new system is that requirements enter the operational suites several times a day (just before every analysis cycle), so users can change their requirements at any time during the day and the change will go straight into production.

6. MEMBER STATES' PRE-DEFINED BIT MAPS

Member States can request dissemination products on pre-defined bit-maps. These bit-maps are used as a basis for dissemination products. Each pre-defined bit-map is defined using command-like definitions and is kept in a separate file. Bit-map files are maintained by users themselves through a web interface.

Bit-maps are:

- checked against request;

- processed (from the command definitions);

- converted into operational format;

Bit-map file name uniqueness is supported by the directory structure of the new system (users are free to choose any bit-map (file) name).

7. DISSEMINATION LANGUAGE

Dissemination products are requested in the form of the MARS language. New dissemination system has overcame differences between old dissemination language and MARS language and has made dissemination language parameters identical to those of MARS.

Parameters, which were supported by the previous dissemination system but not by MARS, are fully supported by the new system. These include parameters like FRAME, BITMAP, OPTIONS, as well as parameter values like TYPE=MT (metgram data) and TYPE=WP (weather parameter products).

Several parameters have been discarded, like parameters FILE and REPRESENTATION. Also, several new parameters have been added to the language like parameter PRIORITY where values of the parameter PRIORITY indicate the transmission priority of dissemination stream.

Experiment version number, EXPVER, is supported by the new dissemination system. EXPVER is used not only for Optional Projects but to enable test data generated by pre-operational e-suites available to Member States prior to operational implementation. Just, recently before the introduction of the new 60 level model into operations we offered test data to dissemination users. In the same way, Y2K test data are now available to Member States.

MARS language parameters, as they are used in dissemination are described in detail in the new Meteorological Bulletin 3.1. ECMWF has implemented all language changes to Member States' requirements prior to operational implementation of the new system. ı

8. DISSEMINATION FILES

Dissemination files in the form in which they are disseminated to Member States are generated by the Products Generation on ECMWF's mainframe machine and stored in the Products Database. From the Products Database files are ftp-ed to the telecommunication machine (each operational suite task ftps a set of files for one forecast/analysis time step).

Only one file format of dissemination files, pure binary, is supported by the new system.

Dissemination files are kept on-line on the telecommunication host (HP cluster) for 24 hours. Directories are cleaned prior to each new operational cycle.

Total number of files generated in one operational cycle is at the moment about 3 500. Total volume of

data in dissemination at the moment is about 3Gb. Changes in the volume of products as well as changes in total number of products are shown in Figure 2. New system has been designed to be flexible and to support ever growing demand for new products.
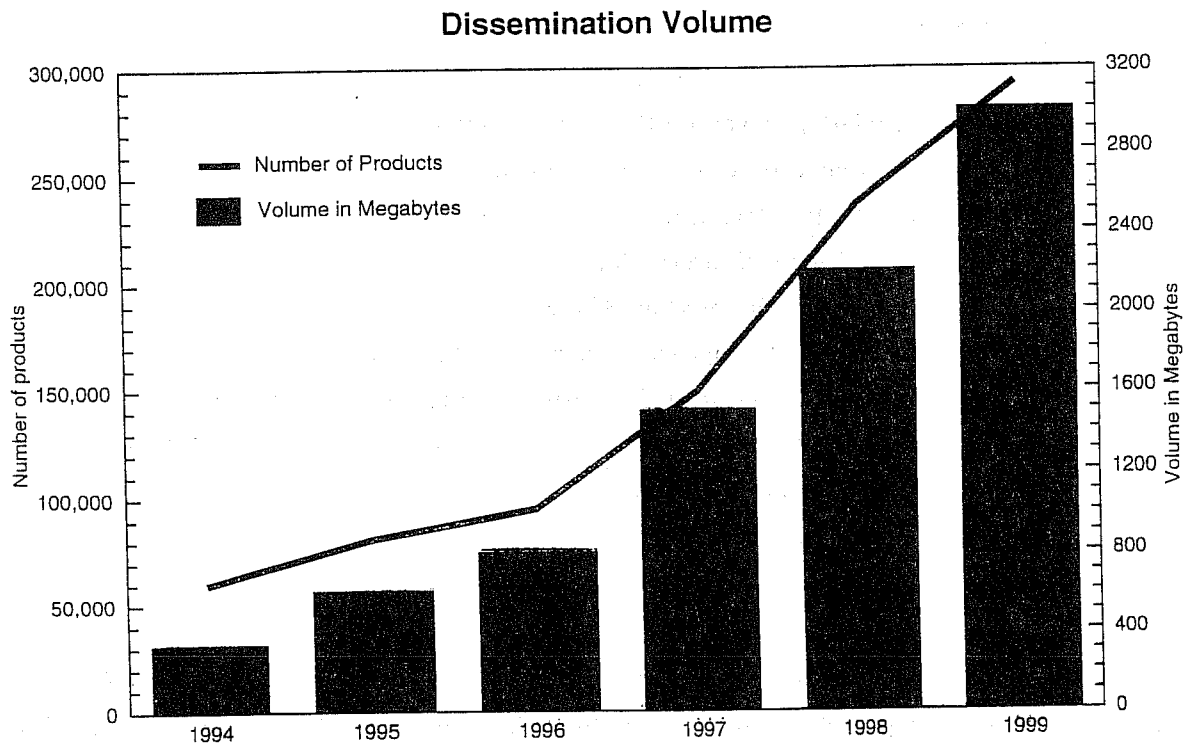
Dissemination Volume

Figure 2. Dissemination Volume

8.1 New dissemination file naming convention

Together with developing the new system, the dissemination file naming convention was changed. The following are the main reasons for introducing the new naming convention

- limits of the old file naming convention were almost reached due to an increased number of forecast time steps;

- it was not possible to support 'experiment version number' by the old convention as well as the real date and time for analysis products;

-old naming convention couldn't properly describe new language definition of analysis products;

-dissemination file naming convention had to be changed in 1999 to make dissemination file naming convention consistent for 12UTC based and 00 UTC based products;

Below is the new ECMWF's dissemination file naming convention. It is often difficult to define a naming convention which would last for some years. This time we decided to use more general categories like date and time in the file name rather then some artificial numbers:

c_1c_2SDDMMHHIIddmmhhiiv where:

c_1c_2	is	two letter dissemination stream name;
S	is	the one letter data stream indicator;
		S ="D" for T319 atmospheric model products;
		S ="E" for EPS products;
		S ="P" for Global Wave Model products;
		S ="M" for European Wave Model products;
DDMMHHII	is	day, month, hour and minute at which products within the file are valid;
ddmmhhii	is	day, month, hour and minute of the analyses data on which forecast(s) (analyses) is/are based;
v	is	version number.

This naming convention will give us flexibility to increase the number of time steps, add new streams and we believe would make file names easier to 'read' by humans and also easier to maintain (date dependency).

9. DISSEMINATION PROCESS

The dissemination process takes prepared dissemination files and passes them as entries to the Transmission System for transfer. Files are taken and queued into the Transmission System according to dissemination schedule. Each time step (analyses and forecast), within each data stream (DA-12 UTC based, EPS, WV-Global, WV-Mediterranean, DA-00 UTC based) has its own schedule. Also, there is a separate dissemination process for each data stream. Figure 3 shows the outlook of the operational 'dissemination suite' run by the SMS.

Once passed to the transmission queuing system, dissemination files are moved to the Transmission System directory structure which ensures efficient usage of disk space. Also, monitoring of dissemination processes has been made much more sophisticated than before.
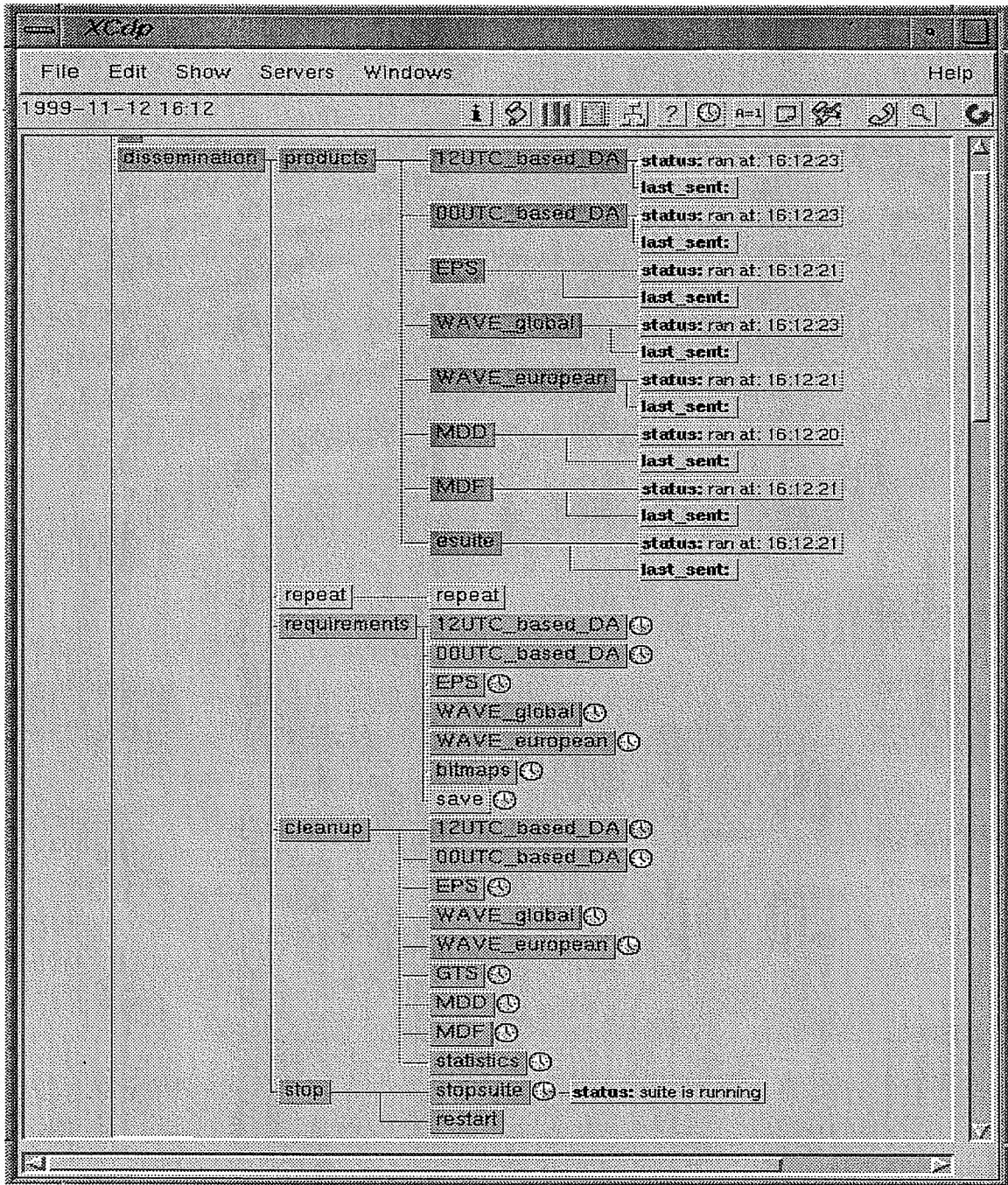
Figure 3 Dissemination Suite

10. TRANSMISSION SYSTEM

Old Transmission System was not in the shape in which it could had been moved to UNIX. We needed a completely new system. Several options were considered and at the end it was decided to develop our own.

The requirements for the new transmission system were to:

- handle more than one entry scheduled at the time;

- support long transmissions (new entries waiting in queues);

- easy to monitor;

- easy to spot problems;

- enable delayed transmission;

- easy to change status of entries;

The basic concept around which the new Transmission System was developed:

- UNIX based;

- easy to install;

- easy to make changes;

- modular;

- easy to monitor;

- reliable;

The components of the new Transmission system are:

- transmission interface;

- QFT daemon (queuing system, transmission method, configuration file)

- log files;

- monitoring tool;

11. REPEAT PROCESS

Repeat tools are used by ECMWF's Member States to repeat dissemination files and to trigger transmission of so called delayed products (products generated at ECMWF, but not automatically transmitted to user). The option of repeating/triggering transmission of dissemination products is made part of the new web based GUI for dissemination requirements maintenance.

Member States with remote job entry facilities are able to request repeat transmission of dissemination products by submitting a 'repeat' job at ECMWF.
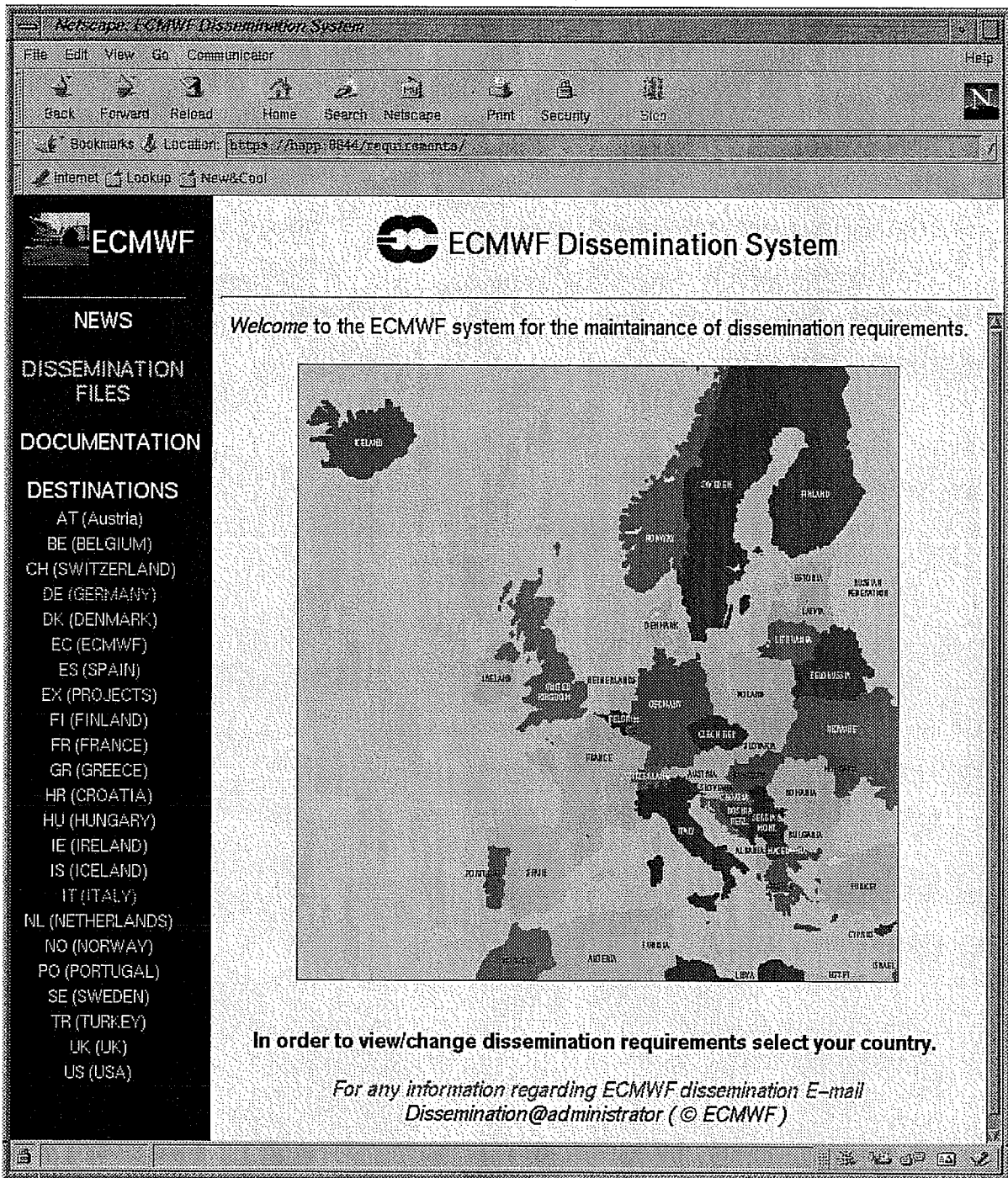
Figure 4Dissemination web page

12. DISSEMINATION WEB PAGE

Maintenance of dissemination requirements in the new system has been moved to the web. Figure 4 shows ECMWF's dissemination web page (https://happ:8844/requirements/). In the new system, dissemination requirements continue to be fully checked against very strict definitions of what is available

through the dissemination system before entering the operational suites. Wrong requirements are reject-
ed and full setup not made operational until all requirements are correct.

A new (http based) GUI has been developed to enable Member State to:

- access own dissemination requirements site;
- choose dissemination stream;
- view requirements;
- make a change;
- inspect impact of the change (number, volume of products);
- make change operational;
- go back if wanted;
- view dissemination queues;
- change priority of transferring of dissemination files;
- repeat files;

13. MIGRATION

Migration is often a long process which requires a lots of coordination and support for users. ECMWF
started moving countries to the new system in June 1999. By the end of September, which was our target
date, all countries but 3 were moved to the new system.

At the end of October 1999, we successfully completed the project by moving all countries to the new
dissemination system.