DATACOMPUTER SUPPORT OF SEISMIC DATA ACTIVITY

Computer Corporation of America

Prepared for:

Defense Advanced Research Projects Agency

31 October 1975

DATACOMPUTER SUPPORT OF SEISMIC DATA ACTIVITY

Quarterly Technical Report

August 1, 1975 to October 31, 1975

Contract No. MDA903-74-C-0227
ARPA Order No. 2613

Submitted to:

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

Attention: Dr. Carl Romney

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Computer Corporation of America	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE

DATACOMPUTER SUPPORT OF SEISMIC DATA ACTIVITY

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Quarterly Technical Report, August 1, 1975 to October 31, 1975

5. AUTHOR(S) (First name, middle initial, last name)

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
8a. CONTRACT OR GRANT NO. MDA903-74-C-0227 b. PROJECT NO. c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)	

10. DISTRIBUTION STATEMENT

Distribution of this document is unlimited

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Defense Advanced Research Projects 1400 Wilson Boulevard Arlington, Virginia 22209

13. ABSTRACT

DD FORM 1 NOV 68 1473 REPLACES DD FORM 1473, 1 JAN 64, WHICH IS OBSOLETE FOR ARMY USE.

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
datacomputer						
seismic data base						
seismic input processor (SIP)						
Ampex TBM						

Computer Corporation of America
575 Technology Square
Cambridge, Massachusetts 02139

Datacomputer Support of Seismic Data Activity

Quarterly Technical Report

August 1, 1975 to October 31, 1975

This research is supported by the Defense Advanced Research Projects Agency of the Department of Defense under 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.

Table of Contents

		Page
1.	Summary	1
	1.1 Project Goals	1
	1.2 Technical Status of the Project	1
2.	The SIP	4
	2.1 The Software	4
	2.2 IMP-TIP Swap	4
	2.3 The Hardware	5
3.	Coordination with Seismic Community	6
	3.1 Comments on Seismic File Descriptions	6
	3.2 Meeting on Seismic File Formats	7
4.	The TBM	9
	4.1 Software Specifications	9
	4.2 Site Preparation	9
	4.3 The Hardware	10
5.	The Datacomputer	11
	Glossary	12
	Figure 1: CCA Configuration	3

1. Summary

1.1 Project Goals

The purpose of this project is to support the ARPA-NMRO seismic data activity by providing data storage and retrieval services. The Arpanet will be used as the primary communications channel. As part of the service, seismic data will be (a) recieved from the Arpanet; (b) stored and indexed in the Datacomputer; and (c) made available to computers on the Arpanet in a convenient and timely manner. These services represent a special application of the Arpanet Datacomputer being developed and maintained by Computer Corporation of America (CCA) under Contract No. MDA903-74-C-0225.

The amount of seismic data to be stored requires the addition of a mass memory to the Datacomputer system. An Ampex Terabit Memory System (TBM) with a capacity of almost two hundred billion bits will be installed at CCA in January 1976 to answer this need. Modifications to CCA's computer site are necessary to accommodate the TBM.

The other hardware item vital to this project, besides the TBM, is a small Seismic Input Processor (SIP). The SIP will perform several functions, the most important of which is to continuously collect data over the network, reformat and buffer the data, and, at intervals, generate a datalanguage update request and burst the data to the Datacomputer via the local CCA network node. The CCA TIP has been replaced as the local network node with a 516 IMP to provide the required bandwidth.

1.2 Technical Status of the Project

Project activity can be divided into four areas: (1) SIP development and network bandwidth considerations; (2) coordination with the seismic community; (3) TBM acquisition and integration into the Datacomputer; and (4) seismic related Datacomputer development.

The SIP consists of a DEC PDP-11/40 with 28K core, two RP04 disks and an Arpanet host interface (see Fig. 1). All of this SIP hardware has been installed, connected to the network, and is working properly. Software development is well underway.

In continuing coordination with the seismic community, CCA reviewed and made comments on the seismic file structures transmitted to it by VSC. A meeting was held to further refine these descriptions.

The TBM memory system is being built by Ampex Corporation as a subcontractor to CCA. All hardware components have been fabricated and are undergoing tests at Ampex. The initial TBM configuration will be one transport driver, two dual transport modules, one data channel and a Communications and Control System (CCS). Generally acceptable software specifications for the TBM have been received by CCA. Ampex has informed CCA that delivery of the complete TBM system will be in January 1976.

A number of specific improvements have been made in the Datacomputer and its underlying TENEX operating system for the seismic application. These improvements appear to be adequate for the initial seismic data rate. Further improvements have been designed.

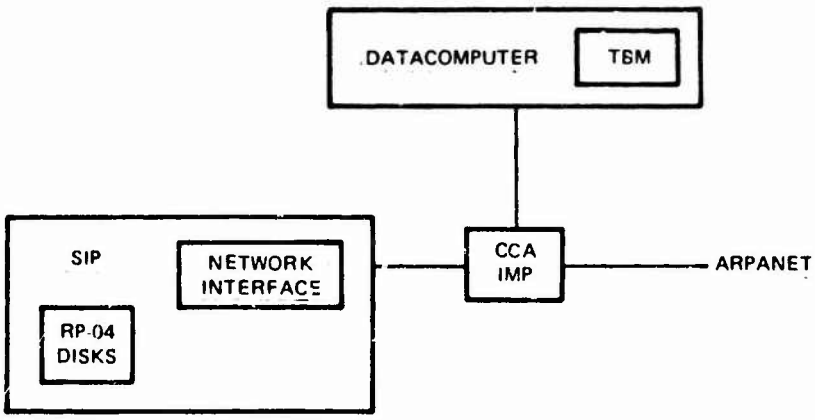

Figure 1--CCA Configuration

2. The SIP

Seismic array data will be collected from the Arpanet, buffered, and reformatted by a small Seismic Input Processor (SIP) which will retransmit the data to the Datacomputer. The SIP is equipped with disk storage adequate for 24-hour buffering of a 15 kilobit per second (kbps) data stream.

2.1 The Software

The CCP-SIP special protocol has been checked out from the SIP end and data was received from the CCP and stored in unparsed form on the SIP's disk. This was done with SIPLAB, a testbed produced by modifying a network interface test program.

A Digital Equipment Corporation (DEC) RSX-11M operating system was installed on August 7. As expected, RSX-11M facilitated the successful development of a Network Control Program (NCP), which has been completed, suitable for the SIP. However, attempts to integrate the rest of the SIP system with the NCP under RSX-11M ran into problems in that the RSX file system takes 4096 more words of memory than is documented. To accommodate this would require not only the purchase of additional memory, but, in addition, since memory would then exceed the direct addressing capability of the SIP computer, memory management hardware and further software to utilize the memory management hardware would be required. One solution to this problem is to proceed with the integration of the SIP by replacing the RSX-11M operating system functions used by the NCP with functions already implemented for the portion of the SIP that communicated with the CCP.

2.2 IMP-TIP Swap

On September 15, the CCA-IMP was replaced by a 516 IMP. A 30% improvement was expected with a 316 IMP after removing the TIP load; however, about a factor of two improvement was found due to the 516 being a faster computer than the 316. Thus the local bandwidth problem has been solved for the initial seismic

data rates. Long distance network traffic is more sensitive to buffer availability and may still present a problem unless the Lincoln Laboratories' VDH is moved to another IMP. Bolt Beranek and Newman Inc (BBN) is also working on reducing the overhead in the network associated with routing messages.

Since several TIP lines were in use at CCA, one as a simple means of communicating with the SIP from CCA-TENEX, some way of supplying more terminal lines at CCA was necessary. Increasing the number of terminal lines on the CCA-TENEX line scanner by adding a DC-10B eight line group seems to be the best way to do this.

2.3 The Hardware

The hardware complement for the SIP is complete except that a permanent status display has not been procured. A 9600 baud terminal is currently being used as a status display. No significant hardware problems have arisen this quarter.

3. Coordination with Seismic Community

In continuing coordination with the seismic community, CCA reviewed the seismic file descriptions communicated to it by VSC, sending comments back to VSC and participating in a meeting concerning the file descriptions as described below. Various tests of the descriptions after this meeting revealed a large number of minor problems, many related to byte sizes, that have been fixed by very minor changes in the descriptions. Most of these byte size problems ultimately stem from attempting to manipulate 8, 16 and 32 bit bytes with maximum efficiency using hardware byte facilities that are sensitive to a 36 bit word size. In considering the difficulty of perfecting the seismic file descriptions, one should keep in mind that they represent at least fifteen different types of files, some with descriptions more than a page long as compared with the largest previous similar application of the Datacomputer (the Arpanet real time survey files produced by MIT-DMS) that involved two different types of files each with a description less than 10 lines long.

3.1 Comments on Seismic File Descriptions

CCA reviewed the seismic file formats submitted by VSC and sent a memo to VSC with suggestions for possible improvements. Some of these suggestions related to efficiency, some to recent and imminent extensions to the Datacomputer that might be of use in the seismic application, and some to clarification of minor points in the organization of the files.

For efficiency reasons, it was suggested that fewer inversions be used and pointed out that some storage was being wasted by certain byte size specifications. It was also pointed out that for the high bandwidth raw data files, it would excessively degrade bandwidth, due to the computational effort involved, to use inversions or large amounts of data conversion and byte realignment and padding. Uniform use of 8 and 16 bit bytes was urged.

Features suggested for use in the seismic files included virtual indexes, which are more general than inversions in some way (see Datacomputer User Manual), the recently introduced ability to give the Datacomputer information about the expected size of a file (as well as minimum and maximum limits), and a minor generalization of the inversion feature. It was suggested that some of the multi-file coordination problems of the seismic application may be helped by the file group feature planned for the Datacomputer.

Finally some suggestions for the directory organization of the seismic files were made along with a number of minor specific changes to certain of the files.

3.2 Meeting on Seismic File Formats

A meeting was held in Alexandria, Virginia on September 11 to try to optimize and settle on seismic file descriptions and a directory organization for them. A number of minor changes were made in the file descriptions to fix problems or enhance capabilities and the following major decisions were made:

(1) The previous "array" raw data files will be broken up into separate files by site. This will make it easier to add and delete sites and make each site file much simpler and easier to manipulate in an optimum manner than the composite file would have been.

(2) To take advantage of the hierarchical file directory of the Datacomputer, the pathnames of the seismic files will be split into type, site (which applicable), year, month, and day (except for lower data rate files that are divided by month).

(3) Virtual indexes were extensively inserted to speed retrieval of data by being able to specify physical positions in files. Physical position, for many seismic files, corresponds directly to the real time of a data item.

4. The TBM

An Ampex Terabit Memory System (TBM) will be part of the Datacomputer to provide the required large amount of on-line storage. The TBM requires site modifications at CCA.

4.1 Software Specifications

New software specifications for the TBM have been received from Ampex. These specifications are generally acceptable for the initial TBM installation at CCA. There are still a few minor problems, particularly regarding the file-id number to be recorded in the tally track for each block and used for error checking.

The new software specifications are, however, inadequate for possible future expansions of the TBM. More dual transport units can be added and more TBM internal data channels for redundancy (only one channel to be used at a time) can be added easily. But, to utilize additional transport drivers or CIU's (for multiple simultaneous data transfers) would require some changes in the software specifications and, according to Ampex, extensive enhancements to the internal TBM software. These matters were discussed by CCA and Ampex personnel at CCA on October 2 and 3.

4.2 Site Preparation

Agreement has been reached on all particulars with Control Data Corporation (CDC), our general contractor, for the site preparation work necessary for TBM installation. Although all of the work, including installation of such new heavy equipment as air conditioning and electric supply transforner will take several weeks, it has been determined that the most disruptive work can be concentrated into one week.

The moving, replacing, and cutting of wall, ceiling and false floor elements, generating metallic particles and filter penetrating plaster dusk, are the most disruptive and potentially damaging of the planned site work. This sort of work will all be concentrated in one week during which time all computer equipment in the area will be turned off and after which the area will be thoroughly cleaned before any equipment is turned on again. This week of down time for site preparation will interfere with other aspects of this project and other CCA activities.

4.3 The Hardware

All of the hardware components of the TBM, as described in the last quarterly technical report, are fabricated and have been integrated. They are undergoing tests at Ampex.

5. The Datacomputer

Experimentation with the type of requests that the SIP will make to insert information in place in seismic array files revealed several inefficiencies that had not previously been noticed in dealing with smaller files. After removal of the worst of these inefficiencies, it was demonstrated that seismic data could be stored faster than the 60 kbps which appears adequate for now.

Two modifications were made to the Datacomputer's underlying TENEX operating system in anticipation of software integration of the TBM. First, track-at-a-time input from and output to the large-scale disk to be used, in some cases, for staging of TBM data was implemented. This represents approximately a three-fold increase in bandwidth over the record-at-a-time disk I/O that had previously been the only type available. Second, TENEX was modified so as to be able to use more than 262,144 words of main memory. More than this previous memory limit has been installed and is necessary for TBM buffers.

Modifications to the Datacomputer's underlying TENEX operating system have been designed to allow efficient handling of 32 bit network connections. Previously it had been assumed that 36 bit (the Datacomputer's word size) network data connections with 4 bits of padding per 36 bits would have to be used between the SIP and the Datacomputer. The seismic data is fundamentally structured in 16 bit units. If this modification is implemented and proves workable, an approximate 10% overall saving in bandwidth between the SIP and the Datacomputer will be made.

Glossary

ARPA	Defense Advanced Research Projects Agency
BBN	Bolt Beranek and Newman Inc
CCA	Computer Corporation of America
CCP	Communications and Control Processor--at SDAC
CCS	Communications and Control System--part of the TBM
CDC	Control Data Corporation
CIU	Channel Interface Unit--part of the CCS
DEC	Digital Equipment Corporation
IBM	International Business Machines Corporation
IMP	Interface Message Processor
Kbps	Kilobits per second
MIT-DMS	Massachusetts Institute of Technology--Dynamic Modelling System
NCP	Network Control Program
NMRO	Nuclear Monitoring Research Office
RSX	Resource Sharing Executive
SDAC	Seismic Data Analysis Center--Alexandria, Virginia
SDAX	Special Disk Area Index
SIP	Seismic Input Processor--at CCA
TBM	Terabit Memory System
TIP	Terminal Interface Processor
VDH	Very Distant Host
VSC	Vela Seismological Center

DD FORM 1-74

CAUTION — REMOVE PROTECTIVE SHEET BEFORE TYPING

"TO BE STORED IN A COOL DRY LOCATION"

MATERIAL INSPECTION AND RECEIVING REPORT

1. PROC. INSTRUMENT IDEN(CONTRACT)	(ORDER) NO.	6. INVOICE NO.	7. PAGE 1 OF 1
MDA903-74-C-0227		DATE	8. ACCEPTANCE POINT D

2. SHIPMENT NO.	3. DATE SHIPPED	4. B/L	5. DISCOUNT TERMS
CCA00024	27JAN76	TCN	

9. PRIME CONTRACTOR CODE 6A046

Computer Corporation of America
575 Technology Square
Cambridge, Massachusetts 02139

10. ADMINISTERED BY CODE S2202A

Mr. J. McDonough
Defense Contract Administration
Services Region, Boston
666 Summer Street
Boston, Massachusetts 02210

11. SHIPPED FROM (If other than 9) CODE FOB:

Same as 9. Above

12. PAYMENT WILL BE MADE BY CODE S2202A

Disbursing Officer
Defense Contract Administration
Services Region, Boston
666 Summer Street
Boston, Massachusetts 02210

13. SHIPPED TO CODE W73QQB

Defense Advanced Research Projects Agy
Architect Building
1400 Wilson Boulevard
Attn: Col D.C. Russell
Arlington, Virginia 22209

14. MARKED FOR CODE

Same as 13

15. ITEM NO.		16. STOCK/PART NO. DESCRIPTION (Indicate number of shipping containers - type of container - container number.)	17. QUANTITY SHIP/REC'D *	18. UNIT	19. UNIT PRICE	20. AMOUNT
0002	AB	Quarterly Technical Report October 31, 1975	2	EA		NSP

21. PROCUREMENT QUALITY ASSURANCE

A. ORIGIN

☐ PQA ☐ ACCEPTANCE of listed items has been made by me or under my supervision and they conform to contract, except as noted herein or on supporting documents.

DATE | SIGNATURE OF AUTH GOVT REP

TYPED NAME AND OFFICE

B. DESTINATION

☐ PQA ☐ ACCEPTANCE of listed items has been made by me or under my supervision and they conform to contract, except as noted herein or on supporting documents.

DATE | SIGNATURE OF AUTH GOVT REP

TYPED NAME AND TITLE

22. RECEIVER'S USE

Quantities shown in column 17 were received in apparent good condition except as noted.

DATE RECEIVED | SIGNATURE OF AUTH GOVT REP

TYPED NAME AND OFFICE

* If quantity received by the Government is the same as quantity shipped, indicate by (✓) mark, if different, enter actual quantity received below quantity shipped and encircle.

23. CONTRACTOR USE ONLY