

AD A037460

Computer Corporation of America
575 Technology Square
Cambridge, Massachusetts 02139

12

6

Datacomputer Support of Seismic Data Activity.

Final Technical Report: 22 Apr 76 - 31 Dec 76.

April 22, 1974 to December 31, 1976

11 28 Feb 77

12 22 p.

15 MDA903-74-C-0227
✓ VARP/ Order - 2013

DDC
RECEIVED
MAR 29 1977
REGISTERED

OW A

February 28, 1977

This research was supported by the Defense Advanced Research Projects Agency of the Department of Defense, Contract No. MDA903-74-C-0227. 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 Defense Advanced Research Projects Agency or the U.S. Government.

DISCUSSION DOCUMENT A
Approved for public release;
Distribution Unlimited

387285

JP

DDC FILE COPY,

DATA COMPUTER SUPPORT OF SEISMIC DATA ACTIVITY
Final Technical Report

April 22, 1974 to December 31, 1976

Contract No. MDA903-74-C-0227

ARPA Order No. 2613

February 28, 1977

ACQUISITION	
DATE	
TIME	
<i>Letter on file</i>	
A	

Submitted to:

Defense Advanced Research Projects Agency
1400 Wilson Boulevard
Arlington, Virginia 22209

Attention: John Filson

1. Introduction

The goal of this contract has been to provide seismic data storage and retrieval services in a convenient and timely manner to computers on the Arpanet. These services were provided via the Datacomputer, a network data utility developed and maintained by CCA for ARPA under a separate Contract No. MDA-903-74-C-0225.

The seismic data storage and retrieval services provided must meet the following difficult requirements: 1. very large online storage capacity; 2. very high bandwidth across the Arpanet; and 3. real time availability of some data streams. The online storage requirements were met by the acquisition and integration of an Ampex Tera-Bit Memory System (TBM) as described in section 3 below and changes in the Datacomputer and CCA's host TENEX system, on which the Datacomputer runs, as described in section 2 below. The Arpanet bandwidth requirements were generally met through changes in the physical configuration of the network and careful protocol design, as detailed in section 4 below. The real time requirements were met through the acquisition and programming of a reliable minicomputer based Seismic Input Processor (SIP) to collect, reformat, and buffer the real time streams of data and periodically forward this information to the Datacomputer. The SIP is described in section 5 below.

During the course of this contract, CCA was also responsible for coordination of Datacomputer use with the seismic community. This

primarily involved detailed consultation with seismic users on efficient file formats and the construction of test procedures. Coordination activities are described in section 6 below.

The report summarizes activities throughout the contract period. Particular emphasis is placed on the last two months (November and December 1976) since this period has not been described in previous technical reports.

The activities described herein are continuing under a new contract, MDA-903-77-C-0183.

2. The Datacomputer

The Datacomputer is a network data utility developed by CCA and designed to handle large files and communicate with multiple remote using programs over the Arpanet. The Datacomputer is the primary online repository for seismic data.

2.1 Staging

The seismic application, and the large relatively slow access mass memory it uses, lead to the largest change in the Datacomputer: a set of enhancements to "stage" data between secondary disk storage and tertiary TBM storage. The routines which accomplish this are called SDAX. (SDAX originally stood for special disk area index.)

SDAX keeps track of the location of different versions of an active file through maps. A map provides a translation between logical locations in a file version and the physical devices and locations in which the corresponding data resides. For efficiency, sections of a map can be flagged to indicate that space is being reserved but has not yet been written.

All files, active or not, have associated with them a chain of "home file" maps of complete file versions on TBM tape. This chain may be of length one. Only the most recent version is normally accessible, but new versions are occasionally created for redundancy by the Datacomputer.

In addition to the home file map chain, active files have a chain of SDAX maps to some parts of the file which have been staged to secondary storage. SDAX is in charge of moving data from the TBM to disk when unstaged data is referenced by a user, copying back data to TBM tape when the staging area is crowded, and creating and maintaining the various map chains as necessary to reflect this activity. SDAX coordinates the case of two users referencing the same unstaged data and assures that only one copy is staged. In the common case of several completed updates occurring to part of a file while staged, several SDAX maps will be created. For efficiency, SDAX will merge maps for old SDAX versions that have no readers and thus reclaim staging device space. Furthermore, while copying back data to TBM tape, SDAX is able to recover from the rare bad block that is encountered on TBM tape by placing the data that was to have been written there in a newly allocated block and appropriately modifying the file map. To avoid files from becoming overly fragmented by being allocated in small pieces and by bad block recoveries, SDAX also tries to compact small sections of a file, if it is being copied back to a different location on TBM tape. (No compaction is possible if the TPM tape version is being "updated in place" by a copy back.) Finally, all of the above activity must survive a crash at any point. SDAX takes great care in the order in which it manipulates data and state flags so as to be safely restartable wherever interrupted.

With SDAX, user requests actually run against copies of files on fast access disk with only occasional delays to read data from TBM

tape. Copying back of modified data to TBM normally occurs as a background task not seen by the user. However, the occasional delays the user can face may be quite lengthy if several TBM operations are queued. A set of user messages was added to the Datacomputer to advise users when they are about to be subjected to TBM delays.

During November and December of 1976, the latest enhancement added to SDAX was to allow multiple physical volumes to be used for staged data. This allowed a tripling of the space available for staging file image data from one 3330 type disk to three.

2.2 Datacomputer Evolution

As the Datacomputer evolved to meet the needs of the seismic application it went through several versions. Version 1 was a disk based system available from August 1975 to October 1976. An experimental mass memory based system was available in parallel with Version 1 in August and September 1976. Based on the lessons learned through seismic use of this experimental Datacomputer, the Version 2 Datacomputer was developed and became the first generally available TBM based Datacomputer on October 1, 1976.

Among the new features of Version 2 were a general file backup utility and volume interlocks. The file backup utility is used to automatically copy most permanent files for redundancy. It is also used either automatically or manually to copy parts of files that are on worn areas of TBM tape. Volume interlocks are

necessary to avoid a TBM drive spending excessive time seeking back and forth among different areas of a tape. Version 2 provides a volume lock which is seized for a user between natural interruption points and prohibits other users from interfering on a drive.

The most recent Datacomputer version, Version 3, became available in January 1977. A major new capability of this version is the File Group feature. Because of the enormous volume and continuous nature of the seismic data streams, it became apparent that they would have to be divided into sequences of physical files. These files are divided by month or day, and by the type and source of the data stream. The File Group feature was designed and implemented to manipulate these sequences of physical files conveniently.

The File Group feature enables a user to create a pseudo-file called a "group". A group is a collection of physical files. A user may include or exclude files from a group and may specify logical constraints on the contents of each group member. For example, file X in group Y may be defined to contain "date" values only between 1 October 1976 and 31 October 1976. A retrieval request against the group acts similarly to a series of requests against its constituent files. However, not all the physical files in a group will necessarily be referenced in processing a request. The Datacomputer checks the logical constraints of each file against the conditions given in a retrieval request and avoids those physical files that are constrained to have no data

meeting the request. In the seismic application, where groups will contain hundreds or thousands of files on many TBM tapes, this optimization is a necessity.

Version 3 also offers a new accounting feature which collects information on system usage and summarizes it for certain billable nodes in the Datacomputer's directory. Among the information collected is file space occupancy in TBM block days and dynamic resource usage including processor time, connect time, network traffic, and secondary and tertiary data transfers.

Space charges for all files are aggregated at their superior billable node and dynamic charges for all users go to the billable node at or above their login node. The Datacomputer also collects information on all file references so that the "owners" of files can see the actual extent of use made by those to whom they have permitted access. For maximum flexibility all accounting information is written into a journal file and processed by a separate accounting program.

2.3 CCA TENEX

The Datacomputer runs on the CCA host system which uses a version of the TENEX operating system. Changes to this operating system were made necessary by the seismic application.

CDC 3330 type disks had been installed on the CCA host computer previously for Datacomputer use. In 1975, track-at-a-time input and output was implemented to these disks in preparation for their

use in staging TBM data. This change made available a three-fold increase in bandwidth over the record-at-a-time disk I/O that had previously been the only type available.

Also in 1975 TENEX was modified to use more than 262,144 words of main memory. This expanded memory was necessary for TBM buffering.

In 1976 CCA TENEX was further modified to improve its Arpanet and TBM interfaces. The Arpanet interface was modified by increasing the number of buffers available and the number of network connections that can exist simultaneously. The TBM interface was improved by adding system calls to activate additional TBM features and increasing the fault tolerance of TENEX with respect to TBM errors.

3. The Mass Memory Subsystem

To provide on-line storage approaching the large amount required, a mass memory, the Ampex Tera-Bit Memory System (TBM), was acquired and integrated into the Datacomputer.

3.1 Hardware Configuration and Site Work

The Datacomputer TBM consists of four tape transports, one transport drive, one TBM Data Channel, a System Control Processor, and a channel interface unit (CIU). The CIU presents a standard IBM 370 block multiplexor channel device interface to the CCA TENEX system which uses a Systems Concepts, Inc., SA-10 to simulate such a channel.

Each transport holds one TBM tape on-line with nearly 50 billion bits of storage capacity. Each TBM tape has a unique number associated with it and is pre-formatted and subdivided into fixed numbered blocks. 43,800 of these blocks, per tape, are user accessible. Other blocks are reserved for hardware maintenance. Besides its number, each block has associated with it, in a separately recorded "tally track", a file data identification number which is used for block address error checking and various other information including counts of operations performed to the block.

Extensive site work was involved in making the CCA computer room suitable for the TBM. This included additional air conditioning and electrical capacity as well as work on the walls, floor, and ceiling.

3.2 Software Specifications

In mid 1975, Ampex prepared initial software specifications based on the hardware connection of the TBM system as described above. Ampex's initial specification was defective in failing to provide any way to directly read tally track information and in failing to provide any way to execute certain error recovery steps automatically under the control of CCA TENEX. Since the Datacomputer is to have centralized automatic error recovery and substantial unattended operation, these were serious defects.

Ampex added a tally read command and the other defects were finally resolved by an agreement under which Ampex would provide two enhancements to the TBM separately after the primary software was accepted. These enhancements are called Automatic Alignment and Read Recovery. Automatic Alignment activates an automatic sequence which adjusts all the read parameters for a drive to try to optimize reading for the tape currently mounted on it. Read Recovery automatically sequences through a series of steps in an attempt to recover data from a block which can not be read normally.

These enhancements had not been successfully provided by the end of 1976 and CCA continues to withhold funds from Ampex pending their completion.

3.3 Delivery, Testing, Acceptance, and Maintenance

Delivery of the TBM was originally contracted for August 15, 1976. In July 1975 Ampex informed CCA that, due to problems with the CIU, delivery would be made in January 1976.

After further slips, the TBM was delivered to CCA in February 1976. By mid 1976, some data had been successfully transferred to and from each of the four tape transports. Extensive negotiations with Ampex led to agreement in June 1976 on an exact acceptance test procedure.

A one week continuous acceptance test was concluded July 31, 1976. During this test, several hundred thousand operations were performed on all four drives. The hardware behaved satisfactorily, however a number of Ampex software problems were encountered. It was agreed that a further software acceptance test should be held. This further primary software acceptance test was passed in early September 1976.

To facilitate testing and maintenance of the TBM, three user programs were written to run under CCA's TENEX system.

(1) A general device test and exercise program called TESTIT. This program uses the non-standard-device routines in CCA TENEX to interact with the TBM and allows numerous patterns of test activity and test data to be operator-specified.

(2) A program to produce formatted printouts of TBM internal dumps on DECTape. This program is intended to aid Ampex in diagnosing problems with TBM internal software.

(3) A TBM checkout program that is normally run daily on all drives after TBM maintenance and on each drive when a new tape is mounted or when a high error rate is being encountered.

In October 1976 this last TBM checkout program was modified so that its TBM use was interruptible at several points and it could be run concurrently with the Datacomputer.

At the end of 1976, the worst operational problem remaining with the TBM was an occasional corruption or improper writing of the tally track information for some blocks. If this occurs when a block is about to be written by the Datacomputer, it can usually recover by writing the information elsewhere. If this occurs when a block is being read, manual intervention to normalize the tally contents is necessary.

4. Arpanet Considerations

Successful use of the Arpanet is vital to the seismic use of the Datacomputer. All of the data to and from the Datacomputer passes through the network and the local node through which CCA is connected to the network. The continuous flow of real time data into the SIP and occasional bursting of this data from the SIP into the Datacomputer place particularly heavy demands on the network. Problems have been encountered with network bandwidth and network hangups.

Initial tests indicated a bandwidth through the network node local to CCA of 50 to 80 kilobits per second rather than the 300 kilobits expected from BBN reports. Fifteen to twenty kilobits per second of seismic array data will flow continuously from the CCP to the SIP. To provide for catch-up and error recovery it is desirable that the SIP-Datacomputer data path be capable of operating at 4-5 times this rate. The 4 kilobits per second of non-array seismic data sent directly to the Datacomputer is subject to a similar compression factor. Thus, excluding seismic data retrievals, non-seismic Datacomputer traffic, and network through traffic, at least 140 kilobits per second capacity is necessary.

Investigation of the previously mentioned saturation of the CCA IMP at 50 to 80 kilobits by CCA and BBN led to the conclusion that the problem was lack of both processor power and buffering in the CCA IMP. Some processor power was being sapped by direct terminal

line handling since the CCA IMP also served as a TIP with direct dial up lines. Buffering was being impacted by the fact that the CCA IMP had a VDH connection to Lincoln Laboratories which required memory space for the VDH code and dedicated VDH buffers.

To help this problem, the CCA IMP was replaced by a model 516 IMP and its direct terminal lines moved away on September 15, 1975. This produced a factor of two improvement overall and solved the local message bandwidth problem at that time. Long distance messages are more sensitive to buffer availability and still presented problems.

In late 1976, the Lincoln Laboratories VDH was moved away from the CCA IMP providing temporary relief by freeing more buffer space.

Some network hangups and buffering problems were traced to the PLURIBUS IMP at SDAC. In some cases the PLURIBUS was reserving excessive numbers of buffers at the CCA IMP so that other traffic was frozen out.

Within the next year, seismic bandwidth is expected to increase significantly and it would appear that the only long term solution is the installation at CCA of an appropriately configured PLURIBUS IMP. Processor power and buffer memory can be added modularly to a PLURIBUS.

5. The Seismic Input Processor

Seismic data from seismic arrays is routed through a Communications and Control Processor (CCP) in Alexandria, Virginia and sent to CCA in real time. This data must be accepted within a few seconds 24 hours a day. The Datacomputer, which runs on a large general purpose computer and requires daily down time for preventive maintenance, cannot be available on this schedule.

To deal with this problem the SIP, a small reliable dedicated computer system, was developed by CCA. The SIP hardware is primarily composed of a Digital Equipment Corporation PDP-11/40 computer with two 3330 type disk drives and an Arpanet interface. The real time data stream is accepted by the SIP which reformats it and buffers it on its disks. The SIP then periodically bursts its accumulated data to the Datacomputer.

The SIP hardware was delivered early in 1975. Arpanet connection was accomplished in June 1975.

5.1 CCP <—> SIP Protocol

The SIP uses standard host-host protocol to communicate with the Datacomputer but uses a special protocol for the real time path to the CCP. This special protocol eliminates the normal host-host handshaking and connection setup overheads and results in simpler, more efficient communication. Furthermore, the special protocol eliminates the standard protocol's requirement that no more than one message be in the network in one direction at a time. This

modification increases bandwidth and decreases network blocking.

By mid-1976 three shortcomings had become apparent in the protocol. First, the maximum efficiency was not yet being achieved in network usage as logical messages and physical messages were forced to correspond. As a result physical messages were not their maximum size and packets less than full size were being sent through the net. Second, the CCP was implemented in such a way that the SIP could not stop accepting messages from the CCP without either disrupting the CCP or bringing down its host ready line which disrupted communication with the Datacomputer. This made it hard to debug the SIP-Datacomputer path. A SIP going down message to the CCP had been provided in the protocol but not specified to silence the CCP to SIP path. Third, the unique message ID numbers used were arbitrary sender's choice. If they had been specified as sequential, they would have been of greater use in duplicate and out-of-order message detection.

During November and December 1976, a new protocol was formulated and a meeting held at SDAC to refine it. This new protocol maximizes network efficiency by packing logical messages into full size physical messages and also uses sequential message ID numbers. The new protocol is being implemented and is expected to be put into service in the first half of 1977.

5.2 Software Development

By mid 1975 utilities for loading programs into the SIP and debugging them were developed. The main components of the SIP software system were completed in late 1975 and for the next several months the SIP was extensively used in checkout of the CCP and tests of the seismic network.

By mid 1976, the SIP was fully integrated and hundreds of hours of real time seismic data had been stored in small test files in the Datacomputer.

In the third quarter of calendar 1976, improvements were made to the SIP software. These improvements made it possible for the SIP to survive the failure of one of its two disk drives under all foreseeable circumstances and to continuously monitor the status of its disk drives. Also, the SIP was modified to avoid initializing Datacomputer files and to do several transfers for one Datacomputer file before moving to the next.

During August and September 1976, the SIP stored data into full size files in an experimental Datacomputer and, on October 1, 1976, the SIP-Datacomputer system became fully operational with data being stored into final files in a standard TBM based Datacomputer.

6. Coordination with the Seismic Community

During this contract, CCA has provided coordination with the seismic community on the use of the Datacomputer. In this regard CCA's efforts have been primarily directed to assisting users in designing and using Datacomputer files and to setting up test files for experimental use. The users CCA has assisted include Vela Seismological Center, Seismic Data Analysis Center (and its contractors, Teledyne Geotech, Texas Instruments, and BBN), Lincoln Laboratories Applied Seismology Group, and Albuquerque Seismological Laboratory (and its contractor Lisle Computer).

CCA has also aided the operational use of the Datacomputer by providing an online status service and, in December 1976, by rescheduling all preventive maintenance on the TBM and CCA TENEX to before 9AM Eastern time for the convenience of seismic users.

6.1 Test Files

In late 1974 some ALPA long period array data from the International Seismic Month was stored. In early 1975, two sets of data corresponding to the Preliminary Event Summary File were stored, one with over 110,000 records, for Lincoln Laboratories.

For both of these datasets, the CCA Datacomputer user program called SMART was modified to access the files so experience in using them could be gained.

6.2 File Formats and Usage Assistance

At a meeting in early 1975, the basic organization of the Preliminary Event Summary File and the Preliminary Signal Waveform File and the linkage between them was settled.

The proposed seismic file structures were studied in detail by CCA and in August 1975 CCA made a number of suggestions including uniform use of 8 and 16 bit bytes, fewer inversions for efficiency, and use of the new highly efficient virtual index feature of the Datacomputer. On December 8th, 1975, a meeting was held at CCA at which the file formats were finalized.

As actual files became available in the Datacomputer, CCA assisted users in manipulating and making use of them. In a few cases, the size and complexity of the seismic files caused users to encounter limitations in the Datacomputer. In all cases either the Datacomputer was appropriately expanded or the user was shown a simple way to avoid the limitation.

In November, 1976, CCA gave an intensive one day Datacomputer Training seminar at SDAC for the seismic community.

6.3 Status Reporting

The dispersed and varied seismic users of the Datacomputer created a need for CCA to communicate to them the operational status of the Datacomputer.

A Datacomputer Status server program was developed and installed during November and December 1976, to provide users with Datacomputer status information.

Information supplied includes Datacomputer operational status, information on all active users and availability of service. Notification is also automatically given if the CCA local Arpanet node or CCA TENEX are expected to go down soon or if the Datacomputer system is heavily loaded.