THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 56, No. 2, February 1977
Printed in U.S.A.

1A Processor:

Memory Systems

By C. F. AULT, J. H. BREWSTER, T. S. GREENWOOD, R. E. HAGLUND, W. A. READ, and M. W. ROLUND

(Manuscript received July 17, 1976)

The memory hierarchy of the 1A Processor consists of call/program stores, file stores, and an auxiliary data system utilizing magnetic tape. The call/program store is a 65,536-word, 26-bit-per-word memory system using a two-wire, coincident-current, ferrite-core array as the memory medium, and is capable of operating with a 1.4-μs read/write cycle time. The file store (FS) with its disk-file memory provides the 1A Processor with a high-performance bulk-storage system. Each FS controller can control from one to four disk files to provide a maximum storage of 2.56 × 10^6 words. Two FSs make up a file-store community of 5.12 × 10^6 words of duplicated storage. The auxiliary data system (ADS) provides a means for efficiently transferring data between the 1A Processor call- or program-store memories and magnetic tape. The tape medium is used for inputting generic programs, office data, and trouble-location information and for outputting billing, traffic, and error-analysis data. The ADS hardware consists of data unit selectors (DUSs) and tape frames (TFs).

I. INTRODUCTION

1.1 Memory hierarchy

The 1A Processor memories are call stores (CSs), program stores (PSs), file stores (FSs), and tape. The program and call stores are very similar in hardware. The main difference is a two-word or 52-bit output from the PS and a one-word or 26-bit output from the CS.

The PSs contain programs that are normally resident and paging areas that are used for only occasional programs. The CSs are for transient data and for translation data. Each CS containing transient data is duplicated

in core while translation CSs are simplexed with the data duplicated in disk. In case of a PS failure, a roving PS spare is substituted and pumped up with program from the file store. Failure of one of a fully duplicated CS pair does not involve a pump up from the FS. Failure of a simplex CS calls for replacement by one of a duplicated CS pair and a pump up from the FS.

The file store is used as a source of data and programs and, in addition, it is used to accumulate data. The file store is essentially nonvolatile; i.e., the contents can be trusted when first powered up. A new unit of memory must be written when first installed. The source can be either the duplicate unit or tape.

The tape units are part of the auxiliary data system (ADS). The tape units are used to accumulate data and to supply program and translation data to the store.

1.2 Memory buses

There are three memory buses: program store, call store, and auxiliary unit (AU).[1] The PS bus services from 2 to 22 stores of 65,536 (26-bit) words. The CS bus services from 2 to 44 stores of 65,536 (26-bit) words. Both of these buses are fully duplicated and operate at multiples of 700 ns. The AU bus services the FS, ADS, and other types of equipment. This bus also operates at multiples of 700 ns. The AU bus communicates with the program stores and call stores via the central control under hardware control. This direct memory access to CS/PS is one of the major differences between the 1A Processor and the No. 1 ESS processor. The file stores are paired, but operate as simplex stores. Each FS can store up to 2,560,000 24-bit words.

There can be two or more pairs of FSs depending on system needs and on other uses of the AU bus. The tape units communicate over the AU bus via the data unit selector (DUS). Both the FS and DUS operate at multiples of 700 ns.

1.3 Cycle times

The 1A Processor is designed to operate with CSs/PSs that operate at either 700-ns or 1400-ns cycle times. The units on the AU bus can transmit data as often as one word every 2.1 μs. The PSs/CSs that are currently in use are magnetic core units operating at 1400 ns. The file stores use disk memories with a transfer rate of 10 μs/word and the tapes are 800 b/in. with a transfer rate of 150 μs/word. Each pair of DUSs can service up to 16 tape units simultaneously.

II. CALL AND PROGRAM STORES

2.1 Introduction

The CS/PS is a 65,536-word, 26-bit-per-word memory system using a two-wire, coincident-current, ferrite-core array as the memory medium, and capable of operating in a 1.4-μs read/write cycle time. As implied in the name, the same basic store unit is used in either the CS or PS communities. When used as a PS, the unit responds to read commands with a two-word (26 bits/word) reply. The differences between call- and program-store operations are determined by minor circuit-pack and bus-cabling options.

When used in a No. 4 ESS, the 1A Processor will contain typically 13 PSs and 12 CSs. A maximum of 22 PSs may be connected to the PS bus, and a maximum of 44 CSs to the CS bus.

The heart of the store design is the 20 A/B memory module. This unit has been in volume manufacture since 1970 for its No. 1 ESS 32K-word call-store application.[2] It is a proven, reliable building block with a relatively low, established cost.

The store will be described in four sections—controller, access circuits, memory module, and power unit.

2.2 Controller

The controller interfaces with the central control (CC) and the memory-access circuits. The controller responds to system requests by determining if the request is for it, deciding what type of order is requested, and generating all the timing pulses and control functions necessary to access the memory circuits and to receive data from or reply data to the CC. The store can perform control or memory type orders. The control orders are used to configure or change the status of the unit, or to return the status to the CC. Additional control orders are generally used by the diagnostic program to verify bus integrity, aid in determining if the unit is working correctly, and, if failing, to resolve faults.

The address format is shown in Table I. The four mode and order bits, R, W, C, and M, determine whether the store will respond via a normal,

Table I — Address format.

Address bit	Function
25	Parity
24–21	Mode & order (R, W, C, M)
20–16	Name
15–0	Memory address

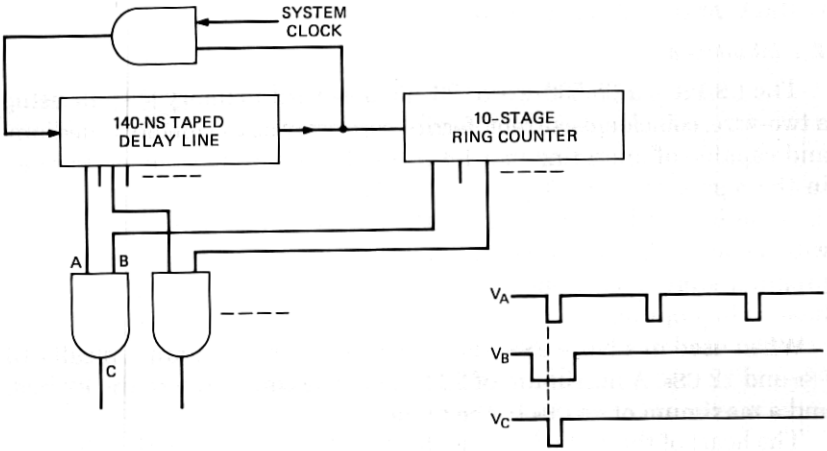

Fig. 1—Clock circuit.

control, or maintenance operation. Each store is assigned a 5-bit name or K-code, which may be changed by a control-write operation. Certain program stores may be set to K-code 20 with a single private pulse. This initializes one store for system recovery.

The store cycle is initiated by receiving a set of sync pulses from the CC. An address window is opened and the internal clock started. If a name match and valid operation (NVOP) is not obtained, the store will shut down, awaiting the next set of syncs. These are repeated every 700 ns. If an NVOP is obtained, the store cycle continues for 1200 ns. The store clock generates precise timing edges with 3-ns resolution over the 1200-ns store cycle. This is realized using a 10-stage ring counter, a 140-ns delay line with 3-ns taps, and logic as shown in Fig. 1. The ring-counter stages define windows during the store cycle. The tapped-delay-line pulses are logically ANDed with the windows to yield precise timing edges for control of store operation.

The controller performs many diagnostic and error-checking functions. A variable-timed strobe, controlled by diagnostic software, is internally generated. It enables the diagnostic to observe the state of critical nodes within the store as a function of time. In addition, incoming address and data are verified for correct parity, the presence of all system clocks is checked, and internal clocks are checked for proper operation. Memory maintenance is also performed by monitoring signals returned from the memory module for level, timing, and double selection. Errors early in the cycle will result in error signals being returned to the CC. These early errors, as well as later errors, are trapped in registers internal to the unit to permit fault resolution.

All control signals to the access circuits are buffered using balanced, differential, emitter-coupled, logic-type drivers in the logic section and

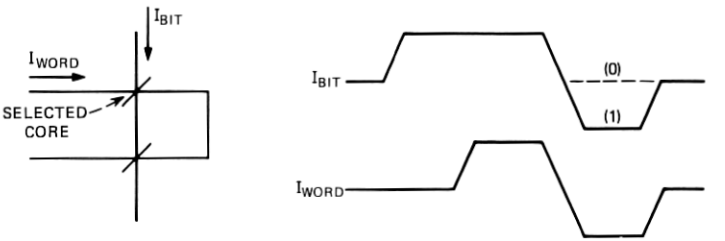

Fig. 2—Bit and word current.

corresponding receivers at the access circuit. This is required to provide adequate noise margin between the logic and the high-voltage (72 V), high-current (1.4 A) access circuits.

2.3 Memory module

The memory module consists of two back-to-back core planes each containing an 832 by 512 array of 23-mil-diameter ferrite cores. Each core has two orthogonal wires passing through it: the bit (or B line) and the word (or A line). The core is switched by coincident currents of 300 mA in each line. During the first part of a read or write memory cycle, the core is read out; i.e., if in a 1 state, switched to a 0. If a 1 is to be written back, the bit and word currents are reversed during the last part of the cycle. The word current is always driven, because it drives more than one bit. During write back, the bit current exists only if a 1 is to be written. As shown in Fig. 2, the bit current is brought up first in the read portion of the cycle. After about 150 ns, the word current is then activated. This is done since the readout signal (the core switching voltage of about 30 mV) is sensed on the bit line. The delay permits the bit line to recover from the high-amplitude drive transient. As also shown in Fig. 2, the word line is looped back through another core on the bit line. This core is selected by reversing the polarities of word current. Since bipolar current is required for one core, the selection of this core is provided only by reversing the timing sequence of the word current. Cancellation of core shuttle noise due to word current also occurs.

2.3.1 B-current selection

For each bit in the memory, a bit line is driven on both the front and back plane. Half of the A loops are on each plane. These two bit lines are driven from a Balun transformer as shown in Fig. 3. The Balun assures the front and back current are each approximately equal to one-half of the input current. The core signal (V_{out}) is sensed across this same transformer. This Balun and the Balun for the adjacent bit are driven from a set of transformer-coupled read and write current sources.

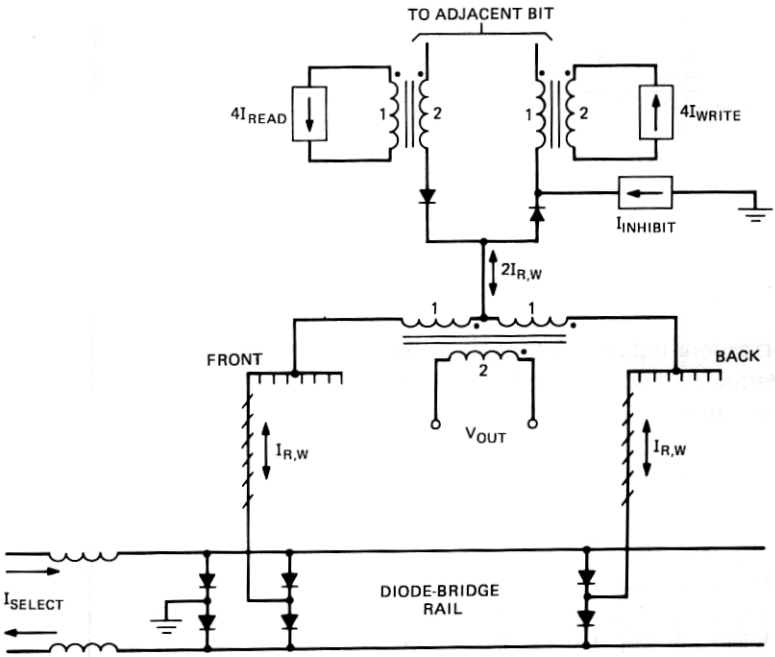

Fig. 3—B-current selection.

Transformer coupling allows all bits to be driven in series, which reduces the number of current drivers. It also ensures that all bit currents are equal. The B-current is inhibited in the memory lines during the write portions when storing a 1 by driving the bit-inhibit current (also transformer coupled), which then provides the write current rather than the bit line.

There are 32 pairs of bit lines (front and back) associated with each bit of a word. Eight pairs are multipled to a given Balun transformer with current-routing circuits selecting one of four such groups of Baluns. The other ends of the bit lines are selected by a balanced multi-armed diode-bridge arrangement, as shown in Fig. 3. By making I_{SELECT} sufficiently larger than the total memory currents, all diodes on the selected rail are forward-biased, providing a virtual ground at the bottom ends of the bit lines. This technique isolates virtually all of the parasitics of the selection circuits, allowing them to be placed remote from the module without adversely affecting the control of the bit-current waveshape. This is important in a two-wire memory since the core switching signal is sensed across the same bit lines that are carrying the high-amplitude bit current.

As with the read and write currents, the rail-select current is also transformer-coupled. Eight sets of rail-selection circuits in conjunction with the four Balun selectors provide the necessary 1-out-of-32-bit pair selection.

2.3.2 Word selection

The word access consists of 512 loops. A one-of-eight select is performed on the nondriven end of the word loop using the same type of bridge-rail circuit as described in the B-current section. Sixty-four drive transformers arranged in an 8-by-8 matrix complete the selection. Memory current is provided on one of eight inputs on one axis of the matrix (called verticals). A one-of-eight selection is performed on the other axis using the bridge-rail circuit (called horizontals). (See Fig. 4.) As shown in Fig. 4, a transformer is located in each vertical to provide a sample of the A current to the maintenance and readout strobing circuits.

2.4 Controller to memory module access

2.4.1 Bit-current selection

All active circuits used to access the core are contained external to the core memory module to meet the reliability and repair time requirements for the store. The memory module contains transformers, diodes, and resistors. The basic drive element used to drive the memory

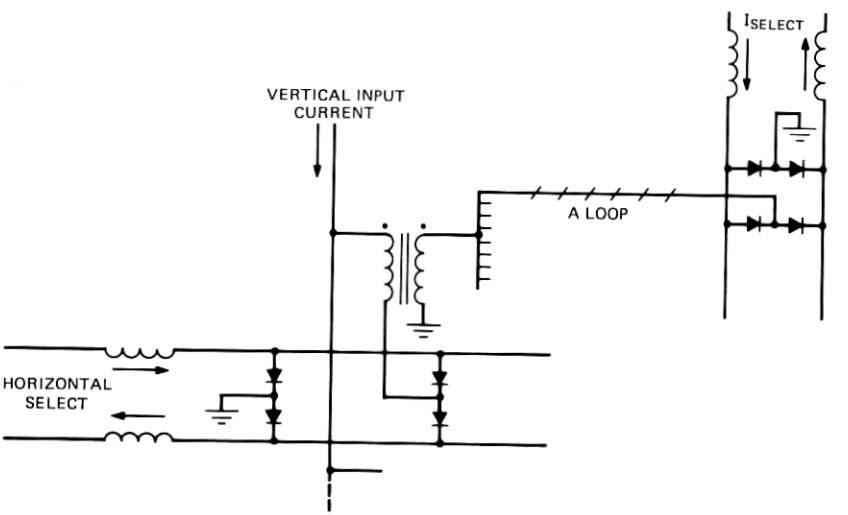

Fig. 4—Word line selection.

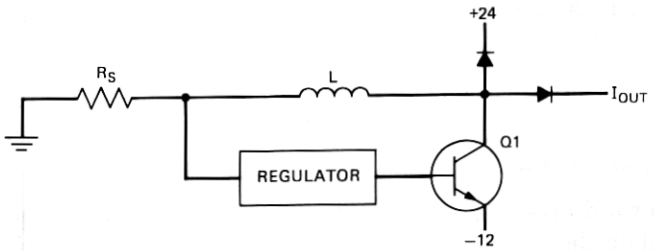

Fig. 5—Current driver.

module is a regulated current driver whose current level is controlled by a reference voltage (VREF). VREF is a memory module output voltage and is determined by the temperature of the core cover plate internal to the module. This is necessary for the drive current to track the core characteristics as a function of temperature. To drive the high-inductance load and obtain the necessary rise and fall times of the memory current, drive voltages of up to 65 volts are required. To reduce power dissipation in the driver, a switching regulator is used to control the current in a large (1-mH) inductor. This large inductor then can drive the inductive load of the memory (approximately 4 μH) and maintain good current control. The basic regulator is shown in Fig. 5, where the inductor is switched to −12 V or clamped at +24 V to control the current. When driving the memory, regulator action is inhibited by turning off Q1.

The output of the drivers must be switched to the desired module inputs. The basic floating high-current switch that performs this function is shown in Fig. 6. This circuit must switch 1.4 A with a 65-V driver level. Switching speeds in the tens of nanoseconds are necessary to meet the system timing requirements. Switch transitions must be fast (<15 ns)

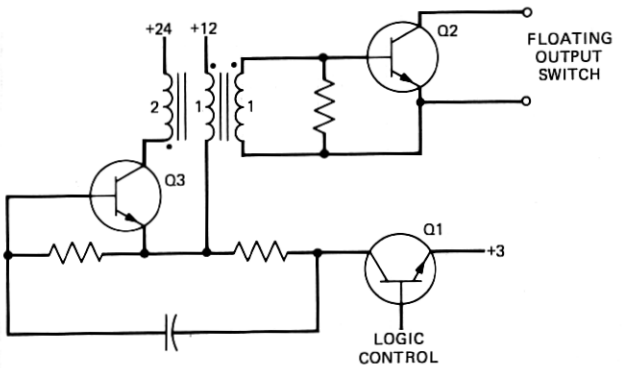

Fig. 6—High-current switch.

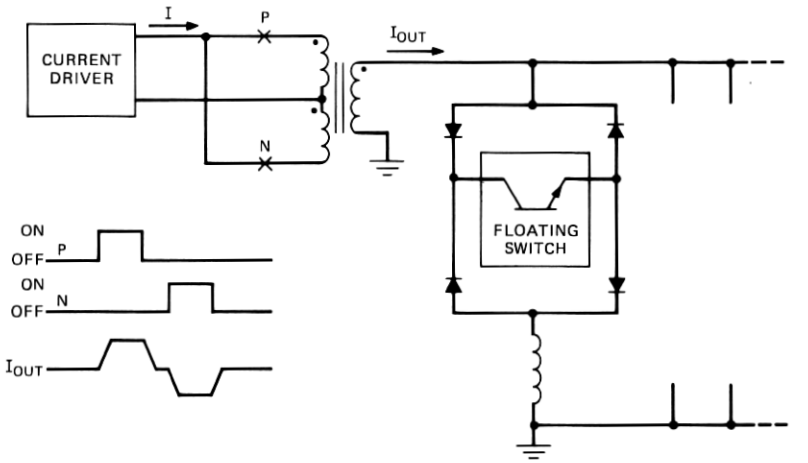

Fig. 7—Vertical current.

to minimize power dissipation. The circuit of Fig. 6 employs a high-speed turn-off of Q2 using Q3. This permits Q2 to be saturated to reduce power dissipation.

To perform the rail selection described in 2.3.1, it is necessary to drive one-of-eight B-rail drive strings. This is accomplished using a single current driver, which is routed into the selected transformer string via a set of eight of the floating, high-current switches described previously.

The one-out-of-four B-read and B-write drive strings are driven in a similar manner. In this case, one current driver is used for both the read and write current. A series switch determines whether read or write current will be driven and isolates the current driver. Selection switches then select one of four drive strings during the B-read and B-write portion of the cycle. Although the memory bit-line currents are on the order of 300 mA, the various transformers used in the selection and routing process provide an effective 4-to-1 ratio from current driver to memory. Hence, the bit-access switches handle currents of the order of 1.2 A.

2.4.2 Word-current selection

Except for lower current, word or A-current selection is similar to the B access described previously. In this case, however, bipolar current is necessary to drive the A vertical inputs. As shown in Fig. 7, a transformer and two switches are used to generate the bipolar current from a unipolar current driver. The operation of switch P results in positive current, whereas N results in negative current since the primary is re-

versed. The selection of 1-out-of-8 switches then completes the A vertical selection. These switches are diode-bridge switches to handle the bipolar current.

The total 1-of-1024 A-word line selection is achieved via 1-of-8 bridge rails, 1-of-8 horizontal rails, 1-of-8 vertical drives, and a 1-of-2 A vertical current polarity.

2.4.3 Signal detection

As described in the module section, the core signal appears across the secondary of the driven B-current Balun. As shown in Fig. 8, two secondaries are connected in series to reduce the number of detectors required without significantly loading the core signal. The resulting two pairs of leads are applied to the data channel. The input stage then selects one of the two pairs. The pair selected is determined by which B current was driven. The core signal rides on top of a significant offset voltage caused by the diode-bridge uncertainties and variations in resistive voltage drops on the driven and nondriven B lines in the memory. In addition, during the bit-current rise time, large voltages due to inductive unbalance of the B line and partial switching of nonselected cores will exist on the readout line. To establish a ground reference for de-

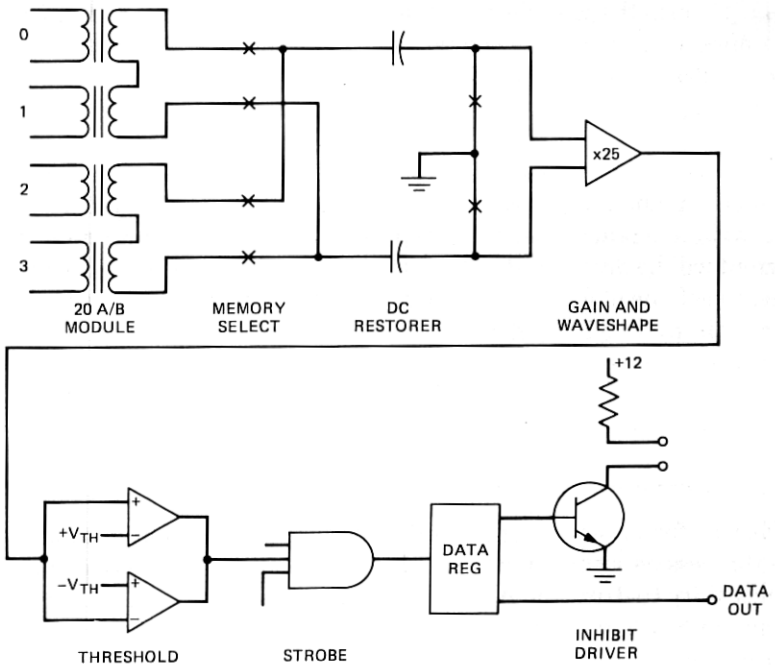

Fig. 8—Data channel.

tecting the core signal, the offset voltage is stored on the restorer capacitors. Just prior to the start of the core switching voltage, the switch holding one side of the capacitor to ground opens and allows the following signal to be amplified. The memory signals are balanced up to the amplifier. Therefore, two restorer switches and capacitors are required. The output of the amplifier then drives two detectors whose outputs are ORed. One detects positive 1s and the other negative 1s. The detector output is then strobed into a data register. The strobe is timed a fixed delay from the 50-percent point of the A current to match the peak of the core signal. The A-current timing is obtained from the A-current sample leads of the 20 A/B memory. The data register then controls the inhibit driver, which when active (data = 0) inhibits the write back of a 1.

A test mode is available where the detector threshold level may be raised or lowered to detect marginal operating conditions.

2.5 Power

In addition to the −48 V system power, the store uses five regulated voltages. These are supplied by a power unit, which also provides high-voltage protection, current limiting, voltage monitors, and various diagnostic test features.

III. FILE STORE—A DISK MEMORY SYSTEM (FIG. 9)

3.1 Introduction

In the No. 1 ESS processor, all programs are stored in the permanent magnet twistor (PMT). The cost of storing a program or data in the PMT is the same whether it is vital or ancillary. Changes are also slow, requiring rewritten magnet cards, which are manually changed. To reduce cost and provide flexibility for changes, the objective in the 1A Processor was to have an electrically alterable memory that would hold data during power loss or hardware faults. A disk admirably meets these objectives. The cost is low and the information resides in the magnetic state of the disk medium and is not easily destroyed. The disk, however, is sometimes too slow to operate out of directly. Therefore, it is used instead as a source of programs, and data, which are moved to the faster random-access stores before being used.

The combination of disk and random-access memory (RAM) is less expensive than the equivalent PMT and is much more flexible. One reason for the low cost is that much more is stored on disk than is in RAM at any instant of time. For instance, most diagnostic programs are brought to RAM as needed and the same RAM space is shared among different programs.

The cost of a bit stored on disk is significantly less than a bit stored in RAM. If the performance of a disk memory (its delay in returning

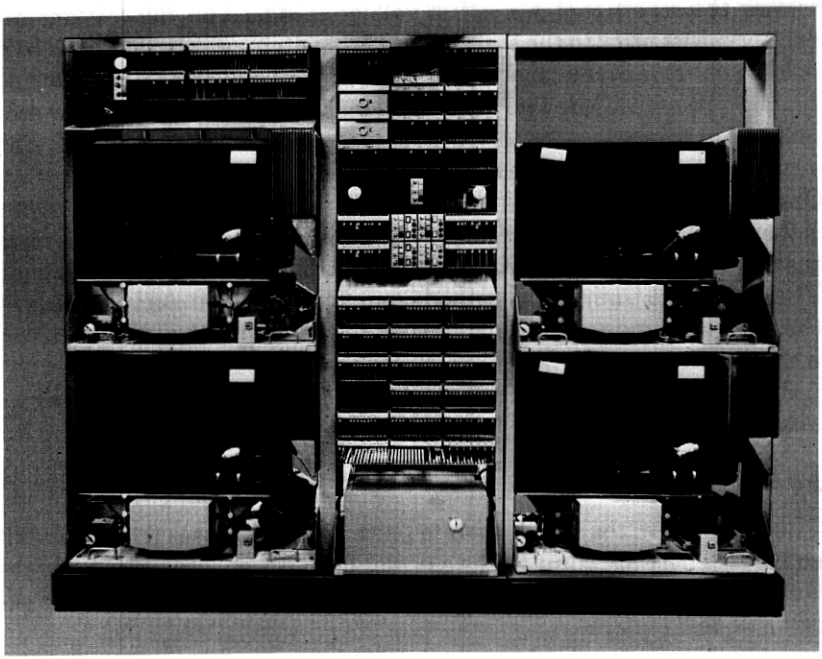

Fig. 9—File-store frame.

needed data) is good enough for a given application, then that class of data will not have to be stored in RAM. If the delay is too long, then the data must be stored in RAM and on disk.

A major objective in the design of the file store was high performance. High performance means low latency, i.e., data is returned with the minimum delay. If the system were to request one piece of data at a time, not much could be done in hardware to reduce latency. Each request on the average would have to wait a half revolution of the disk.

If a group of requests is available, significant improvements in performance can be made. It is possible for the FS to handle 50 separate requests (jobs) in one revolution, giving a delay of less than 1 ms between jobs.

The disk memories are reliable and nonvolatile, but they and the controller can fail. To assure a continuous source of data and program, FSs are duplicated. A community of FSs has two controllers each with from one to four disks. The FS system is designed to operate with disk failures in each controller as long as duplicate pairs of disks have a maximum of one failure. Each controller is on a separate AU bus.

The two controllers in a community communicate the phase relationship of duplicate disk pairs. This information is used to maintain a 180-degree phase relationship. This has two purposes. A job can be submitted to both controllers simultaneously. The first disk to reach the address does the job. The 180-degree phase relationship assures that the job will be completed before the other can start. Completion cancels the job for the other disk. This protection is needed to prevent the latter controller from overwriting a data set which has just been modified. There is another obvious advantage to the 180-degree relationship. It reduces latency by assuring that one of the pair of disks is always within a half rotation of any address.

3.2 Disk-file memory and support circuitry

The commercial disk chosen for this application meets the main requirements of reliability and availability. The characteristics are:

1800 RPM
16×10^6 bits (gross storage)
200 heads
2.4-MHz data rate
Plated media
Continuous air filtration.

The memory units are organized as follows:

100 tracks/face
2 faces/disk
101 sectors
34 words/sector (32 data, 1 preamble, 1 error code)
24 bits/word.

One hundred sectors are used for data and the last sector in each track is used for diagnostic testing.

All power must be derived from either +24 dc or −48 dc. The disk motor is a single-phase, 208-V induction motor with a capacitor start winding. The motor is powered by an inverter, which is part of the 180-degree servo loop. The frequency of the oscillator controlling the inverter is continuously changed to control the 180-degree phase relationship.

The motor continuously accelerates and decelerates the disk; i.e., supplies or accepts energy. When accepting energy, the voltage across the silicon-controlled rectifier (SCR) in the inverter that is being turned on can be zero. This will delay the actual start of conduction. Two things were done to assure reliable commutation when accepting energy and when the start winding kicks out. The inverter output transformer is clamped with diodes. The effect of this is to reduce the reverse voltage

across an inductor in the emitter circuit of the SCRs. This low reverse voltage causes a large circulating current to flow. This current tends to offset the effects of the reverse flow of energy from the motor. Commutation is assured by supplying many trigger pulses to the SCRs. Each trigger has a fast rise time to force the SCRs into conduction rapidly and the pulse is repeated, so that conduction will start reliably whenever the voltage across the SCR becomes positive.

The FS must operate over an appreciable temperature range and it is desirable to minimize adjustments to circuit packs, particularly in the field. To meet these objectives, both timing and gain were made adaptive. No adjustments are required when a disk or a circuit pack is replaced.

The timing circuit was designed with phase adaption in mind. The disk has two clock tracks, bit and sector. All active timing is derived from a phase-lock-loop locked to the bit track. This circuit consists of an oscillator many times the frequency of the bit clock. The oscillator output is counted down to the frequency of the bit clock. The phase of the two are compared and the frequency of the oscillator adjusted to maintain phase lock. One advantage of this approach is that accurate timing is available at periods much shorter than a 1-bit clock cycle. Another advantage is that minor defects in the bit clock do not affect the accuracy of the derived timing. Failure is detected by matching the derived count with the count stored in the sector clock.

Writing on the disk is timed directly from the timing circuitry. When a sector is written, a preamble is written before any data. This preamble is used to first establish a decision level for the read circuitry and then to establish accurate strobe timing. This accurate strobe timing is derived by forcing a counter (Fig. 10) not to count down the oscillator frequency. This causes the preamble to be shifted into a register at a very high rate. When the appropriate pattern (preamble) is recognized in the register, the counter is allowed to operate normally. The next output of the counter will coincide with the center of the first bit of data.

The 180-degree servo loop between duplicated disk files does not require high accuracy, but was difficult to design for several reasons. For economic reasons, it is desirable to minimize the exchanging of position data between communities. The minimum possible is used, one bit per disk. The motor is single phase, and it produces a marginal excess torque at low battery. The motor is coupled to the disk via a belt. Large torque changes at high battery had to be avoided for good belt and motor-mount lifetime.

The servo is a master-slave type capable of operating in either mode. The disk that is the master controls its speed and the disk that is the slave maintains phase lock to the master. If the slave disk detects an out-of-speed indication, it switches to master mode. If a disk is out of speed for three revolutions, the heads are retracted. This three-revolu-

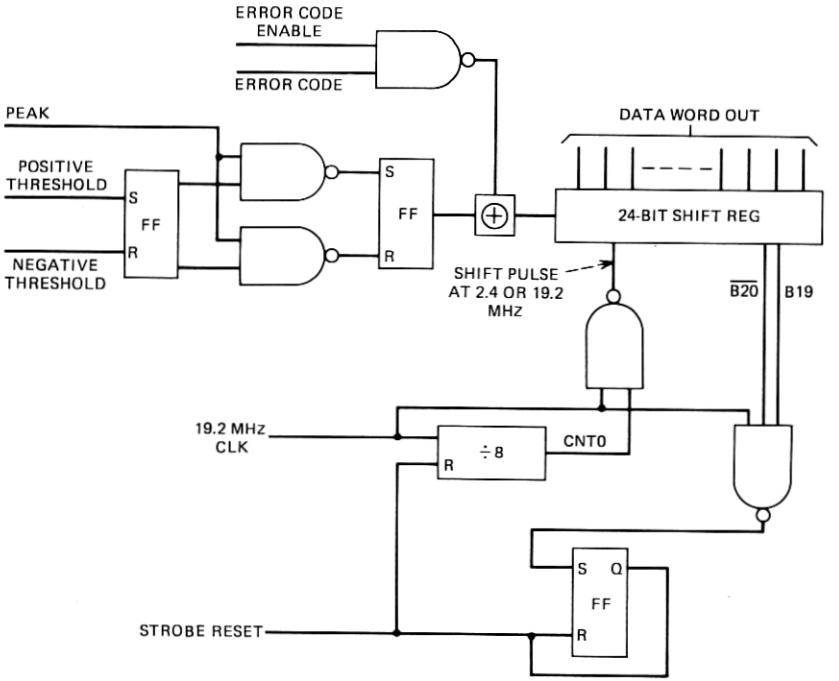

Fig. 10—Adaptive strobe circuit.

tion delay is sufficient to allow the slave-to-master transition to return to the correct speed without retracting heads.

The servo is a digital circuit that minimizes torque excesses and acquires and holds phase lock under all possible conditions:

(*i*) One or two flip-flops are set, depending on the magnitude of the phase error, if the error is increasing.

(*ii*) The phase polarity and the two FFs drive a D-to-A converter with three states increasing and three decreasing the frequency.

(*iii*) The rate of decreasing phase error is made small so that little overshoot occurs, and for all small errors the restoral force is the least output of the converter (making for low torque changes while phase locked).

The servo will acquire the 180-degree phase lock in less than one minute after the heads fly. It will hold lock from 41.75 V to 52 V. There is no measurable increase in wear on the motor or belt due to the servo.

3.3 Job setup and execution

Because the disk revolution rate is extremely long compared to the system cycle time, requests for disk jobs are queued in the FS. Because

of cost considerations, only the minimum amount of information necessary to find a disk job is stored in each register of the job queue. Corresponding to each queue register is a disk request (DR) block in core memory where the job is fully specified. Administering the queue and the DR blocks is a disk-administration program.

Clients submit jobs to the administration program, where an idle queue register in the mate FS is simultaneously loaded. An active FS that is not doing a disk job is continuously searching the queue for work. Once it detects a job that starts within the next one or two sectors, searching stops and the FS autonomously accesses core to retrieve the DR block.

Disk-read jobs, which are no longer than 1024 words, are sent to both FSs simultaneously. Given that corresponding disk files in the duplicate FS are servoed 180 degrees out of phase and the fact that a 1024-word job is only one-third of a revolution in length, one FS can do the job and return status to the DR block before the duplicate FS accesses the DR block. When the duplicate FS reads the status word and finds the read job has been done, it cancels the job in its queue and restarts the search sequencer. A disk-write job is not considered done until both FSs have done it successfully.

When the status word in the DR block is retrieved during job setup, it is checked using the above rules to determine if the job should be done. If it should be done, job setup (retrieving the remaining words in the DR block) continues. To do the disk job, core store is again autonomously accessed to obtain data to write on disk or to write core with data read from disk. At the conclusion of the job, the FS writes job status in the status word of the DR block, idles the job queue register, and starts searching again.

It is the disk-administration program function to notify clients when their job is completed and whether it was done successfully or not.

3.4 Auxiliary unit bus interaction

The FS communicates with CS/PS via the CC and the auxiliary unit (AU) bus. Each AU utilizes a handshaking routine with CC to perform information transfers to and from CS/PS.

An AU first requests use of the bus by sending a bus request. When the request is granted, the AU acknowledges receiving the use of the bus with a verify signal. The unit then sends address and data (if a write to CS/PS) and waits for an acknowledgment that the transfer has been completed. After acknowledgment, the FS returns an acknowledgment-verification signal to the AU sequencer in CC.

AU bus activity is controlled by CC according to a priority structure and to a CC "give up" algorithm. The priority structure gives CC the highest priority followed by FS 0, FS 1, FS 2, FS 3, and then the other AUs.

The priorities for FSs are programmable within the allocated priority groups, while those of other AUs are hardwired. The "give up" algorithm basically says that CC must give up 1-out-of-4 cycles to AUs.

If bus blockage occurs, the FS will continue to request the bus every 700 ns until the request is granted or the FS buffers overflow and the job is resubmitted under program control starting at the point of overflow.

CC "give up" of the AU bus when continued blockage occurs is based on the following criteria:

(*i*) If AUs have been blocked three times, and if during the blockage no AU successfully completed a store operation, CC will give up the next available cycle to an AU.

(*ii*) If an AU to CS/PS operation is in progress, and the CC wants to read or write an AU, the CC inhibits enables to all AUs and waits for the operation to complete before accessing the AU. Enables are inhibited for a maximum of three successive CC cycles.

3.5 Controller organizations

The file-store controller is a 25,000-gate machine with 22 functional circuit areas. The seven major circuit areas are described below.

3.5.1 Disk-request registers

These registers are the hardware job queue for the FS. The queue is made up of either 24 or 48 registers, depending on system requirements. The registers are 13 bits wide and contain only a subset of the information needed to process the transfer to or from disk. This information includes disk sector address, disk file, disk face, read/write bit, an update bit, and the busy/idle bit.

3.5.2 Address translator

The address translator translates a 17-bit record address into two 7-bit fields (sector and track) plus a 3-bit field (disk and face) and allows the system to view the FS memory as having a closed binary address field. The translator converts that binary address to provide the addressing for a disk with 100 tracks, 100 sectors/track, and 32 words/sector.

3.5.3 Search-and-match sequencer

This sequencer searches through all the job requests in the queue until it finds one that is for a sector close to the present position of the read/write heads. Once a match has occurred, the job is passed to the store sequencer for processing. After the store sequencer has processed

the job, a signal is returned to the search sequencer. The signal then sets the idle bit in the job queue and starts searching for the next job.

3.5.4 Job-control registers

These registers contain the detailed job-control information. The four registers are loaded from the DR block and contain the following information: number of words, identification tag, record address, starting word in sector, core-store starting address, plus status information from the other FS in the community if the job has been set to both FSs.

3.5.5 Store sequencer

This sequencer manages information transfers to and from core stores. In addition, it processes all errors detected by the FS. Depending on the nature of the error, the sequencer will either stop the FS, take it off-line and notify the system, or it will report job failure status to the system via the status word in the DR block and then start the search sequencer.

3.5.6 File-store timing

The FS operates on two timing systems. The controller operates on timing sychronized with the CC while the disks operate on disk timing. Information transfers cross this timing interface asynchronously. Transfers from the disks are accepted asynchronously at the timing interface and synchronized with CC timing in the controller. Information transfers to disk are placed in a buffer and moved asynchronously via disk timing to the disk.

3.5.6.1 *Controller timing.* The system-related timing circuit in the controller generates all the system sychronized timing pulses needed in the FS. Fourteen 100-ns and fourteen 50-ns timing pulses are generated for normal-mode timing. This timing chain is duplicated, and the duplicate chain is used to drive FS maintenance circuitry. The use of a mask circuit to inhibit normal-mode clock pulses for diagnostic testing makes the duplicate unmaskable chain a necessity.

CC sends a sychronization pulse to the FS every 700 ns where it is used to start a 20-MHz gated oscillator. This oscillator drives the clock generator, decoder, and fanout circuits. Near the end of the 700-ns cycle, the gated oscillator control is reset. The oscillator then waits for another synchronization pulse before generating another string of timing pulses. The oscillator has been designed to hold its operating tolerance over a 40-year lifetime.

3.5.6.2 *Decision time sequencer—disk timing.* Reading or writing a disk is done with disk timing, which is independent of processor timing. Each disk has its own decision time generator consisting of a voltage-

controlled oscillator. It is phase locked to the bit clock track on the disk, a phase-locked counter, and associated decoding circuitry to provide disk timing pulses.

3.5.7 Automatic exerciser

The automatic wired logic exerciser reads the disk file when the FS is not processing a disk read or write. Thus, the integrity of the recorded data is routinely verified without system intervention. Read errors are counted until an overflow of the error counter (modulo 64) occurs. On overflow, the system is notified by the FS and the FS then stops in the maintenance mode. The last failing disk address is stored in the exerciser register.

3.6 Achieving system design objectives

The following sections enumerate FS design objectives and how each objective was met.

3.6.1 Performance requirements—high throughput, low latency

Performance requirements for the FSs were set by the worst-case throughput required in a No. 4 ESS. This required the FSs to complete 450 requests per second with an average delay of 30 ms per request. The 30 ms represents the elapsed time from job insertion in the hardware queue until job completion. It does not include task-dispenser time used to inform the requesting program that the transfer has been completed.

To achieve high throughput and low latency, the FS performs a closest-sector-address match on jobs in the hardware queue. Thus, the job closest to the present position of the read/write heads is always done first. The job queue is searched for jobs at the rate of one register every 700 ns. Thus, a 48-register queue is searched nine times during each disk sector.

Latency is reduced by a factor of two for read jobs which are no larger than 1024 words by the 180-degree phase relationship of duplicate disks. These jobs are submitted to both FSs and the first FS to process the job has time to complete and return status to the DR block before the second FS acquires the job.

Throughput requirements were met with one community of FSs operational. With only one controller in operation, the requirements were met by limiting low-priority jobs and restricting the number of large jobs in the queue.

Simulation results at 450 requests/second show 2 FSs have an average message delay of 19 to 33 ms (depending on average job size) while four FSs exhibit delays of 12 to 14 ms. Stress testing of the FSs show them meeting system requirements at 600 requests/second.

3.6.2 Reliability

Since the 1A Processor system is expected to provide essentially continuous on-line service, the file store, as part of that system, must be highly reliable.

File stores are duplicated for reliability, but mates must be able to operate independently, thus running in step and matching cannot be used as the error-detecting scheme. As a result, the FS must be self-checking. Many self-checking circuits are provided to assure proper operation, particularly to minimize the probability of mutilating CS/PS data.

One important circuit area that is checked is the registers containing the job-control information, the DR block. These work registers are duplicated and as each word is received, it is gated to duplicate registers over different buses at different times. The contents of the registers are then matched to verify that words arrived in the work registers.

The record address loaded in the job queue is compared with the record address in the DR block to ensure that the correct job was accessed. Since this address has to be translated both times it is received, this procedure also checks the translator.

Addresses generated by the FS to access CS/PS are checked within the FS. Each address sent to the output-address register is returned to the copy in the work registers and matched. This checks the bus path within the FS. Also, the duplicate copies of the CS/PS address are independently incremented and then matched to determine that the addresses were correctly incremented.

Data is protected in the FS by parity (when it is moving over an internal bus) or by a cyclic redundancy code (when on disk). The cyclic code also contains address information and thus provides a check on track-selection failures.

System reliability requirements for the FSs have been demonstrated. With an excess of 500 thousand hours on the disk files, they are exhibiting a meantime-before-failure of four years.

3.6.3 Low system overhead

Low system overhead is achieved by allowing the FS to autonomously process jobs. Once a job has been loaded in the hardware queue, the FS processes the job without system intervention. Following a sector-address match, the FS accesses CS to obtain the complete job-control information. Direct memory access then permits data transfers with a minimum of data-processing cycles stolen from CC. Only when a job status word has been returned to CS by the FS is system action again required.

3.6.4 Diagnosability

The objective of the FS diagnostic program is to detect at least 95 percent of all faults and to resolve these to either the controller or a disk file. For the case of controller faults, fault resolution to two replaceable circuit packs on the average is the objective.

To meet these objectives, diagnostic-test design was done in parallel with the hardware design. This resulted in the addition of maintenance-access circuits that are necessary for testing the error-detection circuits as well as for testing basic operations. Maintenance access includes 300 nodes that can be read to check circuit operation plus 600 control nodes that can be activated to test FS circuitry. Maintenance circuitry plus error-detection circuitry account for over 30 percent of the FS circuitry.

3.6.5 Equipage

The FSs are required to be growable to meet the memory requirements of a range of systems. The FSs can grow from a minimum equipage of one disk (duplicated) to a maximum system composed of two communities of FSs with four disks in each file store. The typical No. 4 ESS will have one community with four disks per FS and the typical No. 1A ESS will have one community with three disks per FS.

IV. AUXILIARY DATA SYSTEM

The auxiliary data system (ADS) provides the 1A Processor with a flexible and efficient means of transferring data from call- or program-store memory to magnetic tape or from magnetic tape to memory. The ADS structure permits other data input or output devices to be attached to the 1A Processor in a manner similar to that of the magnetic tape.

The initial uses of the auxiliary data system are:

(*i*) Automatic message accounting—The recording of call data for subsequent processing for billing purposes.

(*ii*) System reinitialization and system update—System reinitialization is required for initial loading of generic programs and office data into CS/PS memories. Subsequent system reinitialization could be required if data multilation in those memories should occur simultaneously with a duplex failure of the backup information in the file-store memories. Because of its severity, system reinitialization (SR) can only be initiated manually.

System update, on the other hand, is used to load new generic program or office data into one copy of the backup memory while the system continues to operate on the original program and data. After the system

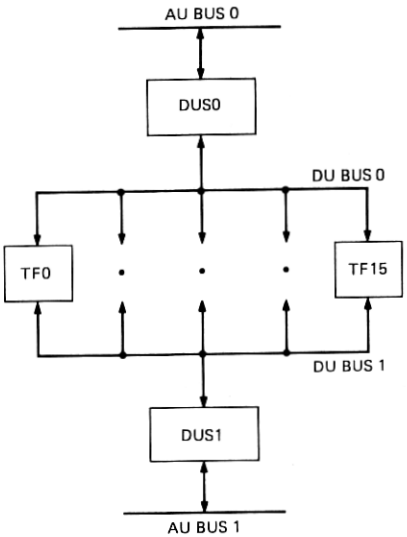

Fig. 11—ADS connection to AU bus.

update is complete and verified, it is loaded into the appropriate areas of active call and program-store memory.

(*iii*) Trouble location—Data bases used for generating lists of suspected circuit packs corresponding to diagnostic failure patterns are stored on magnetic tape.

(*iv*) General use—A number of uses of the tape system are made in addition to the above. These involve outputting of operational or maintenance data from call- or file-store memory, loading special nongeneric programs, verifying the resident copy of generic program and office data, and writing new generic and office-data tapes.

4.1 ADS organization

As indicated in Fig. 11, the ADS consists of data unit selectors (DUS) and tape frames (TF). The 1A Processor can have up to two pairs of DUSs. The even numbered DUS of the pair is connected to AU bus 0 and the odd-numbered DUS is connected to AU bus 1. Each DUS has access, via the data unit (DU) bus, to up to 16 TFs. Both DUSs of a pair access the same TFs. Several TFs can have data-transfer jobs in progress simultaneously since the DUS serves to multiplex their access to memory. The maximum data-transfer rate through a DUS depends upon blockage that can occur due to memory conflicts with higher-priority AU bus users. In

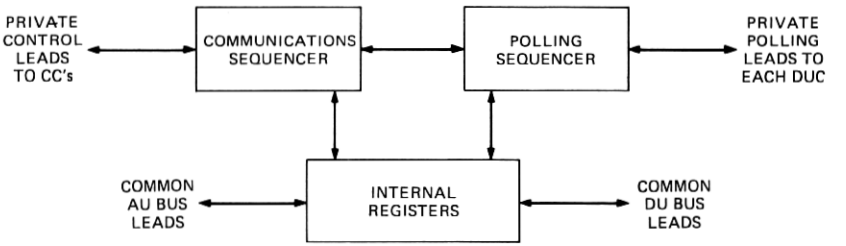

Fig. 12—Data-unit selector organization.

the absence of blockage, DUS throughput can exceed 2×10^6 bits/second.

The organization of the DUS is shown in Fig. 12. The AU bus and DU bus communications are under the control of sequencing logic that gates address, data, and control information from the AU or DU bus to internal DUS registers and vice versa. Polling control is also shown in Fig. 11. The polling function is normally performed by the DUS if it is not in the process of relaying data between memory and a TF. By administering the polling-control logic, the 1A Processor programs can cause the DUS to poll all, none, or any subset of the equipped TFs. The TFs response to the poll will indicate whether a data transfer is being requested (read or write) or if other attention is required. Other attention might be of a normal operational nature; for example, as a result of a successfully completed job, or of a maintenance nature, such as a result of a failure of some operation in the TF. If the poll response indicates a read or write request, the DUS will request access to core memory through the central control. When the central control permits access, the DUS initiates the memory transaction. This is accomplished without program intervention in the central control and normally does not require stealing cycles from program execution. If the poll response indicates a need for operational or maintenance attention, the DUS will cause an operational or maintenance-request flag to be set within central control so that appropriate programs will take action.

The TF contains a tape unit controller (TUC) and a tape transport. The TUC contains circuits for sequencing commands, checking for errors, and maintaining an orderly data flow to the transport.

The salient operational features of the TUC are depicted in Fig. 13. The address register specifies the memory address that the TUC is to access. The character count register contains the count of the number of 8-bit characters remaining to be transferred during a read or write job. The add 1 logic is used to increment the address register each time a word is transferred to or from memory. It is also used to decrement the character count each time a character is transferred to or from the tape unit.

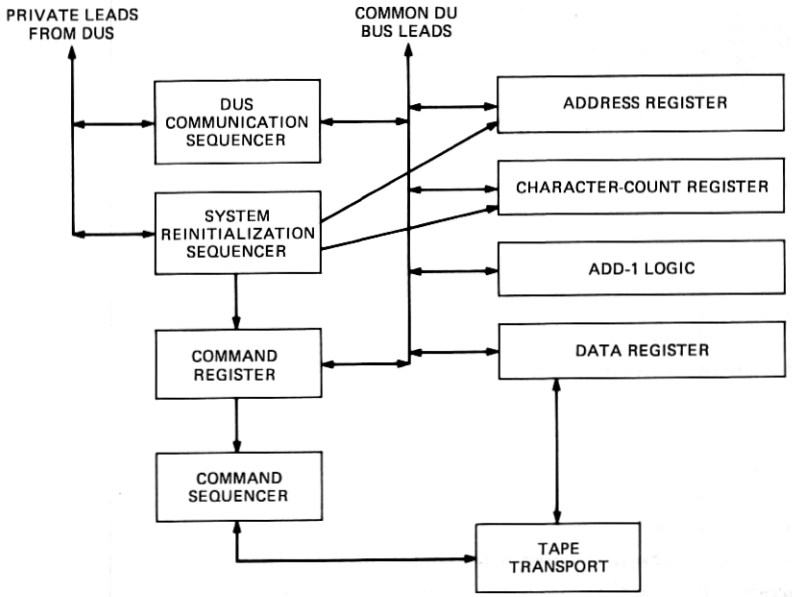

Fig. 13—Tape-unit controller organization.

The command register specifies the job such as read, write, backspace, etc., that the TUC is to perform. The data register serves as a buffer for data en route between the tape unit and the DUS. The TUC-to-tape-transport communications are in units of eight data bits plus parity while communications with the DUS are in units of 24 bits plus parity. Thus, the data register provides a 3-to-1 mapping of characters to words and vice versa.

When the system program initiates a read or write job in TUC, it writes the starting memory address into the address register, the job size into the character-count register, and the command code into the command register. The TUC will then process the job under control of the command and communication sequencers. Successful job completion is indicated by returning an operational interject response to the DUS poll. The interject response causes the DUS to set an interject source bit in the central control, which notifies the program that attention is required.

As mentioned previously, system reinitialization is a function of the ADS. An SR request causes the SR sequencer in the selected TUC to initialize the address, character count, and command registers in a manner that will cause a bootstrap program to be loaded into memory. Once the bootstrap program is loaded, it controls the loading of subsequent blocks of program and call-store memory.

The tape transport is a 9-track, 800-character-per-inch unit, which

reads and writes at a speed of 25 in./s. The data rate (excluding parity) is 160 kb/s. Fast forward and rewind speeds are 100 in./s. The tape unit is powered from a converter that enables it to operate from −48 V office battery.

4.2 Trouble-detection-and-reporting techniques

The ADS incorporates a number of techniques to detect errors. These include parity over data and addresses, sequencer- and decoder-validity checks, echo checks on select and data signals between DUSs and TUCs, and implementation of the standard tape error-detection codes and read-after-write data checks. One of the most important error-detection features is incorporated in the add-1 circuit since incrementing failures could cause incorrect address generation and result in multilation of call or program memory. The add-1 circuit is self-checking and can detect any classical fault (gate input open, output stuck at 1 or 0) that would affect the outcome of the add-1 operation. When an error is detected, the ADS will suspend further accesses of memory and notify the program either via the operational-interject response or by the maintenance-interject response that is also used by the file store for trouble reporting. The program will then test the unit reporting the error and possibly remove it from service.

V. CONCLUSION

The 1A Processor memory systems herein described have demonstrated their reliability and their ability to meet performance goals. Performance has been verified by the systems now in operation; five systems are operating at Bell Laboratories, Indian Hill, Illinois and five No. 4 ESSs with their 1A Processors have been placed in commercial service.

VI. ACKNOWLEDGMENTS

This report is based on the work of many people in Bell Laboratories. All store designs were achieved by close cooperation between store-design groups and the diagnostic, fault-recognition, and system-design groups.

REFERENCES

1. A. H. Budlong et al., "Control System," B.S.T.J., this issue, pp. 135–179.
2. J. G. Chevalier and M. W. Rolund, "New Memory Reduces No. 1 ESS Cost and Size," Bell Laboratories Record, *50*, No. 4 (April 1972), p. 121.