

LA-UR-76-2502

Contract #DE-AC03-74-107

TITLE: COMPUTER INTERFACING AT LASL

AUTHOR(S): John A. Brockmeyer & A. Gene Dornhoff

SUBMITTED TO: AESOP XV/SCIE Conference
September 28-30, 1976
Denver, Colorado

NOTICE
This report was prepared as an account of work sponsored by the United States Government. Neither the United States nor the United States Energy Research and Development Administration, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness or usefulness of any information, apparatus, product or process disclosed, or represents that its use would not infringe privately owned rights.

By acceptance of this article for publication, the publisher recognizes the Government's (license) rights in any copyright and the Government and its authorized representatives have unrestricted right to reproduce in whole or in part said article under any copyright secured by the publisher.

The Los Alamos Scientific Laboratory requests that the publisher identify this article as work performed under the auspices of the USERDA.



An Affirmative Action/Equal Opportunity Employer

Form No. 806
St. No. 2629
1/75

UNITED STATES
ENERGY RESEARCH AND
DEVELOPMENT ADMINISTRATION
CONTRACT W-7405-ENG-36

COMPUTER INTERFACING AT L.A.S.L.

INTRODUCTION

The Los Alamos Scientific Laboratory (LASL) has experienced a tremendous growth in time-shared computing, remote job entry (RJE) and common file storage in the past few years. This in turn has lead to the development of an Integrated Computer Network (ICN) at its Central Computing Facility (CCF). This report describes the network, the hardware interfaces used to implement it, and plans for future improvements.

THE CENTRAL COMPUTING FACILITY (CCF)

The CCF at LASL contains 9 large-scale computers and numerous minicomputers. Seven of these large computers run users' jobs and as such are called Worker computers. Another acts as a file storage controller, and the ninth is in the process of being evaluated for possible procurement and use as a Worker. See Fig. 1.

Two 7600's run the LASL-developed Chili Ridge Operating System (CROS). This is a monoprogramming system with RJE capability and common file storage access. It is optimized for executing large codes for long periods of time. The other two 7600's run the Lawrence Livermore Laboratory developed Livermore Time-Sharing System (LTSS). This operating system is very sophisticated and

allows both interactive and batch execution. It, too, has RJE and common file storage capability. The Control Data Corporation (CDC) Network Operating System (NOS) is run on three CDC 6000 series machines and is a time-sharing system, but simpler than LTSS. It will soon support common file storage usage. One 6400 (actually a CYBER 73) runs the Hydra operating system, which primarily handles the RJE and common file system. This machine connects over an IBM interface to an IBM 1360 Photostore, which can contain up to one trillion bits of archival storage. Hydra also has large amounts of disk storage for more dynamic file storage.

User interaction with this group of machines is done through 500 low-speed terminal ports and 23 RJE stations scattered throughout the laboratory. In addition there are both terminal and RJE dial-up facilities.

The primary hardware links which presently interconnect the various parts of the network are shown in Fig. 2. The PDP-11 Link, which runs serial full duplex up to 300 Kbps, consists of a PDP-11, a Remote Terminal Line Interface (RTLI), two modems (optional), another RTLI, and a second PDP-11. The 6000 Link, which operates 12-bit parallel half duplex up to 4-6 Mbps, is made up of a PDP-11, PDP-11 Coupler, 6000 Synchronizer, 6000 I/O Channel, and a 6000 Peripheral Processor (PP). The 7600 Link is also 12-bit parallel half duplex, runs up to 10 Mbps with a short link cable, and consists of a PDP-11 and Coupler, 7600

Synchronizer, 7600 I/O Channel, and a 7600 PP.

PDP-11 UNIBUS INTERFACE

The ICN nucleus consists of a group of PDP-11's which connect to each other, the terminal ports, the RJE stations, and the Worker computers. Since the PDP-11 is part of so many interfaces, the hardware interface description will begin here.

The PDP-11 I/O is centered around its UNIBUS. All devices, including memory, connect to this Bus. All devices have control registers which are considered part of the machine's address space, and can be directly addressed by the computer. Normally the computer is the Bus Master and controls data transfers on the Bus, but any device can become Bus Master and use the Bus for Direct Memory Access (DMA) transfers into or out of memory. Arbitration for this depends on the physical ordering of the devices on the daisy-chained Bus. In addition, a device can become Bus Master and request that the processor interrupt to a chosen address. There are four levels of interrupt, and for each level priority is also by physical location on the Bus.

PDP-11 LINK - THE RTLI

Connection between Network PDP-11's and between the Network and RJE stations is the task of the Remote Terminal Line Interface

(RTLI). The RTLI allows two PDP-11's to communicate over a full-duplex link. This device is a full-duplex, DMA-oriented interface at the UNIBUS and a serial full-duplex synchronous EIA or differential device at the link end. The serial protocol is 16-bit oriented, with a 16-bit sync word and Cyclic Redundancy Check (CRC) character. The output (TX) from the Bus and the input (RX) to the Bus are completely independent. Each half has its own command-status register, memory address register, and word count register. These are set up by the PDP-11 for each block transfer, and read to determine the success of the transfer. When connected by modems, the TX section starts sending data as soon as its start bit is set. The RX end will only accept data if its start bit is set and it has seen the SYNC word. If the RX is not ready when the TX starts, the data is lost. When connected by wires, another pair of lines signals the TX to wait for the other end's RX start bit to be set before sending data, and no data is lost. When either section completes the transfer, an interrupt occurs. The software sees whether the other end's RX section was ready in time and received the block. Thus each end is in control of its own output and input; there is no master control. The RTLI can run up to 300 Kb full duplex, though the RJE stations run at 9600 baud through modems. In the CCF the differential lines are used, and the transfer rates are 50 or 150 Kb. There is also a self-test feature, which connects the TX to the RX section for local testing.

THE 6000 LINK - THE PDP-11 COUPLER

The PDP-11 end of the 6000 Link is similar in some respects to the RTLI. It, too, has a DMA capability, similar registers, separate TX and RX sections, end-of-block interrupts and self-test capability. However, the Coupler's connection to the 6000 computer, called the 6000 Synchronizer, is 12-bit parallel plus parity, half duplex, with handshaking data transfer and end-of block signaling. As shown in Fig. 3, the Coupler reorganizes three 16-bit PDP-11 words into four 12-bit 6000 PP words or vice-versa, and swaps the bytes going to or from the Bus. This occurs because the PDP-11 puts its bytes in memory the opposite of the other computers.

Since the interface is half duplex, one end or the other must be in control of the link. In this case the 6000 PP is Master, and sends Functions through the 6000 Synchronizer to the Coupler which determine the operation to be done. The Coupler has to approve of the Function or nothing will happen. This approval is done by bits in the Coupler's command-status registers.

THE 6000 LINK - THE 6000 SYNCHRONIZER

The CDC 6000 series I/O is handled by a series of 10 (or 20) Peripheral Processors with a 12-bit word and 4K memory. These PP's access the 12 (or 24) I/O Channels through a multiplexer.

The I/O Channels are synchronous, 12-bit, half duplex, with signals in the form of 25 ns pulses. See Fig. 4. The channels have 12-bit data in and out, data control signals (FULL and EMPTY), block control signals (ACTIVE and INACTIVE), and the special signals FUNCTION and MASTER CLEAR. The FUNCTION controls the operation of the link, and MASTER CLEAR resets the 6000 Synchronizer when the 6000 deadstarts.

The 6000 Synchronizer connects to the I/O channel over two 75 ft. cables. For compatibility the 6000 Synchronizer uses CDC QJ receiver-transmitter modules to connect to the cables. They use 6000 RTL, while the rest of the Synchronizer is TTL logic. The 6000 Synch interfaces the 6000 pulses to the static link signals, sends the functions and controls the operations, generates and checks parity for the link data, and disconnects the PP from the I/O Channel in case of hangups. The 6000 Synchronizer has a 4-6 Mbps data rate, depending on the link cable length, which can be up to 300 feet long.

As shown in Fig. 3, the functions (operations) are few. The only odd one is the Buffer Request, which loads two bits in the RX command-status Register and generates an interrupt in the PDP-11. This feature can be used by the software for link control.

THE 7600 LINK - THE 7600 SYNCHRONIZER

This link uses the same Coupler at the PDP-11 end. However,

the 7600 PP's and I/O channels are somewhat different. The 7600 PP's are also 12-bit with 4K of memory, but each PP has its own eight I/O channels. These channels use differential ECL signalling, with 27 ns pulses for the control signals and static levels for the data. Because these channels are asynchronous, any cable length up to 50 ft. can be used between the PP and its Synchronizer. Like the 6000 Channel, the 7600 Channel has 12 data bits in and out, data and block control signals, but no FUNCTION signal (see Fig. 4). Therefore, the OUTPUT RECORD FLAG followed by an OUTPUT WORD FLAG serves as the FUNCTION signal. The only reset for the Synchronizer is manual. The Synchronizer interfaces the channel signals to the link signals, sends the function and controls the operations, generates and checks parity for the link data, and unhangs the PP on most signalling errors. The PDP-11 to 7600 link can run up to 10 Mbps, depending on the link cable length.

GENERAL COMMENTS

What have we learned from this design experience? The first is that design should be kept simple and universal. The links should have little or no knowledge of the structure of the data they pass. This enables them to be used with anyone's software. Always make a link full duplex or establish one end as master. LASL bought quite a number of DEC DA-11B Unibus Links which turned

out to be unusable because it was a half duplex device with neither end chosen as master. As a result, the software protocol was difficult and 5 interrupts were needed between the two ends to terminate a block. Subsequently, the RTLI was chosen to provide this link instead.

Reliability, recoverability, and replaceability are musts in a computer network. Who needs a seven million dollar computer sitting idle because its thousand dollar link to its users went bad? Because the CDC PP's can hang on a channel operation and lock up a whole system, the links must be able to recover from transmission errors and keep the channels working. Very little of the gear that has been designed has been multiplexed, that is, several links in one piece of hardware. It was felt that even though this is more expensive, it is easier to check them out and replace faulty gear without knocking down any more of the network than necessary.

If self-test can be put into an interface easily, it should be, since it speeds up checkout and fault localization. Also helpful are detailed front-panel indicators for the same reason. PDP-11-driven 6000 and 7600 Channel simulators are very useful, as a 6000 or 7600 is a costly hardware checker.

Discrete TTL, especially LS series recently, packaged in wire-wrapped cards has been used almost exclusively. TTL is good for its MSI capability and simple connection requirements. Wire-wrap is preferable for our operations because we build small

numbers of devices which very often get modified because of software requests. No microprocessors have as yet been used² in any designs because they have not been fast enough.

FUTURE ICN IMPROVEMENTS

The LASL CCF and ICN are constantly undergoing changes designed to improve performance and capacity. Two of the most significant improvements will be the addition of the File Transport (FT) subsystem and the addition of one or more Class VI computers. One possible configuration of the future ICN is shown in Figure 5.

FILE TRANSPORT (FT) SUBSYSTEM

The File Transport (FT) equipment is an example of a subsystem which may be added to the LASL ICN to upgrade the network in a gradual manner, with no extended downtime or inconvenience to users. Existing links will be gradually replaced as the FT is placed on line. The FT will consist of two Systems Engineering Laboratories (SEL) Model 32/55 midicomputers and the interfaces necessary to communicate with the various other machines in the ICN.

Each SEL 32/55 accommodates up to 128k 32-bit words of directly addressable memory. The maximum I/O transfer rate is approximately 26.7 Mbps over any one high speed I/O channel, and slightly greater than 100 Mbps over four or more channels simultaneously. The addition of the FT to the network will

provide the capability to transmit large data files not feasible with the PDP-11's presently in the network. in addition to supporting data transmission rates much faster than is possible with a PDP-11.

All high-speed connections between external devices and the SEL machines will be made using Model 9132 High-Speed Data Interfaces (HSD's). This is a change from the practice followed in the past by LASL in the design of PDP-11 interfaces, where it has been the practice to connect directly to the DEC UNIBUS. The change is due to the basic differences in the busses of the two machines. The UNIBUS is asynchronous and allows long interface cables with few critical timing requirements. The SEL bus is a much faster synchronous design with more rigorous requirements on the bus length and logic timing. In addition, it is not as well documented as the UNIBUS at this time. To minimize development time and the related risks, the standard product HSD's will be used. The HSD, as shown in Figure 6, provides a simple handshaking interface to the external device, simplifying controller logic design. All memory address registers and transfer count registers are provided by the HSD. The HSD is a modified version of SEL's I/O Microprogrammable Processors (IOM's); future designs may be able to take advantage of the microprogrammable feature by using modified firmware to replace part of the external device logic.

The HSD is capable of running at speeds in excess of 25 Mbps over 50 foot cable lengths, including reasonable delays for the external device logic. The goal for all future LASL-designed interfaces is to maintain the 25 Mbps rate for all devices capable of supporting that speed. For slower devices, the external device will determine the speed of the link, with the device being allowed to run at its maximum speed.

Lower speed data links which require speeds of approximately one Mbps or less will be interfaced using the SEL General Purpose Multiplexer Controller (GPMC) and General Purpose Device Controller (GPDC) subsystem. The GPMC is an IOM which is installed on the SEL bus, representing one load to the bus. Up to 16 GPDC's may then be connected to the GPMC on a second bus (generated by the GPMC) referred to as the GPDC bus. The GPMC acts as a multiplexer, increasing the number of external interfaces available for low-speed devices.

Each GPDC contains a simple microprogrammable logic section and provides space for device dependent logic to be added. The primary use by LASL will be for communications with PDP-11's.

Class VI Computer Development

A CRAY-1 computer is currently undergoing evaluation at LASL for possible procurement. Concurrent with that evaluation, the I/O channels have been studied for consideration of ways to interface the CRAY-1 into the LASL ICN should the evaluation be favorable and procurement follow. As this information is of interest to many other ERDA laboratories, some of the general characteristics of the channels and conclusions reached will be presented.

The CRAY-1 I/O section contains 24 channels, of which 12 are input channels and 12 are output channels. The channels are designed so that a pair of channels (one input and one output) provides high-speed full-duplex I/O capability. Two basic types of channels are of interest for interfacing. Each provides 16 data lines and three control lines for each of the input and output channels. Data parity will be included on future machines. The slower type requires a complete control handshake for each 16-bit parcel transferred. This slows the maximum usable transfer rate to something on the order of 25-50 Mbits/second (Mbps). depending on cable lengths. without allowing any delays for the external equipment.

A block diagram of the higher-speed version of the channel is shown in Figure 7. This channel is the primary interest of LASL. and largely overcomes the cable delay problem by requiring a

complete handshake only once for each group of four 16-bit parcels. This raises the absolute maximum transfer rate to a speed in excess of 100 Mb/s in each direction, and a possible rate with usable cable lengths of something on the order of 80 Mb/s. This is comfortably higher than any network interface requirement foreseen by LASL, allowing reasonable design margins in interface equipment.

In contrast to channels on other large scientific computers, specifically CDC 6000 and 7000 machines, the CRAY channels are much simpler in their modes of operation and capabilities: flexibility has apparently been traded for speed. The CRAY-1 does not contain any Peripheral Processors (PP's), and all transfers are directly between the I/O port and main memory. Any area of memory may be accessed, and no restrictions are placed on block size by the channel hardware, other than that the size must be an integral number of 64-bit words. No separate status or function commands are available, and all such information must be handled as data.

As shown in Figure 8, the high speed output channel provides four 50 ns Data Ready pulses synchronously with four 16-bit data parcels on the output data lines. The peripheral device must sample the data synchronously with the rising (trailing) edge of each Data Ready pulse. At the end of the pulse burst, the channel will wait for the peripheral device to respond with an Output Resume pulse (also 50 ns wide). The device may respond

asynchronouslv when it is ready. For the maximum transfer speed, the pulse must be received by the CRAY channel in the 200 ns window shown. The 200 ns delay between data bursts is a minimum time determined by the channel.

At the end of the block, the CRAY channel sends a Disconnect pulse synchronouslv with the final Data Ready pulse. No more data will follow in that block. The CPU will be interrupted following the receipt by the CRAY channel of the final Resume pulse.

It should be noted that no provisions are made to read status or send function commands separate from the data. This may be a problem if the CRAY attempts to send data to an inactive peripheral. Software timeouts should be provided for protection. The CRAY channel hardware will detect some error conditions, but the exact response of the machine is unclear at this time. In some long vector operations, the channel interrupt may be locked out for long periods of time. As the studies made by LASL are as yet preliminary and CRAY documentation in this area is as yet incomplete, no effort has been made to pin down this area. Because of the simple channel hardware, some error conditions will probably require software detection.

The input channel is similar in operation to the output channel, in that there are also 16 data lines and three control lines. The timing sequence begins with a burst of four Data Request pulses from the CRAY to the peripheral. The peripheral then responds by asynchronously sending four 16-bit parcels of

data to the CRAY channel along with four Data Ready pulses. The rate may not exceed one parcel every 100 ns. Following receipt of the fourth parcel, the CRAY will send another burst of four Data Request pulses. At the end of the block, the peripheral device sends a Disconnect pulse to the CRAY channel.

As with the output channel, no provisions are made for separate status or function commands. The CRAY must initiate the transfer, but no limit is placed on the time the peripheral may take to respond. The block size must be controlled by software protocol or be restricted to a maximum size consistent with buffer space reserved in CRAY memory. The channel may be locked out from memory for extended periods of time by long vector operations.

SUMMARY

The conclusions reached as part of the preliminary evaluation of the CRAY channels include the following points.

1. The speed of the High Speed channels is adequate for our requirements with reasonable margin for conservative interface design.
2. More command and control capabilities of the channels for status and function transfers would be a convenience and would probably simplify the software requirements, but the channel

provided is adequate.

3. Due to the possible lockout of the channels by long vector operations, buffering must be provided for any device (such as a disc) that cannot be asynchronously stopped and re-started without loss of data.

4. A full-duplex link to our network would best take advantage of the full-duplex nature of the channels.

5. A general purpose full-duplex communications protocol with many of the features of IBM's SDLC or DEC's DDCMP seems best suited to the channel. The actual protocol to be used, however, must also consider the software impact on the remainder of the LASL network.

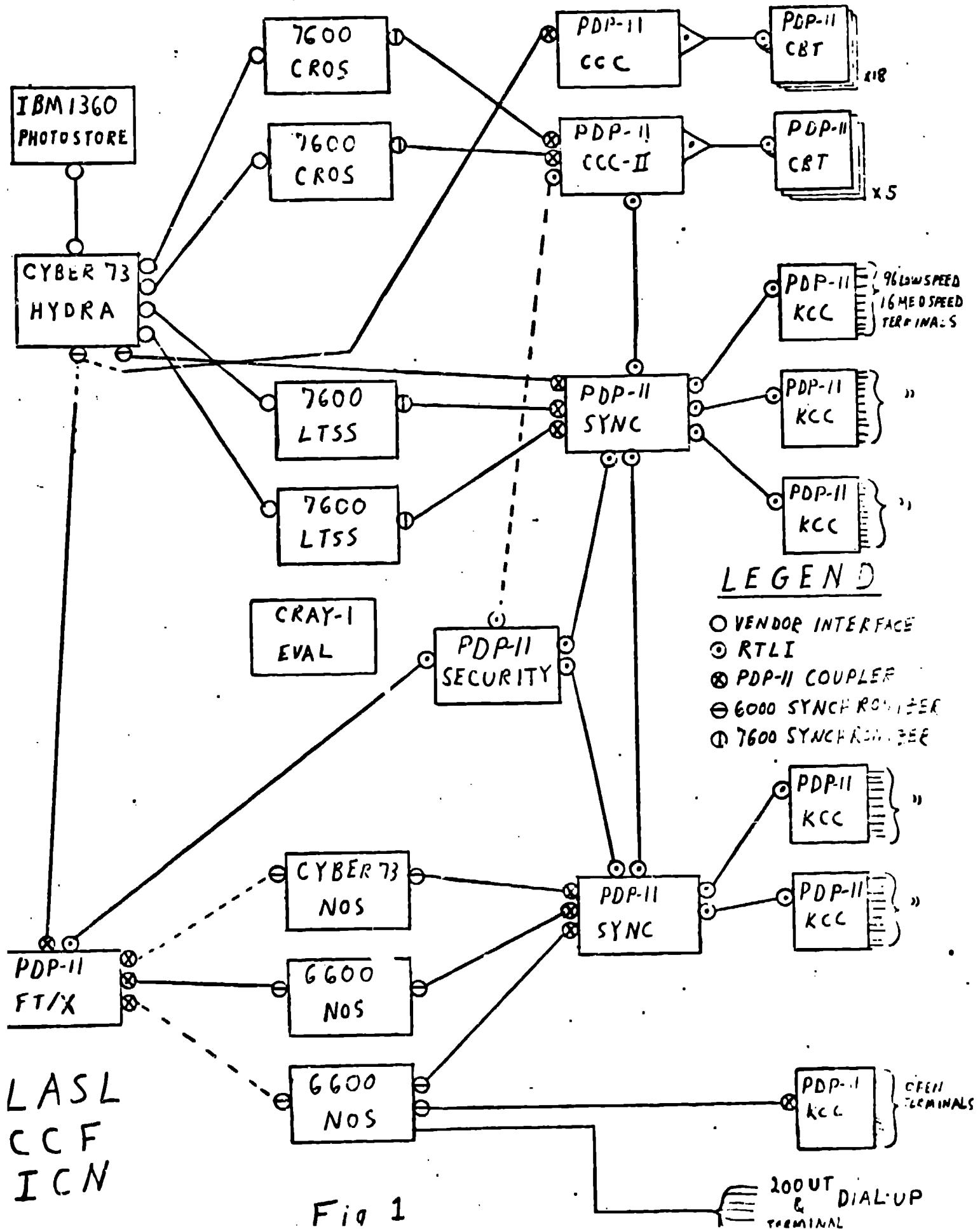
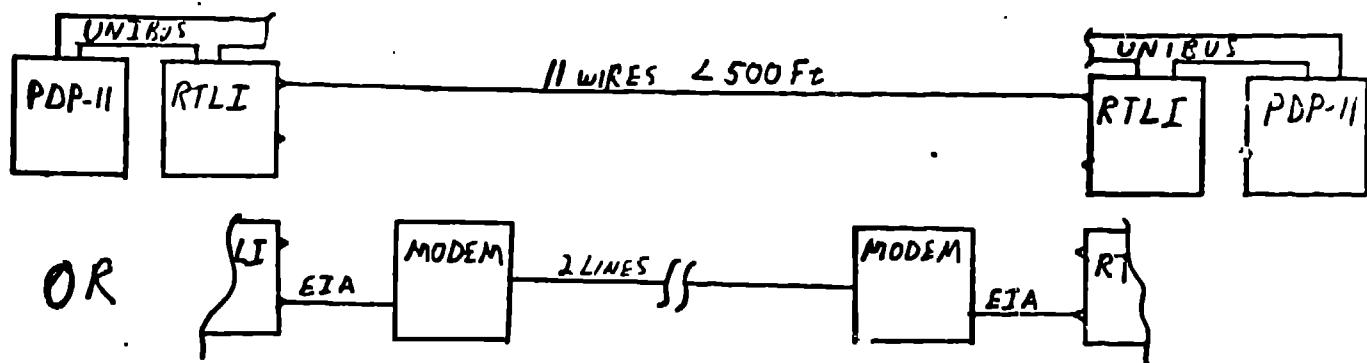
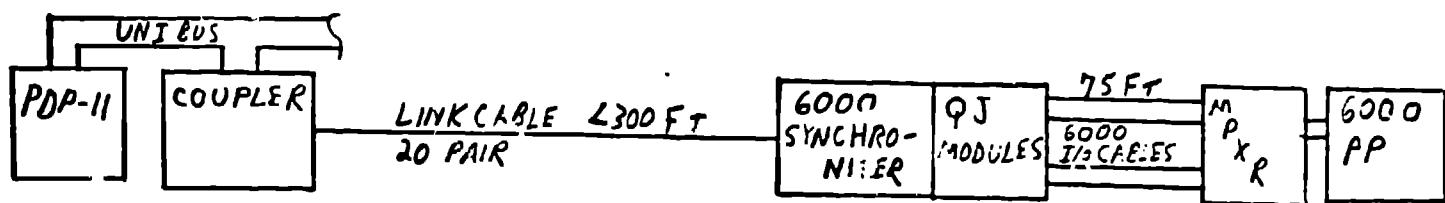


Fig. 1

PDP-11 to PDP-11 LINK



PDP-11 to 6000 LINK



PDP-11 to 7600 LINK

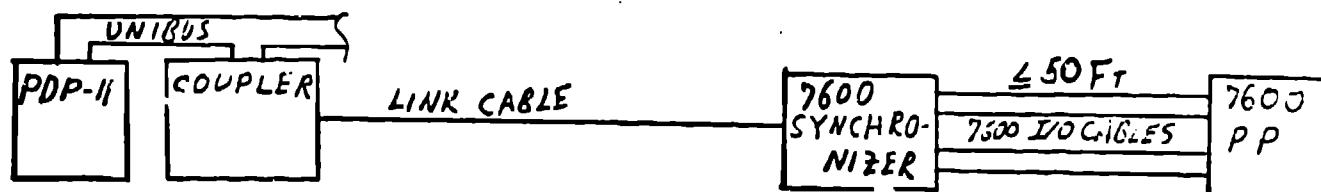
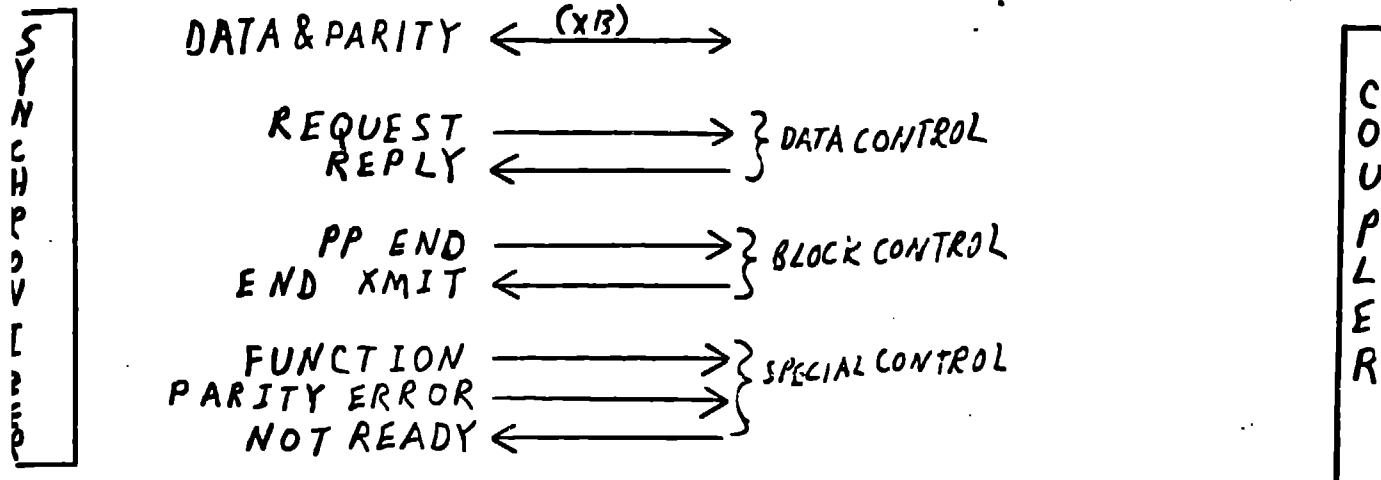


Fig 2

PDP-II TO SYNCHRONIZER LINK SIGNALS



FUNCTIONS

00 READ DATA FROM PDP-II (TX)

10 WRITE DATA TO PDP-II (RX)

02 READ TX STATUS

03 READ RX STATUS

04 BUFFER REQUEST 00

05 01

06 10

07 11

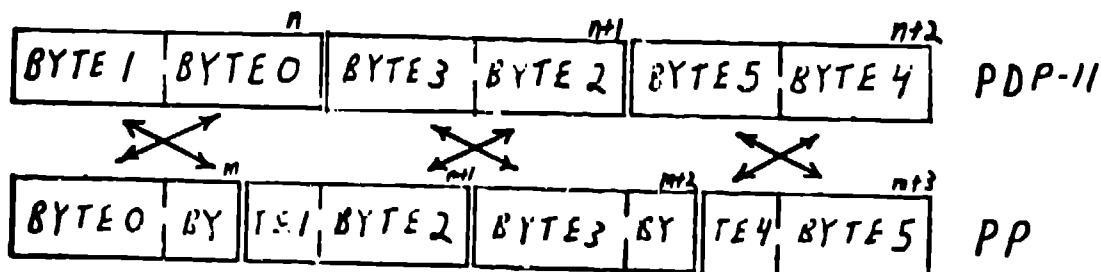
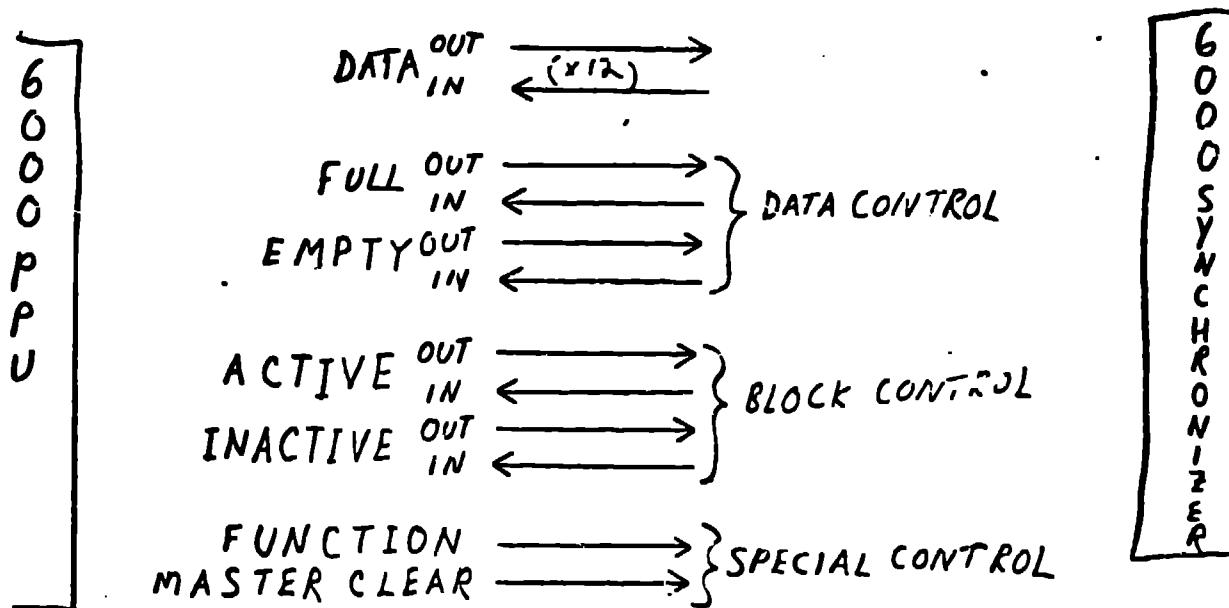


Fig 3

CDC 6000 I/O SIGNALS



CDC 7600 I/O SIGNALS

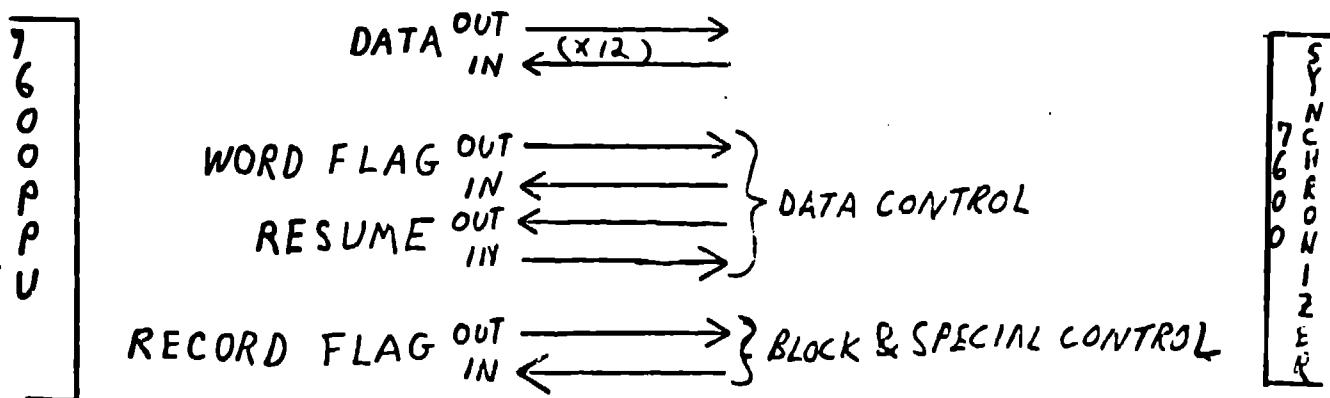


Fig 4

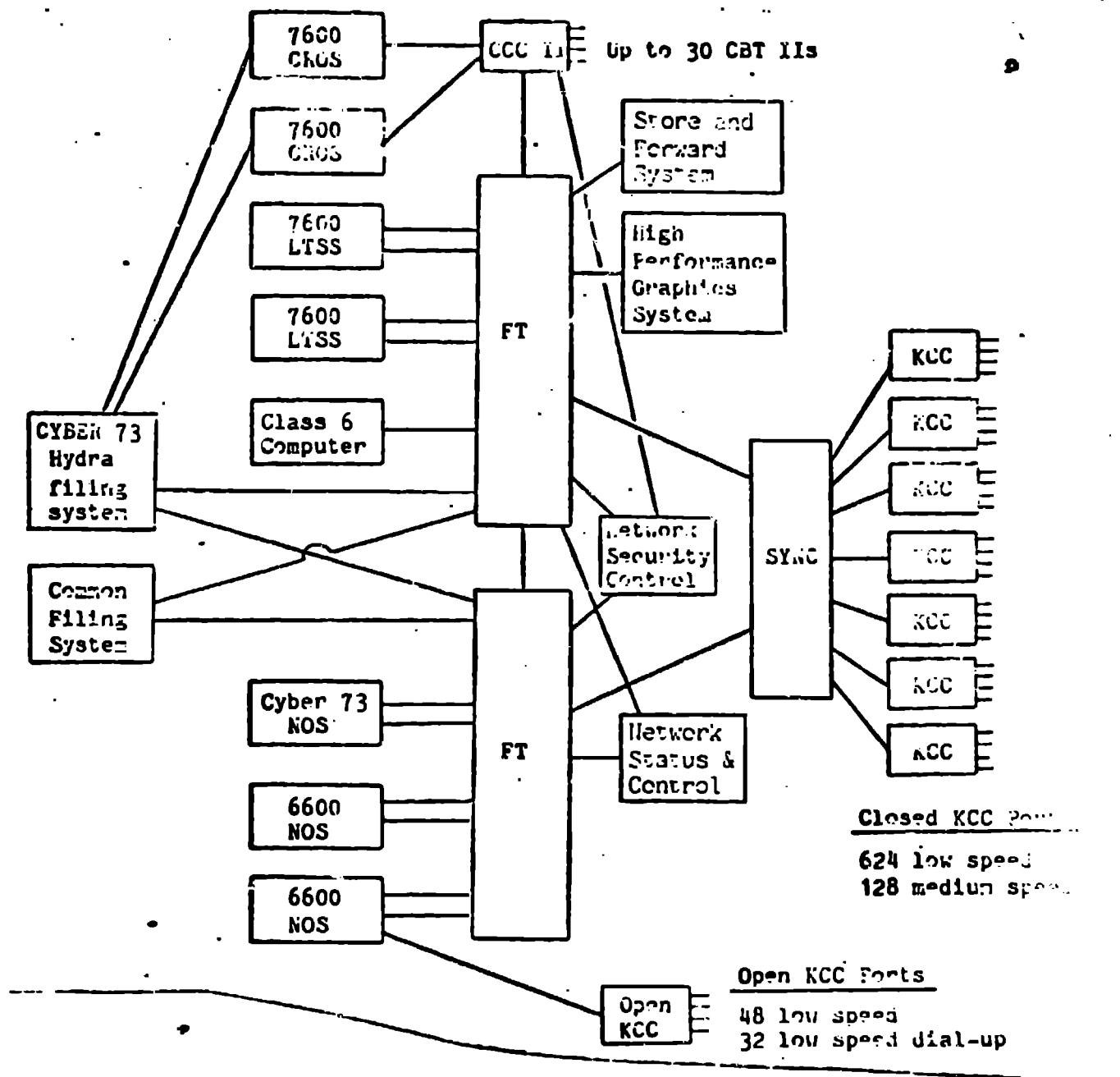


Figure 5. Possible Future ICN Configuration

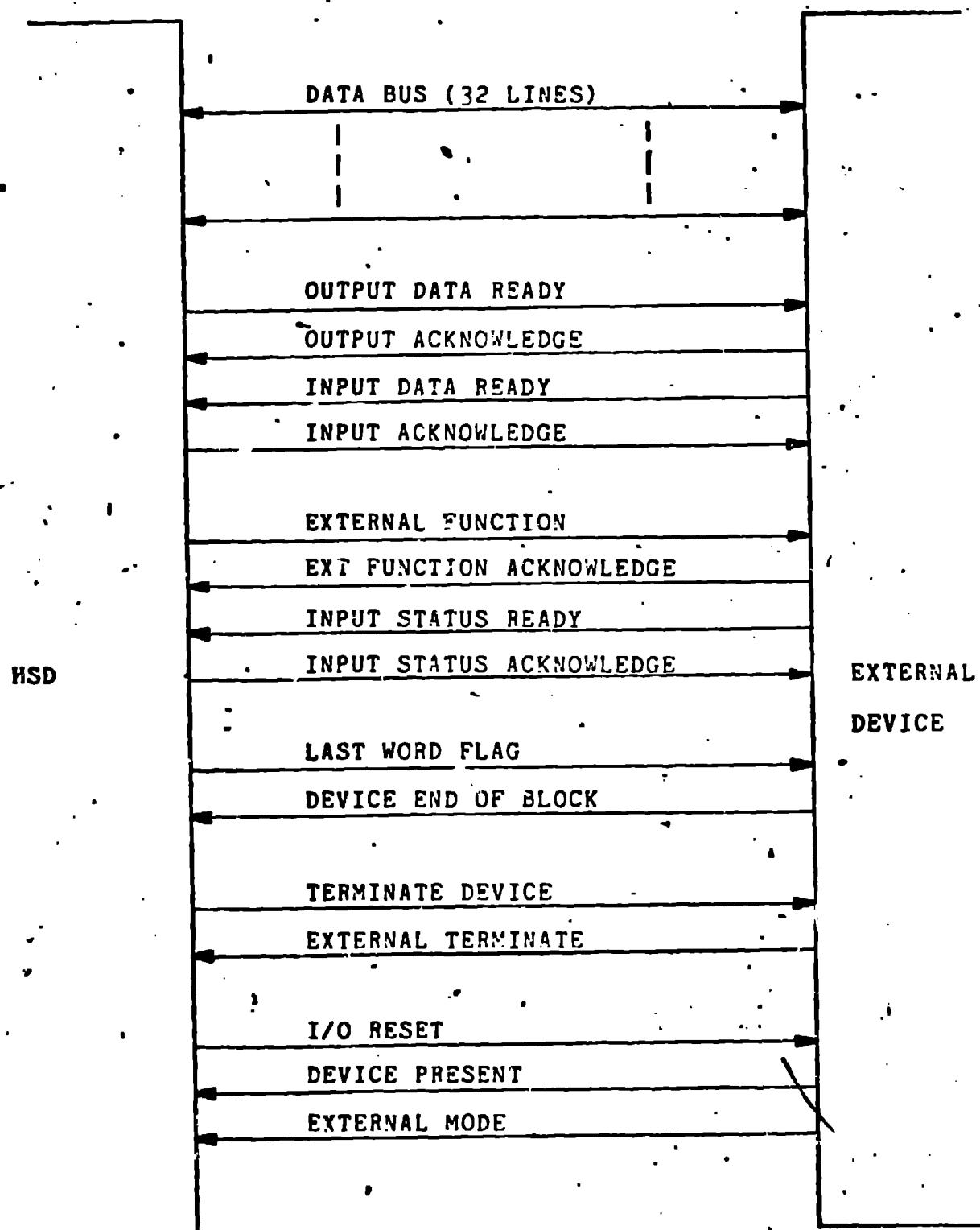


Figure 6. HSD Block Diagram

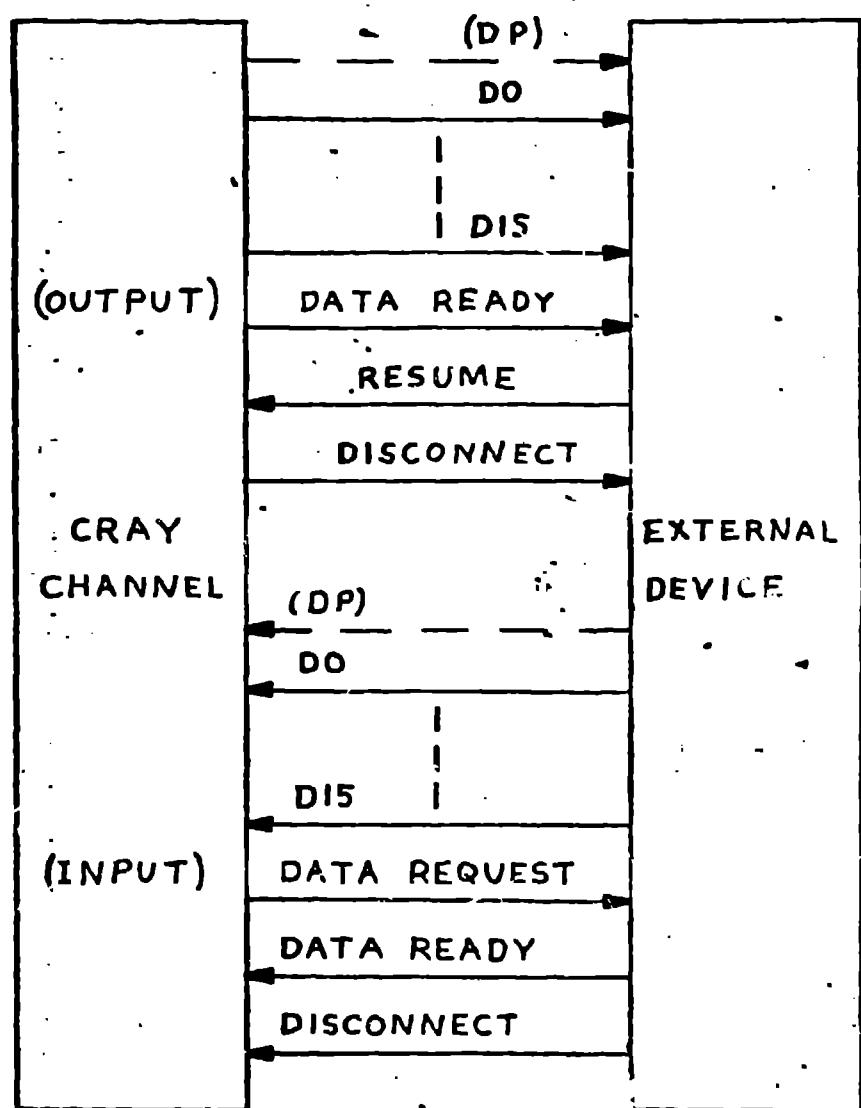
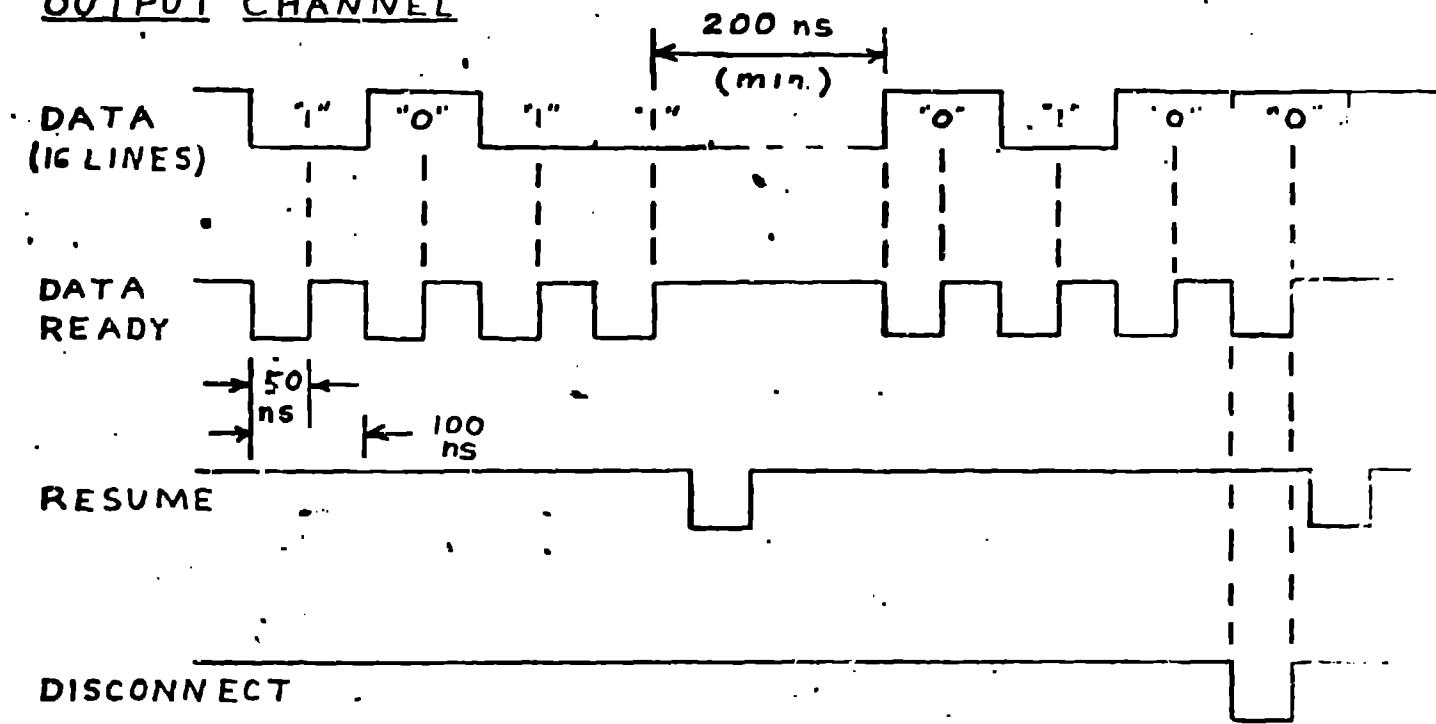


Figure 7. CRAY Channel Block Diagram

OUTPUT CHANNEL



INPUT CHANNEL

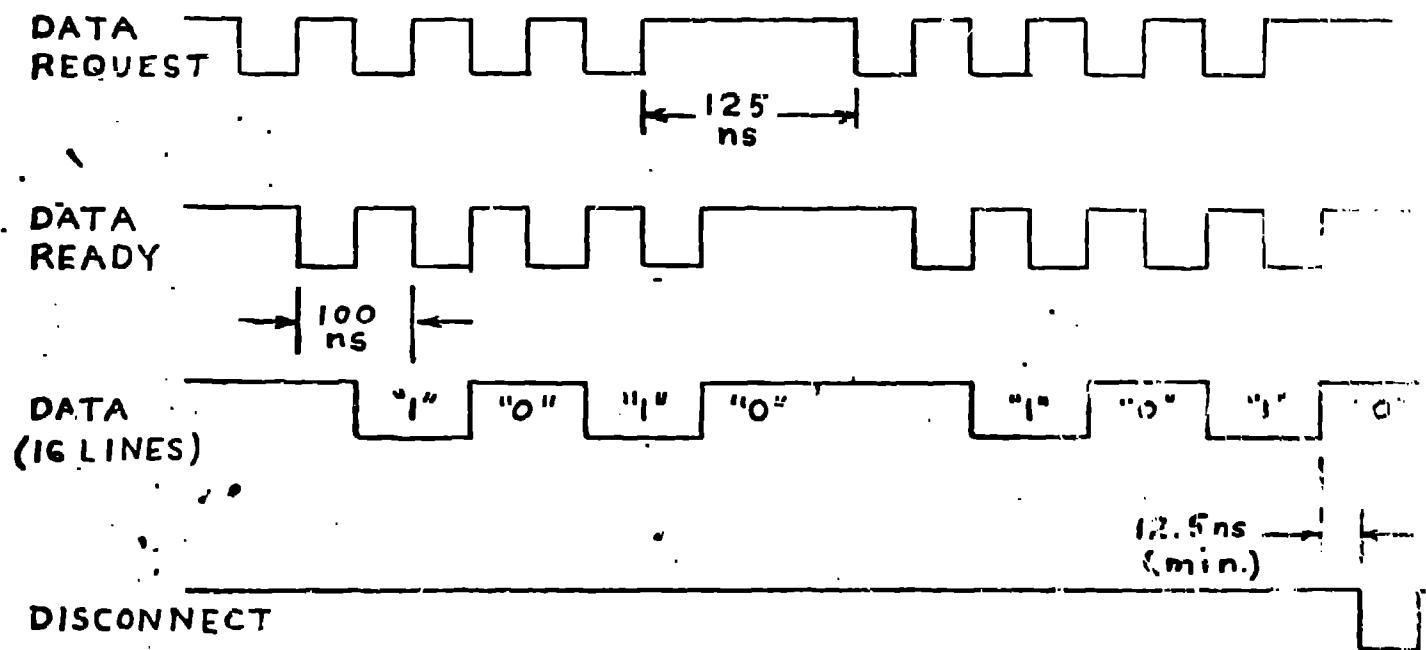


FIGURE 8. SIMPLIFIED CRAY CHANNEL TIMING DIAGRAM