THE UNIX INTEGRATED ENVIRONMENT AT
THE CANADIAN METEOROLOGICAL CENTRE

Robert Mailhot, Garry Toth, José Garcia and Serge Filion
Canadian Meteorological Centre
Dorval, Canada

1. INTRODUCTION

In 1989, the Canadian Meteorological Centre (CMC) decided to acquire a new generation of both front end and back end computers along with graphical workstations. The switch to a single widely available operating system, UNIX, was an integral part of the plan. The following three years were a period of intense effort by all of CMC's personnel. As a result, UNIX has been the sole operating system in use at CMC since December 1992.

It is impossible to cover all the work accomplished by a large team of colleagues and which covered all fields of activity at CMC during that three-year period. We will therefore focus on some aspects that, we believe, were then most important.

First, we will briefly present our current hardware configuration. This will be followed by some considerations that strongly influenced the design of our operational system. The last part will describe the design and/or structure of some important operational subsystems.

2. CURRENT CONFIGURATION

Figure 1 shows our current hardware configuration (November 1993) and its evolution over the past four years. Conversion took place in two steps. The first part was completed in June 1992 after the replacement of our two former CYBER 830 front end computers, running NOS, by the MIPS 4680s. Part two was achieved six months later with the replacement of the CRAY-XMP by a NEC HNSX-3.

Evolution of Computers and Operating Systems

	December 1989	June 1992	December 1992
FRONT END	CYBER 830 (NOS)	2 MIPS 4680 SV 2 CDC 920 SV 60 CDC 910 WS 10 INDIGO WS 1 CRIMSON (with Reality engine) 75 X-TERMINALS 20 SUN WS 7 H.Packard 9000 WS 50 PC(as PCs or dumb terminals) (UNIX)	
BACK END	CRAY XMP 4-32 (COS)	CRAY XMP 4-32 (COS)	1 NEC HNSX- 3 4-256 (UNIX)

SV : server
WS : workstation
PC : personal computer

Fig. 1 Configuration refers to November 1993.

The two MIPS 4680s are the workhorses. Each of them has four 80 mips processors. One MIPS (coined CIDSV07) is dedicated to CMC operations while the second one (CIDSV08) is used for development and research. It also serves, if necessary, as the backup for the operational machine. The definition of "front end operations" may vary from Centre to Centre. At CMC, the term includes data reception, bulletin parsing, decoding, monitoring, quality control, pre-processing for the objective analysis, chart/bulletin/GRIB production, file maintenance and archiving. The NEC is used solely for the execution of objective analyses and numerical models. To summarize, everything that is not research or development or numerically intensive is executed on CIDSV07. The two MIPS 4680s will be replaced by more powerful machines by summer 1994.

3. DESIGN CONSIDERATIONS

We retained UNIX as our operating system for several reasons. First, it is becoming a worldwide standard, available on all kinds of platforms. In addition, its removal of a dependency on vendors is likely to ensure economy in the future. Finally, the replacement of the NOS-COS tandem by a single (though multi-flavoured) operating system means easier code maintenance and shorter development times.

Important factors in CMC's design decisions included deadlines linked to the end of contracts for old machines and obsolescence of existing software. Most software running on the front ends in 1989 was designed and written around or before 1980 during a period of vastly different priorities. For example, monitoring and quality control were a low priority then. Fulfilling a mandate in quality control on an overloaded machine with inflexible software would have been an impossible mission. No less important was the fact that our operational database was suffering from a few shortcomings. Another consideration was that the new workstations provided an excellent opportunity to develop graphical applications and, consequently, to redefine our plotting package.

Since one of our objectives was a better integrated environment, we paid particular attention to the needs expressed by various users in research, development and operations, resulting in strengthened relationships among the three groups. The fact that we all share common program libraries and graphical applications is only one benefit of this approach.

Such considerations were at the origin of changes that took place at CMC during the conversion period. These factors triggered a serious reflection about the way we were doing things and led us to examine carefully all aspects of our operations. We had not only to change JCL and Fortran codes but also had to adopt new approaches and find better ways of improving performance. We spent, not surprisingly, about 70% of the time developing and only 30% converting applications. In summary, the conversion, while important, presented us with the ideal opportunity to redesign many of our applications with their better integration as a goal.

4. DESIGN OF SOME IMPORTANT OPERATIONAL SUBSYSTEMS

In this and the following sections, we will focus on four important aspects of our new operational system. Specifically, they are:

- the processing of incoming meteorological data;
- the structure and formats of our operational meteorological database;
- the development of a datafile management system;
- the development of a common set of graphical applications.

4.1 **The processing of incoming meteorological data**

Meteorological data are obtained from a variety of sources including an internal network (for Canadian data) and the Global Telecommunications System (GTS). Bulletins are received and accumulated on our Telecommunications computer, a TANDEM TXP Non-Stop. It is connected to CIDSV07 via a Personal Computer (PC), which is a cheap, convenient and reliable interface.

The PC creates packets from bulletins obtained from the TANDEM. A packet comprises bulletins received during a three-minute period, although this is occasionally shorter as the maximum size of a packet is 100 Kbytes. The PC normally transfers such a packet to CIDSV07 every three minutes using FTP, while as many as eight packets can be transferred at a time during catchup periods. This is efficient buffering to cope with the irregular arrival of the bulletins or any downstream failure.

A daemon running on CIDSV07 senses the presence of packets every thirty (30) seconds. If one is detected, it calls a job which directs bulletins to one of the three appropriate subsystems: GRIB, BUFR or ASCII. The decoding process can now start. It is followed by insertion into the operational database. When a packet is present, all relevant decoders are launched at once resulting in the use of all available processors. One process per data type is currently allowed to run at a given time, although this could be increased. A second SYNOP decode/insert process will not start unless the previous one is finished. In order to avoid locking problems, UNIX file locking is used by the insertion programs, so that they only may write to operational files.

4.2 **The structure and format of our operational meteorological database**

4.2.1 *The structure*

A subset of the operational database is illustrated in Figure 2 (next page). We take advantage of the tree-like structure of UNIX. Files are defined by data type (SYNOP, DRIFTER, etc.). Each file contains six hours of data that are centred on a main synoptic hour. This was chosen to make the connection to our data assimilation cycle easier. All files share similar identification ("YYMMDDHHHH_"), composed of the "year_month_day_hour". The file name is meaningless unless one knows its pathname.

4.2.2 *The "Standard Format"*

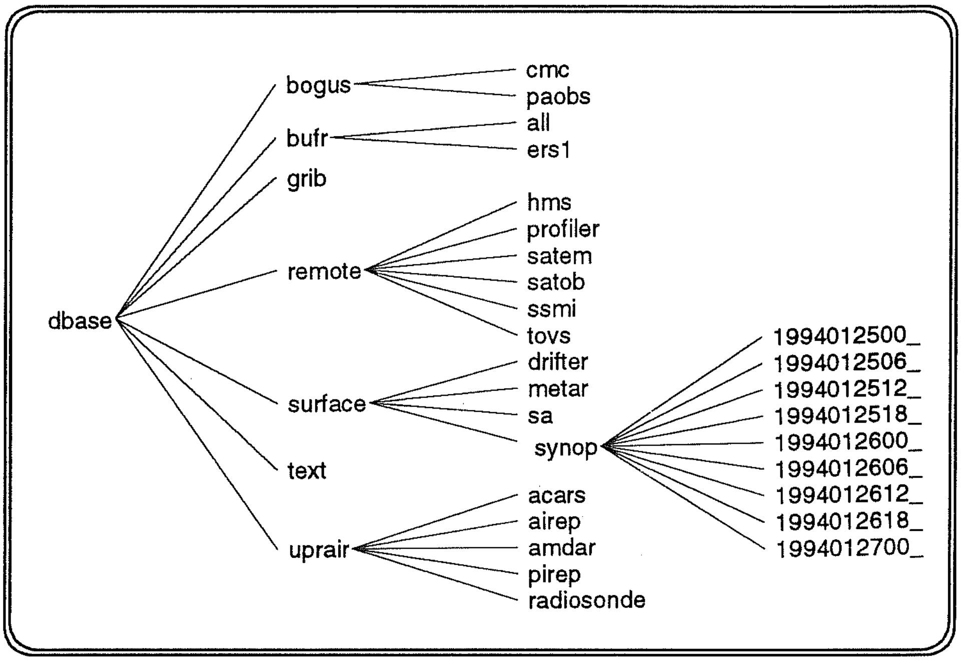

Fig. 2 Subset of our database illustrating its tree-like structure. For clarity, only SYNOP is fully shown. The last ten days of data are available on disk.

We identify two kinds of data to be saved: model outputs and data from stations. Model outputs are saved at CMC in the RPN ("Recherches en Prévisions Numériques") Standard Format ("RPNSF"). This format has been in use for 11 years. One record contains all grid points associated with one variable at one level (height at 850 hPa, for example). Files are portable to 32 bit, 64 bit and 60 bit (because of the former CYBER 830 machines). This format, GRIB-like by design, has proved to be very compact, quite efficient and very appreciated by users. It is a well-established standard within CMC and is widely used within the Atmospheric Environment Service.

4.2.3 *The "BURP" format*

We were in need of a complementary format, one in which a record would contain all data from all levels from one point ("station"). We had the choice to build it from scratch or from a commercial DataBase Management System (DBMS). Our main objectives and constraints were the following:

Objectives:

1. One station per record with capability of station grouping.
2. Files portable to 32 bit and 64 bit machines.
3. Flexibility (to meet present and future needs).
4. Compactness of files.
5. Rapidity of access to files.
6. Adequate management of all types of variables pertaining to stations (observations, statistics, quality control flags, forecasts, analyses, etc.).
7. Minimization of the number of input/output operations.

Constraints:

8. Conversion had to be completed by March 1992. This meant that our database had to be ready by September 1991 since this activity was quite far upstream in the sequence of events.
9. We had to satisfy operational, development and research needs.
10. The performance of the expected front end was in the range of 20 to 30 mips, single processor.
11. We had to cope with three different sets of operating systems during the conversion (NOS-COS, UNIX-COS, UNIX).

Additional arguments militated against the use of commercial DBMSs: no expertise at CMC, no experience within any other similar operational Centre, high costs (in terms of disk space, CPU and site license) and relatively low speeds. Given all this, we believed (and still do) that commercial DBMSs will belong to the domain of research and development for another five to ten years before becoming useful in a complex environment like ours. We opted for the development of a local format which was named BURP (Binary Universal Report Protocol). Inspired from BUFR, its development was a joint venture with RPN. The most important feature we retained from BUFR was the use of the Table B universal descriptors.

Both RPNSF and BURP can be described as direct access formats that use a look-up table accessed sequentially. Each entry of this table is a subset of a record header and contains the various parameters needed to identify the contents of a record. Since the table is copied to main memory at the opening of a file, very good performance is achieved.

Overall, we are very satisfied with our choice. The performance (size of files and access time) of our local formats meets all our expectations. We now have the proper tools to work efficiently.

5. DATAFILE MANAGEMENT SYSTEM (DMS)

We archive about 250 operational files in each six hour period. This includes all files read/written by the objective analyses and/or models executed in the various runs. A corresponding number of files must be deleted to maintain equilibrium. This deletion is one of the two main roles (archiving being the other) of the DMS. Archiving was an important problem initially because no satisfactory commercial tape archiving software seemed to exist on UNIX. We finally chose one and added bug fixes.

For the DMS, the life of a file can be divided in three parts: the identification, the life itself and the deletion of the file, with all three taking place on disk. Identification and deletion are mandatory while the reduction of the file and its archiving as the two potential actions during its life, are optional.

5.1 Reduction of the file

Reducing a file implies one of the following:

- removing unnecessary or redundant records; this is done on binary files.
- compressing a text file; this is done on ASCII files.

In both cases the resulting file overwrites its source.

5.2 Archiving a file

Archiving a file is an action that moves it to a temporary file, called a bucket. Once a bucket reaches 200 Mbytes (the capacity of a cassette, our archiving medium), it is dumped onto a cassette and emptied. As cassettes are all backed up, we use about fifteen cassettes per day. Very shortly after it becomes available, a file may be archived as is, or after reduction. The reduced file may or may not overwrite the original one.

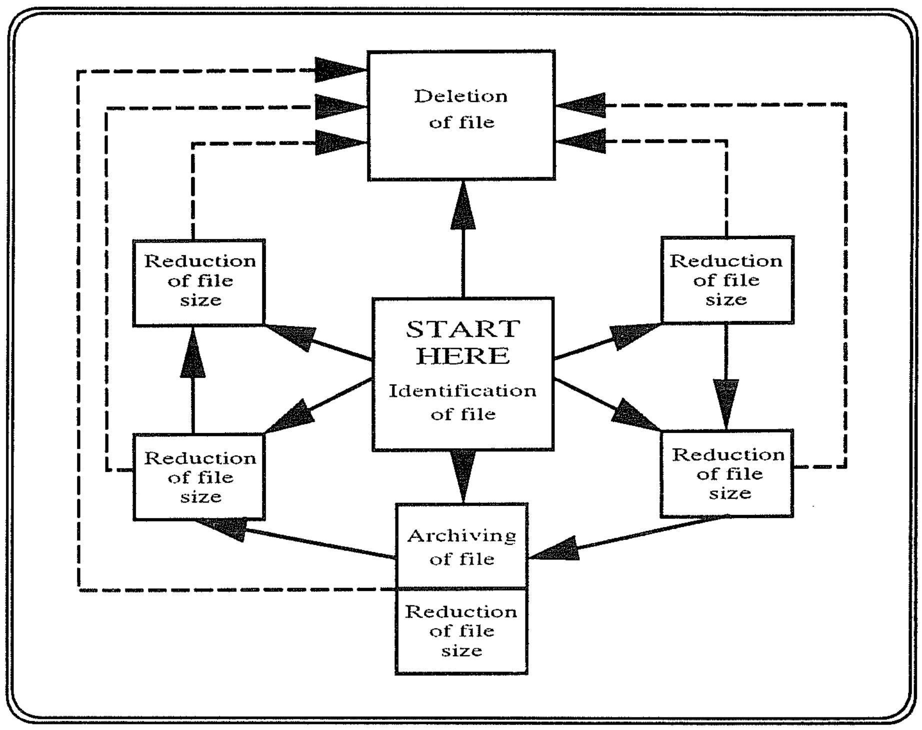

Fig. 3 Schematics of the various combinations of actions on files. Starting from the centre, one moves one step in any direction; from there, you are allowed 1, 2 or 3 steps in the direction indicated by the arrow before ending up to D.

5.3 Control files

Permanent files called "protocole" tell the system what to do and when to take action on every file. An action file is created when a file is identified and it contains a timetable of all the actions to be executed sequentially on the file. Finally, a status file informs the system of the location of every operational file known by it: on disk, in a bucket or on cassettes. Figure 3 shows the various interactions between identification, reduction, archiving and deletion of files.

All users have access to the contents of "protocole" files. We also maintain files online giving information about the availability of all historical files. Users may check this before launching a job to extract archived files or subsets of files.

6. DEVELOPMENT OF GRAPHICAL APPLICATIONS

One of our objectives during the conversion was to avoid duplication of work. Our goal was that people working in operations, development or research would share, as much as possible, the same applications, since their needs would be either similar or complementary. Besides, some of our main libraries of routines had been common for several years.

We did not have enough time to develop one global application which could handle all needs. Nevertheless, we felt it was necessary to make sure that there would not be two applications performing similar tasks. Again this led us to building several applications that did what people needed. They read essentially the two formats, RPNSF and BURP, used operationally at CMC and within the Atmospheric Environment Service.

6.1 **Plotting package**

Our plotting package was revamped. The new one, SIGMA ("Système intégré de graphisme météorologique avancé") reads the two operational formats, BURP and RPNSF. Based on it, we developed several specialized graphical applications:

1. DATAMON...: reads BURP files and displays in real time the quantity of observations available by type and geographical region. It is used by operational meteorologists as a monitoring tool.

2. EDIGRAPH....: allows us to create subjective prognoses on workstations and transmit them directly to the communications system. From RPNSF we will import forecast fields, such as mean sea level pressure, and modify them as necessary. Objects, such as fronts, clouds, text, etc. can be added.

3. MAX..............: MAX reads BURP and RPNSF files. This application is used for contouring analyzed or forecast fields and plotting data. We can also enter bogus data. Macros can be defined. There is a predefined set for displaying derived fields and verifications. A powerful calculator does arithmetic, trigonometric and logarithmic calculations on fields. Various derivatives can also be calculated. All these operations may be combined as desired. Finally, several verification scores can be produced. MAX also displays vertical profiles and cross-sections.

4. REC...............: REC allows us to display and animate the contents of RPNSF files. Parameters, such as color palette and animation speed, are controlled on the fly.

5. TEPHI...........: As its name implies, it is used to display observed, analyzed and forecast tephigrams. Hodographs can be displayed as well.

6. XDISPLAY...: This application was developed to display and animate raster files. Images from other applications can be captured, temporarily saved and then displayed side by side for easier comparison.

Raster images can be saved in "carousels" and used later for presentations or case studies. Applications #3 to #6 are extensively used by meteorologists working either in operations, development or research. One of our objectives was fast display. The readers who attended the workshop realized, through the video we presented, that this objective was indeed achieved.

Though this exercise is exactly what any sound consultant would have vigorously recommended NOT to do, we did it. We were "forced" to succeed for several general reasons: external constraints (termination of contracts for example), "quantum leap" in technology and new needs to satisfy. Several more specific factors contributed to the success of the whole exercise. The three most important were proper training in UNIX and the C language for all personnel involved in the conversion, highly motivated staff and unwavering commitment from our management team.

7. CONCLUSION

During a three year period extending from December 1989, CMC undertook major operational changes. We obtained a new generation of computers and switched to UNIX from the former tandem of NOS-COS and we seized this opportunity to reconsider how we were doing things. This led us to redesign most of our system and to take advantage of UNIX properties, the new open network and the availability of workstations.

Much work remains to be done. However, we are satisfied that, during this period, we built the tools that are necessary to fulfil our mandate as a National Meteorological Centre. We look forward to the future with confidence.

8. ACKNOWLEDGEMENTS

The authors would like to sincerely thank Gilles Verner and Charles Anderson for reading this paper and providing judicious comments.