Cray Research - Network Update

Gary Klesk, Cray Research, Inc., 655-F Lone Oak Drive
Eagan, Minnesota 55121

ABSTRACT: *This paper outlines the various industry standard network interfaces used with Cray Systems. The interfaces to be discussed are Ethernet, FDDI, HIPPI and ATM. Topics covered are interface characteristics, upgrade paths, interface issues, performance and availability.*

Introduction

Supporting standards based network interfaces are the focus for Cray Research. By doing so, Cray customers have greater flexability in expanding or upgrading their networks. Cray Research offers network support for Ethernet, FDDI, HIPPI and ATM interfaces.

Ethernet

Ethernet was introduced on Cray EL systems and is available on J90 VME systems. Ethernet is a low-end network technology which is rated at 10Mb/sec with a 1500 byte MTU. It is a shared media and uses CSMA/CD to handle collisions. The Fast Ethernet Alliance created 10BaseT's successor, 100BaseT fast ethernet with switched capabilities. 100BaseT is an attractive networking upgrade possibly requiring no infrastructure changes(100BaseT4), it is compatible with 10BaseT for approximately twice the cost.

With the interest that 100BaseT ethernet has generated, the Gigabit Alliance was formed to carry ethernet even faster, to 1000Mb/sec. The Alliance is submitting all technical proposals to the IEEE 802.3 standard commitee for approval towards a formal interoperable standard

The Gigabit Ethernet architecture proposal is backwards compatible and retains CMSA/CD mode for shared media and a full duplex mode for inter-switch connections. One major difference is the physical layer chosen. The Fiber Channel PHY standard has been selected which supports fiber optics in both multimode and single mode as well as copper.

Fast Ethernet network interface cards are available today as a dual purpose 10/100BaseT card with autosensing capabilities to switch modes based on it's connection type.

Dependent on the finalization of the standard. Gigabit Ethernet will most likely become available in the 1997 timeframe.

FDDI - Fiber Distributed Data Interface

FDDI connections to Cray systems were initially available by using low-speed Cray channels connected to FDDI capable routers made by vendors NSC and CNT. Cray introduced it's integrated FDDI adapter on the model-E IOS to reduce customer network costs and improve network throughput. FDDI is rated at 100Mb/sec, with a 4500 byte MTU, it's token based and uses multimode fiber as well as copper media.

It also has redundent dual rings and the ability to dynamically remove or insert a node on the ring. It is significantly more expensive than Ethernet and is a mature technology which has been primarily used for LAN backbone environments. Although there is a copper version(CDDI) available which has decreased the cost of an interface card. With the introduction of

Interface	Rated Bandwidth	Sustained Bandwidth TCP/IP	J90-VME	Model-E	GigaRing
Ethernet	10 Mb/sec	8 Mb/sec	Vendor	N/A	MPN
Fast Ethernet	100 Mb/sec	TBD	1Q97	N/A	MPN - 12/96
Gigabit Ethernet	1000 Mb/sec	TBD	TBD - 1997	N/A	TBD - 1997

ATM and the recent technology upgrades to Ethernet, FDDI does not offer anything to warrant upgrading to it.

HIPPI - High Performance Parallel Interface

Cray Research was an active participant in the development and standardization of this interface. It is rated at 800Mb/sec single channel or 1600Mb/sec in dual channel mode with no MTU limit. It is used in a direct connect or switched environment and it's limited to a 25 meter cable length unless fiber optic extenders are used which will support distances of 2km.

Originally designed for networking. HIPPI is now used for disk and tape I/O and supports copper or fiber media. HIPPI is relatively expensive and primarily used for bulk data transfer.

HIPPI technology is moving forward with the ANSI X3T11 HIPPI working group developing the specifications for HIPPI-6400 which is rated at 6400Mb/sec. It is expected run over parallel copper or parallel fiber-optic cables and there is no HIPPI-PH layer compatability with HIPPI-800 and HIPPI-1600.

HIPPI-6400 is targeted for availability in the 1998 timeframe.

ATM - Asynchronous Transfer Mode

Cray supports ATM at the OC-3 rate(155Mb/sec) with a default MTU of 9180, in PVC mode with Classical IP over ATM. It is available on the J90 as an integrated VME NIC card. For Model-E based systems, Cray offers the BBG(Bus Based Gateway). It is an SBus based platform which connects via a HIPPI channel to the Cray and outputs to one ATM interface. Both FORE Systems and Interphase Corporation ATM NIC's have been supported in the past. Presently, only Interphase Corporation NIC cards continue to be supported.

ATM RFC support is as follows;

RFC 1483 - Multiprotocol Encapsulation using Logical Link Control (LLC's) or Virtual Circuit(VC) based. Cray supports the TCP/LLC portion of this RFC.

RFC 1577 - Support for 802.2 headers and ARP.

RFC 1626 - Provides for the default MTU value of 9k.

RFC 1755 - Specifies requirements for signaling support for IP over ATM.

Upgrade Paths and Issues

Traditionally a network technology fills an anticipated bandwidth requirement. That technology matures as it becomes accepted and deployed, followed by work on the next generation. This has clearly not been the case with FDDI. Both Ethernet and HIPPI are extending there bandwidth capabilities

Interface	Rated Bandwidth	Sustained Bandwidth TCP/IP	J90-VME	Model-E	GigaRing
FDDI	100 Mb/sec	64 Mb/sec	Vendor		
FDDI	100 Mb/sec	70 Mb/sec		FCA-1	
FDDI	100Mb/sec	TBD			MPN

Interface	Rated Bandwidth	Sustained Bandwidth TCP/IP	J90-Memory	Model-E	GigaRing
HIPPI-800	800 Mb/sec	520 Mb/sec	HI-P		
HIPPI-800	800 Mb/sec	655 Mb/sec		HCA-3/4	
HIPPI-800	800Mb/sec	TBD			HPN-1
HIPPI-1600	1600Mb/sec	TBD			HPN-2
HIPPI-6400	6400Mb/sec	TBD			1998

Interface	Rated Bandwidth	Sustained Bandwidth	J90-VME	Model-E	GigaRing
ATM-OC3	155 Mb/sec	100 Mb/sec		BBG	
ATM-OC3	155 Mb/sec	110 Mb/sec	Vendor		
ATM-OC3	155 Mb/sec	TBD			MPN
ATM-OC12	622 Mb/sec	TBD	N/A	BBG - 1Q97	MPN - 1Q97

that have or, will overlap the capabilities of other predecessor technologies.

ATM's deployment costs, standards issues and lack of applications have slowed ATM's growth and acceptance in the LAN environment. This has led to huge opportunities for the Ethernet and HIPPI communities.

ATM has not grown as rapidly as had been assumed by proponents of ATM. The speed at which the standards have been ratified and, lack of applications specific to ATM capabilities have greatly added to the slow adoption of ATM. In most cases, ATM is being used as a backbone and is growing in the WAN market.

Support for UNI 4.0 will be implemented as it relates to Cray's Internet suite and to the features that will be supported by the hardware vendors. Q.2931 signaling/SVC support is scheduled for 1Q97 availability.

OC12 development plans are to integrate the anticipated Sbus product from SUN Microsystems into Cray's product offering for the GigaRing Channel connection and in the BBG. Availability of OC12 is anticipated to be in early 1997.

ATM beyond OC-12 speeds continues as a research project for specialized requirements and not as a general purpose interface.

At the desktop, 25Mb/s ATM was supposed to displace 10BaseT Ethernet. But with a four-fold increase in bandwith at about twice the cost of 10BaseT. 100BaseT has clearly slowed down 25Mb/s ATM's acceptance and will most likely dominate the desktop market for the next few years.

The one area that 100BaseT needs work is for Quality of Service provisions. ATM currently provides for it, the Ethernet Alliance(s) are working towards providing RSVP (ReSerVation Protocol) which supports Quality of Service classes for IP applications. 1000BaseT Gigabit Ethernet promises to carry Ethernets dominance to the desktop even further. Possibly out to five years.

At the higher LAN speeds, ATM currently uses OC12(622Mb/s) at the switch level. OC12 at the network interface level is not available. The latest estimates are for availability within the next three months. Although OC12 will provide a 4x increase over OC3. The ANSI X3T11 working group has addressed the need for even higher speeds in the large/bulk data transfer area by promoting the HIPPI-6400 specification. The HNF is addressing such issues as flow control, Quality of Service and latency for small packets.

NOTES:

MPN(Multi Purpose Node) SBus based platform which has eight SBus slots available for networking and I/O cards.

BBG(Bus Based Gateway) SBus based platform which has one HIPPI channel and one ATM OC-3 interface.

HPN-1(Single Purpose Node) 800 Mb/sec HIPPI channel.

HPN-2(Single Purpose Node) 1600 Mb/sec HIPPI channel. Can also be configured as two 800 Mb/sec channels.