5. Link State

Each router extracts certain information from the Adjacency database, adds locally derived information, and constructs a Link State Packet (LSP) that describes its immediate neighbors. The totality of these LSPs constructed by all the routers in the routing area make up the Link State database for the area. The NLSP™ specification aims for each router to maintain a copy of the Link State database and to keep these copies synchronized with each other.

5.1. Overview of the Protocol

The Link State database is synchronized by propagating LSPs reliably from a router that observes a topology state change throughout the routing area. The format of an LSP is described on page 5-29.

If an LSP is too large to fit in lspBufferSize bytes, the router originating it splits the information into several LSPs, each identified by an LSP Number (think of this as a fragment number). Each of these LSPs propagates from the originator independently of the others. When there is a topology change, only the LSP (that is, the fragment) affected by the change is propagated. Each has an independent sequence number (incremented when the LSP content changes) to allow a recipient router to discriminate between new and outdated information.

Each fragment is complete and self-contained. From here onward the term *LSP* refers to one of these fragments, because this is the basic unit of information propagated in the routing area. When referring to the set of fragments originated by a router, the term *LSP series* is used.

There are two parts to the propagation method: flooding and receipt confirmation.

Flooding is instigated when a router detects a topology state change. The router constructs a new LSP and transmits it to each of its neighbors. On a WAN, this is a directed packet; on a LAN, it is multicast. Upon receiving an LSP, a router first decides whether it is newer; that is, whether the sequence number is higher than the one in its current copy of the database. (Other factors enter into the newness decision; complete details follow later.) If it is new, the router retransmits the LSP to all its neighbors except on the circuit over which the LSP was received.

Receipt confirmation is different for LANs and WANs. A router receiving an LSP on a WAN replies with a Partial Sequence Number Packet (PSNP). This serves as an acknowledgment for the LSP (or LSPs) and sequence numbers identified in the PSNP.

On a LAN, there is no explicit acknowledgment. A router multicasts the new LSP and (if all goes well) the other routers on the LAN receive the LSP and absorb it into their databases. To allow for the possibility that the LSP was not successfully received by all the LAN's routers, the Designated Router periodically multicasts a Complete Sequence Number Packet (CSNP) containing all the LSP identifiers and sequence numbers it has in its database for the entire routing area. (It does not send the LSPs; it sends just enough information to allow another router to detect whether it is out of step with the Designated Router.) If all has gone well, the other routers on the LAN verify that they have the same set of LSPs and sequence numbers indicated in the CSNP.

If there is a discrepancy, what happens next depends on which router has the higher sequence number (that is, newer information) for the LSP: the Designated Router or the CSNP-recipient router. If the Designated Router has newer information, the

CSNP-recipient transmits a PSNP, which acts as a request for the missing LSPs. The Designated Router receiving the PSNP replies by multicasting each requested LSP. If the CSNP-recipient has newer information, it multicasts the up-to-date LSP on the LAN, and the Designated Router updates its database (as do the other routers on the LAN).

The remainder of this chapter describes the algorithms in more detail. A number of timing considerations govern when and how often to send certain packets. To accommodate these considerations, it is convenient to describe operation of the protocol in terms of two flags. For each LSP and for each circuit over which routing messages are exchanged, a router maintains two local flags:

- *Send Routing Message* (SRMflag): When set, this flag indicates that LSP should be sent on that circuit. On a LAN, SRMflag is cleared as soon as the LSP is transmitted. On a WAN, it is cleared only on receipt of a PSNP (acknowledging the LSP) or an LSP (indicating that the neighbor was already at least as current as this router). A router regularly scans the Link State database for LSPs for which

  a) SRMflags are set

  b) The LSP was propagated no more recently than minimumLSPTransmissionInterval

  When such an LSP is found, the router transmits it on all circuits having SRMflags set, and updates the lastSent time for it.

- *Send Sequence Number* (SSNflag): When set, this flag indicates that information about that LSP should be included in a PSNP sent on that circuit. When the PSNP has been sent, the flag is cleared.

5.2. Generating and Checking the LSP Checksum

The checksum allows detection of transmission errors and memory corruption, to prevent both (a) the use of incorrect routing information, and (b) propagation of incorrect routing information to other routers. The router originating an LSP computes the checksum when the LSP is generated. The checksum is never modified by any other system as the LSP propagates.

The checksum is computed over all fields in the LSP after the Remaining Lifetime field. This field (and those before it) are excluded so that the LSP can be aged by systems without recomputing the checksum.

This specification makes use of the checksum function defined in ISO 8473 (Reference [ISO88]). The following subsections define the computation.

5.2.1. Symbols and Conventions

`C0` and `C1` are variables used in the computation.

`i` is the number (that is, the position) of a byte in the data block having the checksum performed; the first byte has `i=1`.

`B[i]` is the value of byte `i` in the block being checked.

`L` is the length of the data block being checked, in bytes.

`X` is the first byte of the calculated checksum.

`Y` is the second byte of the calculated checksum.

Addition is performed in one of the following modes:

- Modulo 255 arithmetic
- Eight-bit one's complement arithmetic in which, if any of the variables has the value minus zero (that is, 255), it is regarded as though it had the value plus zero (that is, 0)

5.2.2. Generating a Checksum

To perform a checksum on an LSP being generated, proceed as follows:

```
Construct the complete packet, with the value of the Checksum field set
to zero.

C0 ← C1 ← 0.

For i incrementing by 1 sequentially from 1 to L:

       BEGIN   -- Loop

       C0 ← C0 + B[i]

       C1 ← C1 + C0

       END    -- Loop

Calculate:

       X ← ( L - 8 ) × C0 - C1    (mod 255)

       Y ← ( L - 7 ) × ( - C0 ) + C1    (mod 255)

If X = 0 then X ← 255

If Y = 0 then Y ← 255

Place the values of X and Y in the Checksum field of the packet.
```

There is one exception: if the Remaining Lifetime field of the LSP is zero, the correct value of the checksum is zero.

5.2.3. Checking a Checksum

To verify the checksum of an arriving LSP or an LSP stored in memory, proceed as follows:

```
If either byte (or both bytes) of the Checksum field is zero,

     then the checksum is incorrect; do not proceed further.

C0 ← C1 ← 0

For i incrementing by 1 sequentially from 1 to L:

       BEGIN   -- Loop

       C0 ← C0 + B[i]

       C1 ← C1 + C0

       END    -- Loop

If C0 = C1 = 0 (mod 255) when all bytes have been processed,

     then the checksum calculation has succeeded.

Otherwise, the checksum calculation has failed.
```

There is one exception: if the Remaining Lifetime field of the LSP is zero, the correct value of the checksum is zero.

5.2.4. Partial Precomputation

The portion of the LSP covered by the checksum starts with the Source ID portion of the LSP ID field.

Partial Precomputation is an extra fault-tolerance method that a router can include as an implementation option. It is described in the next paragraph. It protects against a failure in which a router's memory of its own systemID becomes corrupted. Even if very rare, such a failure could cause widespread disruption if not detected. Partial Precomputation contains the failure by causing LSPs generated by the failing router to have checksum errors.

An originating router precomputes the first few steps of the checksum and saves the result for use with each LSP. Specifically, it precomputes the checksum of the systemID portion of the Source ID when the router is activated or whenever the systemID changes. When performing a checksum on a packet, then, it initializes the variables `C0` and `C1` to the saved values, and resumes the computation after the systemID portion of the Source ID field; that is, starting with the Pseudonode ID portion of the LSP ID.

5.3. The Need for Multiple LSPs

Packets for LSPs are limited in size to lspBufferSize. It might not be possible to include all the information about one's neighbors in one packet. In such cases, a router uses multiple LSPs to convey this information. (Even if the information can fit into one packet, a router can optionally use more than one LSP to convey it.) Each LSP of the set carries the same Source ID, but sets its LSP Number individually. Each of the several LSPs is propagated independently in the routing area, allowing pipelining of activity. On the other hand, the Decision Process that builds the Forwarding database recognizes that they constitute a particular router's LSP series because they have the same Source ID.

When some event requires changing the LSP information for a router, that router reissues the one (or more) LSPs that have different contents. It is not required to reissue unchanged LSPs.

The first LSP of a series has LSP Number zero. It is treated in a special way.

a) The following fields are meaningful to the Decision Process only when they are present in the LSP number zero:

   i) The setting of the LSP Database Overload bit.

   ii) The value of the IS Type field.

   iii) The Attached Flag.

   iv) The Area Addresses option field. This option is present only in LSP zero.

   v) The Multi-homed Non-routing Server (NR) field.

   vi) The Management Information option field. This option is present only in LSP zero, and must precede any Link Information options.

   Note: There must be exactly one Area Addresses option and one Management Information option in each LSP number zero. If a router receives one of these options in an LSP with a nonzero

number, or if a router receives a second instance within LSP zero, these additional options are ignored.

b) When any of the preceding items are changed, a router reissues LSP zero to inform other routers of the change. Other LSPs need not be reissued.

Once a particular Option field has been assigned to a particular LSP Number, it is desirable (but not required) that it not be moved to a different LSP number. This is because moving an Option field from one LSP to another can cause temporary loss of connectivity to the entity (for example, a neighbor, service, or external route) represented by the field. This can occur if the new version of the LSP that originally contained the field (but which does not now contain it) is propagated before the new version of the other LSP (which now contains the field).

Note: In most situations, it improves performance to ensure that the number of LSPs generated by a router is close to the optimal number that would be required if the LSPs were densely packed with Link Information options. This can be accomplished by reusing space in LSPs with a lower LSP Number for new adjacencies. Fewer LSPs means consuming less processing capacity, less network bandwidth, and less router memory. However, there are situations in which a link alternates often between active and inactive. Putting such a link in a small LSP consumes less network bandwidth when flooding the LSP to other routers each time it changes. Consequently, implementers are faced with a trade-off on this issue.

If an Option field moves from one LSP to another, the SRMflags of the two updated (or new) LSPs must be set as an atomic action.

Note: If they are not set atomically, a race condition exists, in which one of the two LSPs can propagate quickly, while the other waits for an entire propagation cycle. If this occurs, entities are erroneously eliminated from the topology and routing can become unstable for a period of time potentially as large as maximumLSPGenerationInterval.

5.4. Determining Which of Two LSPs is Newer

At certain points in the algorithms described later, a router has two copies of an LSP with the same LSP ID and must compare the two to determine which is newer, or whether the two are the same. At certain other times, the router does not have an LSP in hand, but has information about a neighbor's copy of an LSP, and must make the comparison with its own copy of the LSP. (The information arrives in sequence number packets from the neighbor.) These are the steps in making the comparisons:

a) If the two have different sequence numbers, the higher sequence number is considered newer.

b) If

   i) the sequence numbers are the same and

   ii) exactly one of the two has zero lifetime,

   the one with zero lifetime is considered newer.

c) If

   i) the sequence numbers are the same and

   ii) the remaining lifetimes are either both zero or both nonzero and

   iii) the checksums match,

the two are considered the same.

d) If, on the other hand,

i) the sequence numbers are the same and

ii) the remaining lifetimes are either both zero or both nonzero **but**

iii) the checksums **do not** match,

the LSPs are considered different, and the one with the higher numerically valued checksum (treated as a four-byte unsigned number) is considered to be newer. This is the *Preference of Checksums* rule.

Note: If two routers are misconfigured with the same systemID, and they both generate LSPs with the same sequence number, this provision at least ensures that the LSPs are treated in a consistent way by all routers.

5.5. Pseudonodes and Designated Routers

There are two types of pseudonodes: *LAN Pseudonodes* and *WAN pseudonodes*.

LAN pseudonodes have been discussed already. They streamline NLSP operation when many routers are attached to the same LAN. There is an election process by which one of the routers is selected to be the Designated Router representing the LAN pseudonode for LSP generation.

Every NLSP router on a LAN retains in its local memory all the information needed for it to become the Designated Router, should the need arise. This local information is referred to as a *Potential Pseudonode*.

WAN pseudonodes serve a different purpose. Section 3 discusses methods by which an NLSP router can establish communication with a RIP router over a WAN (virtual) circuit. For each such circuit, a pseudonode represents the RIP routes and SAP services discovered over that circuit. The NLSP router on the circuit is the WAN Designated Router representing the WAN pseudonode. There is no election process.

Section 3 actually specifies two ways of using RIP: numbered and unnumbered. The first assigns an IPX™ network number to the circuit; the second does not. Depending on this choice, there is either a *Numbered WAN Pseudonode* for the circuit or an *Unnumbered WAN Pseudonode*. The latter appears if any unnumbered router-router protocol is used on the link; for example Unnumbered RIP or On-Demand Static Routing.

Figure 5-1 illustrates the three kinds of pseudonodes.

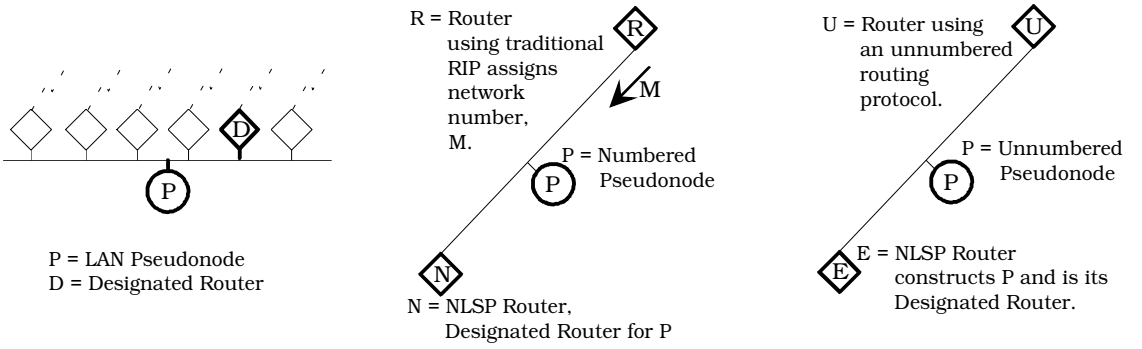


Figure 5-1: Pseudonodes

As in the numbered pseudonode case, Unnumbered RIP uses RIP and SAP on the WAN, not NLSP. As far as the traffic on the WAN is concerned, Unnumbered RIP operation really has no connection with NLSP. If—like node E in the figure—either or both ends of the (virtual) circuit are NLSP routers, that router constructs an unnumbered pseudonode (P) representing the WAN to the NLSP routing area (not shown in the diagram) to which it is attached. The same is true with On-Demand Static Routing, except that the network numbers and services reported in the pseudonode are configured statically, rather than learned by receiving periodic RIP and SAP packets on the link. To other routers in the NLSP routing area, an On-Demand Static Routing link is indistinguishable from an Unnumbered RIP link.

5.6. Aging Out an LSP and Purging Superseded LSPs

When the source generates an LSP, it sets the Remaining Lifetime field to maxAge.

When a router holds an LSP before successfully transmitting it to a neighbor, it decrements the Remaining Lifetime field by the holding time. Before transmitting an LSP to a neighbor, a router decrements the Remaining Lifetime field by at least one.

When the Remaining Lifetime of an LSP in memory becomes zero, the router takes the following actions:

a) Retains only the LSP header (that is, the fixed-length fields), and changes the Remaining Lifetime field to zero

b) Updates the Packet Length and Checksum fields to reflect the changes indicated in Step a)

c) Records the time at which the Remaining Lifetime for the LSP became zero

d) Purges the LSP header from the database when zeroAgeLifetime has expired since the Remaining Lifetime became zero

Any time a router purges an LSP with **nonzero** Remaining Lifetime, it performs the four steps just indicated, but the retention interval in Step d) is maxAge.

> **Note:** This could happen for one of several reasons; for example, (a) resignation as Designated Router, (b) compaction of LSPs to reduce their number, (c) receiving from another router one's own LSP from a prior incarnation (Section 5.12.3, Step b.2).

5.7. Periodic LSP Generation

The timer maximumLSPGenerationInterval governs periodic LSP generation. Within every time interval of length maximumLSPGenerationInterval, a router regenerates every LSP at least once. Jitter is applied, as described in Section 2. It is not required to synchronize regeneration of the individual LSPs.

The purpose of periodic LSP generation is to remove stale information from the routing area at a coarse time granularity. For example, suppose two large, usually disjoint campuses, A and B, temporarily link with each other using a switched network. While the link persists, the Link State information of each area flows into the other. Then the link is taken down. Fine-grained operation of NLSP does not remove B's link state information from routers in A. In case the link is reestablished, it would be costly to reacquire all that information. However, the two campuses might not link with each other again for a very long time. So, at a coarse time scale,

a method is needed to purge B's link state information from routers in A. The method is refreshing LSPs by periodic regeneration, and aging out LSPs that are not refreshed.

A second result of periodic regeneration and aging of LSPs is to handle the very rare event of an LSP corruption not being detected by the checksum check. Ultimately, the corrupted information is aged out of the LSP database.

5.8. Event-Driven LSP and CSNP Generation

A router generates an LSP when an event occurs that would cause the LSP content to change. Events that can cause such a change include

- An Adjacency or Circuit Up/Down event
- A change in a circuit's cost
- A change in manualAreaAddresses
- A change in a router's systemID or name
- A change of Designated Router on a LAN
- A change of network number
- A change in SAP services or external RIP routes
- A change in the waiting status (entering or leaving an overloaded state)

When such an event occurs, the router regenerates changed LSPs with a new sequence number. If the event necessitated generation of an LSP that had not been generated previously (for example, an Adjacency Up event for an adjacency that cannot be accommodated in an existing LSP), the Sequence Number is set to one. The router then propagates the LSPs on every circuit by setting the SRMflag for each circuit. The timer maximumLSPGenerationInterval is reset.

When the local router either becomes or resigns as the Designated Router for a LAN, it generates a lanL1DesRouterChange event. In addition, when it becomes the Designated Router, it

a) Generates and transmits Level 1 pseudonode LSPs

b) Purges the pseudonode LSPs issued by the previous Designated Router (if any)

When a router resigns as the Designated Router for a LAN, it purges the pseudonode LSPs as described on page 5-7.

There is a hold-down timer, minimumLSPGenerationInterval, on the generation of each individual LSP.

When a WAN circuit starts or restarts, that is, whenever an adjacency changes to the "Up" state

c) The router sets the SRMflag for that circuit on all LSPs

d) The router sends a complete set of CSNPs on that circuit

5.9. Generation of Level 1 Non-Pseudonode LSPs

Each router generates a Level 1 non-pseudonode LSP on its own behalf. The LSP describes the router itself and the router's adjacencies.

A Level 1 non-pseudonode LSP contains the following information in its variable-length fields:

- In the Area Address Mask option, the set of manualAreaAddresses for this router (that is, the source router for the LSP).
- In the Management Information option, configuration and version information about this router:
  - Network Number is the router's internalNetworkNumber.
  - Node Number is the router's internal node number (0x000000000001).
  - IPX Version is one.
  - Router/Server Name is a textual identifier of the router originating the LSP. If the system is also a NetWare® server, this field contains the server name. This is the same name that appears in the IPX WAN version 2 (IW2) Information Request/Reply packets.
- The External Routes option does not appear in non-pseudonode LSPs.
- The Services Information options contain information about services that reside on the same system as the router, and that advertise themselves using SAP. All such services must advertise themselves as residing on the internalNetworkNumber.
- An instance of the Link Information option for each neighboring router, formed from the following:

  a) The set of WAN NLSP adjacencies in the "Up" state

  b) The LAN pseudonode for each directly attached active LAN circuit

  c) The WAN pseudonode for each RIP/SAP circuit for which this router is the WAN Designated Router

  Fields of the Link Information option contain values as follows:

  - The cost of the link from this router's viewpoint is reported in the Cost field. There is a default cost, based on the link's data rate (throughput). Figure 5-3 shows how to derive the default Cost from the throughput of the link. For WAN media, the throughput value is determined empirically during the IW2 protocol exchange described in Section 3, and a router determines the default cost from that value using Figure 5-3. For both WANs and LANs, a network administrator can configure a nondefault cost manually.
  - The Neighbor ID for a LAN adjacency is the neighbor's circuitID (the neighboring pseudonode ID); that is, this value is the LAN ID of the LAN Level 1 Hello packets sent on that circuit.
  - The Neighbor ID for a WAN NLSP adjacency is the neighbor's neighborSystemID with appended zero byte (indicating non-pseudonode).
  - The Neighbor ID for a WAN pseudonode adjacency is the WAN Designated Router's (that is, this router's) systemID followed by a one-byte pseudonode ID, unique among all pseudonodes for which this router is Designated Router.
  - MTU Size is the maximum number of bytes that can be transmitted by this router on this link, including the IPX header but not data-link headers.
  - Delay is the period of time (in microseconds) that it takes to transmit one byte of data (excluding protocol headers) to a destination, if the media were free of other traffic. For

a WAN circuit, this value is determined as part of the IW2 protocol. The value used for a LAN circuit is 100 microseconds.

- Throughput is the amount of data, in bits, that can flow through the media and be received at the other side in one second, if there is no other traffic using the interface. For a WAN circuit, this value is determined as part of the IW2 protocol. For a LAN circuit, the value is the bit rate defined for the particular media technology.

  Note: For example, if router A is attached to an X.25 network with an interface speed of 64 Kbps, and router B is attached to the same network with an interface speed of 19.2 Kbps, then both report the adjacency between them as having a throughput of 19.2 Kbps.

- Media Type contains the media code described in Figure 5-2. Additional types will be added in the future. Note that WAN media have codes with the high-order bit set; LAN media have it reset.

Media Code	*Description*
0x0000	Generic LAN for use when no applicable medium is defined
0x8000	Generic WAN for use when no applicable medium is defined
0x0001	LocalTalk*
0x0002	Ethernet II
0x0003	IEEE 802.3 with IEEE 802.2 without SNAP
0x0005	IEEE 802.3 with IPX header immediately following the 802.3 header
0x000A	IEEE 802.3 with IEEE 802.2 and SNAP
0x0004	IEEE 802.5 with IEEE 802.2 without SNAP
0x000B	IEEE 802.5 with IEEE 802.2 and SNAP
0x0006	IEEE 802.4
0x0007	IBM* PC Network II
0x0008	Gateway's G/Net*
0x0009	Proteon's ProNET*
0x000C	Racore's LANPAC
0x800D	ISDN
0x000E	ARCnet*
0x000F	IBM PC Network II with 802.2 without SNAP
0x0010	IBM PC Network II with 802.2 and SNAP
0x0011	Corvus OmniNet at 4 Mbps
0x0012	Harris Adacom
0x0013	IP tunnel
0x8013	IP Relay
0x0014	FDDI with 802.2 without SNAP
0x0015	Commtex IVDLAN
0x0016	Dataco OSI
0x0017	FDDI with 802.2 and SNAP
0x0018	IBM SDLC tunnel
0x0019	PC Office frame
0x001A	Hypercommunications WAIDNET
0x801C	PPP (IETF Point-To-Point Protocol)
0x801D	Proxim RangeLAN
0x801E	X.25
0x801F	Frame Relay
0x0020	Integrated Workstations BUS-NET
0x8021	Novell® SNA Links™

Figure 5-2: Media Codes

*Throughput **		*Default*	
At least	*Strictly less than*	*Cost*	*Typical Media*
0 K	16 K	61	
16 K	32 K	55	
32 K	48 K	55	
48 K	64 K	45	ISDN (U.S.)
64 K	128 K	45	ISDN (European)
128 K	256 K	40	
256 K	512 K	35	
512 K	1 M	30	
1 M	2 M	27	T1 (1.5 M), Corvus Omninet (1 M)
2 M	4 M	26	E1 (2 M), ARCnet (2.5 M)
4 M	8 M	25	Token Ring (4 M), Corvus Omninet (4 M)
8 M	10 M	23	
10 M	16 M	20	Ethernet
16 M	32 M	19	Token Ring (16 M)
32 M	64 M	15	
64 M	128 M	14	FDDI (100 M), CDDI (100 M)
128 M	256 M	9	
256 M	512 M	9	
512 M	1 G	6	
1 G	2 G	6	
2 G	4 G	6	
4 G	8 G	3	
8 G	16 G	3	
16 G	32 G	3	

Figure 5-3: Default Costs

* K = 10^3 bits/second
M = 10^6 bits/second
G = 10^9 bits/second

5.10. Generation of Level 1 Pseudonode LSPs

A router generates a pseudonode LSP on behalf of each circuit for which it is a Designated Router. The LSP specifies the following information in its variable-length fields:

- The Area Addresses option is **not** present. (The set of area addresses for the issuing router is already available from its own non-pseudonode LSP.)
- The Management Information option contains configuration and version information about the circuit:
  - Network Number is the IPX network number for LAN pseudonodes and for numbered WAN pseudonodes. This field is zero for unnumbered WAN pseudonodes.
  - Node Number is the IPX node number (typically, the MAC address) of the Designated Router generating the LSP on the LAN to which the LSP refers. This field is zero for WAN pseudonodes.

    Note: The zero network number is merely a placeholder to hold XRoutes and Services, and is not an actual IPX network number.
  - Router/Server Name gives a textual name for the circuit. It need not be present.
- The External Routes and Services Information parts contain information gleaned from (respectively) RIP and SAP packets received by this Designated Router over the (LAN or WAN) circuit represented by this pseudonode.

  Note: The RIP/SAP packets considered for this purpose are only those from RIP/SAP-only systems. NLSP implementations must be sure not to report here other NLSP routers that are sending RIP/SAP packets for backward compatibility. These "compatibility" RIP/SAP packets are recognizable because they originate from nodes that have NLSP adjacencies with this router. Adjacencies in both the "Up" and "Initializing" states are included. The RIP and SAP packets are intended for consumption by RIP/SAP-only devices sharing the circuit with this router.

  Details of RIP/SAP compatibility are in Section 7.

  If this is an On-Demand Static Routing link, the External Routes and Services Information parts are derived not from periodically received RIP/SAP packets, but from the locally configured RSI database. Details are in Section 3.
- An instance of the Link Information option for each neighboring router, formed from
  a) The Designated Router itself.
  b) If this is a LAN pseudonode, other NLSP routers for which an adjacency with this Designated Router exist in the "Up" state over the circuit represented by the pseudonode.

  Fields of the Link Information option contain values as follows:
  - Zero Cost.
  - The Neighbor ID for the Designated Router itself is its own systemID with appended zero byte (indicating non-pseudonode).
  - The Neighbor ID for each of the other NLSP routers is the other router's systemID with appended zero byte (indicating non-pseudonode).

- The MTU Size and Delay and Throughput fields are zero for pseudonode LSPs.
- Media Type codes are indicated in Figure 5-2.

5.11. Preparing to Initiate Transmission

When generating an LSP as in the preceding sections—periodic or event-driven, pseudonode or non-pseudonode—a router

a) Stores the LSP in its own Link State database, overwriting any previous LSP with the same LSP Number generated by this router

b) Increments the sequence number

c) Calculates the checksum

d) Sets all the SRMflags for that LSP, indicating that it is to be propagated on

- All active LAN circuits (whether or not adjacencies exist for the circuit)
- All WAN circuits having adjacencies in the "Up" state

Circuits running RIP only are excluded, as are On-Demand Static Routing links.

A router must ensure (by reserving resources or otherwise) that it can always store and internalize its own non-pseudonode LSP number zero. If it cannot store and internalize one of its own LSPs, it enters the overload state as described on page 5-24.

Note: It is recommended that a router ensures (by reserving resources or otherwise) that it can always store and internalize all its own LSPs: zero and nonzero, pseudonode and non-pseudonode.

Sometimes an existing LSP is retransmitted, with the same or different sequence number, but with the same information content; that is, the same variable-length part. There is no change of information content because there have been no local topology changes. When this happens, the order of information in the variable-length part must be the same as in the previously transmitted LSP.

Note: This provision allows the receiver to detect that there has been no change of information content by making a byte-wise comparison of the variable-length part. This is an efficient way to check that it is unnecessary to rerun the Decision Process. If a sequence of topology changes results in the local topology returning to some previous state, there is no requirement to preserve the ordering. Preservation is required only if there have been no changes at all.

5.12. Receipt and Propagation of LSPs

A router accepts LSPs as large as actualMaxPacketSize on each circuit. Upon receipt of an LSP on circuit C, a router first performs the following acceptance tests:

a) The general packet acceptance tests described in Section 2 under "General Processing of Incoming IPX Packets."

b) If C is a LAN, and the source data-link address of the LSP does not match the neighborNICAddress of an existing adjacency on C in the "Up" state, the router discards the LSP without generating an event.

Note: A LAN might contain routers from more than one routing area. Checking for existence of an adjacency ensures that only LSPs for the router's own routing area are accepted.

c) If C is a WAN and there is no adjacency on C, the router discards the LSP without generating an event.

d) If the Checksum is zero, set the Remaining Lifetime to zero and proceed.

e) If the Checksum is incorrect (including the case of either or both checksum bytes being zero, and including the zero Remaining Lifetime case described previously), log a badChecksum event and discard the packet.

f) If the Router Type field is not "Level 1" or "Level 1 and Level 2," discard the packet.

g) If the length of the packet, as described in the header, is greater than the buffer in which it was received, log a packetRxSmall event and discard the packet.

h) If the variable part of the LSP is malformed, or contains malformed options, log the malformedOption event and discard the packet.

Note: Do not test the Network Number of the LSP Management Information field against area addresses. Accept any Network Number that arrives in Management Information from an NLSP neighbor.

Depending on two tests of the received LSP, there are four different ways to process the LSP. The tests are as follows:

- Examine the Remaining Lifetime field. If it contains zero, the LSP is *expired*.
- Compare the Source ID field with this router's own systemID. If they match, this is a case of the *same system ID*.

The following four subsections specify the actions to take in the four cases. These subsections use the following shorthand:

- "Send an LSP on circuit C" means
  "Clear SSNflag and set SRMflag for that LSP for C."
- "All circuits" means
  "All WANs having adjacencies in the 'Up' state, and all active LANs regardless of adjacencies."
- "Acknowledge an LSP on circuit C" means
  "Clear the SRMflag for the LSP for C. If C is a WAN, also set the SSNflag for the LSP for C."

When comparing two LSPs for newness, the test on page 5-5 is applied.

5.12.1. Expired LSP with the Same System ID

a) If an LSP with this LSP ID is in the LSP database

  a.1) If the local LSP is also expired

    a.1.1) If the LSPs are equal, acknowledge the LSP on C.

    a.1.2) If the received LSP is older, send the local LSP on C.

    a.1.3) If the received LSP is newer, update the local LSP's sequence number to be one more than the incoming LSP's and send the resulting LSP on all circuits.

  a.2) If the local LSP is not expired

    a.2.1) If the received LSP is older, send the local LSP on C.

a.2.2) If the received LSP is newer or equal, update the sequence number to be one more than the incoming LSP and send the local LSP on all circuits.

b) If an LSP with this LSP ID is **not** in the LSP database, acknowledge the LSP on C without storing it.

5.12.2. Expired LSP with a Different System ID

a) If an LSP with this LSP ID is in the LSP database

a.1) If the LSPs are equal, acknowledge the LSP on C.

a.2) If the received LSP is older, send the local LSP on C.

a.3) If the received LSP is newer, store the new LSP, replacing the local copy, send it on all circuits except C, and acknowledge it on C.

b) If you do not have an LSP with this ID in the LSP database, acknowledge the LSP on C without storing it.

5.12.3. Unexpired LSP with the Same System ID

a) If there is an LSP with this LSP ID in the LSP database

a.1) If the LSPs are equal, acknowledge the LSP on C.

a.2) If the received LSP is newer than the local LSP

a.2.1) Update the local LSP's sequence number to be one more than the incoming LSP and flood the local LSP.

a.2.2) If this is a non-pseudonode LSP and it is LSP number 0, and it contains the Management Information option, compare the server name and the network number fields with those parameters of the local system. If the network numbers do not match, log the duplicateLSPSystemID event. If the network numbers match but the server names do not, log the duplicateInternalNet event.

a.3) If the received LSP is older than the local LSP, send the local LSP on C.

b) If there is no LSP with this ID in the LSP database

b.1) Create an LSP with a sequence number one more than the incoming LSP.

b.2) Follow the procedures for aging an LSP with nonzero remaining lifetime on page 5-7.

5.12.4. Unexpired LSP with a Different System ID

a) If there is an LSP with this ID in the LSP database

a.1) If the LSPs are equal, acknowledge the LSP on C.

a.2) If the received LSP is older, send the local LSP on C.

a.3) If the received LSP is newer

a.3.1) Store the new LSP, replacing the local copy

a.3.2) Send the new LSP on all circuits except C and acknowledge it on C.

b) If you do not have an LSP with this ID in the LSP database

b.1) Store the incoming LSP.

b.2) Send it on all circuits except C and acknowledge it on C.

5.13. Storing a New LSP

When storing a new LSP, the router first ensures that it has enough memory to store the LSP and whatever internal data structures will be required to process the LSP. If the resources are unavailable, the LSP is ignored—it is neither stored nor acknowledged. When an LSP is ignored for this reason, the router enters the *Waiting State* (see page 5-24).

When attempting to store a new version of an existing LSP whose length has not increased, it is recommended that removing the old and storing the new occur as a single atomic action. There is no increase in the resources required. This ensures that such an LSP (which might be carrying the LSP Database Overload indication from an overloaded router) is never ignored for lack of memory resources.

5.14. Receipt of Sequence Number Packets

A router sends Sequence Number Packets (SNPs) to its immediate neighbors and they are not propagated. By sending an SNP, the router communicates the current LSP sequence numbers in its database. A PSNP reports a subset of the LSPs and their sequence numbers. A complete set of CSNPs reports on the total LSP database. A complete set of CSNPs is a set whose Start LSP ID and End LSP ID ranges cover the complete possible range of LSP IDs. That is, every possible LSP ID value appears within the range of one of the CSNPs in the set.

Upon receipt of a PSNP or CSNP on circuit C, a router first performs the following acceptance tests:

a) The general packet acceptance tests described in Section 2 under “General Processing of Incoming IPX Packets.”

b) If C is a LAN, and the source data-link address of the SNP does not match the neighborNICAddress of an existing adjacency on C in the “Up” state, the router discards the SNP without generating an event.

Note: SNPs from a neighbor in the “Initializing” state are ignored. A persistent “Initializing” state might be the result of a neighbor with a faulty network interface—it can send but not receive. The SNPs are ignored to avoid repeated transmissions of LSPs to such a neighbor.

c) If C is a WAN and there is no adjacency on C, the router discards the SNP without generating an event.

d) If a PSNP is received on a LAN for which this system is not the Designated Router, discard the packet.

e) If a CSNP is received on a LAN for which this system is the Designated Router, discard the packet.

f) If the length of the packet, as described in the header, is greater than the buffer in which it was received, log a packetRxSmall event and discard the packet.

g) If the variable part of the SNP is malformed, or contains malformed options, log a malformedOption event and discard the packet.

For each LSP reported in the SNP, use the method on page 5-5 to compare the reported LSP with the corresponding LSP in the local database, and take the following action for C:

Note: No check is made against the receiving router's own systemID. Refer to "Resolving LSP Confusion" later for a rationale.

h) If C is a WAN and the reported LSP is the same as the one stored in the local Link State database, clear the SRMflag.

i) If the reported LSP is older than the local LSP, clear the SSNflag and set the SRMflag.

j) If the reported LSP is newer than the database LSP

j.1) Set the SSNflag.

j.2) If C is a WAN, clear the SRMflag.

k) If no database entry exists for the LSP, and if the reported Remaining Lifetime, Checksum, and Sequence Number fields of the LSP are all nonzero

k.1) Create an LSP entry with sequence number zero and lifetime equal to zeroAgeLifetime.

k.2) Set the SSNflag for that entry for C.

Note: SRMflag must never be set for an LSP having a zero sequence number. Possessing a zero sequence number is equivalent semantically to having no information about that LSP. If such an LSP were propagated by setting the SRMflag, it would result in unnecessary consumption of both bandwidth and memory resources.

l) For an SNP to be valid, the LSP Entries must me in increasing LSP ID order. If a system detects out-of-order entries, it generates a misorderedSNP event. When this happens, which (if any) of the SNP's LSP Entries to process is a local implementation choice.

If the received packet is a CSNP, set the SRMflag for C for LSPs for which

m) The LSP is in this router's database.

n) The LSP has nonzero Sequence Number and nonzero Remaining Lifetime.

o) The LSP ID is within the range specified by the Start LSP ID and End LSP ID fields of the CSNP, inclusive.

p) The LSP is not mentioned in the CSNP.

Note: Because entries in the CSNP are ordered, the LSP screening process can be accomplished in a single pass through the CSNP.

5.15. Transmitting the Packets

In the preceding discussions, packets were composed and generated, flags were set, and the router made preparations to send packets. But they were not actually transmitted. The actual transmission of packets is governed by certain timers. The subsections to follow provide the details.

5.15.1. Expiration of a Complete SNP Interval

For each LAN circuit, a router maintains a local flag: SendCSNPflag.

Note: On WAN links, CSNPs are sent only at initialization.

When a system elects itself Level 1 Designated Router on a circuit, or when a system that is Level 1 Designated Router forms a new adjacency (in the "Up" state) on a LAN, it sets the SendCSNPflag for that circuit.

The router performs the following actions every minimumCompleteSNPInterval; with jitter applied as described in Section 2.

a) for each LAN circuit C for which this system is the Level 1 Designated Router

a.1) If either

- SendCSNPflag is set

Or

- at least maximumCompleteSNPInterval has elapsed since the last CSNP was sent on C

transmit a complete set of Level 1 CSNPs on C. (Ignore the setting of SSNflag on the LSPs.)

a.2) Clear SendCSNPflag for C.

A complete set of CSNPs is a set whose Start LSP ID and End LSP ID ranges cover the complete possible range of LSP IDs. That is, every possible LSP ID value appears within the range of one of the CSNPs in the set. When sending a complete set of CSNPs, observe the following detailed rules.

- The Start LSP ID of the first CSNP of a set is zero. The End LSP ID of the last CSNP of the set is 0xFFFFFFFFFFFFFFFF (all ones).
- The Start LSP ID value of a CSNP is exactly one greater than the End LSP ID value of the preceding CSNP. Do not duplicate the end point LSP IDs. Do not leave gaps.
- If there is a gap in the list of known LSP IDs, there is no relationship between the gap in the LSP ID number space and the boundaries of CSNPs. The Start LSP ID and End LSP ID values might or might not be known LSP IDs.
- As indicated on page 5-18, there can be LSPs in the local database with sequence number zero. These are not included in any CSNP LSP Entries; they are a local bookkeeping device only.
- Include in each CSNP the greatest number of LSP Entries allowed by the circuit's actualMaxPacketSize. An exception is the set's last CSNP, which can have fewer. This rule implies that each CSNP's End LSP ID is strictly greater than its Start LSP ID.
- Do not compute the entire set of LSPs to report at the outset. Rather, each CSNP after the first should reflect those changes in the database that might have occurred since sending the previous CSNP that affects the LSP range not yet covered in this complete CSNP set.

For example, suppose a router holds LSPs with LSP IDs 2, 3, 4, 5, 6, 7, 11, 12, 33, 34, 35, 36, 37, 45, 46, 55. A valid complete set of CSNPs is:

	Start LSP ID	End LSP ID	LSP Entries
CSNP #1	0	5	2, 3, 4, 5
CSNP #2	6	12	6, 7, 11, 12
CSNP #3	13	36	33, 34, 35, 36
CSNP #4	37	0xFFFFFFFFFFFFFFFF	37, 45, 46, 55

Another, equally valid, complete set of CSNPs for the same example is as follows:

	Start LSP ID	End LSP ID	LSP Entries
CSNP #1	0	5	2, 3, 4, 5
CSNP #2	6	32	6, 7, 11, 12
CSNP #3	33	36	33, 34, 35, 36
CSNP #4	37	84	37, 45, 46, 55
CSNP #5	85	0xFFFFFFFFFFFFFFFF	(none)

The empty CSNP #5 could result if, for example, the router held LSP ID 85 at the time CSNP #4 was transmitted, but purged it from the Link State database before CSNP #5 was transmitted.

When multiple CSNPs are transmitted on a circuit, they are separated by an interval of at least 1/18 second. Because of this provision, the interval between transmitting the last CSNP packet of a set and transmitting the first CSNP of a new set might actually be less than minimumCompleteSNPInterval.

> **Note:** Minor variations are sometimes inevitable when adhering to the 1/18 second rule, owing to the timer granularity of the router. Such variations are not material.

5.15.2. Expiration of a Partial SNP Interval

The router performs the following actions every partialSNPInterval for each circuit C, with jitter applied as described in Section 2:

If either

a) C is a WAN

Or

b) C is a LAN and this router is **not** the Designated Router for C

then transmit a PSNP on C containing entries for as many LSPs with SSNflag set as fit in the packet. Then clear the SSNflag for these entries. To avoid starvation, the scan of the LSP database (for those with SSNflag set) starts with the next LSP that was not included in the previous scan. If there are no LSPs with SSNflag set, do not transmit a PSNP.

5.15.3. Expiration of Minimum LSP Transmission Interval

A router takes the following action every minimumLspTransmissionInterval, with jitter applied as in Section 2:

The router scans the LSP database and sends on each circuit C those LSPs having the SRMflag set for C. If C is a LAN, clear the SRMflag.

Where multiple LSPs are transmitted on a circuit, they must be separated by a minimum gap of 1/18 second.

Note: Minor variations are sometimes inevitable when adhering to the minimum gap rule, owing to the timer granularity of the router. Such variations are not material.

Note: In practice, it would be inefficient to scan the entire database this often, particularly when only a few LSPs had SRMflags set. Implementations would typically use additional data structures to optimize this operation.

There is an optional enhancement to the way flooding works. It is an exception to the general rule of waiting for the minimumLspTransmissionInterval to expire. When a system receives a new LSP on a circuit, it floods the LSP immediately (instead of setting SRM flags) on all other circuits. On WAN circuits, it sets the SRMflag as well. (The SRMflag is cleared on a WAN when an SNP acknowledges the LSP.) The objective of this enhancement is faster flooding. Locally generated LSPs are not sent this way. Also, a system that implements faster flooding must still observe circuit pacing (see next subsection). In other situations (for example the receipt of an SNP), or if the circuit is a WAN, the router should follow the normal procedure: set the SRMflag, and actually send the LSP only on expiration of the minimumLspTransmissionInterval. In implementing faster flooding, circuit pacing should be applied, as described in the next subsection; however, do not observe the 1/18 second minumum gap of the previous paragraph.

5.15.4. Circuit Pacing to Avoid Circuit Congestion

Circuit pacing applies to transmission of LSPs and CSNPs on each circuit. This rule is in addition to the minimal 1/18 second gap.

Circuit pacing specifies that no more than paceRate CSNPs are sent per second on the circuit, and no more than paceRate LSPs. The value of paceRate is calculated based on the circuit's Throughput in bits per second (bps), as follows:

paceRate = Throughput / 10,000

with a minimum of 1. For example, at 19,200 bps or less, paceRate is one packet per second. At 19,200 bps and 512-byte packets, no more than 25% of the circuit bandwidth is consumed by LSPs or by CSNPs.

5.15.5. Multiple Circuits Connected to the Same LAN

As an implementation option, a router can take special measures to reduce overhead when there are multiple circuits connected to the same LAN. It is not necessary that distinct routers make the same choice of supporting (or not supporting) these measures.

When a system has two or more separate connections to the same LAN, it can designate one as the *primary circuit* and the others as *secondary circuits*. A router should only do so if the primary circuit is running NLSP and has at least one adjacency in the "Up" state. These circuits must all be configured with the same IPX network number.

A system should treat secondary circuits like all other circuits with the one exception: it need not send LSPs that it is flooding immediately (see page 5-21) on that circuit. Instead of sending the LSP, it takes no action on that circuit. However, it should send LSPs on the secondary circuits in the normal way if SRMflag has been set on that circuit.

Packets received on secondary circuits are processed as usual. For example, SRMflags are cleared on this circuit when receiving an LSP.

5.16. Determining the Latest Information

Several provisions govern manipulation of the values that appear in LSPs. The subsections to follow provide details.

5.16.1. Operation of Sequence Numbers

The Sequence Number is a four-byte unsigned value. When a system initializes, it starts with sequence number one for its own LSPs.

The intent is for the Sequence Number to be incremented monotonically. If a router restarts after being deactivated, it jumps to using sequence numbers beyond the range it used in its previous incarnation. See the next section, "Resolving LSP Confusion," for the details about how this is accomplished.

The sequence numbers a router generates for LSPs of different LSP Number are independent. The algorithm for choosing the numbers is the same, but the numbers advance separately, without being synchronized.

It might happen that a sequence number reaches the maximum value representable in the four-byte sequence-number space. The router owning the LSP, and needing to regenerate it, purges the LSP by sending the LSP with the same sequence number but zero Remaining Lifetime. This causes the LSP to be aged by all other routers. It sets a timer to regenerate the given LSP after maxAge + zeroAgeLifetime. Then the source router starts regenerating the LSP with sequence number 1.

Note: NLSP is less severe than IS-IS in this regard; here, the router need not be disabled. (However, if the router is disabled, a wait of maxAge + zeroAgeLifeTime is necessary before restarting it. In NLSP, recovery from the overflow is automatic. Also, because the router is not shut down, there is only a partial loss of connectivity to the services, external routes, and adjacencies known by the router, rather than a total loss of all connectivity through the router.

5.16.2. Resolving LSP Confusion

This subsection does not add specification logic to the document. It shows how steps in preceding sections fit together to keep LSPs progressing monotonically in an orderly way throughout the routing area despite router restarts.

When a router restarts, losing memory of its previous state, it is possible that an LSP generated by the router in a previous incarnation (that is, by the same system but before the restart) is alive in the routing area. In this case, the collective memory of the other routers advances the sequence numbers of the restarted router's LSPs into a range beyond the numbers used by the router's previous incarnation. Many of the details in processing incoming LSPs and CSNPs are motivated by the need to resolve LSP Confusion.

Suppose router A has restarted and it has immediate neighbors B and C. A sends its first LSP with sequence number one (Section 5.16.1). B still has an old LSP from A with a higher sequence number. So B, considering A to be out of date, helpfully provides A with the "newer" (actually older) LSP (Section 5.12.4, Step a.2). When A receives that LSP from B, it discovers how far its sequence number reached prior to the restart. In then advances its sequence

number for the LSP to make it newer still (Section 5.12.3, Step a.2.1). When the restarted router receives its own LSP from a neighbor and the received LSP has an LSP Number not in use, it purges the LSP from the routing area (Section 5.12.3, Step b.2).

It is even possible for an LSP generated by a system in a previous incarnation to be alive in the routing area and have the same sequence number as the current LSP. To deal with this case, the comparison described on page 5-5 includes the Preference of Checksums rule. With this rule, routers detect that the LSPs are different and all routers make the same decision about which is newer. With this rule, resolution of LSP confusion works the same way for equal and unequal sequence numbers.

Another aspect of LSP confusion is a misconfiguration in which two routers, A and B, are given the same systemID. One of the two, say A, generates an LSP number zero that all routers consider newer than B's LSP number zero. By operation of the protocol, A's LSP eventually reaches B. When this happens, B can readily detect the problem and report A's Network Number and Router/Server Name from the Management Information field of the LSP (Section 5.12.3, Step a.2.2).

References in the preceding paragraphs cited steps in LSP processing. To ensure that LSPs are propagated to the routers that execute these steps, SNP processing must operate in a supportive way. For example, to resolve LSP confusion, a router A requires its neighbors to send A their versions of A's own LSPs. This is the reason that SNP processing does not discriminate in Section 5.14 between this router's and other routers' LSPs.

5.16.3. Synchronizing LSP Expiration

This subsection does not add specification logic to the document. It shows how steps in preceding sections fit together to make old LSPs age out of different routers' Link State databases at (approximately) the same time.

If a router detects that another router has become unreachable, the other router's LSPs are kept anyway. If this were not done, repair of a brief network partition could prompt extensive consumption of network bandwidth and computing resources to reacquire and reabsorb connectivity information that has not changed since before the partition.

Nonetheless, stale information must eventually be removed because the routers in question might not become reachable again. As long as routers in a routing area remain mutually reachable, their LSPs remain in each other's Link State databases. They are refreshed by periodic regeneration (page 5-7). If a router holds an LSP whose remaining lifetime reaches zero, it purges the LSP, as described on page 5-7.

To keep the routers in the routing area consistent with each other, it is not enough that they act independently in purging stale LSPs. Routers' timers are not synchronized with each other. Routing could be disrupted if there was an extended period of mismatched databases. So when the first router ages out an LSP, it conveys this fact to other reachable routers by truncating the LSP and propagating it with zero lifetime (Section 5.6 Step a).

When a router receives such a truncated LSP, it accelerates the aging of the LSP in its own database (Section 5.4 Step b and Section 5.12.4 Step a.3). Once one router has aged an LSP, this provision ensures that all other routers reachable from it purge the LSP as soon as possible.

After purging the LSP, a router keeps it in truncated form in its Link State database for a time (page 5-7). This way, unexpired versions of the same LSP cannot propagate widely. Such

unexpired versions might be in the process of being propagated by a router that is aging it slower than the one instigating the purge. The expired version is considered newer than an unexpired counterpart (Section 5.4 Step b). Keeping the truncated LSP for a while acts as a lock-down on the expiration of the LSP. Including the Remaining Lifetime in Sequence Number Packets reinforces the rapid, firm spread of the expiration news (Section 5.14 Step i).

5.17. Validation of Databases

Corruption of information in router memory is possible from a failure in hardware or software. A router must not continue to operate for an extended period with corrupted routing information. Routers operate in a fail-stop manner. If a failure is detected, the router is disabled until the failure is corrected. Developers can take implementation-specific steps to guarantee this. In the absence of such a guarantee, the router performs the following actions at least every maximumLSPGenerationInterval:

a) To detect corruption of the LSP database while in memory, recheck the checksum of every LSP in the database, except those with zero Remaining Lifetime. If any checksum is incorrect

   i) Log a badIntChecksum event.

   ii) Delete the entire LSP database.

   iii) Reset every WAN adjacency to the "Down" state, as described in Section 4.

   iv) Cause the database to be reacquired.

   Note: One way to accomplish ii, iii, and iv is to disable the router and restart.

b) When the checks are completed, notify the Decision Process, even if all checksums were correct. This causes the Decision Process to be run and the Forwarding database to be recomputed. This protects against corruption of the Forwarding database in memory, which might otherwise persist undetected in memory with a stable topology.

c) Reset the timer for maximumLSPGenerationInterval, with jitter applied as described in Section 2.

5.18. Managing LSP Database Overload

If there are insufficient memory resources to store a received LSP, the receiving router's Link State database might become inconsistent with those of other routers. The router takes the following steps to ensure that other routers do not rely on forwarding paths through the overloaded router:

a) Ignore the LSP that cannot be stored.

b) Enter the Waiting State; start a timer for waitingTime, to limit the duration of the Waiting State.

c) Generate an enteringL1DatabaseOverload event.

d) Generate and flood the router's own LSP having zero LSP Number with the LSP Database Overload bit set.

e) Lower to zero the priority of becoming the Designated Router for every LAN circuit for which the router is not already the Designated Router. (Circuits for which the router is

already the Designated Router are unaffected.) The lowering of priority lasts as long as the overload state, after which the priority reverts to the initial value.

Even if received LSPs can be stored, it is possible that resources become exhausted while running the Decision Process. If this happens, the router enters the Waiting State (as just described) until resources are available and waitingTime has elapsed since the last received LSP was ignored.

While in the Waiting State

a) If an LSP cannot be stored, the router ignores it, restarts the timer for waitingTime, and generates a lnL1DatabaseOverload event.

b) If the Decision Process cannot be run because of insufficient resources, the router restarts the timer for waitingTime and generates a remainingInL1DatabaseOverload event.

c) Propagation of LSPs, running the Decision Process, and forwarding data traffic proceed as normal.

d) When the waitingTime expires, the router

   i) Generates an exitingL1DatabaseOverload event

   ii) Clears the LSP Database Overload bit in its own Level 1 LSP with zero LSP Number and reissues it

   iii) Resumes normal operation

5.19. Link State Database

5.19.1. Configured Values

maxAge
: The time interval after which an LSP is considered to expire. Default is 7500 seconds.

zeroAgeLifetime
: The time interval that a router retains an LSP in its database once the LSP has expired. The expiration can be locally detected, or it can be learned from a neighbor by receiving an expired LSP. Default is 60 seconds.

maximumLSPGenerationInterval
: The maximum interval, in seconds, between the generation of the same LSP. Default is 7200 seconds.

lspBufferSize
: The maximum size of Level 1 LSPs originated by this router. Default is 512 bytes (including the IPX header but not the data-link headers). Every router must be able to send LSPs of at least 512 bytes and accept LSPs of actualMaxPacketSize bytes on a circuit.

  Note: The end user should be sure to configure lspBufferSize to be no larger than the smallest MTU size of any router for any circuit in the area.

minimumLSPTransmissionInterval
: The minimum interval, in seconds, between sending LSPs on a circuit. There can be two separate values: one for LANs and one for WANs. Default is 2 seconds.

minimumLSPGenerationInterval
: The minimal interval, in seconds, between regenerations of the same LSP. Default is 5 seconds.

minimumCompleteSNPInterval
: The smallest interval, in seconds, between generation of Complete Sequence Number Packets by a Designated Router on a LAN. Default is 5 seconds.

maximumCompleteSNPInterval
: The largest interval, in seconds, between generation of Complete Sequence Number Packets by a Designated Router on a LAN. Default is 30 seconds.

partialSNPInterval
: The minimum interval, in seconds, between transmission of Partial Sequence Number Packets. Default is 1 second.

waitingTime
: The number of seconds to remain in the waiting (LSP database overload) state before returning to normal operation. Default is 60 seconds.

5.19.2. Configured Values per Circuit

cost
: The measure attached to a circuit used to make routing decisions. The Decision Process determines the smallest aggregate cost to each reachable NLSP destination. Default values are listed in Figure 5-3.

5.19.3. LSP Database

LSPdatabase
: The set of LSP packets received by this router and still in memory.

SRMflag
: For each LSP and each circuit, a bit indicating that the LSP should be sent on that circuit.

SSNflag
: For each LSP and each circuit, a bit indicating that the LSP should be reported in a PSNP sent on that circuit.

SendCSNPflag
: This flag is set when an event occurs indicating that a complete set of Level 1 CSNPs should be sent on a circuit when the next maximumCompleteSNPInterval elapses.

lastSent
: For each LSP and each circuit, a record of the time at which that LSP was most recently sent on that circuit.

5.19.4. NLSP Events

lanL1DesRouterChange
: The role of Designated Router on a LAN circuit is transferred from one NLSP router to another.

badChecksum
: An LSP was received for which checksum verification failed.

badIntChecksum
: A checksum error was detected in an LSP stored in router memory, indicating a memory corruption error.

misorderedSNP
: An arriving Sequence Number packet contains LSP Entries out of order. The entries must be in increasing LSP ID order.

duplicateLSPSystemID
: An arriving LSP reports a systemID conflicting with this system's systemID, but with a different internal network number.

duplicateInternalNet
: An arriving LSP reports an internalNetworkNumber conflicting with this system's internalNetworkNumber, and with the same systemID and different Router/Server Name.

enteringL1DatabaseOverload
: The router state changes from normal operation to a situation in which the LSP database is overloaded.

remainingInL1DatabaseOverload
: While in an overloaded state, another event occurred to prolong the LSP database overload state.

exitingL1DatabaseOverload
: The router is no longer in the LSP database overload state. Normal operation is resuming.

5.20. Packet Structures

The following packet types are relevant to the link state methods in this section.

- Level 1 LSP (page 5-29)
- Level 1 CSNP (page 5-33)
- Level 1 PSNP (page 5-35)

If this router is capable of receiving packets addressed to the selective link-level multicast address,

- This router's LAN Hello packets are sent with zero as the No Multicast bit.
- The three types of Link State packets are sent initially by multicast when transmitted on a LAN.
- If a LAN Hello packet has been received with the No Multicast bit set, a router must use the broadcast address on that circuit (instead of multicast).

If this router is **not** capable of receiving packets addressed to the selective link-level multicast address,

- This router's LAN Hello packets are sent with the No Multicast bit set to one.
- The three types of Link State packets are sent by broadcast on a LAN.

These packets ride in the data portion of IPX packets. The IPX header fields are encoded as shown in Figure 5-4.

	LSPs and SNPs sent on a WAN	LSPs and SNPs sent on a LAN
Destination Network	zero	zero
Destination Node	0xFFFFFFFFFFFF	0xFFFFFFFFFFFF
Destination Socket	0x9001	0x9001
Source Network	internalNetworkNumber of the router sending the packet	Network number of the LAN on which the packet is being transmitted
Source Node	Node number (0x000000000001) of the router sending the packet on its internalNetworkNumber	IEEE MAC address of the LAN interface through which the packet is being transmitted
Source Socket	0x9001	0x9001
Packet Type	zero	zero

Figure 5-4: IPX Header Fields

These Level 1 LSPs never exceed lspBufferSize bytes in size, including the IPX header but excluding the data-link header.

The maximum size of sequence number packets on any circuit is determined by the media type and the buffer sizes used by neighbors on that circuit. Each Hello packet includes the Local MTU option, containing the sender's localMaxPacketSize for the circuit. This is the way routers communicate the packet size they can receive on that circuit. Every router calculates the minimal value of Local MTU for active adjacencies on each circuit. It then takes the smaller of that value and its own localMaxPacketSize for that circuit. The result is the actualMaxPacketSize value for that circuit. SNPs can be as large as that value, but no larger.

The valid option fields in the variable parts of these packets can appear in any order, except where otherwise noted in the text.

5.20.1. Level 1 LSP

Level 1 LSP

	Number of Bytes
Protocol ID	1
Length Indicator	1
Minor Version	1
Reserved	1
NR | Res | Packet Type	1
Major Version	1
Reserved	2
Packet Length	2
Remaining Lifetime	2
LSP ID	8
Sequence Number	4
Checksum	2
P | Attached Flag | LSP DBOL | Router Type	1
Variable Length Fields	Variable

Protocol ID: 0x83, identifies the NLSP routing layer.

Length Indicator: The number of bytes in the fixed portion of the header (up to and including the Router Type field).

Minor Version: 1, ignored on receipt.

Reserved: 0, ignored on receipt.

Multi-homed Non-routing Server, abbreviated NR in the diagram, (1 bit): When this field contains one, the system has more than one network interface, but does not forward traffic from one network segment to another.

Res (2 bits): 0, ignored on receipt.

Packet Type (5 bits): 18.

Major Version: 1.

Reserved: 0, ignored on receipt.

Packet Length: The entire length of this packet, in bytes, including the fixed portion of the NLSP header.

Remaining Lifetime: The number of seconds before this LSP is considered to expire.

LSP ID: A field composed of three parts:

- Source ID: The system ID of the router that originated the LSP.
- Pseudonode ID: Is zero if this is a non-pseudonode LSP; otherwise, it is a unique (for this Source ID) number designating this pseudonode.
- LSP Number: If a would be LSP is too large to send, the source breaks it into fragments identified by this monotonically increasing number.

LSP ID

	Number of Bytes
Source ID	6
Pseudonode ID	1
LSP Number	1

Sequence Number: The sequence number of the LSP.

Checksum: The checksum of the LSP contents from Source ID (the first part of LSP ID) to the end.

P: (one bit): zero—Indicates that this router does not support partition repair.

Attached Flag (four bits): Indicates whether the router can provide a path to other routing areas.

0 = "No," other routing areas cannot be reached through this router.

1 = "Yes," other routing areas can be reached through this router.

Other values are reserved.

LSPDBOL: (one bit): Set to one when the LSP database is overloaded.

Router Type: (two bits):

1 = "Level 1 Router;" operation is specified in this document.

3 = “Level 1 and Level 2 Router;” accept this value from other routers, for forward compatibility.

Other values are reserved.

Variable Length fields: A series of optional fields, each of which has the following three-part code/length/value Option form:

Option

	Number of Bytes
Code	1
Length	1
Value	Length

Currently defined codes, and the corresponding values, are

- Area Addresses: The set of manualAreaAddresses of the sending router; not present in pseudonode LSPs.

  Code = 0xC0. Length = Total length of the value field, in bytes; either 8, 16, or 24.

  Value = Up to three area addresses. Each area address consists of a four-byte network number, followed by a four-byte address mask. The mask contains from zero to 32 (inclusive) most-significant one bits to indicate which bits of the network number make up the address prefix identifying the routing area. The remaining bits are zero. The bit-wise AND of an Address and Mask pair must be equal to Address; for example, it would be a mistake to send Address = 0x84300000 paired with Mask = 0xFF000000.

Area Addresses

	Number of Bytes
Address	4
Mask	4
... ...	
Address	4
Mask	4

- Management Information: Several fields with information about the router originating the LSP.

  Code = 0xC1. Length = 12 to 60, or more.

  Value = The following five subfields:

  Note: To allow future versions of this protocol to add fields at the end and remain compatible with routers implementing this version, routers this version, routers receiving this option ignore any fields after those listed here.

Management Information

	Number of Bytes
Network Number	4
Node Number	6
IPX Version Number	1
Name Length	1
Router / Server Name	Name Length

  - Network Number: The internal IPX network number of the router generating the LSP. For a LAN pseudonode or a numbered WAN pseudonode, it is the IPX network number of the network segment that the pseudonode represents. For an unnumbered WAN pseudonode, the value is zero.
  - Node Number: For a non-pseudonode LSP, the internal IPX node number (0x000000000001) of the router generating the LSP. For a LAN pseudonode, the node number (typically, the MAC address of the point of attachment) of the Designated Router generating the LSP on the LAN to which the LSP refers. For a WAN pseudonode, the value is zero.
  - IPX Version Number: 1.
  - Name Length: Length of the Router/Server Name field; zero if none is present.

- Router/Server Name: A string of 1 to 47 bytes identifying the router originating the LSP. If this is a pseudonode LSP, this field might not be present; but if it is, it identifies the network segment to which the LSP refers. It is one byte per character, seven-bit ASCII.

  Note: Although the server name is null-terminated in many NetWare protocols, this string is not null-terminated.

- Link Information: Several subfields with information about one adjacency of the source router. This option occurs several times, in general, once for each adjacency.

  Code = 0xC2. Length = 25 or more.

  Value = The following subfields.

  Note: To allow future versions of this protocol to add fields at the end and remain compatible with routers implementing this version, routers receiving this option ignore any fields after those listed here. When sending, use 25 as the Length. When receiving, accept any number 25 or larger.

  - S1 (1 bit): Zero, indicating that the Cost is present.
  - Internal/External, abbreviated I/E in the diagram (1 bit): Zero, indicating that the Cost is an internal metric.

    Note: With hierarchical routing, costs internal to one's own routing area are not usually comparable with those outside it. An internal route is chosen over an external one, if there is a choice. The I/E bit provides information to the Decision Process

    Process that allows the choice to be made.

Link Information — Number of Bytes

S1	I/E	Cost	1
Reserved			3
Neighbor ID			7
MTU Size			4
Delay			4
Throughput			4
Media Type			2

  - Cost (6 bits): The cost of a link to the listed neighbor; an unsigned positive integer.
  - Reserved (three bytes): 0, ignored on receipt.
  - Neighbor ID: For a non-pseudonode neighbor, the neighboring router's system ID plus one byte of zero; for a pseudonode neighbor, the first six bytes are the Designated Router's system ID and the seventh byte is the unique nonzero Pseudonode ID value assigned to this pseudonode by the Designated Router.
  - MTU Size: The maximum number of bytes that can be transmitted on this link by the originating router, including the IPX header but not including the data-link headers. It is zero for a pseudonode LSP.
  - Delay: The time, in microseconds, that it takes to transmit one byte of data (excluding protocol headers) to a destination, if the media is free of other traffic. It is zero for a pseudonode LSP.
  - Throughput: The amount of data, in bits, that can flow through the media and be received at the other side in one second, if there is no other traffic using the interface. It is zero for a pseudonode LSP.
  - Media Type: A code identifying the type of circuit; the most significant bit is one for WAN media, zero for others. See Figure 5-2 for the assigned values.

- Services Information: Describes services that advertise themselves by the SAP protocol.

  Code = 0xC3. Length = 16 to 62.

  Value = The following subfields.

- Hops: The number of hops to reach the service.
- Network Number, Node Number, Socket: The IPX address at which the service is available.
- Type: The type of service offered, see Reference [Nov92] for a partial list of defined types.
- Service Name: The name of the service; its length is determined implicitly by the length of this option—the name is **not** null-terminated.

Services Information

	Number of Bytes
Hops	1
Network Number	4
Node Number	6
Socket	2
Type	2
Service Name	1 to 47

- External Routes: Describes routes obtained by the source router through non-NLSP protocols; for example, RIP (see Reference [Nov92]).

  Code = 0xC4. Length = $7 \times n$; $n \geq 0$.

  Value = The following subfields.

  - Hops: The number of hops reported by the non-NLSP protocol.
  - Network Number: The IPX network number to which this entry refers.
  - Ticks: The RIP Delay (that is, the number of RIP timer ticks) from the source router to network Network Number, as reported by the non-NLSP protocol.

External Routes

	Number of Bytes
Hops	1
Network Number	4
Ticks	2

5.20.2. Level 1 CSNP

CSNP

	Number of Bytes
Protocol ID	1
Length Indicator	1
Minor Version	1
Reserved	1
Reserved / Packet Type	1
Major Version	1
Reserved	2
Packet Length	2
Source ID	7
Start LSP ID	8
End LSP ID	8
Variable Length Fields	Variable

Protocol ID: 0x83, identifies the NLSP routing layer.

Length Indicator: The number of bytes in the fixed portion of the header (up to and including the End LSP ID field).

Minor Version: 1, ignored on receipt.

Reserved: 0, ignored on receipt.

Reserved (3 bits): 0, ignored on receipt.

Packet Type (5 bits): 24.

Major Version: 1.

Reserved: 0, ignored on receipt.

Packet Length: The entire length of this packet, in bytes, including the fixed portion of the NLSP header.

Source ID: The six-byte system ID of the sending router, followed by one byte of zero.

Start LSP ID, End LSP ID: The first and last LSP in the range covered by this CSNP. Each is a field composed of three parts:

- Source ID: The systemID of the router that originated the LSP being reported.
- Pseudonode ID: Zero if this is a non-pseudonode LSP; otherwise, it is a unique (for this Source ID) number designating this pseudonode.
- LSP Number: If a would be LSP is too large to send, the source breaks it into fragments identified by this monotonically increasing number.

LSP ID

	Number of Bytes
Source ID	6
Pseudonode ID	1
LSP Number	1

Variable Length fields: A series of optional fields, each of which has the following three-part code/length/value Option form:

Option

	Number of Bytes
Code	1
Length	1
Value	Length

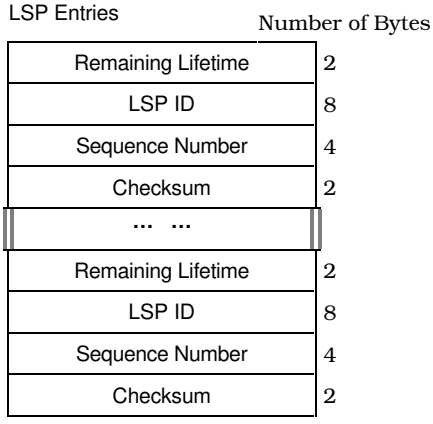
Currently defined codes, and the corresponding values, are

- LSP Entries: This option can appear more than once; if so, instances must be sorted in ascending LSP ID order.

  Code = 9. Length = $16 \times n$; $n > 0$.

  Value = A list of LSP entries, sorted in ascending LSP ID order (the LSP Number byte of the LSP ID is the least significant byte). Each entry has four subfields:

  - Remaining Lifetime: Seconds remaining until the indicated LSP expires.
  - LSP ID: Identifies the LSP referred to by this LSP Entry. It has the same three-part Source ID/Pseudonode ID/LSP Number form described on page 5-33.
  - Sequence Number: The sequence number of the indicated LSP.
  - Checksum: The checksum reported in the indicated LSP.


LSP Entries
Number of Bytes
Remaining Lifetime
2
LSP ID
8
Sequence Number
4
Checksum
2
... ...
Remaining Lifetime
2
LSP ID
8
Sequence Number
4
Checksum
2

5.20.3. Level 1 PSNP

PSNP

	Number of Bytes
Protocol ID	1
Length Indicator	1
Minor Version	1
Reserved	1
Reserved / Packet Type	1
Major Version	1
Reserved	2
Packet Length	2
Source ID	7
Variable Length Fields	Variable

Protocol ID: 0x83, identifies the NLSP routing layer.

Length Indicator: The number of bytes in the fixed portion of the header (up to and including the Source ID field).

Minor Version: 1, ignored on receipt.

Reserved: 0, ignored on receipt.

Reserved (3 bits): 0, ignored on receipt.

Packet Type (5 bits): 26.

Major Version: 1.

Reserved: 0, ignored on receipt.

Packet Length: The entire length of this packet, in bytes, including the fixed portion of the NLSP header.

Source ID: The six-byte system ID of the sending router, followed by one byte of zero.

Option

	Number of Bytes
Code	1
Length	1
Value	Length

Variable Length: A series of optional fields, each of which has the following three-part code/length/value Option form:

Currently defined codes, and the corresponding values, are

- LSP Entries: This option can appear more than once; if so, instances must be sorted in ascending LSP ID order.

  Code = 9. Length = $16 \times n$; $n > 0$.

  Value = A list of LSP entries, sorted in ascending LSP ID order (the LSP Number byte of the LSP ID is the least significant byte). Each entry has four subfields:

  - Remaining Lifetime: Seconds remaining until the indicated LSP expires.
  - LSP ID: Identifies the LSP referred to by this LSP Entry. It has the same three-part Source ID/Pseudonode ID/LSP Number form described on page 5-33.
  - Sequence Number: The sequence number of the indicated LSP.
  - Checksum: The checksum reported in the indicated LSP.

LSP Entries

	Number of Bytes
Remaining Lifetime	2
LSP ID	8
Sequence Number	4
Checksum	2
... ...	
Remaining Lifetime	2
LSP ID	8
Sequence Number	4
Checksum	2