Network Working Group
Internet-Draft
<draft-ietf-ip1394-ipv4-10.txt>
Expires: February, 1999

P. Johansson
Congruent Software, Inc.

IPv4 over IEEE 1394

STATUS OF THIS DOCUMENT

ABSTRACT

This document specifies how to use IEEE Std 1394-1995, Standard for a High Performance Serial Bus (and its supplements), for the transport of Internet Protocol Version 4 (IPv4) datagrams. It defines the necessary methods, data structures and codes for that purpose and additionally defines a method for Address Resolution Protocol (ARP).

TABLE OF CONTENTS

1. INTRODUCTION

This document specifies how to use IEEE Std 1394-1995, Standard for a High Performance Serial Bus (and its supplements), for the transport of Internet Protocol Version 4 (IPv4) datagrams. It defines the necessary methods, data structures and codes for that purpose and additionally defines a method for Address Resolution Protocol (ARP).

The group of IEEE standards and supplements, draft or approved, related to IEEE Std 1394-1995 is hereafter referred to either as 1394 or as Serial Bus.

1394 is an interconnect (bus) that conforms to the CSR architecture, ISO/IEC 13213:1994. Serial Bus permits communications between nodes over shared physical media at speeds that range, at present, from 100 to 400 Mbps. Both consumer electronic applications (such as digital VCR's, stereo systems, televisions and camcorders) and traditional desktop computer applications (*e.g.*, mass storage, printers and tapes), have adopted 1394. Serial Bus is unique in its relevance to both consumer electronic and computer domains and is expected to form the basis of a home or small office network that combines both types of devices.

The CSR architecture describes a memory-mapped address space that Serial Bus implements as a 64-bit fixed addressing scheme. Within the address space, ten bits are allocated for bus ID (up to a maximum of 1,023 buses), six are allocated for node physical ID (up to 63 per bus) while the remaining 48 bits (offset) describe a per node address space of 256 terabytes. The CSR architecture, by convention, splits a node's address space into two regions with different behavioral characteristics. The lower portion, up to but not including 0xFFFF F000 0000, is expected to behave as memory in response to read and write transactions. The upper portion is more like a traditional IO space: read and write transactions in this area usually have side effects. Control and status registers (CSR's) that have FIFO behavior customarily are implemented in this region.

Within the 64-bit address, the 16-bit node ID (bus ID and physical ID) is analogous to a network hardware address---but 1394 node ID's are variable and subject to reassignment each time one or more nodes are added to or removed from the bus.

The 1394 link layer provides a packet delivery service with both confirmed (acknowledged) and unconfirmed packets. Two levels of service are available: "asynchronous" packets are sent on a best-effort basis while "isochronous" packets are guaranteed to be delivered with bounded latency. Confirmed packets are always asynchronous but unconfirmed packets may be either asynchronous or isochronous. Data payloads vary with implementations and may range from one octet up to a maximum determined by the transmission speed (at 100 Mbps, named S100, the maximum asynchronous data payload is 512 octets while at S400 it is 2048 octets).

NOTE: Extensions underway in IEEE P1394b contemplate additional speeds of 800, 1600 and 3200 Mbps.

2. DEFINITIONS AND NOTATION

2.1 Conformance

When used in this document, the keywords "may", "optional", "recommended", "required", "shall" and "should" differentiate levels of requirements and optionality and are to be interpreted as described in RFC 2119.

Several additional keywords are employed, as follows:

expected: A keyword used to describe the behavior of the hardware or software in the design models assumed by this standard. Other hardware and software design models may also be implemented.

ignored: A keyword that describes bits, octets, quadlets or fields whose values are not checked by the recipient.

reserved: A keyword used to describe objects---bits, octets, quadlets and fields---or the code values assigned to these objects in cases where either the object or the code value is set aside for future standardization. Usage and interpretation may be specified by future extensions to this or other standards. A reserved object shall be zeroed or, upon development of a future standard, set to a value specified by such a standard. The recipient of a reserved object shall not check its value. The recipient of an object defined by this standard as other than reserved shall check its value and reject reserved code values.

2.2 Glossary

The following terms are used in this standard:

address resolution protocol: A method for a requester to determine the hardware (1394) address of an IP node from the IP address of the node.

bus ID: A 10-bit number that uniquely identifies a particular bus within a group of multiple interconnected buses. The bus ID is the most significant portion of a node's 16-bit node ID. The value 0x3FF designates the local bus; a node shall respond to requests addressed to its 6-bit physical ID if the bus ID in the request is either 0x3FF or the bus ID explicitly assigned to the node.

encapsulation header: A structure that precedes all IP data transmitted over 1394. See also link fragment.

IP datagram: An Internet message that conforms to the format specified by RFC 791.

link fragment: A portion of an IP datagram transmitted within a single 1394 packet. The data payload of the 1394 packet contains both an encapsulation header and its associated link fragment. It is possible to transmit datagrams without link fragmentation.

multicast channel owner: A multicast source that has allocated a channel for one or more multicast addresses and transmits MCAP advertisements to communicate these channel mapping(s) to other participants in the multicast group. When more than one source transmits MCAP advertisements for the same channel number, the source with the largest physical ID is the owner.

node ID: A 16-bit number that uniquely identifies a Serial Bus node within a group of multiple interconnected buses. The most significant 10 bits are the bus ID and the least significant 6 bits are the physical ID.

node unique ID: A 64-bit number that uniquely identifies a node among all the Serial Bus nodes manufactured worldwide; also known as the EUI-64 (Extended Unique Identifier, 64-bits).

octet: Eight bits of data.

packet: Any of the 1394 primary packets; these may be read, write or lock requests (and their responses) or stream data. The term "packet" is used consistently to differentiate 1394 packets from ARP requests/responses or IP datagrams.

physical ID: On a particular bus, this 6-bit number is dynamically assigned during the self-identification process and uniquely identifies a node on that bus.

quadlet: Four octets, or 32 bits, of data.

stream packet: A 1394 primary packet with a transaction code of 0x0A that contains a block data payload. Stream packets may be either asynchronous or isochronous according to the type of 1394 arbitration employed.

2.3 Abbreviations

The following are abbreviations that are used in this standard:

ARP	Address resolution protocol
CSR	Control and status register
CRC	Cyclical redundancy checksum
EUI-64	Extended Unique Identifier, 64-bits
IP	Internet protocol (within the context of this document, IPv4)
MCAP	Multicast channel allocation protocol

2.4 Numeric values

Decimal and hexadecimal numbers are used within this standard. By editorial convention, decimal numbers are most frequently used to represent quantities or counts. Addresses are uniformly represented by hexadecimal numbers. Hexadecimal numbers are also used when the value represented has an underlying structure that is more apparent in a hexadecimal format than in a decimal format.

Decimal numbers are represented by Arabic numerals or by their English names. Hexadecimal numbers are prefixed by 0x and represented by digits from the character set 0 – 9 and A – F. For the sake of legibility, hexadecimal numbers are separated into groups of four digits separated by spaces.

For example, both 42 and 0x2A represent the same numeric value.

3. IP-CAPABLE NODES

Not all 1394 devices are capable of the reception and transmission of ARP requests/responses or IP datagrams. An IP-capable node shall fulfill the following minimum requirements:

- the *max_rec* field in its bus information block shall be at least 8; this indicates an ability to accept write requests with data payload of 512 octets. The same ability shall also apply to read requests; that is, the node shall be able to transmit a response packet with a data payload of 512 octets;

- it shall be isochronous resource manager capable, as specified by 1394;

- it shall support both reception and transmission of asynchronous streams as specified by P1394a;

- it shall implement the NETWORK_CHANNELS register; and

- it shall be network protocol manager (NPM) capable.

4. NETWORK_CHANNELS REGISTER

This register is required for IP-capable nodes. It shall be located at offset 0xFFFF F000 0234 within the node's address space and shall support quadlet read and write requests, only. The format of the register is shown below.

```
                     1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|1|v|   npm_ID  |              reserved             |  channel  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1 - NETWORK_CHANNELS format

Upon a node power reset or a bus reset, the entire register (with the exception of the most significant bit and the *npm_ID* field) shall be cleared to zero; the *npm_ID* field shall be set to ones.

The most significant bit (a constant one) differentiates the presence of the NETWORK_CHANNELS register in an IP-capable node from the value (all zeros) possibly returned when offset 0xFFFF F000 0234 is read at node(s) that do not implement this register.

NOTE: Nodes compliant with P1394a return an address error response when unimplemented addresses are accessed---but some 1394 implementations are known to return zeros.

The *valid* bit (abbreviated as *v* above), when set to one, indicates that the *channel* field contains meaningful information. IP-capable nodes shall transmit neither ARP requests/responses nor broadcast IP datagrams while the *valid* bit is zero.

The *npm_ID* field identifies the physical ID of the network protocol manager (NPM). When *npm_ID* is equal to 0x3F the physical ID of the NPM is not specified; otherwise it shall be initialized (by the NPM) to the 6-bit physical ID assigned during the self-identification process.

The *channel* field shall be initialized by the NPM (see below) to identify the channel number shared by IP-capable nodes for ARP and IP broadcast.

Only the *valid* bit and the *npm_ID* and *channel* fields may be changed by quadlet write requests; the data value in the write request shall be ignored for all other bit positions.

5. NETWORK PROTOCOL MANAGER (NPM)

In order for ARP or broadcast IP to function on 1394, a prerequisite is the presence of a network protocol manager (NPM). The domain of the NPM is limited to the local Serial Bus; the functions of the NPM are as follows:

- the allocation of a channel number for ARP and broadcast IP; and

- the communication of that channel number to all IP-capable nodes on the same bus.

All IP-capable nodes shall be capable of functioning as the NPM.

Subsequent to a Serial Bus reset a single NPM shall be determined by a distributed algorithm executed by all the NPM-capable nodes. The algorithm is straightforward: the NPM-capable node with the largest 6-bit physical ID shall be the NPM. The steps in the algorithm are as follows:

a) An NPM-capable node shall also be a contender for the role of isochronous resource manager. The C (contender) and L (link active) bits in its self-ID packet shall be set to one;

b) Subsequent to a bus reset, isochronous resource manager contention takes place during the self-identification process specified by 1394;
c) An NPM-capable node that wins the contention process referenced in b) is the NPM and shall proceed with g). Other NPM-capable node(s) not selected as the isochronous resource manager (hereafter referred to as candidates) shall continue with d);
d) A candidate NPM shall delay before it attempts to become the NPM. The delay time shall be equal to 15 ms * (*irm_ID* - *candidate_ID*), where *irm_ID* and *candidate_ID* are the physical ID’s of the isochronous resource manager and the candidate NPM, respectively. After the delay time has elapsed, the candidate NPM shall examine the *npm_ID* field in its own NETWORK_CHANNELS register; if it is not equal to 0x3F, another node is the NPM. The losing candidate shall wait for the *valid* bit of its own register to be set before transmitting any ARP requests/responses or IP datagrams;
e) Otherwise, the candidate NPM shall attempt to read the NETWORK_CHANNELS register of any contenders with a larger physical ID (these nodes were identified by the C bit in their self-ID packets). The candidate NPM shall read the NETWORK_CHANNELS register in the contender with the largest physical ID and progress downward. If the register is implemented, the NPM is determined to be a different node. The losing candidate shall ignore the contents of NETWORK_CHANNELS returned in the read response and shall wait for the *valid* bit of its own register to be set before transmitting any ARP requests/responses or IP datagrams;
f) If no contender with a physical ID larger than the candidate NPM's physical ID implements the NETWORK_CHANNELS register, the search is complete and the candidate becomes the new NPM;
g) Once elected, the NPM shall update the *npm_ID* field in the NETWORK_CHANNELS register of all the IP-capable nodes on the bus (including itself) with its own physical ID. This signals to other candidates that an NPM has been elected but may not have allocated a channel. Either a broadcast write request or a series of write requests addressed to individual nodes may be used;
h) The NPM shall attempt to allocate a channel number from the CHANNELS_AVAILABLE register (note that the NPM may also be the isochronous resource manager). If no channel number had been allocated prior to the bus reset, the NPM shall wait one second before it attempts to allocate a channel number. Otherwise, the NPM shall attempt to reallocate the same channel number in use before the bus reset; if the same channel number is not available, the NPM may allocate a different channel number. If no channel number is available, the NPM shall take no additional action (all *valid* bit(s) were cleared by the bus reset);

NOTE: Parts of the preceding step are still under discussion within the working group; there is as yet no consensus as to what time interval the NPM shall wait (if any) before attempting to allocate a new channel number if the previously allocated channel number is unavailable after a bus reset.

i) Otherwise, the NPM shall update its own NETWORK_CHANNELS register with the allocated channel number and set the *valid* bit to one. The NPM shall then write the updated value of the entire register to the NETWORK_CHANNELS register of all the IP-capable nodes on the bus. Either a broadcast write request or a series of write requests addressed to individual nodes may be used to propagate the information.

In the case that the NPM is unable to allocate a channel number for ARP and broadcast IP, a warning should be communicated to a user that IP initialization could not complete because of a lack of Serial Bus resources. The user should be advised to reconfigure or remove other devices if she wishes to make use of IP.

NOTE: If the NPM is unable to allocate a channel number, IP-capable nodes are unable to use the ARP and broadcast IP methods specified by this document. If other methods (*e.g.*, a search of configuration ROM) permit IP-capable nodes to discover each other, they may be able to send and receive IP datagrams.

An IP-capable node that is not the NPM typically awaits a write to its NETWORK_CHANNELS that sets the *valid* bit to one; this indicates that the *channel* field is valid for ARP and IP broadcast. If some time-out elapses without this occurrence, an IP-capable node may attempt to locate the NPM and retrieve valid information from the NETWORK_CHANNELS register. If the *npm_ID* field in its own NETWORK_CHANNELS register is not equal to 0x3F, the address of the NPM is known; otherwise the node may search for the NPM as described in e) above. In either case, it is recommended that reads of the NETWORK_CHANNELS register not be performed within a tight loop, as this could adversely affect both IP and overall 1394 performance on the local bus.

6. LINK ENCAPSULATION AND FRAGMENTATION

All IP datagrams (broadcast, unicast or multicast), as well as ARP requests/responses, that are transferred *via* 1394 block write requests or stream packets shall be encapsulated within the packet's data payload. The maximum size of data payload, in octets, is constrained by the speed at which the packet is transmitted.

Table 1 - Maximum data payloads

Speed	Asynchronous	Isochronous
S100	512	1024
S200	1024	2048
S400	2048	4096
S800	4096	8192
S1600	8192	16384
S3200	16384	32768

The maximum data payload for asynchronous requests and responses may also be restricted by the capabilities of the sending or receiving node(s); this is specified by *max_rec* in either the bus information block or ARP response.

For either of these reasons, the minimum capabilities between IP-capable nodes may be less than the 1500 octet maximum transmission unit (MTU) specified by this document. This requires that the encapsulation format also permit 1394 link-level fragmentation and reassembly of IP datagrams.

6.1 Encapsulation header

All IP datagrams transported over 1394 are prefixed by an encapsulation header with one of the formats illustrated below.

If an entire IP datagram may be transmitted within a single 1394 packet, it is unfragmented and the first quadlet of the data payload shall conform to the format illustrated below.

```
                        1                   2                   3
    0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   | lf|          reserved         |           ether_type          |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 2 - Unfragmented encapsulation header format

The *lf* field shall be zero .

The *ether_type* field shall indicate the nature of the datagram that follows, as specified by the following table.

ether_type	Datagram
0x800	IPv4
0x806	ARP

NOTE: Other network protocols, identified by different values of *ether_type*, may use the encapsulation formats defined herein but such use is outside of the scope of this document.

In cases where the length of the datagram exceeds the maximum data payload supported by the sender and all recipients, the datagram shall be broken into link fragments; the first two quadlets of the data payload for the first link fragment shall conform to the format shown below.

```
                     1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| lf|rsv|      buffer_size      |          ether_type           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              dgl              |          signature            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 3 - First fragment encapsulation header format

The second and subsequent link fragments (up to and including the last) shall conform to the format shown below.

```
                     1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| lf|rsv|      buffer_size      |  rsv  |    fragment_offset    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              dgl              |          signature            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 4 - Subsequent fragment(s) encapsulation header format

The definition and usage of the fields is as follows:

The *lf* field shall specify the relative position of the link fragment within the IP datagram, as encoded by the following table.

lf	Position
0	Unfragmented
1	First
2	Last
3	Interior

buffer_size: The size of the buffer, expressed as *buffer_size* + 1 octets, necessary for the recipient to reassemble the link fragments.

ether_type: This field is present only in the first link fragment and shall have a value of 0x800, which indicates an IPv4 datagram.

fragment_offset: This field is present only in the second and subsequent link fragments and shall specify the offset, in octets, of the fragment from the beginning of the IP datagram. The first octet of the datagram (the start of the IP header) has an offset of zero; the implicit value of *fragment_offset* in the first link fragment is zero.

dgl: The value of *dgl* (datagram label) shall be the same for all link fragments of an IP datagram. The sender shall increment *dgl* for successive, fragmented datagrams; the incremented value of *dgl* shall wrap from 65,535 back to zero.

signature: The sender shall set this field to the most significant 16-bits of its own NODE_IDS register.

All IP datagrams, regardless of the mode of transmission (block write requests or stream packets) shall be preceded by one of the above described encapsulation headers. This permits uniform software treatment of datagrams without regard to the mode of their transmission.

6.2 Link fragment reassembly

The recipient of an IP datagram transmitted *via* more than one 1394 packet shall use both *signature* and *dgl* to identify all the link fragments from a single datagram. Subsequent to reassembly, the recipient shall verify the IP header checksum of the datagram.

NOTE: The use of *signature* for any purpose other than link fragment reassembly is fraught with error and is strongly discouraged.

Upon receipt of a link fragment, the recipient may place the data payload (absent the encapsulation header) within an IP datagram reassembly buffer at the location specified by *fragment_offset*. The size of the reassembly buffer may be determined from *buffer_size*.

If a link fragment is received that overlaps another fragment for the same *signature* and *dgl*, the fragment(s) already accumulated in the reassembly buffer shall be discarded. A fresh reassembly may be commenced with the most recently received link fragment. Fragment overlap is determined by the combination of *fragment_offset* from the encapsulation header and *data_length* from the 1394 packet header.

Upon detection of a Serial Bus reset, recipient(s) shall discard all link fragments of all partially reassembled IP datagrams and sender(s) shall discard all not yet transmitted link fragments of all partially transmitted IP datagrams.

7. ADDRESS RESOLUTION PROTOCOL (ARP)

ARP requests ~~and responses~~ shall be transmitted by the same means as broadcast IP datagrams; ARP responses may be transmitted in the same way or they may be transmitted as unicast IP datagrams. An ARP request/response is 56 octets and shall conform to the format illustrated below.

```
                        1                   2                   3
    0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |    hardware_type (0x0018)     |    protocol_type (0x0800)     |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |  hw_addr_len  |  IP_addr_len  |            opcode             |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                                                               |
   +---                    sender_unique_ID                     ---+
   |                                                               |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |         sender_node_ID        |     sender_unicast_FIFO_hi    |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                    sender_unicast_FIFO_lo                     |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   | sender_max_rec|      sspd     |            reserved           |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                      sender_IP_address                        |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                                                               |
   +---                    target_unique_ID                     ---+
   |                                                               |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |         target_node_ID        |     target_unicast_FIFO_hi    |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                    target_unicast_FIFO_lo                     |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   | target_max_rec|      tspd     |            reserved           |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                      target_IP_address                        |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 5 - ARP request/response format

Field usage in an ARP request/response is as follows:

hardware_type: This field indicates 1394 and shall have a value of 0x0018.

protocol_type: This field shall have a value of 0x0800; this indicates that the protocol addresses in the ARP request/response conform to the format for IP addresses.

hw_addr_len: This field indicates the size, in octets, of the 1394-dependent hardware address associated with an IP address and shall have a value of 20.

IP_addr_len: This field indicates the size, in octets, of an IP version 4 (IPv4) address and shall have a value of 4.

opcode: This field shall be one to indicate an ARP request and two to indicate an ARP response.

sender_unique_ID: This field shall contain the node unique ID of the sender and shall be equal to that specified in the sender's bus information block.

sender_node_ID: This field shall contain the most significant 16 bits of the sender's NODE_IDS register.

sender_unicast_FIFO_hi and *sender_unicast_FIFO_lo*: These fields together shall specify the 48-bit offset of the sender's FIFO available for the receipt of IP datagrams in the format specified by section 8. The offset of a sender's unicast FIFO shall not change, except as the result of a power reset .

sender_max_rec: This field shall be equal to the value of *max_rec* in the sender’s configuration ROM bus information block.

sspd: This field shall be set to the lesser of the sender’s link speed and PHY speed. The link speed is the maximum speed at which the link may send or receive packets; the PHY speed is the maximum speed at which the PHY may send, receive or repeat packets. The encoding used for *sspd* is specified by the table below; all values not specified are reserved.

Table 2 - Speed codes

Value	Speed
0	S100
1	S200
2	S400
3	S800
4	S1600
5	S3200

sender_IP_address: This field shall specify the IP address of the sender.

target_unique_ID: In an ARP request, the value of this field is not specified; it shall be ignored by the recipient. In an ARP response, it shall be set to the value of *sender_unique_ID* from the corresponding ARP request.

target_node_ID: In an ARP request, the value of this field is not specified; it shall be ignored by the recipient. In an ARP response, it shall be set to the value of *sender_node_ID* from the corresponding ARP request.

target_unicast_FIFO_hi and *target_unicast_FIFO_lo*: In an ARP request, the value of these fields is not specified; they shall be ignored by the recipient. In an ARP response, they shall be set to the value of *sender_unicast_FIFO_hi* and *sender_unicast_FIFO_lo* from the corresponding ARP request.

target_max_rec: In an ARP request, the value of this field is not specified; it shall be ignored by the recipient. In an ARP response, it shall be equal to the value of *max_rec* from the corresponding ARP request.

tspd: In an ARP request, the value of this field is not specified; it shall be ignored by the recipient. In an ARP response, it shall be equal to the value of *sspd* from the corresponding ARP request.

target_IP_address: In an ARP request, this field shall specify the IP address from which the responder desires a response. In an ARP response, it shall be set to the value of *sender_IP_address* from the corresponding ARP request.

8. IP UNICAST

A unicast IP datagram may be transmitted to a recipient within a 1394 primary packet that has one of the following transaction codes:

tcode	Description	Arbitration
0x01	Block write	Asynchronous
0x0A	Stream packet	Isochronous
0x0A	Stream packet	Asynchronous

Block write requests are suitable when 1394 link-level acknowledgement is desired but there is no need for bounded latency in the delivery of the packet (quality of service).

Isochronous stream packets provide quality of service guarantees but no 1394 link-level acknowledgement.

The last method, asynchronous stream packets, is mentioned only for the sake of completeness. This method should not be used for IP unicast, since it provides for neither 1394 link-level acknowledgment nor quality of service---and consumes a valuable resource, a channel number.

Regardless of the IP unicast method employed, asynchronous or isochronous, it is the responsibility of the sender of a unicast IP datagram to determine the maximum data payload that may be used in each packet. The necessary information may be obtained from:

- the SPEED_MAP maintained by the 1394 bus manager, which provides the maximum transmission speed between any two nodes on the local Serial Bus. The bus manager analyzes bus topology in order to construct the speed map; the maximum transmission speed between nodes reflects the capabilities of the intervening nodes. The speed in turn implies a maximum data payload (see Table 1);

- the *target_max_rec* field in an ARP response. This document requires a minimum value of 8 (equivalent to a data payload of 512 octets).

Nodes that operate at S200 and faster are encouraged but not required to implement correspondingly larger values for *target_max_rec*; or

- other methods beyond the scope of this standard.

The maximum data payload shall be the minimum of the largest data payload implemented by the sender, the recipient and the PHYs of all intervening nodes (the last is implicit in the SPEED_MAP entry indexed by sender and recipient).

NOTE: The SPEED_MAP is derived from the self-ID packets transmitted by all 1394 nodes subsequent to a bus reset. An IP-capable node may observe the self-ID packets directly.

8.1 Asynchronous IP unicast

Unicast IP datagrams that do not require any quality of service shall be contained within the data payload of 1394 block write transactions addressed to the *target_node_ID* and *target_unicast_FIFO* obtained from an ARP response .

If no acknowledgement is received in response to a unicast block write request, the state of the target is ambiguous.

NOTE: An acknowledgment may be absent because the target is no longer functional, may not have received the packet because of a header CRC error or may have received the packet successfully but the acknowledge sent in response was corrupted.

8.2 Isochronous IP unicast

Unicast IP datagrams that require quality of service shall be contained within the data payload of 1394 isochronous stream packets.
The details of coordination between nodes with respect to allocation of channel number(s) and bandwidth are beyond the scope of this standard.

9. IP BROADCAST

Broadcast IP datagrams are encapsulated according to the specifications of section 6 and are transported by asynchronous stream packets. There is no quality of service provision for IP broadcast over 1394. The channel number used for IP broadcast is specified by the NETWORK_CHANNELS register.

All broadcast IP datagrams shall use asynchronous stream packets whose channel number is equal to the *channel* field from the NETWORK_CHANNELS register.

Although 1394 permits the use of previously allocated channel number(s) for up to one second subsequent to a bus reset, IP-capable nodes shall not transmit asynchronous stream packets at any time the *valid* bit in their NETWORK_CHANNELS register is zero. Since the *valid* bit is automatically cleared to zero by a bus reset, this prohibits the use of ARP or broadcast IP until the NPM allocates a channel number.

10. IP MULTICAST

Multicast IP datagrams are encapsulated according to the specifications of section 6 and are transported by stream packets. Asynchronous streams are used for best-effort IP multicast while isochronous streams are used for IP multicast that requires quality of service.

CAUTION: The working group has yet to define facilities and methods for the provision of quality of service for IP multicast.

By default, all best-effort IP multicast shall use asynchronous stream packets whose channel number is equal to the *channel* field from the NETWORK_CHANNELS register. In particular, datagrams addressed to 224.0.0.1 and 224.0.0.2 shall use this channel number. Best-effort IP multicast for ~~particular~~other multicast group addresses may utilize a different channel number if such a channel number is allocated and advertised prior to use, as described below.

IP-capable nodes may transmit best-effort IP multicast only if one of the following two conditions is met:

- the channel number in the stream packet is equal to the *channel* number field in the NETWORK_CHANNELS register and the *valid* bit in the same register is one; or

- for other channel number(s), some source of IP multicast has allocated and is advertising the channel number used.

The remainder of this section describes a multicast channel allocation protocol (MCAP) employed by both IP multicast sources and recipients whenever a channel number other than the default is used. MCAP is a cooperative protocol; the participants exchange messages over the broadcast channel used by all IP-capable nodes on a particular Serial Bus.

CAUTION: The working group has yet to define facilities and methods for shared use of a single channel number (other than the default channel number specified by the NETWORK_CHANNELS register) by more than one IP multicast address.

10.1 MCAP Message Format

MCAP messages, whether sent by a multicast ~~source~~channel owner or recipient, have the format illustrated below. The first eight octets of the message are fixed; the remainder consists of variable-length tuples, each of which encodes information about a particular multicast group. Individual MCAP messages shall not be fragmented and shall be encapsulated within a stream packet as *ether_type* 0x8861.

```
                     1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             length            |       checksumreserved        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|        bus_ID         |  phy_ID   |    reserved   |    opcode     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
+                          message data                         +
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 5 - MCAP message format

Field usage in an MCAP message is as follows:

length: This field shall contain the size, in octets, of the entire MCAP message.

~~*checksum*: This field shall contain a checksum calculated on the entire MCAP message. The checksum shall be the one's complement of the one's complement sum of all the 16-bit words in the message. For the purpose of calculating the checksum, the *checksum* field is treated as if zero.~~

bus_ID: This field shall specify the 10-bit bus ID for which information in the MCAP message is valid. The value of *bus_ID* shall be equal to the most significant 10 bits of the sender's NODE_IDS register.

phy_ID: This field is valid when *opcode* indicates an MCAP advertisement. In this case, *bus_ID* and *phy_ID* together specify the 16-bit node ID of the owner of all channel number(s) present in the message data.

opcode: This field shall have one of the values specified by the table below.

opcode	Name	Comment
0	Advertise	Sent by a multicast ~~source~~channel owner to broadcast the current mapping(s) from one or more group addresses to their corresponding channel number(s).
1	Solicit	Sent to request multicast ~~source~~channel owner(s) to advertise the indicated channel mapping(s) as soon as possible.

message data: The remainder of the MCAP message is variable in length and shall consist of zero or more group address descriptors with the format illustrated below.

```
                     1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    length     |     type      |           reserved            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  expiration   |    channel    |     speed     |   reserved    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           bandwidth                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
+                         group_address                         +
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 5 - MCAP group address descriptor format

length: This field shall contain the size, in octets, of the MCAP group address descriptor.

type: This field shall have a value of one, which indicates a group address descriptor.

expiration: The usage of this field varies according to *opcode*. For solicit messages the *expiration* field shall be ignored. Otherwise, for advertisements, this field shall contain a time-stamp, in seconds, that specifies a future time after which the channel number specified by *channel* may no longer be used.~~ Time is expressed in terms of the CYCLE_TIME.*seconds*; a match occurs when *expiration* equals the most significant seven bits of the CYCLE_TIME register.~~

channel: This field is valid only for advertise messages, in which case it shall specify a valid channel number, in the range zero to 63 inclusive. All other values are reserved.

speed: This field is valid only for advertise messages, in which case it shall specify the speed at which stream packets for the indicated *channel* are transmitted. The encoding used for *speed* is specified by Table 2.

bandwidth: This field shall be ~~ignored~~zero; it is allocated in the group address descriptor to accommodate future extensions to MCAP that specify quality of service and utilize the isochronous capabilities of Serial Bus.

group_address: This variable length field shall specify the IP address of a particular multicast group. The length of *group_address*, in octets, is derived from the length of the group address descriptor by subtracting 12 from the *length* field.

10.2 Multicast receive

An IP-capable device that wishes to receive multicast data not transmitted on the default channel shall first ascertain the channel mapping (if any) that exists between a group address and a channel number. Such a device may observe the MCAP ~~messages~~advertisements on the broadcast channel for the desired channel mapping~~ or it may transmit a solicitation request with the desired channel mapping~~(s).

An intended multicast recipient may transmit MCAP solicitation requests in order to request multicast channel owner(s) to broadcast advertisements sooner than the next ten second interval. Originators of MCAP solicitation requests shall limit the rate at which they are transmitted. Subsequent to sending a solicitation request, ~~neither~~ the originator ~~nor any other node that observes the request~~ shall not send another MCAP solicitation request ~~that specifies any of the group addresses contained in the first~~ until ~~either a)~~ 10 seconds have expired ~~or b) an MCAP advertisement has been observed~~.

In either case, if a valid mapping exists for the group address an MCAP advertise message is expected within ten seconds. Upon receipt of an MCAP advertise message that describes one or more valid channel mappings, the intended multicast recipient may receive IP datagrams on the indicated channel number(s) until the *expiration* time.

If multiple MCAP advertise messages are observed that specify the same group address, the valid channel number shall be obtained from the advertisement message with the largest *phy_ID*.

If no MCAP advertise message is received for the desired group addresses, no multicast sources are active and there is no data to receive.

10.3 Multicast transmit

An IP-capable device that wishes to transmit multicast data on other than the default channel shall first ascertain whether or not another multicast source has already allocated a channel number for the group address. The intended multicast source may transmit an MCAP solicitation request with one or more group address descriptors.

Whether or not a solicitation request has been transmitted, the intended multicast source shall monitor the broadcast channel for MCAP advertisements. If a valid channel mapping already exists for the group address, an MCAP advertisement should be received within ten seconds. In this case the intended multicast source may commence transmission of IP datagrams on the indicated channel number(s) and may continue to do so until their *expiration* time. The multicast source shall monitor MCAP advertisements in order to refresh the *expiration* time of channel number(s) in use.

When no other multicast source has established a valid channel mapping for the group address, the intended multicast source may attempt to allocate a channel number from the isochronous resource manager's CHANNELS_AVAILABLE register according to the procedures described in IEEE Std 1394-1995. If the channel number allocation is successful, the multicast source shall advertise the new channel mapping(s); once such advertisement has been made, the multicast source may transmit IP datagrams using the channel number obtained.

Multicast IP datagrams may be transmitted on the default channel until the sender observes (or transmits) an advertisement that specifies non-default channel mapping(s) for the multicast addresses. This permits the smooth transition of multicast from the default channel to an explicitly allocated channel.

10.4 Advertisement of channel mappings

Each multicast source shall periodically broadcast an advertisement of all multicast group addresses for which it has allocated a channel number different from the default multicast channel number. An advertisement shall consist of a single MCAP message with an *opcode* of zero which contains one or more group address descriptors (one for each group address assigned a channel number other than that specified by the NETWORK_CHANNELS register).

Within each group address descriptor, the *group_address* and *channel* fields associate a multicast group address with a Serial Bus channel number. More than one multicast group address may be mapped to a single Serial Bus channel number by means of separate group address descriptors. The *speed* field specifies the maximum 1394 speed at which any of the senders within the multicast group is permitted to transmit data. The *expiration* field specifies a future time after which the channel mapping(s) are no longer valid.

No more than ten seconds shall elapse from the transmission of its most recent advertisement before a the owner of a channel mapping initiates transmission of the subsequent advertisement.

10.5 Overlapped channel mappings

When two intended multicast sources wish to transmit to the same multicast group and no valid channel mapping exists for the group address, there is a chance that both will allocate channel numbers and both will advertise the channel mappings. These channel mappings overlap, *i.e.*, the same group address is mapped to more than one channel number.

Multicast ~~sources~~channel owners shall monitor MCAP advertisements in order to detect overlapped channel mappings. When an overlapped channel mapping is detected, the owner ~~of the largest channel number~~ with the largest *phy_ID* is not required to take any action. ~~The owners of all smaller channel number(s) mapped to the same group address shall invalidate their own (overlapped) channel mapping(s) as soon as possible by transmitting an MCAP advertisement message with the *expiration* time no more than ten seconds in the future. Once the~~ The owner(s) with smaller physical ID's shall cease transmission of MCAP advertisements for the overlapped channel number. As soon as these channel mapping(s) are no longer valid, their owners shall deallocate any unused channel numbers as described in 10.7 below.

10.6 Transfer of channel ownership

The owner of a channel mapping may cease multicast transmission on a particular channel, in which case it should invalidate the channel mapping and in some cases deallocate the channel number. Because other multicast sources may be using the same channel mapping, an orderly process is defined to transfer channel ownership.

The owner of an existing channel mapping that wishes to release the mapping shall transmit an MCAP advertisement with an *expiration* time at least 30 seconds in the future. The owner shall commence a timer to measure *expiration* seconds; until the timer counts down to zero, the owner shall continue to transmit MCAP advertisements for the affected channel but shall adjust *expiration* in the second and subsequent advertisements to reflect the time remaining until the channel is to be deallocated. If another multicast source is using the same channel mapping, it shall commence transmitting MCAP advertisements for the channel mapping with refreshed *expiration* times that maintain the validity of the channel mapping. If the original owner observes an MCAP advertisement for the channel to be relinquished ~~within 30 seconds of the *expiration* time~~before its own timer has expired, it shall not deallocate the channel number.

Otherwise, if ~~30 seconds elapse after the most recent *expiration* time~~the owner's timer expires without the observation of a valid MCAP advertisement by another node, the owner of the channel number shall deallocate the channel as described below.

10.7 Expired channel mappings

A valid channel mapping expires when ~~CYCLE_TIME.*seconds* matches the *expiration* time in~~ *expiration* seconds have elapsed since the most recent MCAP advertisement. At this time, multicast recipients shall stop reception on the expired channel number(s). The owner of the channel mapping(s) shall ~~wait an additional 30 seconds before deallocating~~deallocate the channel number and ~~indicating~~indicate its availability in the isochronous resource manager's CHANNELS_AVAILABLE register.

~~If 30 seconds elapse subsequent to the *expiration* time and no MCAP advertisement is observed that refreshes the expired channel mapping, the owner of the channel mapping shall deallocate the channel number so long as the channel number is not in use by any other channel mapping.~~

11. SECURITY CONSIDERATIONS

This document specifies the use of an unsecured link layer, Serial Bus, for the transport of IPv4 datagrams. Serial Bus is vulnerable to denial of service attacks; it is also possible for devices to eavesdrop on data or present forged identities. Implementers who utilize Serial Bus for IPv4 should consider appropriate counter-measures within application or other layers.

12. ACKNOWLEDGEMENTS

This document represents work in progress by the IP/1394 Working Group. The editor wishes to acknowledge the contributions made by all the active participants, either on the reflector or at face-to-face meetings, which have advanced the technical content.

13. REFERENCES

[1] IEEE Std 1394-1995, Standard for a High Performance Serial Bus

[2] ISO/IEC 13213:1994, Control and Status Register (CSR) Architecture for Microcomputer Buses

[3] IEEE Project P1394a, Draft Standard for a High Performance Serial Bus (Supplement)

[4] IEEE Project P1394b, Draft Standard for a High Performance Serial Bus (Supplement)

14. EDITOR’S ADDRESS

Peter Johansson
Congruent Software, Inc.
3998 Whittle Avenue
Oakland, CA 94602

(510) 531-5472
(510) 531-2942 FAX
pjohansson@aol.com