DATA FLOW IN THE METEOROLOGICAL COMPUTER SYSTEM AT DWD

J. Böttcher
German Weather Service
Offenbach, Federal Republic of Germany

Summary: A short description of the new computer system at the German Weather Service is followed by a representation of the data processing system to be developed to run on the new computer system.

1. INTRODUCTION

During this year, the German Weather Service got its new computer system. Routine programs for the new system are not yet fully developed. Therefore, the following will describe what we expect the future system should look like.

2. THE NEW METEOROLOGICAL COMPUTER SYSTEM AT DWD

The new German Weather Service computer system consists of an ETA 10, a CDC Cyber 180-850 and a CDC Cyber 180-860.

The ETA 10 has 4 CPUs, two and two cooled in fluid nitrogen. This permits the use of at least 2 CPUs, even in case of maintenance. Therefore, e. g. routine forecast models may be designed to use 2 CPUs in parallel. A 1 GBytes Sheared Memory reduces access to Input/Output Units to the absolute minimum.

The two Cybers serve as front ends to the ETA10 through the Loosely Coupled Network (LCN) which has a maximum capacity of 50 Mbit/s. The Cyber 180-850 is reserved for routine programs. As long as routine programs are not fully developed under the NOS-VE operational system or for use on the ETA10 the Cyber 180-850 is running in dual state under both operational systems NOS-BE and NOS-VE. NOS-BE on Cyber 180 permits running of most of our Cyber 173 programs after minor

changes. The Cyber 180-860 may be switched to dual state in case of maintenance or failure of the Cyber 180-850. Normally the Cyber 180-860 will be available for program developing and testing. A relational database system (IM/DM) is also running on the Cyber 180-860. Interactive terminals, printers etc. are connected to the Cybers through CDCNET which has a maximum capacity of 10 Mbits/s.

The VAX computers of our telecommunication centre are connected with the LCN.

3. DECODING OF METEOROLOGICAL DATA

The contents of meteorological data received from the telecommunication centre must be detected, checked and translated to BUFR code. This procedure is called decoding of meteorological data.

The new decoding program is using tables containing information on data structure and code forms. These tables are read by the program in parallel to the data. Using tables will reduce maintenance of the program to a minimum while maintenance of tables can easyly be done by using a relational database. A comment in each line of the tables simplifies maintenance and may be used to produce explanations in control output.

A Headline Table is used to detect what kind of data follow. It contains data designator/geographical designator/number of bulletin (TTAAii), international four-letter location indicator (CCCC) or the word "SHIP", the WMO FM number or a free number for codes with no FM number, and an indicator to the subroutine to be called. In most cases the first one or two letters of TTAAii are sufficient to detect the code form of the messages following. However, we have national bulletins where checking of CCCC in addition to TTAAii is necessary to detect the code form. Table lines containing the word "SHIP" in lieu of the international four-letter location indicator

cause checking of the geographical designator to have a valid value for bulletins containing ships' weather reports (including reports from automatic marine stations).

A Decoding Table is used to detect and extract specific quantities from reports. Lines in the table describing coding procedures for code groups contain a table line key, a pointer to special decoding procedures, if nessessary, the length of the group, if fixed, an indicator for test output and length and contents of an group indicator, if applicable. Lines in the table describing coding procedures for specific quantities in the code contain a table line key, a term for addition and a multiplier to allow adjustment from reported values to BUFR units, a pointer to special decoding procedures, if nessessary, length, starting position within a group, minimum and maximum value of the quantity to be decoded, and the information where to store the value of the quantity within the program. Special lines containing a pointer to the starting point of a repetition permit processing of parts of code which may be reported more than once.

A table for the preparation of BUFR messages from extracted values, is not yet developed. It should contain for any BUFR format to be prepared a list of codes leading to the BUFR format and a list of elements to be included in BUFR containing program adresses of elements, lengths of values in BUFR and, if necessary, information on repetition of data.

Separate BUFR formats are prepared for those parts of reports which are used only by a few stations or only at specified times of a day. Separate BUFR formats are also prepared for those parts of reports where an encoding procedure is not provided. The corresponding data are transferred to the appropriate number of one character elements described by operator 205001 of BUFR table C. Using tables reduces remarkably the need of changing programs if coding procedures change. Only the special decoding procedures in the program are related to coding procedures of meteorological codes.

However, the special decoding procedures are designed for use at various places in meteorological codes, if possible.

The decoding program is preparing BUFR formats consisting of a sequence descriptor (defined nationally), and BUFR Section 4 (Data Section) only. Therefore, supplementary information like the local tables version number or information which might be useful for error tracing has to be included in the Data Section.

4. A HIGH SPEED DATABASE AND ARCHIVAL SYSTEM

A relational database system or similar commmercial database systems are not suitable to be used as an operational meteorological database for storage of observations. The high rate of new data during each day would slow down the speed of operations of commercial systems below acceptable limits in a few hours.

The design of a high speed database must be adjusted to invariant attributes of data. Invariant attributes of meteorological data are time and geographical location of observation. The order of the two invariant attributes depends on whether the data are to be used for synoptical (Request: All data at one time of observation) or climatological (Request: All data of one station) purposes.

The high speed database system for observations, primarily designed for synoptical use, is designed as follows:

> Observations made during a predefined time interval (3 hours or surface, 6 hours for upper-air observations and 1 month for climatological reports) are stored in one data file. Two additional files contain geographical pointers and and pointers to data in the data file.

All predefined time intervals cover a predefined number of days (at present 2 days for observations and 2 years for climatological reports). If, for actual data, a new data file has to be opened the file containing eldest data will be deleted. Delayed data are stored in a separate file for later archiving.

The database is defined in a fixed geographical area. Normally the fixed area is the whole world. However, for special data sets the fixed area may be limited to a part of the world and data outside the area will be rejected by the database.

Observations are stored in a data file as delivered by the decoding program. Therefore, if one of the two pointer files would have been destroyed the database could simply be repaired by reloading its own data file. The user may wish to get data in an expanded format either one computer word per quantity (normally integer format, however, character format is used for data stored in CCITT IA5 format) or in a character-oriented form (each quantity occupies a fixed number of characters). The second form may be used as input for a commercial database system or for data exchange to computer systems not capable to handle BUFR code. The user may also get a format which is a combination of more than one internal BUFR format. In this case the non-meteorological part of each report (station data, time of observation etc.) would have the same format regardless the type of the station.

The compressed form of BUFR is used to archive observations. This will reduce the number of tapes, however, restoring of archived data to database requires expansion of the data.

Grid point data of numerical analysis and forecast are stored in GRIB code in a database file. An additional file contains pointers to data in the data file. The data base may be defined for as many days as needed. If, for actual data, no

space is available in the file the eldest data will be deleted. Grid point data are archived in the same format as they are stored in the database.

The high speed database system of DWD may be used either interactively through menus or by program call. Files are used to transfer data from database to user and vice versa.