Cray Scalable I/O Overview [2]

The new GigaRing scalable input/output (I/O) and networking channel from Cray Research supports the next generation of I/O peripherals for the CRAY T90 series, CRAY T3E series, and CRAY J90se series of systems.

The GigaRing I/O and networking channel provides improvements in the following areas:

- Performance
- Reliability
- Maintainability
- Availability
- Scalability

2.1 Introduction to Cray scalable I/O

The Cray scalable I/O (SIO) architecture consists of a number of I/O nodes (IONs) connected by a new high-speed system channel called the *GigaRing channel*. GigaRing technology is implemented as a ring-based channel that connects multiple clients together with high-speed, point-to-point links. GigaRing clients consist of system nodes (Cray Research mainframes) and IONs that support I/O peripherals.

Figure 1. GigaRing channel topology

The GigaRing channel topology is modified from the Standard Coherent Interface (SCI) to incorporate a pair of counter-rotating rings. Each node on the ring receives and transmits data on two 32-bit rings referred to as positive and negative rings. Each node provides a host interface (system node or ION) to the rings through a 32-bit or 64-bit full-duplex client interface.

The GigaRing channel is actually a pair of counter-rotating rings. This ring topology, along with sharing of a common channel by nodes on the ring, enhances interoperability.

Each system node and ION on a GigaRing channel has a unique identifier called the *physical node address*. The physical node address is a 13-bit unique physical node ID that consists of seven ring identifier bits and six node identifier bits.

The basic unit of transfer on a GigaRing channel is called a *packet*. A GigaRing packet is limited to 32 64-bit words of payload.

2.1.1 I/O nodes

Cray Research has developed a variety of I/O nodes (IONs) to support a wide range of connectivity requirements. The two main types of IONs are single-purpose nodes and the multipurpose node. Single-purpose nodes (SPNs) support specific channel interfaces and/or devices (for more information on SPNs, see Section 2.1.1.1, page 12). The multipurpose node (MPN-1) provides an interface based on the SBus standard to support industry-standard I/O channels (for more information on MPNs, see Section 2.1.1.2, page 14).

IONs based on GigaRing technology provide access to mass storage devices such as disks and tapes, as well as to industry-standard computer networks such as High Performance Parallel Interface (HIPPI) networks. For more information on I/O products and equipment, see Table 1 and Table 2.

The Cray scalable I/O (SIO) design provides resiliency and the ability to perform *hot swaps*, in which power supplies, IONs, I/O cables, and other components that are designated field replaceable units (FRUs) can be replaced without powering down or interrupting other equipment in the PC-10 cabinet (peripheral cabinet) or on the GigaRing channel. The capability to perform a hot swap depends heavily on the system configuration; for example, the configuration of alternate paths to disks is required if IONs are to be swapped.

The CRAY SSD-T90 solid-state storage device also connects directly to the GigaRing channel, providing dynamic random access memory (DRAM) secondary storage. The CRAY SSD-T90 storage device is used primarily with CRAY T90 systems.

Table 1. Cray scalable I/O product names and descriptions

Model number	Product name	Channel capacity	Description
IPN-1	IPI-2 I/O Node	Five IPI	Can be configured to five independent IPI disk channels or as a four data plus one parity RAID-3 disk array.
BMN-1	Block Mux Tape I/O Node	Two block mux tape	Connects tape drive and tape storage subsystems.
HPN-1	HIPPI Channel I/O Node	100 Mbyte/s (32-bit)	Each node contains two 100 Mbyte/s HIPPI channels.
HPN-2	HIPPI Channel I/O Node	200 Mbyte/s (64-bit)	Can be configured to one 200 Mbyte/s HIPPI channel or one 100 Mbyte/s HPPI channel.
FCN-1	Fibre Channel I/O Node	Five fibre channel arbitrated loops (FCALs)	Supports 100 Mbyte/s burst rate per loop; each loop can operate several disks concurrently; supports RAID–3 and RAID-5.
ESN-1	ESCON I/O Node	Four independent channels per node	Supports four ESCON channels with bandwidth of 17 Mbyte/s per channel.
MPN-1	Multi-Purpose I/O Node	Up to eight SBus channels	Contains two SBuses, each supporting up to four SBus cards. SBus-based channel adapters support Ethernet network connections, FDDI network connections (requires 2 SBus channels), ATM network connections, SCSI disks, and Supervisory channel SBus (SC01).

Table 2. Cray scalable I/O support equipment

Model number	Product name	Description
ION support equipment		
PC-10	I/O peripheral cabinet	Single air-cooled I/O cabinet with DE-100 type skins for housing ION assemblies. (There are two types of PC-10 cabinets, designated as PC-10A and PC-10B. The only functional difference between these two cabinets is that the PC-10B cabinet supports WACS while the PC-10A cabinet does not.
NSR-1	ION subrack	Contains space for one to four channel adapters (does not support MPN-1).
DSF-1	Fibre disk subrack	Contains n+1 power and space for up to 10 Fibre Channel interface disk drives.
DSS-1	SCSI disk subrack	Contains n+1 power and space for up to 8 SCSI-2 interface disk drives.
DD-308	Fibre Channel interface	8 Gbyte Fibre Channel interface 3.5 inch drive.
GigaRing channel support equipment		
FOX-1	Fibre Optic GigaRing channel extender	Optical extension to the standard Cray GigaRing channel.

2.1.1.1 Single-purpose nodes

Single-purpose nodes (SPNs) connect I/O peripherals to the GigaRing channel to provide various I/O services to system nodes (CRAY T90 series, CRAY T3E series, and CRAY J90se series of systems) on the ring. The GigaRing architecture enables full connectivity between nodes on a ring.

Each SPN contains a SPARC processor, which runs the VxWorks real-time operating system with Cray Research node-specific I/O software. This provides I/O capabilities to a system on a GigaRing channel.

SPNs are housed in a stand-alone, air-cooled peripheral cabinet (the PC-10) and support the connection of various I/O peripherals such as intelligent peripheral interface (IPI) disk arrays, fibre channel disk arrays, block mux tape, and so on, to the system.

For a list of peripherals supported by SPNs, see Table 3.

Certain I/O-related peripherals do not require a unique I/O node (ION). For example, tape robots (autoloaders) are controlled through Transmission Control Protocol/ Internet Protocol (TCP/IP) network connections to the GigaRing channel. Messages are sent to the loaders through whatever node is carrying network traffic (Fiber Distributed Data Interface (FDDI), HIPPI, Ethernet, and so on).

Table 3. Peripherals supported by single-purpose nodes

ION	Devices	Description
IPN-1	DD-60	IPI-2 disk drive
IPN-1	DA-60	IPI-2 disk array (RAID-3)
IPN-1	DD-62	IPI-2 disk drive
IPN-1	DA-62	IPI-2 disk array (RAID-3)
IPN-1	DD-301	IPI-2 disk drive
IPN-1	DA-301	IPI-2 disk array (RAID-3)
IPN-1	DD-302	IPI-2 disk drive
IPN-1	DA-302	IPI-2 disk array (RAID-3)
FCN-1	DD-308	Fibre Channel disk drive
FCN-1	DD-308	Fibre Channel disk array (4 + 1 RAID-3)
BMN-1	IBM 3420	9-track tape (256-Kbyte block limit) or (IBM 3420 compatible 9-track reel tapes, such as STK 4670)
BMN-1	IBM 3480	18-track tape
BMN-1	IBM 3490	36-track tape
BMN-1	STK 4480	18-track tape
BMN-1	STK 4490	36-track tape
BMN-1	STK 4400	Libraries and robots
BMN-1	STK 9310	Libraries and robots
BMN-1	STK 9360	Libraries and robots
ESN-1	IBM 3490E	36-track tape (extended capacity)

ESN-1	IBM 3590	Magstar
ESN-1	STK 4490	36-track tape
ESN-1	IBM 3490E	36-track tape (extended capacity
ESN-1	STK 9490	Timberline
ESN-1	STK SD-3	RedWood
ESN-1	IBM 3494	Libraries and robots
ESN-1	STK 4400	Libraries and robots
ESN-1	STK 9310	Libraries and robots
ESN-1	STK 9360	Libraries and robots
HPN-1 and HPN-2		HIPPI network connections
HPN-1 and HPN-2	ND-12	Network disk
HPN-1 and HPN-2	ND-14	Network disk
HPN-1 and HPN-2	ND-40	Network disk with semaphore

For more information on SPNs, see *Cray Scalable I/O Functional Overview*, Cray Research publication SD–2208.

2.1.1.2 Multipurpose node

The peripheral cabinet also houses a special I/O node (ION) called the Multipurpose I/O Node, MPN-1. The MPN-1 provides SBus-based I/O connections for industry standard peripherals such as small computer system interface (SCSI) and Asynchronous Transfer Mode (ATM) to the GigaRing channel interface. An MPN-1 supports up to eight SBus peripheral connections. It also enables communication between the system workstation (SWS) and the GigaRing channel.

Each MPN-1 contains a SPARC processor, which runs the VxWorks real-time operating system with Cray Research node-specific I/O software. This provides industry standard I/O capabilities to a system on a GigaRing channel. MPNs are also housed in the PC-10 cabinet.

Table 4 lists peripherals supported by the Multipurpose I/O Node, MPN-1, for Cray Research mainframes.

Table 4. Peripherals supported by MPN

ION	SBus controller	Devices	Description
MPN-1	SCS-10	DD-314	SCSI-2 disk drive
MPN-1	SCS-10	DD-318	SCSI-2 disk drive
MPN-1	SCS-10	STK 4781	SCSI 18-track tape
MPN-1	SCS-10	STK 4791	SCSI 36-track tape
MPN-1	SCS-10	STK 4890	SCSI tape, Twin Peaks
MPN-1	SCS-10	STK 9490	SCSI tape, TimberLine
MPN-1	SCS-10	STK SD-3	SCSI tape, RedWood
MPN-1	SCS-10	IBM 3490E	SCSI 36-track tape
MPN-1	SCS-10	IBM 3590	SCSI tape, Magstar NTP
MPN-1	SCS-10	EXABYTE 8505	SCSI tape, 8mm
MPN-1	SCS-10	HP C1533A	SCSI tape, 4mm
MPN-1	SCS-10	Quantum DLT 4000	SCSI digital linear tape
MPN-1	SCS-10	Quantum DLT 7000	SCSI digital linear tape
MPN-1	SCS-10	STK 4400	SCSI libraries and robots
MPN-1	SCS-10	STK 9310	SCSI libraries and robots
MPN-1	SCS-10	STK 9360	SCSI libraries and robots
MPN-1	SCS-10	STK 9710	SCSI libraries and robots
MPN-1	SCS-10	IBM 3494	SCSI libraries and robots
MPN-1	ETN-10		Ethernet
MPN-1	FDI-10		FDDI networks
MPN-1	ATM-10		ATM OC3

For more information on multipurpose nodes (MPNs), see *Cray Scalable I/O Functional Overview*, Cray Research publication SD–2208.

2.1.1.3 Solid-state storage device

The next generation of Cray Research solid-state storage device, the CRAY SSD-T90 device, provides secondary storage to systems on a GigaRing channel. The CRAY SSD-T90 device uses the GigaRing channel as its only interface to other devices.

2.1.2 System workstation for GigaRing environments

The Cray system workstation (SWS) is a SPARC 5 workstation that serves in these capacities:

- As the operator workstation for CRAY T90 series, CRAY T3E series, and CRAY J90se series of systems.
- As a maintenance platform, providing support for diagnostics, ring management and maintenance, node monitoring, and hardware failure detection and isolation.
- As the system workstation, running management and maintenance software that provides support for both IONs and system nodes on a GigaRing channel.

The SWS operational model describes how a single SWS is designed to operate multiple system and IONs, which can run various levels of software.

The IONs are connected to the SWS by an Internet Protocol (IP) network connection.

For more information on the SWS and SWS operational model, see *SWS-ION Administration and Operations Guide*, Cray Research publication SG–2204.