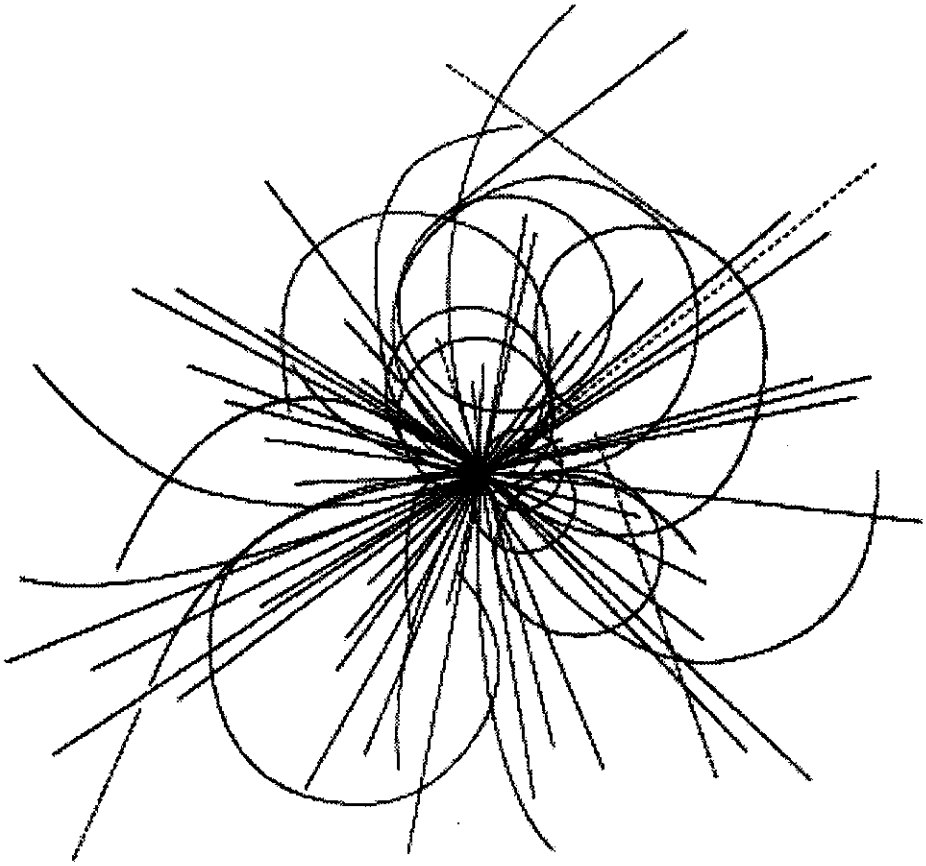
SSCL-N-861

SSCL-N-861
May 1994
Distribution Category: 400

D. Gurd

Communications Level IV Development Specification



Superconducting Super Collider Laboratory

SSCL-N-861

Communications Level IV Development Specification

D. Gurd

Superconducting Super Collider Laboratory*
2550 Beckleymeade Ave.
Dallas, TX 75237, USA

May 1994

*Operated by the Universities Research Association, Inc., for the U.S. Department of Energy under Contract No. DE-AC35-89ER40486.

CONTENTS

1.0 SCOPE 1
2.0 REQUIREMENTS 1
2.1 Response Time 1
2.2 Throughput 1
2.3 Video 1
2.4 Audio 1
2.5 Data Acquisition 1
2.6 Control Signals 1
2.7 Synchronization 1
3.0 CHOICE OF COMMUNICATIONS ARCHITECTURE 2
3.1 Single Mode Fiber-Optic Cable 2
3.1.1 Advantages of Fiber over Copper 2
3.1.2 Fiber-Optic Conduit Layout 3
3.2 SONET 4
3.2.1 SONET Signal Hierarchy 4
3.2.2 SONET Overhead Channel 4
3.2.3 Multiplexing 5
3.3 The SONET Backbone 5
3.4 T1 6
3.5 Framing 7
3.6 Messages 7
3.6.1 Control Messages 7
3.6.2 Data Messages 8
3.6.3 Diagnostic Messages 8
3.6.4 Broadcast Messages 8
3.7 Control Room Communications 8
4.0 HIGHER LEVEL PROTOCOLS 9

1.0 SCOPE

The global controls communications specification covers the systems used to link computer equipment controlling the accelerator systems at the Superconducting Super Collider Laboratory (SSCL).

2.0 REQUIREMENTS

The control system of the SSCL is different from many existing high-energy facilities in both size and complexity of the systems used to control the accelerator equipment.

The control system is required to be available to commission and operate each machine, and to coordinate individual machines into a complex that operates at maximum performance, in terms of availability and quality of beam to experiments.

The system must pass data and control messages between any of its 2000 real-time processors and 100 Unix workstations.

2.1 Response Time

Response time of the communication system is defined as the time for a control or data message to pass from a user application in one machine to another user application in another machine. Response times vary by machine and system, from ~ 1 ms for the quench protection and magnet power supply systems to ~ 1 s for process control equipment. These response times are "worst case."

2.2 Throughput

Total throughput of the systems is ~ 1 Gbyte/s.

2.3 Video

The system must be able to transmit real-time video signals from any location in the complex to any other. Total capacity would be less than 100 channels. These signals are primarily used for surveillance monitoring in the tunnels.

2.4 Audio

The control communications system will not transmit audio signals.

2.5 Data Acquisition

Data from systems such as beam position monitors have to be transported for use in feedback systems, analysis, and archival.

2.6 Control Signals

Control signals have to be transferred between systems.

2.7 Synchronization

Requirements for synchronization, which vary among machines and systems will be satisfied by three overlapping systems: general comms, where synchronization is required to an accuracy of not better than 10 ms; message broadcast, where synchronization is required to an accuracy of not better than 100 μs; and the fast timing system, where synchronization is required to an accuracy of not better that 10 μs.

3.0 CHOICE OF COMMUNICATIONS ARCHITECTURE

Due to the requirements for a very fast deterministic response time and a desire for a single network infrastructure to link all systems, it was decided to install a Time Division Multiplexing network infrastructure.

The communication system therefore consists of a number of layers:

- Messages
- T1
- SONET
- Single Mode Fiber-Optic Cable.

3.1 Single Mode Fiber-Optic Cable

Fiber has been selected because of its high bandwidth capacity, compact size, immunity to electro-magnetic interference, and low cost.

3.1.1 Advantages of Fiber over Copper

Optical fibers offer several advantages over conventional copper wire facilities and microwave transmission systems. These advantages are discussed below.

3.1.1.1 Bandwidth capacity. Bandwidth capacity (Table 1) of fiber-optic cable is much greater than that of copper.

Table 1. Bandwidth Capacity.

Media	Level	Data Rates	Voice Circuits
Copper	DS0	64 kb/s	1
	DS1	1544 kb/s	24
	DS2	6312 kb/s	96
	DS3	44.736 Mb/s	672
Fiber	OC-1	51.84 Mb/s	672
	OC-3	155.52 Mb/s	2016
	OC-12	622.08 Mb/s	8064
	OC-24	1244.16 Mb/s	16128
	OC-48	2488.32 Mb/s	32256

3.1.1.2 Low signal attenuation. Little loss of signal during transmission allows longer distances between repeaters.

Copper—loss increases with frequency, which limits the practical distance between repeaters.

Fiber—optical fiber exhibits low inherent attenuation at the lightwave region of the spectrum. With low attenuation, an acceptable signal level can be maintained over longer distances without the use of repeaters to regenerate the signals at various locations along the cable. Thus higher transmission rates can be achieved.

3.1.1.3 High data integrity. Data integrity is a measure of performance based on the rate of undetected errors. A typical fiber-optic transmission system has a bit-error rate of better than 10–11, whereas the bit-error rate usually found in metallic cabling is about 10-9. Thus fiber-optic systems do not require the extensive use of error checking protocol. This offers two benefits:

1. Much of the error-checking overhead may be eliminated, resulting in even higher data transmission rates.
2. The number of retransmissions is reduced, improving overall system performance.

3.1.1.4 Immunity to interference. Electro-magnetic interference (EMI) and radio frequency interference (RFI) are principal causes of data errors in transmissions over metallic cable systems. Optical fibers are also immune to the effects of adverse weather conditions that plague microwave transmissions. Such immunity simplifies distribution, since fiber-optic cables need not be rerouted around elevators, machinery, auxiliary power generators, fluorescent lighting, and other potentially disruptive sources of EMI and RFI. This also makes fiber more economical to install, since there is no need to build special conduits to shield fiber from the external environment. And because optical fibers do not emit electro-magnetic radiation, which can cause crosstalk, multiple fibers may be bundled into a single cable.

3.1.1.5 High security. Optical fibers do not radiate signals, so they are a more secure transmission medium than unshielded metallic cable or microwave. Radiated signals can be collected through a variety of means, but since optical fibers do not use electro-magnetic or radio frequency energy, there is nothing to collect. Therefore, other means of tapping must be used, which entail physically breaking the core and fusing a connection to it. Such means are routinely used to add nodes to the fiber cable. But during the procedure, no light can be transmitted past that point, which would make unauthorized access easy to detect. Moreover, the increased signal loss from the tap could be detected.

3.1.1.6 Greater durability. A cross section of an optical fiber reveals four segments:

- The core, which transmits the light signal
- The cladding, which keeps the light signal within the core
- The coating, which protects the glass from abrasion
- The strengthening member, which provides the cable with strength.

Optical fiber has the same tensile strength as steel wire of the same diameter. Cables containing optical fibers are typically reinforced with both a strengthening member inside the cable and a protective jacket around the outside of the cable. This gives optical fiber the necessary durability to withstand being pulled through walls, floors, and underground conduits without being damaged. In addition, fiber can withstand greater temperatures than copper, making it better able to survive potentially disastrous fires. Fiber is also more resistant to corrosion than copper wire and, consequently, has a longer useful life.

3.1.1.7 Ease of Splicing. New advances in splicing have made accurate splicing of optical fiber a fairly routine matter. Today, there are few if any disadvantages associated with using optical fiber rather than copper wiring, except for the higher installation cost.

3.1.1.8 Other Advantages of Fiber. Additional advantages to using optical fiber include:

- Better operating environment. Absence of electrical current minimizes the risk of fire from a short circuit.
- No need for a common ground, permitting greater routing flexibility.
- Size and weight advantages over metallic cable, which simplifies shipping, handling, and storage.

3.1.2 Fiber-Optic Conduit Layout

Fiber cable has to be pulled to each niche in the Collider and High Energy Booster (HEB), and to each surface equipment building. The fiber is passed through PVC schedule 40 conduit and buried under the concrete of the tunnel floor. The concrete provides shielding from beam radiation.

At each equipment niche (every 450 m in the Collider), a pull box is placed to allow access to the fiber cable in order to route some of the fiber into the niche. From this pull box, fiber is routed to a fiber distribution rack at the back of the niche.

At each sector building, fiber is passed up the utility shaft, in order that SONET multiplexing equipment for the sector can be located in an area that can be accessed while the beam is on.

3.2 SONET

SONET (Synchronous Optical NETwork) is the set of standards for passing TDM (Time Division Multiplexing) data over a fiber-optic network at rates of 55 Mbits and above. It is presently defined at rates from 55 Mbits/s (OC1) up to 10 Gbits/s (OC192). Multiplexers are presently available from OC1 to OC48 (2.5 Gbits/s).

SONET has been chosen as the data transport standard at the SSC for the Control System. It is a widely accepted international standard. Many vendors make SONET devices and test equipment. The standards includes many built-in diagnostic and network management features. It can transport many types of data, from 64 Kbit/s digitized voice or data up to FDDI.

Many of these data types can be transported on the same fiber, even at different rates. SONET is upgradeable, and does not have an upper bandwidth limit. It will have a long lifetime, and is expected to form the fabric of the telephone system for at least the next 25 years.

SONET was originally proposed by Belcore for a family of interfaces for use in Operating Telephone Company optical network. With single-mode fiber becoming the medium of choice for high-speed digital transport, the lack of signal standards for optical networks inevitably led to a proliferation of proprietary interfaces. Thus, the fiber-optic transmission systems of one manufacturer cannot optically interconnect with those of any other manufacturer, and the ability to mix and match different equipment is restricted. SONET defines standard optical signals, a synchronous frame structure for the multiplexed digital traffic, and operations procedures.

3.2.1 SONET Signal Hierarchy

The basic building block and first level of the SONET signal hierarchy is called Synchronous Transport Signal–Level 1 (STS-1). The STS-1 has a bit rate of 51.84 Mb/s and is assumed to be synchronous with an appropriate network synchronization source. The STS-1 frame structure can be drawn as 90 columns and 9 rows of 8-bit byte. The order of transmission of the bytes is row by row, from left to right, with one entire frame being transmitted every 125 μs. The first three columns of the STS-1 contain section and line overhead bytes. The remaining 87 columns and 9 rows are used to carry the STS-1 Synchronous Payload Envelope (SPE). The SPE is used to carry SONET payloads, including 9 bytes of path overhead. The STS-1 can carry a clear channel DS3 signal (44.736 Mb/s) or, alternatively, a variety of lower-rate signal such as DS1, DS1C, and DS2.

The Optical Carrier–Level 1 (OC-1) is obtained from the STS-1 after scrambling (to avoid long strings of ones and zeros and to allow clock recovery at receivers) and electrical-to-optical conversion. The OC-1 is the lowest-level optical signal to be used at SONET equipment and network interface.

3.2.2 SONET Overhead Channel

The SONET overhead is divided into section, line, and path layers. The section layer contains those overhead channels that are processed by all SONET equipment, including regenerators. The section overhead channels for an STS-1 include two framing bytes that show the start of each STS-1 frame, an STS-1 identification byte, an 8-bit Bit-Interleaved Parity (BIP-8) check for section error monitoring, an orderwire channel for craft (network maintenance personnel) communication, a channel for unspecified network user (operator) ap-

plications, and three bytes for a section-level data communications channel to carry maintenance and provisioning information.

The line overhead is processed at all SONET equipment except the regenerator. It includes the STS-1 pointer bytes, an additional BIP-8 for line-error monitoring, a two-byte Automatic Protection Switching (APS) message channel, a 9-byte line-data communication channel, byte reserved for future growth, and a line orderwire channel.

The path overhead bytes are processed at SONET STS-1 payload terminating equipment; that is, the path overhead is part of the SONET STS-1 payload and travels with it. The path overhead includes a path BIP-8 for end-to-end payload error monitoring, a signal label byte to identify the type of payload being carried, a path status byte to carry maintenance signal, a multiframe alignment byte to show DS0 signaling bit phase, and others.

3.2.3 Multiplexing

Higher rate SONET signals are obtained by first byte-interleaving N frame-aligned STS-1s to form an STS-N. To transport payloads requiring less than an STS-1 payload capacity, the STS-1 SPE is divided into payload structures called virtual tributaries (VTs). There are four sizes of VTs: VT1.5, VT2, VT3, and VT6, where each VT has enough bandwidth to carry DS1, CEPT-1 (2.048 Mb/s), DS1C, and DS2 signal, respectively. Each VT occupies several 9-row columns within the SPE. The VT1.5 is carried in 3 columns (27 bytes), the VT2 in 4 columns (36 bytes), the VT3 in 6 columns (54 bytes), and VT6 in 12 columns (108 bytes).

A VT group is defined to be a 9-row-by-12-column payload structure that can carry four VT1.4s, three VT2s, two VT3s, or one VT6. Seven VT groups (84 columns), one path overhead column, and two unused columns are byte-interleaved to fully occupy the STS-1 SPE. VT groups carrying different VT types can be mixed within one STS-1.

Two different modes have been adopted for transporting payloads within a VT. The VT operating in the "floating" mode improves the transport and cross-connection of a VT payload. A floating VT is so called because a VT pointer is used to show the starting byte position of the VT SPE within the VT payload structure. In this sense, the operation of the VT pointer is directly analogous to that of the STS-1 pointer, and has the same advantage of minimizing payload buffers and associated delay when mapping signals into the VT. The other VT mode is the "locked" mode. The locked VT does not use the VT pointer, but instead locks the VT payload structure directly to the STS-1 SPE. The locked mode improves the transport and cross-connection of DS0 signals by maintaining the relative phase and frequency of DS0 signals carried in multiple locked VTs. When VT-organized, each STS-1 SPE carries either all floating or locked VTs.

More than one specific payload mapping is possible with each of the VT modes described above. Asynchronous mappings are used for clear channel transport of nominally asynchronous signals using the floating mode of operation; conventional positive bit-stuffing is used to multiplex these signals into the VT SPE. "Byte synchronous" mapping has been defined in both the locked and floating modes for the efficient, synchronous transport of DS0 signals and their associated signaling; conventional fixed-position mappings are used to carry the DS0s in the VT SPE (floating mode) or VT (locked mode). "Bit synchronous" mappings are used in both the locked and floating modes for the clear-channel transport of unframed, synchronous signals.

3.3 The SONET Backbone

We consider the SSC accelerator complex as consisting of 14 sectors (10 sectors for the Collider, one sector for the HEB, and one sector for each of the other machines—linear accelerator (LINAC); Low Energy Booster (LEB); and Medium Energy Booster (MEB). Each sector uses one pair of 155 Mb/s (SONET OC-3) single-mode fiber cables to provide full duplex communication for the machine control. The communication across the section boundaries is established by adding one Add-Drop multiplexer at both ends of the sections. There are several advantages to this architecture:

- Lower Cost. The SONET OC-3 Add-Drop-Multiplexer is available now at approximately $25 000 per unit, and there are many different commercial systems on the market. The higher-level systems (OC-12, OC-48) are several times more expensive than the OC-3 ADM, and there are very few commercial systems on the market now.
- Ease of design. The very high speed digital links (OC-12, OC-48) are primarily used by telephone companies to carry traffic across the nation. The links will demultiplex to lower speed at the local central office and distribute to the end user. The SSC accelerator niches are similar to the end users, and each section is similar to the local central office of the telephone company. We do have a lot of data exchange between the control room and niche, but do not have much communication between niches. It is quite simple to group a whole section as a unit and to communicate with the control room. If we use a two-level hierarchy, it requires multiplex and demultiplex at each sector.
- Minimum time delay. The multiplex and demultiplex process will introduce time delay. The maximum nodes in this scheme are limited to 17 (17 niches per sector); therefore the maximum time delay will be 16 times the propagation delay per ADM. If we use more than a one-layer hierarchical design, we have to add the propagation delay at every multiplex and demultiplex between high-speed and low-speed systems. It may be a nightmare for the synchronous application (such as the broadcast system).
- Feasible technology. In the control room, we can process 155 Mb/s (20 Mbyte/s) data using existing technology (VME, Futurebus) without too much difficulty. But processing Gigabit data without demultiplex is extremely difficult using existing technology.
- Increased reliability and maintainability. Each section has a maximum of 17 nodes, so it is much easier to isolate the fault than with a 200-node ring structure. The complete communication failure rate in this scheme should be ten times better than using one large ring structure. Putting Gigabit equipment in the hazardous environment would decrease MTBF and further cause reliability problems.
- Reduced operation cost. We stock just one kind of spare TDM equipment in this architecture, thereby reducing the operation cost.

3.4 T1

T1 is a standard that defines TDM data transport at a rate of 1.554 Mbits/s. It is the standard used for data transport of Wide Area Networks (WANs). It consists of a serial byte stream arranged on a 125-μs boundary. Within this 125-μs time slot are passed 24 bytes of data and one synchronization bit. The 24 bytes of data can be individual channels (each of 64 kbits/s) or a single channel. T1 is a point-to-point network. Standard SONET multiplexers have interfaces available to transport many T1 streams, and many have the ability to rearrange (groom) data streams down to the 64-kbit/s level.

T1 was chosen for the SSC controls communication for the following reasons. It is deterministic—that is, data is transported at a known data rate and response time, and the load on one channel does not affect the load on other channels. The appropriate number of 64-kbit channels can be allocated for any device. It does not have a minimum data length. Computer interfaces are available from industry. Other data streams such as 19.2 kbaud can be transported using standard commercial hardware. It is a mature technology that has a proven reliability record.

T1 is the digital, two-way transmission of voice, data, or video over a single high-speed circuit. T1 refers to the equipment used to transmit these signals. The transmission rate of that signal is based on the bandwidth of one voice signal in digital form. This one voice signal, called DS-0, consumes 64kbits/s. When T1 technology was first introduced, it was based on sending 24 voice frequency channels over two pairs of twisted cable. The T1 transmission rate was calculated by multiplying the 24 channels times 64kbits/s for a total of 1 536 000

bits/s. A necessary "control" bit was added to keep T1 multiplexers synchronized. This control bit takes up another 8000 bits/s, for a total of 1 544 000 bits/s. This is the transmission speed of T1.

3.5 Framing

In order for the receiving multiplexer to recognize bits as they come in, the bits have to be organized in some way. This is accomplished with framing, which separates the transmitted bits into identifiable groups by sending the 24 channels 8 bits at a time. Each 8-bit data word occupies a time slot in the frame. The frames repeat at a rate of 8000 times per second. So one time slot equals 64 Kb/s.

All 24 channels grouped together form a frame consisting of 192 bits (8×24), with the addition of a control bit. The 192 bits and the 193rd control bit add up to a combined speed of 1.544 Mb/s.

Modern T1 equipment uses one of two line coding schemes: super-frame (SF) format or extended super-frame (ESF) format. In the SF format, each super-frame consists of 12 frames. The 193rd bit alternately acts as a framing bit (known as "Ft" bit in odd-number frames) and a signaling bit (known as "Fs" bit in even-number frames) that provide multi-frame alignment information.

The SF format does have one limitation. It cannot measure the error rate on a T1 circuit while that circuit is in operation.

ESF processes the 193rd bit differently than the SF format does. Instead of using 12 frames, it extends the super-frame boundaries to include 24 frames. In addition, more efficient digital processing techniques allow frame synchronization to be performed in only 25% of the 8000-bit bandwidth required by SF for synchronization. ESF uses only one 2000-bit channel for framing (known as "FPS" bit). Another 2 Kb/s is used for error checking, and the remaining 4 Kb/s, called the Facility Data Link (FDL), can be used to control diagnostics and transmit performance characteristics.

In T1, data is transmitted in an alternate mark inversion (AMI) format, which allows clock signals to be derived from data, eliminating the need for separate clock transmission. The clock signal is expected from the AMI waveform using phase-locked loops or LC tank circuitry. Clock extraction circuitry requires a minimum density of "1's" to operate correctly. To meet this density requirement (networks cannot transmit a code consisting solely of zero), existing T-carrier equipment changes bit 7 of any channel consisting solely of "0's" to 1. Bit 7 stuffing does not affect the quality of voice transmission, but it does corrupt data significantly. An alternative to bit 7 stuffing is bipolar eight zero substitution (B8ZS), which replaces any transmitted zero octet with a B8ZS code word. If the last "1" transmitted was positive, the inserted word is 000+–0–+. If the last "1" transmitted was negative, the code word inserted is 000–+0+–. Bipolar violations occur in the fourth and seventh bit position; the receive side detects the code word and replaces it with all zeros.

3.6 Messages

Passing on these T1 links will be the information necessary to control the SSC accelerator complex. This information is broken down into a number of message types: Control Messages, Data Messages, Diagnostic Messages, and Broadcast Messages.

3.6.1 Control Messages

Control Messages are the commands that are sent to and between accelerator equipment modules, and the responses that return. The commands will be of a fixed format. Response would normally be at least an acknowledgement, but might include error status or information. For instance, a command to read an ADC channel one time might return the ADC value.

3.6.2 Data Messages

Data Messages are the normal way of passing channel data to supervisory computers. This would normally be arranged so that each channel would have a dedicated time slot on the TDM for its data. This would eliminate any chance of congestion. At the receiver of this channel, data can be stored in a FIFO or circular buffer. As the TDM system is deterministic, data collected at the same time within the front-end electronics can be made to be coherent at the supervisory computers.

3.6.3 Diagnostic Messages

Diagnostic Messages include unsolicited error reporting, down-loading, and debugging of programs. Some of these may well use the TCP/IP protocol, as this is supported by many debuggers.

3.6.4 Broadcast Messages

Broadcast Messages are messages that are intended for many destinations and/or are repetitive in nature. Each broadcast message is from 2 to 13 bytes long. Some messages signify the state of a machine; for instance, the next Linac pulse is intended for the medical facility. Others denote an event; for instance, a quench has occurred in the HEB. Others are for regular synchronization; for instance, 720-Hz tick message. Yet others carry data such as the present magnet field or current value.

Some of these message types will be able to share the same T1 links in certain circumstances. Most Data Messages will be from the equipment to the supervisory computers, and Broadcast Messages will be in the other direction. Also commands might be transmitted in the idle time between Broadcast Messages. (See Broadcast Message system documentation).

3.7 Control Room Communications

Communication between UNIX computers in the control room will be Asynchronous Transfer Mode (ATM) networks, running TCP/IP protocols. ATM is a fast packet switching technology using short (53-byte) packets. Interface to the TDM network for access and control of real-time systems will be via commercial routers.

Control functions exist primarily at the main control room area, at the sector service buildings, and at alcoves strategically located adjacent to the tunnel.

The central computer system resides near the main control room and contains the complete database for the entire accelerator complex. It is connected to control consoles, terminals, personal computing equipment, workstations, and the wide-area ring network that communicates with the sector and injection complex control computers; it is also connected to a link dedicated to the transmission of high-bandwidth beam diagnostic (primarily BPM) information. This information will be needed at the control room area but will follow a separate path from that of the lower rate data.

At each of the ten sector/cluster service buildings the control functions are complex, involving the operation of a large refrigerator and two main bus power supplies in addition to communication with all of the tunnel devices of that sector.

Each niche contains a number of relay racks containing control system and accelerator subsystem equipment. The controls equipment at these locations consists of the more sophisticated subsystems such as quench protection, beam position, and vacuum. Also present is the low-level hardware for the less involved subsystems—power supply control, A/D, timing.

There are three types of data and control flows:

- point-to-point between the main control room and each of the sector service buildings
- ring-wide, with connections at the sector service buildings
- sector-wide, for the information that will be transmitted to the service building from each niche.

4.0 HIGHER LEVEL PROTOCOLS

Many software systems, including the EPICS software used in the control system and commercial debuggers, use the TCP/IP protocol to communicate between equipment nodes. This leads to a requirement for the Accelerator Control Communications system to transport these protocols over the TDM channels that comprise the communications infrastructure.

The Point to Point Protocol (PPP), a part of the Internet suite of protocols, defines a system for establishing communications over a serial point-to-point link. This protocol is widely supported by commercial routers. It enables the transport of IP packets over serial links.

Although broadcast is not supported over PPP links, we use the capabilities of the SONET add-drop multiplexers to broadcast, using drop and continue technology, on a separate serial TDM channel that can be transmitted to many crates.

For more information see documentation on the Point to Point Protocol driver that has been written.