Dick Dixon

It is with great sadness I have to report the death of Dick Dixon, Chief Technical Advisor to the Computer Division, on 13 May at the age of 49. Dick became a victim of cancer after a long fight, a fight conducted with great tenacity, dignity, and courage right up to the end.

Dick joined ECMWF in 1978 from Leicester University Computer Laboratory where he had been a systems analyst. From the start he showed great technical prowess and skill that was to become the hallmark of his career at ECMWF. Around 1983 he took up the challenge of providing ECMWF with a reliable and robust data server that would cope with the ever burgeoning storage needs of the meteorologists, and in 1993 he became the project officer responsible for providing today's Data Handling System. He succeeded in this brilliantly, as he had in other challenges, becoming in the process one of the world's acknowledged experts in large scale data storage.

But computers and computing were not his only forte. His passion for fine malt whiskies was renowned. He was proud of his early membership number in the Scotch Malt Whisky Society, and made many whisky pilgrimages to Scotland with friends from ECMWF. Over the years he also became an accomplished fly fisherman, resulting on many occasions in an invitation to Centre staff to help themselves to a free trout from the fruits of his endeavours. A keen cyclist, he regularly took part in the annual London to Brighton cycle run, ever proud that each time he had managed to scale the notorious Ditchling Beacon hill without dismounting and that he had relieved many of his colleagues of money to support his favourite charity. On a distinctly less energetic level he regularly joined the canal boat trips organised from within the Centre. Much planning and coordinating was required in order to arrive at the next pub just 5 minutes before opening time, a feat regularly and reliably achieved!

Dick's technical expertise was substantial. He was well read, keeping up to date across the board in computing matters. That knowledge will be sorely missed, by his colleagues both at ECMWF and IBM with whom he worked extensively. It was a privilege to have known and worked with Dick, to have shared his companionship, his love of life, his gentle humour, and his good natured friendliness. ECMWF has lost much by his untimely death.

David Burridge (Director ECMWF) and Staff

In this issue

European Centre for
Medium-Range Weather Forecasts

Shinfield Park, Reading, Berkshire RG2 9AX, UK

Fax: . . . . +44 118 986 9450
Telephone: National 0118 949 9000
International . . . . +44 118 949 9000
Internet URL . . . . http://www.ecmwf.int

Cover

The front page shows the Breitling Orbiter 3 balloon soon after take-off from Châteaux–d'Oex, Switzerland.

(Photograph courtesy of Jean-François Luy)

Editorial

The article on page 2 by Pierre Eckert (Switzerland) and Luc Trullemans (Belgium) describes the problems of weather forecasting for the first successful round-the-world flight by a manned balloon that took place earlier this year. It highlights the importance of the guidance provided by forecast products from the ECMWF and the NCEP medium-range forecasting models, and it graphically illustrates the need for the forecasters to think of the atmosphere as an evolving three dimensional system, in contrast to the usual aviation forecasting requirement for predictions at constant-pressure flight levels.

Details the revised system for scheduling, monitoring and supervising tasks of the operational suite (XCDP) are given on page 7. The new system has been designed to address the differing requirements of operators, analysts and researchers, and includes improved graphical displays and many additional features. Plans for increasing the Centre's computer power are now in place (page 15). Next year the total computing capability of the VPP system will be more than 1.3 TFlops; this is over 3.5 times what is available now on the VPP700 and the VPP700E.

Changes to the Operational Forecasting System

Recent changes

On 13 July 1999, the following changes were introduced (Cy 21r2):

- a new coupling of the physics and dynamics components of the numerical scheme;
- the use of a new screen-level humidity and temperature analysis for the optimum interpolation analysis of soil temperature and humidity;
- a slight relaxation of the quality control of the TOVS/ATOVS radiance data, and the use of Meteosat AMV (winds) at 90-minute frequency using EUMETSAT quality control information;
- US profilers' hourly data are now active;
- ship wind data make use of WMO tables for height assignment, while ship humidity data are blacklisted;
- the bias correction of radiosonde mass information is reintroduced in a revised form to apply temperature corrections using the sonde-type information provided in the TEMP message;
- a modified wave energy balance equation.

Planned changes

The next change (Cy21r4) is planned for 12 October; it includes an increase in the number of vertical levels to 60 (most of the extra resolution being near the surface), with modifications to the cloud and convection scheme, and revised mean and subgrid-scale orography fields.

Brian Norris

Breitling Orbiter 3
Meteorological aspects of the balloon flight around the world

Over the past few years a passion has developed for what has been called “the last aeronautical adventure of the century”: uninterrupted flight of a balloon around the world. The rules for such a flight have been drawn up by the Fédération Aéronautique Internationale (www.fai.org/-ballooning/rtw298.htm). In particular, they stipulate that the start meridian must be crossed a second time and a minimum distance of about 28000 km must be covered, so as to avoid short flights close to the poles.

In 1993 a transatlantic balloon race took place and was won by a team composed of the Belgian pilot Wim Verstraten and the Swiss pilot Bernard Piccard. The meteorological assistance for the flight was provided by the Royal Meteorological Institute of Belgium under the leadership of one of the authors of this article (Luc Trullemans).

The Swiss Meteorological Institute has, with the help of the second author of this article (Pierre Eckert), established a solid reputation in routing yacht races, most notably three Whitbread around-the-world races. It was, therefore, quite natural for Bertrand Piccard and Wim Verstraten to come and see us in 1995 to weigh up the meteorological possibilities for tackling the challenge of a round-the-world balloon flight. A climatological study showed that the best season is winter, in order to take advantage of the regularity of the subtropical jet stream that blows at a latitude of about 35°N. It is also important to be able to change altitude frequently in order to benefit from variations in wind direction with height, so as to be able to steer the balloon in the desired direction. If, moreover, one wants the flight not to take too long, and to avoid closed circulation systems which often occur at low altitude, the ability to fly at heights up to 10,000 m, or even 12,000 m, is essential. All this led to the construction of a Rozier type balloon whose basic lift was assured by a large quantity of helium gas that maintained a constant altitude by day, additional lift being provided by a propane-heated air cone (required especially at night). A pressurised capsule allowing for more than 20 days duration also had to be built. The balloon and capsule, both representing considerable technological feats, were designed and built by the British company Cameron Balloons. Technical aspects of the project are described on the web site www.breitling-orbiter.ch.

The Swiss alpine resort Château-d’Oex, situated 1000 m above sea level, was chosen as the starting point for the flight. This location is well known for its many activities relating to hot-air ballooning, including hosting an international meeting usually held in January. The advantage of Château-d’Oex is the satisfactory control of wind at the bottom of the valley; about six hours of calm was required to inflate the balloon. The disadvantage is the departure from a rather high latitude for joining the subtropical jet stream, but climatological simulations showed that this was frequently possible.

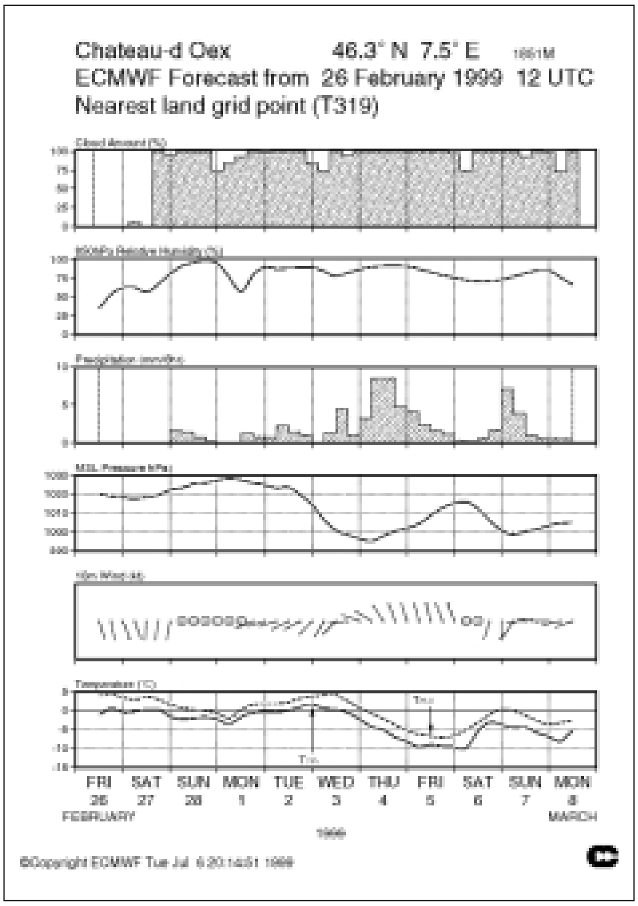


Figure 1: The meteogram covering the launch period. One can see the predicted short ‘window’ clear of precipitation during the morning of 1 March 1999; in reality, the window was a little longer and the cloud cover much less.

Our team attempted the project unsuccessfully in both 1997 and 1998; about twenty other competitors did so also. The winter of 1998-1999 was again considered for a possible start. The two pilots chosen for this attempt were Bertrand Piccard (Switzerland) and Brian Jones (Britain). There was a strong geopolitical constraint to be satisfied also, namely to cross China south of latitude 26°N (another window north of 43°N was also available, but we did not consider this). From November 1998 we concentrated on situations with north-westerly winds over the Alps, which could take us first to Egypt and then south of the Himalayas. To do this we calculated trajectories based on the 12 UTC model of ECMWF and on the MRF model of NOAA. Local weather conditions (precipitation, cloud, and wind) were also forecast in the medium term and their probabilities were evaluated with the help of the ECMWF ensemble forecasts.

We noticed that situations of strong north-westerly winds would indeed take the balloon quickly to latitudes around 30°N. However, since they were usually gener-

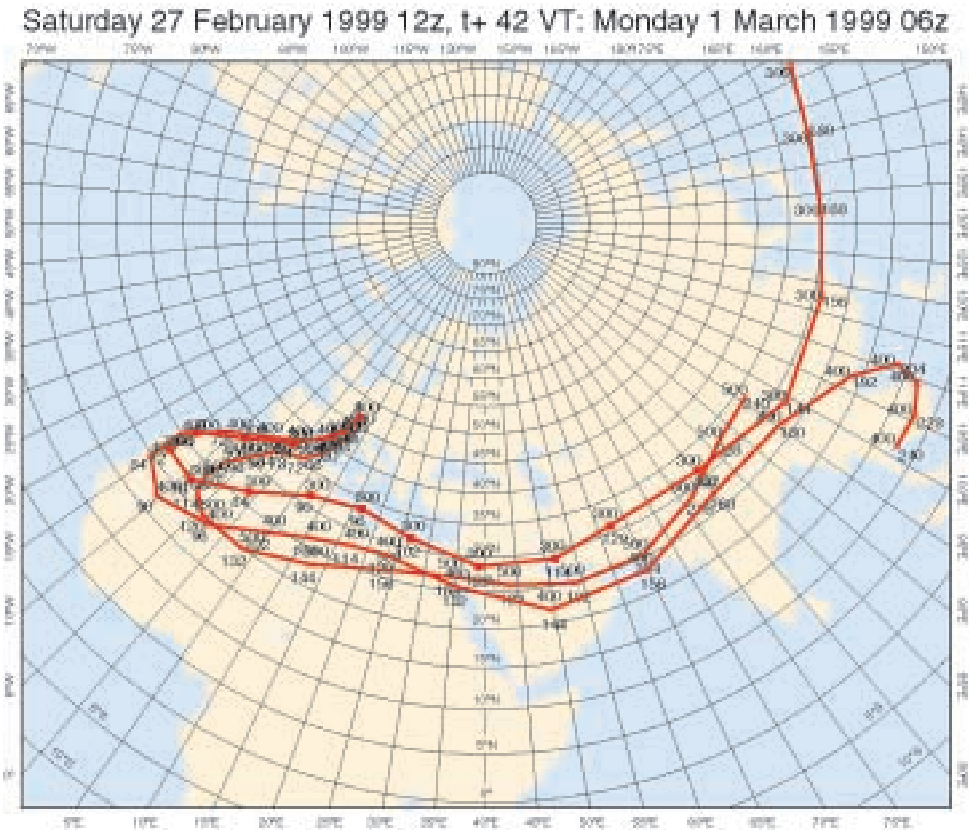


Figure 2: The trajectories we had at our disposal when we took the final decision to start. The numbers above the trajectories are the pressure levels in hPa. The numbers below are the times (hours) after the forecast initialisation time (12 UTC 27 February 1999). We chose to fly at heights between 500 and 400 hPa in order to reach Oman.

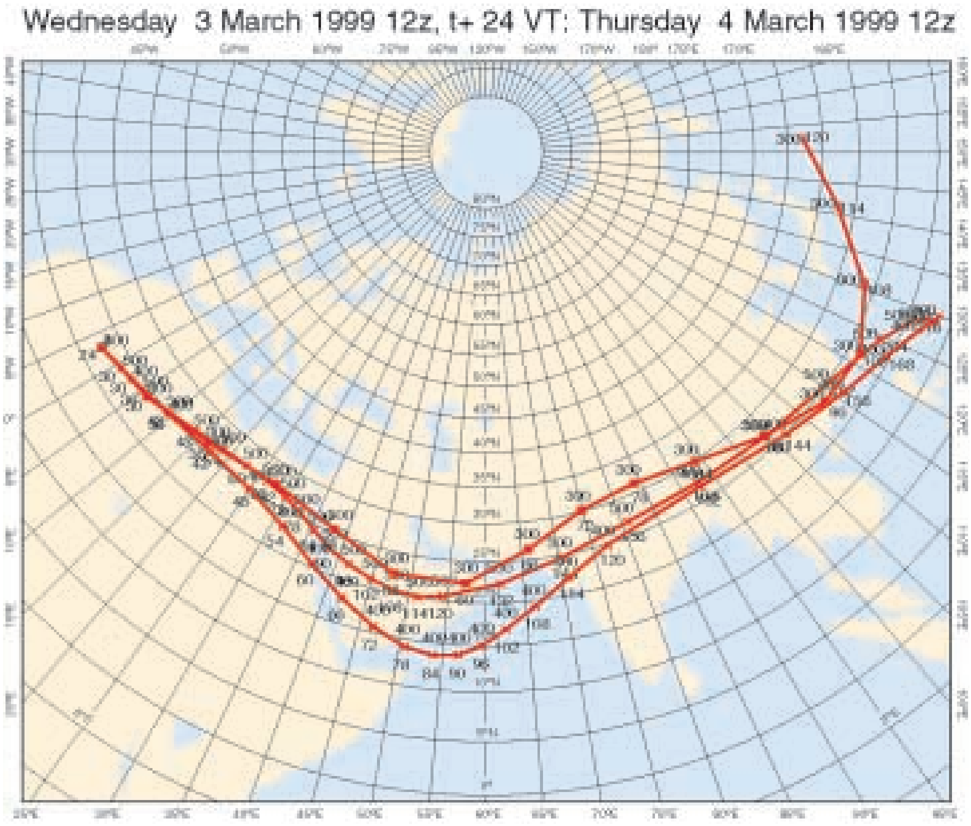


Figure 3: Trajectories at 500 hPa (5,500 m), 400 hPa (7,000 m) and 300 hPa (9,000 m) starting from northern Mali. One can see that flying too high (9,000 m) would have led to a track north of the Himalayas (contravening the flight clearance given by the Chinese authorities). Consequently, between Africa and India we chose a trajectory between 500 and 400 hPa.

Breitling Orbiter 3 crew *(photo courtesy of Jean-François Luy)*

ated by a low to the east of the Mediterranean, they systematically recurved into the Caspian Sea region. The only solution was to leave in a situation of northerly (or even north-easterly) winds with the possibility that the depression might either move east or dissipate. Such a situation occurred on the morning of 1 March 1999. The window was very short, from about 00 UTC to 12 UTC, and the initial route had to avoid too high an altitude (Figures. 1 and 2).

The start went according to plan, the balloon crossing southern France, the Balearic Islands and Morocco. After a diversion of 17° longitude in the wrong direction, the turn to the east was made above Mauritania with an entry into the subtropical jet stream. The winds above Africa, however, were still blowing in a direction that made it impossible to fly above 6,000 m, otherwise there was a risk of being taken north of the Himalayas (Figure 3).

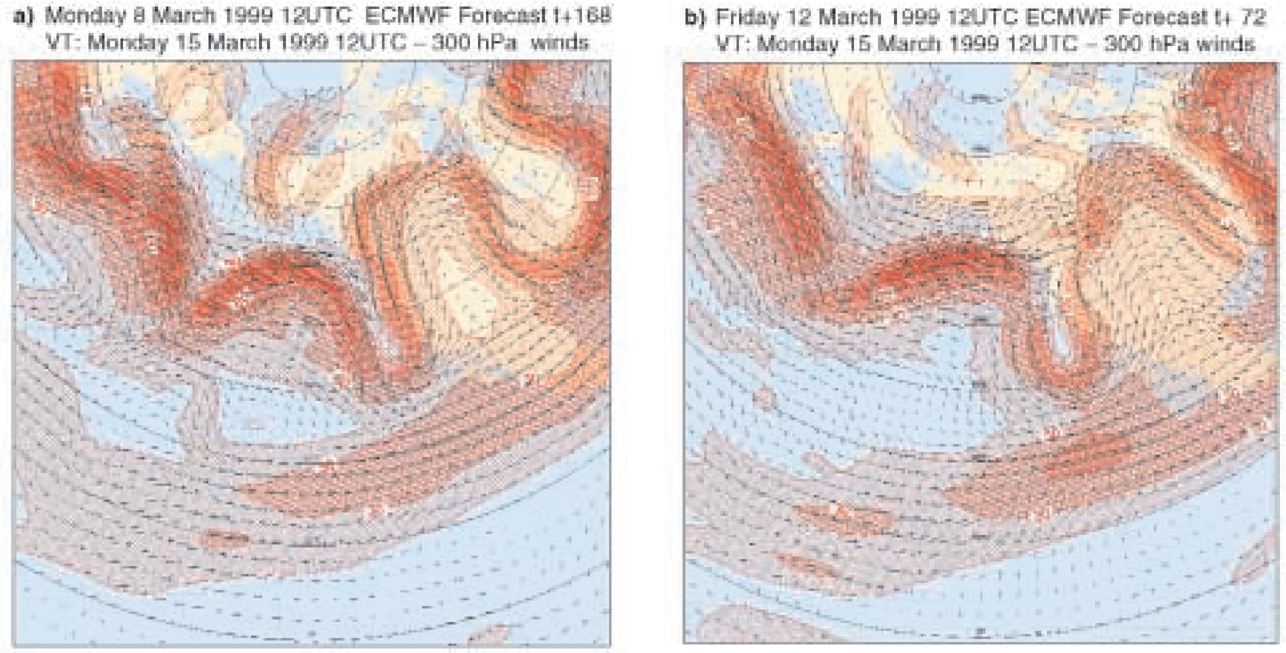
The passage over Oman at 17°N achieved our first intermediate target, and the route towards India and Bangladesh was undertaken at an altitude of 6,000 m. The entry into Chinese air space at 25°15' N, and the transit of the balloon over a track of more than 2,000 km between the 25th and 26th parallel (coming closer than 40 km to the no-fly zone), show the degree of meteorological precision that was attained during this phase of the flight (Figure 4).

After crossing a third of the Pacific, a crucial decision arose: whether to fly at about 7,000 m and join a branch of the polar jet stream, or whether to fly above 9,000 m with a view to catching the subtropical jet stream that was forming south of Hawaii at around 10°N. We decided to exclude the former, not only because it would have taken the balloon into zones of turbulence and icing, but also because it would have led to a rather high probability of crossing the North Atlantic within a second no-fly zone, forbidden because of the volume of civilian air traffic. The southern option, however, gave us good confidence; a strong flow from a position south of Hawaii towards Mexico had been forecast for some time by our two models (Figures 5 and 6). Moreover, the forecast models did not predict too much storm activity at these latitudes, which are quite close to the equator.

By choosing the southern option, Breitling Orbiter 3 travelled at speeds of about 30 knots in a south-easterly direction. On 15 March it reached the surprisingly low latitude of 9°N and was travelling at speeds of about 60 knots. Later, heading in a more northerly direction towards Mexico, the balloon reached speeds of up to 100 knots, very close to the heart of the jet stream. It was then that the numerical models deceived us by predicting a trajectory directed towards the centre of Mexico, Yucatan and Cuba; in reality the balloon went more to the south and passed over Guatemala, Belize and the Honduran coast. This development was caused by an error of only about 10° in wind direction, but the loss of time was considered critical because only four of the 32 propane containers remained.

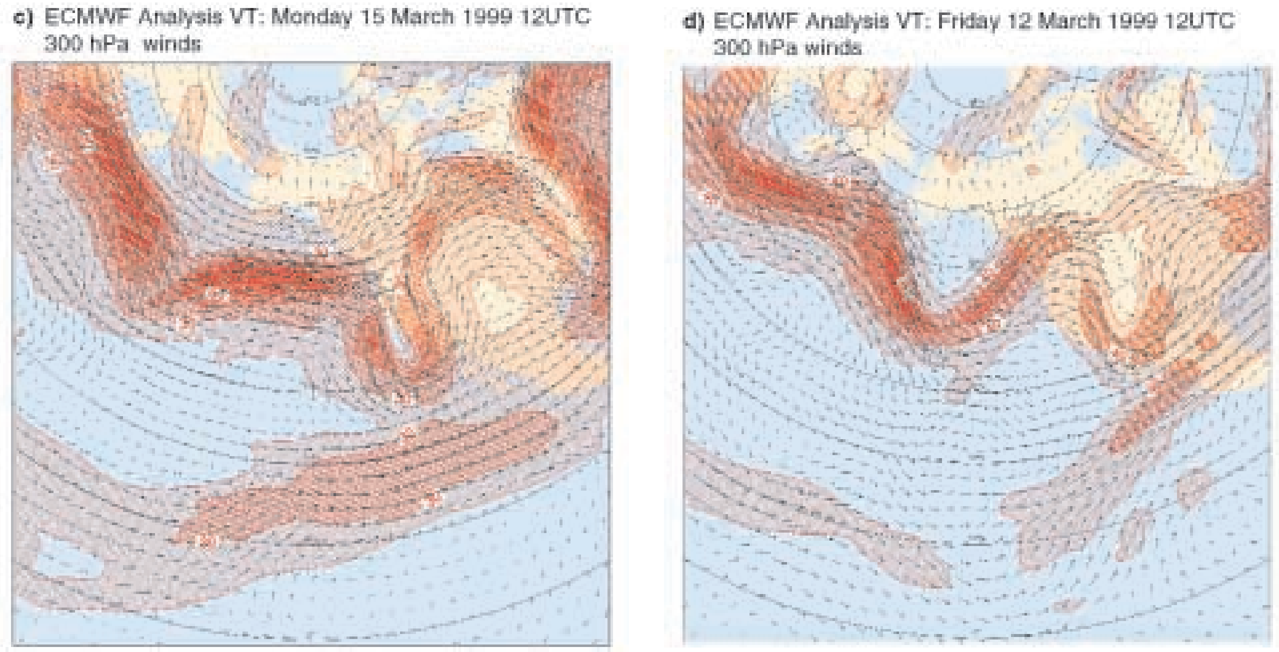
The ECMWF model was the first to correct the error. It did show a direction of about 100° at 300 hPa (9,000 m), but it also showed that the situation could be rescued by climbing to 250 hPa (11,000 m). The balloon was then already flying at an altitude of more than 10,000 m and

Figure 4: The passage over China; the full, red line is the path followed by the balloon, and the dotted, blue line is the 26°N latitude limit, north of which flight approval was denied. This very precise trajectory, between 25°N and 26°N at speeds of 80 to 100 knots, justified delaying the launch until 1 March 1999.



a) Monday 8 March 1999 12UTC ECMWF Forecast t+168
VT: Monday 15 March 1999 12UTC – 300 hPa winds

b) Friday 12 March 1999 12UTC ECMWF Forecast t+ 72
VT: Monday 15 March 1999 12UTC – 300 hPa winds



c) ECMWF Analysis VT: Monday 15 March 1999 12UTC
300 hPa winds

d) ECMWF Analysis VT: Friday 12 March 1999 12UTC
300 hPa winds

Figure 5: Analyses and predictions of the subtropical jet stream at 300 hPa south of Hawaii and towards Mexico; the balloon caught the jet stream on 15 March at 9°N. (a) The 7–day forecast, (b) the 3-day forecast and (c) the analysis for 12 UTC 15 March. The good forecast was a considerable help in avoiding the polar jet stream. (d) The analysis three days earlier, the day we took the decision to go south; at this time, the jet stream had only about half the strength it took three days later.

was continuing its progression too far southward. In a telephone call to the pilots we suggested to them that they rise as high as possible (to the balloon's maximum height) to recapture the correct trajectory towards Africa. In the event a climb of merely 300 m was sufficient for the direction to change from 102° to 75° (Figure 7) – that's what you call a wind shear! Now the race was almost won because we could cross the Atlantic Ocean at speeds of between 60 and 100 knots.

The finishing line for the flight (9°30' W) was crossed at 10 UTC on 20 March over Mauritania, but since it was not practical to land a balloon of this size in the heat-generated turbulence of the afternoon, we continued our flight for another night. As the subtropical jet stream was quite strong at that moment and the Orbiter could climb to 12,000 m because it had become lighter, it was possible to cross all of north Africa in 18 hours, and the balloon landed at dawn on 21 March in the Egyptian desert.

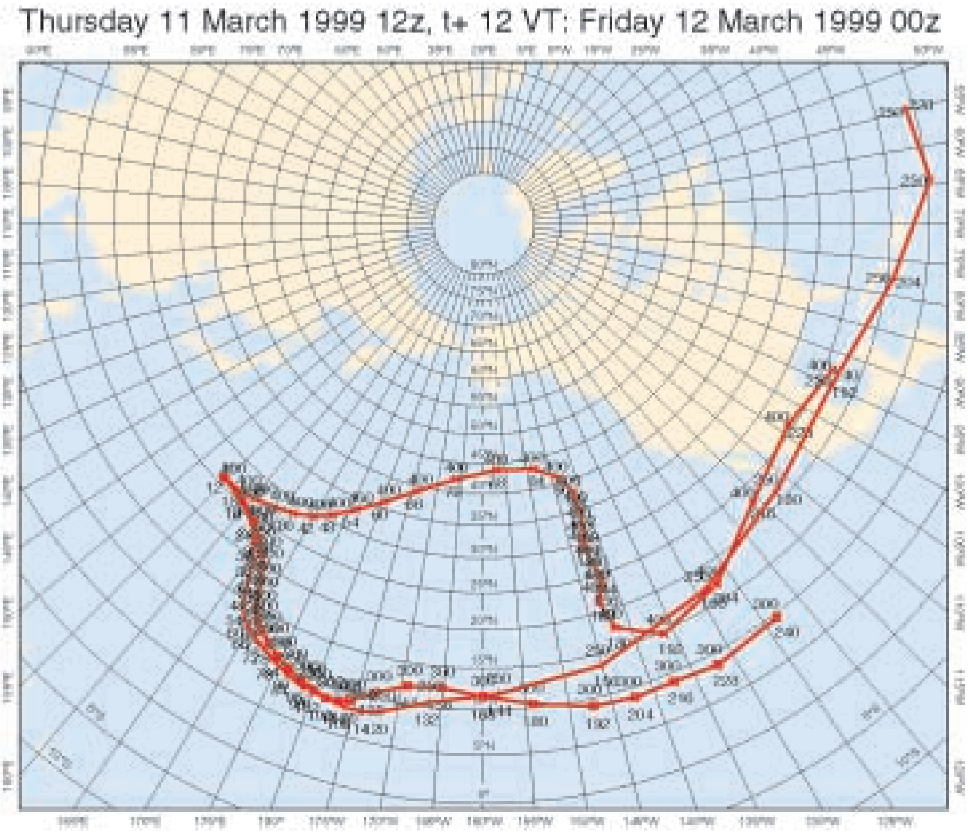


Figure 6: Trajectories at 400 hPa (7,000m), 300 hPa (9,000m) and 250 hPa (11,000m) from a point south-east of Japan. We can see how critical the flying height was. A trajectory at 7,000m would have led the balloon track into the undulating polar jet stream; the higher flight level that was followed led to a track within the subtropical jet stream that had just formed south of Hawaii.

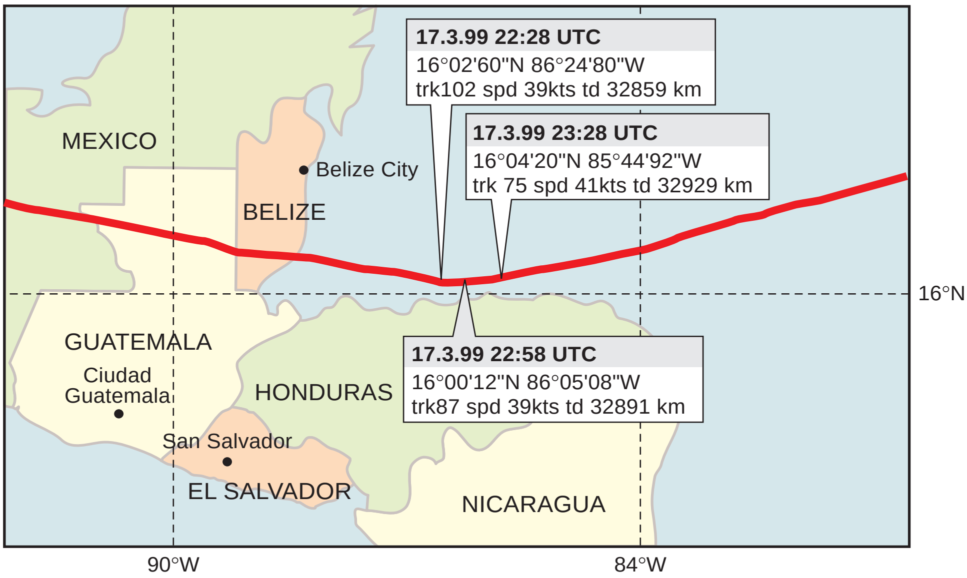


Figure 7: The situation over the coast of Honduras. Over a period of one hour, between 22:28 UTC and 23:28 UTC on 17 March, the balloon went 300 m higher and the track changed from 102° (towards the direction of Venezuela) to 75° (towards the more favourable direction of north Africa).

This expedition was, above all, and extraordinary adventure. On a scientific level, the numerical models showed their capacity to forecast trajectories for relatively long periods, even though the trajectory calculation tends to enhance the uncertainties of the forecast. But the human aspect was also essential – a team spirit and a mutual understanding between the constructors of the balloon, the pilots and the meteorologists became established from the start. The media impact of this expedition has been world-wide; we can safely claim that on this occasion the science of meteorology has been shown in an extremely positive light.

Pierre Eckert (Institut Suisse de Météorologie), e-mail: pek@sma.ch,
Luc Trullemans (Institut Royal Météorologique de Belgique), e-mail: luc.trullemans@skynet.be

A new version of XCDP

SMS is an application that was developed at ECMWF to schedule, monitor and supervise the tasks of the operational suite. SMS is controlled by a program called CDP (command display program). When ECMWF acquired its first workstations, a graphical version of CDP was developed: XCDP.

Background

Seven years ago, when the first XCDP was designed, the ECMWF environment and working methods were quite different from now. There were two SMSs, one running on the super-computer, a Cray, and one running on the acquisition and dissemination machines, a VAX cluster. This architecture was inherited from previous versions of the supervisor. There were, at the time, only four suites: an operational suite (o), a test suite (e), a dissemination suite and an acquisition/pre-processing suite. Furthermore, the only people using the system were the operators, to monitor the suites, and the Meteorological Applications Section analysts for troubleshooting and maintenance of the suites.

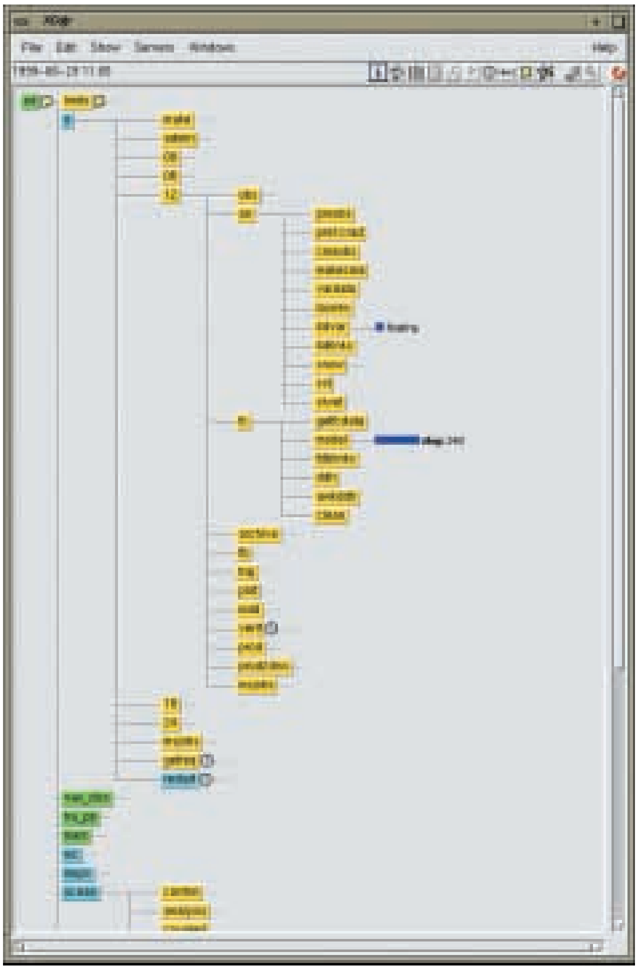


Figure 1: XCDP main window showing the status of the operational suite.

Later, the reanalysis project was launched, and the reanalysis group set up a large and complex suite to run their project. The number of tasks they were monitoring (more that 8000) was too large for XCDP; the application became too slow to be useful, and most of the display routines had to be rewritten.

Then SMS was moved to a workstation to relieve the super computer. In the meantime, the Research Department started using SMS and XCDP in order to run and monitor experiments, as part of the PREPAN, PREPEX and PREPIFS applications. There was an explosion in the number of suites, and every researcher began using XCDP to follow her/his experiments. The operators could then monitor research experiments and operational suites in the same way, and could optimise the usage of the super-computer.

Although SMS and XCDP scaled well to accommodate this new usage, the workstation hosting SMS was running out of resources in trying to serve more than fifty users simultaneously.

The natural way to overcome this problem was to spread the load over several machines, running an SMS on each of them and managing a few suites for a few users. Nevertheless, this idea conflicted with the fact that the operators had to monitor the activity of research experiments. The only way they could monitor several SMSs at once was to run the same number of XCDPs on their workstation, but that was not practical.

Therefore, we decided to modify XCDP so it could monitor several SMSs in a single window. Unfortunately, the original design was such that is was nearly impossible to do so without introducing numerous bugs.

During this time, SMS itself had been enhanced (meters, labels, repeats . . .) and had had a lot of new functionality added to it, and making the corresponding changes in XCDP was more and more difficult.

We then decided that it was time for a complete rewrite of the application, in an open way, so new features could be easily added. At the same time, most of the user wishes could be incorporated.

Design

The main requirements for the new XCDP were:

- Remove any restriction that made SMS improvements difficult to implement in XCDP.
- Design an interface that caters for three kinds of users: operators, analysts and researchers. These three sorts of users have different requirements and different working habits.
- The amount of information to be displayed is so great that designing a clear graphical interface is difficult. The solution to this problem is to design a very consistent interface.

To achieve these goals, XCDP was rewritten in C++, using an object-oriented design:

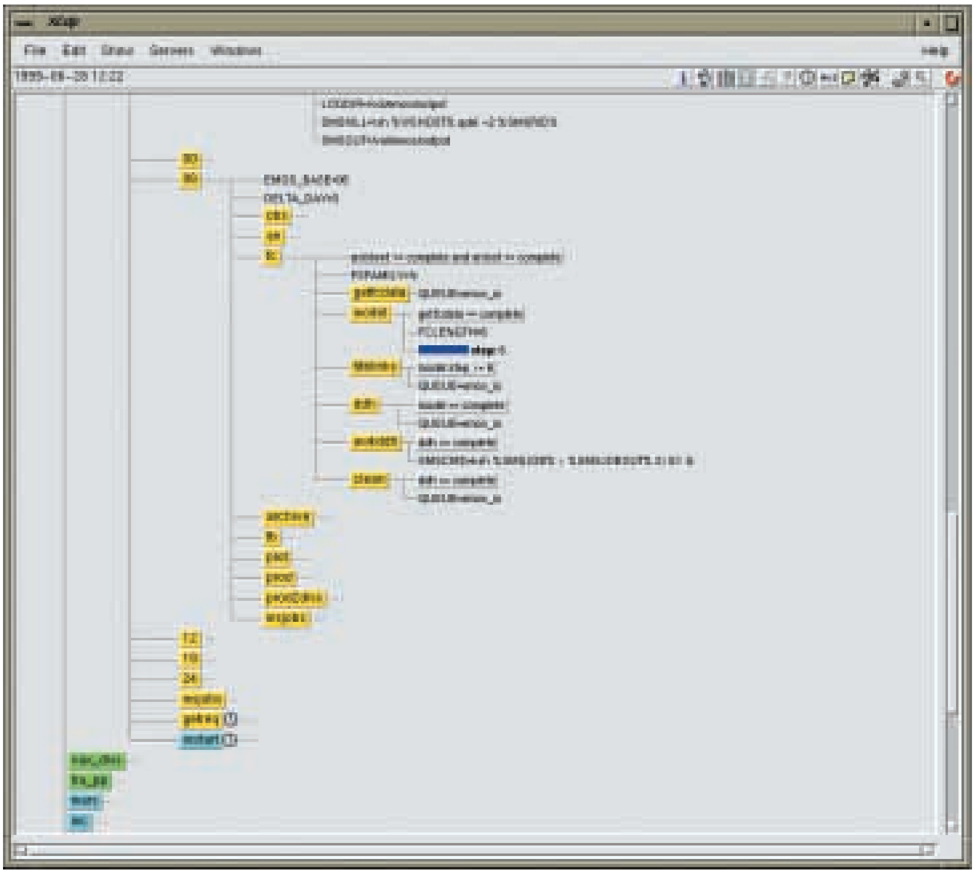


Figure 2: Viewing variables and triggers on the tree window.

- ◆ Every piece of information is now handled in a similar fashion. Variables, triggers, and any other attributes are also visible on the user interface, in addition to suites, families and tasks. Adding a new kind of information simply means designing a new class.
- ◆ Each piece of information can be viewed and edited using various windows. For example, given a particular task, the user can choose to view its output, its script, as well as its dependencies on other tasks. Again, designing a new class adds a new view to a piece of information.

Overview of the new features

Configuration

We were often asked to give the user more control over the look of the user interface. The way X11 handles fonts and colours is usually too complex for normal users, so we decided not to use the X11 resources to control the XCDP look and feel, but allow people to change their setting from within XCDP.

Visibility

As before, XCDP displays the hierarchy suite/family/task in a tree fashion, using colour coding to reflect the status of each node (active is green, complete is yellow, aborted is red, and so on . . .).

In the new XCDP, every attribute can now be shown in the tree window, if required. The user can then view or edit those attributes, as they would do for the suites, families or tasks. This includes variables, time and date dependencies, triggers, labels, meters, limits and much more.

The user can configure what is visible or hidden. He or she can also perform actions on nodes, such as rerunning jobs or changing variables. By selecting nodes, the user can get extended information concerning their selection.

Multiple windows

We had two conflicting requirements: not cluttering the screen with windows, and viewing more than one aspect of the suite (e.g. several outputs) at once.

In the previous XCDP, we had a different window for a task script, its output, its trigger dependencies, and so on. The user would select a node and the windows would change their content accordingly. It was not possible to view two outputs simultaneously.

In the new XCDP, the windows are no longer specialised, but can display any kind of information. The user can select a node and open a window in which he or she can switch between script, output, triggers and more. Having only one window opened not only helps to reduce screen clutter, but will greatly improves the response of XCDP when used over a slow link such as a PPP connection. As before, selecting another node updates the content of the window. By opening several windows displaying a different content, the user can mimic the behaviour of the previous XCDP.

On the other hand, the user can also open a window that is tied to a specific node and not following the selection. This new facility enables the user to view information from two different nodes at once, such as two outputs.

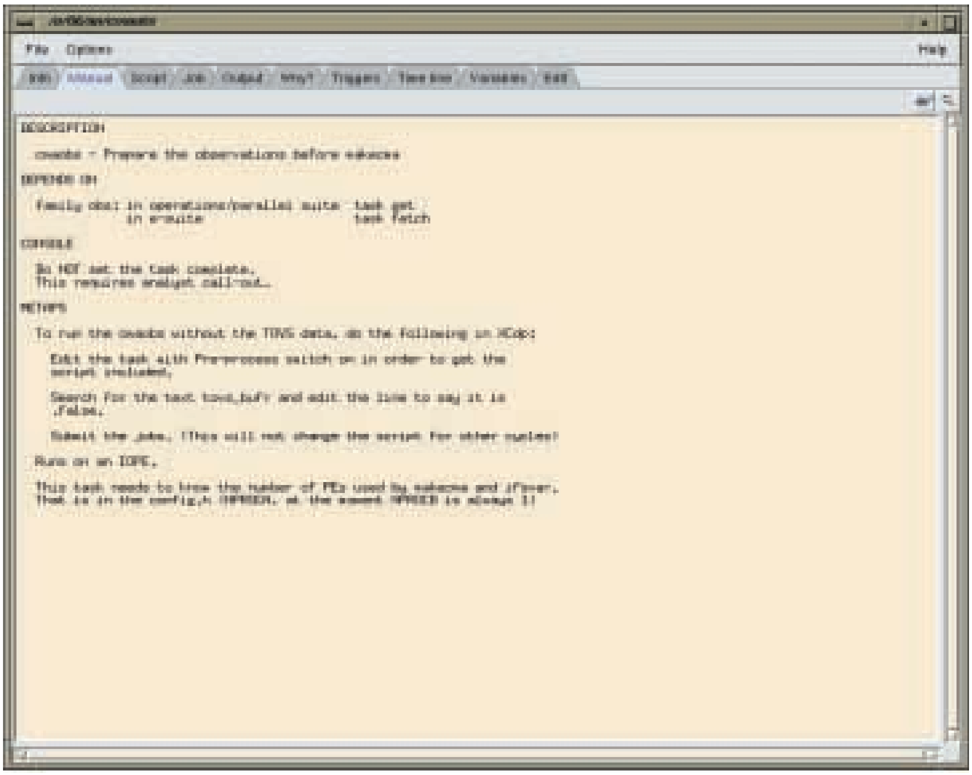


Figure 3: Displaying the manual page of one task. The user can change the content of this window by clicking on one of the tabs.

Multiple hosts

This was the original requirement that triggered the redesign of XCDP. It is possible with the new XCDP to monitor several SMSs at once. This helps operators to supervise suites from different sources. We can now spread the load of the SMSs over several hosts.

Navigation

XCDP uses several techniques to display the information available: graphs (tree, triggers), lists (limits, histories) and hypertexts (dependencies). Each time a node is referenced in one of these windows, the user can click on the reference to get some information about the node.

Commands

Menu

To perform any actions on a node, such as suspending a task or editing a label, the user uses a menu whose content varies according to the node under the mouse pointer. This menu is built from a file that describes what commands can be performed on what node. Through this menu, the user has access to the complete range of available CDP commands.

Collector

A very common wish from XCDP users was the ability to issue a command on several nodes, such as resuming all suspended tasks at once. This was never implemented because this is one of those things that is more easily done with a command-line interface such as CDP than with a graphical user interface such as XCDP. In CDP, you can use wildcards to match several nodes.

The solution we have chosen in the new XCDP is to collect a set of nodes in a “collector” window. The user can collect the result of the find command, or select individual nodes in the user interface, then type in a CDP command that will be executed for each node of the selection. As in the command menu, the user has access to the complete range of available CDP commands.

Searching

Because of the sheer number of nodes that make the suites we are monitoring, XCDP provides a powerful way of searching nodes. Not only can the user search nodes by name, but also she or he can search the variables and the triggers for occurrences of a given string. It is now possible to find what jobs run on a particular machine. The search facility is linked to the node collector so the user can perform an action on the result of a search.

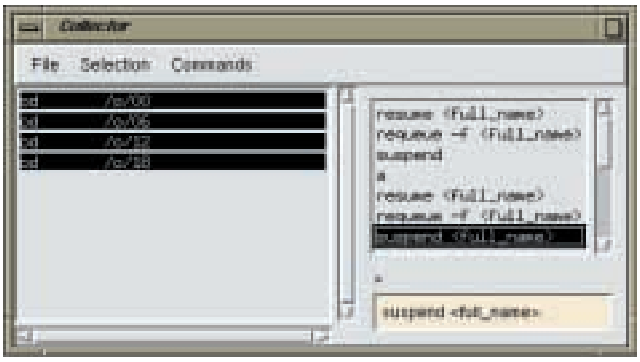
Dependencies

By defining triggers between task, families and suites, the analyst can create a complex graph of dependencies. XCDP tries to display this information in three different ways:

Figure 4: The user can suspend or rerun the analysis using a popup menu.

Why?

When a job has not been scheduled by SMS, it could be for various reasons. XCDP tries to work out the reasons and answer the question "Why is this task not running?" To do so, XCDP will scan the triggers, the date and time dependencies, the limits and the status of various nodes.

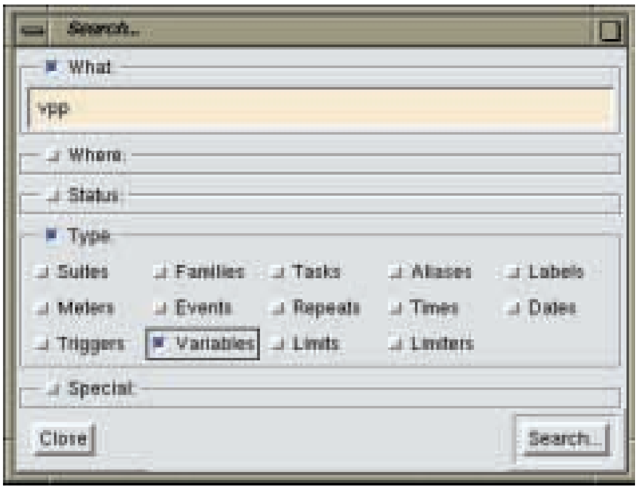


```
od    /o/00
od    /o/06
od    /o/12
od    /o/18
```

```
resume <full_name>
requeue -f <full_name>
suspend
#
resume <full_name>
requeue -f <full_name>
suspend <full_name>
```

```
suspend <full_name>
```

Figure 5: A command can be performed on several nodes at once using the collector.



☒ What

vpp

☐ Where

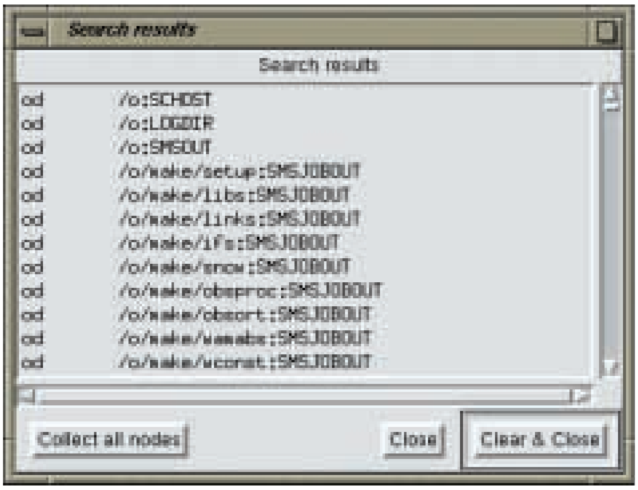
☐ Status

☒ Type

☐ Suites ☐ Families ☐ Tasks ☐ Aliases ☐ Labels
☐ Meters ☐ Events ☐ Repeats ☐ Times ☐ Dates
☐ Triggers ☒ Variables ☐ Limits ☐ Limiters

☐ Special

Figure 6: Searching for all variables containing "vpp"



```
od         /o:SCHOST
od         /o:LOGDIR
od         /o:SMSOUT
od         /o/wake/setup:SMSJOBOUT
od         /o/wake/libs:SMSJOBOUT
od         /o/wake/links:SMSJOBOUT
od         /o/wake/ifs:SMSJOBOUT
od         /o/wake/snow:SMSJOBOUT
od         /o/wake/obsproc:SMSJOBOUT
od         /o/wake/obsort:SMSJOBOUT
od         /o/wake/wamabe:SMSJOBOUT
od         /o/wake/wconst:SMSJOBOUT
```

Figure 7: The results of the search.

Triggers

XCDP can display triggers, as they were defined by the analyst, in a hypertext window or in a graph window. The user can browse through the graphs to follow the triggers. Triggers can also be displayed directly in the tree.

Dependencies

A lot of effort has gone into the analysis of the dependencies between nodes, as well as date and time dependencies. XCDP can work out hidden dependencies, such as dependencies between two families via some of their tasks. Figure 10 shows the dependencies of the operational suite. On the right-hand side are the nodes waiting for some data from the operational suite, such as the ensemble forecast (mc) and the limited-area wave model (law). On the left-hand side are the nodes the operational suite is waiting for, in this case the four extractions of the pre-processing (ha_pp), as well as some time dependencies and limits.

Time-line

By reading SMS log files, XCDP can build a time map of the various events recorded in the log files (starting jobs, completing jobs, distance between two events...). These events can be sorted in various ways. For each task, XCDP can give a summary of its activity. This is a brand new feature of XCDP, and we are still investigating how it can be improved. It has already proved to be very useful in spotting some delays in the archive jobs due to poor data transfer rates. It has also been used to organise the frequency of acquisitions jobs within an hour. Interesting queries can be answered, such as "what jobs are running at the same time as the forecast?". In conjunction with the dependency information, this will help to improve the design of the operational suite.

New SMS features

Various features have been added to SMS. These are the main ones, and how they are presented to the user by XCDP:

Limits

After years of using SMS, we discovered that suite designers were defining two kinds of task dependencies: "data dependencies" and "courtesy dependencies". The former are genuine dependencies; one task needs the data produced by another one, and cannot run until this data has been produced. The latter are dependencies that are defined to prevent the SMS from flooding the systems with jobs. Good examples are the plot jobs. Each cycle of the operational suite will produce tens of plots. Although those jobs are inherently independent, they are actually chained using triggers, one job starting when the previous one has finished. It is sometimes very important for an analyst to be able to distinguish between those two sorts of dependencies. So a new concept was introduced in SMS: the "limit". A limit is an SMS object that keeps track of the state of certain tasks, preventing new tasks from being scheduled once a threshold has been reached.

Repeats

It is possible to ask SMS to schedule a group of tasks a given number of times, looping over a set of strings, dates or numbers. This facility is used by the research experiments to run their assimilations and forecasts over a range of dates. The user can change this range to extend or shorten an experiment using XCDP.

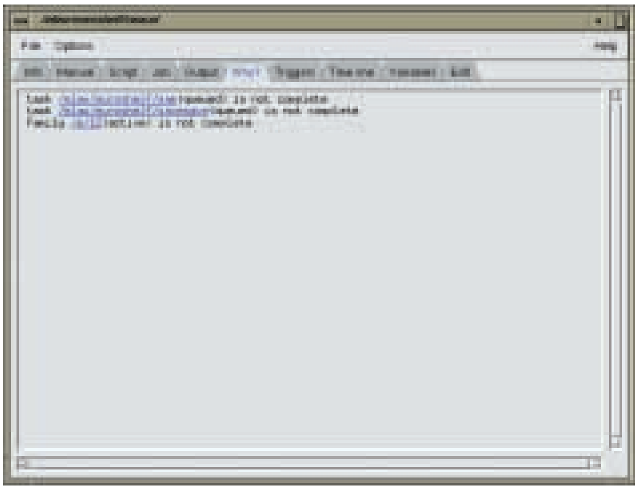


Figure 8: Why is this task not running? The user can click on the hyperlinks to get more information.

Aliases

A very useful addition is the ability to clone a task, perform some minor modifications to its script, and run it as an "alias". Such a task can run under the control of the SMS but has no impact on the activity of the suite itself. This feature is used to rerun a task on a previous date, or to solve transient problems such as full file systems. In this case, the task can be run using a different disk, in order to guarantee the completion of the suite on schedule. The analyst then has time to understand why this condition arises and take the necessary actions.

Moving suites between SMSs

Some users schedule their experiments from an SMS running on their desktop workstation. Doing that, they gain in performance but they cannot have their suite monitored by the operators. When those users go on holiday, they simply transfer their suites from their SMS to one that is monitored. An SMS can "adopt" a suite that started on another SMS, even if the suite is still active. The user can perform this operation from XCDP in a few clicks.

Chat

When there are some problems with the operational suites, the operators are alerted by XCDP. If they cannot solve the problem, they call the on-call analyst. He usually logs in to the Centre using the dialback facility. Once connected, he has no means of communicating with the operators, as his telephone line is busy. It is now possible for the analyst to use XCDP to chat, via SMS, to the operator.

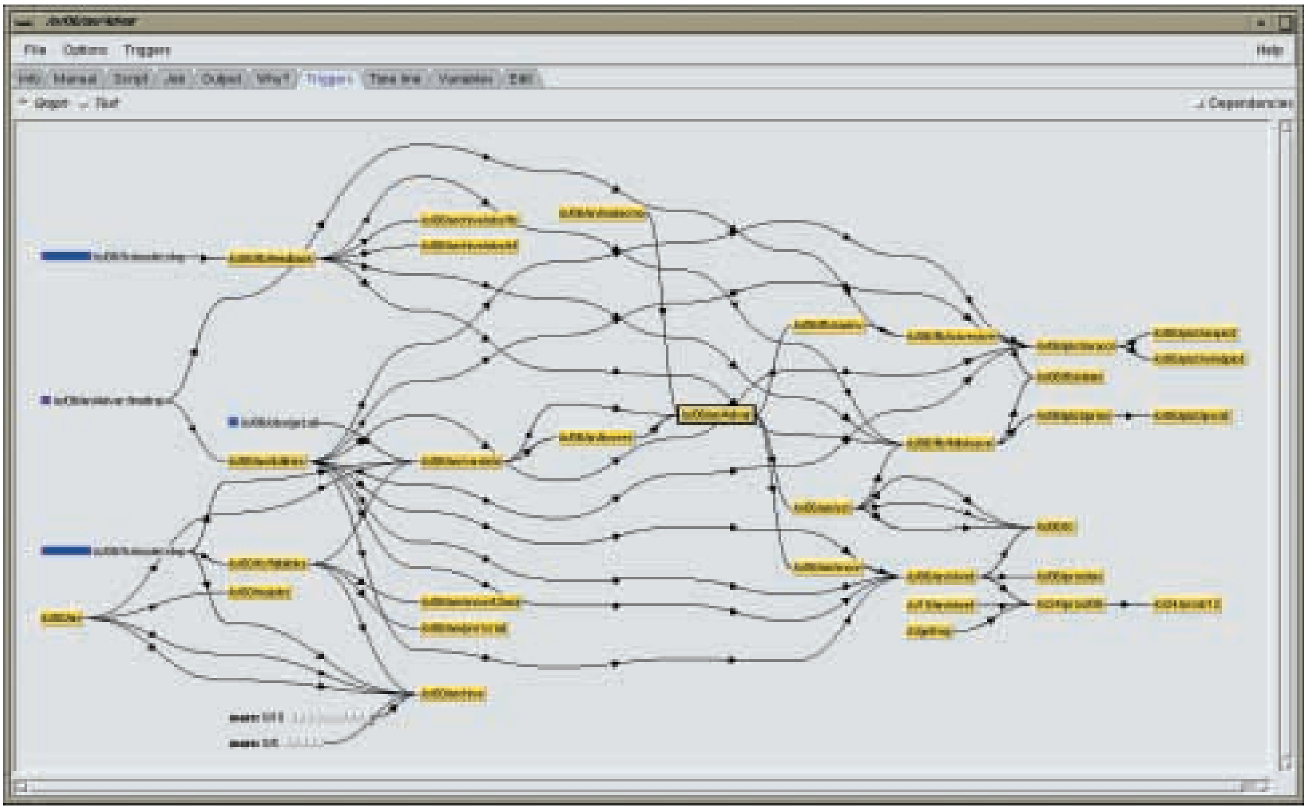


Figure 9: Triggers dependencies can be very complex. The selected node is the 12Z assimilation.

This facility can also be called by a task, to inform the operators of some critical conditions, or to request a human action.

Zombies

Sometimes, after some network problems or if a suite is cancelled while active, some jobs lose contact with the SMS that originally scheduled them. They become orphans (or "zombies"), XCDP provides a way to handle these jobs by either cancelling them or having an SMS adopt them.

Texts

Work has been done on the way XCDP displays text files. XCDP can display the scripts, jobs and outputs of tasks. It will now try to use NFS to load these files when possible, relieving SMS of a lot of work. If the files are not accessible (i.e. on another machine) XCDP will still receive them from SMS. XCDP has also been modified to be able to display very large files (several megabytes) by memory-mapping the files.

The user can also choose to view the files using other applications, such as the Unix "more" command. Some users are more efficient when using tools they have mastered.

Searching text strings has been enhanced to support POSIX regular expressions.

Editing

Most of the objects visible in XCDP are editable (apart from triggers). Labels, limits, variables, meters, repeats can all be changed by the user. As before, the user can edit the script and the variables of a script. Nevertheless, those changes are only valid for one run of the task, changes that are more permanent must be done directly on the script files.

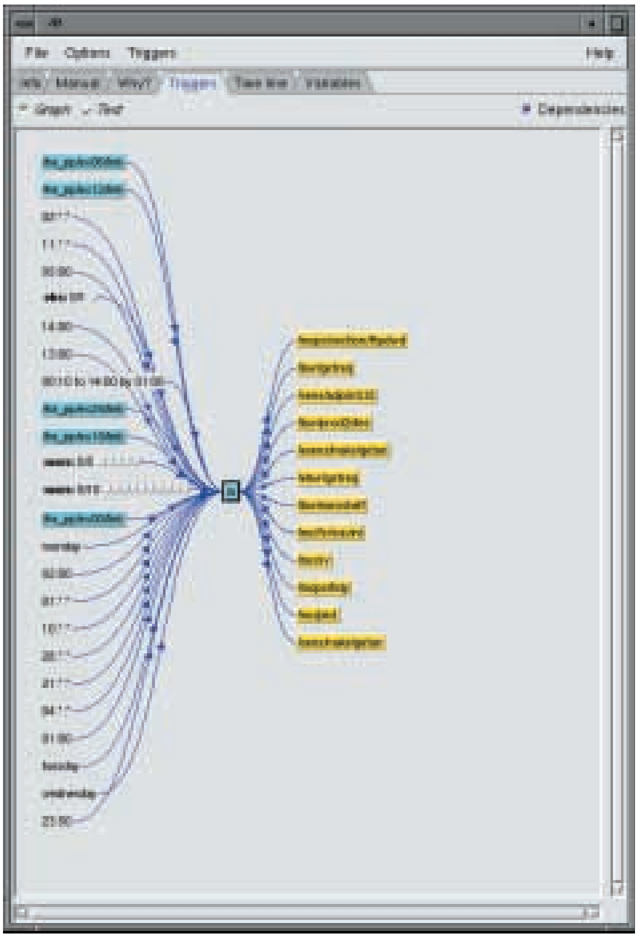


Figure 10: Dependencies of the operational suite.

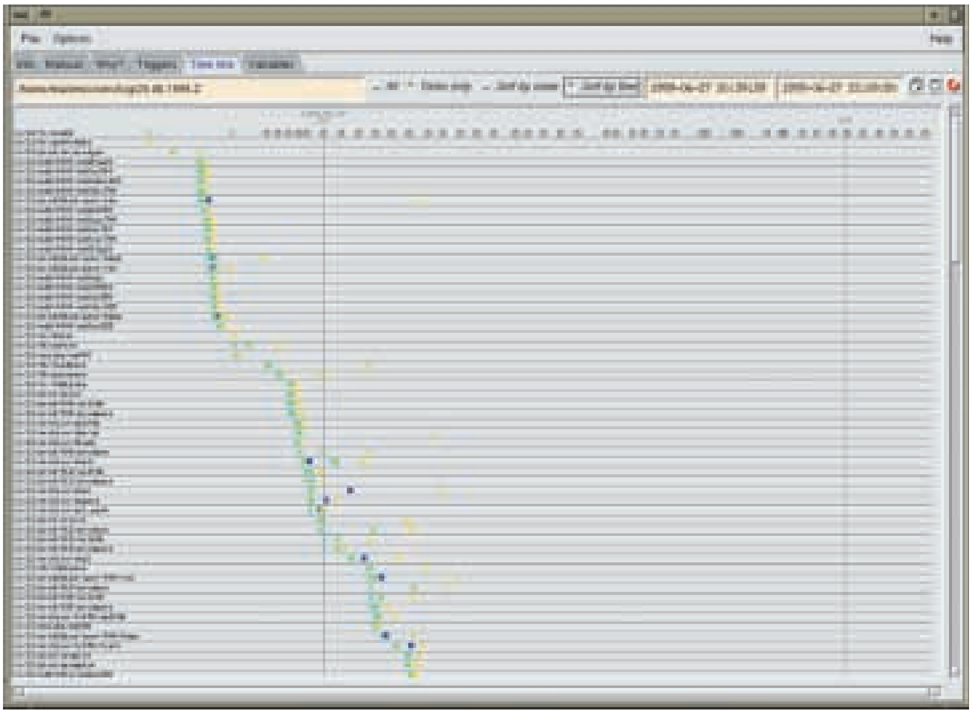
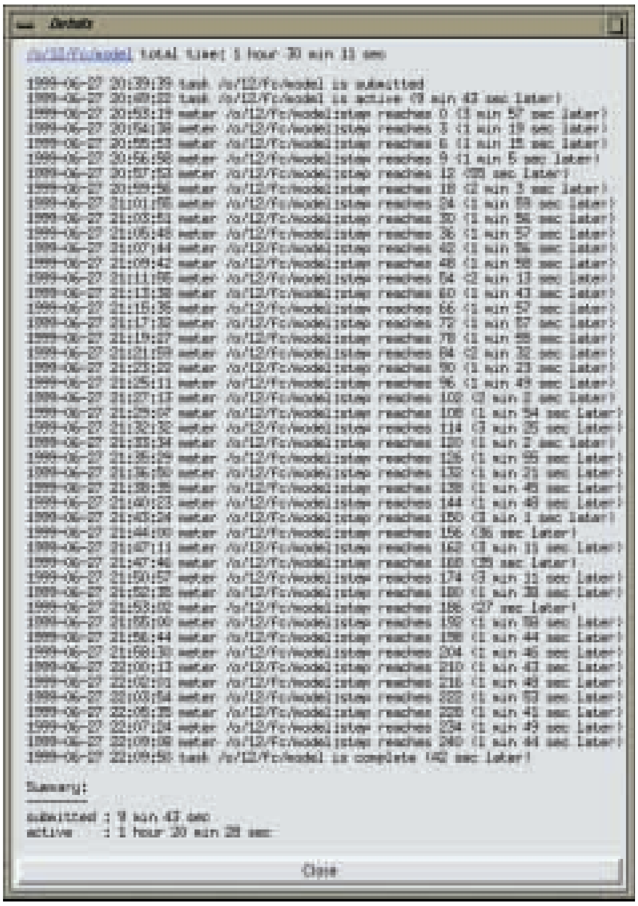


Figure 11: Time line centred on the run of the 12Z forecast. Each circle is a write-up step.



```
/o/12/fc/model total time: 1 hour 30 min 11 sec

1999-06-27 20:39:39 task /o/12/fc/model is submitted
1999-06-27 20:49:22 task /o/12/fc/model is active (9 min 43 sec later)
1999-06-27 20:53:19 meter /o/12/fc/model:step reaches 0 (3 min 57 sec later)
1999-06-27 20:54:38 meter /o/12/fc/model:step reaches 3 (1 min 19 sec later)
1999-06-27 20:55:53 meter /o/12/fc/model:step reaches 6 (1 min 15 sec later)
1999-06-27 20:56:58 meter /o/12/fc/model:step reaches 9 (1 min 5 sec later)
1999-06-27 20:57:53 meter /o/12/fc/model:step reaches 12 (55 sec later)
1999-06-27 20:59:56 meter /o/12/fc/model:step reaches 18 (2 min 3 sec later)
1999-06-27 21:01:55 meter /o/12/fc/model:step reaches 24 (1 min 59 sec later)
1999-06-27 21:03:51 meter /o/12/fc/model:step reaches 30 (1 min 56 sec later)
1999-06-27 21:05:48 meter /o/12/fc/model:step reaches 36 (1 min 57 sec later)
1999-06-27 21:07:44 meter /o/12/fc/model:step reaches 42 (1 min 56 sec later)
1999-06-27 21:09:42 meter /o/12/fc/model:step reaches 48 (1 min 58 sec later)
1999-06-27 21:11:55 meter /o/12/fc/model:step reaches 54 (2 min 13 sec later)
1999-06-27 21:13:38 meter /o/12/fc/model:step reaches 60 (1 min 43 sec later)
1999-06-27 21:15:35 meter /o/12/fc/model:step reaches 66 (1 min 57 sec later)
1999-06-27 21:17:32 meter /o/12/fc/model:step reaches 72 (1 min 57 sec later)
1999-06-27 21:19:27 meter /o/12/fc/model:step reaches 78 (1 min 55 sec later)
1999-06-27 21:21:59 meter /o/12/fc/model:step reaches 84 (2 min 32 sec later)
1999-06-27 21:23:22 meter /o/12/fc/model:step reaches 90 (1 min 23 sec later)
1999-06-27 21:25:11 meter /o/12/fc/model:step reaches 96 (1 min 49 sec later)
1999-06-27 21:27:13 meter /o/12/fc/model:step reaches 102 (2 min 2 sec later)
1999-06-27 21:29:07 meter /o/12/fc/model:step reaches 108 (1 min 54 sec later)
1999-06-27 21:32:32 meter /o/12/fc/model:step reaches 114 (3 min 25 sec later)
1999-06-27 21:33:34 meter /o/12/fc/model:step reaches 120 (1 min 2 sec later)
1999-06-27 21:35:29 meter /o/12/fc/model:step reaches 126 (1 min 55 sec later)
1999-06-27 21:36:50 meter /o/12/fc/model:step reaches 132 (1 min 21 sec later)
1999-06-27 21:38:35 meter /o/12/fc/model:step reaches 138 (1 min 45 sec later)
1999-06-27 21:40:23 meter /o/12/fc/model:step reaches 144 (1 min 48 sec later)
1999-06-27 21:43:24 meter /o/12/fc/model:step reaches 150 (3 min 1 sec later)
1999-06-27 21:44:00 meter /o/12/fc/model:step reaches 156 (36 sec later)
1999-06-27 21:47:11 meter /o/12/fc/model:step reaches 162 (3 min 11 sec later)
1999-06-27 21:47:46 meter /o/12/fc/model:step reaches 168 (35 sec later)
1999-06-27 21:50:57 meter /o/12/fc/model:step reaches 174 (3 min 11 sec later)
1999-06-27 21:52:35 meter /o/12/fc/model:step reaches 180 (1 min 38 sec later)
1999-06-27 21:53:02 meter /o/12/fc/model:step reaches 186 (27 sec later)
1999-06-27 21:55:00 meter /o/12/fc/model:step reaches 192 (1 min 58 sec later)
1999-06-27 21:56:44 meter /o/12/fc/model:step reaches 198 (1 min 44 sec later)
1999-06-27 21:58:30 meter /o/12/fc/model:step reaches 204 (1 min 46 sec later)
1999-06-27 22:00:13 meter /o/12/fc/model:step reaches 210 (1 min 43 sec later)
1999-06-27 22:02:01 meter /o/12/fc/model:step reaches 216 (1 min 48 sec later)
1999-06-27 22:03:54 meter /o/12/fc/model:step reaches 222 (1 min 53 sec later)
1999-06-27 22:05:35 meter /o/12/fc/model:step reaches 228 (1 min 41 sec later)
1999-06-27 22:07:24 meter /o/12/fc/model:step reaches 234 (1 min 49 sec later)
1999-06-27 22:09:08 meter /o/12/fc/model:step reaches 240 (1 min 44 sec later)
1999-06-27 22:09:50 task /o/12/fc/model is complete (42 sec later)

Summary:

submitted : 9 min 43 sec
active    : 1 hour 20 min 28 sec
```

Figure 12: Summary for the 12Z forecast.

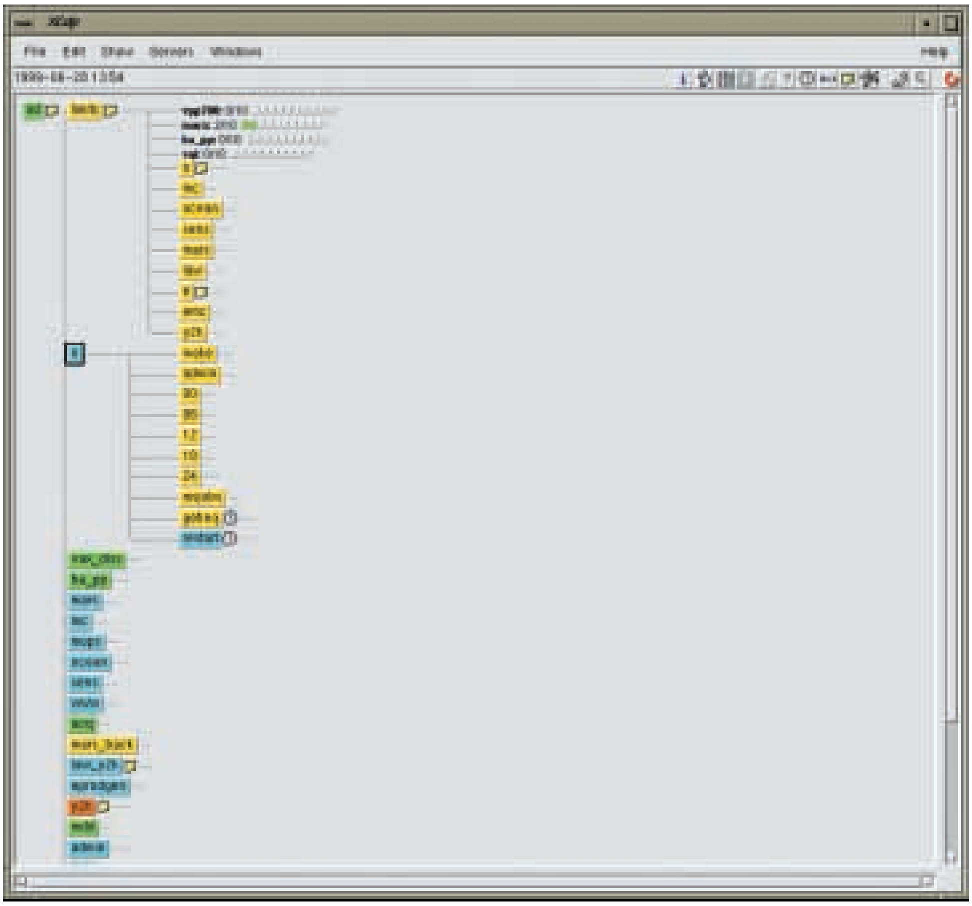


Figure 13: We try to limit the access to MARS (the ECMWF archive system) to ten concurrent jobs. Currently two tasks are using MARS.

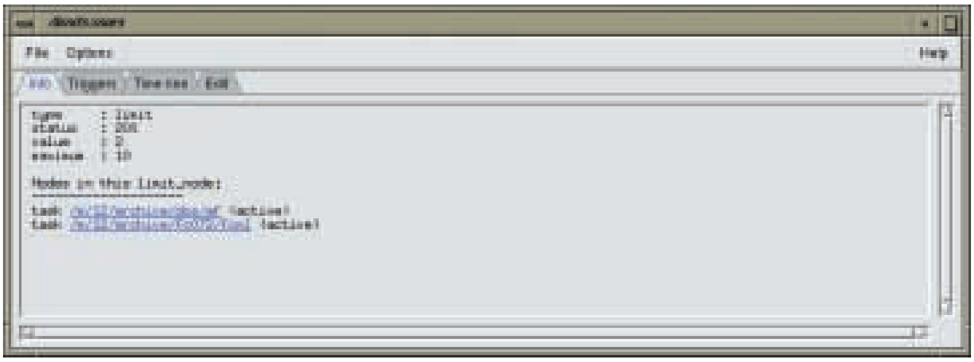


Figure 14: The information window for the limit "mars" will show which ones.

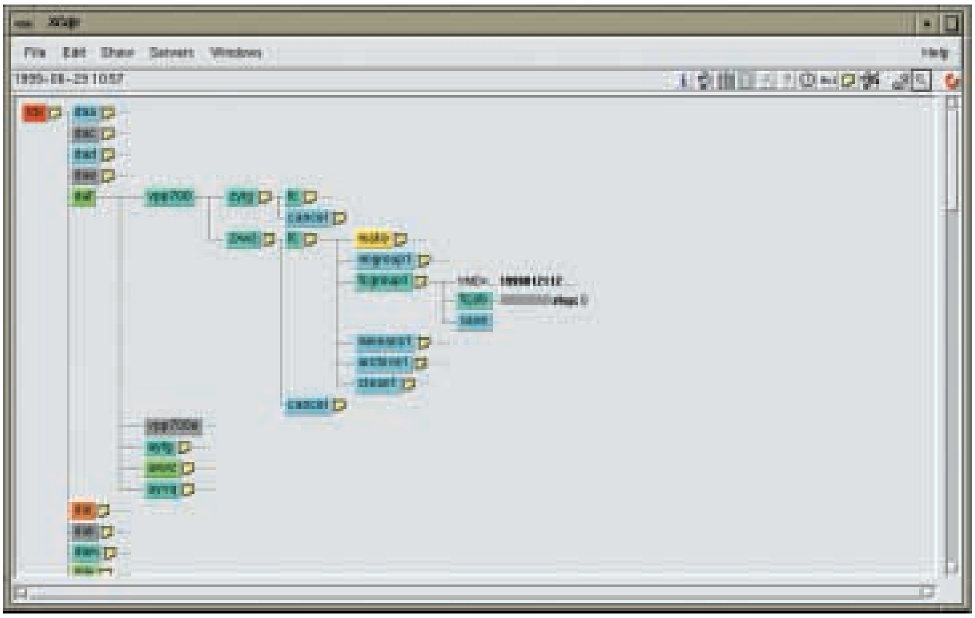


Figure 15: Repeats are heavily used by the research experiments. Experiment "zxwz" of user "daf" is currently processing the 21st of January 1999 at 12Z.

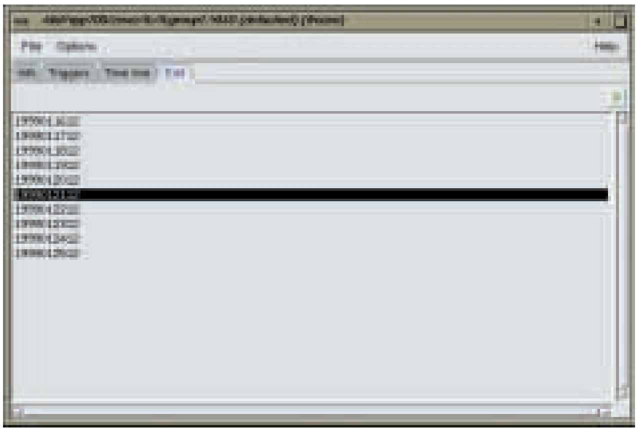


Figure 16: The user may alter this value using XCDP.

As for the output, users can invoke their favourite editor instead of the one provided by XCDP.

Future developments

There are still a few ideas we would like to add to XCDP. Some of them are actually being worked on, while others may simply remain ideas. Here are some of the developments we would like to do, without any specific order:

Scripting of XCDP

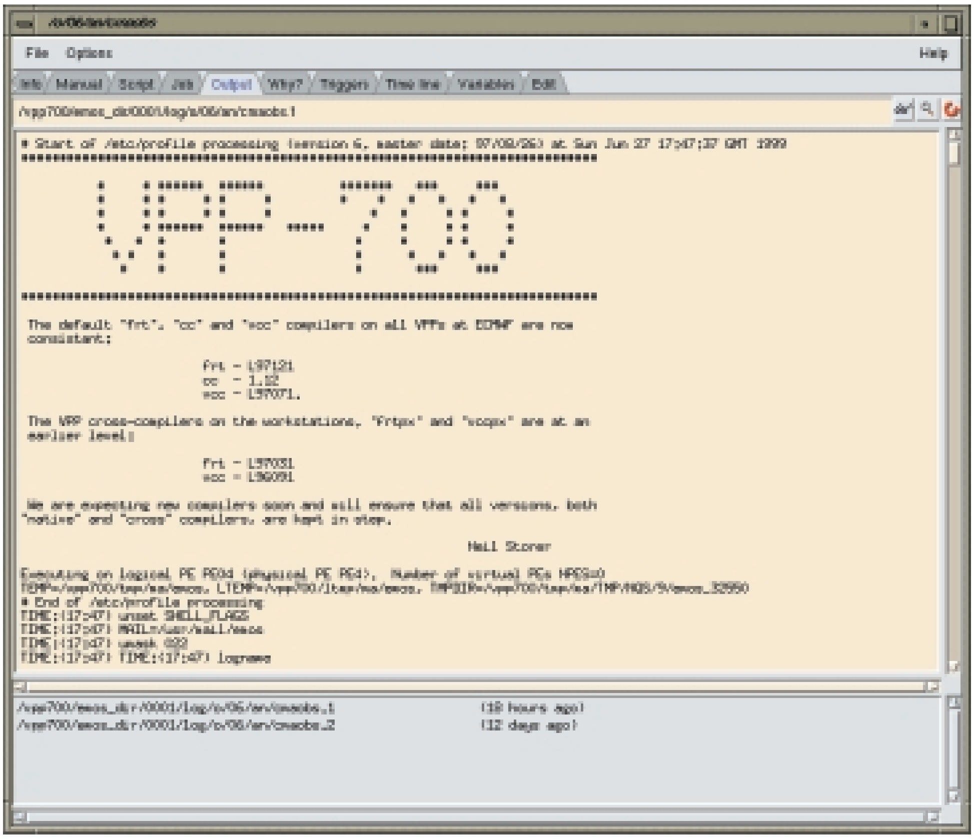
XCDP, being based on CDP, has access to a powerful scripting language. One idea would be to provide the ability to drive XCDP using this language. The user could open windows, search the outputs and perform common tasks automatically by running user-defined scripts.

Critical path analysis

XCDP already performs some sort of graph analysis to display the dependencies between various tasks. One very useful feature would be to find the critical path of a suite (Gantt or PERT methods). With this information, modification in the suite architecture could be done to make the maximum use of the Centre's resources, such as its super-computer.

HTML documentation

Nowadays, the natural way to edit and distribute documentation is the HTML format. Work is being done to provide SMS and XCDP documentation in this format. By linking XCDP and a Web browser, the user could have access to contextual help.



```
/vpp700/emos_dir/0001/log/o/06/an/cmaobs.1

# Start of /etc/profile processing (version 6, master date: 97/09/26) at Sun Jun 27 17:47:37 GMT 1999

 The default "frt", "cc" and "vcc" compilers on all VPPs at ECMWF are now
 consistent:

                        frt - L97121
                        cc  - 1.12
                        vcc - L97071.

 The VPP cross-compilers on the workstations, "frtpx" and "vccpx" are at an
 earlier level:

                        frt - L97031
                        vcc = L96091

 We are expecting new compilers soon and will ensure that all versions, both
"native" and "cross" compilers, are kept in step.
                                                    Neil Storer

Executing on logical PE PE04 (physical PE PE4),  Number of virtual PEs NPES=0
# End of /etc/profile processing
TIME:(17:47) unset SHELL_FLAGS
TIME:(17:47) MAIL=/usr/mail/emos
TIME:(17:47) umask 022
TIME:(17:47) TIME:(17:47) logname

/vpp700/emos_dir/0001/log/o/06/an/cmaobs.1                (10 hours ago)
/vpp700/emos_dir/0001/log/o/06/an/cmaobs.2                (12 days ago)
```

Figure 17: The output of an analysis job. The previous outputs are listed at the bottom of the window.

Web-based XCDP

ECMWF will provide, through a Web-based version of PREPIFS, a way to submit experiments to the Centre's super-computer. In order for the users to monitor these experiments, a Web version of XCDP is being developed. It is a lot simpler that the X/Motif XCDP, but will allow remote users to monitor their experiments using a standard Web browser. This application will also be useful for ECWMF analysts to solve problems remotely.

Defining and editing of suites

It is not possible to define or edit suites from XCDP. This was a design decision, as command line interfaces are more powerful than graphical user interfaces when it comes to processing complex suite definitions. The user has to use CDP to define suites, and XCDP would only provide a monitoring and a troubleshooting tool. With the new design of XCDP, it is possible to conceive a suite editor.

Conclusion

The new XCDP has been tested successfully on SGI, SUN and HP platforms. All the changes were well received by the users. Nowadays, there are more than 25 SMSs, running over twenty thousand tasks daily, for all sorts of jobs and on all sorts of machines. Apart from the operational suites, SMS is also running to produce statistics, migrate files, generate web pages, test software changes, handle data services orders and much more.

SMS and XCDP are one of the most used applications at ECMWF.

Baudouin Raoult, Otto Pesonen

Increased computing power at ECMWF

At the Council meeting in December 1998 the decision was taken to extend the Service Contract Agreement with Fujitsu for a further 2 years, to December 2002. This extension provides for an increase in computer power this year and a further increase in the year 2000. The contract amendment calls for the installation of a VPP5000 system ready for acceptance on 1st December 1999, followed by an upgrade to this system by 1st August 2000. The exact number of Processing Elements (PEs) has yet to be formally decided, as the numbers of PEs are relative to the power of the processors. However, it is safe to say that the initial VPP5000 system will comprise 38 PEs, while the upgraded system next year will contain about 100 PEs.

Configurations

The VPP5000 is the follow-on system from the VPP700 and VPP700E. Table 1 compares these systems.

The acquisition of the VPP5000/38 this year will almost double the peak performance of the VPP systems at ECMWF, to over 0.75 TFLOPS. Next year the total peak performance of the VPP systems will be in excess of 1.3 TFLOPS.

Details

Like the VPP700, the VPP5000 is a Vector Parallel Processing (VPP) system consisting of identical Processing Elements, or PEs. Each PE (including the IOPEs and P-PE) is capable of executing floating-point operations at the rate of 2.2 GFLOPS in the case of the VPP700 and 9.6 GFLOPS for the VPP5000. The higher performance of the VPP5000 comes from a decrease in the cycle time over the VPP700, from 7 to 3.33 ($^{10}/_{3}$) nano-seconds, and from a doubling of the pipes, i.e. the number of results that can be produced concurrently in 1 CPU cycle. For a 'Multiply & Add' instruction the VPP700 can produce 16 results per cycle, while the VPP5000 can produce 32. Hence the theoretical speed-up of the VPP5000 over the VPP700 is:

$$(7 \times (3 \div 10)) \times (32 \div 16) = 4.2$$

Extra features of the VPP5000 over the VPP700 mean that for many single–PE codes the speed-up may be better than the extra pipes and faster clock alone would indicate. In this article we will try to explain those extra features and show that the VPP5000 is not just a 'faster VPP700', but a machine designed to remedy shortcomings of the VPP700.

Scalar Processor Enhancements

It is an acknowledged fact that the 'weak link' of the VPP700 PE is its Scalar Processor. The main cause of this is the small size of the primary cache, and the lack of a secondary cache. Initial testing at ECMWF of the VPP showed that compilation speeds and TCP/IP activity were especially hit by this, both of these being scalar in nature. Fujitsu quickly sorted out the TCP/IP problem, by making more use of the vector hardware in the Operating System code, and reducing the amount of memory copying within the TCP/IP stack. Compilation

	VPP700	**VPP700E**	**VPP5000 initial**	**VPP5000 final**	**Yr 2000 Total***
No. of PEs	116	48	38	~100	264
No. of S-PEs*	105	43	33	?	
No. of IOPEs*	10	4	4	?	
PE performance (GF)	2.2	2.4	9.6	9.6	
Peak performance/ system (GF)	255	115	365	960	1330
PE memory (GB)	2	2	4	4	
Crossbar speed (MB/s)	550 (x2)	550 (x2)	1600 (x2)	1600 (x2)	
Parallel jobs/PE*	2	2	4	4	
Disk Space (TB)	2	1	2	5	
Types of Disk	"Allegro" "Hitachi" RAID GEN-5 (HIPPI)	"Allegro" "Hitachi RAID GEN-5 (HIPPI)	"Allegro" "Hitachi" RAID NOBLE (Fbr-chnl)	"Allegro" "Hitachi" RAID NOBLE (Fbr-chnl)	
Network Interfaces	FDDI (8) HIPPI (4)	FDDI (4) HIPPI (2)	FDDI (2) HIPPI (4) Gbit-Ethernet (6)	HIPPI (?) Gbit-Ethernet (?)	

* **Yr. 2000 Total**: This takes into account the fact that the VPP5000/100 system is an **upgrade** to the 38-PE system. It is not an **additional** system.

** **S-PEs**: Secondary-PEs. These can only be used to execute programs, as their only I/O interface is the non-blocking bi-directional crossbar, which connects all PEs together.

*** **IOPEs**: These can be used to execute programs, as they contain exactly the same CPUs as the S-PEs, but they are primarily used to do I/O to directly connected disks and network devices. On each of the systems there is also a Primary-PE (P-PE). This is also an IOPE but it has special functions to perform, such as booting the system, running NQS, the batch subsystem and scheduling PE allocation etc, though it too can execute programs.

**** **Parallel jobs/PE**: The VPP5000 has an extra facility that allows a single parallel job to have up to 4 of its parallel processes on a single PE. This feature is not available on the VPP700(E).

Table 1: Detailed Comparison of ECMWF's Major VPP Systems

was much more of a problem. A ‘workaround’ solution was to install a cross-compiler that could run on an SGI system, whose (scalar) processor is more suited to this task. The scalar CPU of the VPP5000 has been considerably improved over that of the VPP700. Table 2 shows a comparison of the two.

Vector Processor Enhancements

It is in the vector hardware that much of the performance enhancement is concentrated. Table 3 compares the vector units of the VPP700 and VPP500.

Other VPP5000 improvements over the VPP700 include:

- Reduced vector register access time from 4.5 cycles to 1.5 (on average);
- Sub-array mask generation, done by hardware;
- High speed vector summation operation, which runs in ¼ the number of cycles as are taken on the VPP700;
- Vector element ‘slide’ operation that eliminates the need for some memory references;

The reduced vector start-up time, coupled with a much better design that allows more functional unit concurrency, and a better memory subsystem, means that the sustained performance of the VPP5000 is much closer to the theoretical peak, than can be achieved by the VPP700. During the Multiply-Add sequence of matrix multiply, the Multiply and Add functional units of the VPP700 are only used concurrently about 75% of the time, while the new single ‘Multiply & Add’ functional unit of the VPP5000 will effectively reach 100% concurrency.

Figure 1 shows how concurrency of memory loads and the multiply unit on the VPP5000 compares to the same sequence on the VPP700. The elimination of start-up costs is obvious.

The increase in technology (from 0.35 micron CMOS to 0.22 micron) since the design of the VPP700 has allowed more complexity to be included on the chips of the VPP5000. Many of the new features of this machine were thought about at the time when the VPP700 was designed, but the reduced technology meant that they could not be implemented on that machine. Consequently, the VPP700 had its limitations and deficiencies. We feel confident that most, if not all, of these deficiencies are solved by the VPP5000.

We have concentrated on the hardware differences between the two machines, and have said very little about new features in the software of the VPP5000. This new machine runs a 64-bit operating system based on UXP/V, the 32-bit system running on the VPP700. The VPP5000 can run in a binary-compatible mode with the VPP700, but some of the new hardware features, such as hardware SQRT cannot then be used. We hope to bring you more information about software in an article in a later newsletter.

With the acquisition of the new system this year and its upgrade next year, the Centre will continue to provide the European meteorological community with the necessary computing resources to ensure that they continue to remain at the forefront of developments in meteorological modelling research.

For more information please look at the Fujitsu web pages at URL:
http://www.fujitsu.co.jp/hypertext/Products/Info_process/hpc/vpp5000e/index.html

Table 2: Comparison of the Scalar Unit

	VPP700	**VPP5000**
Concurrent Instruction Execution	3 (Long Instruction Word)	4 (Very Long Instruction Word)
Size of Primary Cache (KB)	64 (32 D + 32 I) Direct	128 (64 D + 64 I) 2-way associative
Size of Secondary Cache (MB)	0	2
Peak Performance (MOPS)	428	1320
Size of Pointers (Bits)	32	32 or 64
Floating-point registers	32	64
FLOPs/cycle	1	2
SQRT instruction	Software	Hardware
Pre-fetch to cache	N/A	Hardware
Branch target buffer	N/A	Hardware

Size of Pointers*: The VPP5000 has 64-bit addressing, but can run in 32-bit (VPP700-compatible) mode.

	VPP700	VPP5000
Rate of instruction Issue	1 per 4 cycles	1 per cycle
Pre-load issue	N/A	Supported
Instruction Start-up Latency	10 cycles	8 cycles
FLOPSs/cycle	16	32
Number of multiply/add pipes	1 M + 1 A, each 8 wide	1 “M & A”, 16 wide
SQRT	Software	Hardware
Vector Registers	256 * 64 (8 Bytes wide)	256 * 64 (8 Bytes wide)
Number of load/store pipes	1 L + 1 S, each 8 wide	1 “L or S”, each 16 wide
Indirect Address Match	4 pairs at cycle ‘ n’	56 pairs at cycle ‘ n’ , 16 pairs at ‘ n-1’
Branch target buffer	N/A	Hardware

Table 3: Comparison of the Vector Unit

Indirect Address Match*: This reduces memory accesses when the same memory addresses are used during indirect memory calculations. This condition occurs a lot in IFS subroutine ‘LASCAW’.

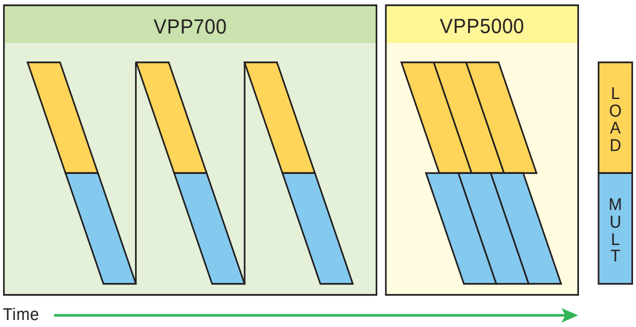


Figure 1: Concurrency of Memory Loads & Multiply Unit

STOP PRESS

Recently it has been possible to run IFS on a VPP5000 available in Japan for benchmarking. The performances are shown in Fig. 2, with comparative timings for the VPP700 and VPP700E.

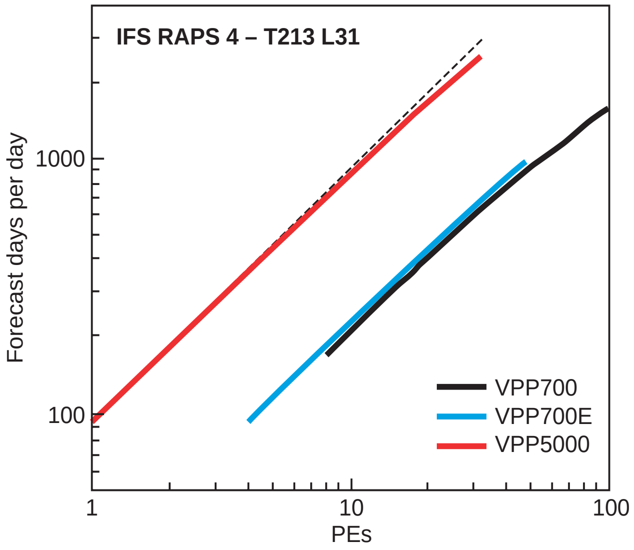


Figure 2: The IFS (cycle 12) at T213 L31 performance on various FUJITSU vector computers.

Neil Storer

Seventh workshop on meteorological operational systems

ECMWF, Shinfield Park, Reading, Berkshire, United Kingdom

15 - 19 November 1999

The planned biennial Workshop on Meteorological Operational Systems, to be held at ECMWF 15-19 November 1999, will be the seventh in the series.

The workshop will review the state of the art of meteorological operational systems and will address future trends in the use of medium-range forecast products, data management and meteorological applications on workstations.

Use and interpretation of medium-range forecast guidance

The session will address the problems and solutions related to the use of numerical guidance in medium-range weather forecasting. The operational ECMWF Ensemble Prediction System has been developed further in recent years and the product range has been extended. With the introduction of the coupled ocean-wave/atmosphere forecasting system, additional wave EPS products will become available. The operational procedures for the evaluation of probabilistic forecasts have been developed further to allow a user-orientated assessment of the quality and the value of the forecast.

Operational centres will present their approaches to medium-range weather forecasting and report on their experiences with a combined use of output from different models, including ensemble prediction systems. Presentations on preliminary attempts to issue or verify forecasts on the seasonal range will also be welcome. The issue of forecasting extreme weather events in the medium range will be addressed and discussed in a working group.

Operational data management systems

The development of operational database and archive systems, as well as data delivery systems, will be reviewed. Particular consideration will be given to the use of the Internet for accessing data and for the provision of on-line data descriptions and documentation. Tools for controlling operational meteorological systems and data flow will receive special attention, as will the benefits of distributed databases and archives.These issues will be addressed and discussed in a working group.

Meteorological workstation applications

Updates to existing 2/3D visualization applications and new developments in this area will be presented and demonstrated at the exhibition during the workshop. At previous workshops the focus was on traditional UNIX systems. This time topics related to porting meteorological UNIX applications to other systems such as NT desktops, LINUX based systems and JAVA will also be presented. Strategies for porting meteorological visualization software to other systems will be discussed in a working group. Experience with available porting tools will also be addressed.

ECMWF Publications

Technical Memoranda

No.268 **Teixeira, J.**: The impact of increased boundary layer vertical resolution on the ECMWF forecast system *February 1999*

ECMWF Re-Analysis Project Report Series

No. 7 Comparison of the land-surface interaction in the ECMWf re-analysis model with the 1987 FIFE data by **Alan K Betts**, **Pedro Viterbo** and **Anton C M Beljaars** (also published in *Monthly Weather Review*, 1998, **126**, 186-198).

ECMWF Workshop Proceedings

Workshop on Predictability *20-22 October 1997*

Forecast and Verification charts up to and including June 1999.