In this issue

European Centre for Medium-Range Weather Forecasts

Shinfield Park, Reading, RG2 9AX, UK

Fax: (0118) 986 9450

Telephone: National 0118 949 9000

International . . . . +44 118 949 9000

Internet URL . . . . http://www.ecmwf.int

Cover

The implementation of its 4D-Var analysis system is a major milestone for ECMWF, see the article on page 2.

Editorial

On 25 November 1997 ECMWF brought into production its 4D-Var analysis scheme. This is another milestone in the implementation of variational analyses at ECMWF. Following the use of 1D-Var in satellite retrievals, and 3D-Var in global analyses, 4D-Var has now been made operational. The initial implementation uses data in a six-hour time window, centred around the analysis time. *See page 2 for full details.*

ECMWF Technical Memorandum No. 238 on the 'Revision of the J_b term in 3D- /4D-Var' is summarised on page 5.

On page 6 is the annual review of ECMWF's computer system status and plans. A major part of these plans is the continuing development of the new data handling system, page 8 has an update on progress in this area.

Fortran is widely used in the meteorological field as the programming language of choice. With a move towards parallel systems, a version called High Performance Fortran (HPF) is emerging to take advantage of these new architectures. On page 10 is an article about HPF, including ECMWF's involvement in an European Community funded project aimed at exploring extensions to HPF.

Changes to the Operational Forecasting System

A number of changes to the physical parametrization scheme were introduced on 16 December 1997 (model cycle 18r3):

1. a modification in the treatment of the water vapour absorption in the long-wave part of the radiation scheme;
2. a new method of moisture convergence closure;
3. a new treatment in the ice fall-out in the cloud scheme.

Planned changes

- An improved formulation of the two-time-level semi-Langrangian scheme.
- A new, more accurate resolved model orography.
- Coupling of the atmospheric and ocean-wave model.
- Revised initial perturbations for the EPS, to increase spread, particularly in the early part of the forecast range.

François Lalaurette

The operational implementation of 4D-Var

On 25 November 1997, the ECMWF operational system was switched to use a 4D-Var assimilation algorithm. 4D-Var is a four-dimensional variational data assimilation technique that performs a statistical interpolation between a distribution of meteorological observations in time and space (figure 1) and an *a priori* estimate of the model state (called the background). A special property of 4D-Var is that it takes into account the dynamics and the physics of the forecast model in order to ensure that the observations are used in a meteorologically consistent way. As will be illustrated below, this implies that asynoptic data can be used at their appropriate observation times, the structure functions of the assimilation are flow-dependent, and the physical imbalances in the model are reduced.

This is the first ever operational application of the 4D-Var technique successfully applied to a high-resolution assimilation and forecast system. After the initial proposal in 1985 (Lewis and Derber 1985, Courtier and Talagrand 1987, Talagrand and Courtier 1987), this was made possible by more than ten years of scientific and technical developments in and around ECMWF's Integrated Forecast System (IFS) (Pailleux, 1997) as well as the availability of a powerful new computer system organized around a Fujitsu VPP700 with 116 processors. The pre-operational and early operational experience with 4D-Var has shown a very clear improvement in the performance of the forecasting system. This is an impressive yet young assimilation system that offers an exceptional scope for future improvements.

From 3D-Var to 4D-Var

The variational analysis system was designed to allow a smooth transition from 3D-Var to 4D-Var; technically speaking, the software changes only slightly with 4D-Var (although it is computationally more expensive). Both systems work with a T213L31 forecast system, and compute analysis increment fields at a lower resolution of T63L31, as well as analysis and forecast error estimates at T42L31. The observations used are almost the same: conventional synoptic stations (including Australian pseudo-observations), buoys, radiosondes, aircraft reports, cloud-motion winds, cloud-cleared TOVS radiances and ambiguous scatterometer winds. In both systems, the analysis is produced every 6 hours, e.g. for the 12 UTC analysis (the initial condition for the medium-range forecast and the Ensemble Prediction System (EPS)), the observations between 9 UTC and 15 UTC are used. The formulation of the background term is identical in 3D- and in 4D-Var, as is the penalization of the tendency of the gravity wave modes. The variational analysis problem is solved at a low resolution using the incremental technique in order to reduce the computer costs: first, the observations are compared to the high-resolution T213 background model state, but the minimization of the cost function of the analysis is then carried out at a lower T63 resolution using the observation operators linearized in the vicinity of the high-resolution background. Finally, the T63 increments are converted back into T213 increments using normal mode initialization and added to the T213 background. This is roughly equivalent to replacing the minimization of a T213 cost-function by the minimization of its T63 quadratic approximation in the vicinity of the background. As explained below, the technique is slightly modified in 4D-Var.

There are, however, some important new features. Due to the nature of the 4D-Var algorithm, the model evolution during the 6-hour interval, or assimilation 'window', is explicitly taken into account. While in 3D-Var the observations were assumed to have been produced at the central time of the interval (except for a tendency correction of the surface pressure observations), in 4D-Var the observations are compared to a model trajectory at 1-hour intervals.

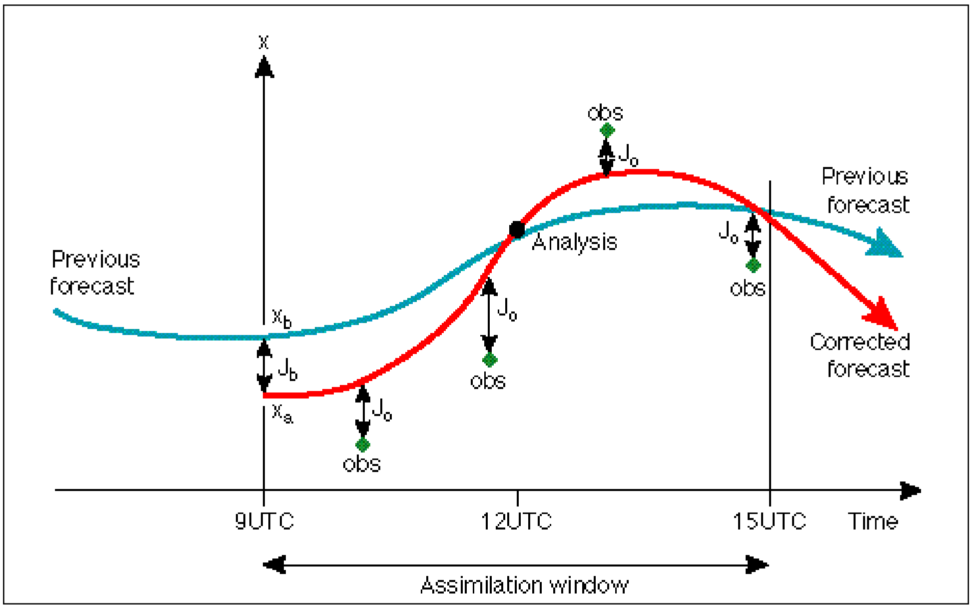


Figure 1: *Simplified view of the 4D-Var assimilation technique for a single parameter x. Over a given time window (six hours here), the observations are compared at their appropriate time with a short-range forecast issued from the previous analysis. The model state at the initial time of the window is then adjusted to achieve a statistically good compromise x_a between the fit J_b to the previous forecast x_b, and the fit J_o to the observations.*

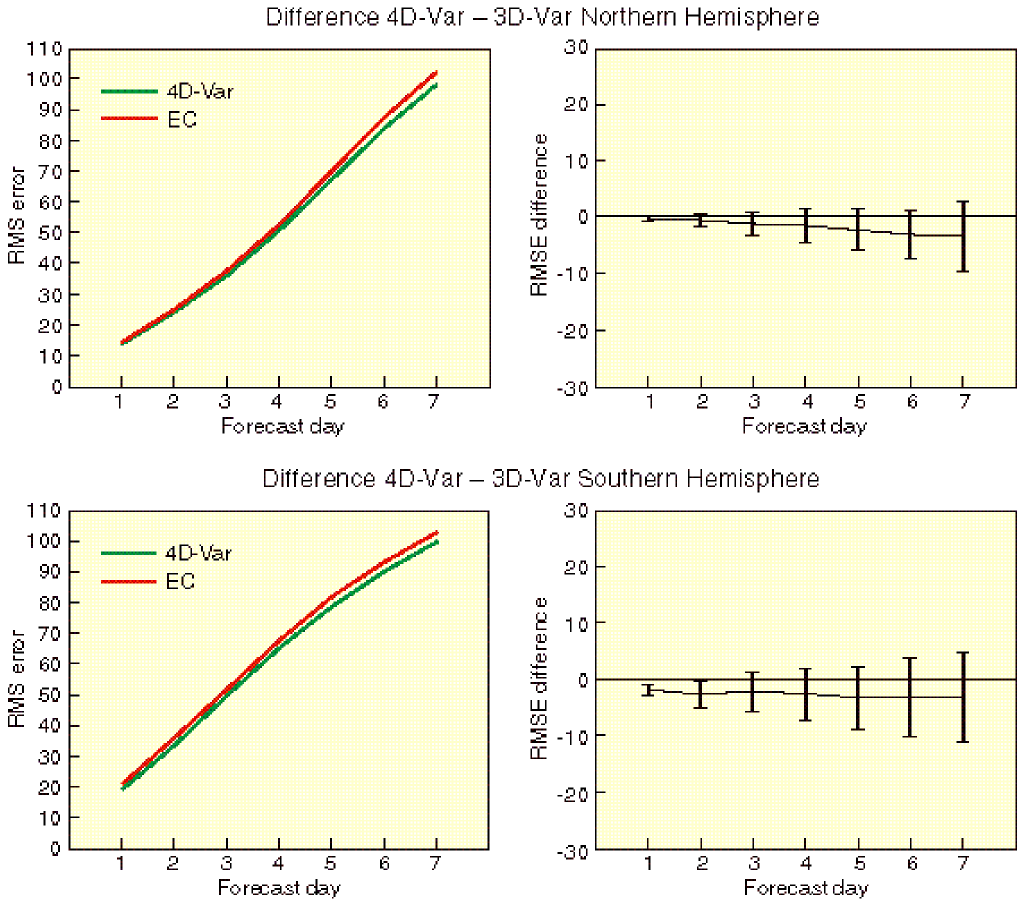


Figure 2: *Root-mean square of the differences between analyses and forecasts of the 500hPa geopotential height from 4D-Var (green lines) and from 3D-Var (red lines), averaged over 40 cases of the preoperational suite in the Northern (top) and the Southern (bottom) Hemispheres. The left panels show the scores, the right panels show the differences between the scores with vertical bars depicting the 95% confidence level of the Student t-test.*

This means that the asynoptic observations are compared more precisely with the model forecast, thus improving the timing of meteorological events. There is some potential for using more data than in 3D-Var by allowing the inclusion of more than one report from the same observing station; however, experimentation has shown that temporal observation error correlation must be explicitly taken into account to use the added information correctly.

Because the observations are compared with a sequence of model states over the 6-hour window, the concept of analysis time is blurred in 4D-Var. At the beginning of the minimization, the observations are compared with the previous forecast; at the end, they are compared with a corrected forecast that starts from a modified model state at the beginning at the assimilation window, as shown in figure 1. Since the corrected forecast is the result of an adjustment to observations distributed over the whole window, the forecast errors do not grow with time as they would with a forecast issued from a static analysis. It can be shown that the forecast errors are at a minimum near the middle of the window, at 12 UTC in our example. For this reason, the 3-hour forecast state is labelled as the 'official' atmospheric analysis at 12 UTC. Together with the analyzed surface fields, this is the initial state for the medium-range forecast. This procedure has the added advantage that the change of the assimilation technique is transparent for the dissemination of products to the Member States. When the 4D-Var window is extended to 12 or even 24 hours, the choice of cut-off technique will need to be reassessed.

In 4D-Var, the model state at the beginning of the assimilation window is optimized in order to minimize the distance to the background (defined at the same time) and to the observations; whenever it is modified, the new fit to the observations is calculated by running the model forecast until the time of each observation, interpolating the forecast state to the observation location and variable, weighting the (obs - model) mismatch by the assumed observation errors, and computing the adjoint of this process (i.e. the transpose of its derivatives, with a suitable inner product) in order to obtain the sensitivity of the mismatch with respect to the model state at the beginning of the window. This introduction of the forecast model and its adjoint into the assimilation algorithm introduces flow-dependent structure functions for all observations, except the ones made right at the beginning of the assimilation window, which are used exactly as they would be in 3D-Var. When the flow is unstable, the

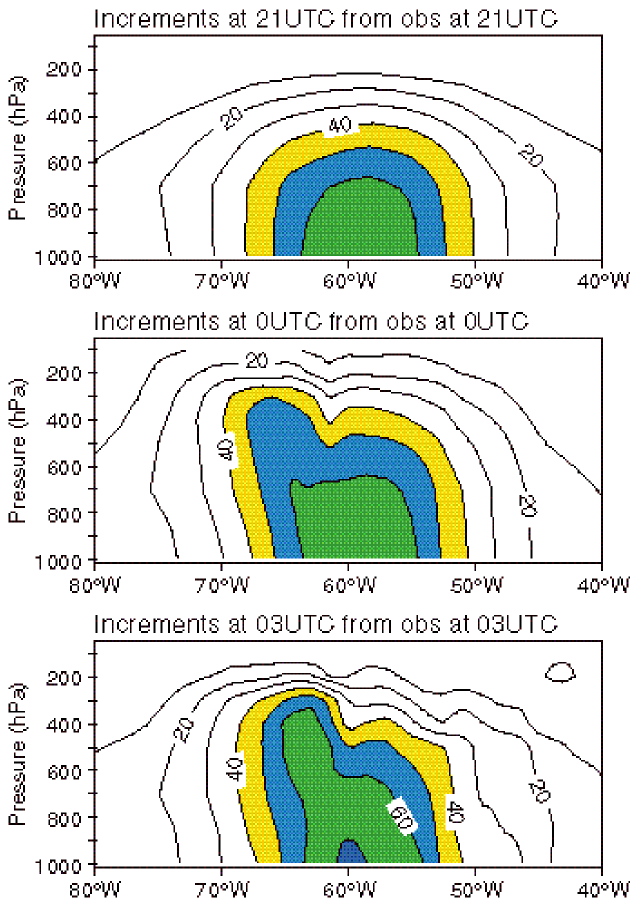


Figure 3: *Corrections to the previous forecast caused by a height observation at 850hPa, (40N,60W). Isolines show the resulting increment, in geopotential units, for an observation at three different times, as explained in the panel captions.*

observations are given more weight; when a wave develops, the analysis increments caused by the observations tend to follow the shape of the wave; when the advection of a meteorological feature (e.g. the humidity field) does not agree with the sequences of observations made in the area, the wind is suitably corrected, even if it is not directly observed. This flow dependence is stronger for observations made near the end of the assimilation window; over a 6-hour window as used now, the design of the background term J_b is still quite important in determining the structure functions in most parts of the analysis.

The introduction of the model into the assimilation procedure brings flow-dependence of the structure functions and use of observations at the appropriate time, but it comes at a price. Firstly, since typically 80 evaluations of the variational cost-function and its gradient have to be done to minimize the 4D-Var cost function, the computer cost of about 160 6-hour T63 model forecasts has to be paid on top of the cost of the other software components which already existed in 3D-Var. In practice, the model itself vectorizes and distributes rather well on the Fujitsu computer, and most of the cost actually comes from the management of observations and from the fact that 4D-Var uses the tangent-linear and adjoint versions of the model (adiabatic core, plus simplified physics for part of the minimization). Secondly, the realism of the model used inside 4D-Var is important because it determines how perturbations to the initial state of the model are translated into perturbations in terms of observed parameters. It is limited by the low resolution (T63L31) and by the simplification of the physics; consequently, problems related to those model limitations are likely to affect the use of observations near the end of the assimilation window and near the ground, as well as wherever there are non-linearities, threshold processes or interactions with the water cycle. This could limit the usefulness of 4D-Var. For the time being, the 4D-Var incremental procedure is carried out in two low-resolution minimization steps (instead of one in 3D-Var):

- a high-resolution T213 comparison with observations to linearize the problem in the vicinity of the background forecast,
- a long minimization at T63 with minimal linearized physics,
- an update of the high-resolution T213 model with this first set of increments in order to correct the T63 linearization around a state closer to the analysis,
- another short minimization at T63 with a much more complete version of the linearized physics,
- finally, this second set of increments is added to the current high-resolution initial estimate of the initial conditions to produce the final T213 analysis.

The observation quality control procedure has been adapted from 3D- to 4D-Var in order to take into account the use of observation at the appropriate time. This has meant only slight changes to the assumed observation errors. The screening and variational quality control procedures have not been changed, although some improvements are planned.

Impact on operational performance

Extensive pre-operational validation of 4D-Var was carried out before deciding on the operational implementation. The impact of going from 3D-Var to 4D-Var (everything else being kept identical) has been assessed with several configurations in the selection of observations, the use of physics, and the number of high-resolution updates in the incremental formulation. It has led to the choice of the system described above. The impact on some forecast scores is presented in figure 2 (cf. Rabier et al, 1998, for more details). The fit of the background fields to the observations is improved as well, and the short-range spin-up of precipitation is reduced.

The impact might look small, but it is an essential step toward further improvements of the system. Since the implementation, the operational scores have compared very well with other forecast centres, although this is certainly a combination of effects between 4D-Var and previous changes like the revision to the background term J_b (cf. article by F. Bouttier in this newsletter), the extension of the use of TOVS radiances, and the slightly later changes in the physics package, among others.

A principal feature of 4D-Var is the flow-dependence of the structure functions, already documented by *Thépaut* *et al* (1993). Although a 6-hour time window is rather short for it to take its full development, it is already benefitting the assimilation in dynamically active areas, as shown in figure 3 (taken from Rabier et al, 1997).

Future evolution

4D-Var is viewed as an essential tool for future improvements of the assimilation system. First of all, it is a very new assimilation system which needs to be retuned: because the observations are used differently and the forecasts are improved with respect to 3D-Var, the assumed observation and background errors and the rejection thresholds need to be readjusted using objective methods. More data than in 3D-Var can be used now because several reports from the same station can be used, assuming that serial correlation of observation error is properly taken into account. The dynamical initialization problem is different from the one in 3D-Var because some model properties are included in the 4D-Var increments. More efficient incremental strategies can be devised to accelerate the convergence of the 4D-Var variational assimilation problem. All these could be implemented soon at almost no computational cost, but with potentially significant impacts on the forecast quality.

In 4D-Var, the model is naturally integrated within the assimilation. As the forecast model is improved, the 4D-Var configuration will evolve too. It is planned to improve the vertical model resolution in the stratosphere and to include ozone as a prognostic variable. The same changes will be brought into 4D-Var, paving the way for a coupled ozone/wind assimilation in the stratosphere. The realism of the low-resolution model used in the minimization will be improved by the inclusion of better physics and a better horizontal resolution, which are expected to improve the use of data over land as well as observations related to clouds and precipitation. In return this should improve the assimilation and forecast of actual weather parameters.

In the longer term, the full promise of 4D-Var will only be realized if we are able to handle sophisticated flow-dependent structure functions. In other words, the future system should be able to decide automatically to which observations the quality of the ensuing forecast will be sensitive, and to use them in a meteorologically optimal way. The natural way of doing this will be to extend the time window of 4D-Var to 12 and perhaps even 24 hours. In order for the flow dependence to be managed better in the assimilation system, an extension to 4D-Var called Simplified Kalman Filter is being developed, which will introduce flow-dependence into the J_b term. The improved estimates of analysis error it provides will also improve the generation of ensemble members in the EPS.

References

Bouttier, **F.**, **J. Derber** and **M. Fisher**, 1997: the 1997 revision of the Jb term in 3D/4D-Var. ECMWF Research Dept. Technical Memorandum no. 238.

Lewis, **J.** and **J. Derber**, 1985: The use of adjoint equations to solve a variational adjustment problem with advective constraints. *Tellus*, **37A,** 309-322

Courtier, P. and **O. Talagrand**, 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation. II: Numerical results. *Quart. J. Roy. Meteor. Soc.*, **113**, 1329-1347.

Pailleux, J., 1997: 1987-1997, Ten years of research and operational activities with the integrated forecasting system (IFS). ECMWF Newsletter no. 75, Spring 1997.

Rabier, F., **J.-F. Mahfouf**, **M Fisher**, **H. Järvinen**, **A. Simmons**, **E. Andersson**, **F. Bouttier**, **P. Courtier**, **M. Hamrud**, **J. Haseler**, **A. Hollingsworth**, **L. Isaksen**, **E. Klinker**, **S. Saarinen**, **C. Temperton**, **J.-N. Thépaut**, **P. Undén** and **D. Vasiljevic**, 1997: The ECMWF operational implementation of four-dimensional variational assimilation. ECMWF Research Dept. Technical Memorandum no. 240.

Talagrand, O. and **P. Courtier**, 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory. *Quart. J. Roy. Meteor. Soc.*, **113**, 1311-1328.

Thépaut, J.-N., **R. Hoffman** and **P. Courtier**, 1993: Interactions of dynamics and observations in a four-dimensional variational assimilation. *Mon. Wea. Rev.*, **121**, 3393-3414.

F. Bouttier and F. Rabier

Summary of ECMWF Technical Memorandum 238

The 1997 revision of the Jb term in 3D- / 4D-Var

F. Bouttier, J. Derber and M. Fisher

The spatial structures of the corrections made by the analysis on the meteorological fields, called structure functions, are determined by the formulation of the background term J_b in the cost function of the 3D-Var analysis. This is also true to a large extent in 4D-Var. An optimal J_b design would ideally reflect the covariances of the short range forecast errors in the assimilation. In practice, the covariances must be modelized, based on the average variances, autocorrelations and balance properties of background error fields. These error fields are themselves approximated by forecast differences.

A revision to the operational background term was implemented in May 1997. The main change is in the way the multivariate correlations are implemented. Some related improvements to the management of the back ground error variances, or 'cycling', were implemented at the same time. The revision led to a substantial improvement of the assimilation and forecasts of tropical winds, and to a smaller improvement of the extratropical forecast scores. Consequently, the quality of the wave forecasts was improved, and a better use of TOVS radiances could be implemented with a revised bias correction a couple of months later. The revision to the background term was also shown to be as beneficial for 4D-Var as it is for 3D-Var, and its mathematical properties are facilitating the ongoing development of a Simplified Kalman Filter (SKF).

The revision of the background term did not affect directly the analysis of humidity, and the structure functions for all variables are still non-separable (i.e. the vertical structures depend on the horizontal scale). However, the structure functions are now a function of latitude, so that they are very different in the mid-latitudes than in the tropical regions. The balance constraint is incorporated into the preconditioner of the variational analysis which is now carried out in terms of vorticity, specific humidity, unbalanced divergence, unbalanced temperature and unbalanced surface pressure. The definition of the balance constraint is based on linear regressions between the forecast errors of the different variables. Physically it is very similar to an assumption of geostrophic mass/ wind balance for 90% of the errors in the extratropics. However, the balance is weaker near the ground and near the tropopause, where there is a significant correlation between vorticity and divergence errors. Near the Equator, the geostrophic character of the balance vanishes and more importance is given to the connection between divergence, temperature and surface pressure.

The autocovariances are specified in terms of the variables implied by the balance operators; they are assumed to be non-separable, homogeneous and isotropic. Physically it means that the autocovariances of vorticity and specific humidity are globally homogeneous (except that the field of vorticity variances is made to reflect the geographical density of the observing network). This is approximately true for the divergence as well. However, the autocovariances of temperature and surface pressure depend on latitude, so that the background term revision has a large impact on the tropical structures.

The simple mathematical formulation of the new background term allows a detailed meteorological examination of the implied structure functions. This has been used to demonstrate the need for future improvement: the introduction of some geographical variability in vorticity covariances, a thorough revision of the management of humidity, a special treatment of the very large scales, and the extension to new variables like ozone, for which an experimental balance with vorticity has al ready been developed.

In the Simplified Kalman Filter the static structure functions described above will be combined with flow-dependent ones in order to improve further the use of observations in meteorologically sensitive areas.

ECMWF's computer status and plans

(This article is based on a talk given to the Member States Computing Representatives on 7 October 1997)

Overview

There have been the following major changes in1997:

- an enlargement of the Fujitsu system to 116 PEs
- the acquisition of a second SGI Origin 2000 server
- the new data handling system, based on four RS/6000 servers, has been brought into use
- the old data handling system is now largely read only.

The current configuration is shown in Figure 1.

Regarding Computer Division staff, the structure of the division has remained the same. However, Geerd Hoffmann left ECMWF on 30 June 1997, and I was appointed to the position as of 1 July. As a result of this, Richard Fisker is now acting head of the Security, Internal Networks, and Workstation Section.

High performance computing facility

This consists of three Fujitsu machines:

- the 116 processor VPP700 system as the major production machine (upgraded in September)
- a 9 processor VPP300 for seasonal forecasting and re-analysis (installed 8 March 1997)
- a 4 processor VPP300 for system testing (installed 6 February 1997).

The upgrade of the main system was done in two stages. A 70 PE system was installed on 22 August, followed by a period of parallel running when all the major functions were duplicated on to that system. After a week of such running, the 46 PE system was powered off on 10 September and immediately dismantled. The next day, the 46 PEs were installed into the 70 PE machine, making a 116 PE system. There was a 10-hour downtime slot to accomplish all of this, which was easily met. Our great thanks go to Fujitsu for carrying out such an upgrade without problem.

As far as the users were concerned they saw 2½ days without service, that time being taken to recreate all the file systems and to run the operational forecast on schedule each night. The user service resumed on 12 September, on a machine that has 2½ times the capacity of the original system.

The 116 PE system has one primary PE (plus a spare), 10 I/O PEs, 8 secondary PEs with 2 Gbytes of memory each, and 97 secondary PEs with 1 Gbyte of memory each. The secondary PEs with 2 Gbytes of memory are proving very useful to accommodate those user jobs which are difficult to change to use less than 1 Gbyte of memory.

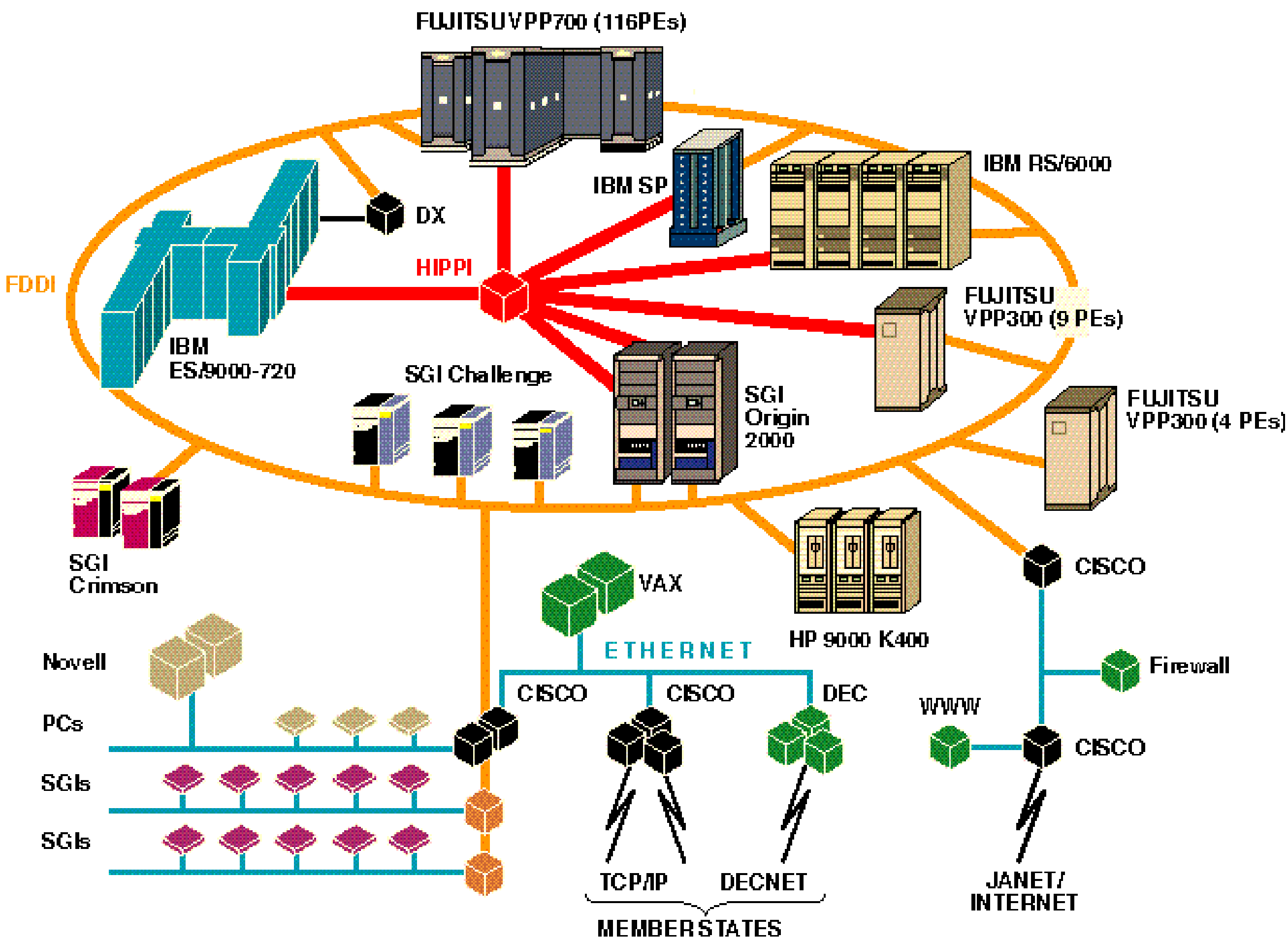


Figure 1: *The ECMWF current computer configuration*

Regarding usage of the original 46 PE system the utilisation had steadily increased from about 40% after installation to about 70% just before it was upgraded. This increase has been due to several factors:

- we have learnt how to operate the machine more efficiently
- Fujitsu has introduced software improvements
- users have improved the efficiency of their jobs.

Following the upgrade to the 116 PE system we now have to go through a similar learning and improvement cycle.

Comparing user usage statistics for 1996 (Cray) and 1996/7 (Fujitsu) there is a great increase by Data Assimilation (4D-Var work), and a relative decrease by both Diagnostics / Predictability and by Member States. The Member State usage level is not that unexpected given that it takes time and effort to convert work from a Cray to a Fujitsu environment. Also some Member State users have opted to move their work to a Cray elsewhere rather than convert to Fujitsu.

File servers

These are based on various SGI servers, including two Origin 2000 systems. The installation of the first of those Origin systems did not go as smoothly as we would have hoped. There were some initial hardware problems which resulted in several components having to be replaced. However, the second machine (installed in June 1997) had no such problems and has proved very stable from day one. Both are now working well.

The Member State machine is ecgate1, a two-processor SGI Challenge system. There were problems earlier this year when it was overloaded, 100% CPU utilisation was frequently reached, and the service deteriorated. To alleviate this we have

- moved MARS related processes off on to the Origin systems (behind the scenes so that users were not required to make any modifications)
- added ecgate2 (another SGI Challenge)
- added NQS load levelling so that jobs are always started on the least active machine.

Data Handling Systems (DHS)

As stated already, the old IBM ES/9000 based system is now largely in read only mode. There is no new user data going into it, only some of the operational data in parallel with the new DHS.

The volume of data on that system is roughly 65 TBbytes in 12 million files.

Phase 1 of the new DHS consisted of three IBM RS/6000 systems and a 25 Tbyte IBM 3494 automated tape library, running the HSM and ADSM archive management systems. Phase 1 was accepted in December 1996.

Subsequently, a trial MARS service was started on it in parallel with the ES/9000 MARS service, going into production on 17 June 1997. Currently, we are now migrating both MARS data and user files from the old system to the new. Hence in the background there is a substantial copying load running almost continuously. Unfortunately it is not possible just to move the cartridges from the old to the new system, they are in entirely different formats.

Phase 2 of the new DHS has also been installed. It provided a fourth RS/6000 processor, a second tape library, and more disks. It was accepted in May.

Currently, we have 8 Tbytes of user data in ecfs (the successor to ECFILE) and 16 Tbytes in MARS. Both these are growing rapidly, due to the copying of the old archive and new data from the 116 PE Fujitsu system.

Initially, the new DHS was not very stable, especially towards the end of 1996 and early 1997. However, since then the system has become a lot more stable and is now providing a reliable service.

Telecommunications

Only one Member State remains wholly reliant on DECnet, all others have converted to either TCP/IP exclusively or to TCP/IP plus DECnet.

There have been no line speed changes in the past 12 months, except the Internet link which is now 8 Mbps.

Plans

There will be a doubling of the Fujitsu capacity in early 1999. The contract with Fujitsu terminates in December 2000, thus a replacement strategy has to be decided by the end of 1998, with a possible ITT in 1999 and installation of the new system in 2000 (with parallel runs during 2000).

Phase 3 of the new DHS has been negotiated, and the contract signed. In the light of experience gained to date various changes to the original phase 3 configuration were made, it is now based on an SP2 system. In addition, we will modify and attach the existing StorageTek tape silos to the new system once they have finished service on the old DHS. The old DHS is scheduled to be closed down in 3Q98.

Later we will need to investigate how the new DHS can grow further to handle the ever increasing load. For example, we can introduce double density cartridges, giving 40 Gbytes per cartridge compared to the 10 and 20 Gbytes on today's cartridges.

The Member State server ecgate1 will be upgraded to a 4 processor SGI Challange with 640 Mbytes of memory.

Eventually we hope to terminate the VAX service. We have moved all the data acquisition and pre-processing off the VAXs. They currently only handle the incoming data, and we would like to move that to the HP servers. Then product dissemination will have to be moved. Finally, it depends on how soon all Member States can move off DECnet before we can close the VAX service completely.

The major item in wide-area network development is the RMDCN project, where it is proposed to merge the leased line network of ECMWF with the GTS network of WMO Region VI, and convert it all to a managed network. The next stages involve an ITT early in 1998 with first services in 1999.

A thorough review will be required of the internal network in 1998 to see how it can be developed to cope with the ever increasing load, especially bearing in mind the possible increases in capacity in the main compute server and data handling system.

Various activities have started regarding possible Year 2000 problems. A project group has been established, reports from other sites have been studied, and inventories are now being prepared. In parallel, various application groups have started to slowly check their code and then to test (as far as they can) in an isolated environment. If feasible in early 1999 we would want to create a testbed (say on the VPP300 9 processor system) which we could run for a couple of weeks in dedicated mode.

Walter Zwieflhofer

The new data handling service

Overview

In 1981, ECMWF had a data archive consisting of 1000 open-reel ½" tapes, with a total data volume of approximately 100 Gbytes. The management of it was proving difficult so we looked for an alternative, as we knew it would grow. A worldwide search of similar sites revealed that Los Alamos were using central file management system (CFS) which seemed suitable. We got it working at ECMWF in 1983 and, with the excellent co-operation from Los Alamos, it has proved a very effective system over the past 14 years.

By 1993, the CFS archive had grown to 10 Tbytes and we could see that, although it had been very good, it would not cope for ever with the likely expansion of the archive.

Los Alamos themselves were no longer supporting CFS and we had not got the resources to maintain and develop such a large system. The expected expansion in the years to come could not be accommodated on the 5¼" cartridges (~1 Gbyte per cartridge) in use in 1993. CFS would also require a fairly major revision to support another media type, hence we had to find another solution.

New data handling system

In 1994 we issued an ITT for a large data management system, the outcome of which was the selection and installation of an IBM based system (see ECMWF Newsletter 70, pp. 15-18). Today that system is in service. It consists of four servers (RS6000 based) running IBM's ADSM software as the tape manager. The CFS system used four StorageTek tape robots, the new system uses two IBM robots (16 x IBM 3590 drives each). They have 12 times the capacity per tape cartridge of the CFS drives and media. The larger of the two IBM robots, when full, will alone hold 50 Tbytes.

It took some time to get the initial configuration accepted because of deficiencies in IBM's software. However, there were major improvements during 1996 and since then we have progressed reasonably well. We now have replacement (known as ECFS) for the old ECFILE service, plus a MARS service. The objective is to provide a data handling service to support the Fujitsu system and its planned upgrades.

Earlier this year we installed the second phase, consisting of an additional server together with increased disk and robot capacity. The new MARS went into service, on schedule, in the middle of 1997. We have approximately 40 Tbytes of data archived already.

The next development was originally timed to coincide with the planned Fujitsu upgrade in 1998. Because of changes to that timetable the data handling system phase III has also been amended slightly. Since the project began our projections of data volume increases have largely been borne out in practice, and these suggest that we will have 250 Tbytes by early 2000.

ECFILE is now largely read only, and the major load on CFS is the copying out of it all the data that is required to be kept. We expect to finish that copying by 3Q98, when the old system will be turned off.

ECFS

This service is for users'private files, it is a traditional file archive where data is stored by file name using a UNIX-like user interface. Today it holds 8 Tbytes in 1 million files.

It runs on one server and, so far, the service is generally satisfactory. There have been the occasional peaks in demand but the system has coped so far and in fact it is a lot more responsive than the old ECFILE system.

The first limit we will hit on this service will be the number of files rather than data volume or throughput capacity. Thus in early 1998 we expect the largest data partition to become unmanageable because of the number of files it will hold. Thus we have a development plan in place to overcome this.

The current growth is almost linear, running at about 20 Gbytes per day. The biggest partition now has 650K files in it. We believe it will run into problems when this partition reaches the ¾ million mark, although the software has been verified out to 1 million files. To overcome this we will create a series of slave servers, each of which carries only part of the overall file system. A data base behind the scenes will keep track of what is stored on which server, so that the whole process will be transparent to users. Using this technique, we believe we can expand ECFS indefinitely without the users knowing about the underlying mechanism.

MARS

The user interface for MARS remains as it was before. The MARS client side is unchanged, except now it accesses both the old CFS based and the new DHS based MARS servers.

Today we have 17 Tbytes (plus 15 Tbytes backups) held in 150K files. It runs on two servers, but even so it is only just coping. It is handling all the current archive/retrieve requests plus a significant copying load from the old CFS based MARS. The current critical point in this service is the meta data base (which holds details of where the actual data is stored). It uses the whole of one processor, and is the current limiting factor on the number of requests we can handle at any one time.

The growth rate is currently about 150 Gbytes per day in 400 files. Thus each file holds a lot of data, average size 400 Mbytes compared with 4 Mbytes per file in the old CFS based MARS. This large file size was deliberately planned as we knew that ADSM can handle a lot of data, but not a lot of individual files. Thus we are confident that the new MARS can expand considerably beyond its present size, to at least 250 Tbytes.

Phase III

The next phase of the new DHS is to add further capacity so as to be able to support the 116 PE Fujitsu in a better manner. First of all, we propose to re-use some of the StorageTek tape silos currently on the CFS system. These silos have proved very effective and very reliable in the past. More server capacity will be provided adding an IBM SP system having four nodes in it.

The complex will then consist of four RS6000 systems plus the existing two IBM tape robots with 16 drives each, and the four node SP machine with extra disks and two StorageTek tape silos with 12 drives each.

It is planned that the MARS service will be transferred onto the SPmachine during 1998. The additional capacity and expandability of the SP will provide a satisfactory platform for MARS for some time to come.

R. Dixon

High performance FORTRAN

High Performance Fortran (HPF) is a language that facilitates parallel programming on a wide range of parallel computers, without the need for explicit message passing. This is achieved by adopting a single thread / global index space model, which is the same model as seen by a serial program. This article introduces the main features of HPF. It also describes work done in an EC funded project called HPF+ to make HPF more useable by complex applications such as ECMWF's IFS model. Finally, some issues of programming the IFS in HPF are discussed.

Background

HPF arose from the desire to have a common Fortran-based language to program the growing number of parallel computer architectures that existed in the late 1980's. These architectures included SIMD (Thinking Machines, MasPar) and MIMD (Cray C90, IBM SP1, Intel Paragon, KSR1, SGI Challenge) shared and distributed memory machines.

While the MIMD systems could be programmed using Fortran with message passing for inter-processor communication and synchronisation (e.g. PVM), this typically involved a large programming effort. In addition, there was a maintenance problem, as often both serial and parallel code versions were developed and both had to be supported. The SIMD systems (do any still exist?) were even worse as they required vendor specific compilers and libraries.

In May 1993, the first specification of HPF (v1.0) was published by the High Performance Fortran Forum (HPFF). HPFF is a U.S. body containing members from industry, universities and U.S. government laboratories. In November 1994, HPF v1.1 was released, mainly incorporating corrections and some clarifications to the language. This specification resulted in the development of a number of HPF v1.1 compilers. However, it was also recognised that the v1.1 standard required further development to be applicable to large industrial and scientific applications. One such effort was the HPF+ project, which investigated HPF language extensions based on the requirements of a small number of large applications, one of which was the IFS model from ECMWF.

The contributions made by the HPF+ project and other members of the HPFF resulted in the publication of the HPF v2.0 specification in January 1997. This and other HPFF documents are available at the HPF web site at http://www.crpc.rice.edu/HPFF/home.html.

Since that time the HPF+ project has identified the need for additional language features that will be proposed for a future specification of HPF. These are discussed later in the HPF+ project section.

HPF language (v2.0)

The HPF language is based on Fortran 90 and uses **!HPF$** directives to express parallelism. These directives appear as comments to Fortran 90 compilers, and to a large extent the same source can be compiled by such compilers for serial execution.

A good part of the HPF language consists of constructs which may be used to specify the mapping of data in the program to an abstract set of processors. This is achieved via a two-level mapping: first, data arrays are aligned with other objects, and then groups of objects are distributed onto an array of abstract processors. These processors are then mapped to the physical processors of the target machine in an implementation dependent manner.

The **!HPF$ DISTRIBUTE** directive is provided to control distribution of the dimensions of arrays onto an abstract set of processors. The distribution of a dimension may be described by selecting one of a set of predefined primitives which permit **BLOCK**, **CYCLIC**, **GEN_BLOCK** or **INDIRECT** mapping of the elements.

One of the most common form of distribution is the **block distribution**. As illustrated by the example in Figure 1, this means that each processor contains a single contiguous block of the array. The size of each processor's block is equal to the number of elements in the array divided by the number of processors. If the number of elements is not exactly divisible by the number of processors, some blocks will have fewer elements than others. This type of distribution is used for regular domain decomposition in fields like computational fluid dynamics and QCD, where most operations on an array element only involve its nearest neighbours. It ensures that neighbouring elements are normally on the same processor and therefore minimises costly inter-processor communication. In the IFS model, block distribution could be used to distribute grid-point space if a one-dimensional distribution only were required.

Another regularly used distribution is the **cyclic distribution**. The first element is on the first processor, the second on processor 2 and so on. If a 16 element array is distributed across 4 processor, processor 1 has elements 1, 5, 9 and 13, as shown in Fig. 2. This type of distribution is useful for applications where the work required for each element varies significantly throughout the array and locality is not important, for example, ray-tracing. Cyclic distribution could be used to distribute IFS grid-point space, and could result in an improved computational load balance arising from convective processes.

The **gen_block distribution** can be viewed as a generalisation of the block distribution by allowing non equal sized blocks, as shown in Figure 3. This distribution could be particular useful in the IFS to distribute fourier and spectral spaces, as these spaces involve a unequal distribution of Fourier coefficients and spectral waves, respectively.

Finally, an **indirect distribution** (Figure 4) allows each element of an array dimension to be mapped individually using a mapping array. This distribution

```
      REAL, DIMENSION(16) :: a
!HPF$ DISTRIBUTE (BLOCK) :: a

        P1  P2  P3  P4

        1   5   9   13
        2   6   10  14
        3   7   11  15
        4   8   12  16
```

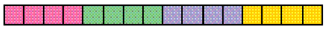
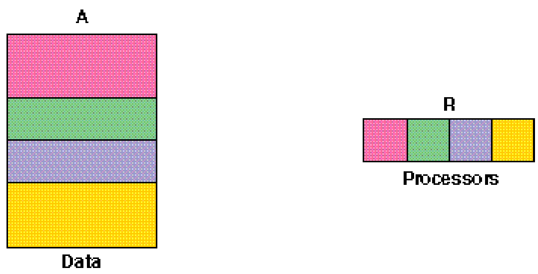


Figure 1: *Block distribution*

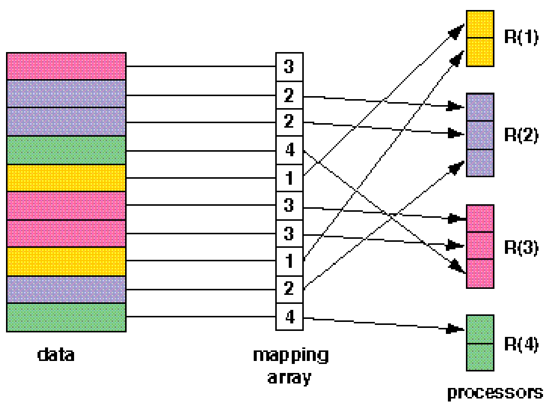
```
      REAL, DIMENSION(16) :: a
!HPF$ DISTRIBUTE (CYCLIC) :: a

        P1  P2  P3  P4

        1   2   3   4
        5   6   7   8
        9   10  11  12
        13  14  15  16
```

Figure 2: *Cyclic distribution*




```
!HPF$ PROCESSORS R(4)
      integer, parameter    :: bs(4) = (/ 30,20,20,30/)
      real, dimension(100)  :: a
!HPF$ DISTRIBUTE A(GEN_BLOCK(BS)) ONTO R
```

Figure 3: *Gen_Block distribution*




```
!HPF$ PROCESSORS R(4)
   integer, parameter :: map(10)=(/ 3,2,2,4,1,3,3,1,2,4 /)
   real, dimension(10,5) :: b
!HPF$ DISTRIBUTE B(INDIRECT(MAP),:) ONTO R
```

Figure 4: *Indirect distribution*

could have an application in the IFS to represent grid-point space where a 2-dimensional distribution of physical space is used to minimise communication required for the semi-Lagrangian calculation.

The **!HPF$ ALIGN** directive is used to align elements of data arrays to other data arrays. The rules for alignment are more flexible: a linear function may be used to specify the relationship between the mappings of two different data arrays. Mechanisms are also provided to enable dynamic modification of both alignment and distribution; these are the **!HPF$ REALIGN** and **!HPF$ REDISTRIBUTE** directives, respectively.

Once large data arrays are distributed, HPF compilers need to establish whether communication is required prior to executing DO loops, using an owner computes rule. If they can determine this at compile time then performance for such loops will be comparable to coding with explicit message passing. However, if these inter-processor communications (schedules) cannot be determined at compile time, then they have to be determined at run-time. This typically occurs when indirect addressing is used in the loop. This overhead is often referred to as the inspector overhead, and can be as much as an order of magnitude greater than the computational cost of the loop, even if no communication is required. Thus as a general rule it is a good idea to use array syntax in HPF programs, as it is much easier for HPF compilers to determine the required communication schedule at compile time.

Besides communication for DO loops, HPF supports redistribution of data on subroutine calls if the callers distribution differs from the subroutines distribution. A typical way of avoiding such redistribution is to use the **$HPF INHERIT** directive for the distributed dummy argument, which simply inherits the distribution from the caller.

The main expression of data parallelism in HPF is provided by the FORALL statement, the INDEPENDENT directive, and through the use of new intrinsic functions.

The **FORALL** statement is a data parallel construct which defines the assignment of multiple elements in an array but without enforcing an order on the assignments to individual elements. The reason behind providing FORALL is to create a parallel construct which guarantees identical results if applied in serial or in parallel. It can be thought of as a generalisation of the Fortran 90 array assignment, though focusing on the more fine-grain parallelism of the problem. FORALL is included in the Fortran 95 language standard.

The **!HPF$INDEPENDENT** directive is used to provide the compiler with additional information about the execution of a FORALLconstruct or a DO loop. It is a promise by the user to the compiler that the result of the loop statements will be the same whether executed in serial or in parallel. If the user breaks this promise and the iterations of the loop interfere with each other, then the code is not HPF conforming. This really means that the program will either produce wrong results or abort, but compiler writers prefer to use the 'non conforming'terminology! It should be noted that the INDEPENDENT directive should not alter the meaning of the code, it simply gives the compiler new information. This new information removes some restrictions from the compiler and allows it to generate more efficient code.

Afinal language feature to note in HPF is the **EXTRINSIC** mechanism. This allows different models of parallelism, LOCAL and SERIAL, and different languages, C and Fortran 77, to be supported. In particular, an **HPF_LOCAL** extrinsic procedure can be used to reduce the overheads of HPF programs by entering a procedure where only local data on each processor is accessible. This can be extremely useful for large sections of code where no communication is required (e.g. IFS physics calculations), or as an interface to a parallel I/O library.

HPF+ Project (Jan 96 - Dec 97)

The purpose of this project was to improve the current version of HPF and related compiler technology (initially HPF-1) by extending the functionality of the language and developing optimising compilation strategies, depending on the requirements of a set of advanced application codes. The key objectives of the project were:

- to develop a set of project benchmarks, reflecting advanced application problems,
- to develop a specification of an extended HPF language, 'HPF+', which addressed the requirements of the project benchmarks, and
- to participate in the standardization effort for HPF-2,
- to extend and implement optimising compiler technology for HPF+ in the framework of the Vienna Fortran Compilation System (VFCS),
- to extend the Measurements Description, Evaluation and Analysis Tool (MEDEA) for performance analysis of the project benchmarks,
- to evaluate the new language and compiler technology by using the project benchmarks and comparing the outcome with implementations based on HPF-1 and explicit message passing, and
- to support a transfer of technology from academia to industry.

The HPF+ consortium consisted of developers of large industrial applications (AVL List GmbH, ECMWF, and Engineering Systems International SA), plus both academic and commercial language, compiler, and tool developers (Universities of Vienna and Pavia, C&C Research Laboratories, and NA Software). The approach taken to achieve the goals of the project was based upon a close cooperation between all partners.

More specifically, the three application partners played a crucial role in the design and evaluation phases by selecting the benchmarks, identifying critical requirements, and establishing the evaluation criteria. On the other hand, the four language and tool developers based their work on the stated requirements and provided feedback to the application designers about the best ways of implementing critical code sections using HPF+ and commercial compilers for HPF.

The benchmarks produced in the HPF+ project reflected kernels of the applications developed by the application partners:

- **FIRE** (Fig. 5) from AVL is a fully interactive general purpose computational fluid dynamics program, designed specially for computing compressible and incompressible turbulent fluid flows as encountered in engineering environments.
- **IFS** (Fig. 6) from ECMWF is an Integrated Forecast System, comprising all the principal components for producing medium-range global weather forecasts using NWP (Numerical Weather Prediction) techniques.
- **PAM-CRASH** (Fig. 7) from ESI is a finite-element program for solving highly nonlinear dynamic problems arising in car crashworthiness simulation.

All three applications represent state-of-the-art codes that are characterised by irregular data structures. For example, the IFS forecast model employs a reduced grid to represent physical space, where the number of grid-points per latitude reduce as the poles are approached.

One of the key developments in HPF+ has been the support for communication schedule reuse. In a large number of applications it has been recognised that iterative processes exist which result in the generation of the same communication schedule for successive executions of INDEPENDENT DO loop nests. As stated earlier, this results in a high overhead, as the inspector phase is executed each time the same DO loop nests are executed. However, by a simple syntactic extension to the INDEPENDENT directive, the HPF compiler can be directed to save the communication schedule and reuse it on subsequent occasions. This mechanism has been used in a number of the HPF+ benchmark kernels and has typically reduced HPF DO loop overheads by a factor of five.The REUSE specification also supports a logical expression as an argument, thereby giving the user control on when the communication schedule should be recomputed. Without such an expression, the default is to compute the communication schedule on the first occasion, and then re-use the schedule on all subsequent occasions.

Another important development in the HPF+ project was the extension of an application called MEDEA (Univ. Pavia) which can be used to perform a statistical analysis of performance trace files produced by an executing HPF+ program. Metrics measured include all generated message passing calls and also the inspector and work distributor overheads of INDEPENDENT loops.

Ongoing work in HPF+ is focussing on the parallelisation of nested INDEPENDENT loops and the support of a HALO mechanism. The HALO mechanism is important in the IFS Semi-Lagrangian calculation as this involves a large amount of inter-processor communication to determine the departure point and mid-point of the trajectory and to interpolate quantities at these locations. Without a HALO mechanism, the semi-Lagrangian calculation would involve an excessive

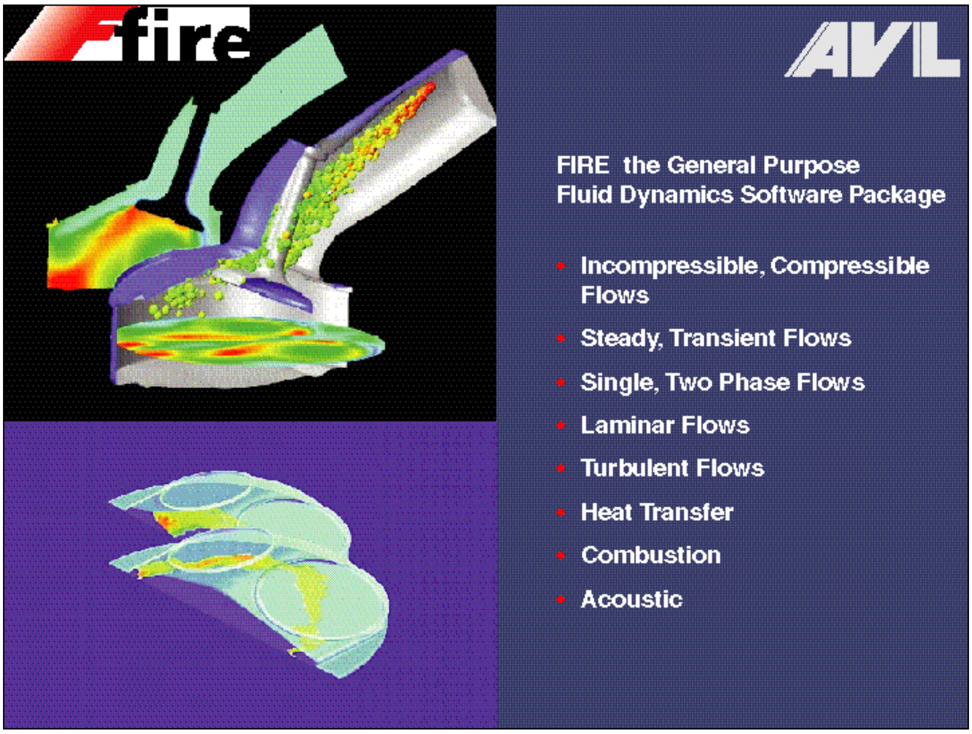


Figure 5 (left): *FIRE from AVL is a fully interactive general purpose computational fluid dynamics program, designed specially for computing compressible and incompressible turbulent fluid flows as encountered in engineering environments.*

Figure 6 (below): *IFS from ECMWF is an Integrated Forecast System, comprising all the principal components for producing medium-range global weather forecasts using NWP (Numerical Weather Prediction) techniques.*

inspector overhead for each variable being interpolated. This could be resolved by using a HALO specification within a distribute directive which would result in a single communication schedule and thereby improved performance.

The HPF+ project web page can be found at http://www.par.univie.ac.at/hpf+, together with benchmark kernel descriptions, evaluation reports and the latest HPF+ language specification.

HPF+ is supported by the European Commission under the ESPRIT IV, Long Term Research Project nr. 21033.

IFS HPF Issues

Experience gained from the HPF+ project has shown that it should be possible to code an HPF version of IFS, given a HPF 2.0 compiler. This version of the IFS should theoretically be easier to maintain, however the performance is unlikely to be acceptable in the production environment at ECMWF.

The development of such an HPF version would also require several man-years of effort, and therefore is not a viable proposition for the present IFS. In addition, HPF 2.0 compilers are only just becoming available and are typically provided as pre-processors to existing f90 compilers. This can make debugging more challenging, as the transformed source presented to the underlying f90 compiler is often substantially different from the original HPF source. However, we should keep an open mind on HPF for future applications at ECMWF, especially when such applications are written from scratch.

The structure of new HPF applications should attempt to utilise the good features of HPF (global view, easy distributions), and avoid the bad (run-time overhead of INDEPENDENT loops and cost of redistributions).

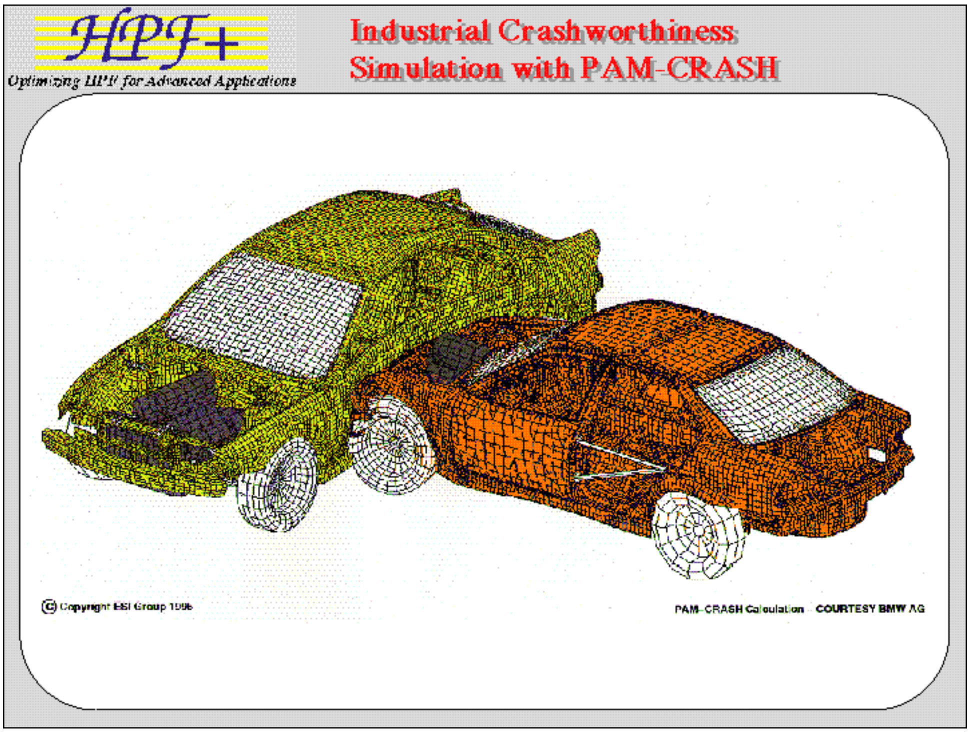


Figure 7: *PAM-CRASH from ESI is a finite-element program for solving highly non-linear dynamic problems arising in car crashworthiness simulation.*

The use of array syntax and FORALL will be the key to obtaining good parallel performance. Likewise, the use of the extrinsic mechanism HPF_LOCAL should be considered for large sections of code where locality is evident.

The Future

The HPF 2.0 language specification is now well defined and compilers are becoming available. However, this specification consists of two parts, the base HPF 2.0 language and HPF 2.0 approved extensions. HPF 2.0 compilers are required to implement the base language, but not the approved extensions. The approved extensions include GEN_BLOCK, INDIRECT distributions and extrinsic HPF_LOCAL. So users of HPF compilers need to consider whether to use any approved extensions for fear of being locked into a particular compiler vendor or computer manufacturer.

This issue has been recognised by the Japanese supercomputer vendors (Fujitsu, Hitachi, NEC), who, together with university and industrial partners, have initiated a Japan Association of High Performance Fortran (JAHPF). In particular, this group is some way towards defining a HPF v2.0 superset language called HPF/JA that will be implemented by all JAHPF supercomputer vendors. This will be an interesting development to follow. The web page for JAHPF can be found at http://www.tokyo.rist.or.jp/~shunchan/index-e.html.

In recent years, a new generation of high performance architectures has become commercially available. Many of these machines are either symmetric shared-memory architectures SMPs or clusters of SMPs, where an interconnection network connects a number of nodes, each of which is an SMP. Thus these machines display a hybrid structure integrating shared-memory with distributed memory parallelism. One of their dominating characteristics is their use of a deep memory hierarchy, often involving multiple levels of cache.

As a consequence, these architectures have not only to deal with the locality problem typical for distributed memory machines - which is addressed by HPF - but also with cache locality. A cache miss in a program executing on a cluster of SMPs may be more expensive than a non-local memory access.

HPF and its compilers are currently not designed to deal with such issues. The future will show whether a (possible extended) HPF paradigm will be able to cope with such architectures, or whether other programming methods will prove more adequate.

George Mozdzynski

The ECMWF Re-Analysis Sample Data CD-ROMs 1979-1993

ECMWF has provided global atmospheric analyses from its archive for many years. The ECMWF Re-analysis (ERA) project was devised in response to wishes expressed by many users for a data set generated by a modern, consistent, and invariant data assimilation system. The result is a validated 15-year data set assimilated data for the period 1979-1993.

Two CD-ROMs have been produced containing selected fields in GRIB format for the ERA period.

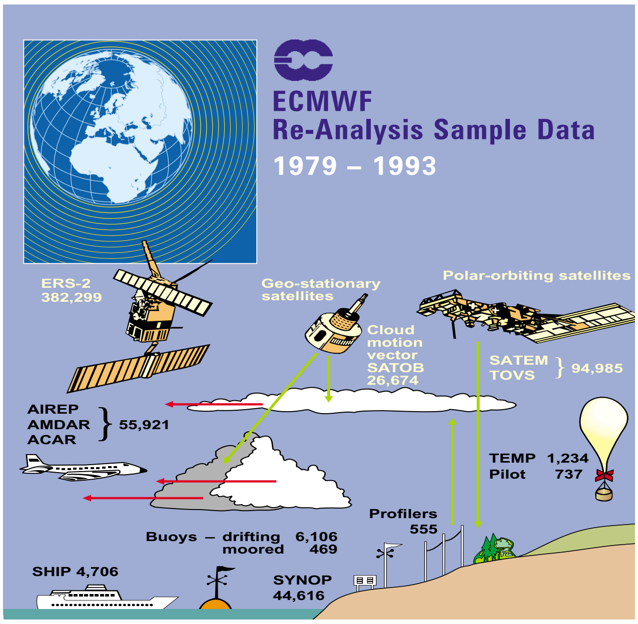
The following daily fields are included:

- Instantaneous fields valid at 12UTC
- Geopotential (z) on levels 1000 and 500 hPa, 8 bit,
- Temperature (t) on 850 and 200 hPa, 8 bit,
- U- and V-Wind components (u, v) on 850 and 200 hPa, 8 bit,
- U10- and V10-Wind components (u10, v10) at 10 metre height, 12 bit,
- total cloud cover (tcc), 12 bit,
- two meter temperature (t2), 12 bit,
- two meter dew point temperature (d2), 12 bit.

Fields of quantities accumulated over 24-hour periods centred on 12UTC:

- evaporation (e), 12 bit,
- total precipitation (large-scale plus convective) (tp), 12 bit.

All of the fields on these CD-ROMs have been converted to a uniform 2.5 degree resolution latitude-longitude grid.



One CD-ROM contains the fifteen years of surface fields while the other contains fifteen years of upper air data, the ECMWF Re-Analysis Project Report and a picture gallery of seasonal means, monthly means and cross-sections.

Graphics display systems and GRIB decoding tools are provided on both CD-ROMs.

Keith Fielding

ECMWF Calendar 1998

Feb 23 - 6 Mar	**Computer User Training Course**		
23 - 25 Feb	**COM1** - *Introduction for new users*		
25 - 27 Feb	**COM2** - *Fujitsu optimisation*		
2 Mar	**COM3** - *MARS*		
3 - 6 Mar	**COM4** - *Graphics*		
Mar 16 - 12 Jun	**Meteorological Training Course**		
16 - 24 Mar	**MET1**	*Numerical methods, adiabatic formulation*	
24 Mar - 3 Apr	**MET2**	*Data assimilation & use of satellite data*	
20 - 24 Apr	**MET3**	*General circulation, systematic model errors and predictability*	
27 Apr - 8 May	**MET4**	*Parametrization of diabatic processes*	
11 - 21 May, *[2 - 12 Jun]	**MET5**	*Use & interpretation of ECMWF products*	

* two sessions if required

May 6 - 7	Policy Advisory Committee	*9th*
May 11 - 12	Security Representatives meeting	
May 18 - 19	Finance Committee	*59th*
Jun 29 - 2 Jul	Workshop - *Modelling and data assimilation for land-surface processes*	
Jun 15 - 16	Expert meeting on EPS	
Jun 17 - 18	Seasonal Forecasting Users meeting	
Jun 23 - 24	Council	*48th*
Sep 7 - 11	Seminar - *Recent developments in numerical methods for atmospheric modelling*	
Sep 28 -30	Scientific Advisory Committee	*27th*
Oct 7 - 9	Technical Advisory Committee	*26th*
Oct 13 - 14	Finance Committee	*60th*
Nov 2 - 4	Workshop - *Diagnosis of Data Assimilation Systems*	
Nov 9 - 13	Workshop - *WGNE/GCSS/GMPP - Cloud processes in large-scale models*	
Nov 16 - 20	Workshop - *Parallel Processors*	
Dec 2-3	Council	*49th*

ECMWF Annual Seminar

Recent developments in numerical methods for atmospheric modelling

7-11 September 1998

The topic for the 1998 annual seminar is Recent Developments in Numerical Methods for Atmospheric Modelling, and will be held from 7 to 11 September.

The seminar will provide a pedagogical review of these developments which are crucial to the continuing evolution of NWP and climate models. Many operational centres are using semi-Lagrangian advection schemes which provide substantial economies in computing cost. These schemes continue to be refined and there is activity in methods based on conservation of potential vorticity. The development of non-hydrostatic models for operational use grows steadily. Also, there is consideration being given to alternative horizontal representations (e.g. icosahedral grids, wavelets) and the choice of vertical staggering. Finally, there is now much experience in the suitability and adaptability of methods on new computer architectures.

Posters providing further information on the programme and application forms will be distributed around May 1998.

Els Kooij-Connally

Member State computer resource allocations 1998

Member State	Fujitsu (kunits)	Data (Gbytes)
Belgium	78	428
Denmark	63	348
Germany	378	2084
Spain	126	697
France	260	1436
Greece	56	308
Ireland	48	262
Italy	220	1216
Yugoslavia*	51	280
Netherlands	95	524

Member State	Fujitsu (kunits)	Data (Gbytes)
Norway	60	332
Austria	72	398
Portugal	55	304
Switzerland	83	459
Finland	56	307
Sweden	74	411
Turkey	67	369
United Kingdom	210	1161
Special projects	228	676
Total	2280	12000

* In accordance with UN Security Council Resolution 757 (1992) of 30 May 1992, the Council instructed the Director to suspend the telecommunications connection to Belgrade with immediate effect. This took place on 5 June 1992. As a consequence no operational products are disseminated to Belgrade and access to the Centre's computer system is not available to Belgrade.

Special Project allocations 1998

Member State		Institution	Project title	Fujitsu units	Data storage Gbytes
Continuation Projects					
Austria	**1**	Univ für Bodenkultur, Vienna (Kromp-Kolb)	Vertical ozone transports in the Alps	500	4.2
	2	Univ Vienna (Ehrendorfer)	Covariance evolution and forecast skill	700	5.0
	3	Univ Vienna (Hantel)	Estimating the global mean sub-gridscale energy conversion term	30	2.0
France	**4**	L.A.M.P., Aubière (Cautenet)	Chemistry, cloud and radiation interactions in a meteorological model	93	2.0
	5	CERFACS (Thual)	Universal software for data assimilation: variational method for global ocean	1,120	7.0
Germany	**6**	Inst. for Geo. and Met. (Speth)	Interpretation and calculation of energy budgets	2	6.0
	7	MPI Hamburg (Roeckner)	Modelling the earth's radiation budget and evaluation against ERBE data	4,500	15.0
	8	MPI, Hamburg (Bengtsson)	Numerical experimentation with a coupled ocean/atmosphere model	10,625	60.0
	9	MPI, Hamburg (Bengtsson)	Simulation and validation of the hydrological cycle	8,500	50.0

Member State		Institution	Project title	Fujitsu units	Data storage Gbytes
	10	Univ. of Munich (Wirth/Egger)	The behaviour of cut-off cyclones in ECMWF analysis: impact of diabatic processes on their development and decay	70	1.0
	11	FU, Berlin (Fischer/Thoss)	Comparison of the ECMWF cloud scheme with multi-spectral satellite data in the Baltic Sea	70	5.0
	12	GKSS, Geesthacht (Rockel)	Energy and water cycle components in regional forecasts, remote sensing and field experiments	50	0.2
Ireland	13	Met Éireann (Lynch)	The HIRLAM 4 project	3,500	5.0
Italy	14	Ist. per lo Studio della Dinamica delle Grandi Masse, Venezia (Cavaleri)	Testing and applications of a third generation wave model in the Mediterranean Sea	3,000	1.5
	15	Univ Bologna (Rizzi)	TOVS 1B radiances and model simulations	50	20.0
	16	CINECA, Bologna (Molteni)	Influence of springtime land-surface conditions on the Asian summer monsoon	6,000	7.0
Netherlands	17	KNMI, De Bilt (Siegmund)	Stratosphere-troposphere exchange	400	4.0
	18	KNMI (van Velthoven)	Chemistry and transport studies with a 3D off-line tracer model	3,000	25.0
	19	KNMI (Komen)	North Sea wave climate	17,000	5.0
	20	KNMI (Drijfhout)	Agulhas	10,500	0
	21	KNMI (Komen)	Validation of re-analysed A/S fluxes	17,000	5.0
	22	KNMI (Siebesma)	Large Eddy Simulation (LES) of boundary layer clouds	6,000	10.0
	23	KNMI (Opsteegh)	Short term, regional probabilistic forecasting using IFS	4,000	10.0
	24	KNMI (Kelder)	The relation between satellite ozone observation errors and dynamical errors in the ECMWF model	1,400	7.0
	25	KNMI (Burgers)	OGCM mixed-layer modules	7,500	0.0
Norway	26	Geophysical Inst.,Univ. of Bergen (Grønås/Kvamstø)	Parametrization of clouds in general circulation models	100	0.2
	27	Univ Oslo (Isaksen)	Ozone as a climate gas	700	5.0
Turkey	28	Wea Forec Dept (Yildirim)	Numerical study of small-case disturbances formed in Mediterranean Sea and their impact on Turkish provinces	2,500	2.0
UK	29	Univ Reading (Hoskins)	Routine back trajectories	7,000	4.0
	30	Univ Oxford (Sutton)	Tropical seasonal cycle in the coupled atmosphere-ocean system	3,734	25.0
	31	CGAM, Univ Reading (Slingo)	Predictability experiments for the Asian Summer Monsoon	6,000	7.0
	32	Br Antarctic Survey, Cambridge (Turner and Leonard)	Assessment of ECMWF forecasts over the high latitude areas of the S. Hemisphere	0	1.0
New projects					
Austria	33	Univ Innsbruck (Mayr)	Heavy convective precipitation over and along mountains - numerical simulations	1,800	2.0
Belgium	34	Univ Louvain (Van Ypersele)	Modelling the climate and its evolution at the global and regional scales (CLIMOD Network)	12,600	4.0
Germany	35	Univ Freising (Stohl)	Validation of trajectory calculations	200	3.0
	36	DLR Institut für Physik der Atmosphäre (Hoinka)	The climatology of the global tropopause	50	10.0
Switzerland	37	Univ Berue (Schüpbach)	The climatology of global transport (long-range transport) to 'background' stations	21,000	85.0
France	38	CNES (Courtier)	MERCATOR	42,500	160.0
			Total requested	**203,794**	**565.1**
			Reserve (to be allocated by ECMWF)	**24,206**	**110.9**
			Total overall	**228,000**	**676.0**