

UNITED STATES
DEPARTMENT OF THE INTERIOR
GEOLOGICAL SURVEY

Data Format Conversion Programs for the PC and SUN Computer

by

John A. Rogers¹

Open File Report # 89 - 662

This report is preliminary and has not been reviewed for conformity with the U. S. Geological Survey editorial standards. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U. S. Government. Although these programs have been used by the U. S. Geological Survey, no warranty, express or implied, is made by the USGS as to the accuracy and functioning of the programs and related program material nor shall the fact of distribution constitute any such warranty, and no responsibility is assumed by the USGS in connection therewith.

Menlo Park, California¹

1989

ACKNOWLEDGEMENTS

I want to thank John Lahr for his patience in helping test the PC2SUN program and Chris Stephens for his work in modifying the PC2SUN program and helping integrate the IRIGE subroutine into the program.

ABSTRACT

A program has been written to convert from a format used to collect data on the PC computer to a popular data format used for earthquake recording on the SUN workstation. The program takes as input an event file generated on the PC and produce as output a file in a format compatible with the SUN computer. The conversion of formats makes possible the use of an analysis program on the SUN. Another program converts a PC data file to a text file.

INTRODUCTION

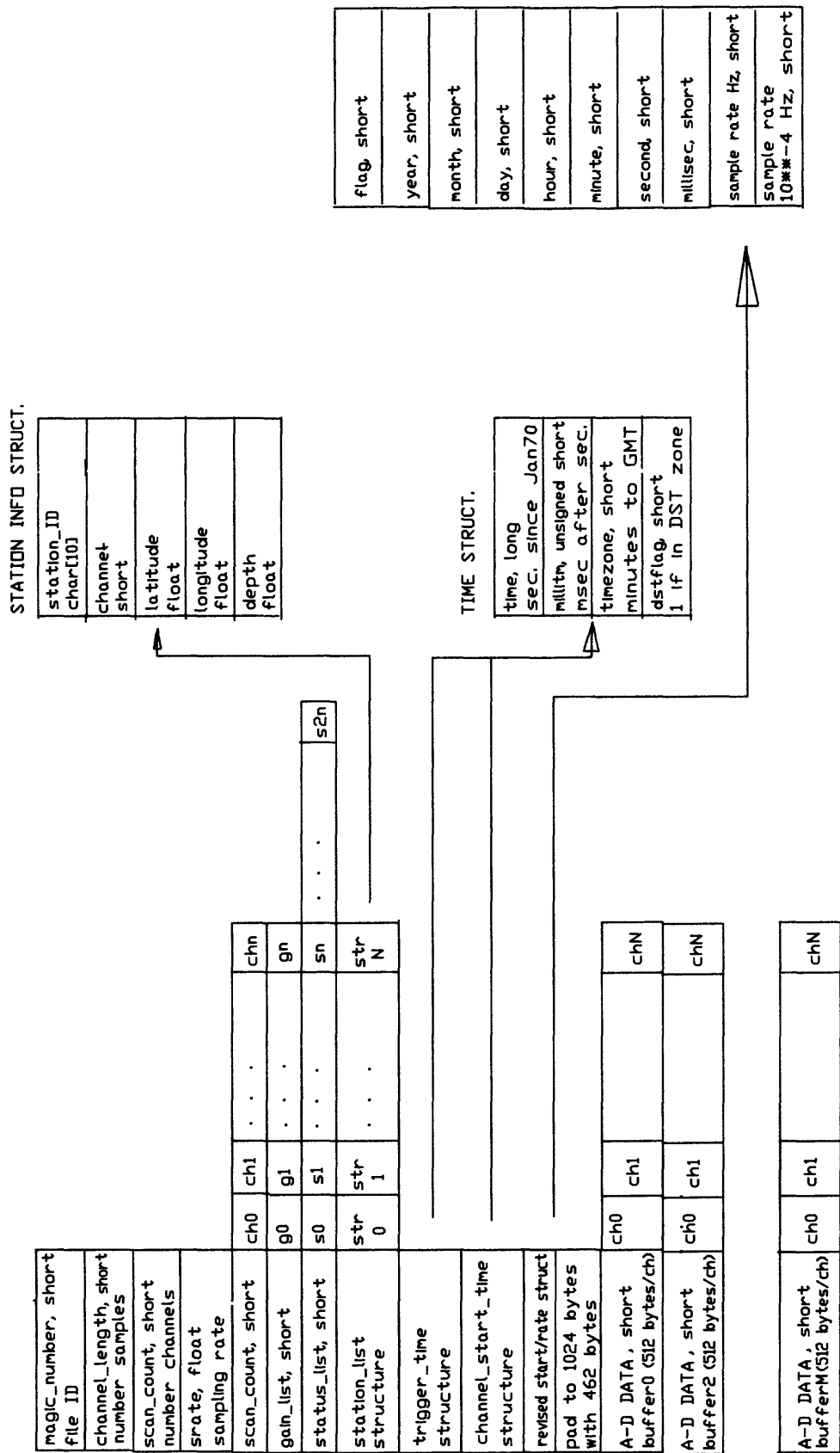
The phase-out of analog recording of seismic data in favor of digital event recording has offered the possibility of easier data exchange between researchers. This exchange was never convenient with the older analog systems due to technical difficulties of copying and playing back analog tapes. Unfortunately the switch to digital event data has introduced a new set of problems which still limit the interchange of data. Although copying or transmitting data is now very simple, the data itself is often encrypted in poorly documented formats usually unique to the computer system of the original recording. The result is that data recorded on one computer cannot be analyzed on a different computer.

Eventually a seismic data format recording standard will probably emerge, but in the interim there will be a need to convert from one format to another. In Alaskan seismology this situation has already arisen since part of the recording is done in the Lamont "AH" (ad hoc) format while another part is done in the PC (Lee, 1989) format. By converting the PC data to the AH format, data can be merged and then analyzed using a single analysis program on the SUN computer.

CONVERTING PC FILES

The Microsoft C language program "PC2SUN" converts event files generated on the PC to the AH format. Figures 1 and 2 show graphical representations of the two formats. One fundamental difference in formats is that the PC format uses one file per event where as the AH format uses one file per channel for each event. Although all of the data files are binary and use common data types (integer, floating point etc.) the data is stored differently on the PC and SUN computers. For 16-bit integers (called short throughout), the two 8-bit bytes must be swapped before the data can be used. For 32-bit long and float types the two 16-bit words must be swapped as well as the two 8-bit bytes of each word. Character data does not need to be converted. This conversion is performed by the three functions `intswp`, `lngswp` and `fltswp` with each returning the converted value. If it were not for this difference in representation, format conversion would amount to changing the order of the data in a file.

The PC2SUN program first reads the PC header and stores the converted values (using the three routines described above) in the AH header structure for each channel data file. Since the PC data is multiplexed in 512-byte chunks, data for each channel is written to the appropriate AH channel file as many times as there are buffers in the PC header. Thus, if a PC event is 13 buffers long, each AH data file created will be written to 13 times with waveform data, once for each buffer.



KEY: short=16 bits, long=32 bits, float=32 bits
channel_list is digitizing order; channel_stat_list is ON or OFF for trigger

Figure 1. PC Event Structure

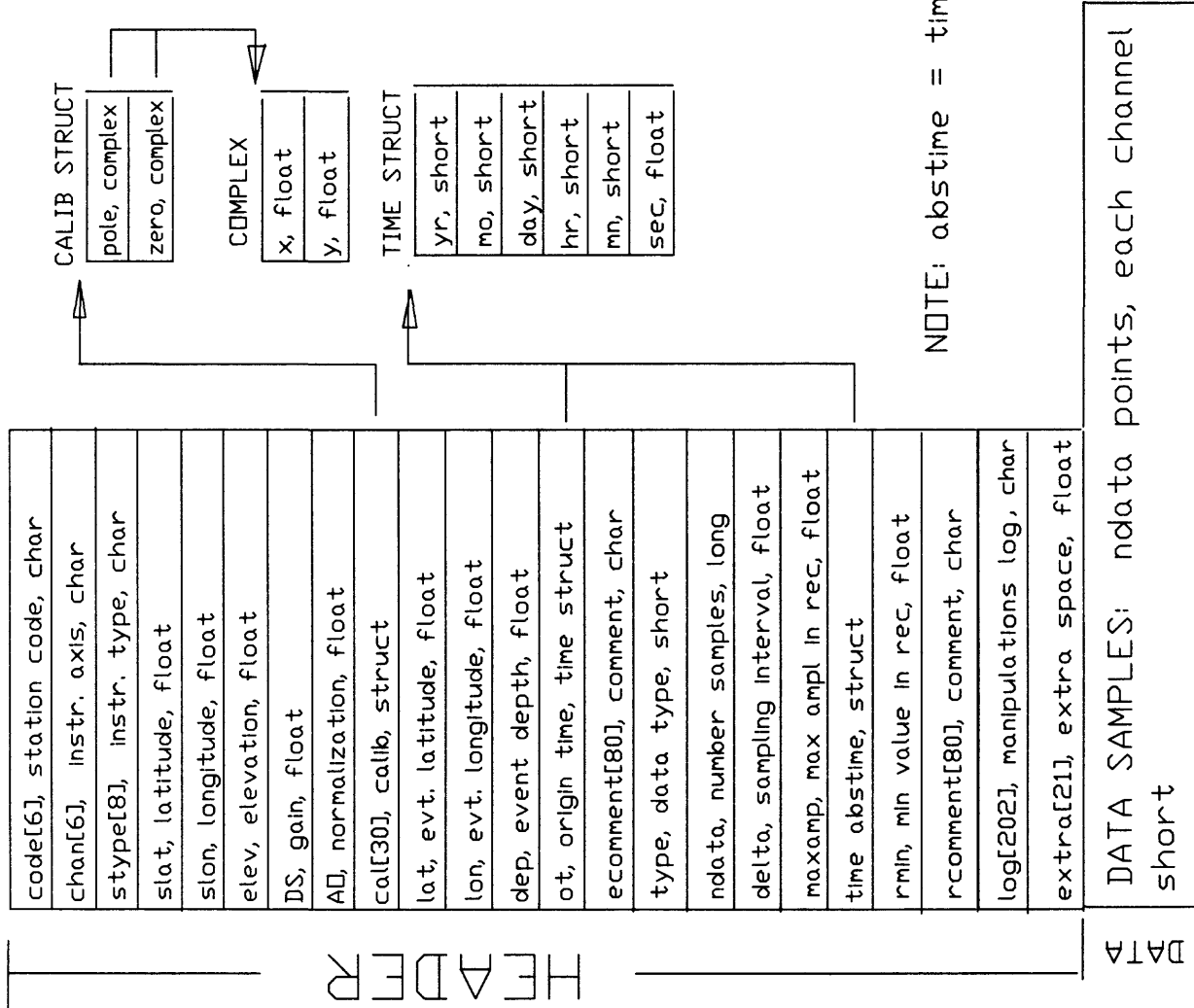


Figure 2. AH Event Structure

PC2SUN program also solves a timing accuracy problem inherent on the PC. The PC clock uses an unsynchronized crystal oscillator which drifts on the order of 1 second per day. The PC2SUN program can replace this inaccurate time with the time provided by an IRIGE time code. This time code is available from the GOES satellite and is accurate to 10 milliseconds. This replacement is achieved by calling the FORTRAN subroutine "IRIGE()" (Carl Johnson, USGS CUSP, unpublished data) and passing to it several pieces of data, one being an array containing the digitized IRIGE serial time code. The time returned is then used to overwrite the inaccurate PC time. In addition, the sample rate is computed from the IRIGE time code.

To run the PC2SUN program, type "PC2SUN" at the DOS command level. The program asks for the name of the "list" file (see below) and the path to the directory containing the list file and the PC files to be converted. The path name should end with a "\". Next the program asks for the name of the directory beneath which subdirectories will be created for the output files. For example, if the OC event file is "89092799.wva", the created subdirectory would be named "89092700". A new sub-directory is created for each event.

The list file mentioned above is a DOS directory listing of the waveform files to be converted. This file is conveniently created by using DOS ">" re-direction in conjunction with the DIR command and giving the name of the list file. For example, to create a list file with all the waveform files, type: "dir *.wvm > pc2sun.lst". All files which should not be converted must be eliminated from the list file. It is not necessary to eliminate the heading and footing which DOS provides. PC2SUN ignores all lines which do not start with the number eight or nine.

As the program runs, the user is kept informed of its progress. The header conversion takes less than one second, but the data conversion, which involves multiple reads and writes takes several minutes per input file. The program outputs to the screen the current channel and buffer being converted. Since the data is multiplexed in 512-byte blocks, each block must be processed one time for each channel of data. This is the process the user sees on the screen as the program runs.

It should also be noted that PC2SUN works to create a file compatible with the seismic analysis program XPICKS. It can be modified to produce a file for the original SUNPICKS analysis program. Another change from the SUNPICKS analysis program. Another change from the SUNPICKS format is the inclusion of a "data type" variable which can be set to 1 or 7 to indicate either float or short data respectively. Finally the first three "extra" spaces indicate analog-to-digital (A-D) pin number, the bias value removed from the data samples and millivolts per count for the A-D convertor. The last item is set equal to zero if it is undefined.

CONVERTING PC FILES TO ASCII

Occasionally, it is useful to be able to view a data file directly using a text editor or other DOS utility. This can be invaluable when problems occur or data needs to be examined point-by-point. The program "BIN2ASC" takes as input a PC file (as shown in Figure 1) and outputs a set of files, one per channel, in ASCII (text) format.

To run the program, type "BIN2ASC" at the DOS command level. The program creates a set of text files in the form "W_ASC.X", where X is the channel number. The most important information on the file's contents are output to the screen as the program runs. The text files are put in the same directory as the PC input file. The ASCII files are much larger than the binary files.

OBTAINING THESE PROGRAMS

The programs including both source and executable code are available on a 5 1/4" DOS formatted 360 Kbyte disk. To obtain a copy send a blank disk with a stamped, self-addressed envelope to the author (U.S.G.S, 345 Middlefield Rd., MS 977, Menlo Park, CA 94025) with a brief note.

REFERENCES

Toolbox for seismic data acquisition, processing and analysis, (1989) Edited by W. H. K. Lee, IASPEI Software Library Vol. 1, Seismological Society of America, El Cerrito, CA