HOW SUCCESSFUL HAVE COMMERCIAL TOOLS AND PACKAGES BEEN IN THE MANAGEMENT OF AUSTRALIAN OPERATIONAL DATA?

J.B.Lorkin
Bureau of Meteorology
Melbourne, Australia

1. INTRODUCTION

The Bureau of Meteorology distributes its operational functions between a Head Office in Melbourne and 7 Regional Offices, one in each State. The computing infrastructure is similarly distributed. In 1994, HO was converted to a network of UNIX computers and now has 2 Crays, 6 HP9000/800 servers, a StorTek mass store and many workstations linked by Ethernet and FDDI LANs. Conversion of ROs to UNIX is underway, from Tandem Non-Stop to IBM AIX. These conversions have been catalysts for significant changes to operational data management systems. Some HO systems jumped at least two decades in software architecture by using or building on software developed elsewhere. I will describe the major building blocks, some commercial and some from other Met services, and discuss their value to the Bureau.

2. SOFTWARE BUILDING BLOCKS

2.1 Databases

2.1.1 *NEONS*

NEONS was chosen because it was a proven architecture for operational met data. It has been used very successfully to store global observations and fields from Australian models and also EC, UK and US models. Data access methods used by old programs were rewritten so that they retrieved data from the NEONS database. This simplified the conversion of the programs but the development effort was badly underestimated and that approach would not be taken again; instead the programs would be modified to use NEONS calls. Overall, NEONS has been very useful for HO systems and will be used in ROs as they are converted to UNIX. To increase the transportability of programs, the RO RTDBs (Real-Time Data Base) are planned to look like the HO RTDB.

The major complaint about NEONS is that the observations cannot be read by SQL because they are stored in near-BUFR. If general users had been able to use SQL packages to read the observations, there would quickly have been a second complaint, namely the use of SI units. These will be serious problems in ROs. The proposed solution is to duplicate the storage of conventional observations - once in NEONS and once in SQL tables. The SQL tables will be similar to those in the Climate Database and will use conventional observing units instead of SI units. Additionally, a parameterised extractor program is being used to put the observations into a self-describing flat file in ASCII; this file is readily addressible by general packages.

The centrality of the NEONS RTDB is limited to computers with RDBMS licences, notably excluding our supercomputers. Tranfer files can then become the basic interfaces for software systems instead of NEONS. We are interested in the development of netNEONS to provide greater access to NEONS.

2.1.2 *EMPRESS*

Empress 6.2 was chosen in 1993 because it was cheap and known to work well with NEONS. Overall, it has proved to be satisfactory although its client/server has been very slow and developers have not been happy with the 4GL. Some critical jobs were moved from client computers to the database server to speed them up. Empress 6.4 was not used for the DB because its client/server failed after many update transactions in a way which Empress was unable to explain. Support from the Australian agent and from Empress in Canada has been poor; prices have risen to be comparable to Oracle.

2.1.3 *ORACLE*

ORACLE V6 was chosen for the climatological DB (ADAM) in the early 90s because of its wealth of features and tools. ORACLE CASE was used for the DB definition but is not in regular use because the ADAM structure is relatively simple; however it will be used for a complex system to maintain information about stations. ORACLE FORMS is used for the entry and QC of observations. Development from scratch was very slow and much of ADAM was incomplete when the FACOM computers were decommissioned; conversion midway to Oracle V7 was very painful. Two-thirds of ADAM is now in place and is being fed from the RTDB every few hours. The challenges at present are to settle on one RDBMS and to make data in the RTDB and ADAM DBs accessible to outside organisations.

2.2 **Offline storage**

EPOCH was chosen in 1993 to migrate files from disk to tape and was implemented in July 1994. Each day, 6GB of operational data and 3GB of research data are archived through EPOCH to the StorTek mass store; 3GB are retrieved. EPOCH runs on a SUN SPARC II with 34MB of RAM and 20GB of disk.

The Bureau's experience with EPOCH has been unsatisfactory. The problems are

. an inefficient Hierarchical Storage Manager which takes too long to search its DB of file pointers
. (once) a large loss of data when the DB of file pointers was corrupted
. very slow speed of retrieval
. insufficient disk space for migration and retrieval
. the need for careful nursing by sysadmin staff

It has been described by some users as a WORN archive (Write-Once Read-Never). Support from EMC, the new owners of EPOCH, has been poor. EPOCH was not ported well from IBM to UNIX and EMC does not seem to understand UNIX. EPOCH will be replaced within 12 months.

2.3 Scheduling

SMS/XCDP was purchased from the ECMWF in 1992 after an abortive attempt to build a scheduler on an expert-system shell. It runs on a SUN SPARC II, uses little resources and has been an excellent package. We have not upgraded yet from V4.0 mainly because of some difficulty in porting it to the HPs, but this will be done soon.

2.4 Graphics

NCAR GRAPHICS V3 is now the engine for display of all HO operational charts and observations. The BMRC has extended the NCAR Diagnostic Processor to interface to the Bureau's files and to the RTDB. McIDAS-X is used for interactive display of fields and observations - although IDL is gaining momentum as an interactive display tool, but it does not interface adequately to the NEONS RTDB yet.

CIMCAD was chosen in 1991 as the basis for an application to draw significant weather charts for aviation. This has been useful because it ensures legibility, but it is not fast to use and it does not interface to observations and fields in the database. We have not been able to devote enough programming effort to make it really good.

2.5 yacc/lex

Functions for decoding character-coded observations have been written using mixtures of yacc, lex and C. The yacc code resembles the WMO Codes Handbook and this makes it very easy to maintain. This has been a very pleasing development to those of us accustomed to the complexity of decoding in FORTRAN or C. The decoders (character, BUFR and GRIB) are transaction driven and are controlled by tables. Some of this code may be of interest to other met services.

3. SUMMARY

The Bureau has gained much by using commercial software and software developed by other Met services. Staff and financial pressures will strengthen this trend.