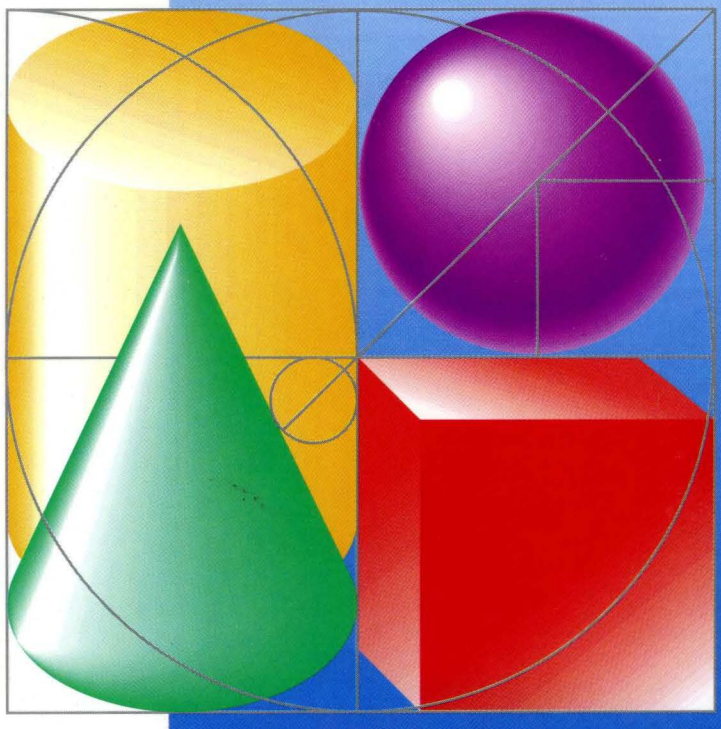
AppleTalk® Phase 2 Protocol Specification

An Addendum to Inside AppleTalk

APDA™ # C0144LL/A



Apple Computer, Inc.
20525 Mariani Avenue
Cupertino, California 95014
(408) 996-1010
TLX 171-576

To reorder products, please call:
Apple Programmers and Developers Association
1-800-282-APDA

AppleTalk® Phase 2 Protocol Specification

An addendum to *Inside AppleTalk*

APPLE COMPUTER, INC.

© Apple Computer, Inc., 1989
20525 Mariani Avenue
Cupertino, CA 95014-6299
(408) 996-1010

Simultaneously published in the United States and Canada.

Contents

Chapter 1 **Introduction**

THIS DOCUMENT DESCRIBES the changes to AppleTalk® protocols that are defined by AppleTalk Phase 2. AppleTalk Phase 2 provides extensions to AppleTalk addressing and enhancements to AppleTalk routing and naming services. These changes are designed to support larger AppleTalk networks while remaining compatible with current AppleTalk network hardware and software. (AppleTalk routers, however, must be upgraded to support AppleTalk Phase 2.)

Apart from the changes in AppleTalk protocols described in this document, all other aspects of the AppleTalk protocol suite remain as defined in *Inside AppleTalk*, published by Addison-Wesley. The original AppleTalk protocol suite will hereafter be referred to as *AppleTalk Phase 1*.

Since many of the changes found in AppleTalk Phase 2 involve routing and related protocols, much of this document is concerned with the role of AppleTalk routers and the issues encountered by developers of such routers. Readers who are not developing routers may not require all of this information. ■

Goals of AppleTalk Phase 2

AppleTalk Phase 2 was designed to meet four key design goals:

- the ability to address more nodes per network than the 254 node limit of AppleTalk Phase 1
- the ability to create multiple zones per network
- improved performance in a large, heterogeneous network environment
- a smooth upgrade path from AppleTalk Phase 1 with maximum compatiblity for users

◆ *LocalTalk™ is unchanged:* The changes in AppleTalk addressing and zones defined in this document are required only for EtherTalk™ and TokenTalk™ networks. LocalTalk networks are unchanged from AppleTalk Phase 1 (with the exception of changes in routers).

Some of the changes described in this document can optionally be implemented in LocalTalk nodes to fully conform to AppleTalk Phase 2. These changes are listed in the Appendix to this document, "Changes in LocalTalk Nodes."

Chapter 2 **New Terms in AppleTalk Phase 2**

A NUMBER of new terms are introduced in AppleTalk Phase 2. These terms are described briefly in this chapter and are dealt with in more detail in later sections of this document. ■

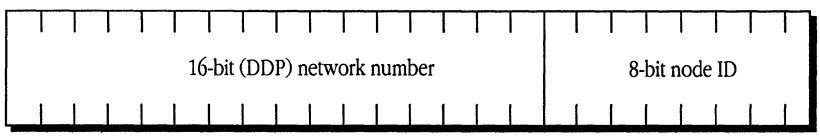
Extended AppleTalk network

To allow the addressing of more than 254 nodes on a network, AppleTalk Phase 2 defines the concept of an **extended AppleTalk network,** which is identified by a *range* of network numbers rather than a single network number. A network number range is made up of a series of contiguous network numbers (for example, 1000–1099). An extended network can also have a range of one (for example, 3–3).

Delivery of packets to a node on an extended network must now be done by using both the node's 8-bit node ID *and* a 16-bit network number that is within the range assigned to the network (see *Figure 2-1).* This expands the prior 2^8 address limit (254 nodes) to approximately 2^{24} addresses, or more than 16 million nodes, that can be uniquely addressed on an extended network.

AppleTalk Phase 2 places further restrictions on the values of network numbers that can be assigned by network administrators to AppleTalk networks. As with AppleTalk Phase 1, network numbers 0 and $FFFF cannot be used by any AppleTalk network. Furthermore, the range of numbers $FF00 to $FFFE (referred to as the *startup range,* described below) is now reserved; it is not available for assignment to actual networks in an internet environment.

■ **Figure 2-1** Node addressing in AppleTalk Phase 2



Nonextended network

An AppleTalk network that does not implement extended addressing, such as a LocalTalk network or an EtherTalk 1.0 network, is referred to as a **nonextended network.**

Provisional address

The dynamic process by which a node selects its node ID is now extended to include selecting its network number as well.

This is done in two steps. In order to pick a network number within the range of numbers assigned to its network, the node must collaborate with the routers connected to the network. However, this communication itself requires that the node select an address, known as a **provisional address,** including a provisional network number. It is essential that this provisional network number value not conflict with any other network numbers already in use on an internet. The selection of the provisional address involves the use of a specially reserved range of network numbers known as the startup range.

Network number startup range

As in AppleTalk Phase 1, a node saves its most recently used address in parameter RAM (pRAM) and attempts to use this address *as a provisional address* when restarted. However, if no address is found in pRAM, the node selects its provisional network number from a **startup range,** which is specified as $FF00–$FFFE. *This is a reserved range of network numbers that should never be assigned to any AppleTalk network in an internetwork environment.*

Final address

If an internet router is running on a network, or if one is started after the node has started up, the provisional address that a node had acquired upon startup is replaced by a **final address.** The network number in this address is selected from the range of network numbers actually assigned to the network by a seed router. This final address is unique on the internet and can be used without risk of address conflict.

Network-wide broadcast

AppleTalk Phase 2 defines a **network-wide broadcast,** which is a broadcast intended for all nodes on a given network. It is sent to destination network $0000, node $FF, and received and accepted by every node on the network.

Zones list

On an extended AppleTalk network, *there is no strict relationship between zone names and network numbers.* In fact, two nodes having the same network number can be in different zones.

In AppleTalk Phase 2, a list of zone names is associated with each extended network. A node on an extended network can belong to any one of the zones in that network's **zones list.** The zones list for an extended network is assigned by a network administrator when setting up seed routers for the network.

Default zone

The selection of the zone to which a node will belong is typically done by that node's user. For certain nodes, such as servers, which do not have a user interface or display capabilities, this is not feasible. For such nodes, AppleTalk Phase 2 identifies a zone to which any node on an extended network will automatically belong until a different zone is explicitly selected for that node. This zone is known as the **default zone** for that extended network. A node can obtain the name of its default zone from a router.

Zone multicast address

To prevent NBP LookUp broadcasts intended for nodes belonging to a given zone from interrupting *all* AppleTalk nodes on an extended network, each node on an extended network is assigned a **zone multicast address.** The zone multicast address is a data-link–dependent multicast link-level address at which the node will receive the NBP broadcasts directed to its zone. This link-level multicast address is provided by the network's routers.

Zone-specific broadcast

A **zone-specific broadcast** is a broadcast intended only for nodes belonging to a given zone. It is sent to a link-level zone multicast address and the AppleTalk network address $0000, node $FF, and is accepted only by nodes belonging to the zone indicated in the broadcast packet. (See the section "NBP" in Chapter 4.)

IEEE 802.2 and SAP

The Institute of Electrical and Electronics Engineers (IEEE) **802.2** standard defines service interfaces and packet formats for both connectionless and connection-based data-link service. The AppleTalk Phase 2 EtherTalk and TokenTalk link access protocols (ELAP and TLAP) use the 802.2 Type 1 format, which corresponds to the 802.2 connectionless service. In this standard, a **service access point (SAP)** is defined to differentiate between the different client protocol stacks using 802.2. The SAP to which AppleTalk packets are sent is $AA, which is reserved for use by protocols that are not defined by the IEEE.

SNAP

To distinguish between the different protocol families using the $AA SAP, all packets sent to this SAP must begin with a 5-byte protocol discriminator. The use of a protocol discriminator for the $AA SAP is known as the **Sub-Network Access Protocol (SNAP).** The format for an 802.2 Type 1 SNAP packet is shown in the descriptions of ELAP and TLAP in Chapter 4, "Protocol Details."

Chapter 3 **Operational Overview**

THE IMPLEMENTATION of extended addressing involves the following sequence of operations:

- obtaining a provisional address
- obtaining network information
- checking the validity of the provisional address and zone
- distinguishing between packets intended for delivery on the local network and those to be sent to a router for internet delivery

These operations are summarized below, and they are further described in the next chapter, "Protocol Details." ■

Obtaining a provisional address

Obtaining a provisional address upon startup is accomplished in two different ways depending on whether the node has previously saved an address in parameter RAM. An address saved in pRAM consists of two parts: the 16-bit network number, denoted $nnnn, and the 8-bit node ID, denoted $yy. The concatenated 24-bit address value of [network number, node ID] is denoted $nnnnyy.

If no address was previously saved in pRAM

Upon starting up when no information is saved in parameter RAM, a node will randomly select a provisional network number $FFnn in the startup range. This range is specified to be $FF00 to $FFFE inclusive (most significant byte first). The node then also randomly selects a node ID $yy (yy cannot be $00, $FE, or $FF). As in AppleTalk Phase 1, the node must use AARP to ensure that $FFnnyy is not in use by any other node on the network. If another node is already using this address, the node should try all other possibilities for $FFnnyy until a valid provisional address is obtained.

If an address was saved in pRAM

If there is a saved 24-bit address of the form $nnnnyy in pRAM, the node can use it as the provisional address. The node must use AARP to ensure that this address is not in use by any other node. If another node is already using this address, the node should try all other possibilities for yy (yy cannot be $00, $FE, or $FF) keeping nnnn the same until all possibilities are exhausted (nnnn is probably a valid network number for this network unless the node has been moved from another network).

If all possibilities are exhausted, the node must select a new address as if none was previously saved in pRAM (as described in the previous section).

Obtaining network information

Once a node has obtained a provisional address, if a router is running on the network, the node can acquire information about its network from the router. The node learns the range of valid network numbers and confirms that its saved zone name is valid for that network. If it does not currently have a saved zone name, it can obtain the list of available ones from a router and pick one in an implementation-dependent manner.

Network information is obtained by using a new request packet, ZIPGetNetInfo. If the node has a saved zone name for the zone to which it wishes to belong, this name is also included in the request. The request is retransmitted until a router responds or the retry count is exceeded. If no response is received, this is taken to imply that *no* router is currently running on the network.

The response will provide the following information:

- the address of the responding router (to be saved as the node's initial A-ROUTER parameter)
- the range of valid network numbers for this network (to be saved as the node's THIS-NETWORK-RANGE parameters)
- the zone multicast address to use for the desired zone name (if this zone name was valid)
- the default zone name (and multicast address) to use if the requested zone name is not valid

Upon receiving the response, a node registers with its data-link layer (in a data-link–dependent manner) to receive packets sent to the zone multicast address; routers will send NBP broadcasts (for the node's zone) to this link-level address.

Obtaining a valid address and zone

After obtaining the desired network information, if a node's provisional network number is not in THIS-NETWORK-RANGE, the node can proceed to select an address in this range, and then ensure through the use of AARP that the address is not in use by another node on the network. This new address then becomes the node's final address, and is saved in pRAM. Note that, except when a node is first started on a network, its provisional address will typically be in the correct range, and it should not have to repeat the address acquisition process.

If the zone to which a node wishes to belong is not valid for its network, or if the node has no saved initial zone name, it is possible for the node to obtain the zones list for the network. This is obtained from a router by using the new ZIP GetLocalZones request, and then selecting a zone from that list in an implementation-dependent manner (on the Macintosh, a dialog box is displayed asking the user to choose a zone). The node can then issue the ZIP GetNetInfo request to obtain the node's zone multicast address, and to register that address with its data-link layer.

Note that in a network with just one zone name in its zones list, a user need not be made aware of the node's zone name at all. Such a network will require no user intervention to select the node's zone name, and should appear no different from a nonextended network.

Operating without a router

Most extended networks are likely to be connected by routers into internets. However, extended addressing is implemented in a way that also permits operation on a network where no router is present. Such extended networks will in particular be used when more than 254 nodes are to be connected to a single AppleTalk network.

Furthermore, an extended network may operate without a router as a transitional state between operating with and without routers.

If no router is present

If no router responds to a ZIP GetNetInfo request, the node uses the provisional address as its final address. Any previously saved zone name is ignored; the node is, for now, in zone "*" and has no zone multicast address. If a router comes up later, the node will be able to communicate with the internet as long as its final address is still within the network number range for the network (this will typically be the case if the provisional address was obtained from pRAM and the node has not been moved to another network).

When a router first comes up

When a router first comes up on a network, the address of a node already operating on the network might now *not* be valid for the network (its network number may be outside the range entered in the router). This condition must be corrected before the node can continue to communicate on the internet. The following is the procedure a node should use when first detecting a router (via RTMP packets).

If the node's network number is in the startup range

If a node's address is of the form $FFnnyy, that is, in the startup range, the router must be ignored until the node's AppleTalk implementation is reinitialized to allow the node to acquire a valid internet address. The node should not send any packets for forwarding on the internet, nor name lookup requests to A-ROUTER. It can, however, continue to access nodes on the local network in the normal fashion. The node remains in zone "*", but will continue to see all nodes on the local network regardless of their zone (since it will still send NBP lookups as AppleTalk broadcasts).

If possible, the node should alert its user that the node's AppleTalk implementation must be reinitialized in order to continue to access the internet. Nodes using network numbers not in the startup range will not be able to see those nodes using network numbers in the startup range. This is true regardless of the zone, since the network number startup range has no zone multicast address. Likewise, nodes outside this network also will not see its nodes that are using network numbers in the startup range.

If the node's network number is not in the startup range

If the node's network number is *not* in the startup range, upon detecting the first router, the node needs to determine whether its own address is in the correct network number range for the network. This correct network number range is determined from the header of the RTMP packet that alerted the node of the router's existence.

If the node's network number is in the correct range

If the node's network number is in the correct network number range, it can proceed to verify that its desired zone name is valid for the network, obtain its zone multicast address through a ZIP GetNetInfo request, and set its data-link layer to listen on that address. If its desired zone name is not valid for the network, the node can either issue a ZIP GetLocalZones request and ask the user to select a zone or decide to belong to the default zone for that network. A valid zone name should be saved in long-term storage upon receipt.

If the node's network number is in neither the startup range nor the correct range

If the node determines that its network number is neither in the startup range nor in the correct network number range, it should behave exactly as if its address was in the startup range. However, nodes in the valid network number range that had been communicating with this node until the router came up will now lose access to the node. The node's AppleTalk implementation should be restarted as soon as possible.

When the last router goes down

When a network on which one or more routers have been running loses all routers, the normal mechanism for aging the A-ROUTER entry in nodes should prevail. In addition, the node should also expand THIS-NETWORK-RANGE to $0001–$FFFE, making all network numbers valid. The node will continue to operate with the address that it was using. Therefore, the network condition reverts back to one where no routers are available at startup. A node's zone name should revert to "*", and the node's zone multicast address should be deleted (if one has been set).

Extended addressing in operation

Sending a packet remains much the same as on a nonextended network. A node determines whether the desired destination is on the local network, and if so, obtains the destination node's hardware address using AARP and sends the packet out. If the destination node is not on the local network, AARP can be used to obtain the hardware address of a router and the packet can be sent to that router.

This is much like the AppleTalk Phase 1 DDP algorithm of comparing the destination network number to THIS-NET and sending the packet to A-ROUTER if it does not match. Instead, the node now compares the destination network number to THIS-NETWORK-RANGE and sends it to A-ROUTER if it is not in this or the startup range. (This scheme can, however, result in an extra hop for off-network traffic; a more efficient scheme to choose the "best" router is described in the section "DDP" in the next chapter.)

Note that, on an extended network, DDP broadcasts can now occur in three forms:

- A **network-wide DDP broadcast** is addressed to destination network $0000, node ID $FF, and received and accepted by every node on the network.
- A **network-specific DDP broadcast** is intended for all nodes on the network whose DDP address includes the specified network number. Such a datagram is addressed to a specific network *number* (not the entire network number range), and node ID $FF. All nodes on the network will receive this as a data-link broadcast and must discard it if not intended for them (that is, if it does not match their network number).
- A **zone-specific DDP broadcast** is intended for all nodes on the network that belong to the specified zone. Such a broadcast is sent at the data-link level to a zone multicast address with DDP destination network $0000, node ID $FF. Only routers transmit zone-specific broadcasts.

The operational aspects of extended addressing are further specified in the next chapter, "Protocol Details."

Chapter 4 **Protocol Details**

THE FOLLOWING SECTIONS DESCRIBE new aspects of AppleTalk protocols and changes in protocol packet formats brought about by the architectural changes of AppleTalk Phase 2. ■

AppleTalk data links

AppleTalk data links must support some form of broadcast mechanism, a method of sending packets to *all* nodes connected to the data link.

It is desirable that the link support multicast addressing. In particular, one of these multicast addresses would be used as a means to send a packet to all AppleTalk nodes on a network. The use of this multicast address, known as the **AppleTalk broadcast data-link address,** will ensure that the packet is received by AppleTalk nodes only, without disruption of non-AppleTalk nodes.

Data links that support multicast addresses should also define a number of these to be used as zone multicast addresses. Each AppleTalk zone on the data link must be mapped to one of these addresses, as specified in the description of ZIP later in this chapter.

An AppleTalk node should be able to receive packets sent to

- the node's data-link address
- the AppleTalk broadcast address
- the multicast address for the node's zone

ELAP and TLAP

The EtherTalk and TokenTalk link access protocols (ELAP and TLAP) provide connectionless service using IEEE 802.2 Type 1 packet formats. Packets are sent to SAP $AA. The SNAP protocol discriminator for AppleTalk is $080007809B (Apple's company code followed by the EtherTalk protocol type).

Figure 4-1 shows AppleTalk packet formats for ELAP and TLAP. The data portion of the packet begins with the long format DDP header. There is no LLAP-style header.

■ **Figure 4-1** AppleTalk packet formats

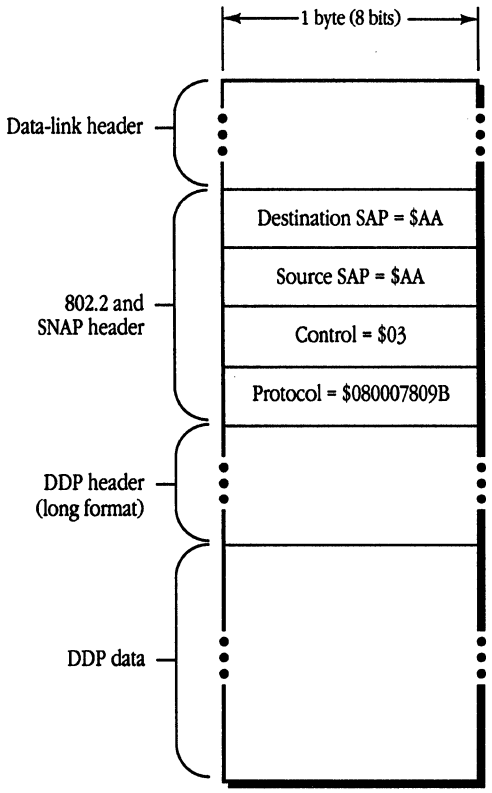


Table 4-1 specifies the AppleTalk broadcast and zone multicast addresses used by ELAP and TLAP.

■ **Table 4-1** Broadcast and zone multicast addresses for ELAP and TLAP

	ELAP	TLAP
AppleTalk broadcast address	$090007FFFFFF	$C00040000000
Zone multicast addresses When used with the address assignment algorithm described in this chapter, the first address in each list represents a_0.	$090007000000 *253 addresses* $0900070000FC	$C00000000800 $C00000001000 $C00000002000 $C00000004000 $C00000008000 $C00000010000 $C00000020000 $C00000040000 $C00000080000 $C00000100000 $C00000200000 $C00000400000 $C00000800000 $C00001000000 $C00002000000 $C00004000000 $C00008000000 $C00010000000 $C00020000000

AARP

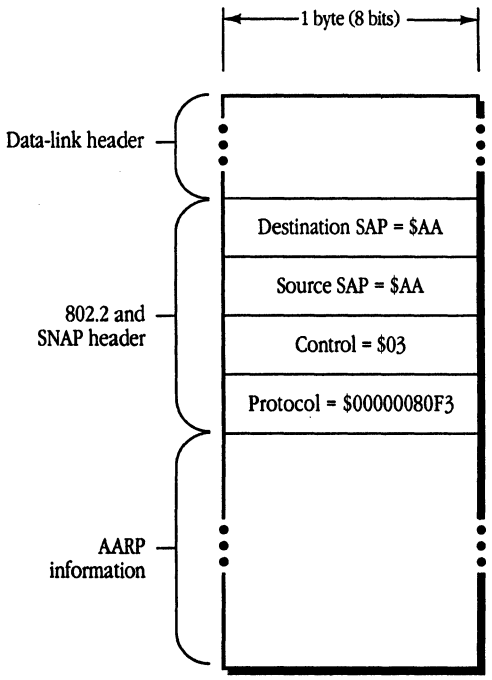
AARP packets are also encapsulated in the IEEE 802.2 format on data links that support this standard. The SNAP protocol discriminator for AARP packets is $00000080F3. The AARP packet format is shown in *Figure 4-2.*

AARP *cannot* use the following reserved addresses when probing for an AppleTalk address:

- $nnnn00
- $nnnnFE
- $nnnnFF
- any address using $0000 or $FFFF as the network number (nnnn)

When probing, AARP should send ten probes, at one-fifth second intervals. AARP probes and requests should be sent to the AppleTalk broadcast address.

■ **Figure 4-2** AARP packet format



The exact sequence in which AARP tries 24-bit addresses while probing is not specified here. Subject to the constraints of the range within which AARP is choosing the address, it is free to choose addresses in any order.

DDP

The DDP packet format remains unchanged from AppleTalk Phase 1; however, only the *long* DDP header format is used. A destination network number of $0000 is accepted by all nodes on the network (see "Packet Filtering" in this section).

Sending packets to a router

DDP provides an added service in routing nodes on extended networks, for use primarily by the NBP routing algorithm. A DDP destination address of the form $nnnn00 is allowed, meaning "any router connected to a network whose network number range includes nnnn (or whose network number is nnnn)." A packet with such an address is forwarded through an internet via the normal DDP forwarding mechanism until it reaches such a router. This router accepts the packet as if it were sent to the packet's destination address.

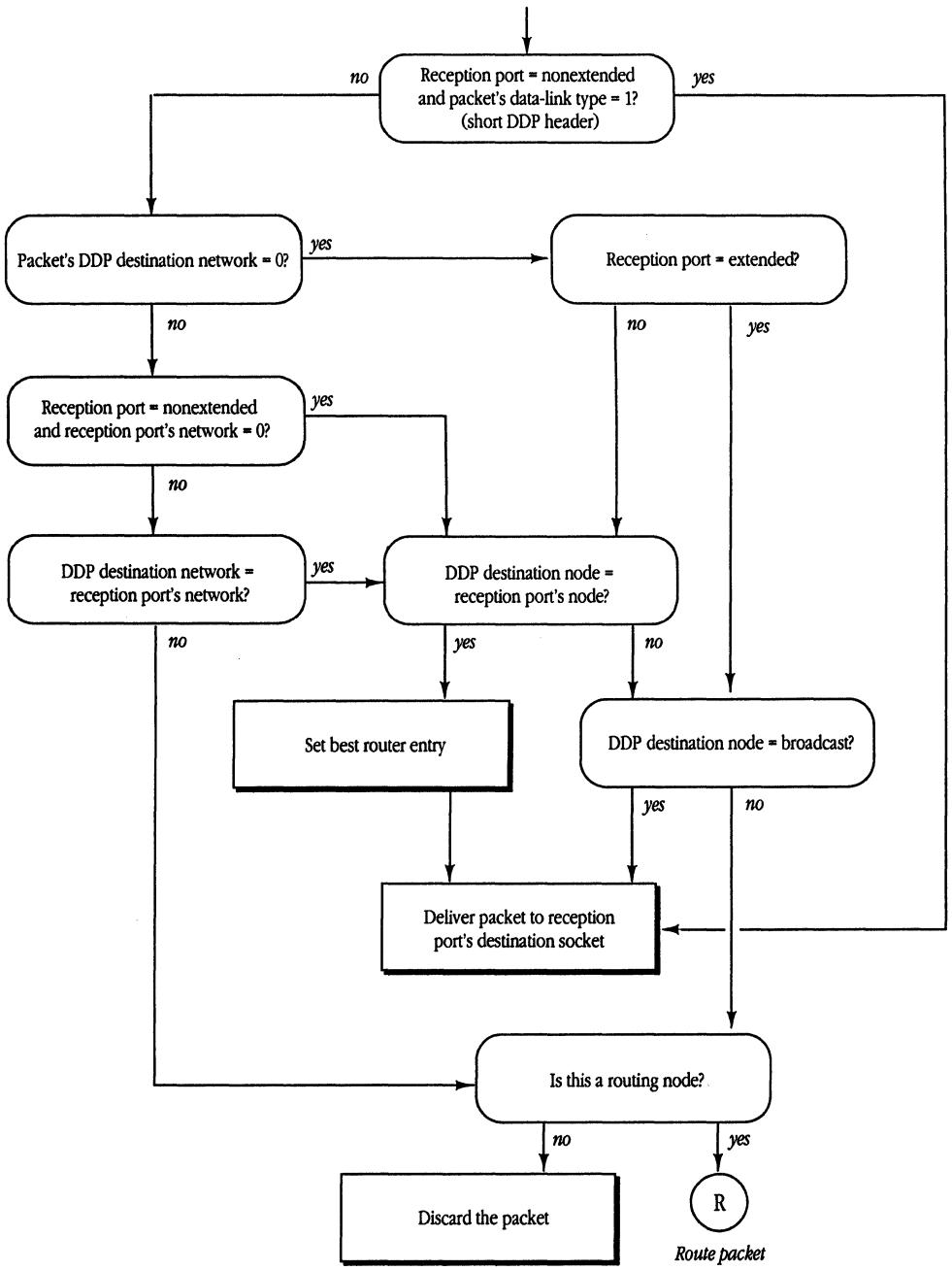
Support for zone multicasts

DDP in a routing node must provide the node's router with the ability to send a packet to a specific zone multicast address. This ability is used by NBP. Such an ability should *not* be provided in nonrouting nodes.

Packet filtering

DDP in nonrouting nodes on extended networks should always accept datagrams addressed to destination network number zero, node ID $FF (these are network-wide or zone-specific broadcasts). However, DDP on extended networks should *not* accept datagrams destined for network zero and any node ID other than $FF (even the node's own). DDP on nonextended networks should accept both. The DDPRead algorithm shown in *Figure 4-3* illustrates this process.

■ **Figure 4-3** The DDPRead algorithm

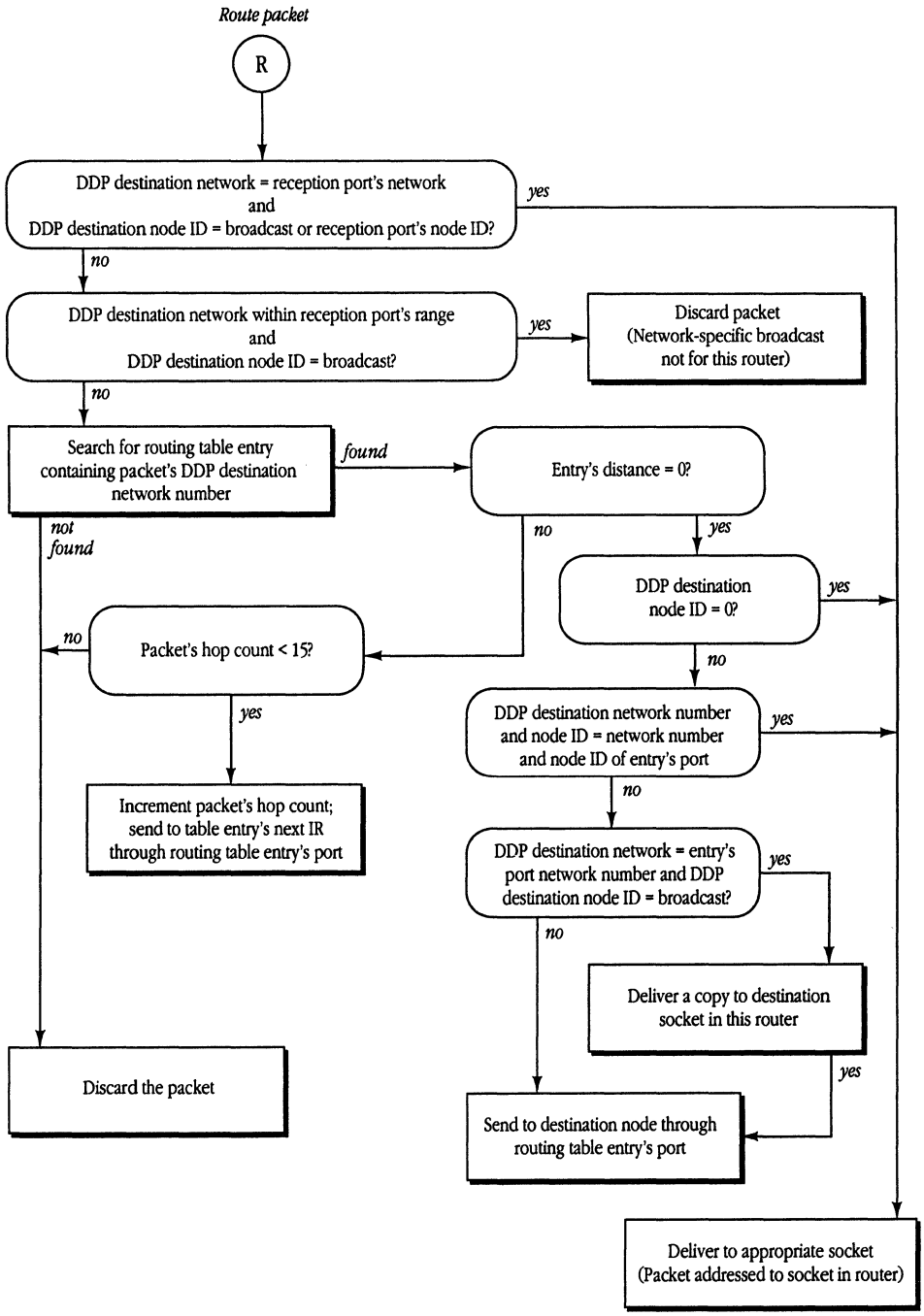


DDP routing details

The routing algorithm for *all* routers has changed (particularly on extended networks). *Figure 4-4* illustrates these changes.

As stated above, all routers must accept datagrams received for a node ID of zero if they are connected to the network indicated in the datagram. In addition, a router on an extended network should discard any broadcast packet that would otherwise be forwarded out through the same port that it was received. This can be done by examining the destination network number of datagrams whose destination node ID is $FF. If the destination network is within the network number range of the network on which the datagram was received, the datagram should be discarded; otherwise it should be forwarded. (If the destination network is zero or equal to THIS-NET, the packet is delivered internally and not forwarded.)

■ **Figure 4-4** Datagram routing algorithm for a router



Selecting the best router

It is desirable for nonrouting nodes to send each datagram to the router yielding the shortest route to the datagram's destination network—the "best" router. Although the current specification does not require this, an optional strategy follows for implementing a "best routing" DDP algorithm in nonrouting nodes. This strategy requires a capability in nonrouting nodes to send a packet to a specific data-link level address.

When a packet arrives from an off-network node, DDP reads the data-link–level source address. This is the data-link address of the last router on the route from the datagram's originating network. This router should generally be the optimal "next router"—in terms of hops—in the route back to that network. DDP maintains a cache of "best routers" for networks that have recently been heard from and sends datagrams to those routers for forwarding to those networks. If there is no cache entry for a network, DDP sends the datagram to A-ROUTER, expecting that a response packet will provide the information necessary to make an entry in the cache.

Note that a node's best-router cache entries need to be aged fairly quickly so that when a router goes down, an alternate route can be adopted (if one is available) before connections are broken.

The RTMP stub

The following changes are specified for the RTMP stub in nodes on an extended network.

Aging router information

The router aging time is reduced to 50 seconds. When a router is aged out, the RTMP stub must expand THIS-NETWORK-RANGE back to 1–$FFFE and set the node's zone name back to "*". It should also delete the node's zone multicast address, if one has been set.

Processing incoming RTMP packets

If the node's A-ROUTER parameter is nonzero (that is, the node knows about a router), the RTMP stub should accept an RTMP packet only if the network range indicated by the packet exactly matches the node's values for THIS-NETWORK-RANGE. Thus, if a node's network range were 3-4, it would not process a packet indicating a network range of 3-6.

If, however, A-ROUTER is zero (the node is *not* aware of any router), the RTMP stub should accept an RTMP packet if the packet indicates a range within which the node's network number is valid. If the node's network number is not within the range indicated by this tuple, the RTMP packet is rejected (see "When a Router First Comes Up" in Chapter 3). When an RTMP packet is accepted, THIS-NETWORK-RANGE is set to the value indicated by the packet and A-ROUTER is set to the sender's 24-bit AppleTalk address.

Note that this scheme allows a network range to be expanded without restarting all the nodes on that network. If a range is originally 3-4, but the routers are reconfigured to set it to 3-6, nodes will first age out the 3-4 value of THIS-NETWORK-RANGE (since none of the routers would be sending it) and then replace it with the range 3-6 obtained from subsequent RTMP packets.

RTMP

The following changes are specified for implementing RTMP.

Packet formats

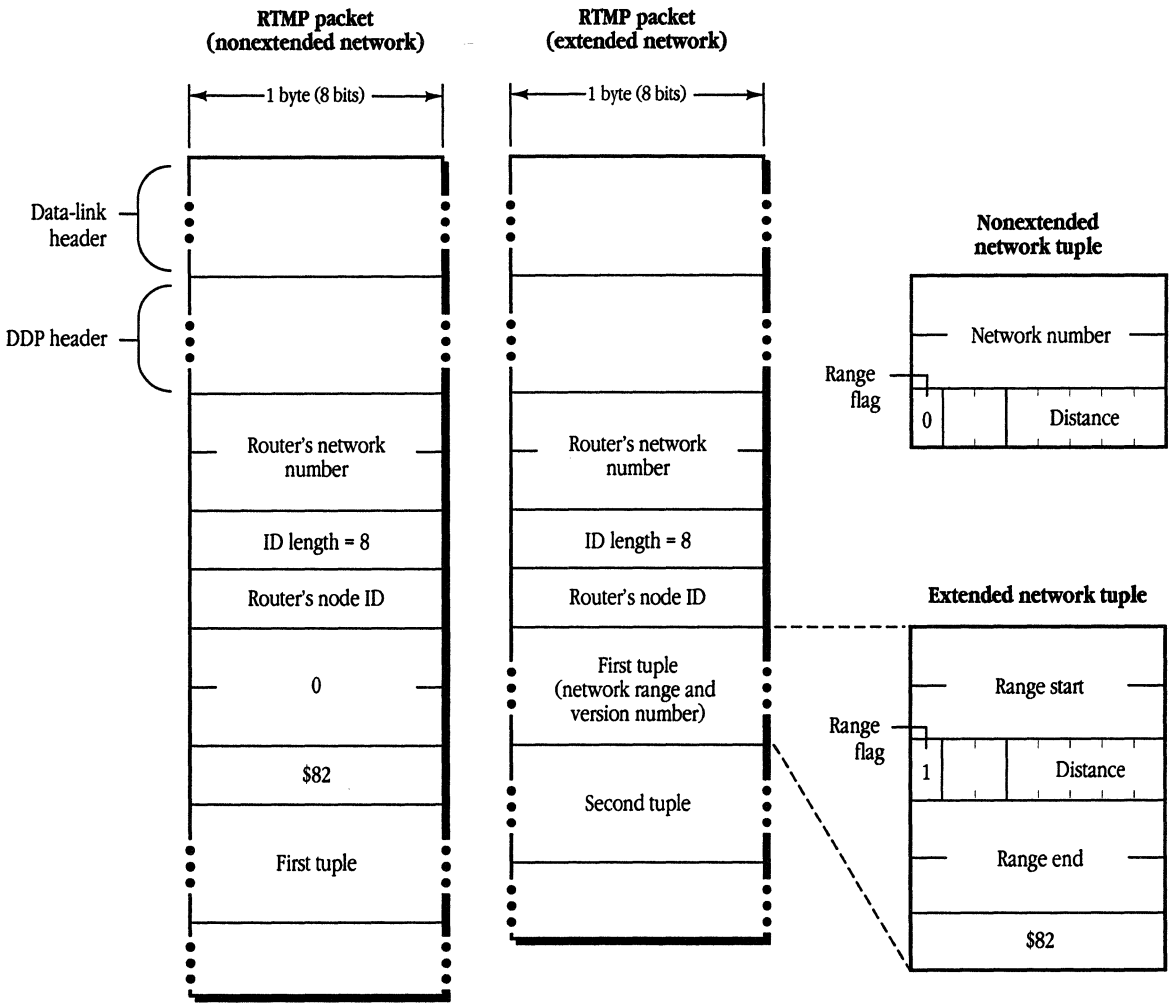
RTMP data packets on both extended and nonextended AppleTalk networks can now contain tuples in two forms: a network number or a network range. The first tuple in a packet sent on an *extended* network is the range for that network, allowing nodes that receive this packet to determine their network range. This tuple may be repeated later in the packet. Figure 4-5 illustrates RTMP packet formats for extended and nonextended networks.

RTMP data packets on extended networks are sent to destination address $0000FF using a network-wide broadcast. The "sender's network number" field always indicates the high 16 bits of the sender's address. The "sender's node ID" field indicates the low 8 bits of that address. *On both extended and nonextended networks,* each tuple has a flag indicating whether it is a 6-byte tuple (range start, distance, range end, and a version number byte) or a 3-byte tuple (network number, distance). A tuple for an extended network with a range of one (for example, 3–3) should still be sent as a network range tuple.

The RTMP header of packets on both extended and nonextended networks has been expanded to include additional information:

- On an extended network, the header includes a tuple that indicates the network range of this network. This tuple contains a distance of zero and the version number of RTMP being used, which for AppleTalk Phase 2 is $82.
- On a nonextended network, the header includes 2 bytes of zero (to keep tuple alignment the same as AppleTalk Phase 1 RTMP packets), followed by the version number of $82.

■ **Figure 4-5** RTMP packet formats



RTMP request packets remain the same as in AppleTalk Phase 1 (however, they can now be used to request three types of responses, as described later in this chapter). RTMP responses on extended networks should include the initial network range tuple.

Maintaining routing tables

Routers must maintain their routing table on a range-by-range basis for extended networks. Each range has a list of zone names associated with it (the order of zone names in this list is unspecified). Entries for nonextended networks should contain only a single network number, so that 3-byte tuples can be sent out for these entries.

The routing table update process must be extended to handle network range mismatches. For example, an incoming RTMP packet might contain a tuple indicating network range 3–57, while the router's current routing table might contain an entry for range 2–58 (or 2–3, or 22–33). Such overlapping network ranges clearly indicate an internet configuration error.

In order to minimize the effect on the internet of a misconfigured router being brought on-line, the following is specified: If an incoming RTMP packet contains a tuple for which there is no *exact* match in the routing table (that is, if there is no entry with the same starting *and* ending network numbers), but that tuple does overlap with part of some entry's range, that tuple should be disregarded. (In evaluating network range matches, a nonextended network number should be considered as a range of one, for example, 3–3.)

Split horizon

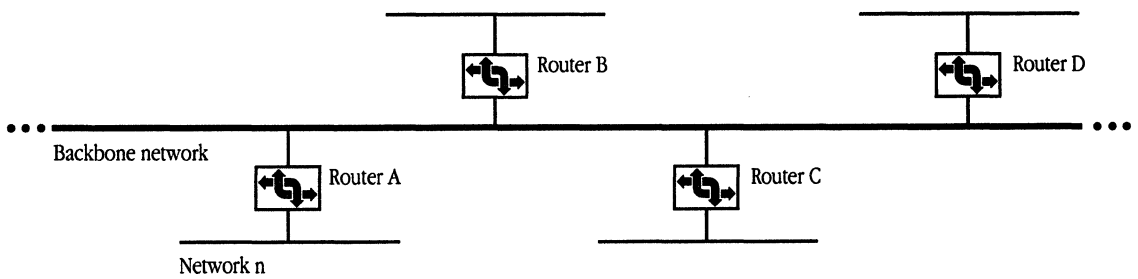
AppleTalk Phase 2 specifies a technique known as **split horizon** for use in sending routing tables on both extended and nonextended networks. This technique significantly reduces internet traffic caused by large numbers of routers exchanging their routing tables.

Split horizon reduces the number of redundant routing table entries exchanged by routers. As illustrated in *Figure 4-6,* router B need not transmit the tuple for network n to router A, since router A itself is on the path to that network. Similarly, it is unnecessary for router B to transmit this information to other neighboring routers (C, D), as they will acquire it from router A. In such a case, particularly on a backbone, most of the routing table need not be sent out at all.

To implement split horizon, a simple modification is made to the RTMP routing table sending algorithm: In a routing table sent out a given port,

- omit all entries for which the next router is on the network connected to that port
- include information for directly connected networks (for example, router A must transmit the tuple for network n out *both* its ports)

- **Figure 4-6** A scenario for split-horizon processing



Notify neighbor

RTMP now includes a new aging procedure called **notify neighbor.** Under notify neighbor, whenever a routing table's entry state is "Bad," instead of omitting that entry from the broadcasted routing table, the entry is sent *with a distance of 31.* This entry indicates to receiving routers that the entry's network is no longer reachable through the sending router, and that an alternate route, if available, should be adopted. This entry is sent only if it would not otherwise be eliminated by split-horizon processing.

Upon receiving a tuple with an entry of distance 31, *if the router sending that tuple is the forwarding router for the associated network,* a router should immediately set the state of that entry to "Bad." In this way, the fact that a network is unreachable will propagate at the RTMP table sending rate. Note that, to minimize the effect of lost packets, the time for the "Bad" state is extended from one Validity timer to two Validity timers (from 20 to 40 seconds).

Routers that are knowingly going down or deactivating a port should, as a courtesy, use "notify neighbor" to inform other routers of this fact.

RTMP requests

The RTMP request has been extended to allow for a variation called the **RTMP Route Data Request** or **RDR.** The RDR is useful for routing and network management purposes to obtain routing information from any router *on demand.*

The different types of RTMP request packets are distinguished by the value of the Command field, which is 1, 2, or 3:

- A Command field of 1 indicates a Network Information Request. (This is equivalent to the former RTMP Request.)
- A Command field of 2 indicates a Route Data Request using normal split-horizon processing.
- A Command field of 3 indicates a Route Data Request *without* using split horizon, causing the entire table to be returned.

Upon receiving an RDR, a router responds by sending its routing data *directly to the source socket* of the requesting node. (Normally, broadcasted RTMP data packets are sent to the well-known RTMP socket, 1—not to the requester's source socket.) An interested node can therefore obtain routing information by opening a socket and sending an RDR to a router through this socket. Note that a full response can consist of a number of packets, and that a tuple can never span two packets.

NBP

NBP packet formats remain unchanged from AppleTalk Phase 1. However, several protocol changes are specified. NBP in a routing node must change *even if the router is connected only to nonextended networks.*

The THIS-ZONE variable

NBP in nodes on extended networks must maintain a variable known as THIS-ZONE. In nonrouting nodes on extended AppleTalk networks, NBP must now verify that a LkUp packet is intended either for the zone to which the node belongs or for zone "*"; that is, NBP must match object, type, *and zone.* This is because more than one zone may correspond to the same zone multicast address (in fact, if the data link does not support multicast, all zones will correspond to the same zone multicast address, the AppleTalk broadcast address).

Broadcast requests and forward requests

NBP on an *extended* network should never send out a BrRq for zone "*" (the router has no way of determining the node's zone, since there is no correspondence between network number and zone name). However, in the absence of a router, NBP should always broadcast LkUps to zone "*".

A router receiving a BrRq must first change any zone name of "*" (received from a nonextended network) to the node's actual zone name, and then change the packet type from BrRq to a new type, FwdReq (forward request). The value in the control field for a FwdReq packet is 4.

The router then uses DDP to send a FwdReq packet to each network whose zone list contains the desired zone. These request packets are sent to the NIS at address $nnnn00, where nnnn is the first network in each range.

The packet proceeds through the internet until it reaches a router directly connected to the intended destination network, where it is passed to the NBP process. The NBP process then changes the packet type to LkUp and sends this packet as a zone-specific multicast on the intended network. Specifically,

- If the packet is intended for an extended network, NBP changes the destination address to $0000FF, determines the multicast address associated with the intended zone, and calls DDP to send the packet to that zone multicast address through the appropriate port.
- If the packet is intended for a nonextended network, NBP simply changes the destination address to $nnnnFF and broadcasts the packet on that network.

A router on *nonextended* networks that receives a BrRq for zone "*" and has not yet discovered the zone name associated with the sender's network should broadcast a LkUp packet on that network with a zone name of "*". It should not, however, send out any FwdReq packets.

◆ *Note:* A router receiving a BrRq for a zone that is in the zone list for one or more of the networks to which the router is directly connected should not send out FwdReq's for these networks. It should instead send out LkUp's to the appropriate zone multicast addresses.

Special characters

AppleTalk Phase 2 disallows the use of a character with value $FF as the first byte in an NBP object, type, or zone string. This value is reserved for future flexibility.

NBP has also been enhanced to provide additional wildcard support. The character ≈ ($C5) is now reserved in the object name and type strings and is used in a lookup to mean "a match of zero or more characters." Thus

- ≈abc matches abc, xabc, xxxabc, and so on
- abc≈ matches abc, abcx, and so on
- abc≈def matches abcdef, abcxdef, and so on

At most one ≈ is allowed in any one string. As a single, standalone character, ≈ has the same meaning as a single =, which must also continue to be accepted. The ≈ character has no special meaning in zone names.

ATP

No changes in ATP are necessary to support the architectural changes in AppleTalk Phase 2. Clients using and implementing ATP will continue to work under AppleTalk Phase 2 as they did under AppleTalk Phase 1. However, the following change has been made to provide additional flexibility in ATP Exactly Once (XO) service.

There are 3 unused bits in the Command field of the ATP header. For XO request packets only, these 3 bits are used as an indicator of the length of the TRel timer that the transaction responder should use. In this way, the requester can indicate to the responder an approximate measure of how long to wait for the TRel. The values in the following list are specified.

Value	Timer
000	30 seconds
001	1 minute
010	2 minutes
011	4 minutes
100	8 minutes

Other values for the XO request command field are reserved.

An ATP requesting client could use the echo protocol or other means to estimate the time for a TRel to be received and set this timer value accordingly. This value also depends, however, on the client's retransmission rate: If retransmitting slowly, the TRel timer should be set higher, in case a retransmitted request is lost. Certain session level protocols, such as PAP and ASP may wish to set this timer value to the value of their connection timer. Clients of ATP must be aware that nodes running AppleTalk Phase 1 may not process this new timer feature correctly.

ZIP

ZIP includes several additional functions as well as new packet formats. Two new ZIP requests are used on extended networks in AppleTalk Phase 2:

- ZIP GetNetInfo is sent out by a node upon starting up to determine its network range and zone multicast address.
- ZIP GetLocalZones is used by a node to acquire the list of zone names that are valid for the node's network.

ZIP GetNetInfo and NetInfoReply

ZIP GetNetInfo is an example of a *port-dependent* request. Like the RTMP request packet, such a packet requests information associated with a specific port—the port through which a packet is received.

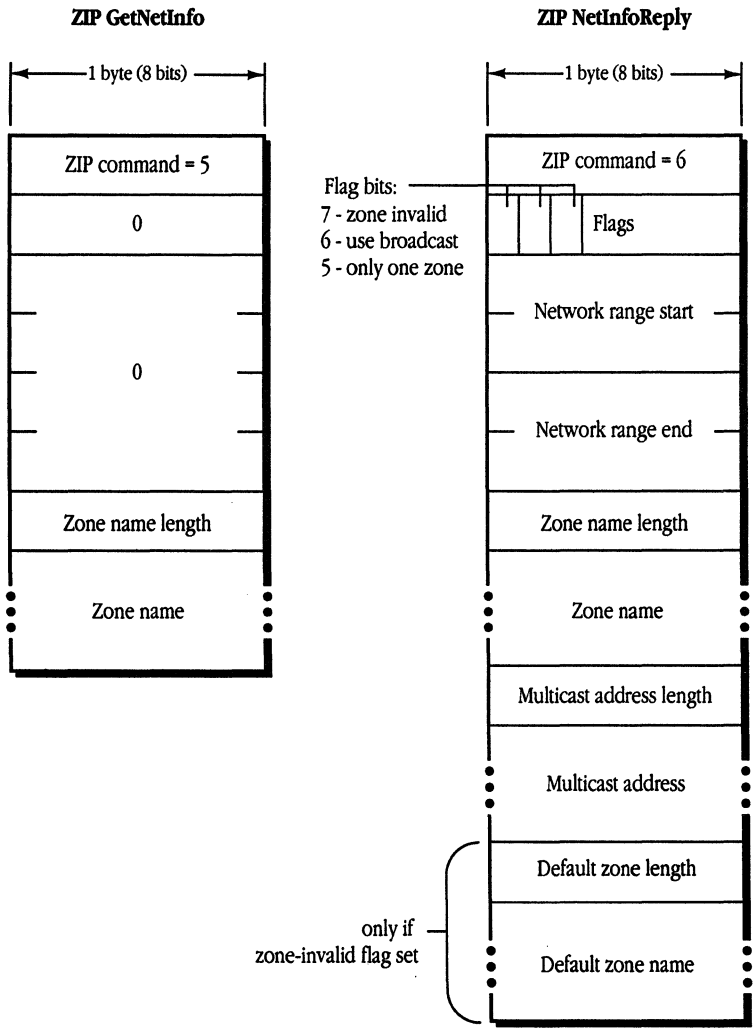
ZIP GetNetInfo is generally sent as a network-wide broadcast to the ZIS when a node is first started up. If the node has a saved A-ROUTER value, the GetNetInfo request can be sent to this router. The response, ZIP NetInfoReply, is generally directed to the requesting node and socket.

If, however, the request was broadcasted network-wide and its source network number is invalid for the port through which it is received (and it is not in the startup range), the response is broadcasted back network-wide through the requesting port. This allows the router to communicate with a node that has started up with an invalid network number saved in pRAM.

In either case, the response contains the network range followed by a copy of the zone name from the request (so that the response can be matched with the request). This is followed by the zone multicast address for the node. If the request contained a valid zone name, this address is the zone multicast address for that zone. If the zone name was either invalid or NIL, this address is the zone multicast address of the default zone for that network, followed by the default zone name. (The default zone name for an extended network is set as part of the seed information in a router's port descriptor for the network. It must be one of the zones in that network's zone list. Nonseed routers discover this information as part of their startup process through a GetNetInfo request.)

As shown in *Figure 4-7,* the NetInfoReply also contains a number of flags, which can indicate that the requested zone name is invalid, that there is only one zone name for the network (in which case a GetLocalZones request is not needed), and that the data link does not support multicast (in which case the multicast address length in the packet will be set to zero).

■ **Figure 4-7** ZIP GetNetInfo and NetInfoReply packet formats



Assignment of zone multicast addresses

Upon receiving a ZIP GetNetInfo request, the ZIP process in a router verifies that the specified zone name is valid for the network. If it is, ZIP obtains the associated multicast address and returns it in the reply.

To obtain the zone multicast address, ZIP first converts the zone name to uppercase (since zone names are case-insensitive). This conversion function is specified in *Inside Macintosh* and also *Inside AppleTalk* (Appendix D). The router then obtains a number, *h*, in the range 1–$FFFF, associated with this zone name by performing the DDP checksum algorithm (as documented in *Inside AppleTalk)* on each byte of the zone name (excluding the length byte). This number *h* is passed to the data link, which is assumed to provide *n* multicast addresses: a_0 through a_{n-1}. (Addresses for EtherTalk and TokenTalk are specified in Table 4-1.) The multicast address for that zone, returned by the data link, is $a_{(h \bmod n)}$.

ZIP Query and Reply

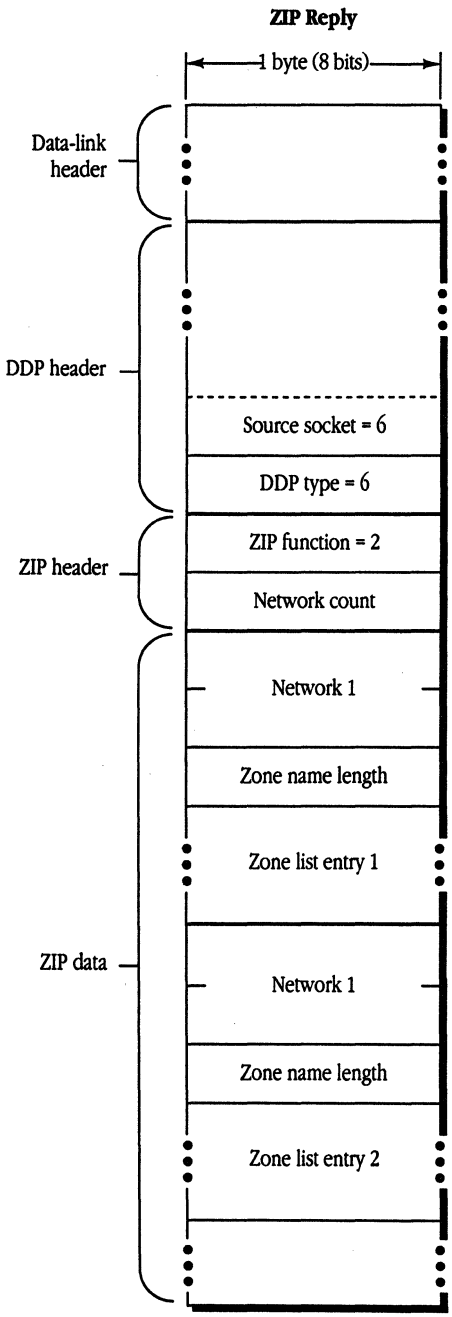
ZIP Query packets remain unchanged from AppleTalk Phase 1. The Network Number field for an extended network is set to the first network number in that network's range. Multiple networks can still be included in one packet (both extended and nonextended networks can be mixed).

ZIP Reply packets remain exactly the same as previously—in cases where a network's zone list will fit in one packet (17 maximum size zone names, or 36 zone names with an average length of 16 bytes). However, an extended network can now have more zones than can fit in a single packet. An extended network's zone list is indicated by multiple entries. In each entry, the Network Number field is set to the first network number of the range for that network. Response packets again can contain multiple network entries, provided that an entry is completely contained within the packet. *Figure 4-8* shows the Zip Reply packet format.

If a network's zone list *cannot* fit in one ZIP response packet, a series of new packets are returned. These packets, called **Extended ZIP Reply** packets, have a new ZIP command byte (with a value of 8). Their format is the same as a ZIP Reply packet, except that the Network Count field has a new meaning. Instead of indicating the number of network/zone tuples in the packet (which can be determined by reading entries until the packet ends), this field indicates the total number of zones for the extended network. It will be the same for each Extended ZIP Reply packet for a given network and has a maximum value of 255. The queried router sends as many of these packets as is necessary. The querying router collects all the responses and can determine whether any have been lost. If any are lost, all the information must be requested again (the router must send another ZIP Query for that network). A router may use an Extended ZIP Reply packet even for a network whose zone list does fit in one packet; in this case, only one network's zone list can be sent in the packet.

Note that until a router has all the zone information for a given network, it must respond to other routers' ZIP Queries for that network as if it had *none* of the information.

■ **Figure 4-8** ZIP Reply packet format



ZIP ATP requests

ZIP GetZoneList, which uses ATP, remains unchanged from AppleTalk Phase 1, but must be sent to the full 24-bit A-ROUTER address on extended networks. (GetZoneList requests that require multiple ATP transmissions should all be sent to the *same* A-ROUTER address.)

ZIP GetMyZone should not be sent on an extended network, since the node already knows its zone name (and the router could not determine it from the node's address). ZIP GetLocalZones is used by nodes on an extended network to acquire the network's zone list. The ZIP GetLocalZones packet is nearly identical to ZIP GetZoneList; however, this packet contains a command byte of 9 in the ATP header. The same algorithms used in GetZoneList apply for obtaining the network zone list. A router on a nonextended network will respond with a single-zone reply.

Changing zone names

On an extended network, nodes must be made aware of changes to the name of the zone in which they reside. Nodes may also need to be given a new zone multicast address.

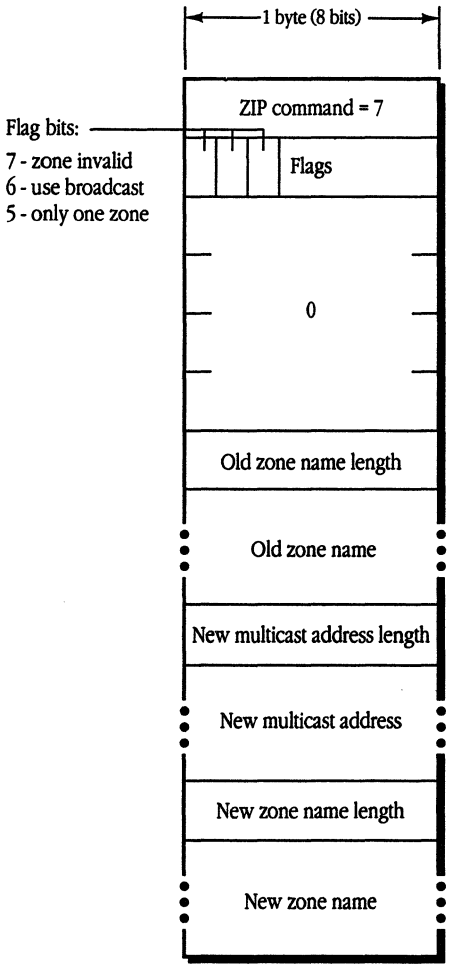
A node on an extended network maintains a ZIP stub on the ZIS. A node's ZIP stub listens for a new ZIP Notify packet, which indicates a change of zone name. The ZIP Notify packet is shown in *Figure 4-9.*

The ZIP Notify packet, which is sent to the ZIS using a zone-specific broadcast, contains the old zone name, the new zone name, and the new zone multicast address. This packet is very similar to a ZIP GetNetInfo reply; however, the ZIP command code is 7, the "zone invalid" flag is never set, and the network number fields are unused.

A node receiving such a packet must check to see whether it is in the zone being changed. If so, the node must change THIS-ZONE and its copy in long-term storage accordingly, delete the old zone multicast address, and register on the new zone multicast address.

The implementation of the ZIP stub and the processing of ZIP Notify packets are optional. However, following a zone name change, NBP names on nodes not implementing ZIP Notify will not appear in the new zone until the node's AppleTalk implementation is reestablished.

■ **Figure 4-9** ZIP notify packet format



◆ *ZIP takedown and bringup:* This document does not specify the method by which zone names associated with active networks are actually changed. ZIP takedown and bringup are not a part of AppleTalk Phase 2, and such packets should be ignored by all AppleTalk Phase 2 routers.

Changing a zone name for a given network involves not only informing the routers (and other nodes) connected to that network, but also informing every router on the internet of that change. AppleTalk Phase 2 removes this function from ZIP and delegates it to network management protocols, to be documented elsewhere. (This process can also be performed by shutting down all routers connected to a network, reconfiguring the seed routers, and then restarting all routers.)

Appendix **Changes in LocalTalk Nodes**

THIS APPENDIX LISTS changes that can be made to nonrouting implementations on LocalTalk to fully conform to AppleTalk Phase 2. While they are not currently required in LocalTalk nodes, future products may require these changes to provide full functionality to LocalTalk nodes.

Each of these changes is described in prior sections of this document.

- Additional NBP wildcard character ≈
- ATP TRel timer
- A-ROUTER aging time of 50 seconds
- "Best router" address cache in nodes ■

The Apple Publishing System

This Apple® manual was written, edited, and composed on a desktop publishing system using Apple Macintosh® computers and Microsoft® Word. Proof pages were created on the Apple LaserWriter® printers; final pages were printed on a Varityper® VT600™. Line art was created using Adobe Illustrator™. PostScript®, the LaserWriter page-description language, was developed by Adobe Systems Incorporated.

Text type and display type are Apple's corporate font, a condensed version of ITC Garamond®. Bullets are ITC Zapf Dingbats.