3. IPX WAN Version 2

This section describes how the Novell® IPX™ protocol and routing protocols for IPX operate over various WAN media.

As its name implies, the IPX WAN version 2 (IW2) protocol specified in this section is the successor to the IPX WAN protocol specified in Reference [All92]. This document supersedes that specification, and is backward-compatible with it.

Reference [All92] specifies how to operate IPX, RIP, and SAP over the relevant media. IW2 adds what is needed to support several additional routing protocols. Part of IW2 is to negotiate which routing protocol is used over a particular link.

Reference [All92] specifies how to operate over PPP and X.25. This specification adds two media types: Frame Relay and IP Relay.

3.1. Support of Several Routing Protocols

The routing protocols covered by this specification are as follows:

- RIP, as specified in Reference [Nov92].
- Unnumbered RIP. This operates the same way as RIP, but the link over which IW2 is being run is not assigned an IPX network number.
- On-demand, static routing. This method serves two purposes related to economizing network cost.
- Client-router connection. This is the method for an individual client personal computer or workstation to connect remotely to an internetwork.
- The NLSP™ protocol, as specified in sections 4 through 8 of this document.

Unnumbered RIP is motivated by administrative convenience. In an internetwork in which there are many WAN circuits, it is inconvenient to set aside IPX network numbers for them all, and to make sure that there is no duplication.

On-demand, static routing applies to transient data-link connections (like dial access and X.25 switched virtual circuits) as well as persistent data-link connections (like dedicated point-to-point links and X.25 permanent virtual circuits). This method serves two purposes. The “on-demand” part means that a router initiates communication to a destination only when there is data to be forwarded to that destination. Here, *initiating communication* includes making a call (where necessary) and conducting the IW2 exchange. A transient connection is closed after a period of inactivity. The *static routing* part means that no routing information is sent over the link—no RIP, no SAP, and no NLSP. Instead, the router at each end is configured with the routes and services accessible through the link.

Client-router connection is fundamentally different from the other routing types described in this chapter. Its operation is asymmetric. It does not connect two routers—the client system performs no routing functions. No routing protocol is used on the link, except when the client makes an explicit RIP or SAP information request. The router assigns a six-byte node number to the client, a number that plays the role of the IEEE MAC address. This way, a router can share a network number among many clients.

Of the router-to-router protocols, only RIP involves assigning a network number to the link running IW2. RIP is called a *numbered* routing protocol type.

NLSP, Unnumbered RIP, and On-demand static routing do not use a network number for the WAN link. They are called *unnumbered* routing protocol types.

Additional routing protocols of numbered and unnumbered types might be specified in the future. The categorization of routing protocols into numbered or unnumbered is an important distinction, which must be considered when planning new designs.

3.2. Implementing Media-Dependent Functions

IW2 strives to treat all media the same way. Different treatment of different media is limited to the minimum needed for effective communication.

Link establishment and termination are procedures that vary from one medium to the next. Other details (for example, encapsulation) are also media-dependent. Recall that the NLSP and RIP packets ride within IPX packets, as do the IPX WAN and IW2 protocols, so no additional information about their encapsulation need be specified, beyond that of IPX.

Operation over individual media types is described one type at a time.

3.2.1. Operation over PPP

IPX and its routing protocols use PPP (Reference [Sim92]) when operating over point-to-point synchronous and asynchronous networks. The normal PPP frame encapsulation applies.

With PPP, link establishment means the IPXCP (IPX Control Protocol, Reference [Sim92a]) reaches the “Open” state. The IW2 protocol must not begin until the IPXCP reaches the “Open” state. (Novell’s implementation of IPXCP rejects all options.) If IPXCP negotiates options conflicting with those negotiated later with IW2, the IW2 negotiation takes precedence.

PPP allows either side of a connection to stop forwarding IPX if one end sends an IPXCP or an LCP Terminate-Request. When a router detects this, it immediately reflects the lost connectivity in its routing information database instead of aging it out naturally.

3.2.2. Operation over X.25 Switched Virtual Circuits

With X.25, link establishment means successfully opening an X.25 switched virtual circuit (SVC). As specified in Reference [Mal92], the protocol identifier 0x800000008137 is used in the X.25 Call User Data field of the Call Request frame, and indicates that the SVC is devoted to IPX.

Each IPX packet is encapsulated directly in X.25 data frame sequences without additional framing.

Either side of the SVC can close it, thereby tearing down the IPX link. When a router detects this, it immediately reflects the lost connectivity in its routing information database instead of aging it out naturally.

3.2.3. Operation over X.25 Permanent Virtual Circuits

The nature of permanent virtual circuits (PVCs) is that no call request is made. When the router is informed that X.25 Layer 2 is up, the router assumes that link establishment is complete.

Each IPX packet is encapsulated in an X.25 data frame sequence without additional framing. Local configuration determines which PVCs are devoted to IPX.

If a router receives a Layer 2 error condition (X.25 Restart, for example), it should reflect lost connectivity for the PVCs in its routing information database and perform the necessary steps to obtain a full IPX connection again.

3.2.4. Operation over Frame Relay

To determine when a PVC has become active or inactive, the router interacts periodically with either a private Frame Relay switch or a public Frame Relay network. The method used depends on the switch or service provider. Some support Section 6 of Reference [DEC90]; others, Annex D of Reference [ANS91]. NLSP supports both these specifications.

When a router is restarted, IW2 exchanges over active Frame Relay PVCs (that is, PVCs that have remained active before and after restart) can begin immediately.

When a router detects that a Frame Relay PVC has switched from an inactive state to an active state, link establishment is considered complete and IW2 exchange over this newly activated PVC begins.

When an active PVC becomes inactive, the router reflects the lost connectivity in its routing information database.

Framing of IPX packets is specified in Reference [Bra92], Section 6. The NLPID (Network layer Protocol ID) used for IPX is 0x800000008137.

3.2.5. Operation over IP Relay

IP Relay is a streamlined method to convey IPX traffic through an IP internetwork. IP Relay operates like a collection of PVCs with two end points, although the traffic travels as UDP/IP datagrams (References [Pos80] and [Pos81]). Each end point of each PVC is an IP address. Each router implementing the IP Relay feature is configured to know the opposite end of each PVC that terminates at its IP address.

When an IPX router is first activated and notices that IP is active on its system (or if IP becomes active while NLSP is already active), all the configured PVCs are considered to be established data-link connections. IW2 exchanges can begin immediately.

A router can receive an ICMP packet (Reference [Pos81a]) indicating nondelivery of an IP Relay datagram. The IP Relay function ignores any such ICMP packets. In particular, ICMP packets do not cause the router to consider a PVC to be closed.

If the IPX router detects that its IP protocol operation has been deactivated, it considers that the IP Relay PVCs have all been cleared and it reflects the lost connectivity in its routing information database.

Each IPX packet comprises the data portion of a UDP packet, with the first byte of the IPX header immediately following the UDP header. The UDP port used is 213 decimal.

Any IP implementation supporting this encapsulation technique must be capable of sending and receiving 1472-byte UDP packets (before any IP fragmentation that might be necessary). The 1472 bytes cover the data portion of the UDP packet, excluding the UDP and IP headers.

A system can optionally implement source validation. When source validation is in force, any packet arriving at IP Relay's UDP port from an IP address not in the receiver's PVC list is discarded.

3.2.6. Operation over other WAN media

Additional WAN media will be added here as specifications are developed.

3.3. Outline of the Stages of IW2 Operation

After the underlying data-link connection is established as described in the preceding media-dependent description, the IW2 protocol is activated to exchange identities and determine certain operational characteristics of the link.

There are three steps of IW2 operation:

- Negotiating the master/slave role and choice of routing protocol. The master/slave roles persist for the IW2 interactions only. The master keeps the rest of IW2 proceeding in an orderly way by initiating the request/reponse exchanges.
- Exchange of packets to determine certain link characteristics empirically. The details of this step depend on the routing protocol negotiated.
- Final exchange of the routers' configuration information

After these steps are concluded, transmission of IPX routing packets begins—using the routing protocol negotiated—as well as transmission of IPX data traffic.

3.4. IW2 Packets and their Usage

IW2 involves exchanges of IPX WAN packets, whose structure is defined on page 3-19. There are nine types of IW2 packets:

- Timer Request—a 576-byte packet sent from both ends of the WAN link to measure packet turnaround time
- Timer Response—response to a Timer Request; this packet is also 576 bytes long
- Throughput Request—a pair of n-byte packets sent back-to-back from the Master to measure the media throughput
- Throughput Response—acknowledgment of the Throughput Request packet pair
- Delay Request—sent by the Master to measure media delay
- Delay Response—response to the Delay Request packet; it has the same size as the request
- Information Request—sent by the Master at the conclusion of the IW2 negotiation to exchange delay and throughput measurements, a network number, and the router names
- Information Response—response packet to the Information Request
- NAK—negative acknowledgment; sent back to a party that sent an illegal packet

Certain packets apply to certain routing types but not others. Details follow.

If the router's software environment includes a priority system, IW2 packets should be processed with higher priority than all other IPX packets.

3.5. Steps of the Initial Negotiation for Router-Router Operation

The several router-router protocols all begin operation the same way, as described in this section. The client-router protocol is a variation on the theme. It is discussed later, starting on page 3-13.

The first exchange of packets decides the master/slave roles and the routing protocol to be used on the link. It also gauges the link delay, by measuring the turnaround time for a moderately sized packet exchange. The initial negotiation is the same for all routing types. It is illustrated in Figure 3-1.

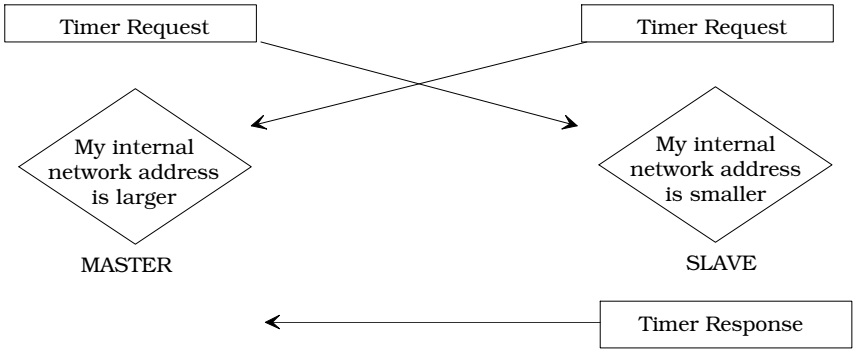

Figure 3-1: Initial Negotiation

After link establishment, both sides of the link send Timer Request packets and start a timer waiting for a Timer Response. See page 3-19 for a specification of packet contents. These Timer Request packets are sent every 20 seconds until a response is received or retry counter is exhausted. Usually, the retry counter is set to 16. If the retry counter is exhausted, the data-link connection is closed.

In composing the Timer Request packet, the router takes into consideration

- Which types of routing protocols it supports.
- Whether it is prepared to assign an IPX network number to the link.
- Whether it can support IPX header compression.
- For client-router connection, whether there is a need to specify the node number on a reconnect.

For each routing protocol supported, place the corresponding option in the Timer Request packet. They appear in the originator's order of preference, with the most preferred first.

For each compression method supported, place the corresponding option in the Timer Request packet.

If the router is prepared to assign an IPX network number to the link, it sends its internalNetworkNumber in the WNode ID field, and omits the Extended Node ID option. On the other hand, if the router is **not** prepared to assign an IPX network number to the link, it sends

zero in the WNode ID field. In the latter case, it also includes its internalNetworkNumber in the Extended Node ID option.

On receiving a Timer Request packet, a router determines its role—Master or Slave—for the remainder of the IW2 exchanges. The Master role does not denote special privilege. It merely means that the router is the requester in the ensuing request/response exchanges. The decision is made as follows:

a) If the WNode ID field is zero in both the sent and the received Timer Requests

   i) If both Timer Requests include an Extended Node ID, the router with the higher numeric value of this field is the Master. If the two Extended Node ID fields are equal, a configuration error has occurred. After reporting the error, the router issues a disconnect on the underlying data-link connection. Manual intervention is needed to correct the error condition.

   ii) If only one Timer Request includes the Extended Node ID, the router sending it is the Master.

   iii) If neither Timer Request includes the Extended Node ID, a connection cannot be established. The data-link circuit is cleared by the system that initiated it.

b) If either the sent or received Timer Request (or both) contains a nonzero WNode ID field, the router with the higher WNode ID is the Master.

c) If the two WNode ID fields are equal and nonzero, a configuration error has occurred. After reporting the error, the router issues a disconnect on the underlying data-link connection. Manual intervention is needed to correct the error condition.

The numeric comparisons are done by considering each byte as an unsigned integer, and the first byte as most significant. Note that in Figure 3-1 the decisions in the diamond-shaped symbols are simplified; the actual decision takes into account the two fields: WNode ID and Extended Node ID.

The Slave responds to the Timer Request with a Timer Response. To do so, it determines the routing protocol to be used on the link according to the following rule:

```
For each routing Routing Type in the received Timer Request,
  in order of appearance

   BEGIN      -- LOOP

   If the Routing Type is one supported by this router

         BEGIN      -- TEST

         If the Routing Type is an unnumbered type,

               choose it and exit.

         If the received WNode ID is nonzero,

               choose this Routing Type and exit.

         END        -- TEST

   END        -- LOOP
```

Note: It is permitted for a router to support a numbered routing type, but not be able to assign the network number. In this case, that routing type can be selected only if the other router supports it and can assign the

network number. The rules for determining the Master/Slave role and routing protocol ensure the following:

- Routers implementing IW2 are interoperable with those implementing Reference [All92].
- If only one router is prepared to assign a network number, it is the Master (the role that does the assignment), so that if a numbered routing type is chosen, the Master has a network number to assign.

The remainder of IW2 depends on the routing type selected. A separate subsection follows for each alternative. The WNode ID value in all subsequent IW2 packets is the sender's actual internalNetworkNumber, not zero.

3.6. Remaining Steps for the Numbered RIP Routing Type

When the Timer Request/Response exchange concludes with RIP as the chosen routing type for the link, the Information Request/Reply is the next (and only) remaining step in IW2 (See Figure 3-2). This reiterates the specification in [All92]. Once the negotiation is completed, the link can be used for RIP, SAP, and IPX data traffic, but not for NLSP packets.

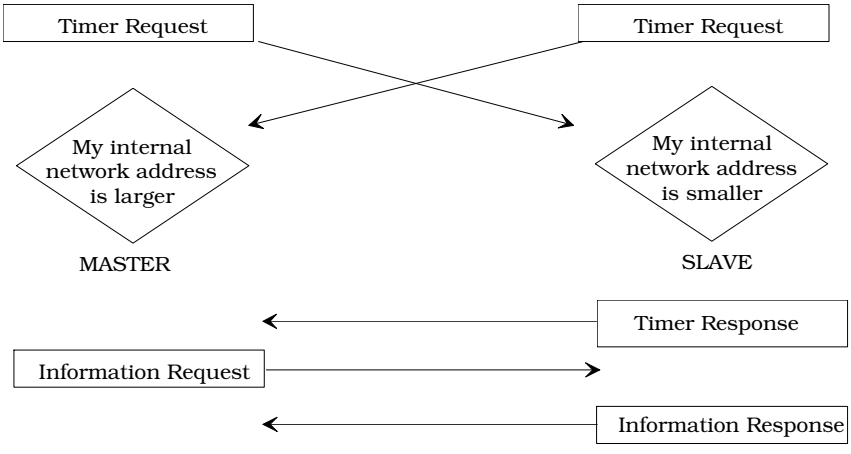

Figure 3-2: IW2 Exchanges for the RIP Routing Type

The following option subfield in the Information Request/Reponse exchange contain operational information used in RIP.

- RIP/SAP Info Exchange option—WAN Link Delay subfield. This link delay is used in the RIP information propagated to the router's other links. The following algorithm determines the IPX WAN link delay, based on the Timer Request/Response exchange. Here, `start_time` is the time that the Timer Request is sent, and `end_time` is the time that the Timer Response is received.

```
link_delay ← end_time - start_time                -- 1/18th second

link_delay ← MAX ( link_delay, 1 )     -- Ensure link delay is at least 1

link_delay ← 6 × link_delay
```

Note: Workstations use the link delay to timeout sessions. The factor of six is in anticipation of queuing delays for the WAN link that can readily impact round-trip time when data traffic starts flowing. It is a biasing value to prevent timeout of multiple workstation sessions sharing the link.

```
link_delay ← 55 × link_delay  -- Convert link delay to milliseconds
```

The Link Delay is used as the network transport time when advertised in the IPX RIP packet. For a consistent network, a common link delay is required at both ends of the link and is calculated by the link Master.

Note: IPX WAN in its initial version does not retry the Information Request exchange. If no Information Response is received after the first Information Request, the data-link connection is cleared and starts again.

Only a Timer Response whose sequence number matches the last sent Timer Request sequence number is acted upon. This rule applies to all the routing types.

Once the IW2 exchanges are complete, RIP operates over the link as specified in Reference [Nov92].

3.7. Remaining Steps for the Unnumbered RIP Routing Type

With RIP and IPX WAN, an IPX network number is assigned to each (virtual) circuit by operation of the protocol. With Unnumbered RIP, there are no network numbers used for circuits. This is an advantage, because it is easier to administer. The Unnumbered RIP routing type exploits this administrative ease, while still using RIP and SAP as the routing protocols for the circuit.

The IW2 packets exchanged for Unnumbered RIP are the same as for RIP. If the Timer Response selects Unnumbered RIP, that mode is in effect on the link. The next exchange is the Information Request/Response. The Common Network Number field in the RIP/SAP Info Exchange field is zero, as a placeholder.

Once the IW2 exchanges are complete, RIP operates over the link as specified in Reference [Nov92], with the following exceptions. The Network Numbers in the source/destination addresses of RIP/SAP packets traversing this link are set to zero. These RIP/SAP packets are originated by the routers at either end. Forwarded packets, by contrast, continue to have the originator's and the final destination's network numbers. Because the two ends of the link are routers, the fact that this link has no network number is inconsequential. If one end were an end node, there would have to be a way for systems far away in the internetwork to address it, requiring a network number. If services are accessible on the same system as the router, or if there are software elements that act as clients, they are attached to the system's internal network number.

3.8. Remaining Steps for On-Demand, Static Routing

On-demand, static routing serves two purposes. The *on-demand* part means that a router initiates communication to a destination only when there is data to be forwarded to that destination. *Initiating communication* includes making a data-link call (where necessary) and performing the IW2 exchange. A transient connection is closed after a period of inactivity.

The *static routing* part means that no routing information is sent over the link—no RIP, no SAP, and no NLSP. Instead, the router at each end is configured with the routes and services accessible through the link.

The IW2 exchange for this routing type is a variation of Unnumbered RIP's. The differences are as follows:

In the Timer Request, the calling router offers ***only*** the on-demand static routing Routing Type. This implies that on-demand static routing is the only routing type that can be negotiated successfully on the link. If the Slave detects a mismatch in routing types, it disconnects the link.

With on-demand, static routing, the routers must be able to identify one another using data-link information. For example, with X.25 switched virtual circuits, the calling DTE address can be used; with IP Relay, the remote IP address can be used; with transient point-to-point links, PPP authentication can be used. With a persistent data-link connection, the physical port identifier or a PVC identifier can be used. The choice of identifier is an implementation decision. Whatever value the called router uses is called a *Remote System Identifier*, or *RSI*.

Note: In Novell implementations of on-demand static routing over PPP, the authentication features of PPP identify the peer router. For interoperability with Novell routers in this context, systems must implement PPP authentication.

A router implementing on-demand, static routing must maintain a database of RSIs, and lists describing the network numbers and services reachable through each RSI. These lists determine the reachability information it transmits to other routers in a routing area. Other routers treat each on-demand, static routing link as though it were permanently available.

For a persistent data link, there might be no discernable link establishment event. For such media, arrival of a Timer Request plays the role of detecting link establishment.

As with Unnumbered RIP, there is no network number assigned to the link. NLSP packets are not sent on the link. Moreover, periodic RIP and SAP packets are not sent on the link. However, a router must respond to RIP and SAP queries received on the link.

3.9. Remaining Steps for the NLSP Routing Type

NLSP employs an exchange of IW2 packets lengthier than those of other routing types.

3.9.1. Normal IW2 Exchanges for NLSP

Figure 3-3 shows the complete IW2 exchange in preparation for using NLSP on the link for the normal case of no error conditions.

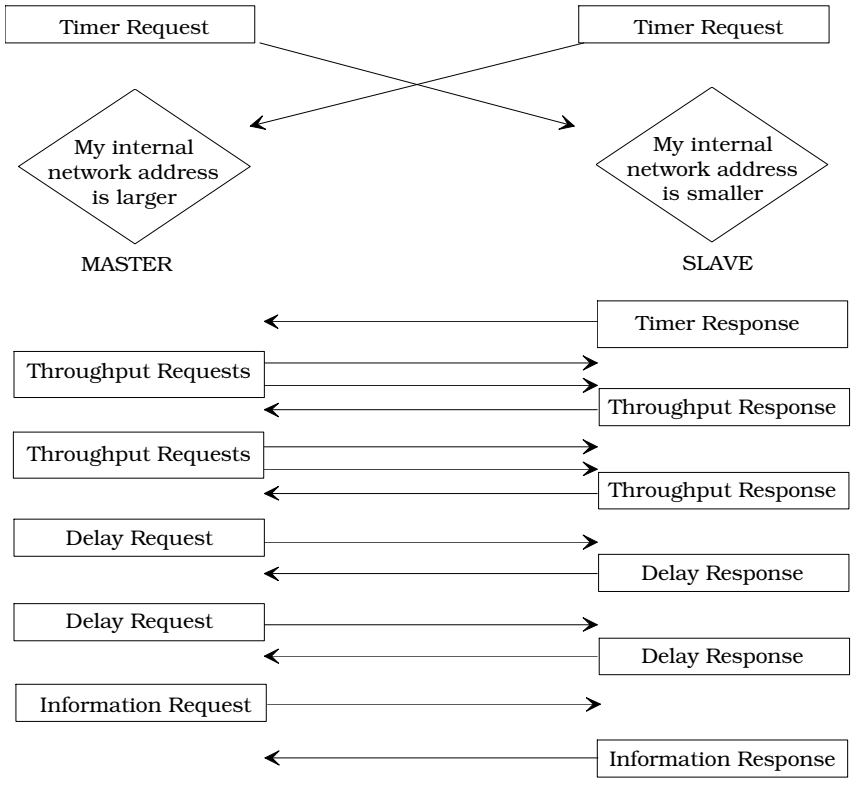

Figure 3-3: IW2 Protocol for NLSP

The Master initiates Throughput Request and Delay Request packets to calculate those two circuit characteristics. See page 3-19 for a specification of packet contents. Throughput Request and Delay Request packets are re-sent every 20 seconds until a response is received or the retry count is exhausted. If no response is received after the 16 retries, the router issues a disconnect to the link. The Sequence Number field matches a response with the corresponding request. For the Throughput Request, the matched pair contains the same sequence number (not an even/odd pair).

The Throughput and Delay packet exchanges provide empirical measurements of those two characteristics of the link. The measurements are used later during NLSP operation. *Throughput* is the amount of data (in bits per second) that can flow across the link. *Delay* is how long it takes to send a zero-byte packet to the destination. (Actually, it is an indication of the wire propagation delay to reach from one point to the other.)

Finally, the IW2 Master sends an Information Request packet to provide the Slave with operational information about the link. The Slave responds to acknowledge receipt. The exchange also informs each router of the other's textual name. If the Master does not receive the reply, NLSP retries every 20 seconds until a response is received or the retry count (16) is exhausted. If the retry counter is exhausted, close the data-link connection.

The following option subfields in the Information Request/Response exchange contain operational information used in NLSP. The following discussion indicates how to calculate the values that appear in the subfields, based on prior exchanges.

- NLSP Information option—Throughput subfield. This value is calculated based on the Throughput Request/Response exchange: two equal-sized Throughput Request packets are sent back to back. Each is *R* bytes. Nominally, R = 512, but the Master might use a different size based on local configuration, network management action, or the results of earlier exchanges on the link. The receiver (Slave) of these back-to-back packets measures the time between two packets and sends a Throughput Response packet with delta time and the Throughput Request packet length. The Master calculates the actual throughput capacity of the circuit by dividing packet size by delta time. The method assumes that packet size is large enough to offset the driver's holding time, and that the lower level driver is sending the packets one after another. The method also expects that an X.25 network delivers packets one after another to the receiving side. As illustrated in Figure 3-3, the request is repeated twice; the results are averaged, as follows:

R-byte packets

--------->

--------->

---------> *t1*

---------> *t2*

throughput1 = R / (t2 – t1)

R-byte packets

--------->

--------->

---------> *t3*

---------> *t4*

throughput2 = R / (t4 – t3)

Measured throughput = (throughput1 + throughput2) / 2

- NLSP Information option—Delay subfield. This Delay measurement appears in the Link Information option of the Link State packet. The quantity conceptually represents the one-way electrical propagation delay between the two points. It is calculated based on the Delay Request/Response exchange. The idea behind the calculation is that the measured time interval between sending a full packet and receiving its echo from the other router consists of two parts: (a) the time to clock the bits onto the wire, and (b) the propagation delay. Now, the time to clock bits onto the wire is the reciprocal of the Throughput. Combining these facts,

  Echo_time / 2 = Delay + (1 / Throughput).

  The division by 2 converts from round-trip to one-way. Now, Delay is the number being sought, Throughput was determined by the previous exchange, and Echo_time is measured by the Delay exchange. (Actually, it is measured twice and the measurements are averaged.) The calculation is made by solving the above equation for Delay.

Delay = Echo_time / 2 – (1 / Throughput)

with the proper conversion of units. This is depicted as follows:

576-byte packet

t1 ----------> ----------->+

t2 <---------- <----------+

echo1 = (t2 – t1) / 2 *(microseconds), the division by 2 is included here*

576-byte packet

t3 ----------> ----------->+

t4 <---------- <----------+

echo2 = (t4 – t3) / 2 *(microseconds), the division by 2 is included here*

Measured delay = (echo1 + echo2) / 2 –
[1,000,000 × 8 × 576 / Throughput]

- RIP/SAP Info Exchange option—WAN Link Delay subfield. This value is calculated as described on page 3-7. It must be present and calculated correctly even when NLSP has been chosen.

On receiving the Information Request, the Slave records the pertinent information and reponds with an Information Response packet, in which its own router name and node identifier are substituted.

After the Slave has received the Information Request and the Master has received the Information Response, NLSP packets and IPX data packets can start being sent over the link.

3.9.2. NLSP Configured Values

A router can permit the two values described in IW2 packets (NLSP Delay and NLSP Throughput) to be configured manually. This can be accomplished either by local action at the router or remotely by network management action.

Whether or not a router is configured manually, it proceeds through the IW2 exchange as specified above, reporting the measured values. After the IW2 exchange, when Link State packets begin flowing, the values in the Link State packets are the manually configured ones. The values appearing in the Information Request/Response are disregarded by routers that have overriding manually configured values.

If either (a) the router at only one end of a link is configured, or (b) the routers at the two ends of a link are configured with different values, the two routers can report different values in their respective Link State packets. This can result in asymmetric routes, but NLSP operates with asymmetric routes. (An asymmetric route is one in which traffic from *A* to *B* traverses different links than traffic from *B* to *A*.) However, asymmetric routes are considered harmful because of troubleshooting difficulties.

In no event does a router attempt to assert its configured values by manipulating the numbers sent in IW2 exchanges.

3.10. Client-Router Connection

Client-router connection is the method for an individual client personal computer or workstation to connect remotely to an internetwork. The IW2 exchanges to initiate client-router operation use the same basic structure as router-router operation. There are additional provisions. Some of these provisions result from the asymmetric nature of this routing type, which makes it quite different from router-router operation. Other additions are designed for the client to learn addressing information from the router. Finally, there are additions are designed to support *reconnection*.

Reconnection is a feature to deal with a practical aspect of dial access. A data-link disconnect can occur if software detects a period of inactivity, or if the line is low quality. If a client-router connection is broken and reestablished, it is desirable that the application software continue as though the interruption had not occurred. In particular, it is important that file server connections be maintained. To support this need, IW2 has provisions that permit the address information assigned in the initial connection to be asserted for the reestablished connection.

Clients must have an IPX network number and a unique IPX node number on that network to be individually addressable. However, clients do not need periodic RIP and SAP broadcasts. This allows routers to reduce background traffic on a client link by not sending any periodic RIP or SAP data. (Sending RIP and SAP does not cause a functional problem, but it does impact performance.)

RIP and SAP information should only be sent if the client makes a specific request for information: a service or route request.

If a router assigns the same IPX network number to multiple clients, broadcasts on that network — whether originated from a workstation or the router—must reach all the (other) workstations sharing the network number. The router duplicates the packet on all those client links.

3.10.1. Agreeing on Master/Slave Roles and Routing Type

Because client-router operation is asymmetric, additional rules apply to the initial negotiation to ensure that a client can only rendezvous with a router. Client-client communication is not supported by IW2, and router-router communication must use a different routing type.

As with router-router operation, the routing type is determined from the Timer Request. A client's Timer Request is recognized by including **only** the client-router Routing Type. A router offering the "client-router" Routing Type must also include at least one additional Routing Type in the Timer Request, to distinguish itself from a client. If the router does not happen to implement any additional routing type, it includes Routing Type zero (RIP) in the Timer Request with "No" as WAccept Option. Because of the "No," it is impossible to agree on "RIP."

If a client receives a Timer Request with either

- the client-router Routing Type missing

or

- only one Routing Type in the packet,

it disconnects the link.

In summary, as long as

- exactly one of the two parties is a client

and

- the other party supports the router side of client-router operation

the provisions of the preceding paragraph ensure that client-router operation is negotiated based on the Timer Request packets.

3.10.2. Choosing the Network Number

If this is the first time the client is connecting to the router, the client sets the WNodeID to zero in the Timer Request, indicating that the router should be the Master and allocate a network number for the new link. The Extended Node ID option is omitted. The client is the Slave.

If the client is reconnecting a link after an earlier disconnect, it must be able to specify its earlier network number. In this case, the client sets its WNodeID field to 0xFFFFFFFF, forcing itself to be the link Master. In this case, the router responds with only the client-router Routing Type acknowledged.

If the router can assign a network number

- for client-router operation

but not

- for router-router operation

it sets its the WNodeID to zero in the Timer Request. With client-router operation, the router must be willing to assign the network number, regardless of the WNodeID value in the Timer Request. This is precisely what happens when the client is connecting to the router for the first time.

All dial-in clients could share the same network number and be assigned unique node numbers by the router, or each workstation could be assigned a different network number. This is a router-specific implementation detail. Use of a single network for all clients is preferred; however, this does involve extra work by the router to replicate broadcast frames.

If a router receives an inappropriate Information Request from a client trying to set the network number, the router overwrites the values with a preferred value. When the client receives the Information Response, it must note the new value. If the client cannot adjust to the new value, it disconnects the link.

3.10.3. Choosing the Node Number

In client-router operation, the router can typically be communicating with many dial access clients simultaneously. To conserve address space, many clients are considered nodes on a single network, with one IPX network number. Because the clients are independent of each other, this requires a mechanism to assign IPX node numbers dynamically at the time a connection is made. (The node number is a six-byte value playing the role of the IEEE MAC address in the IPX header.) The mechanism is a Node Number option in the Information Request and Response packets.

As was explained in the previous section, the router is the Master for an initial client-router connection, but the client is Master for a reconnection. In both cases, the Master determines the IPX network number for use over the link. In the same way, the Master chooses the node number that the client uses on the link. The choice is transmitted in the Node Number field of the Information Request.

The most-significant byte of the Node Number must be the value 0x02. (In the IEEE MAC address space, the two least-sigificant bits of this value indicate that the Node Number is a nonmulticast, locally assigned value.) Other values are reserved. The router's own node number should be avoided.

Note: Because future versions of this specification might use some of the reserved values, the client should accept any Node Number assigned by the router, even if the most-significant byte is not 0x02.

When the router is the Master and allocating node numbers on a single network, it should always use a unique value by incrementing the Node Number for each client connection. This allows a reconnecting client to specify its old network and node numbers. It is unlikely with five bytes free in the node number that a wraparound in the numbers could occur. The router stores in permanent memory the range already consumed, to avoid confusion between reconnecting clients should the router be restarted.

If a router receives an inappropriate Information Request from a client trying to set the node number, the router overwrites the value with a preferred value. When the client receives the Information Response, it must note the new value. If the client cannot adjust to the new value, it disconnects the link.

3.11. Checking and Recovery Features of IW2

Packets received on socket number 0x9004 not having the WIdentifier field set to 0x5741534D are discarded.

If an unknown WPacket Type is received, the receiver should respond with a NAK packet.

If an invalid or unsupported WOption Number is received in a request, the receiver should respond to the packet with a WAccept Option set to "No" for that option.

If a router determines that it cannot support any of the Routing Types included in the Timer Response or Timer Request packet, it issues a disconnect on the underlying data-link connection.

If a Timer Request arrives after the router has sent or received a Timer Response, the exchange reverts to the start. That is, the router reverts to sending its Timer Request packet.

When a router assumes the Slave role, it starts a 60-second timer. If the IW2 exchange has not concluded by the time the timer expires, the router issues a disconnect on the underlying data-link connection.

If a protocol error is detected, the router issues a disconnect on the underlying data-link connection. Protocol errors include (a) packets received out of order, and (b) a Slave receiving a Timer Response. The same happens if no response is detected after the request retries are exhausted. If this happens before the Master role is decided, either router issues the disconnect.

If a router receives a packet with an illegal WPacket Type, or if the options are malformed (for example, they extend beyond the end of the packet), the router sends a NAK packet, and otherwise ignores the offending packet. (There is an exception. If the routing type is RIP, or if it is not yet decided, a NAK is not sent. This is because Reference [All92] does not include the NAK.) When a router receives a NAK packet, it can try something different or it can cause a disconnect.

After a disconnect, the router that initiated the original connection can (at its option) try to establish a new data-link connection. With X.25 switched circuits, this means establishing a new virtual circuit before restarting IW2. With dial-up PPP, this means establishing a new data-link connection before restarting PPP and then IW2. With a dedicated PPP link, this means that both peers restart the PPP protocol from the beginning, then IW2.

For an X.25 PVC, a Frame Relay PVC, or IP Relay, establishing a new data-link connection means restarting IW2 from the beginning. In these cases, the Timer Request can be retried indefinitely.

Under certain circumstances—particularly on X.25 or Frame Relay PVCs—it is only possible to detect that the remote router went away when it comes back up again. In this case, one side of the link receives a Timer Request packet when IPX is in a fully connected state. The side receiving the Timer Request must realize that a problem occurred and start tearing down the link to reestablish the connection.

Note: Resolving call collisions is outside the scope of this specification. Call collisions occur when two routers are configured to maintain a circuit to each other on the same physical connection, and the two call requests are issued at the same time. Resolving call collisions is the responsibility of the underlying Data-link layer.

Note: It might happen that two circuits become established connecting the same two end points. In fact, this might be desirable to increase throughput in some situations. IW2 does not include any protocol provision to discriminate between wanted and unwanted duplicates. Every circuit is kept, unless clearing it is forced by local manual action or by operation of network management.

3.12. Recalibrating Throughput and Delay

The specification does not currently include a provision for the Throughput or Delay to be recalibrated after completion of the IW2 exchange, while the circuit is actively conveying routing traffic and data traffic.

To allow this capability to be added in the future, a router that receives a Throughput Request or Delay Request after completion of the IW2 exchange ignores those packets, without generating an alarm and without disconnecting the circuit.

3.13. Forwarding “Type 20” Packets

Section 5 of reference [Nov92] documents the use of IPX Packet Type 20 to implement a form of internetwork broadcast. That reference specifies a list in the IPX packet of up to eight Network Number fields. The list indicates the path traversed by the packet so far. There are two cases in which a WAN link does not have a network number for inclusion in the list.

- When a system forwards an IPX Type 20 packet on an Unnumbered RIP link, add the internal network number of the sending system to the Network Number list.
- When a system forwards an IPX Type 20 packet on an NLSP WAN link, add the internal network number of the receiving system to the Network Number list. That network number is learned previously as part of the IW2 exchange.

IPX Type 20 packets are not forwarded over on-demand, static routing links.

3.14. IW2 Database

3.14.1. Constant Values

minMTU
: The packet size that every router must be prepared to handle: 576 bytes, including IPX header and data, but not data-link headers or trailers.

3.14.2. Configured Values

internalNetworkNumber
: Every NLSP router and every NetWare 3 or 4 server has an IPX network number internal to the system itself. The network number is within the relevant routing area, but is distinct from the network numbers of the physical network segments in a routing area. Any application services operating on the system are attached to the internalNetworkNumber. Network management packets to a router should be sent to the internalNetworkNumber. (In previous specifications, such as Reference [All92], there is the concept of a *primary network number* for systems not necessarily having an internal network number. A *primary network number* can be either an internal one or the network number of a distinguished permanently attached network segment. The internalNetworkNumber idea supersedes that of a primary network number.)

nodeNumber
: In some situations, a router needs a six-byte IPX node number value for itself on a WAN link, to play the role of the IEEE MAC address. For example, the value would be used as the source node number in the IPX header of an outgoing RIP or SAP packet. The node number is formed from the internalNetworkNumber. It is the four-byte internalNetworkNumber followed by two zero bytes.

routerName
: Every router has a readable textual, symbolic name. It is especially useful for network management purposes to identify systems by a user-defined mnemonic name. The name is 1 to 47 characters, containing uppercase English letters, underscore (_), hyphen (-), and "at" sign (@). When a router coexists on the same system as a NetWare server, the routerName coincides with the server name. For a client-router connection, the client should use a string reflecting the user's name; if the name is not known, use the default string "DIAL-IN-CLIENT".

nlspDelayOverride
: A router can provide to the user a mechanism to (optionally) override the NLSP Delay, usually determined empirically by the IW2 protocol exchange.

nlspThroughputOverride
: A router can provide to the user a mechanism to (optionally) override the NLSP Throughput, usually determined empirically by the IW2 protocol exchange.

3.14.3. Dynamic Values

ipxWanNetworkNumbers
: As a result of IW2 exchanges, each WAN circuit is either (a) determined to run an unnumbered routing protocol, or (b) has an IPX network number assigned. The router keeps track of the network number (or lack thereof) for each circuit. The router can

optionally have a pool of network numbers to assign dynamically during IW2 operation to circuits being brought into service.

3.14.4. Dynamic Values per Circuit

Delay
: The electrical propagation delay for a signal to flow from one end point to the other, in milliseconds.

Throughput
: The amount of data, in bits per second, that can flow through the circuit if there is no other traffic.

RIP Link Delay (Ticks)
: The amount of time it takes a 576-byte packet to reach a location. It is measured in ticks, with 18.21 ticks per second (minimum of one).

3.15. Packet Structures

IW2 employs one kind of packet structure: IPX WAN (page 3-19).

These packets ride in the data portion of IPX packets. The IPX header fields are encoded:

Destination Network	zero
Destination Node	0xFFFFFFFFFFFF
Destination Socket	0x9004
Source Network	zero
Source Node	zero
Source Socket	0x9004
Packet Type	4

IPX WAN packets never exceed minMTU bytes in size, including the IPX header but excluding the data-link headers. All multibyte fields are transmitted with the most significant byte first.

3.15.1. IPX WAN 2

IPX WAN 2

	Number of Bytes
WIdentifier	4
WPacket Type	1
WNode ID	4
WSequence Number	1
WNum Options	1
Variable Length Fields	Variable

WIdentifier: 0x5741534D (ASCII for "WASM") indicates packets of the IPX WAN family of protocols.

WPacket Type: Identifies a specific type of IW2 packet:

0 = "Timer Request"

1 = "Timer Response"

2 = "Information Request"

3 = "Information Response"

4 = "Throughput Request"

5 = "Throughput Response"

6 = "Delay Request"

7 = "Delay Response"

0xFF = "NAK;" the packet content is an exact copy of the received, rejected packet, except that (a) the Packet Type indicates NAK, and (b) the WNode ID is that of the sending router. Only packets with the proper Identifier are replied to with NAK.

Packet Types 0 through 3 are defined in Reference [All92]. Packet types 4 through 7 and 0xFF apply only to IW2.

WNode ID: The internalNetworkNumber of the sending router. There are three exceptions, each applying only to the Timer Request packet.

- If the router cannot assign a network number to the link, this field contains zero and the internalNetworkNumber is conveyed in the Extended Node ID option.
- If a client is connecting initially, it places a zero in this field.
- If a client is reconnecting to a router, it places all-ones in this field.

WSequence Number: This field starts with zero for each request packet type and is incremented by one for each retry of that packet type. Each response echoes the sequence number of the request. This allows each response to be matched with the request that provoked it.

WNum Options: Indicates how many options are present in the variable length fields. Options in a packet can occur in any order.

Option

	Number of Bytes
WOption Number	1
WAccept Option	1
WOption Data Len	2
WOption Data	WOption Data Len

Variable Length fields: A series of optional fields, each of which has the following four-part Option form.

WOption Number: Identifies a particular option. Currently defined codes are in the following bullet list.

WAccept Option: This field is used for negotiating options between the two parties:

0 = "No"

1 = "Yes"

3 = "Not Applicable"

WOption Data Len: The length, in bytes, of the WOptionData field.

WOption Data: Details specific to a particular code.

Currently defined codes, and the corresponding values, are as follows:

- Routing Type

  Code = 0x00 Length = 1

  Value = Type of IPX routing protocol the sender can support (or cannot support if this is a Timer Response and WAccept Option is "No"):

  0 = "RIP"
  1 = "NLSP"
  2 = "Unnumbered RIP"
  3 = "On-demand, static routing"
  4 = "Client-router connection"

- RIP/SAP Info Exchange

  Code = 0x01 Length = 54

  Value = The following three subfields:

  - WAN Link Delay: The metric IPX RIP uses for routing; the RIP Delay in milliseconds.

    Note: The units here are different from the units (Ticks) in RIP packets.

  - Common Network Number: For numbered Routing Types, this field contains the IPX network number assigned to the link; for unnumbered Routing Types, the value zero occupies this field.
  - Router Name: A unique textual identifier for the router, one byte per character, seven-bit ASCII, flush-left, null-filled.

RIP/SAP Info Exchange

	Number of Bytes
WAN Link Delay	2
Common Network Number	4
Router Name	48

- NLSP Information

  Code = 0x02 Length = 8

  Value = Two subfields, Delay and Throughput, containing values estimated by IW2 exchanges.

NLSP Information

	Number of Bytes
Delay	4
Throughput	4

- NLSP Raw Throughput Data

  Code = 0x03 Length = 8

  Value = The following two subfields:

  - Request Size: Size of the Throughput Request packet.
  - Delta Time: The time interval, in microseconds, between receipt of equal size Throughput Request packets.

NLSP Raw Throughput Data

	Number of Bytes
Request Size	4
Delta Time	4

- Extended Node ID

  Code = 0x04 Length = 4

  Value = Sending router's internalNetworkNumber. This option occurs only in Timer Requests and only when WNode ID is zero.

- Node Number

  Code = 0x05 Length = 6

  Value = IPX node number to be used by the client on a client-router connection. It is comparable to the IEEE MAC address.

- Compression

  Code = 0x80 Length = variable

  Value = The first byte indicates a style of compression; the remaining bytes specify parameters pertaining to that style.

  Currently, one compression type is specified:

  Telebit compression (option length = 3 bytes), see Reference [Mat92].
  Byte 0: Contains the value 0 to indicate Telebit compression.
  Byte 1: Compression options.
  Byte 2: Number of compression slots.

- Pad

  Code = 0xFF Length = variable

  Value = A sequence of totally random bytes to fill the packet to the required size. The Value is ignored on receipt.

  Note: The purpose of randomness is to prevent data-compressing modems from defeating accurate estimations of the link's operational characteristics. Randomness is a change from reference [All92].

Figure 3-4 summarizes which options are permitted in which request/response packets. The text of this chapter explains which are actually present under what circumstances. The notes are itemized immediately after the figure.

Timer Request

	Notes
IW2 Header (fixed part)	See page 3-19
Routing Type	R d
Extended Node ID	A i
Compression	O y
Pad	M i

Timer Response

	Notes
IW2 Header (fixed part)	See page 3-19
Routing Type	= 4 c
Compression	E 3 c
Pad	M i

Throughput Request

	Notes
IW2 Header (fixed part)	See page 3-19
Pad	M P i

Throughput Response

	Notes
IW2 Header (fixed part)	See page 3-19
NLSP Raw Throughput Data	M P y

Delay Request

	Notes
IW2 Header (fixed part)	See page 3-19
Pad	M P i

Delay Response

	Notes
IW2 Header (fixed part)	See page 3-19
Pad	M P i

Information Request

	Notes
IW2 Header (fixed part)	See page 3-19
RIP/SAP Info Exchange	M y
NLSP Information	N y
Node Number	D y

Information Response

	Notes
IW2 Header (fixed part)	See page 3-19
RIP/SAP Info Exchange	= y
NLSP Information	= y
Node Number	D y

Figure 3-4: IW2 Option Usage

Notes for Figure 3-4

M Mandatory field.

N Mandatory for NLSP; optional otherwise.

O Optional field; can appear more than once.

= Response fields contain the same WOption Data as the corresponding request fields.

E Response fields correspond one-for-one with request fields; the response can zero out zero or more option bits of the Telebit options field; it must not set bits that were zero in the request.

P Throughput and Delay exchanges occur only with NLSP.

R Routing types appear in order of preference (most preferred first). At least one must be present.

A Appears if and only if the WNode ID field is zero.

3 Each WAccept Option field is set to indicate acceptance or rejection; the responder selects, **at most,** one of the choices offered in the Request.

4 Each WAccept Option field is set to indicate acceptance or rejection; the responder selects **exactly** one of the routing types offered in the Request.

D Negotiation in client-router routing type only.

y WAccept Option must be “Yes.”

d WAccept Option must be “Yes,” except for one special case in client-router operation.

c WAccept Option must be “Yes” or “No.”

i Value of WAccept Option is immaterial.