

Network Working Group
Request for Comments: 2067
Category: Standards Track
Obsoletes: 1374

J. Renwick
NetStar, Inc.
January 1997

IP over HIPPI

Status of This Memo

This document specifies an Internet standards track protocol for the Internet community, and requests discussion and suggestions for improvements. Please refer to the current edition of the "Internet Official Protocol Standards" (STD 1) for the standardization state and status of this protocol. Distribution of this memo is unlimited.

Abstract

ANSI Standard X3.218-1993 (HIPPI-LE[3]) defines the encapsulation of IEEE 802.2 LLC PDUs and, by implication, IP on HIPPI. ANSI X3.222-1993 (HIPPI-SC[4]) describes the operation of HIPPI physical switches. The ANSI committee responsible for these standards chose to leave HIPPI networking issues largely outside the scope of their standards; this document describes the use of HIPPI switches as IP local area networks.

This memo is a revision of RFC 1374, "IP and ARP on HIPPI", and is intended to replace it in the Standards Track. RFC 1374 has been a Proposed Standard since November, 1992, with at least 10 implementations of IP encapsulation and HIPPI switch discipline. No major changes to it are required. However, the ARP part of RFC 1374 has not had sufficient implementation experience to be advanced to Draft Standard. The present document contains all of RFC 1374 except for the description ARP, which has been moved into a separate document.

TABLE OF CONTENTS

1	Introduction.....	2
2	Scope.....	3
2.1	Changes from RFC 1374.....	3
2.2	Terminology.....	4
3	Definitions.....	4
4	Equipment.....	5
5	Protocol	7
5.1	Packet Format.....	7
5.2	48 bit Universal LAN MAC addresses.....	11
5.3	I-Field Format.....	12

5.4	Rules For Connections.....	13
5.5	MTU.....	15
6	Camp-on	16
7	Path MTU Discovery.....	17
8	Channel Data Rate Discovery.....	17
9	Performance.....	18
10	Sharing the Switch.....	20
11	References.....	21
12	Security Considerations.....	21
13	Author's Address.....	21
14	Appendix A -- HIPPI Basics.....	22
15	Appendix B -- How to Build a Practical HIPPI LAN.....	27

1 Introduction

The ANSI High-Performance Parallel Interface (HIPPI) is a simplex data channel. Configured in pairs, HIPPI can send and receive data simultaneously at nearly 800 megabits per second. (HIPPI has an equally applicable 1600 megabit/second option.) Between 1987 and 1991, the ANSI X3T9.3 HIPPI working group drafted four documents that bear on the use of HIPPI as a network interface. They cover the physical and electrical specification (HIPPI-PH [1]), the framing of a stream of bytes (HIPPI-FP [2]), encapsulation of IEEE 802.2 LLC (HIPPI-LE [3]), and the behavior of a standard physical layer switch (HIPPI-SC [4]). HIPPI-LE also implies the encapsulation of Internet Protocol[5]. The reader should be familiar with the ANSI HIPPI documents, copies of which are archived at the site "ftp.network.com" in the directory "hippi", and may be obtained via anonymous FTP.

HIPPI switches can be used to connect a variety of computers and peripheral equipment for many purposes, but the working group stopped short of describing their use as Local Area Networks. This memo takes up where the working group left off, using the guiding principle that except for length and hardware header, Internet datagrams sent on HIPPI should be identical to the same datagrams sent on a conventional network, and that any datagram sent on a conventional 802 network[6] should be valid on HIPPI.

2 Scope

This memo describes the HIPPI interface between a host and a crosspoint switch that complies with the HIPPI-SC draft standard. Issues that have no impact on host implementations are outside the scope of this memo. Host implementations that comply with this memo are believed to be interoperable on a network composed of a single HIPPI-SC switch. They are also interoperable on a simple point-to-point, two-way HIPPI connection with no switch between them. They may be interoperable on more complex networks as well, depending on the internals of the switches and how they are interconnected; however, these details are implementation dependent and outside the scope of this memo.

Within the scope of this memo are:

1. Packet format and header contents, including HIPPI-FP, HIPPI-LE, IEEE 802.2 LLC[7] and SNAP.
2. I-Field contents
3. Rules for the use of connections.

Outside of the scope are

1. Address Resolution (ARP)
2. Network configuration and management
3. Host internal optimizations
4. The interface between a host and an outboard protocol processor.

2.1 Changes from RFC 1374

RFC 1374 described the use of ARP on HIPPI, but because of insufficient implementation experience, the description of ARP has been separated from IP encapsulation and moved to an Informational memo. It may be returned to the standards track in the future if interest and implementations warrant it.

RFC 1374's specification of IP over HIPPI has been changed in this document. Certain packet format options, permitted in RFC 1374, are no longer allowed:

1. Optional short burst first;
2. D1 fill bytes;
3. Nonzero D2 offset.

That is, the header format is no longer variable and is required to be that which is recommended by RFC 1374.

With these changes, it is possible to send packets which conform to the ANSI standards but not to this memo. Because there are no RFC 1374 implementations in use that used these options, we believe that all existing RFC 1374 implementations are compliant with the requirements of this memo, and there should be no interoperability problems associated with these changes.

2.2 Terminology

In this document the use of the word SHALL in capital letters indicates mandatory points of compliance.

3 Definitions

Conventional

Used with respect to networks, this refers to Ethernet, FDDI and 802 LAN types, as distinct from HIPPI-SC LANs.

Destination

The HIPPI implementation that receives data from a HIPPI Source.

Node

An entity consisting of one HIPPI Source/Destination pair that is connected by parallel or serial HIPPI to a HIPPI-SC switch and that transmits and receives IP datagrams. A node may be an Internet host, bridge, router or gateway. This memo uses the term node in place of the usual "host" to indicate that a host might be connected to the HIPPI LAN not directly, but through an external adaptor that does some of the protocol processing for the host.

Serial HIPPI

An implementation of HIPPI in serial fashion on coaxial cable or optical fiber, informally standardized by implementor's agreement in the Spring of 1991.

Switch Address

A value used as the address of a node on a HIPPI-SC network. It is transmitted in the I-field. HIPPI-SC switches may map Switch Addresses to physical port numbers.

Source

The HIPPI implementation that generates data to send to a HIPPI Destination.

Universal LAN Address (ULA)

A 48 bit globally unique address, administered by the IEEE, assigned to each node on an Ethernet, FDDI, 802 network or HIPPI-SC LAN.

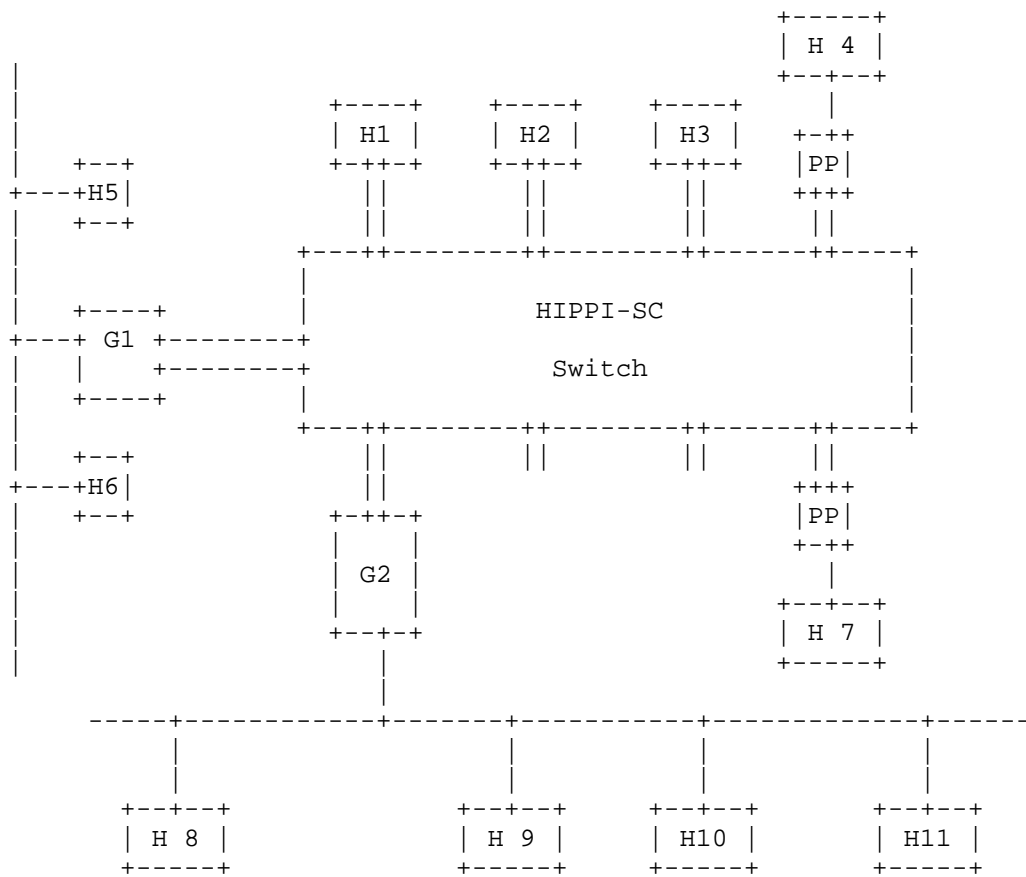
4 Equipment

A HIPPI network can be composed of nodes with HIPPI interfaces, HIPPI cables or serial links, HIPPI-SC switches, gateways to other networks.

Each HIPPI interconnection between a node and a switch SHALL consist of a pair of HIPPI links, one in each direction.

If a link between a node and the switch is capable of the 1600 Megabit/second data rate option (i.e. Cable B installed for 64 bit wide operation) in either direction, the node's HIPPI-PH implementation SHALL also be capable of 32 bit operation (Cable B data suppressed) and SHALL be able to select or deselect the 1600Mb/s data rate option at the establishment of each new connection.

The following figure shows a sample HIPPI switch configuration.



A possible HIPPI configuration

A single HIPPI-SC switch has a "non-blocking" characteristic, which means there is always a path available from any Source to any Destination. If the network consists of more than one switch, the path from a Source to a Destination may include a HIPPI link between switches. If this link is used by more than one Source/Destination pair, a "blocking" network is created: one Source may be blocked from access to a Destination because another Source is using the link it shares. Strategies for establishing connections may be more complicated on blocking networks than on non-blocking ones.

This memo does not take blocking issues into account, assuming that the HIPPI LAN consists of one HIPPI-SC switch or, if the network is more complex than that, it presents no additional problems that a node must be aware of.

5 Protocol

5.1 Packet Format

The HIPPI packet format for Internet datagrams SHALL conform to the HIPPI-FP and HIPPI-LE draft standards, with further restrictions as imposed by this memo. Because this memo is more restrictive than the ANSI standards, it is possible to send encapsulated IP datagrams that conform to the ANSI standards, but are illegal according to this memo. Destinations may either accept or ignore such datagrams.

To summarize the additional restrictions on ANSI standards found here:

- Any short burst must be the last burst of the packet.
Leading short bursts are not permitted.

- Nonzero values for the HIPPI-FP D2_Offset field are not permitted.

- The D1_AreaSize SHALL be 3 (64-bit words). No D1 Fill is permitted.

Note: Although this document is for IP over HIPPI, the encapsulation described below accommodates ARP as well.

The HIPPI-FP D1_Area SHALL contain the HIPPI-LE header. The HIPPI-FP D2_Area, when present, SHALL contain one IEEE 802.2 Type 1 LLC Unnumbered Information (UI) PDU. Support of IEEE 802.2 XID, TEST and Type 2 PDUs is not required on HIPPI, and Destinations that receive these PDUs may either ignore them or respond correctly according to IEEE 802.2 requirements.

The length of a HIPPI packet, including trailing fill, SHALL be a multiple of eight bytes as required by HIPPI-LE.

+-----+	+-----+	+-----+	+-----+	+-----+
HIPPI-FP	HIPPI-LE	IEEE 802.2 LLC/SNAP	IP . . .	0 - 7
(8 bytes)	(24 bytes)	(8 bytes)		bytes
+-----+	+-----+	+-----+	+-----+	+-----+
				fill
+-----+	+-----+	+-----+	+-----+	+-----+

HIPPI Packet Structure

ULP-id (8 bits) SHALL contain 4.

D1_Data_Set_Present (1 bit) SHALL be set.

Start_D2_on_Burst_Boundary (1 bit) SHALL be zero.

Reserved (11 bits) SHALL contain zero.

D1_Area_Size (8 bits) SHALL be sent as 3.

D2_Offset (3 bits) SHALL be zero.

D2_Size (32 bits) Shall contain the number of bytes in the IEEE 802.2 LLC Type 1 PDU, or zero if no PDU is present. It SHALL NOT exceed 65,288. This value includes the IEEE 802.2 LLC/SNAP header and the IP datagram. It does not include trailing fill bytes. (See "MTU", below.)

HIPPI-LE Header

FC (3 bits) SHALL contain zero unless otherwise defined by local administration.

Double_Wide (1 bit) SHALL contain one if the Destination associated with the sending Source supports 64 bit HIPPI operation. Otherwise it SHALL contain zero.

Message_Type (4 bits) contains a code identifying the type of HIPPI-LE PDU. Defined values are:

- 0 Data PDU
- 1 Address Resolution Request PDU (AR_Request)
- 2 Address Resolution Response PDU (AR_Response)
- 3 Self Address Resolution Request PDU (AR_S_Request)
- 4 Self Address Resolution Response PDU (AR_S_Response)

Destination_Switch_Address is a 24-bit field containing the Switch Address of the Destination if known, otherwise zero. If the address comprises less than 24 bits, it SHALL be right justified (occupying the least significant bits) in the field.

Destination_Address_Type (4 bits) and Source_Address_Type (4 bits) contain codes identifying the type of addresses in the Destination_Switch_Address and Source_Switch_Address fields respectively. Defined values (binary) are:

- 0 Unspecified
- 1 HIPPI-SC Source Route (24 bits)
- 2 HIPPI-SC Address (12 bits)

Source_Switch_Address is a 24-bit field containing the Switch Address of the Source. If the address comprises less than 24 bits, it SHALL be right justified (occupying the least significant bits) in the field.

Reserved (16 bits) SHALL contain zero.

Destination_IEEE_Address (48 bits) SHALL contain the 48 bit Universal LAN MAC Address of the Destination if known, otherwise zero.

LE_Locally_Administered (16 bits) SHALL contain zero UNLESS otherwise defined by local administration.

Source_IEEE_Address (48 bits) SHALL contain the 48 bit Universal LAN MAC Address of the Source if known, otherwise zero.

IEEE 802.2 LLC

The IEEE 802.2 LLC Header SHALL begin in the first byte of the HIPPI-FP D2_Area.

SSAP (8 bits) SHALL contain 170 ('AA'h).

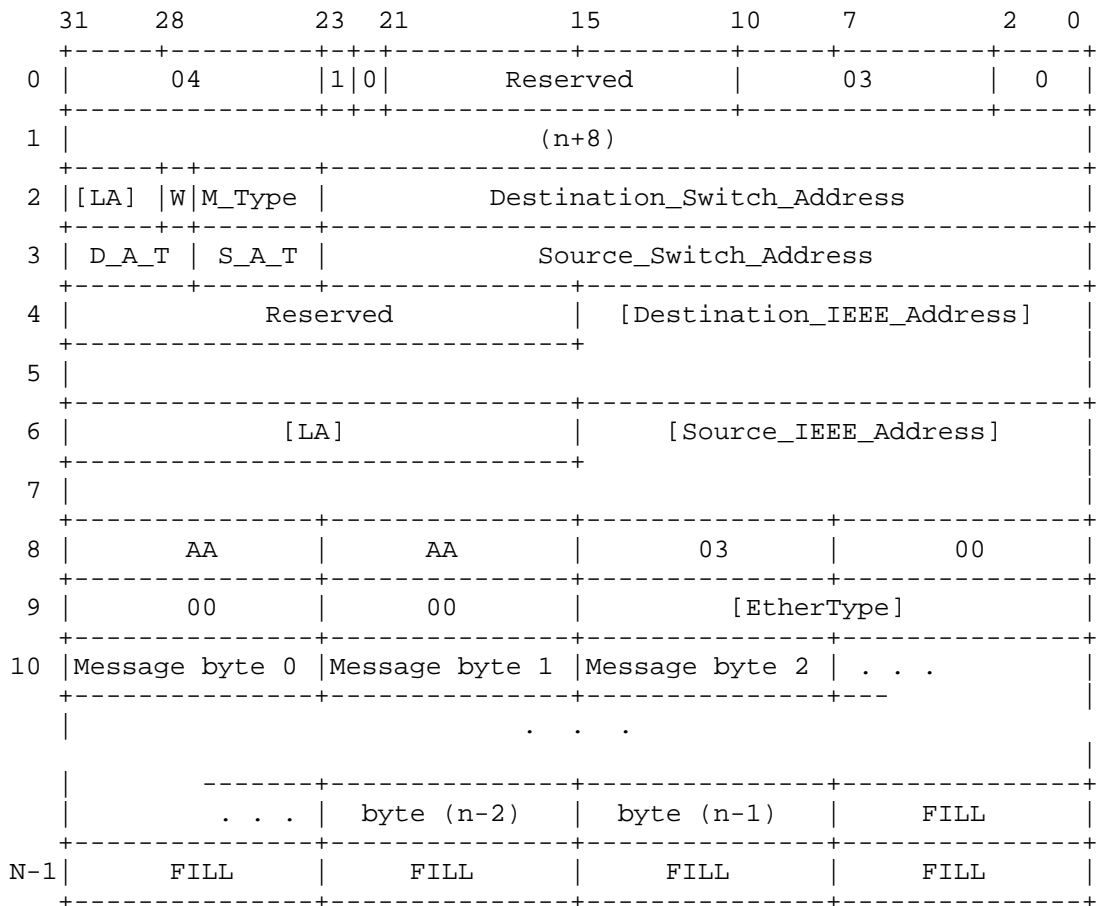
DSAP (8 bits) SHALL contain 170 ('AA'h).

CTL (8 bits) SHALL contain 3 (Unnumbered Information).

SNAP

Organization Code (24 bits) SHALL be zero.

EtherType (16 bits) SHALL be set as defined in Assigned Numbers [8]:
 IP = 2048 ('0800'h), ARP = 2054 ('0806'h), RARP = 32,821 ('8035'h).



HIPPI Packet Format

Words 0-1: HIPPI-FP Header
Words 2-7: D1 Area (HIPPI-LE Header)
Words 8-9: D2 Area (IEEE 802.2 LLC/SNAP)
Words 10-(N-1): D2 Area (IP message)
(n) is the number of bytes in the IP message.
[LA] fields are zero unless used otherwise locally.
Abbreviations: "W" = Double_Wide field;
 "M_Type" = Message_Type field;
 "D_A_T" = Destination_Address_Type;
 "S_A_T" = Source_Address_Type;
[FILL] bytes complete the HIPPI packet to an even
number of 32 bit words. The number of fill bytes
is not counted in the data length.

IEEE 802.2 Data

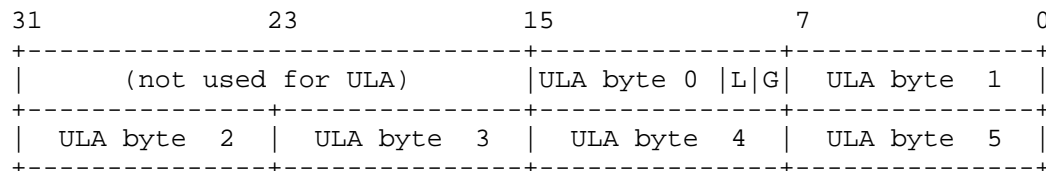
The IEEE 802.2 Data SHALL begin in the byte following the EtherType field. Fill bytes SHALL be used following the Data as necessary to make the number of bytes in the packet a multiple of 8. In accordance with HIPPI-FP, the amount of this fill is not included in the D2_Size value in the HIPPI- FP Header.

The order of the bytes in the data stream is from higher numbered to lower numbered data signal (left to right) within the HIPPI word, as specified in HIPPI-FP Clause 7, "Word and byte formats." With the 1600 megabit/second data rate option (64 bit) bits 32 through 63 are on Cable B, so that the four bytes on Cable B come logically before those on Cable A. Within each byte, the most significant bit is the highest numbered signal.

5.2 48 bit Universal LAN MAC Addresses

IEEE Standard 802.1A specifies the Universal LAN MAC Address. The globally unique part of the 48 bit space is administered by the IEEE. Each node on a HIPPI-SC LAN should be assigned a ULA. Multiple ULAs may be used if a node contains more than one IEEE 802.2 LLC protocol entity.

The format of the address within its 48 bit HIPPI-LE fields follows IEEE 802.1A canonical bit order and HIPPI-FP bit and byte order:



Universal LAN MAC Address Format

L (U/L bit) = 1 for Locally administered addresses, 0 for Universal.

G (I/G bit) = 1 for Group addresses, 0 for Individual.

The use of ULAs is optional, but encouraged. Although ULAs are not used by HIPPI-SC switches, they may be helpful for HIPPI Switch Address resolution, and for distinguishing between multiple logical entities that may exist within one node. They may also be used by gateway devices that replace HIPPI hardware headers with the MAC headers of other LANs. Carrying the ULAs in the HIPPI header may simplify these devices, and it may also help if HIPPI is used as an interface to some future HIPPI based LAN that uses ULAs for addressing.

5.3 I-Field format

fi The I-field bits, as defined in HIPPI-SC, SHALL be set as follows:

Locally Administered (bit 31) SHALL be zero.

Reserved (bits 30, 29) should be zero. Destinations SHALL accept any value for these bits.

Double wide (bit 28) SHALL be set when Source Cable B is connected and the Source wants a 64 bit connection. It SHALL be zero otherwise.

Direction (bit 27) should be sent as zero, however Destinations SHALL accept either zero or one and interpret the Routing Control field accordingly, per HIPPI-SC.

Path Selection (bits 26, 25) SHALL be 00, 01, or 11 (binary) at the Source's option. 00 (source route mode) indicates that the I-field bits 23-00 contain a 24 bit source route; 01 or 11 (logical address mode) indicate that bits 23-00 contain 12 bit Source and Destination Addresses. The value 11 is

meaningful when more than one route exists from a Source to a Destination; it allows the switch to choose the route. Use of 01 forces the switch always to use the same route for the same Source/Destination pair.

Camp-on (bit 24) may be 1 or 0; however, a Source SHALL NOT make consecutive requests without Camp-on to the same Destination while the requests are being rejected. The purpose of this restriction is to prevent a node from circumventing the fair share arbitration mechanism of the switch by repeating requests at a very high rate.

If logical address mode is used:

Source Address (bits 23-12) is not used.

Destination Address (bits 11-0) SHALL contain the Switch Address of the Destination.

If source route mode is used:

Routing control (bits 23-00) SHALL contain the route to the Destination.

5.4 Rules For Connections

The following rules for connection management by Source and Destination are intended to insure frequent, fair share access to Destinations for which multiple Sources are contending. If possible, nodes should transfer data at full HIPPI speeds and hold connections no longer than necessary.

A source may hold a connection for as long as it takes to send 68 HIPPI bursts at what ever speed the two connected nodes can achieve together. The number of packets sent in one connection is not limited, except that the number of bursts over all the packets should not exceed 68. This is not a recommendation to send as many packets as possible per connection; one packet per connection is acceptable. The purpose of this limit is to give each Source an fair share of a common Destination's bandwidth. Without a limit, if there is a Destination that is constantly in demand by multiple Sources, the Source that sends the most data per connection wins the greatest share of bandwidth.

The limit of 68 bursts is not absolute. An implementation may check the burst count after transmission of a packet and end the connection if it is greater than or equal to some threshold. If this is done, the threshold should be less than 68 depending on the typical packet

size, to ensure that the 68 burst limit is not normally exceeded. For instance, a Source sending 64K packets would send two per connection (130 bursts) if it checked for 68 at the end of each packet. In this situation the Source is required to check for a value small enough that it will not send a second packet in the same connection.

Destinations SHALL accept all packets that arrive during a connection, and may discard those that exceed its buffering capacity. A Destination SHALL NOT abort a connection (deassert CONNECT) simply because too many bursts were received; however a Destination may abort a connection whose duration has exceeded a time period of the Destination's choosing, as long as the Source is allowed ample time to transmit its quota of bursts.

The rules admonish the node to do certain things as fast as it can, however there is no absolute measure of compliance. Nodes that cannot transfer data at full HIPPI speeds can still interoperate but the faster the implementation, the better the performance of the network will be.

Assuming that bursts flow at the maximum rate, the most important factor in network throughput is the connection switching time, measured from the deassertion of REQUEST by the Source at the end of one connection to its first assertion of BURST after the establishment of the new connection.

Implementations should keep this time as short as possible. For a guideline, assuming parallel HIPPI and a single HIPPI-SC switch, ten microseconds permits nearly full HIPPI throughput with full-sized packets, and at 60 microseconds the available throughput is reduced by about 10%. (See "Performance", below.)

All HIPPI electrical signaling SHALL comply with HIPPI-PH. In every case, the following rules go beyond what HIPPI-PH requires.

Rules for the Source

1. Do not assert REQUEST until a packet is ready to send.
2. Transmit bursts as quickly as READYs permit. Except for the required HIPPI Source Wait states, there should be no delay in the assertion of BURST whenever the Source's READY counter is nonzero.
3. Make a best effort to ensure that connection durations do not exceed 68 bursts.

4. Deassert REQUEST immediately when no packet is available for immediate transmission or the last packet of the connection has been sent.

Rules for the Destination

1. Reject all connections if unable to receive packets. This frees the requesting Source to connect to other Destinations with a minimum of delay. Inability to receive packets is not a transient condition, but is the state of the Destination when its network interface is not initialized.
2. A HIPPI node should be prepared to efficiently accept connections and process incoming data packets. While this may be best achieved by not asserting connect unless 68 bursts worth of buffers is available, it may be possible to meet this requirement with fewer buffers. This may be due to a priori agreement between nodes on packet sizes, the speed of the interface to move buffers, or other implementation dependent considerations.
3. Accept a connection immediately when buffers are available. The Destination should never delay the acceptance of a connection unnecessarily.
4. Once initialized, a Destination may reject connection requests only for one of the following reasons:
 1. The I-field was received with incorrect parity.
 2. The I-field contents are invalid, e.g. the "W" bit set when the Destination does not support the 1600 megabit data rate option, the "Locally Administered" bit is set, the Source is not permitted to send to this Destination, etc.

Transient conditions within the Destination, such as temporary buffer shortages, must never cause rejected connections.

5. Ignore aborted connection sequences. Sources may time out and abandon attempts to connect; therefore aborted connection sequences are normal events.

5.5 MTU

Maximum Transmission Unit (MTU) is defined as the length of the IP packet, including IP header, but not including any overhead below IP. Conventional LANs have MTU sizes determined by physical layer specification. MTUs may be required simply because the chosen medium

won't work with larger packets, or they may serve to limit the amount of time a node must wait for an opportunity to send a packet.

HIPPI has no inherent limit on packet size. The HIPPI-FP header contains a 32 bit D2_Size field that, while it may limit packets to about 4 gigabytes, imposes no practical limit for networking purposes. Even so, a HIPPI-SC switch used as a LAN needs an MTU so that Destination buffer sizes can be determined.

The MTU for HIPPI-SC LANs is 65280 bytes.

This value was selected because it allows the IP packet to fit in one 64K byte buffer with up to 256 bytes of overhead. The overhead is 40 bytes at the present time; there are 216 bytes of room for expansion.

HIPPI-FP Header	8 bytes
HIPPI-LE Header	24 bytes
IEEE 802.2 LLC/SNAP Headers	8 bytes
Maximum IP packet size (MTU)	65280 bytes

Total	65320 bytes (64K - 216)

6 Camp-on

When several Sources contend for a single Destination, the Camp-on feature allows the HIPPI-SC switch to arbitrate and ensure that all Sources have fair access. (HIPPI-SC does not specify the method of arbitration.) Without Camp-on, the contending Sources would simply have to retry the connection repeatedly until it was accepted, and the fastest Source would usually win. To guarantee fair share arbitration, Sources are prohibited from making repeated requests to the same Destination without Camp-on in such a way as to defeat the arbitration.

There is another important reason to use Camp-on: when a connection without Camp-on is rejected, the Source cannot determine whether the rejection came from the requested Destination or from the switch. The Source also cannot tell the reason for the rejection, which could be either that the Destination was off line or not cabled, or the I-field was erroneous or had incorrect parity. Sources should not treat a rejection of a request without Camp-on as an error. Camp-on prevents rejection due to the temporary busy case; with one exception, rejection of a Camp-on request indicates an error condition, and an error event can be recorded. The exception occurs when a 64 bit connection is attempted to a Destination that does not have Cable B connected, resulting in a reject. This case is covered in "Channel Data Rate Discovery", below.

7 Path MTU Discovery

RFC 1191 [9] describes the method of determining MTU restrictions on an arbitrary network path between two hosts. HIPPI nodes may use this method without modification to discover restrictions on paths between HIPPI-SC LANs and other networks. Gateways between HIPPI-SC LANs and other types of networks should implement RFC 1191.

8 Channel Data Rate Discovery

HIPPI exists in two data rate options (800 megabit/second and 1600 megabit/second). The higher data rate is achieved by making the HIPPI 64 bits parallel instead of 32, using an extra cable containing 32 additional data bits and four parity bits. HIPPI-SC switches can be designed to attach to both. Source and Destination HIPPI implementations can be designed to operate at either rate, selectable at the time a connection is established. The "W" bit (bit 28) of the I-field controls the width of the connection through the switch. Sources with both cables A and B attached to the switch may set the "W" bit to request a 1600 megabit/second connection. If the requested destination also has both cables attached, the switch can connect Source to Destination on both cables. If the requested Destination has only Cable A, the switch rejects the request. Sixty-four bit Sources can connect to 32 bit Destinations by requesting with the "W" bit clear and not using Cable B. Sixty-four bit Destinations must examine the "W" bit in the received I-field and use or ignore Cable B accordingly. Note that both INTERCONNECT signals stay active while a 64 bit HIPPI is used in 32 bit mode.

The following table summarizes the possible combinations, the switch's action for each, and the width of the resulting connection.

		Destination	
		32	64
Source	32	W=0 Accept 32	W=0 Accept 32
	32	W=1 N/A	W=1 N/A
	64	W=0 Accept 32	W=0 Accept 32
	64	W=1 Reject	W=1 Accept 64

HIPPI Connection Combinations

If the path between a 64 bit Source and a 64 bit Destination includes more than one switch, and the route between switches uses a link that is only 32 bits wide, the switch rejects 64 bit connection requests as if the Destination did not have 64 bit capability.

In a mixed LAN of 32 bit and 64 bit HIPPIs, a 64 bit Source needs to know the data rates available at each Destination and on the path to it. This can be known a priori by manual configuration, or it can be discovered dynamically. The only reliable method of discovery is simply to attempt a 64 bit connection with Camp-on. As long as 64 bit connections succeed, the Source knows the Destination and path are double width. If a 64 bit connection is rejected, the Source tries to connect for 32 bits. If the 32 bit connection succeeds, the Source assumes that the Destination or path is not capable of double width operation, and uses only 32 bit requests after that. If the 32 bit request is rejected, the Source assumes that the Destination or path is down and makes no determination of its capability.

The Double_Wide bit in the HIPPI-LE header, if nonzero, gives the node that receives it a hint that the 64 bit connection attempt may be worthwhile when sending on the return path.

Note that Camp-on must be used at least in the 64 bit attempt, because it removes some ambiguity from the meaning of rejects. If the request is made with the "W" bit and no Camp-on, a reject could mean either that the Destination has no Cable B or that it is simply busy, and no conclusion can be drawn as to its status for 64 bit connections.

9 Performance

The HIPPI connection rules are designed to permit best utilization of the available HIPPI throughput under the constraint that each Destination must be made available frequently to receive packets from different Sources. This discipline asks both Sources and Destinations to minimize connection setup overhead to deliver high performance. Low connection setup times are easily achieved by hardware implementations, but overhead may be too high if software is required to execute between the initial request of a connection and the beginning of data transfer. Hardware implementations in which connection setup and data transfer proceed from a single software action are very desirable.

HIPPI connections are controlled by HIPPI Sources; a Destination, being unable to initiate a disconnect without the possibility of data loss, is a slave to the Source once it has accepted a connection.

Optimizations of connection strategy are therefore the province of the HIPPI Source, and several optimizations are permitted.

If the rate of available message traffic is less than the available HIPPI throughput and Destinations are seldom busy when a connection is requested, connection optimizations do not pay off and the simplest strategy of waiting indefinitely for each connection to be made and sending messages strictly in the order queued cannot be improved upon. However if some nodes are slow, or network applications can send or receive messages at a higher aggregate rate than the available HIPPI bandwidth, Sources may frequently encounter a busy Destination. In these cases, certain host output queuing strategies may enhance channel utilization. Sources may maintain separate output queues for different HIPPI Destinations, and abandon one Destination in favor of another if a connection attempt without Camp-on is rejected or a connection request with Camp-on is not accepted within a predetermined interval. Such a strategy results in aborted connection sequences (defined in HIPPI-PH: REQUEST is deasserted before any data is sent). Destinations must treat these as normal events, perhaps counting them but otherwise ignoring them.

Two components of connection setup time are out of the control of both Source and Destination. One is the time required for the switch to connect Source to Destination, currently less than four microseconds in the largest commercially available (32 port) switch. The second component is the round trip propagation time of the REQUEST and CONNECT signals, negligible on a standard 25 meter copper HIPPI cable, but contributing a total of about 10 microseconds per kilometer on fiber optic links. HIPPI-SC LANs spanning more than a few kilometers will have reduced throughput. Limited span networks with buffered gateways or bridges between them may perform better than long serial HIPPI links.

A Source is required to drop its connection after the transmission of 68 HIPPI bursts. This number was chosen to allow the transmission of one maximum sized packet or a reasonable number of smaller sized packets. The following table lists some possibilities, with calculated maximum burst and throughput rates in millions (10^6) of bytes per second:

Maximum HIPPI Throughput Rates

User Data	Number of Packets	Number of Bursts	Hold Time (usec)	Burst Rate MB/sec	-----Max throughput MB/sec-----					
					Connection Setup Overhead (usec)					
					10	30	60	90	120	150
63K	1	64	654	98.7	97.2	94.4	90.4	86.8	83.4	80.3
32K	2	66	665	98.6	97.1	94.3	90.4	86.8	83.5	80.4
16K	4	68	667	98.3	96.8	94.1	90.2	86.6	83.3	80.2
8K	7	63	587	97.8	96.1	93.0	88.7	84.8	81.2	77.8
4K	13	65	551	96.7	95.0	91.7	87.2	83.1	79.4	76.0
2K	22	66	476	94.6	92.7	89.0	84.0	79.6	75.6	72.0
1K	34	68	384	90.8	88.5	84.2	78.5	73.5	75.8	65.3

These calculations are based 259 40 ns clock periods to transmit a full burst and 23 clock periods for a short burst. (HIPPI-PH specifies three clock periods of overhead per burst.) A packet of "n" kilobytes of user data consists of "n" full bursts and one short burst equal in length to the number of bytes in the HIPPI, LLC, IP and TCP headers. "Hold Time" is the minimum connection duration needed to send the packets. "Burst Rate" is the effective transfer rate for the duration of the connection, not counting connection switching time. Throughput rates are in megabytes/second, accounting for connection switching times of 10, 30, 60, 90, 120 and 150 microseconds. These calculations ignore any limit on the rate at which a Source or Destination can process small packets; such limits may further reduce the available throughput if small packets are used.

10 Sharing the Switch

Network interconnection is only one potential application of HIPPI and HIPPI-SC switches. While network applications need very frequent transient connections, other applications may favor longer term or even permanent connections between Source and Destination. Since the switch can serve each Source or Destination with hardware paths totally separate from every other, it is quite feasible to use the same switch to support LAN interconnects and computer/peripheral applications simultaneously.

Switch sharing is no problem when unlike applications do not share a HIPPI cable on any path. However if a host must use a single input or output cable for network as well as other kinds of traffic, or if a link between switches must be shared, care must be taken to ensure that all applications are compatible with the connection discipline described in this memo. Applications that hold connections too long on links shared with network traffic may cause loss of network packets or serious degradation of network service.

11 References

- [1] ANSI X3.183-1991, High-Performance Parallel Interface - Mechanical, Electrical and Signalling Protocol Specification (HIPPI-PH).
- [2] ANSI X3.210-1992, High-Performance Parallel Interface - Framing Protocol (HIPPI-FP).
- [3] ANSI X3.218-1993, High-Performance Parallel Interface - Encapsulation of IEEE 802.2 (IEEE Std 802.2) Logical Link Control Protocol Data Units (802.2 Link Encapsulation) (HIPPI-LE).
- [4] ANSI X3.222-1993, High-Performance Parallel Interface - Physical Switch Control (HIPPI-SC).
- [5] Postel, J., "Internet Protocol", STD 5, RFC 791, USC/Information Sciences Institute, September 1981.
- [6] IEEE, "IEEE Standards for Local Area Networks: Logical Link Control", IEEE, New York, New York, 1985.
- [7] IEEE, "IEEE Standards for Local Area Networks: Logical Link Control", IEEE, New York, New York, 1985.
- [8] Reynolds, J.K., and Postel, J., "Assigned Numbers", STD 2, RFC 1340, USC/Information Sciences Institute, July 1992.
- [9] Mogul, J.C., and Deering, S.E., "Path MTU discovery", RFC 1191, Stanford University, November, 1990.

12 Security Considerations

Security issues are not discussed in this memo.

13 Author's Address

John K. Renwick
NetStar, Inc.
10250 Valley View Road
Minneapolis, MN USA 55344

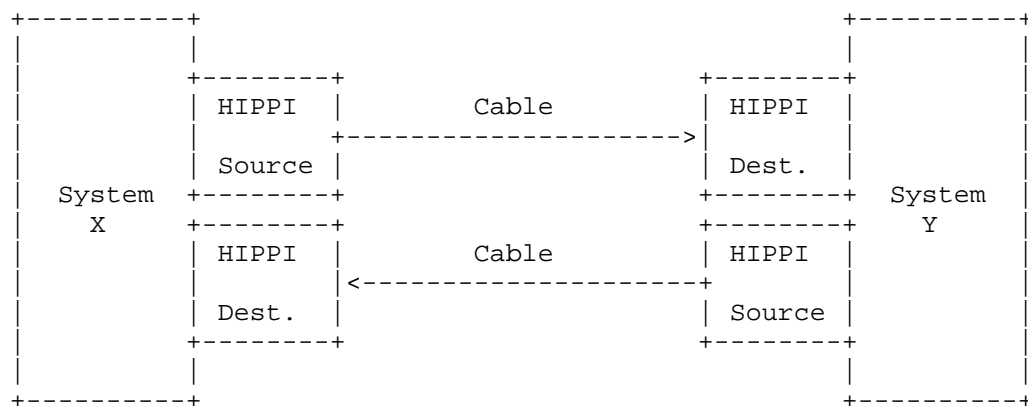
Phone: (612) 996-6847
EMail: jkr@NetStar.com

Mailing List: hippo-ext@think.com

14 Appendix A -- HIPPI Basics

This section is included as an aid to readers who are not completely familiar with the HIPPI standards.

HIPPI-PH describes a parallel copper data channel between a Source and a Destination. HIPPI transmits data in one direction only, so that two sets are required for bidirectional flow. The following figure shows a simple point-to-point link between two computer systems:



A Simple HIPPI Duplex Link

Parallel copper cables may be up to 25 meters in length.

In this document, all HIPPI connections are assumed to be paired HIPPI channels.

HIPPI-PH has a single optional feature: it can use a single cable in each direction for a 32 bit parallel channel with a maximum data rate of 800 megabit/second, or two cables for 64 bits and 1600 megabit/second. Cable A carries bits 0-31 and is used in both modes; Cable B carries bits 32-63 and is use only with the 1600 megabit/second data rate option.

HIPPI Signal Hierarchy

HIPPI has the following hardware signals:

Source to Destination

- INTERCONNECT A
- INTERCONNECT B (64 bit only)
- CLOCK (25 MHz)
- REQUEST
- PACKET
- BURST
- DATA (32 or 64 signals)
- PARITY (4 or 8 signals)

Destination to Source

- INTERCONNECT A
- INTERCONNECT B (64 bit only)
- CONNECT
- READY

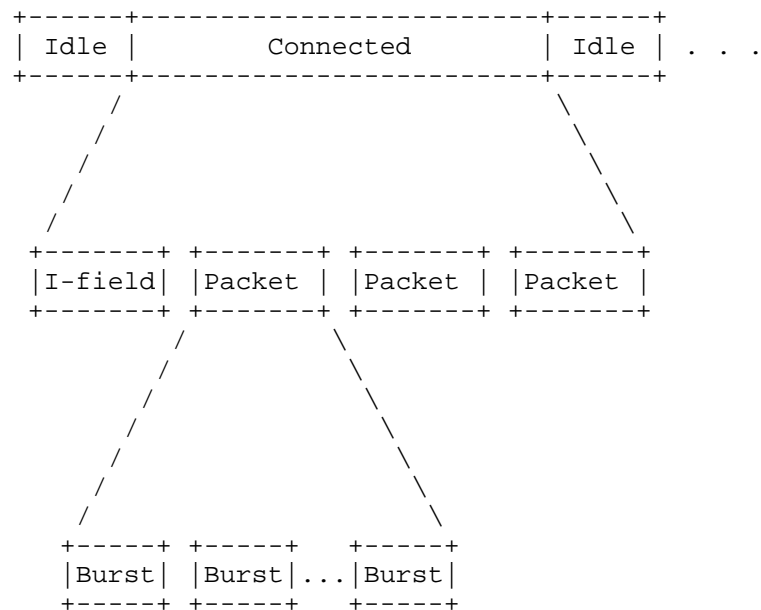
The INTERCONNECT lines carry DC voltages that indicate that the cable is connected and that the remote interface has power. INTERCONNECT is not used for signaling.

The CLOCK signal is a continuous 25 MHz (40 ns period) square wave. All Source-to-Destination signals are synchronized to the clock.

The REQUEST and CONNECT lines are used to establish logical connections. A connection is always initiated by a Source as it asserts REQUEST. At the same time it puts 32 bits of data on DATA lines 0-31, called the I-field. The Destination samples the DATA lines and can complete a connection by asserting CONNECT. Packets can be transmitted only while both REQUEST and CONNECT are asserted.

A Destination can also reject a connection by asserting CONNECT for only a short interval between 4 and 16 HIPPI clock periods (160-640 nanoseconds). The Source knows a connection has been accepted when CONNECT is asserted for more than 16 clocks or it receives a READY pulse.

Either Source or Destination can terminate a connection by deasserting REQUEST or CONNECT, respectively. Normally connections are terminated by the Source after its last Packet has been sent. A Destination cannot terminate a connection without potential loss of data.



HIPPI Logical Framing Hierarchy

The Source asserts PACKET for the duration of a Packet transmission, deasserting it to indicate the end of a Packet. A sequence of Bursts comprise a Packet. To send a burst, a Source asserts the BURST signal for 256 clock periods, during which it places 256 words of data on the DATA lines. The first or last Burst of a Packet may be less than 256 clock periods, allowing the transmission of any integral number of 32 or 64 bit words in a Packet.

The READY signal is a pulse four or more clock periods long. Each pulse signals the Source that the Destination can receive one Burst. The Destination need not wait for a burst before sending another READY if it has burst buffers available; up to 63 unanswered READYs may be sent, allowing HIPPI to operate at full speed over distances of many kilometers. If a Source must wait for flow control, it inserts idle periods between Bursts.



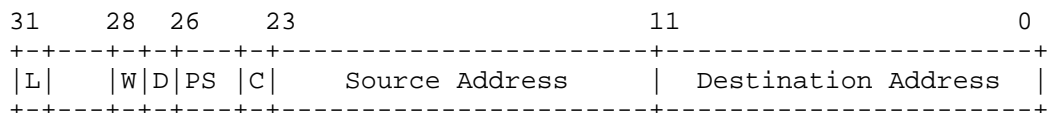
HIPPI Signal Timing Diagram

Serial HIPPI

There is no ANSI standard for HIPPI other than the parallel copper cable version. However an implementors' agreement exists, specifying a serial protocol to extend HIPPI signals on optical fiber or coaxial copper cable. Serial links may be used interchangeably with parallel links to overcome HIPPI distance limitations; they are transparent to the Source and Destination, except for the possibility of longer propagation delays.

I-Field and Switch Control

The REQUEST, CONNECT and I-field features of HIPPI-PH were designed for the control of switches as described in HIPPI-SC. A switch is a hub with a number of input and output HIPPI ports. HIPPI Sources are cabled to switch input ports, and switch output ports are cabled to HIPPI Destinations. When a HIPPI Source requests a connection, the switch interprets the I-field to select an output port and electrically connects the HIPPI Source to the HIPPI Destination on that port. Once connected, the switch does not interact with the HIPPIs in any way until REQUEST or CONNECT is deasserted, at which time it breaks the physical connection and deasserts its output signals to both sides. Some existing switch implementations can switch connections in less than one microsecond. There is the potential for as many simultaneous connections, each transferring data at HIPPI speeds, as there are input or output ports on the switch. A switch offers much greater total throughput capacity than broadcast or ring media.



HIPPI-SC I-field Format (Logical Address Mode)

- L = Locally defined (1 => entire I-field is locally defined)
- W = Width (1 => 64 bit connection)
- D = Direction (1 => swap Source and Destination Address)
- PS = Path Selection (01 => Logical Address Mode)
- C = Camp-on (1 => wait until Destination is free)

HIPPI-SC defines I-field formats for two different addressing modes. The first, called Source Routing, encodes a string of port numbers in the lower 24 bits. This string specifies a route over a number of switches. A Destination's address may differ from one Source to another if multiple switches are used.

The second format, called Logical Address Mode, defines two 12 bit fields, Source Address and Destination Address. A Destination's 12 bit Switch Address is the same for all Sources. Switches commonly have address lookup tables to map 12 bit logical addresses to physical ports. This mode is used for networking.

Control fields in the I-field are:

- L The "Locally Defined" bit, when set, indicates that the I-field is not in the standard format. The meaning of bits 30-0 are locally defined.
- W The Width bit, when set, requests a 64 bit connection through the switch. It is meaningless if Cable B is not installed at the Source. If W is set and either the Source or the requested Destination has no Cable B to the switch, the switch rejects the connection. Otherwise the switch connects both Cable A and Cable B if W is set, or Cable A only if W is clear. This feature is useful if both Source and Destination implementations can selectively disable or enable Cable B on each new connection.
- D The Direction bit, when set, reverses the sense of the Source Address and Destination Address fields. In other words, D=1 means that the Source Address is in bits 0-11 and the Destination Address is in bits 12-23. This bit was defined to give devices a simple way to route return messages. It is not useful for LAN operations.
- PS The Path Selection field determines whether the I-field contains Source Route or Address information, and in Logical Address mode, whether the switch may select from multiple possible routes to the destination. The value "01" selects Logical Address mode and fixed routes.
- C The Camp-on bit requests the switch not to reject the connection if the selected Destination is busy (connected to another Source) but wait and make the connection when the Destination is free.

15 Appendix B -- How to Build a Practical HIPPI LAN

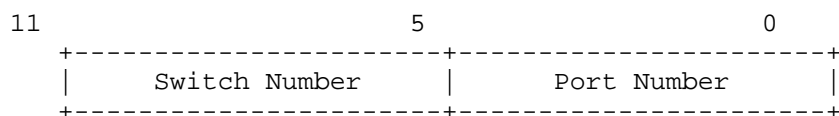
"IP on HIPPI" describes the network host's view of a HIPPI local area network without providing much information on the architecture of the network itself. Here we describe a network constructed from available HIPPI components, having the following characteristics:

1. A tree structure with a central HIPPI-SC compliant hub and optional satellite switches
2. Each satellite is connected to the hub by just one bidirectional HIPPI link.

3. Serial HIPPI or transparent fiber optic HIPPI extender devices may be used in any link.
4. Some satellites may be a particular switch product which is not HIPPI-SC compliant.
5. Host systems are attached either directly to the hub or to satellites, by single bidirectional links in which both HIPPI cables go to the same numbered switch port.

Switch Address Management

Switch addresses use a flat address space. The 12-bit address is subdivided into 6 bits of switch number and 6 bits of port number.



Logical Address Construction

Switches may be numbered arbitrarily. A given host's address consists of the number of the switch it is directly attached to and the physical port number on that switch to which its input channel is attached.

In the singly-connected tree structure, there is exactly one path between any pair of hosts. Since each satellite must be connected directly to the hub, the maximum length of this path is three hops, and the minimum length is one. Each HIPPI-SC compliant switch is programmed to map each of the host switch addresses to the appropriate output port: either the port to which the host is directly attached or a port that is linked to another switch in the path to it.

Special Treatment of Nonstandard Switches

There is one commercially available switch that was designed before the drafting of HIPPI-SC and is not fully compliant. It is in common use, so it is worth making some special provisions to allow its use in a HIPPI LAN. This switch supports only the Source Route mode of addressing with a four bit right shift that can be disabled by a hardware switch on each input port. Addresses cannot be mapped. The switch does not support the "W", "D", or "PS" fields of the I-field; it ignores their contents. Use of this switch as a satellite will require a slight deviation from normal I-field usage by the hosts that are directly attached

to it. Hosts attached to standard switches are not affected.

For a destination connected to a non compliant satellite, the satellite uses only the least significant four bits of the I-field as the address. Since the address contains the destination's physical port number in the least significant bits, its port will be selected. Nonstandard switches should be set to disable I-field shifting at the input from the hub, so that the destination host will see its correct switch address in the I-field when performing self-address discovery. I-field shifting must be enabled on the satellite for each input port to which a host is attached.

Hosts attached to nonstandard satellites must deviate from the normal I-field usage when connecting to hosts on another switch. It is suggested that all host implementations have this capability as long as the nonstandard switches remain in use. The host must know, by some manual configuration method, that it is connected to a nonstandard switch, and it must have its "link port" number; that is, the number of the port on the satellite that is connected to the hub.

The normal I-field format for a 32-bit connection, per the document, is this:

31	26	23	11	0
+-----+-----+-----+-----+-----+				
0 0 0 0 0		x 1 C		
Unused			Destination Address	
+-----+-----+-----+-----+-----+				

The special I-field format is:

31	26	24	15	4	3	0
+-----+-----+-----+-----+-----+						
0 0 0 0 0		x 1 C		Unused		
Destination Address				Link		
+-----+-----+-----+-----+-----+						

This I-field is altered by shifting the lower 24 bits left by four and adding the link port number. Camp-on is optional, and the PS field is set to 01 or 11 (the host's option) as if the switch supported logical address mode. All other I-field bits are set to zero. When the host requests a connection with this I-field, the switch selects a connection through the link port to the hub, and shifts the lower 24 bits of the I-field right by four bits. The link port number is discarded and the I-field passed through to the hub is a proper HIPPI-SC I-field selecting logical address mode.

A host on a nonstandard satellite may use the special I-field format for all connection requests. If connecting to another host on the same satellite, this will cause the connection to take an unnecessarily long path through the hub and back. If an optimization is desired, the host can be given additional information to allow it to use the standard I-field format when connecting to another host on the same switch. This information could consist of a list of the other hosts on the same switch, or the details of address formation, along with the switch number of the local satellite, which would allow the host to analyze the switch address to determine whether or not the destination is on the local switch. This optimization is fairly complicated and may not always be worthwhile.

