Computer Corporation of America

575 Technology Square
Cambridge
Massachusetts 02139

617-491-3670

Assistance in the Design of the SWF System

CCA-79-13

Donald E. Eastlake, III

Joanne Z. Sattley

32p.

January 30, 1979

This research was supported by the Advanced Research Projects Agency of the Department of Defense under Contract No. N00039-78-C-0246, ARPA Order 3540.
The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the Advanced Research Projects Agency or the U.S. Government.

Table of Contents

1. Introduction

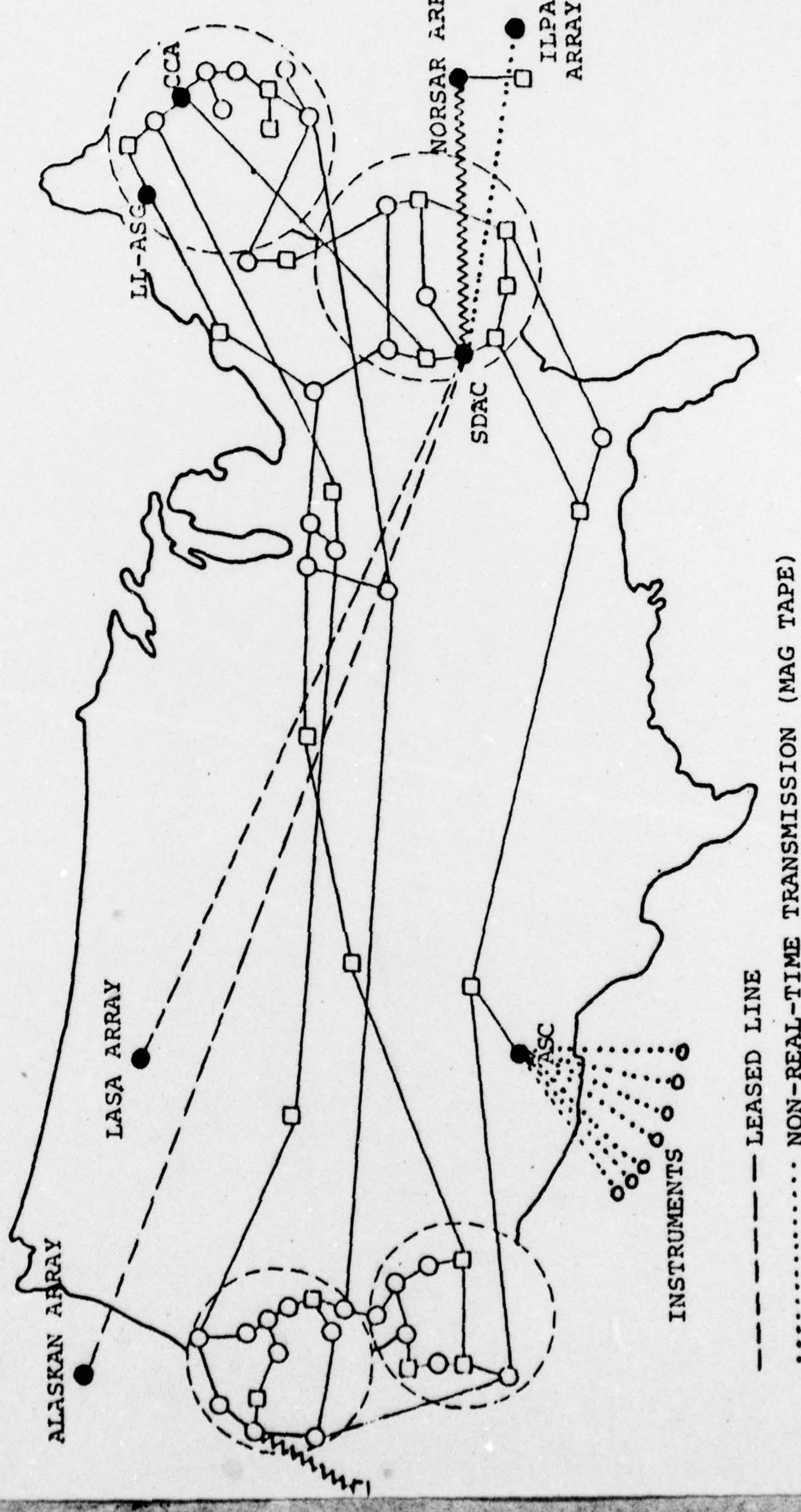
During 1978, Computer Corporation of America assisted the Vela Seismological Center (VSC) and the Seismic Data Analysis Center (SDAC) in the design, creation, and management of the Signal Waveform Files (SWF). The SWF system is intended to make all of the recent "interesting" seismic data of the Vela Seismological Network available in the Datacomputer. This work is continuing in 1979 under a follow on proposal.

The Datacomputer is a unique network data utility [MARILL and STERN] developed and operated by Computer Corporation of America for the Defense Advanced Research Projects Agency. It is designed for support of very large databases and shared remote access by heterogeneous computers [CCA]. The Datacomputer provides very large on-line capacity through an Ampex TBM video tape based mass storage system [EASTLAKE]. This mass storage system is presently configured for approximately 200 billion bits of on-line capacity. Data that has been stored into the system in excess of that amount is kept in off-line TBM tapes.

The Velanet is a worldwide network consisting of seismic instruments, for example, communications channels, data processors, and data storage systems. Figure 1.1 shows the logical configuration of the Velanet at the beginning of 1978. Figure 1.2, provided by SDAC, shows the geographic configuration of the Vela Network. The Datacomputer is the primary storage and retrieval resource in the seismic data activity [DORIN and EASTLAKE].

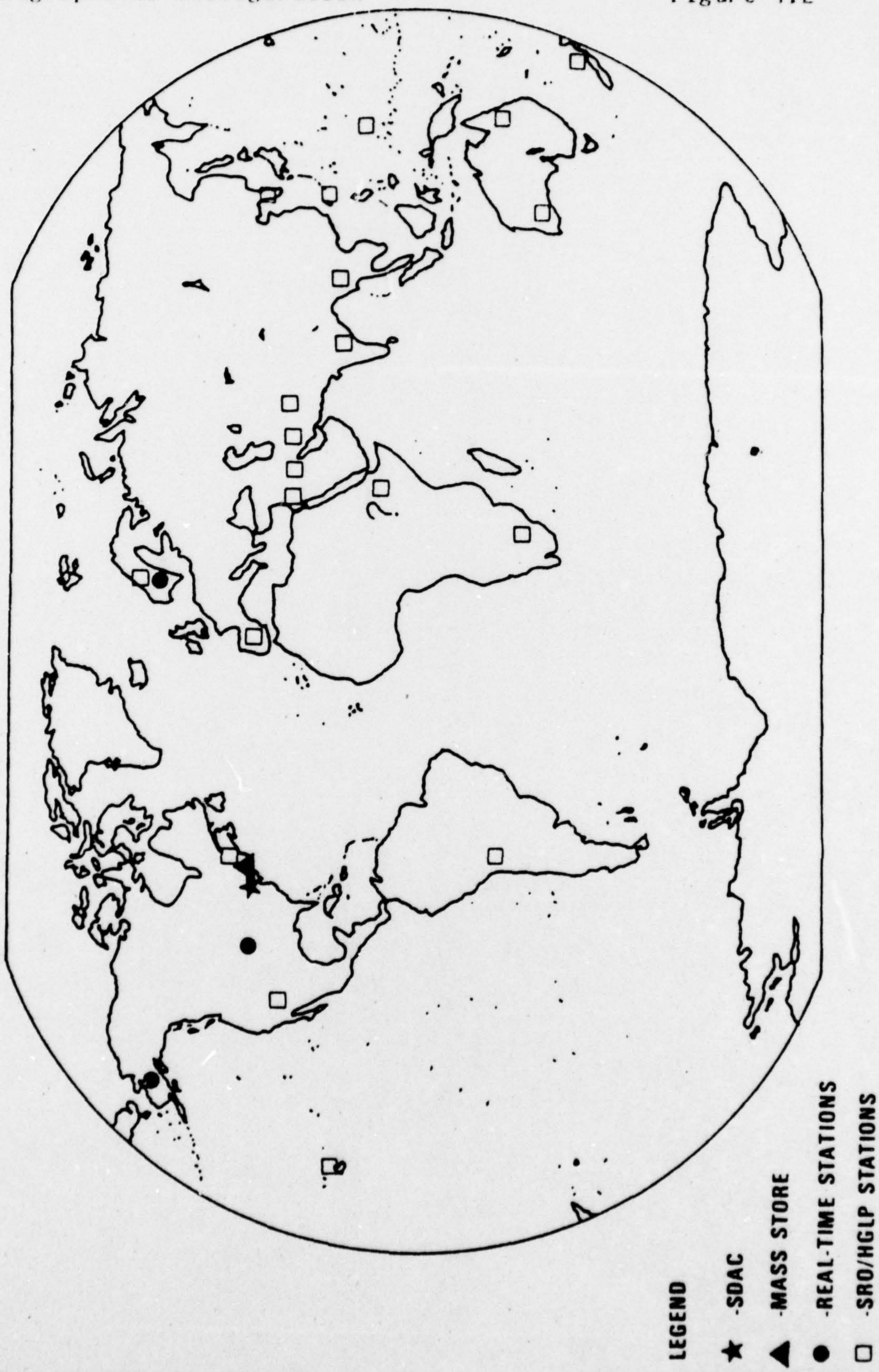
The seismic data stored in the Datacomputer is structured into files of several types. Two of these types, status and event files, are relatively small but valuable and it is presently planned that they be kept on line indefinitely. Status files give information on the status of seismic instruments, their calibration, operationality, and the like. Event files list seismic events of significance and include many pieces of information on each event such as location, depth, severity, method of calculation, error bounds, region, etc. In addition, each event in an Event Summary File is accompanied by a list of "arrivals"; that is, a list of times, seismic stations, and other information that relates the event to the arrival of seismic disturbances at particular seismometers.

Velanet Configuration Figure 1.1



Geographical Configuration

Figure 1.2



Other than the status and event files, there are seismic files in the Datacomputer which consist primarily of time series seismometer readings. For the real time seismic arrays in the Vela network, this data is continuous, with missing data areas only when the real time system was not fully operational. For the Seismic Research Observatory (SRO) data, long period data consisting of one sample per second is recorded and later stored continuously. Short period data, with ten samples per second, is only recorded at the SRO, and later stored in the Datacomputer, if it exceeds a local threshold. The volume of data in these files makes it impossible to keep them on line for more than a couple of months in the current mass storage system configuration before newer data crowds out the old which must be put off-line.

Recently a new type of seismometer data readings file has been created, the Signal Waveform File. It is described in the following section.

2. The Signal Waveform File

The seismic files that are expected to be the most useful for seismological research are the Event Summary Files and the Signal Waveform Files.

The Signal Waveform Files will contain seismic data segments that relate directly to the detected seismic events in the Event Summary File and also noteworthy unassociated seismic arrivals at particular seismic stations. The files will not contain other seimsic data. This represents a reduction in data of over an order of magnitude. There is a corresponding beneficial increase in the amount of time that the data can remain on-line for a particular mass memory configuration and a decrease in the amount of time it takes to perform certain types of processing over the data.

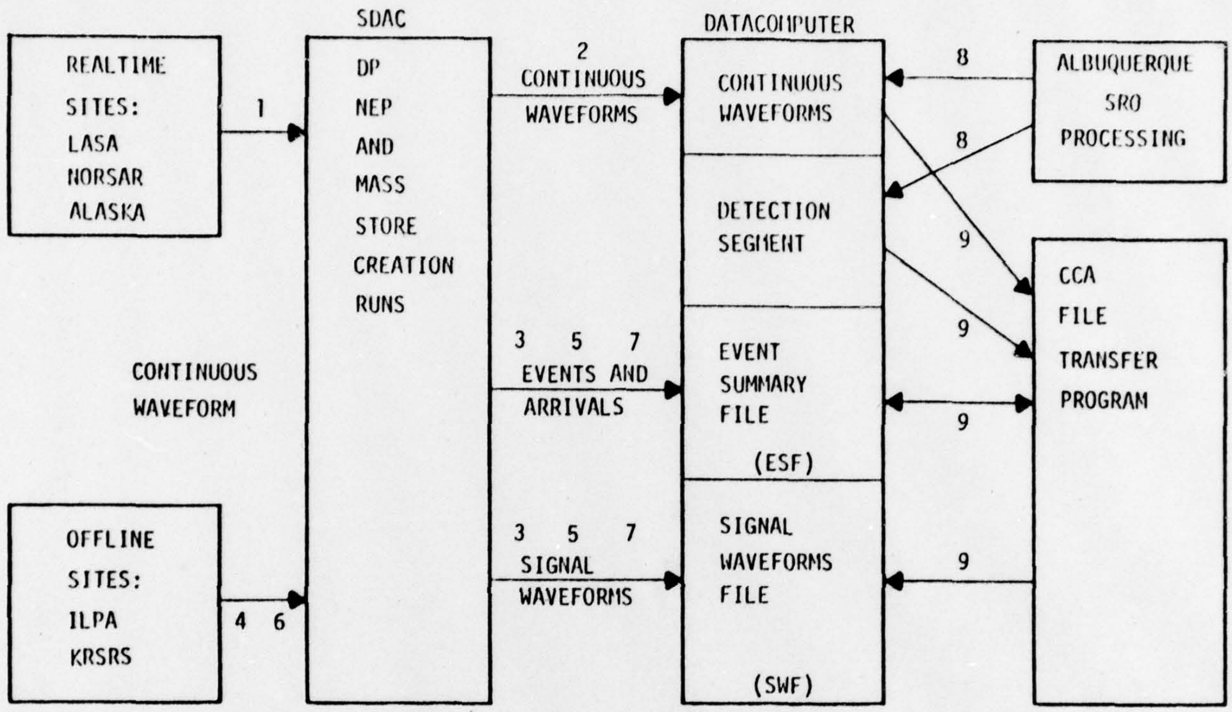
2.1 SDAC Processing

Much of the processing involved in forming the SWF is performed by SDAC. The determination of the events and seismic signal arrival times and stations for the PESF file is SDAC's responsibility. These PESF entries control the subsequent formation of the SWF.

SDAC will also store directly into the SWF in the Datacomputer all of the signal waveforms except those from the non-array seismic research observatories (SROs). The storage SDAC will perform involves the selection of the appropriate data from that arriving in real time from seismic arrays and later the selection and transmission of the array data arriving at SDAC by magnetic tape from ILPA, the Iranian array, and KSRS, the Korean array.

Figure 2.1, provided by SDAC, shows the processing involved from the SDAC viewpoint.

SDAC processing Figure 2.1



1 - waveform data from realtime sites arrives at the SDAC
2 - data from the realtime sites goes to the DATACOMPUTER
3 - ESF/SWF including realtime sjtes signal waveforms go to the DATA COMPUTER
4 - ILPA field tape arrives at SDAC
5 - ILPA signal waveforms go to DATACOMPUTER
6 - KSRS field tape arrives at the SDAC
7 - KSRA signal waveforms go to the DATACOMPUTER
8 - ALBEQUERQUE reviewed SRO data goes to the DATACOMPUTER
9 - CCA FILE TRANSFER PROGRAM completes signal waveforms

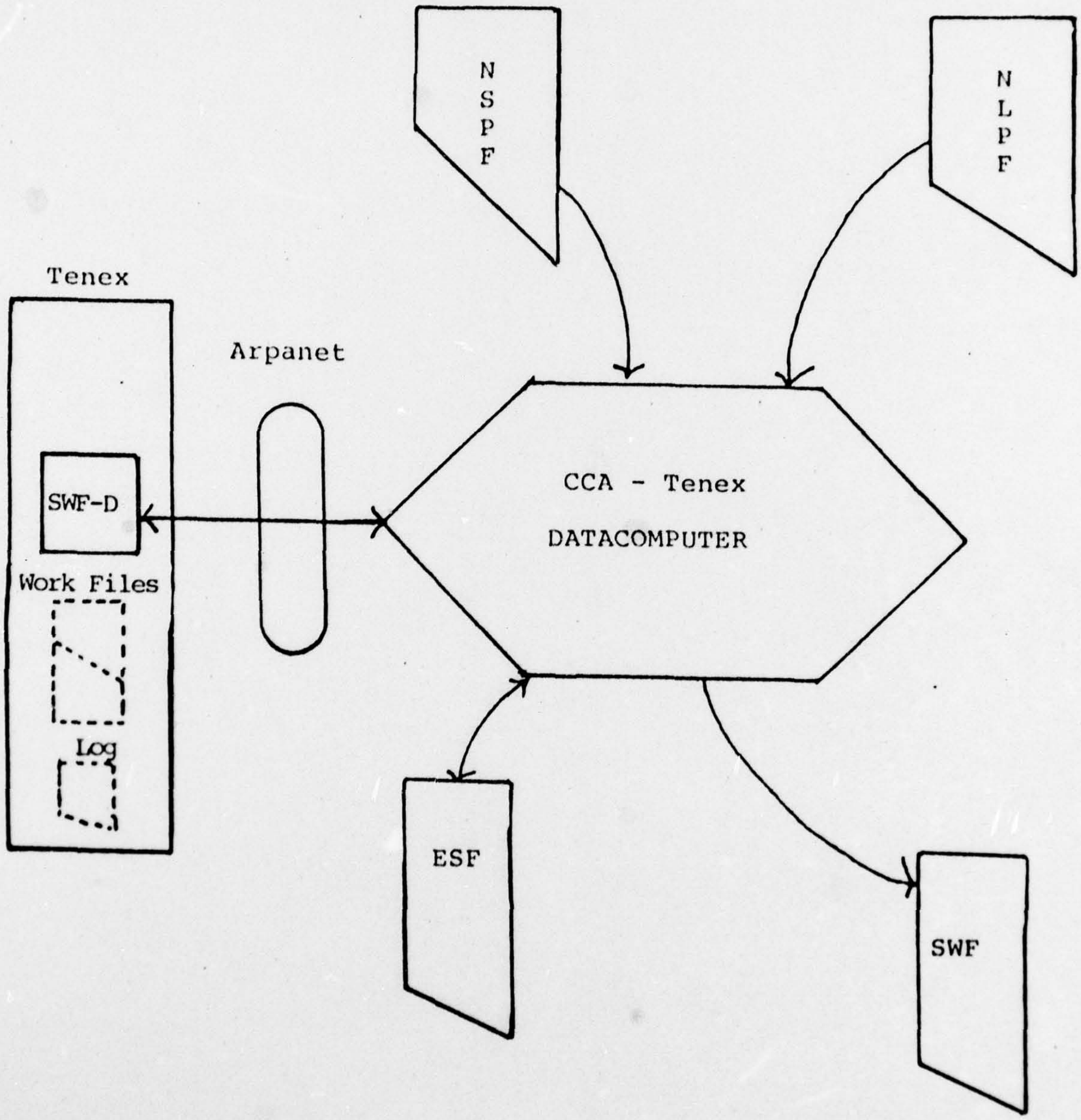
2.2 CCA Processing

CCA processing involved in the creation of the SWF is described in detail in several other technical reports [EASTLAKE and SATTLEY] [SATTLEY a] [SATTLEY b]. In brief, it involves scanning the Event Summary Files for flagged arrivals referring to SRO data, seeking out the data from the SRO files in the Datacomputer, moving the data to the SWF, and updating the ESF to reflect this as shown in Figure 2.2. If the data requested for the arrival does not exist, the ESF flag is reset to an indication that no waveform data will be available.

In order to efficiently locate and extract the discontinuously recorded and stored short period SRO data, the CCA software must first compute a summary file of the available data. Since chunks of short period SRO data are recorded only if the seismic activity exceeds a local threshold at the seismometer involved, the entries in this summary file are referred to as detections.

These short period SRO detections list files are of interest in their own right and are being regularly compiled by CCA, stored in the Datacomputer, and made available to seismic researchers.

CCA & DC Relationship with Seismic Files Figure 2.2



Prior to CCA compilation of these files, the Lincoln Laboratories Applied Seismology Group compiled similar files with less information in them. Because of the Datacomputer's field name matching and selective data extraction features, LL-ASG was able to use the CCA compiled files by simply changing their programs to reference them. No other changes were necessary as the additional data fields that LL-ASG did not want were stripped by the Datacomputer from the data sent to them.

3. Assistance

CCA assisted the Vela Seismological Center and the Seismic Data Analysis Center (SDAC) during the course of this contract, by consultation, in the design, creation, and management of the SWF and in the use of the Datacomputer for that end. This assistance was rendered through meetings held at SDAC, memoranda, telephone consultations, and the like as described below.

3.1 Meeting 16 January 1978

This all day meeting concerned both the SWF and also the specific component of the SWF system that CCA is implementing. This component moves SRO data into the SWF under the control of entries placed in the Event Summary Files by SDAC. CCA was represented by Donald E. Eastlake. The following points were among those covered:

a. There was considerable discussion of the volume of data to be involved in the SWF. CCA commented on the effects of various options on TBM tape usage and on-line retention time.

b. It was decided that CCA would not rotate SRO components (changing them from north-south and east-west to radial and tangental with respect to the event) because the Datacomputer would be much less efficient at doing this than would the specially equipped processors at SDAC.

c. The method of flagging the arrivals in the Event Summary Files which CCA will handle was decided. A field that normally has a "Y" or an "N", to indicate that a waveform is or is not available in the SWF, will be set to "T" to flag arrivals needing processing. The questions as to how much data should be moved into the SWF for the arrival were to be answered by look up in a simple table to be supplied to CCA by SDAC.

d. It was decided that there would be only one type of SWF file. This would initially have the real time waveforms and the SRO and non-real time data would be appended. Earlier plans had called for a preliminary and final version of each SWF file.

e. Whether there would be only one type or both preliminary and final versions of the Event Summary File was left undecided. CCA urged that the initial event files, of which there are one per day, be later condensed into monthly final event summary files. These fewer monthly files would be easier to deal with. In the process of transferring from preliminary to final ESF, predicted arrivals for which waveform data will not be available can be deleted.

f. There was considerable discussion of scheduling and testing for developing the SWF system. Using the SDAC-UNIX system for retrievals to independently test the output of CCA's component of the SWF system was suggested.

3.2 Meeting 13 April 1978

The primary purpose of this meeting, which was held at SDAC in Alexandria Virginia, was to review a Design Plan Draft for the SWF prepared by Joe Greenhalgh of Teledyne Geotech. CCA was represented by Donald E. Eastlake and Joanne Z. Sattley.

After general discussion of SDAC's plans, an explanation was given of CCA's plans to implement its part of the system (as described in Section 2.2 above).

Numerous details of the SDAC plan were decided, many on the basis of seismological considerations. All questions that came up concerning the volume of data to go into the SWF were initially decided in favor of more data.

CCA urged SDAC to get some real predicted SRO arrival entries into the ESF file as soon as possible so assist our development of our part of the SWF system. (Such entries were not successfully entered into the ESF by SDAC until 1979.)

A question arose as to one piece of information to be included in the SWF with the data CCA is to move into the

SWF. This is a scale factor giving the value, in nanometers of displacement, of the least significant bit of the seismic data. SDAC is required to keep up to date on the scale factors of all seismic instruments involved, which change at unpredictable intervals, but CCA personnel generally have no need for this information. It was agreed to have SDAC include the scale factor in an ESF field when it stores arrivals for CCA to handle. This field would be copied to the SWF and then overwritten when CCA updates the ESF to reflect its completed handling of the arrival.

Another question brought up was whether or not abbreviations could be used in the arrivals SDAC will store in the ESF. There are cases where three seismic channels always appear in conjunction and where the waveform data will either be available for all or none of the channels. It was agreed that, in such cases, the use of abbreviations was reasonable and beneficial.

SDAC issued a document on 31 July 1978 based on this meeting and on the design draft circulated at it [GEOTECH].

3.3 Memoranda

Two memoranda were produced by CCA in assistance of the seismic Datacomputer application.

The first was "Comments on the SDAC Mass Store Retrieval Guide", January 17, 1978, by Donald Eastlake. It contained 6 pages of detailed comments on the 31 January 1977 edition of the SDAC Mass Store Retrieval Guide. It particularly urged SDAC to utilize appropriate intermediate programs to access the Datacomputer rather than to use direct human interaction as they were. The following is quoted from this comments memorandum:

> "More to the point, perhaps, this guide [SDAC Mass Store Retrieval Guide] still suffers from an error in the philosophy of how to go about giving ordinary users access to data in the Datacomputer. For example, in 1-2-4-3 and other sections it implies that users need to know about datalanguage. While some users' program and those who wrote it need to know all about datalanguage, the right way to provide access is to provide a program that has a convenient interface (terminal interactive, batch command file drive, or whatever else is appropriate) for the user and that composes the datalanguage requests. This program would respond to, and in some cases translate for the user, datalanguage replies."

The second memorandum was "TBM drive allocation and the SWF", 20 March 1978, by Donald Eastlake. It was a brief two page note that discussed TBM capacity in comparison with expected SWF data volume and reviewed the results of all reasonable strategies for TBM drive allocation to various expected data streams.

3.4 Other Assistance

CCA remained in periodic telephone contact with the Vela Seismological Center, the Seismic Data Analysis Center, and the Lincoln Laboratories Applied Seismology Group. On occasion, off line TBM tapes were temporarily mounted for LL-ASG and assistance was given them in tracking down some Datacomputer user problems.

Telephone contact with VSC and SDAC was primarily related to the SWF. Through such contacts it was decided later in the year to abandon the table look up approach to determining how large a data window CCA was to move into the SWF for an arrival. Instead, the amount of data before and after the expected arrival time will be indicated in ESF fields that are overwritten after successfully moving the indicated data into the SWF.

It was also decided to remove inversions from the fields in the Event Summary File that would be updated by CCA after moving data to the SWF. CCA had previously recommended that these and some other inversions be removed as unnecessary and not worth the additional computation they invoke. Their removal was necessary to make these inner list fields updatable.

References

[CCA]
Computer Corporation of America, Datacomputer Version 5 User Manual, 575 Technology Square, Cambridge, Massachusetts 02139, July 1978.

[DORIN and EASTLAKE]
R. H. Dorin and Donald E. Eastlake, III; "Use of the Datacomputer in the Vela Seismological Network", Technical Report CCA-77-Ø8, Computer Corporation of America, April 15, 1977.

[EASTLAKE]
Donald E. Eastlake, III; "Tertiary Memory Access and Performance in the Datacomputer", Technical Report CCA-77-11, Computer Corporation of America, June 30, 1977.

[EASTLAKE and SATTLEY]
Donald E. Eastlake, III, and Joanne Z. Sattley, Program Design for SWF-D: The Signal Waveform File Demon, Computer Corporation of America, June 30, 1978.

[GEOTECH]
Teledyne Geotech, Velanet ESF/SWF Processing, Contract No. F08606-78-C-0007, 31 July 1978.

[LINCOLN]
Lincoln Laboratories - Applied Seismology Group, ESD-TR-78-64 Semiannual Technical Summary, Seismic Discrimination, 31 March 1978, Lincoln Laboratories, Massachusetts Institute of Technology, Lexington, Massachusetts 02173.

[MARILL and STERN]
Marill, Thomas; and Stern, D.H. "The Datacomputer: A Network Utility", Proceedings AFIPS National Computer Conference, AFIPS Press, Vol. 44, 1975.

[SATTLEY a]
Joanne Z. Sattley, "Implementation and Test of the SWF-D: The Signal Waveform File Demon", Technical Report in progress, CCA-79-Ø9, Computer Corporation of America, January 1979.

[SATTLEY b]
Joanne Z. Sattley, "Program Listings of the SWF-D: The Signal Waveform File Demon", Technical Report CCA-79-10, Computer Corporation of America, January 1979.

A. PESF File Description

The following is a prototype for a one day Preliminary Event Summary File. The monthly Final Event Summary File has the same description except that it has space for more entries.

```
SDAC.VELANET.PROTOTYPES.PESF PORT LIST (10,50,300),B=32
EVENT STRUCT,B=32
   EINDEX BYTE,V=I
   EVENTNUM STR(9),B=8,I=D
   ORIGIN STRUCT
      DATE INT(6),B=8,I=D       /* OYYDDD */
      TIME INT(8),B=8      /* HHMMSSCC */
   END
   STANDEVORIGT INT(4),B=8
   NSTA INT(3),B=8
   HYPOCENTER STRUCT
      COMPSOURCE STR(1),B=8,I=D
      EPICENTERCOMP STR(1),B=8,I=D
      LATITUDE STRUCT
         LAT INT(5),B=8   /* XX.XXX DEGREES */

         HEM STR(1),B=8,I=D
      END
      LONGITUDE STRUCT
         LONG INT(6),B=8   /* XXX.XXX DEGREES */
         HEM STR(1),B=8,I=D
      END
      NSTA INT(3),B=8,I=D
      DEPTH STRUCT
         DEPTH INT(3),B=8   /* XXX KM */
         STANDEV INT(3),B=8   /* XXX KM */
         METHOD STR(1),B=8
      END
      CONFIDREGION STRUCT
         SMAJORAXIS INT(5),B=8   /* XXXX.X KM */
         SMINORAXIS INT(5),B=8   /* XXXX.X KM */
         ANGLE INT(4),B=8   /* XXX.X DEGREES */
```

```
END
REGION STRUCT
   GEOCODE STR(3),B=8,I=D
   SEISNUM STR(2),B=8,I=D
END
BODYWAVE STRUCT
   MBMAG INT(3),B=8    /* X.XX */
   STANDEV INT(3),B=8
   NSTA INT(3),B=8
END
SURFACEWAVE STRUCT
   MSMAG INT(3),B=8    /* X.XX */
   STANDEV INT(3),B=8
   NSTA INT(3),B=8
END
LOCALMAGNITUDE STRUCT
   MLMAG INT(3),B=8
   SOURCE STR(1),B=8
END
PARAM STR(20),B=8
NPHASE INT(3),B=8
NARR INT(3),B=8
ARRIVALS LIST (,500,999),C=NARR
   ARRIVAL STRUCT
      AINDEX BYTE,V=I
      STA STR(5),B=8,I=I
      CHANTYPE STR(1),B=8
      RATE INT(2),B=8,I=I
      CHANID STR(4),B=8,I=I
      GAIN STR(1),B=8,I=I
      COMP STR(1),B=8,I=I
      ASSOCCONF STR(1),B=8,I=I
      DIST INT(4),B=8    /* XXX.X DEGREES */
      AZ INT(4),B=8    /* STA TO EPI XXX.X DEGREES */
      DATASEGSTART STRUCT
         DATE INT(6),B=8
         TIME INT(8),B=8
      END
      PHASEARR STRUCT
         DATE INT(6),B=8,I=I
         TIME INT(8),B=8
      END
      PHASEID STR(6),B=8,I=I
      PHASENUM INT(2),B=8
      PHASECODE STR(2),B=8,I=I
      AMP INT(7),B=8    /* XXXXXX.X NM PK - PK */
      PER INT(3),B=8    /* XX.X SECONDS */
      RES INT(5),B=8    /* XX.XX SECONDS */
      USAGE STRUCT
         LOCATION STR(1),B=8,I=I
```

```
                    MBUSE STR(1),B=8,I=I
                    MSUSE STR(1),B=8,I=I
                END
                WAVEFORMAVAIL STR(1),B=8
            END
        END
END;
```

B. PSWF File Description

The following is a prototype for a one day Signal Waveform File.

```
SDAC.VELANET.PROTOTYPES.PSWF PORT       /* SEGMENTED DAILY */
   LIST (50,200,2200) ,B=32
      /* 50 DETECTIONS PER DAY * 2 ARRIVALS/SITE * 22 SITES */
      SIGNAL STRUCT       /* ONE PER SITE PER PHASE */
         ARRIVAL STRUCT
                EVENTID STRUCT      /*EVENT IDENTIFIER*/
                   EVDATE INT(5),B=8    /*YYDDD */
                   EVNUM INT(4),B=8   /* NUMBER */
            EVENTNUM STR(9),VE=EVDATE!EVNUM
               END
            INDEX BYTE,V=I
            STA STR(5),B=8,I=D        /* STATION */
            CHANTYPE STR(1),B=8
               /* A,B,S,I=ADAPTIVE BEAM,BEAM,SUBARRAY BEAM,
               INDIVIDUAL INSTRUMENT */
            RATE INT(2),B=8
```

```
        CHANID STR(4),B=8,I=D
            /* BBUU FOR INFINITE VELOCITY BEAMS */
        GAIN STR(1),B=8  /* H,L */
        COMP STR(1),B=8
            /* Z,N,E,T,R,1,2,3 NUMERICS FOR ALPA */
        BEAMAZ INT(5),B=8
            /*  0-360 DEGREES: XXX.XX DEGREES */
        BEAMVEL INT(5),B=8   /* XXX.XX DM/SEC */
        DETECTCONF STR(1),B=8
            /* A,B,C, = SIGNAL, S/N ABOUT 1, S/N LESS THAN 1
            AND SIGNAL SUSPECT; ANALYST SAYS */
        CLIPPED STR(1),B=8 /* Y,N, ANALYST SAYS */
        ARVLPAD STR(1),B=8 /* PAD BYTE */
    END
    DATASEGMENT STRUCT
        START STRUCT
            DATE INT(6),B=8       /* OYYDDD */
            TIME INT(8),B=8       /* HHMMSSCC */
        END
        NSAMP INT(4),B=8
            /* DATA SAMPLES IN DATASEGMENT */
        DATAFORMAT STR(1),B=8
            /* I,G: INTEGER, GAINRANGED */
        SCALEFACTOR INT(8),B=8
            /* XXX.XXXXX NANOMETERS PER LEAST SIGNIFICANT BIT */
```

```
            DATAPAD STR(1),B=8
            TIMESERIES LIST(,1600),B=16,C=NSAMP
               DATA STRUCT
               DINDEX BYTE,V=I    /*DATA POINT INDEX*/
               DATUM BYTE,B=16
               /* SP: 20 SEC OF NOISE FOLLOWED BY 60 SEC OF SIGNAL
               FOR 20 SAMPLES/SEC DATA.  SP WITH 10 S/S WILL
               USE ONLY 800 16 BIT BYTES.  LP: 300 SEC OF NOISE
               FOLLOWED BY 1300 SEC OF SIGNAL. */
               END
         END
      END;
```

C. SPDET File Description

The following is a prototype for a month file of SRO short period data detections. (See section 2.2.)

```
SDAC.VELANET.PROTOTYPES.SPDET PORT LIST(0,3500,35000),
                                          IA=51,ID=1200,II=5000
 DETECTION STRUCT INDEX BYTE,V=I
  STA STR(5),I=D
  STANDEX BYTE
  SDATE INT(6),I=D
  STIME INT(8)
  SINDEX BYTE
  EDATE INT(6)
  ETIME INT(8)
  EINDEX BYTE
END;
```