Network Working Group
Request for Comments: 1220

F. Baker, Editor
ACC

April 1991

Point-to-Point Protocol Extensions for Bridging

1. Status of this Memo

This document defines an extension of the Internet Point-to-Point Protocol (PPP) described in RFC 1171, targeting the use of Point-to-Point lines for Remote Bridging. It is a product of the Point-to-Point Protocol Extensions Working Group of the Internet Engineering Task Force (IETF).

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

2. Historical Perspective

Two basic algorithms are ambient in the industry for Bridging of Local Area Networks. The more common algorithm is called "Transparent Bridging" and has been standardized for Extended LAN configurations by IEEE 802.1. IEEE 802.5 has proposed an alternative approach, called "Source Routing", and is in the process of standardizing that approach for IEEE 802.5 extended networks.

Although there is a subcommittee of IEEE 802.1 addressing remote bridging, neither standard directly defines Remote Bridging per se, as that would technically be beyond the IEEE 802 committee's charter. Both allow for it, however, modeling the line as an unspecified interface between half-bridges.

This document assumes that the devices at either end of a serial link

- have agreed to utilize the RFC 1171 line discipline in some form.
- may have agreed, by some other means, to exchange other protocols on the line interspersed with each other and with any bridged PDUs.
- may be willing to use the link as a vehicle for Remote Bridging.
- may have multiple point-to-point links that are configured in parallel to simulate a single line of higher speed or

reliability, but message sequence issues are solved by the transmitting end.

3. General Considerations

3.1. Link Quality Monitoring

It is strongly recommended that Point-to-Point Bridge Protocol implementations utilize Magic Number Loopback Detection and Link-Quality-Monitoring. This is because the 802.1 Spanning Tree protocol, which is integral to both Transparent Bridging and Source Routing (as standardized), is unidirectional during normal operation, with HELLO PDUs emanating from the Root System in the general direction of the leaves, without any reverse traffic except in response to network events.

3.2. Message Sequence

The multiple link case requires consideration of message sequentiality. The transmitting station must determine either that the protocol being bridged requires transmissions to arrive in the order of their original transmission, and enqueue all transmissions on a given conversation onto the same link to force order preservation, or that the protocol does NOT require transmissions to arrive in the order of their original transmission, and use that knowledge to optimize the utilization of the several links, enqueuing traffic to links to minimize delay.

In the absence of such a determination, the transmitting station must act as though all protocols require order preservation; many protocols designed primarily for use on a single LAN in fact do. A protocol could be described to maintain message sequentiality across multiple links, either by sequence numbering or by fragmentation and re-assembly, but this is neither elegant nor absolutely necessary.

3.3. Maximum Receive Unit Considerations

Please note that the negotiated MRU must be large enough to support the MAC Types that are negotiated for support, there being no fragmentation and re-assembly. Even Ethernet frames are larger than the default MRU of 1500 octets.

3.4. Separation of Spanning Tree Domains

It is conceivable that a network manager might wish to inhibit the exchange of BPDUs on a link in order to logically divide two regions into separate Spanning Trees with different Roots (and potentially different Spanning Tree implementations or algorithms). In order to

do that, he must configure both ends to not exchange BPDUs on a link. For the sake of robustness, a bridge which is so configured must silently discard the BPDU of its neighbor, should it receive one.

4. IEEE 802.1 Transparent Bridging

4.1. Overview of IEEE 802.1 Transparent Bridging

As a favor to the uninitiated, let us first describe Transparent Bridging. Essentially, the bridges in a network operate as isolated entities, largely unaware of each others' presence. A Transparent Bridge maintains a Forwarding Database consisting of

{address, interface}

records by saving the Source Address of each LAN transmission that it receives along with the interface identifier for the interface it was received on. It goes on to check whether the Destination Address is in the database, and if so, either discards the message (if the destination and source are located at the same interface) or forwards the message to the indicated interface. A message whose Destination Address is not found in the table is forwarded to all interfaces except the one it was received on; this describes Broadcast/Multicast behavior as well.

The obvious fly in the ointment is that redundant paths in the network cause indeterminate (nay, all too determinate) forwarding behavior to occur. To prevent this, a protocol called the IEEE 802.1(d) Spanning Tree Protocol is executed between the bridges to detect and logically remove redundant paths from the network.

One system is elected as the "Root", which periodically emits a message called a Bridge Hello Protocol Data Unit, or BPDU, heard by all of its neighboring bridges. Each of these modifies and passes the BPDU on to its neighbors, and they to theirs, until it arrives at the leaf LAN segments in the network (where it dies, having no further neighbors to pass it along) or until the message is stopped by a bridge which has a superior path to the "Root". In this latter case, the interface the BPDU was received on is ignored (i.e., it is placed in a Hot Standby status, no traffic is emitted onto it except the BPDU, and all traffic received from it is discarded) until a topology change forces a recalculation of the network.

4.2. IEEE 802.1 Remote Bridging Activity

There exist two basic sorts of bridges - ones that interconnect LANs directly, called Local Bridges, and ones that interconnect LANs via an intermediate medium such as a leased line, called Remote Bridges.

The Point-to-Point Protocol might be used by a Remote Bridge.

There is more than one proposal within the IEEE 802.1 Interworking Committee for modeling the Remote Bridge. In one model, the interconnecting serial link(s) are treated in the same way that a LAN is, having a standard IEEE 802.1 Link State; in another, the serial links operate in a mode quite different from the LANs that they interconnect. For the sake of simplicity of specification, the first model is adopted, although some of the good ideas from proponents of the second model are included or allowed for.

Therefore, given that transparent bridging is configured on a line or set of lines, the specifics of the link state with respect to the bridge is defined by IEEE 802.1(d). The Bridge Protocol Data Unit, or BPDU, is defined there, as well as the algorithms for its use.

It is assumed that, if a Point-to-Point Link neighbor receives IEEE 802.1 BPDUs without rejecting them with the RFC 1171 Protocol-Reject LCP PDU, Transparent Bridging is permitted on the link.

4.3. IEEE 802.5 Source Routing

The IEEE 802.5 Committee has defined a different approach to bridging for use on the Token Ring, called Source Routing. In this approach, the originating system has the responsibility of indicating what path that the message should follow. It does this, if the message is directed off the local ring, by including a variable length MAC header extension called the Routing Information Field, or RIF. The RIF consists of one 16 bit word of flags and parameters followed by zero or more ring-and-bridge identifiers. Each bridge en route determines from this "source route list" whether it should receive the message and how to forward it.

The algorithm for Source Routing requires the bridge to be able to identify any interface by its ring-and-bridge identifier, and to be able to identify any of its OTHER interfaces likewise. When a packet is received which has the Routing Information Field (RIF) present, a boolean in the RIF is inspected to determine whether the ring-and-bridge identifiers are to be inspected in "forward" or "reverse" sense. In a "forward" search, the bridge looks for the ring-and-bridge identifier of the interface the packet was received on, and forwards the packet toward the ring identified in the ring-and-bridge identifier that follows it. In a "reverse" search, the bridge looks for the ring-and-bridge identifier of the OTHER INTERFACE, and delivers the packet to the indicated interface if such is found.

The algorithms for handling multicasts ("Functional Addresses" and "Group Addresses") have been the subject of much discussion in 802.5,

and are likely to be the most troublesome for bridge implementations. Fortunately, they are beyond the scope of this document.

4.4. IEEE 802.5 Remote Bridging Activity

There is no Remote Bridge proposal in IEEE 802.5 at this time, although IBM ships a remote Source Routing Bridge. Simplicity would dictate that we choose the same model for IEEE 802.5 Source Routing that was selected for IEEE 802.1, but necessity requires a ring number for the line in some cases. We allow for both models.

Given that source routing is configured on a line or set of lines, the specifics of the link state with respect to the bridge is defined by the IEEE 802.5 Addendum on Source Routing. The requisite PDUs for calculating the spanning tree (used for assuring that each ring will receive at most one copy of a multicast) are defined there, as well as the algorithms for their use. MAC PDUs (Beacon, Ring Management, etc) are specific to the MAU technology and are not exchanged on the line.

4.5. Source Routing to Transparent Bridge Translation

IEEE 802 also has a subcommittee looking at the interoperation of Transparent Bridging and Source Routing. For the purposes of this standard, such a device is both a transparent and a source routing bridge, and will act on the line in both ways, just as it does on the LAN.

5. Traffic Services

Several services are provided for the benefit of different system types and user configurations. These include LAN Frame Checksum Preservation, LAN Frame Checksum Generation, Tinygram Compression, and the identification of closed sets of LANs.

5.1. LAN Frame Checksum Preservation

IEEE 802.1 stipulates that the Extended LAN must enjoy the same probability of undetected error that an individual LAN enjoys. Although there has been considerable debate concerning the algorithm, no other algorithm has been proposed than having the LAN Frame Checksum received by the ultimate receiver be the same value calculated by the original transmitter. Achieving this requires, of course, that the line protocols preserve the LAN Frame Checksum from end to end. The protocol is optimized towards this approach.

5.2. Traffic having no LAN Frame Checksum

The fact that the protocol is optimized towards LAN Frame Checksum preservation raises twin questions: "What is the approach to be used by systems which, for whatever reason, cannot easily support Frame Checksum preservation?" and "What is the approach to be used when the system originates a message, which therefore has no Frame Checksum precalculated?".

Surely, one approach would be to require stations to calculate the Frame Checksum in software if hardware support were unavailable; this would meet with profound dismay, and would raise serious questions of interpretation in a Bridge/Router.

However, stations which implement LAN Frame Checksum preservation must already solve this problem, as they do originate traffic. Therefore, the solution adopted is that messages which have no Frame Checksum are tagged and carried across the line.

When a system which does not implement LAN Frame Checksum preservation receives a frame having an embedded FCS, it converts it for its own use by removing the trailing four octets. When any system forwards a frame which contains no embedded FCS to a LAN, it forwards it in a way which causes the FCS to be calculated.

5.3. Tinygram Compression

An issue in remote Ethernet bridging is that the protocols that are most attractive to bridge are prone to problems on low speed (64 KBPS and below) lines. This can be partially alleviated by observing that the vendors defining these protocols often fill the PDU with octets of ZERO. Thus, an Ethernet or IEEE 802.3 PDU received from a line that is (1) smaller than the minimum PDU size, and (2) has a LAN Frame Checksum present, must be padded by inserting zeroes between the last four octets and the rest of the PDU before transmitting it on a LAN. These protocols are frequently used for interactive sessions, and therefore are frequently this small.

To prevent ambiguity, PDUs requiring padding are explicitly tagged. Compression is at the option of the transmitting station, and is probably performed only on low speed lines, perhaps under configuration control.

The pseudo-code in Figure 1 describes the algorithms.

5.4. LAN Identification

In some applications, it is useful to tag traffic by the user community it is a part of, and guarantee that it will be only emitted onto a LAN which is of the same community. The user community is defined by a LAN ID. Systems which choose to not implement this feature must assume that any frame received having a LAN ID is from a different community than theirs, and discard it.

Figure 1: Tinygram Compression Pseudo-Code

PPP Transmitter:

```
if (ZeroPadCompressionEnabled &&
    BridgedProtocolHeaderFormat == IEEE8023 &&
    PacketLength == Minimum8023PacketLength) {
 /*
  * Remove any continuous run of zero octets preceding,
  * but not including, the LAN FCS, but not extending
  * into the MAC header.
  */
    Set (ZeroCompressionFlag);             /* Signal receiver */
    if (is_Set (LAN_FCS_Present)) {
        FCS = TrailingOctets (PDU, 4);     /* Store FCS */
        RemoveTrailingOctets (PDU, 4);     /* Remove FCS */
        while (PacketLength > 14 &&        /* Stop at MAC header */
               TrailingOctet (PDU) == 0)  /*  or last non-zero octet */
            RemoveTrailingOctets (PDU, 1);/* Remove zero octet */
        Appendbuf (PDU, 4, FCS);           /* Restore FCS */
    }
    else {
        while (PacketLength > 14 &&        /* Stop at MAC header */
               TrailingOctet (PDU) == 0)  /*  or last zero octet */
            RemoveTrailingOctets (PDU, 1);/* Remove zero octet */
    }
}
```

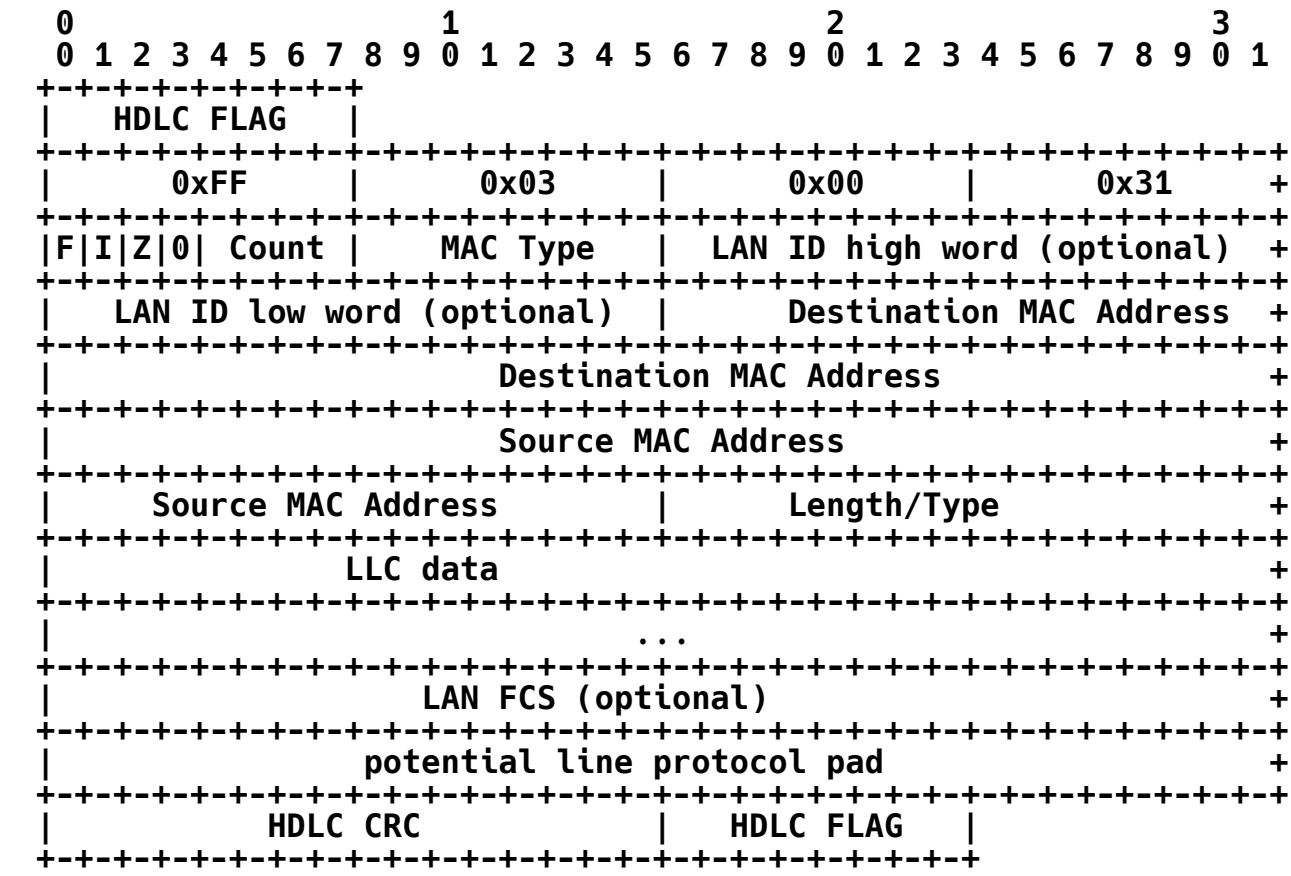
PPP Receiver:

```
if (ZeroCompressionFlag) {                 /* Flag set in header? */
 /* Restoring packet to minimum 802.3 length */
    Clear (ZeroCompressionFlag);
    if (is_Set (LAN_FCS_Present)) {
        FCS = TrailingOctets (PDU, 4);    /* Store FCS */
        RemoveTrailingOctets (PDU, 4);    /* Remove FCS */
        Appendbuf (PDU, 60 - PacketLength, zeroes);/* Add zeroes */
        Appendbuf (PDU, 4, FCS);          /* Restore FCS */
    }
    else {
        Appendbuf (PDU, 60 - PacketLength, zeroes);/* Add zeroes */
    }
}
```

6. Protocol Data Unit Formats

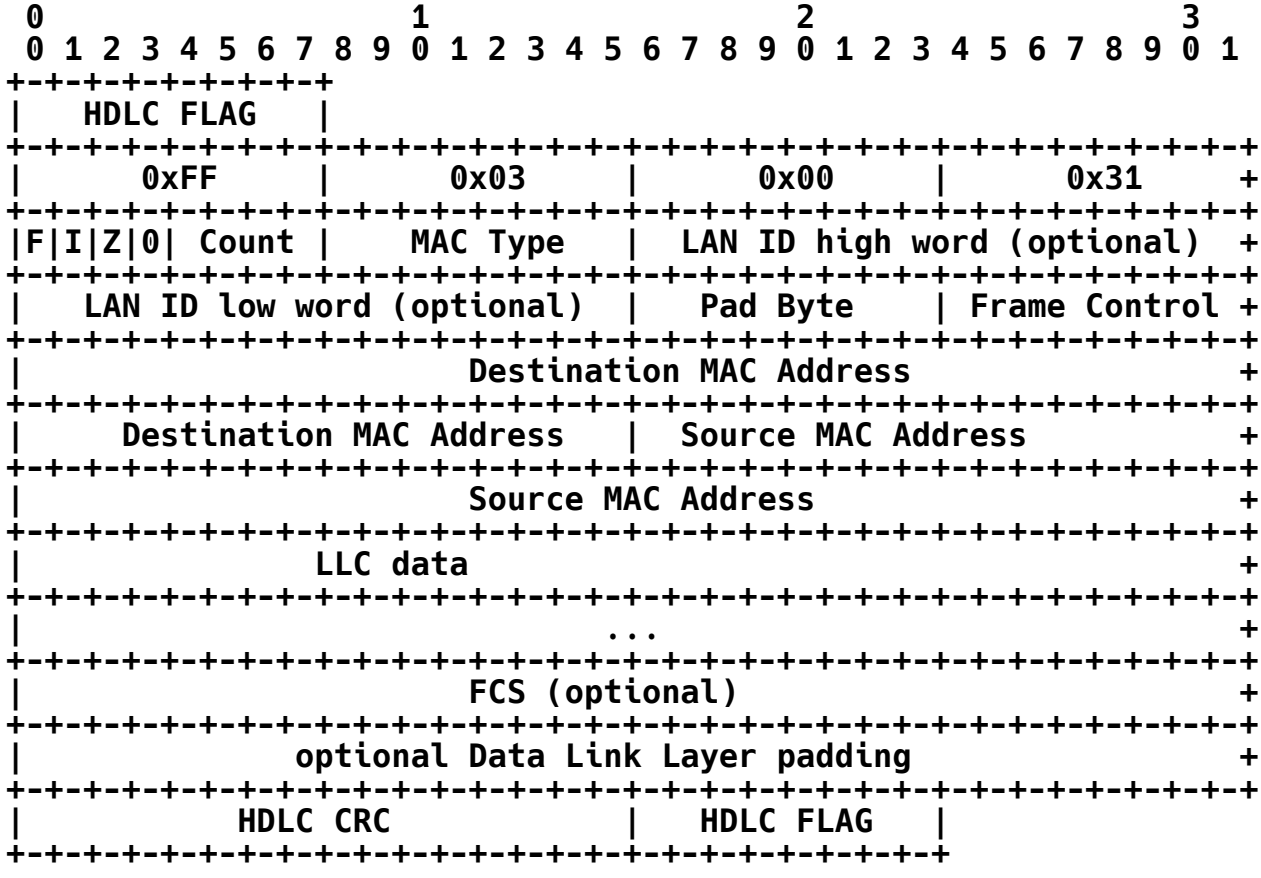
6.1. Common LAN Traffic

Figure 2: 802.3 Frame format



For Bridging LAN traffic, the format of the frame on the line is as shown in Figures 2 or 3. This conforms to RFC 1171 section 3.1 "Frame Format". It also allows for RFC 1172 [2] negotiation of Protocol Field Compression and Address and Control Field Compression. It is recommended that devices which use controllers that require even memory addresses negotiate to NOT USE Protocol Field Compression on other than low speed links.

Figure 3: 802.4/802.5/FDDI Frame format



The fields of this message are as follows:

Address Field and Control Field:
As defined by RFC 1171

Protocol Field:
0x0031

Flags:
bits 0-3: length of the line protocol pad field.
bit 4: Reserved, Set to Zero
bit 5: Set if IEEE 802.3 Pad must be zero filled to minimum size
bit 6: Set if the LAN ID Field is present
bit 7: Set if the LAN FCS Field is present

The "number of trailing "pad" octets is a deference to the fact that any point-to-point frame may have padding at the end. This

number tells the receiving system how many octets to strip off the end.

MAC Type:
0: Reserved
1: IEEE 802.3/Ethernet
2: IEEE 802.4
3: IEEE 802.5
4: FDDI
other: Assigned by the Internet Assigned Numbers Authority

LAN ID:
This optional 32 bit field identifies the Community of LANs which may be interested to receive this frame, as described in section 5.4. If the LAN ID flag is not set, then this field is not present, and the PDU is four octets shorter.

Frame Control:
On 802.4, 802.5, and FDDI LANs, there are a few octets preceding the Destination MAC Address, one of which is protected by the FCS. Since the MAC Type field defines the bit ordering, these are sent in MAC order. A pad octet is present to avoid odd machine address boundary problems.

Destination MAC Address:
As defined by the IEEE. Since the MAC Type field defines the bit ordering, this is sent in MAC order.

Source MAC Address:
As defined by the IEEE. Since the MAC Type field defines the bit ordering, this is sent in MAC order.

LLC data:
This is the remainder of the MAC frame. This is that portion of the frame which is (or would be were it present) protected by the LAN FCS; for example, the 802.5 Access Control field, and Status Trailer are not meaningful to transmit to another ring, and are omitted.

LAN Frame Checksum:
If present, this is the LAN FCS which was calculated by (or which appears to have been calculated by) the originating station. If the FCS Present flag is not set, then this field is not present, and the PDU is four octets shorter.

Optional Data Link Layer Padding
RFC 1171 specifies that an arbitrary pad can be added after the data intended for transmission. The "Count" portion of the flag

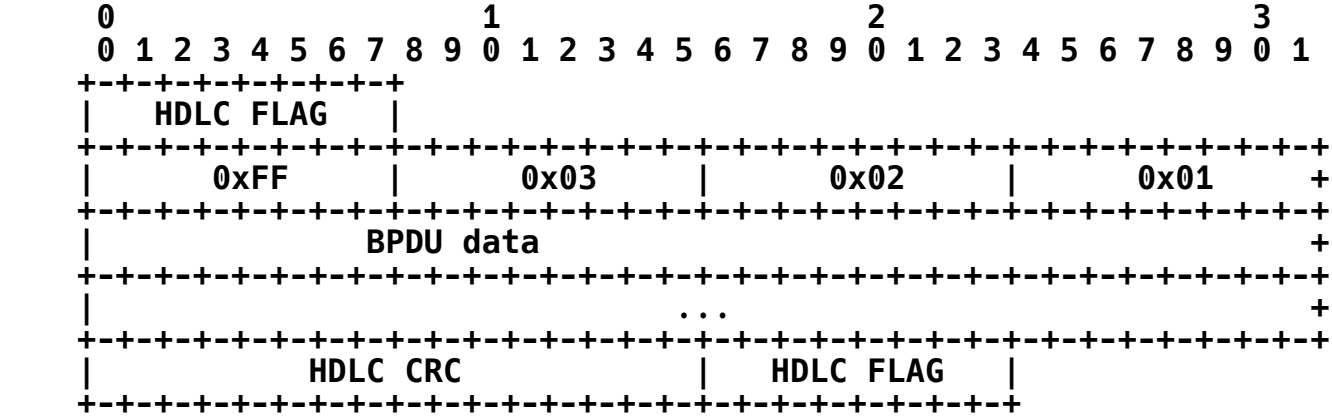
field contains the length of this pad, which may not exceed 15 octets.

CRC-CCITT
Mentioned primarily for clarity. The CRC used on the PPP link is separate from and unrelated to the LAN FCS.

6.2. IEEE 802.1 Bridge

This is the BPDU as defined by IEEE 802.1(d), without any MAC or 802.2 LLC header (these being functionally equivalent to the Address, Control, and Protocol Fields). The LAN Pad and Frame Checksum fields are likewise superfluous and absent. The Address and Control Fields are optional, subject to the Address and Control Field Compression negotiation.

Figure 4: Bridge "Hello" PDU



The fields of this message are as follows:

Address Field and Control Field:
As defined by RFC 1171

Protocol Field:
0x0201

MAC Frame:
802.1(d) BPDU

6.3. IEEE 802 Network Control Protocol

The Bridge Network Control Protocol is responsible for configuring, enabling, and disabling the bridges on both ends of the point-to-point link. As with the Link Control Protocol, this is accomplished through an exchange of packets. BNCP packets may not be exchanged until LCP has reached the network-layer Protocol Configuration Negotiation phase. Likewise, LAN traffic may not be exchanged until BNCP has first opened the connection.

The Bridge Network Control Protocol is exactly the same as the Point-to-Point Link Control Protocol with the following exceptions:

Data Link Layer Protocol Field
: Exactly one Bridge Network Control Protocol packet is encapsulated in the Information field of PPP Data Link Layer frames where the Protocol field indicates type hex 8031 (BNCP).

Code field
: Only Codes 1 through 7 (Configure-Request, Configure-Ack, Configure-Nak, Configure-Reject, Terminate-Request, Terminate-Ack and Code-Reject) are used. Other Codes should be treated as unrecognized and should result in Code-Rejects.

Timeouts
: BNCP packets may not be exchanged until the Link Control Protocol has reached the network-layer Protocol Configuration Negotiation phase. An implementation should be prepared to wait for Link Quality testing to finish before timing out waiting for Configure-Ack or other response.

Configuration Option Types
: The Bridge Network Control Protocol has a separate set of Configuration Options. These permit the negotiation of the following items:

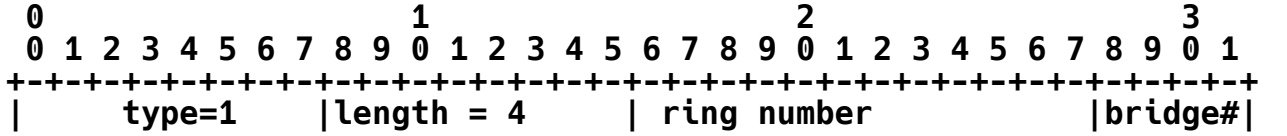
  - MAC Types supported
  - Tinygram Compression support
  - LAN Identification support
  - Ring and Bridge Identification

6.4. IEEE 802.5 Remote Ring Identification Option

Since the Remote Bridges are modeled as normal Bridges with a strange internal interface, each bridge needs to know the ring/bridge numbers of the bridges it is adjacent to. This is the subject of a Link Negotiation. The exchange of ring-and-bridge identifiers is done using this option on the Network Control Protocol.

The Token Ring Ring-and-Bridge Identifier, and its use, is specified by the IEEE 802.5 Addendum on Source Routing. It identifies the ring that the interface is attached to by its configured ring number, and itself by bridge number on the ring.

Figure 5: Remote Ring Identification Option



Type 1 = IEEE 802.5 Source Routing Ring/Bridge Identifier

Length
4 Octets

Ring Number
A 12 bit number identifying the token ring, as defined in the IEEE 802.5 Source Routing Specification.

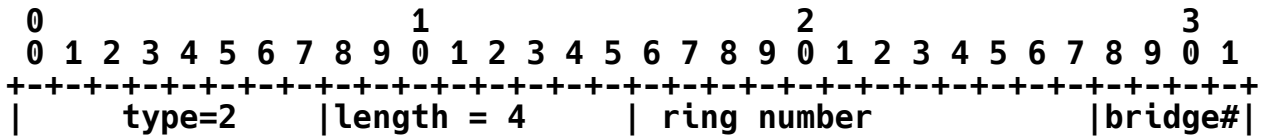
Bridge Number
A 4 bit number identifying the bridge on the token ring, as defined in the IEEE 802.5 Source Routing Specification.

6.5. IEEE 802.5 Line Identification Option

This option permits the systems to treat the line as a visible "Token Ring", in accordance with the Source Routing algorithm. The bridges exchange ring-and-bridge identifiers using this option on the Network Control Protocol. The configured ring numbers must be identical in normal operation.

The Token Ring Ring-and-Bridge Identifier, and its use, is specified by the IEEE 802.5 Addendum on Source Routing. It identifies the ring that the interface is attached to by its configured ring number, and itself by bridge number on the ring.

Figure 6: Line Identification Option



Type 2 = IEEE 802.5 Line "Ring/Bridge" Identifier

Length
4 Octets

Ring Number
A 12 bit number identifying the line, as defined in the IEEE 802.5 Source Routing Specification.

Bridge Number
A 4 bit number identifying the bridge on the token ring, as defined in the IEEE 802.5 Source Routing Specification.

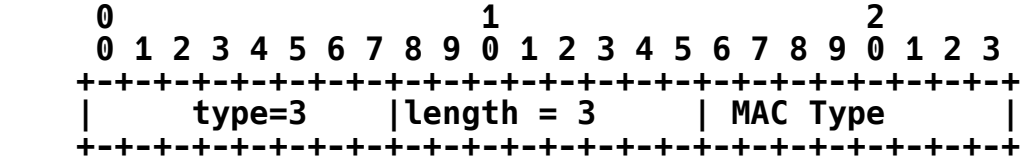
6.6. MAC Type Support Selection

The MAC Type Selection Option is provided to permit nodes to advertise what sort of traffic they are prepared to convey. A device negotiating a 1600 octet MRU, for example, may not be willing to support 802.5 (although it might, with certain changes necessary in the RIFs it passes, and given that the hosts it supports implement the 802.5 Maximum Frame Size correctly), and is definitely not prepared to support 802.4 or FDDI.

A system which does not announce the MAC Types that it supports may be assumed to support all MAC Types; it will discard those that it does not understand. A system which chooses to announce MAC Types is advising its neighbor that all unspecified MAC Types will be discarded. Announcement of multiple MAC Types is accomplished by placing multiple options in the Configure Request.

The Rejection of a MAC Type Announcement (in a Configure-Reject) is essentially a statement that traffic appropriate to the MAC Type, if encountered, will be forwarded on the link even though the receiving system has indicated it will discard it.

Figure 7: MAC Type Selection Option



Type 3 = MAC Type Selector

Length
3 Octets

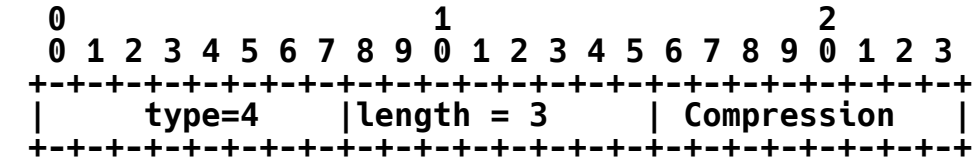
MAC Type Selector
One of the values of the PDU's MAC Type Field that this system is prepared to receive and service.

6.7. Tinygram Compression

Not all systems are prepared to make modifications to messages in transit; on high speed lines, it is probably not worth the effort. This option permits the system to negotiate compression.

Consistent with the behavior of other compression options in the Internet Point-to-Point set of protocols, no negotiation implies no compression. The systems need not agree on the setting of this parameter; one may be willing to decompress and the other not. A system which does not negotiate, or negotiates this option to be disabled, should never receive a compressed packet, however.

Figure 8: Tinygram Compression Option



Type 4 = Tinygram Compression Support Option

Length
3 Octets

Compression Enable/Disable
If the value is 1, Tinygram Compression is enabled. If the value is 2, Tinygram Compression is disabled, and no decompression will occur.

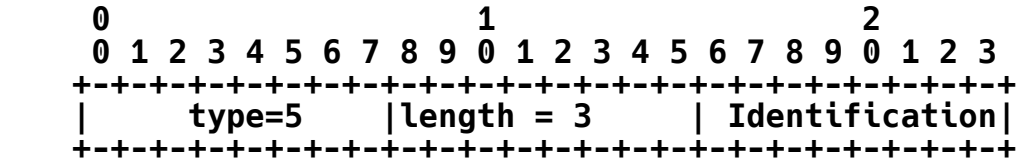
6.8. LAN Identification Support

Not all systems are prepared to make use of the LAN Identification field. This option enables the systems to negotiate its use.

The parameter is advisory; if the value is "enabled", then there may exist labeled LANs beyond the system, and the system is prepared to service traffic to it. if the value is "disabled", then there are no labeled LANs beyond the system, and all such traffic will by definition be dropped. Therefore, a system which is advised that his peer does not service LAN Identifications need not forward such traffic on the link.

The default value is that LAN Identification disabled.

Figure 9: LAN Identification Option



Type 5 = LAN Identification Support Option

Length
3 Octets

Identification Enable/Disable
If the value is 1, LAN Identification is enabled. If the value is 2, LAN Identification is disabled.

7. Acknowledgements

This document is a product of the Point-to-Point Protocol Extensions Working Group. Special thanks go to Steve Senum of Network Systems, Dino Farinacci of 3COM, and Rick Szmauz of Digital Equipment Corporation.

8. Bibliography

[1] Perkins, D., "The Point-to-Point Protocol for the Transmission of Multi-Protocol Datagrams Over Point-to-Point Links", RFC 1171, CMU, July 1990.

[2] Hobby R., and D. Perkins, "The Point-to-Point Protocol (PPP) Initial Configuration Options", RFC 1172, CMU, UC Davis, July 1990.

[3] IEEE Draft Standard P802.1d/D9 MAC Bridges, Institute of Electrical and Electronic Engineers. Also Published as ISO DIS 10038, July 1989.

[4] IEEE Draft Standard P802.5d/D13 Draft Addendum to ANSI/IEEE Std 802.5-1988 Token Ring MAC and PHY Specification Enhancement for Multiple-Ring Networks, Institute of Electrical and Electronic Engineers, May 1989.

9. Security Considerations

Security issues are not discussed in this memo.

10. Author's Address

Fred Baker
Advanced Computer Communications
720 Santa Barbara Street
Santa Barbara, CA 93101

Phone: (805) 963-9431

EMail: fbaker@ACC.COM
Or send comments to: ietf-ppp@ucdavis.edu