

ADA 026380

12

6
DATACOMPUTER SUPPORT OF SEISMIC DATA ACTIVITY.

9
Quarterly Technical Report. 1 Feb - 30 Apr 76.

11
30 Apr 76

February 1, 1976 to April 30, 1976

12
14

15
Contract No. MDA903-74-C-0227

11
ARPA Order No. -2613

DISSEMINATION STATEMENT A

Approved for public release;
Distribution Unlimited

Submitted to:

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

Attention: John Filson

DDC
307

Computer Corporation of America
575 Technology Square
Cambridge, Massachusetts 02139

Letter on 2

A

Datacomputer Support of Seismic Data Activity
Quarterly Technical Report
February 1, 1976 to April 30, 1976

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
2. SIP and Arpanet.....	1
2.1 Operations.....	3
2.2 Arpanet.....	3
2.3 CCP-SIP Protocol.....	4
2.4 The Software.....	5
3. Coordination with the Seismic Community.....	6
4. The TBM	7
4.1 Site Preparation.....	7
4.2 The Hardware.....	7
5. The Datacomputer.....	9

Figures:

Figure 1: Seismic Data Flow.....	2
Figure 2: TBM Error Analysis.....	8

1. Summary

The purpose of this project is to provide seismic data storage and retrieval services. These services are to be provided via the Datacomputer being developed and maintained by Computer Corporation of America (CCA) for ARPA under a separate Contract No. MDA903-74-C-0225. The seismic data is to be received over the Arpanet and made available to a variety of computers on the Arpanet.

To provide the requisite services the Datacomputer must be augmented by a mass memory system to provide additional storage and by a small Seismic Input Processor (SIP) to continuously collect data over the network, reformat and buffer it, and periodically burst it into the Datacomputer. An Ampex Terabit Memory System (TBM) with a capacity of two hundred billion bits has been ordered for the mass memory system and will be installed at CCA in early 1976. A DEC 11/40 computer with RP04 disks and an Arpanet interface has been installed as the hardware for the SIP. Development of the SIP software is nearly complete.

Project activity can be divided into four categories: (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. One section below in this report is devoted to each of these categories.

2. SIP and Arpanet

Seismic array data is collected from the SDAC Communications and Control Processor (CCP) over the Arpanet, buffered, and reformatted, all by the CCA Seismic Input Processor (SIP) which retransmits the data to the Datacomputer. The SIP is equipped with disk storage adequate for 24-hour buffering of a 15 kilobit per second data stream. See Figure 1.

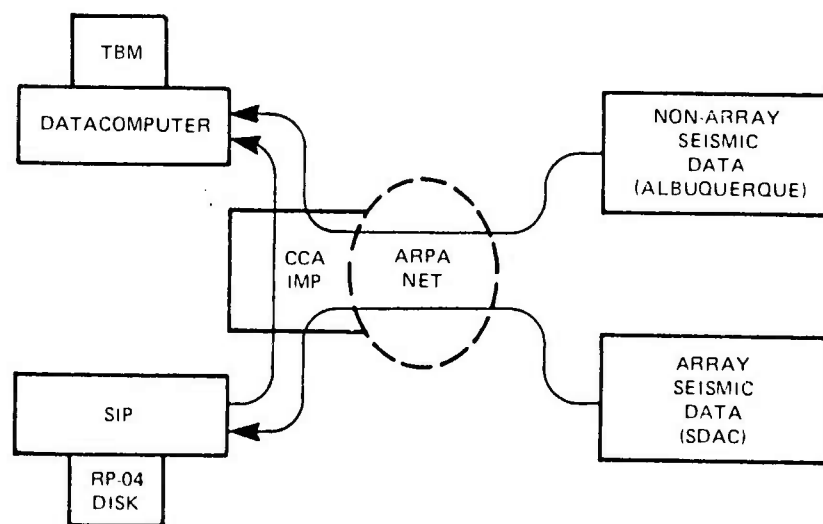


Figure 1--Seismic Data Flow

2.1 Operations

A version of the SIP became operational in this quarter. More than one hundred hours of real seismic array data were received from the CCP, reorganized, formatted, buffered on the SIP's disk, and forwarded to the Datacomputer. Minor improvements are still being made in the SIP.

2.2 Arpanet

For the proper Datacomputer support of seismic data, fairly high Arpanet bandwidths are required. The Arpanet should in principle be able to provide these bandwidths; however, considerable difficulties were encountered during this quarter.

Some of the difficulties appear to have been encountered due to the Pluribus IMP at SDAC. There seem to have been both software and hardware problems with the Pluribus, but there was indication, toward the end of this quarter, that these problems were on their way to being solved.

More intractable problems have been encountered due to limited reassembly buffers at the CCA Arpanet IMP. Achieving high bandwidth over the Arpanet required that large messages be sent which the network breaks into packets and reassembles at the destination. To avoid internal lockups in the network, large messages are not fully accepted until space for reassembly has been reserved at the destination. This scheme places demands on reassembly buffers proportional to the bandwidth of a datastream (number of messages per second) and the delay between source and destination (usually proportional to the number of intervening network nodes).

The CCA IMP has enough buffer space for only three large messages. The two large messages a second from the CCP to the SIP and the traffic from the SIP to the Datacomputer saturate the CCA IMP so that any significant amount of other

traffic causes gross congestion and such difficulties as a high level of incomplete transmissions (messages accepted but not delivered) due to insufficient resources.

This difficulty is likely to become increasingly severe due to seismic data retrievals, direct storage of non-array data from Albuquerque Seismological Center (ASC) into the Datacomputer, increased flow of array seismic data from the CCP through the SIP into the Datacomputer, and other increased use of the CCA IMP.

Short term solutions that have been suggested by CCA for sometime include removal of the Lincoln Laboratories VDH interface from CCA IMP and/or installation of more direct internal network communication lines. However, the extent of relief offered by such short term solutions is limited, and the only long term solution appears to be to install a higher capacity IMP at CCA, presumably a Pluribus IMP.

It should be pointed out that these problems are not within CCA's purview. We must rely on those who have responsibility for proper Arpanet operation to provide adequate service.

2.3 CCP-SIP Protocol

At this point enough experience has been accumulated with the CCP-SIP transmission protocol to point out two ways in which it was inadequately specified in a way that turned out to be disadvantageous to the SIP.

First, the effect of the SIP-to-CCP message (provided for in the protocol) saying that the SIP was going down is not to stop any CCP messages to the SIP. The CCP appears to stop transmitting only if the SIP brings down its host ready line. To avoid disruption to the CCP, the SIP has to assure that the CCP stops sending messages about the time the SIP has stopped listening, and thus the SIP has to bring down

its host ready line even for brief debugging halts. This turns out to be inconvenient, as it destroys information and disrupts communication between the SIP and the Datacomputer.

Second, no information is given in the protocol on how the unique message ID numbers are to be computed. This has no effect on their utility in acknowledgements, which is the primary concern of the protocol. However, these numbers, if they had, for example, been specified to be sequential in one or more series, might have been useful for duplicate and out-of-order message detection, especially for the asynchronous status messages.

Also, in order to alleviate some of the network congestion problems mentioned in 2.2 above, the number of retransmissions for messages was decreased to one. Messages are retransmitted if not rapidly acknowledged and, since network congestion may cause delayed acknowledgements, excessive retransmissions can make the congestion worse.

2.4 The Software

The SIP's software can be divided into three sections: (1) the basic SIP operating system, which is complete and has stabilized, (2) the tasks that manage communications with the CCP by a special protocol and store the information received from it on the SIP's disk, and (3) the NCP and other tasks that manage communications with the Datacomputer and send it information retrieved from the SIP's disk.

It has been demonstrated that, in practice, a very simple algorithm can be used in storing formatted array seismic data on the SIP's disk. Such data is synchronous with one message a second and normally arrives in chronological order. There can be, however, gaps, duplicate messages, messages arriving out of order due to retransmissions, etc. The SIP

is not immediately concerned with such irregularities as it simply stores data for each site and type into hourly files by doing an update in place, reading in the relevant part of the file from disk, modifying it, and writing it back out, before going on to the next message. More complex techniques to accumulate updates in core and reduce disk operations have proven unnecessary.

The SIP periodically sends data in chunks of up to one hour over to the Datacomputer, also using an update-in-place philosophy for the seismic array information. Since there are fixed slots for each site, type, and second of information, no problems are caused by accidentally resending data or sending hours in nonchronological order. At worst, data is overwritten with identical data.

An interesting class of problems encountered in developing the SIP were various sorts of internal buffer lock-ups. One form of congestion involved the SIP filling up with output messages so as to have no room for input while the CCA-IMP was refusing to accept messages until it could deliver one to the SIP. This problem was overcome by more compactly encoding acknowledgement-type output messages and limiting the number that could accumulate before they were discarded. The second lock-up problem involved the SIP accepting a burst of input messages such that no space was left to process and thus free any of the buffers with input messages in them. This problem was solved by modifications to the buffer allocation routines so as not to allocate input buffers where it might cause such a lock-up.

3. Coordination with the Seismic Community

Towards the end of this quarter, real seismic data was available from small test files in both the regular Datacomputer and a new experimental Datacomputer. VSC and SDAC were informed of the availability of this data and assisted in its retrieval.

CCA also cooperated with SDAC in developing files that will be written directly by SDAC. These are anticipated to be the event summary files, seismic waveform files, and the instrument status and calibration file.

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 has required site modifications at CCA.

4.1 Site Preparation

In early February the vacuum and air compressor systems for the TBM arrived and were installed. All of the TBM equipment except for the second dual transport drive arrived February 23.

Some minor cosmetic touch-ups were also done during this quarter. The only remaining site work is the installation of a new electric circuit breaker for some of the TBM equipment.

Earlier site work had involved installation of a new TRANE air conditioner including a humidifier system with a replacable throw-away canister. Difficulties have been encountered with this humidifier and with obtaining a spare canister of the appropriate type. These problems may be related to the very high level of impurities in the water at the computer site in Cambridge, Massachusetts.

4.2 The Hardware

Progress is being made by Ampex toward turning over the TBM to CCA for formal acceptance testing. Some data has been successfully transferred to and from each of the four TBM tape drives.

Negotiations between CCA and Ampex toward agreement on the specifics of the acceptance test has lead to the development of a tree structure, shown in figure 2, for classification

READ ERROR NOMENCLATURE

(Assumes the error occurs on a block that has been successfully written and read verified)

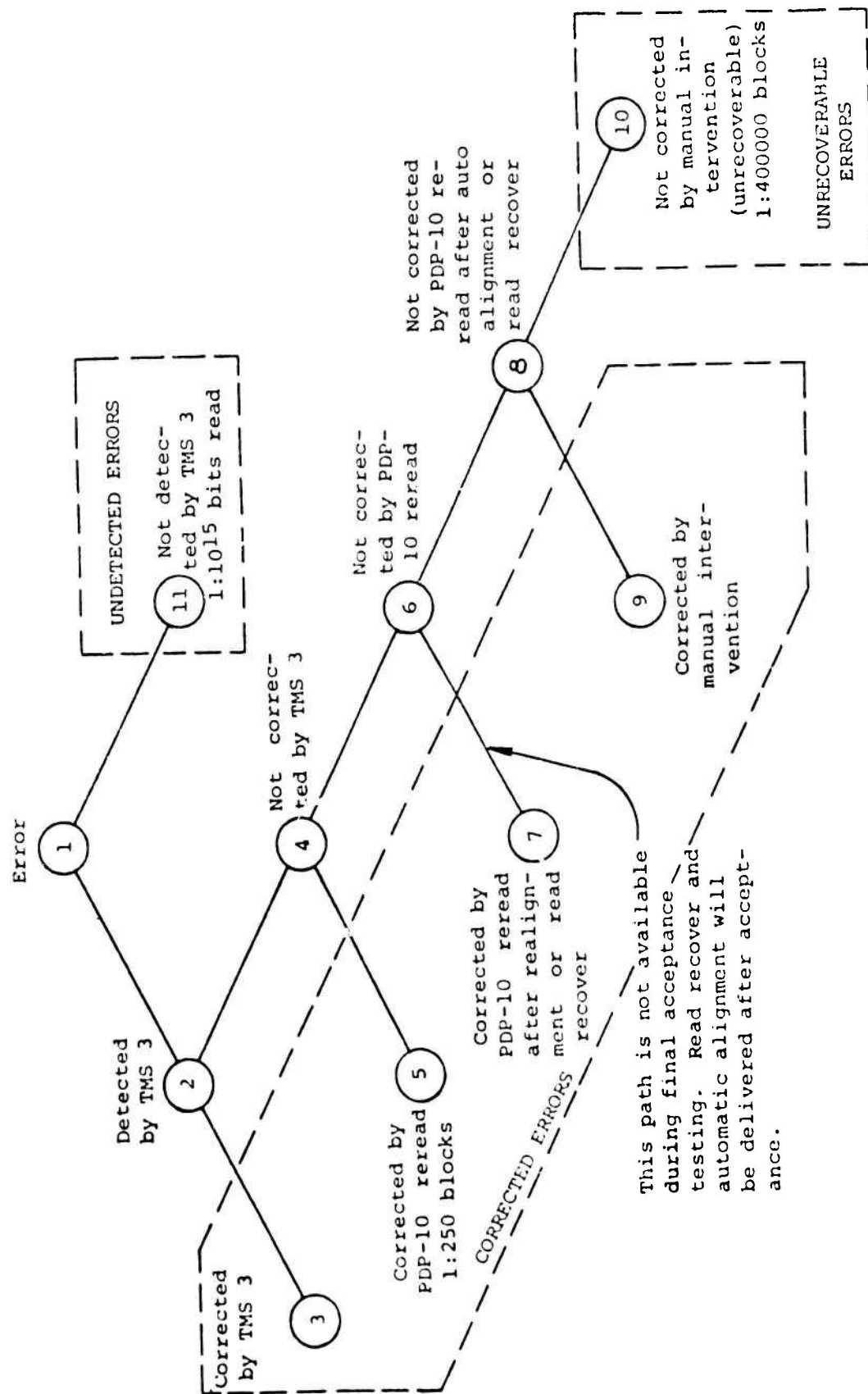


Figure 2-- TBM Error Analysis

of data errors. This scheme might be useful as an example in developing error classifications or acceptance procedures for other devices.

5. The Datacomputer

A new experimental version of the Datacomputer software was put up this quarter in parallel with the regular Version 1 Datacomputer software. This new experimental version has most of the features that Version 2 will have (except that it does not yet use the TBM.)

The main problem that had been encountered in SIP use of Version 1 was an inadequate data rate when storing the Datacomputer. This was traced to bugs in the Version 1 routines that check if data conversion is necessary due to differences, between source and destination, in the described format of data. In this case, the source is the SIP over the Arpanet and the destination is a seismic array data file. The effect of these bugs was that, although the SIP carefully prepares its data so that no conversion is necessary and whole records can be stored from the Arpanet into a file, the Version 1 Datacomputer, erroneously thinking conversion might be necessary, was manipulating the subfields of each record, thus restricting bandwidth. These bugs have been fixed.

The new version of the Datacomputer also contains a preliminary implementation of the SDAX routines that will keep track of multiple version of files, staging of data from TBM to disk, and multiple simultaneous readers of a file along with one writer. Since the new Datacomputer version does not actually use the TBM, SDAX is set up, for testing purposes, to stage from disk to disk. These staging operations produce additional messages to inform the user of their progress. It was necessary to make some very minor modification to the SIP to adapt it to these additional messages.

Unclassified

Security Classification

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, February 1, 1976 to April 30, 1976			
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		8b. ORIGINATOR'S REPORT NUMBER(S)	
b. PROJECT NO.			
c.		9b. OTHER REPORT NO(S) (Any other numbers that may be assigned to this report)	
d.			
10. DISTRIBUTION STATEMENT		DISSEMINATION STATEMENT A Approved for public release; Distribution Unlimited	
11. SUPPLEMENTARY NOTES		12. SPONSORING MILITARY ACTIVITY Defense Advanced Research Projects 1400 Wilson Boulevard Arlington, Virginia 22209	
13. ABSTRACT			

DD FORM 1473

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

Unclassified

Security Classification

Unclassified

Security Classification

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						

Unclassified

Security Classification

CAUTION - REMOVE INDICATOR SHEET BEFORE TYPING
"TO BE STORED IN A COOL DRY LOCATION"

MATERIAL INSPECTION AND RECEIVING REPORT		1. PROC. INSTRUMENT IDEN (CONTRACT) <div style="text-align: center; font-size: 1.2em;">MDA903-74-C-0227</div>		(ORDER) NO. 6. INVOICE NO. DATE		7. PAGE 1 OF 1 8. ACCEPTANCE P D	
		2. SHIPMENT NO. CCA00030		3. DATE SHIPPED 15JUN76		4. B/L TCN	
9. PRIME CONTRACTOR CODE Computer Corporation of America 575 Technology Square Cambridge, Massachusetts 02139				10. ADMINISTERED BY CODE Mr. J. McDonough Defense Contract Administration Services Region, Boston 666 Summer Street Boston, Massachusetts 02210			
11. SHIPPED FROM (If other than 9) CODE Same as 9. Above				12. PAYMENT WILL BE MADE BY CODE Disbursing Officer Defense Contract Administration Services Region, Boston 666 Summer Street Boston, Massachusetts 02210			
13. SHIPPED TO CODE Defense Advanced Research Projects Agy Architect Building 1400 Wilson Boulevard Attn: Co. D.C. Russell Arlington, Virginia 22202				14. MARKED FOR CODE Same as 13.			
15. ITEM NO.	16. STOCK/PART NO. <small>(Indicate number of shipping containers - type of container - container number.)</small>	17. QUANTITY SHIP/REC'D *	18. UNIT	19. UNIT PRICE	20. AMOUNT		
0002 AB	Quarterly Technical Report April 30, 1976	2	EA		NSP		
21. PROCUREMENT QUALITY ASSURANCE <div style="display: flex; justify-content: space-between;"> <div style="width: 45%;"> A. ORIGIN <input type="checkbox"/> POA <input type="checkbox"/> 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. </div> <div style="width: 45%;"> B. DESTINATION <input type="checkbox"/> POA <input type="checkbox"/> 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. </div> </div> <div style="display: flex; justify-content: space-between; margin-top: 10px;"> <div style="width: 45%;"> DATE _____ SIGNATURE OF AUTH GOVT REP _____ TYPED NAME AND OFFICE _____ </div> <div style="width: 45%;"> DATE _____ SIGNATURE OF AUTH GOVT REP _____ TYPED NAME AND TITLE _____ </div> </div>					22. RECEIVER'S USE Quantities shown in column 17 were received in apparent good condition except as noted. DATE RECEIVED _____ SIGNATURE OF AUTH GOVT _____ TYPED NAME AND OFFICE _____ <small>* If quantity received by the Government is less than quantity shipped, it shall be marked "defect". Enter a list, quantity received below quantity shipped on an invoice.</small>		
23. CONTRACTOR USE ONLY							