Network Working Group
Request for Comments: 1142

David Oran, Editor
Digital Equipment Corp.
February 1990

OSI IS-IS Intra-domain Routing Protocol

Status of this Memo

This RFC is a republication of ISO DP 10589 as a service to the Internet Community. This is not an Internet standard. Distribution of this memo is unlimited.

Contents

Figures

Tables

Introduction

This Protocol is one of a set of International Standards produced to facilitate the interconnection of open systems. The set of standards covers the services and protocols required to achieve such interconnection.

This Protocol is positioned with respect to other related standards by the layers defined in the ISO 7498 and by the structure defined in the ISO 8648. In particular, it is a protocol of the Network Layer. This protocol permits Intermediate Systems within a routeing Domain (as defined in ISO/TR 9575) to exchange configuration and routeing information to facilitate the operation of the routeing and relaying functions of the Network Layer.

The protocol is designed to operate in close conjunction with ISO 9542 and ISO 8473. ISO 9542 is used to establish connectivity and reachability between End Systems and Intermediate Systems on individual Subnetworks. Data is carried by ISO 8473. The related algorithms for route calculation and maintenance are also described.

The intra-domain IS–IS routeing protocol is intended to support large routeing domains consisting of combinations of many types of subnetworks. This includes point-to-point links, multipoint links, X.25 subnetworks, and broadcast subnetworks such as ISO 8802 LANs.

In order to support large routeing domains, provision is made for Intra-domain routeing to be organised hierarchically. A large domain may be administratively divided into *areas*. Each system resides in exactly one area. Routeing within an area is referred to as *Level 1 routeing*. Routeing between areas is referred to as *Level 2 routeing*. Level 2 Intermediate systems keep track of the paths to destination areas. Level 1 Intermediate systems keep track of the routeing within their own area. For an NPDU destined to another area, a Level 1 Intermediate system sends the NPDU to the nearest level 2 IS in its own area, regardless of what the destination area is. Then the NPDU travels via level 2 routeing to the destination area, where it again travels via level 1 routeing to the destination End System.

Information technology — Telecommunications and information exchange between systems — Intermediate system to Intermediate system Intra-Domain routeing exchange protocol for use in Conjunction with the Protocol for providing the Connectionless-mode Network Service (ISO 8473)

1 Scope and Field of Application

This International Standard specifies a protocol which is used by Network Layer entities operating ISO 8473 in Intermediate Systems to maintain routeing information for the purpose of routeing within a single routeing domain. The protocol herein described relies upon the provision of a connectionless-mode underlying service.[1]

This Standard specifies:

a) procedures for the transmission of configuration and routeing information between network entities residing in Intermediate Systems within a single routeing domain;

b) the encoding of the protocol data units used for the transmission of the configuration and routeing information;

c) procedures for the correct interpretation of protocol control information; and

d) the functional requirements for implementations claiming conformance to this Standard.

The procedures are defined in terms of:

a) the interactions between Intermediate system Network entities through the exchange of protocol data units; and

b) the interactions between a Network entity and an underlying service provider through the exchange of subnetwork service primitives.

c) the constraints on route determination which must be observed by each Intermediate system when each has a routeing information base which is consistent with the others.

2 References

2.1 Normative References

The following standards contain provisions which, through reference in this text, constitute provisions of this International Standard. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this International Standard are encouraged to investigate the possibility of apply-

[1] See ISO 8473 and its Addenda 1 and 3 for the mechanisms necessary to realise this service on subnetworks based on ISO 8208, ISO 8802, and the OSI Data Link Service.

ing the most recent editions of the standards listed below. Members of IEC and ISO maintain registers of currently valid International Standards.

ISO 7498:1984, *Information processing systems — Open Systems Interconnection — Basic Reference Model.*

ISO 7498/Add.1:1984, *Information processing systems — Open Systems Interconnection — Basic Reference Model — Addendum 1: Connectionless-mode Transmission.*

ISO 7498/Add.3:1989, *Information processing systems — Open Systems Interconnection — Basic Reference Model — Addendum 3: Naming, including Addressing.*

ISO 7498/Add.4:[2], *Information processing systems — Open Systems Interconnection — Basic Reference Model — Addendum 4: Management Framework.*

ISO 8348:1987, *Information processing systems — Data communications — Network Service Definition.*

ISO 8348/Add.1:1987, *Information processing systems — Data communications — Network Service Definition — Addendum 1: Connectionless-mode transmission.*

ISO 8348/Add.2:1988, *Information processing systems — Data communications — Network Service Definition — Addendum 2: Network layer addressing.*

ISO 8473:1988, *Information processing systems — Data communications — Protocol for providing the connectionless-mode network service.*

ISO 8473/Add.3:1989, *Information processing systems — Telecommunications and information exchange between systems — Protocol for providing the connectionless-mode network service — Addendum 3: Provision of the underlying service assumed by ISO 8473 over subnetworks which provide the OSI data link service.*

ISO 8648:1988, *Information processing systems — Open Systems Interconnection — Internal organisation of the Network Layer.*

ISO 9542:1988, *Information processing systems – Telecommunications and information exchange between systems — End system to Intermediate system Routeing exchange protocol for use in conjunction with the protocol for providing the connectionless -mode network service (ISO 8473).*

ISO 8208:1984, *Information processing systems — Data communications — X.25 packet level protocol for Data terminal equipment*

ISO 8802:1988, *Information processing systems — Telecommunications and information exchange between systems — Local area networks.*

ISO/TR 9575:1989, *Information technology — Telecommunications and information exchange between systems — OSI Routeing Framework.*

ISO/TR 9577:[2], *Information technology — Telecommunications and information exchange between systems — Protocol Identification in the Network Layer.*

2.2 Other References

The following references are helpful in describing some of the routeing algorithms:

McQuillan, J. et. al., *The New Routeing Algorithm for the ARPANET*, IEEE Transactions on Communications, May 1980.

Perlman, Radia, *Fault-Tolerant Broadcast of Routeing Information*, Computer Networks, Dec. 1983. Also in IEEE INFOCOM 83, April 1983.

[2] to be published

Aho, Hopcroft, and Ullman, *Data Structures and Algorithms*, P204–208 — Dijkstra algorithm.

3 Definitions

3.1 Reference Model definitions

This International Standard makes use of the following terms defined in ISO 7498:

a) Network Layer
b) Network Service access point
c) Network Service access point address
d) Network entity
e) Routeing
f) Network protocol
g) Network relay
h) Network protocol data unit

3.2 Network Layer architecture definitions

This International Standard makes use of the following terms defined in ISO 8648:

a) Subnetwork
b) End system
c) Intermediate system
d) Subnetwork service
e) Subnetwork Access Protocol
f) Subnetwork Dependent Convergence Protocol
g) Subnetwork Independent Convergence Protocol

3.3 Network Layer addressing definitions

This International Standard makes use of the following terms defined in ISO 8348/Add.2:

a) Subnetwork address
b) Subnetwork point of attachment
c) Network Entity Title

3.4 Local Area Network Definitions

This International Standard makes use of the following terms defined in ISO 8802:

a) Multi-destination address
b) Broadcast medium

3.5 Routeing Framework Definitions

This document makes use of the following terms defined in ISO/TR 9575:

a) Administrative Domain
b) Routeing Domain
c) Hop
d) Black hole

3.6 Additional Definitions

For the purposes of this International Standard, the following definitions apply:

3.6.1 **Area:** A routeing subdomain which maintains detailed routeing information about its own internal composition, and also maintains routeing information which allows it to reach other routeing subdomains. It corresponds to the Level 1 subdomain.

3.6.2 **Neighbour:** An adjacent system reachable by traversal of a single subnetwork by a PDU.

3.6.3 **Adjacency:** The subset of the local routeing information base pertinent to a single neighbour.

3.6.4 **Circuit:**The subset of the local routeing information base pertinent to a single local SNPA.

3.6.5 **Link:** The communication path between two neighbours.

A Link is “up” when communication is possible between the two SNPAs.

3.6.6 **Designated IS:** The Intermediate system on a LAN which is designated to perform additional duties. In particular it generates Link State PDUs on behalf of the LAN, treating the LAN as a pseudonode.

3.6.7 **Pseudonode:** Where a broadcast subnetwork has *n* connected Intermediate systems, the broadcast subnetwork itself is considered to be a *pseudonode*.

The pseudonode has links to each of the *n* Intermediate systems and each of the ISs has a single link to the pseudonode (rather than *n-1* links to each of the other Intermediate systems). Link State PDUs are generated on behalf of the pseudonode by the Designated IS. This is depicted below in figure 1.

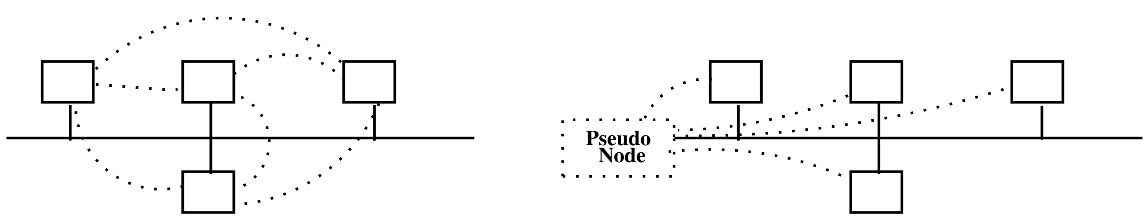


Figure 1 - Use of a Pseudonode to collapse a LAN Topology

3.6.8 **Broadcast subnetwork**: A subnetwork which supports an arbitrary number of End systems and Intermediate systems and additionally is capable of transmitting a single SNPDU to a subset of these systems in response to a single SN_UNITDATA request.

3.6.9 **General topology subnetwork:** A subnetwork which supports an arbitrary number of End systems and Intermediate systems, but does not support a convenient multi-destination connectionless transmission facility, as does a broadcast subnetwork.

3.6.10 **Routeing Subdomain:** a set of Intermediate systems and End systems located within the same Routeing domain.

3.6.11 **Level 2 Subdomain:** the set of all Level 2 Intermediate systems in a Routeing domain.

3.6.12 **Hippity Cost:** the cost, in terms of a routeing metric value, of traversing an Intermediate System. An Intermediate System having an **infinite hippity cost** indicates that no routes are to be calculated passing through that Intermediate System.

4 Symbols and Abbreviations

4.1 Data Units

PDU	Protocol Data Unit
SNSDU	Subnetwork Service Data Unit
NSDU	Network Service Data Unit
NPDU	Network Protocol Data Unit
SNPDU	Subnetwork Protocol Data Unit

4.2 Protocol Data Units

ESH PDU	ISO 9542 End System Hello Protocol Data Unit
ISH PDU	ISO 9542 Intermediate System Hello Protocol Data Unit
RD PDU	ISO 9542 Redirect Protocol Data Unit
IIH	Intermediate system to Intermediate system Hello Protocol Data Unit
LSP	Link State Protocol Data Unit
SNP	Sequence Numbers Protocol Data Unit
CSNP	Complete Sequence Numbers Protocol Data Unit
PSNP	Partial Sequence Numbers Protocol Data Unit

4.3 Addresses

AFI	Authority and Format Indicator
DSP	Domain Specific Part
IDI	Initial Domain Identifier
IDP	Initial Domain Part
NET	Network Entity Title
NSAP	Network Service Access Point
SNPA	Subnetwork Point of Attachment

4.4 Miscellaneous

DA	Dynamically Assigned
DCM	Dynamic Connection Management
DED	Dynamically Established Data link
DTE	Data Terminal Equipment
ES	End System
IS	Intermediate System
L1	Level 1
L2	Level 2
LAN	Local Area Network
MAC	Media Access Control
QoS	Quality of Service
SN	Subnetwork

SNAcP	Subnetwork Access Protocol
SNDCP	Subnetwork Dependent Convergence Protocol
SNICP	Subnetwork Independent Convergence Protocol
SRM	Send Routeing Message
SSN	Send Sequence Numbers Message
SVC	Switched Virtual Circuit

5 Typographical Conventions

This International Standard makes use of the following typographical conventions:

a) Important terms and concepts appear in *italic* type when introduced for the first time;

b) Protocol constants and management parameters appear in sansSerif type with multiple words run together. The first word is lower case, with the first character of subsequent words capitalised;

c) Values of constants, parameters, and protocol fields appear enclosed in “double quotes”.

6 Overview of the Protocol

6.1 System Types

There are the following types of system:

End Systems: These systems deliver NPDUs to other systems and receive NPDUs from other systems, but do not relay NPDUs. This International Standard does not specify any additional End system functions beyond those supplied by ISO 8473 and ISO 9542.

Level 1 Intermediate Systems: These systems deliver and receive NPDUs from other systems, and relay NPDUs from other source systems to other destination systems. They route directly to systems within their own area, and route towards a level 2 Intermediate system when the destination system is in a different area.

Level 2 Intermediate Systems: These systems act as Level 1 Intermediate systems in addition to acting as a system in the subdomain consisting of level 2 ISs. Systems in the level 2 subdomain route towards a destination area, or another routeing domain.

6.2 Subnetwork Types

There are two generic types of subnetworks supported.

a) *broadcast subnetworks*: These are multi-access subnetworks that support the capability of addressing a group of attached systems with a single NPDU, for instance ISO 8802.3 LANs.

b) *general topology subnetworks*: These are modelled as a set of point-to-point links each of which connects exactly two systems.

There are several generic types of general topology subnetworks:

1) *multipoint links*: These are links between more than two systems, where one system is a primary (or master) system, and the remaining systems are secondary (or slave) systems. The primary is capable of direct

communication with any of the secondaries, but the secondaries cannot communicate directly among themselves.

2) *permanent point-to-point links*: These are links that stay connected at all times (unless broken, or turned off by system management), for instance leased lines or private links.

3) *dynamically established data links (DEDs):* these are links over connection oriented facilities, for instance X.25, X.21, ISDN, or PSTN networks.

Dynamically established data links can be used in one of three ways:

(i) *static point-to-point (Static):* The call is established upon system management action and cleared only on system management action (or failure).

(ii) *dynamic connection management (DCM):* The call is established upon receipt of traffic, and brought down on timer expiration when idle.

(iii) *dynamically assigned (DA):* Uses dynamic connection management, but in addition, the address to which the call is to be established is determined dynamically.

All subnetwork types are treated by the Subnetwork Independent functions as though they were connectionless subnetworks, using the Subnetwork Dependent Convergence functions of ISO 8473 where necessary to provide a connectionless subnetwork service. The Subnetwork Dependent functions do, however, operate differently on connectionless and connection-oriented subnetworks.

6.3 Topologies

A single organisation may wish to divide its *Administrative Domain* into a number of separate *Routeing Domains.* This has certain advantages, as described in ISO/TR 9575. Furthermore, it is desirable for an intra-domain routeing protocol to aid in the operation of an inter-domain routeing protocol, where such a protocol exists for interconnecting multiple administrative domains.

In order to facilitate the construction of such multi-domain topologies, provision is made for the entering of *static* inter-domain routeing information. This information is provided by a set of *Reachable Address Prefixes* entered by System Management at the ISs which have links which cross routeing domain boundaries. The prefix indicates that any NSAPs whose NSAP address matches the prefix may be reachable via the SNPA with which the prefix is associated. Where the subnetwork to which this SNPA is connected is a general topology subnetwork supporting dynamically established data links, the prefix also has associated with it the required subnetwork addressing information, or an indication that it may be derived from the destination NSAP address (for example, an X.121 DTE address may sometimes be obtained from the IDI of the NSAP address).

The Address Prefixes are handled by the level 2 routeing algorithm in the same way as information about a level 1 area within the domain. NPDUs with a destination address matching any of the prefixes present on any Level 2 Intermediate System within the domain can therefore be relayed (using level 2 routeing) by that IS and delivered out of the domain. (It is assumed that the routeing functions of the other domain will then be able to deliver the NPDU to its destination.)

6.4 Addresses

Within a routeing domain that conforms to this standard, the Network entity titles of Intermediate systems shall be structured as described in 7.1.1.

All systems shall be able to generate and forward data PDUs containing NSAP addresses in any of the formats specified by ISO 8348/Add.2. However, NSAP addresses of End systems should be structured as described in 7.1.1 in order to take full advantage of IS–IS routeing. Within such a domain it is still possible for some End Systems to have addresses assigned which do not conform to 7.1.1, provided they meet the more general requirements of ISO 8348/Add.2, but they may require additional configuration and be subject to inferior routeing performance.

6.5 Functional Organisation

The intra-domain IS–IS routeing functions are divided into two groups

- *Subnetwork Independent Functions*
- *Subnetwork Dependent Functions*

6.5.1 Subnetwork Independent Functions

The Subnetwork Independent Functions supply full-duplex NPDU transmission between any pair of neighbour systems. They are independent of the specific subnetwork or data link service operating below them, except for recognising two generic types of subnetworks:

- **General Topology Subnetworks,** which include HDLC point-to-point, HDLC multipoint, and dynamically established data links (such as X.25, X.21, and PSTN links), and
- **Broadcast Subnetworks**, which include ISO 8802 LANs.

The following Subnetwork Independent Functions are identified

- **Routeing.** The routeing function determines NPDU paths. A path is the sequence of connected systems and links between a source ES and a destination ES.

  The combined knowledge of all the Network Layer entities of all the Intermediate systems within a routeing domain is used to ascertain the existence of a path, and route the NPDU to its destination. The routeing component at an Intermediate system has the following specific functions:

  - It extracts and interprets the routeing PCI in an NPDU.
  - It performs NPDU forwarding based on the destination address.
  - It manages the characteristics of the path. If a system or link fails on a path, it finds an alternate route.
  - It interfaces with the subnetwork dependent functions to receive reports concerning an SNPA which has become unavailable, a system that has failed, or the subsequent recovery of an SNPA or system.
  - It informs the ISO 8473 error reporting function when the forwarding function cannot relay an NPDU, for instance when the destination is unreachable or when the NPDU would have needed to be segmented and the NPDU requested “no segmentation”.
- **Congestion control**. Congestion control manages the resources used at each Intermediate system.

6.5.2 Subnetwork Dependent Functions

The subnetwork dependent functions mask the characteristics of the subnetwork or data link service from the subnetwork independent functions. These include:

- Operation of the Intermediate system functions of ISO 9542 on the particular subnetwork, in order to
  - Determine neighbour Network entity title(s) and SNPA address(es)
  - Determine the SNPA address(s) of operational Intermediate systems
- Operation of the requisite Subnetwork Dependent Convergence Function as defined in ISO 8473 and its Addenda 1 and 3, in order to perform
  - Data link initialisation
  - Hop by hop fragmentation over subnetworks with small maximum SNSDU sizes

- Call establishment and clearing on dynamically established data links

6.6 Design Goals

This International Standard supports the following design requirements. The correspondence with the goals for OSI routeing stated in ISO/TR 9575 are noted.

- **Network Layer Protocol Compatibility.** It is compatible with ISO 8473 and ISO 9542. *(See clause 7.5 of ISO/TR 9575),*
- **Simple End systems:** It requires no changes to end systems, nor any functions beyond those supplied by ISO 8473 and ISO 9542. *(See clause 7.2.1 of ISO/TR 9575),*
- **Multiple Organisations:** It allows for multiple routeing and administrative domains through the provision of static routeing information at domain boundaries. *(See clause 7.3 of ISO/TR 9575),*
- **Deliverability** It accepts and delivers NPDUs addressed to reachable destinations and rejects NPDUs addressed to destinations known to be unreachable.
- **Adaptability.** It adapts to topological changes within the routeing domain, but not to traffic changes, except potentially as indicated by local queue lengths. It splits traffic load on multiple equivalent paths. *(See clause 7.7 of ISO/TR 9575),*
- **Promptness.** The period of adaptation to topological changes in the domain is a reasonable function of the domain diameter (that is, the maximum logical distance between End Systems within the domain) and Data link speeds. *(See clause 7.4 of ISO/TR 9575),*
- **Efficiency.** It is both processing and memory efficient. It does not create excessive routeing traffic overhead. *(See clause 7.4 of ISO/TR 9575),*
- **Robustness.** It recovers from transient errors such as lost or temporarily incorrect routeing PDUs. It tolerates imprecise parameter settings. *(See clause 7.7 of ISO/TR 9575),*
- **Stability.** It stabilises in finite time to “good routes”, provided no continuous topological changes or continuous data base corruptions occur.
- **System Management control.** System Management can control many routeing functions via parameter changes, and inspect parameters, counters, and routes. It will not, however, depend on system management action for correct behaviour.
- **Simplicity.** It is sufficiently simple to permit performance tuning and failure isolation.
- **Maintainability.** It provides mechanisms to detect, isolate, and repair most common errors that may affect the routeing computation and data bases. *(See clause 7.8 of ISO/TR 9575),*
- **Heterogeneity.** It operates over a mixture of network and system types, communication technologies, and topologies. It is capable of running over a wide variety of subnetworks, including, but not limited to: ISO 8802 LANs, ISO 8208 and X.25 subnetworks, PSTN networks, and the OSI Data Link Service. *(See clause 7.1 of ISO/TR 9575),*
- **Extensibility.** It accommodates increased routeing functions, leaving earlier functions as a subset.
- **Evolution.** It allows orderly transition from algorithm to algorithm without shutting down an entire domain.
- **Deadlock Prevention.** The congestion control component prevents buffer deadlock.

- **Very Large Domains.** With hierarchical routeing, and a very large address space, domains of essentially unlimited size can be supported. *(See clause 7.2 of ISO/TR 9575),*

- **Area Partition Repair.** It permits the utilisation of level 2 paths to repair areas which become partitioned due to failing level 1 links or ISs. *(See clause 7.7 of ISO/TR 9575),*

- **Determinism.** Routes are a function only of the physical topology, and not of history. In other words, the same topology will always converge to the same set of routes.

- **Protection from Mis-delivery.** The probability of mis-delivering a NPDU, i.e. delivering it to a Transport entity in the wrong End System, is extremely low.

- **Availability.** For domain topologies with *cut set* greater than one, no single point of failure will partition the domain. *(See clause 7.7 of ISO/TR 9575),*

- **Service Classes.** The service classes of *transit delay, expense*[3], and *residual error probability* of ISO 8473 are supported through the optional inclusion of multiple routeing metrics.

6.6.1 Non-Goals

The following are not within the design scope of the intra-domain IS–IS routeing protocol described in this International Standard:

- **Traffic adaptation.** It does not automatically modify routes based on global traffic load.

- **Source-destination routeing.** It does not determine routes by source as well as destination.

- **Guaranteed delivery.** It does not guarantee delivery of all offered NPDUs.

- **Level 2 Subdomain Partition Repair.** It will not utilise Level 1 paths to repair a level 2 subdomain partition. For correct operation, a connected level 2 subdomain is required.

  EDITOR'S NOTE: The above statement about "correct operation" is not strictly true, and leaves the impression that the protocol will break if the L2 subdomain becomes partitioned. This is not true — the two pieces each work correctly in isolation. The editor suggests that the second sentence above be changed to read: "For full logical connectivity to be available, a connected...".

- **Equal treatment for all ES Implementations.** The End system poll function defined in 8.4.5 presumes that End systems have implemented the Suggested ES Configuration Timer option of ISO 9542. An End system which does not implement this option may experience a temporary loss of connectivity following certain types of topology changes on its local subnetwork.

6.7 Environmental Requirements

For correct operation of the protocol, certain guarantees are required from the local environment and the Data Link Layer.

The required local environment guarantees are:

a) Resource allocation such that the certain minimum resource guarantees can be met, including

   1) memory (for code, data, and buffers)
   2) processing;

   See 12.2.5 for specific performance levels required for conformance

[3] "Expense" is referred to as "cost" in ISO 8473. The latter term is not used here because of possible confusion with the more general usage of the term to indicate path cost according to any routeing metric.

b) A quota of buffers sufficient to perform routeing functions;

c) Access to a timer or notification of specific timer expiration; and

d) A very low probability of corrupting data.

The required subnetwork guarantees for point-to-point links are:

a) Provision that both source and destination systems complete start-up before PDU exchange can occur;

b) Detection of remote start-up;

c) Provision that no old PDUs be received after start-up is complete;

d) The following events are “low probability”, which means that they occur sufficiently rarely so as not to impact performance, on the order of once per thousand NPDUs

   1) Loss, duplication or corruption of NPDUs;

   2) Routeing PDU non-sequentiality; and

e) Reporting of failures and degraded subnetwork conditions.

The required subnetwork guarantees for broadcast links are:

a) Multicast capability, i.e., the ability to address a subset of all connected systems with a single PDU;

b) The following events are “low probability”, which means that they occur sufficiently rarely so as not to impact performance, on the order of once per thousand PDUs

   1) Routeing PDU non-sequentiality,

   2) Routeing PDU loss due to detected corruption; and

   3) Receiver overrun;

c) The following events are “very low probability”, which means performance will be impacted unless they are extremely rare, on the order of less than one event per four years

   1) Delivery of NPDUs with undetected data corruption; and

   2) Non-transitive connectivity, i.e. where system *A* can receive transmissions from systems *B* and *C*, but system *B* cannot receive transmissions from system *C*.

The following services are assumed to be not available from broadcast links:

a) Reporting of failures and degraded subnetwork conditions that result in NPDU loss, for instance receiver failure. The routeing functions are designed to account for these failures.

6.8 Functional Organisation of Subnetwork Independent Components

The Subnetwork Independent Functions are broken down into more specific functional components. These are described briefly in this sub-clause and in detail in clause 7. This International Standard uses a functional decomposition adapted from the model of routeing presented in clause 5.1 of ISO/TR 9575. The decomposition is not identical to that in ISO/TR 9575, since that model is more general and not specifically oriented toward a detailed description of intra-domain routeing functions such as supplied by this protocol.

The functional decomposition is shown below in figure 2.

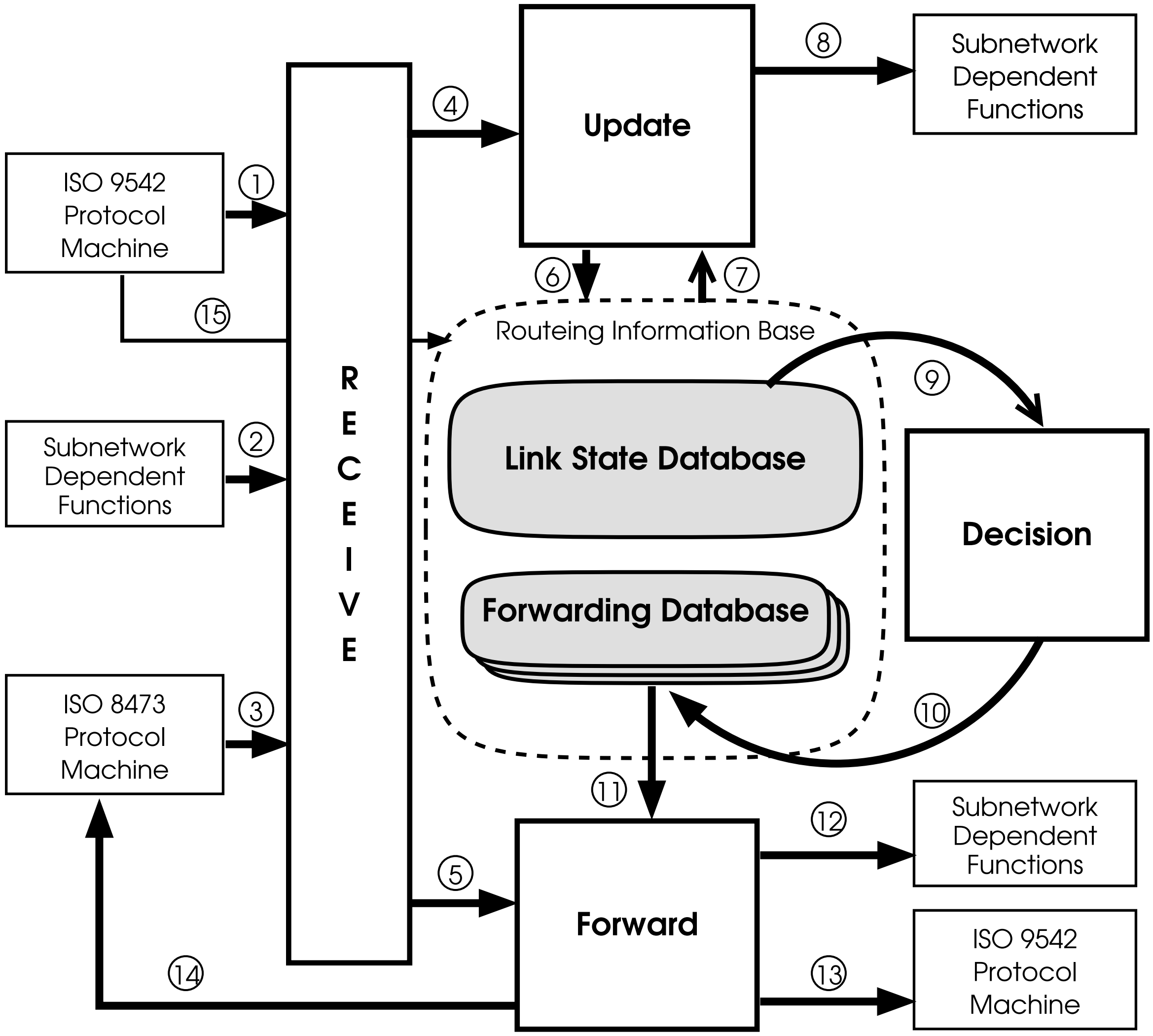


Figure 2 - Decomposition of Subnetwork Independent Functions

6.8.1 Routeing

The routeing processes are:

- *Decision Process*
- *Update Process*

  NOTE — this comprises both the *Information Collection* and *Information Distribution* components identified in ISO/TR 9575.
- *Forwarding Process*
- *Receive Process*

6.8.1.1 Decision Process

This process calculates routes to each destination in the domain. It is executed separately for level 1 and level 2 routeing, and separately within each level for each of the routeing metrics supported by the Intermediate system. It uses the *Link State Database*, which consists of information from the latest Link State PDUs from every other Intermediate system in the area, to compute shortest paths from this IS to all other systems in the area – ⑨ in figure 2. The Link State Data Base is maintained by the Update Process.

Execution of the Decision Process results in the determination of [circuit, neighbour] pairs (known as *adjacencies*), which are stored in the appropriate Forwarding Information base – ⑩ – and used by the Forwarding process as paths along which to forward NPDUs.

Several of the parameters in the routeing data base that the Decision Process uses are set by system management or are determined by the implementation. These include:

- maximum number of Intermediate and End systems within the IS's area;
- maximum number of Intermediate and End system neighbours of the IS, etc.,

so that databases can be sized appropriately. Also parameters such as

- routeing metrics for each circuit; and
- timers

can be adjusted for enhanced performance. The complete list of System Management set-able parameters is listed in clause 11.

6.8.1.2 Update Process

This process constructs, receives and propagates Link State PDUs. Each Link State PDU contains information about the identity and routeing metric values of the adjacencies of the IS that originated the Link State PDU.

The Update Process receives Link State and Sequence Numbers PDUs from the Receive Process — ④ in figure 2. It places new routeing information in the routeing information base — ⑥ and propagates routeing information to other Intermediate systems – ⑦ and ⑧ .

General characteristics of the Update Process are:

- Link State PDUs are generated as a result of topological changes, and also periodically. They may also be generated indirectly as a result of System Management actions (such as changing one of the routeing metrics for a circuit).

- Level 1 Link State PDUs are propagated to all Intermediate systems within an area, but are not propagated out of an area.

- Level 2 Link State PDUs are propagated to all Level 2 Intermediate systems in the domain.

- Link State PDUs are not propagated outside of a domain.

- The update process, through a set of System Management parameters, enforces an upper bound on the amount of routeing traffic overhead it generates.

6.8.1.3 Forwarding Process

This process supplies and manages the buffers necessary to support NPDU relaying to all destinations.

It receives, via the Receive Process, ISO 8473 PDUs to be forwarded – (5) in figure 2.

It performs a lookup in the appropriate[4] Forwarding Database – (11) – to determine the possible output adjacencies to use for forwarding to a given destination , chooses one adjacency – (12) —, generates error indications to ISO 8473 – (14) , and signals ISO 9542 to issue Redirect PDUs – (13) .

6.8.1.4 Receive Process

The Receive Process obtains its inputs from the following sources

- received PDUs with the NPID of Intra-Domain routeing – (2) in figure 2,
- routeing information derived by the ES–IS protocol from the receipt of ISO 9542 PDUs – (1) ; (ISO 9542 places its routeing information directly into the routeing information base — (15)) and
- ISO 8473 data PDUs handed to the routeing function by the ISO 8473 protocol machine – (3) .

It then performs the appropriate actions, which may involve passing the PDU to some other function (e.g. to the Forwarding Process for forwarding – (5)).

[4] The appropriate Forwarding Database is selected by choosing a routeing metric based on fields in the QoS Maintenance option of ISO 8473.

7 Subnetwork Independent Functions

This clause describes the algorithms and associated databases used by the routeing functions. The managed objects and attributes defined for System Management purposes are described in Clause 11.

The following processes and data bases are used internally by the subnetwork independent functions. Following each process or data base title, in parentheses, is the type of systems which must keep the database. The system types are “L2” (level 2 Intermediate system), and “L1” (level 1 Intermediate system). Note that a level 2 Intermediate system is also a level 1 Intermediate system in its home area, so it must keep level 1 databases as well as level 2 databases.

Processes:

- Decision Process (L2, L1)
- Update Process (L2, L1)
- Forwarding Process (L2, L1)
- Receive Process (L2, L1)

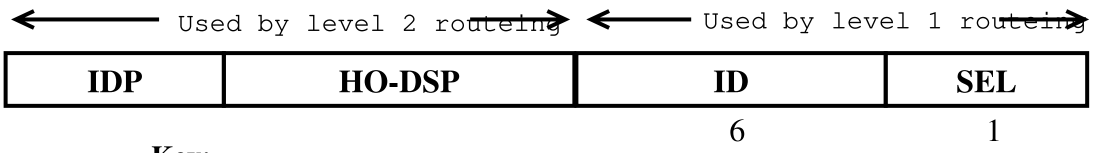
Databases:

- Level 1 Link State data base (L2, L1)
- Level 2 Link State data base (L2)
- Adjacency Database (L2, L1)
- Circuit Database (L2, L1)
- Level 1 Shortest Paths Database (L2, L1)
- Level 2 Shortest Paths Database (L2)
- Level 1 Forwarding Databases — one per routeing metric (L2, L1)
- Level 2 Forwarding Database — one per routeing metric (L2)

7.1 Addresses

7.1.1 Addresses within a routeing domain

Within a routeing domain an address is a variable length quantity conforming to the ISO NSAP address structure described in ISO 8348/Add.2. Addresses used in conjunction with this protocol are further structured as shown below in figure 4. Each field’s definition and use is defined below.



Key:

IDP	Initial Domain Part
HO-DSP	High Order Domain Specific Part
ID	System Identifier
SEL	NSAP Selector

Figure 3 - Address Structure for Intra-domain IS–IS Routeing

IDP, the Initial Domain Part of the address as defined by ISO 8348/Add.2. It consists of subfields AFI and IDI. Any of the AFI values defined by ISO 8348/Add.2 are legal NSAP addresses or Network entity titles when used in conjunction with this protocol.

HO-DSP, the high order part of the DSP as defined by ISO 8348/Add.2. The Intra-domain IS-IS protocol places no restrictions on the usage of this field, however a number of technical considerations apply to the assignment of DSP values in routeing domains conforming to ths protocol. See B.3 for more information.

The concatenation of IDP and high-order DSP is known as the *Area Address*.[5]

ID, a 6 octet *system identification*. Correct operation of the routeing protocol requires that this field be unique within an *area* for End Systems and Level 1 Intermediate systems, and unique within the routeing domain for Level 2 Intermediate systems.

NOTE — The use of an IEEE 802 MAC address, which is guaranteed to be globally unique, as the value of the ID field would ensure that the necessary area-wide uniqueness property is satisfied. If an 802 address is used it may correspond to the actual MAC address of the system on an ISO 8802 LAN, but this correspondence is neither assumed nor required by the routeing algorithms.

EDITOR\S NOTE — The current six octet length of the ID field has yet to be confirmed through member body ballot and comment. The current feeling is that this field should definitely remain of fixed size, in order to ensure that the ability of an Is to do fast relaying is not compromised. Also, the current opinion is that if the field size is changed, it may get larger (perhaps 7 or 8 octets) but not smaller. It is also possible that the length may be fixed on a per routeing domain basis instead of globally by the standard.

SELECTOR *NSAP Selector*, a 1 octet field which acts as a selector for the entity which is to receive the PDU (may be a Transport entity or the Intermediate system Network entity itself.) It is always the last octet of the address. The entire address minus the SELECTOR field is known as the *Network Entity Title* (NET) of the corresponding system.

The Area Address is the entire address minus the last 7 octets. An area address shall be unique within the entire global network addressing domain as defined in ISO 8348/Add.2.

If an address exactly matches one of the area addresses of a system, it is in the system’s area, and is routed by level 1 routeing. Otherwise it is routed via level 2 routeing.

Level 2 routeing acts on address prefixes, routeing towards the longest reachable address prefix that matches the destination address.

7.1.2 Address encoding

All full NSAP addresses and Network entity titles are encoded according to the ISO *preferred binary encoding*. This applies not only to addresses contained in ISO 8473 and ISO 9542 PDUs, but also to addresses carried by this protocol’s PDUs (such as LSPs and IIH PDUs). For details of the encoding of address prefixes see 7.1.4.

[5] An area may have more than one area address. See clause 7.2.11 for more details.

7.1.3 Permissible End System Address Formats

End systems are not required to follow the addressing convention defined above in 7.1.1. However, the following restrictions apply in order for Intermediate systems to route correctly in the presence of such End systems in a routeing domain:

a) Two End systems shall not have addresses that match all but the SELECTOR field. (Intermediate systems will not distinguish between such addresses, since the SELECTOR field is ignored for the purposes of identifying the system to which a PDU is to be delivered. It is used only within the destination system, to identify the entity to which the PDU should be delivered.)

b) An End system whose address does not match the Area Address of a neighbour Level 1 Intermediate system will not be able to use that Level 1 IS as a relay. (Level 1 Intermediate systems route solely based on the ID portion of the address, and only for addresses that match their Area Address.)

c) An End system whose address matches the Area Address of an area in a domain, shall reside in that area of that domain. (Level 1 Intermediate systems in the matching area will assume the location of the system is in that area.)

d) Two End systems residing in the same Level 1 area shall not have addresses with the same ID field.

7.1.4 Encoding of Level 2 Addresses

When a full NSAP address is encoded according to the preferred binary encoding specified in ISO 8348/Add.2, the IDI is padded with leading digits (if necessary) to obtain the maximum IDP length specified for that AFI.

A Level 2 address prefix consists of a leading sub-string of a full NSAP address, such that it matches a set of full NSAP addresses that have the same leading sub-string. However this truncation and matching is performed on the NSAP represented by the abstract syntax of the NSAP address, not on the encoded (and hence padded) form.[6]

Level 2 address prefixes are encoded in LSPs in the same way as full NSAP addresses, except when the end of the prefix falls within the IDP. In this case the prefix is directly encoded as the string of semi-octets with no padding.

7.2 The Decision Process

This process uses the database of Link State information to calculate the forwarding database(s), from which the forwarding process can know the proper next hop for each NPDU. The Level 1 Link State Database is used for calculating the Level 1 Forwarding Database(s), and the Level 2 Link State Database is used for calculating the Level 2 Forwarding Database(s).

7.2.1 Input and output

INPUT

- Link State Database – This database is a set of information from the latest Link State PDUs from all known Intermediate systems (within this area, for Level 1, or within the level 2 subdomain, for Level 2). This database is received from the Update Process.

- Notification of an Event – This is a signal from the Update Process that a change to a link has occurred somewhere in the domain.

OUTPUT

- Level 1 Forwarding Databases — one per routeing metric

- (Level 2 Intermediate systems only) Level 2 Forwarding Databases — one per routeing metric

[6] An example of prefix matching may be found in annex B, clause B.1.

- (Level 2 Intermediate systems only) The Level 1 Decision Process informs the Level 2 Update Process of the ID of the Level 2 Intermediate system within the area with lowest ID reachable with real level 1 links (as opposed to a virtual link consisting of a path through the level 2 subdomain)

- (Level 2 Intermediate systems only) If this Intermediate system is the “partition designated level 2 Intermediate system” in this partition, the Level 2 Decision Process informs the Level 1 Update Process of the values of the default routeing metric to and ID of the “partition designated level 2 Intermediate system” in each other partition of this area.

7.2.2 Routeing metrics

There are four routeing metrics defined, corresponding to the four possible orthogonal qualities of service defined by the QoS Maintenance field of ISO 8473. Each circuit emanating from an Intermediate system shall be assigned a value for one or more of these metrics by System management. The four metrics are as follows:

a) *Default metric:* This is a metric understood by every Intermediate system in the domain. Each circuit shall have a positive integral value assigned for this metric. The value may be associated with any objective function of the circuit, but by convention is intended to measure the *capacity* of the circuit for handling traffic, for example, its throughput in bits-per-second. Higher values indicate a lower capacity.

b) *Delay metric:* This metric measures the *transit delay* of the associated circuit. It is an optional metric, which if assigned to a circuit shall have a positive integral value. Higher values indicate a longer transit delay.

c) *Expense metric:* This metric measures the *monetary cost* of utilising the associated circuit. It is an optional metric, which if assigned to a circuit shall have a positive integral value[7]. Higher values indicate a larger monetary expense.

d) *Error metric:* This metric measures the *residual error probability* of the associated circuit. It is an optional metric, which if assigned to a circuit shall have a non-zero value. Higher values indicate a larger probability of undetected errors on the circuit.

NOTE - The decision process combines metric values by simple addition. It is important, therefore, that the values of the metrics be chosen accordingly.

Every Intermediate system shall be capable of calculating routes based on the default metric. Support of any or all of the other metrics is optional. If an Intermediate system supports the calculation of routes based on a metric, its update process may report the metric value in the LSPs for the associated circuit; otherwise, the IS shall not report the metric.

When calculating paths for one of the optional routeing metrics, the decision process only utilises LSPs with a value reported for the corresponding metric. If no value is associated with a metric for any of the IS’s circuits the system shall not calculate routes based on that metric.

NOTE - A consequence of the above is that a system reachable via the default metric may not be reachable by another metric.

See 7.4.2 for a description of how the forwarding process selects one of these metrics based on the contents of the ISO 8473 QoS Maintenance option.

7.2.3 Broadcast Subnetworks

Instead of treating a broadcast subnetwork as a fully connected topology, the broadcast subnetwork is treated as a pseudonode, with links to each attached system. Attached systems shall only report their link to the pseudonode. The

[7]The path computation algorithm utilised in this International Standard requires that all circuits be assigned a positive value for a metric. Therefore, it is not possible to represent a “free” circuit by a zero value of the expense metric. By convention, the value 1 is used to indicate a “free” circuit.

designated Intermediate system, on behalf of the pseudonode, shall construct Link State PDUs reporting the links to all the systems on the broadcast subnetwork with a zero value for each supported routeing metric[8].

The pseudonode shall be identified by the sourceID of the Designated Intermediate system, followed by a non-zero pseudonodeID assigned by the Designated Intermediate system. The pseudonodeID is locally unique to the Designated Intermediate system.

Designated Intermediate systems are determined separately for level 1 and level 2. They are known as the *LAN Level 1 Designated IS* and the *LAN Level 2 Designated IS* respectively. See Clause 8.4.4.

An Intermediate system may resign as Designated Intermediate System on a broadcast circuit either because it (or it's SNPA on the broadcast subnetwork) is being shut down or because some other Intermediate system of higher priority has taken over that function. When an Intermediate system resigns as Designated Intermediate System, it shall initiate a network wide purge of its pseudonode Link State PDU(s) by setting their Remaining Lifetime to zero and performing the actions described in 7.3.16.4. A LAN Level 1 Designated Intermediate System purges Level 1 Link State PDUs and a LAN Level 2 Designated Intermediate System purges Level 2 Link State PDUs. An Intermediate system which has resigned as both Level 1 and Level 2 Designated Intermediate System shall purge both sets of LSPs.

When an Intermediate system declares itself as designated Intermediate system and it is in possession of a Link State PDU of the same level issued by the previous Designated Intermediate System for that circuit (if any), it shall initiate a network wide purge of that (or those) Link State PDU(s) as above.

7.2.4 Links

Two Intermediate systems are not considered neighbours unless each reports the other as directly reachable over one of their SNPAs. On a Connection-oriented subnetwork (either point-to-point or general topology), the two Intermediate systems in question shall ascertain their neighbour relationship when a connection is established and hello PDUs exchanged. A malfunctioning IS might, however, report another IS to be a neighbour when in fact it is not. To detect this class of failure the decision process checks that each link reported as "up" in a LSP is so reported by both Intermediate systems. If an Intermediate system considers a link "down" it shall **not** mention the link in its Link State PDUs.

On broadcast subnetworks, this class of failure shall be detected by the designated IS, which has the responsibility to ascertain the set of Intermediate systems that can all communicate on the subnetwork. The designated IS shall include these Intermediate systems (and no others) in the Link State PDU it generates for the pseudonode representing the broadcast subnetwork.

7.2.5 Multiple LSPs for the same system

The Update process is capable of dividing a single logical LSP into a number of separate PDUs for the purpose of conserving link bandwidth and processing (see 7.3.4). The Decision Process, on the other hand, shall regard the LSP with LSP Number zero in a special way. If the LSP with LSP Number zero and remaining lifetime > 0, is not present for a particular system then the Decision Process shall not process any LSPs with non-zero LSP Number which may be stored for that system.

The following information shall be taken only from the LSP with LSP Number zero. Any values which may be present in other LSPs for that system shall be disregarded by the Decision Process.

a) The setting of the "Infinite Hippity Cost" bit.

b) The value of the Intermediate system Type field.

c) The Area Addresses option.

[8] They are set to zero metric values since they have already been assigned metrics by the link to the pseudonode. Assigning a non-zero value in the pseudonode LSP would have the effect of doubling the actual value.

7.2.6 Routeing Algorithm Overview

The routeing algorithm used by the Decision Process is a *shortest path first (SPF)* algorithm. Instances of the algorithm are run independently and concurrently by all Intermediate systems in a routeing domain. Intra-Domain routeing of a PDU occurs on a hop-by-hop basis: that is, the algorithm determines only the next hop, not the complete path, that a data PDU will take to reach its destination. To guarantee correct and consistent route computation by every Intermediate system in a routeing domain, this International Standard depends on the following properties:

a) All Intermediate systems in the routeing domain converge to using identical topology information; and

b) Each Intermediate system in the routeing domain generates the same set of routes from the same input topology and set of metrics.

The first property is necessary in order to prevent inconsistent, potentially looping paths. The second property is necessary to meet the goal of determinism stated in 6.6.

A system executes the SPF algorithm to find a set of paths to a destination system in the routeing domain. The set may consist of:

a) a single path of minimum cost: these are termed *minimum cost paths*;

b) a set of paths of equal minimum cost: these are termed *equal minimum cost paths*; or

c) a set of paths which will get a PDU closer to its destination than the local system: these are called *downstream paths.*

Paths which do not meet the above conditions are illegal and shall not be used.

The Decision Process, in determining its paths, also ascertains the identity of the adjacency which lies on the first hop to the destination on each path. These adjacencies are used to form the Forwarding Database, which the forwarding process uses for relaying PDUs.

Separate route calculations are made for each pairing of a level in the routeing hierarchy (i.e. L1 and L2) with a supported routeing metric. Since there are four routeing metrics and two levels some systems may execute multiple instances of the SPF algorithm. For example,

- if an IS is a L2 Intermediate system which supports all four metrics and computes minimum cost routes for all metrics, it would execute the SPF calculation eight times.

- if an IS is a L1 Intermediate system which supports all four metrics, and additionally computes downstream routes, it would execute the algorithm 4×(number of neighbours + 1) times.

Any implementation of an SPF algorithm meeting both the static and dynamic conformance requirements of clause 12 of this International Standard may be used. Recommended implementations are described in detail in Annex C.

7.2.7 Removal of Excess Adjacencies

When there are more than maximumPathSplits minimal cost paths to a destination, this set shall be pruned until only maximumPathSplits remain. The Intermediate system shall discriminate based upon:

NOTE - The precise precedence among the paths is specified in order to meet the goal of determinism define in 6.6.

- **adjacency type:** End system or level 2 reachable address prefix adjacencies are retained in preference to other adjacencies

- **neighbour ID:** where two or more adjacencies are of the same type, an adjacency with a lower neighbour ID is retained in preference to an adjacency with a higher neighbour id.

- **circuit ID:** where two or more adjacencies are of the same type, and same neighbour ID, an adjacency with a lower *circuit ID* is retained in preference to an adjacency with a higher *circuit ID*, where *circuit ID* is the value of:
  - ptPtCircuitID for non-broadcast circuits,
  - L1CircuitID for broadcast circuits when running the Level 1 Decision Process, and
  - L2CircuitID for broadcast circuits when running the Level 2 Decision Process.
- **LANAddress:** where two or more adjacencies are of the same type, same neighbour ID, and same circuit ID (e.g. a system with multiple LAN adapters on the same circuit) an adjacency with a lower LANAddress is retained in preference to an adjacency with a higher LANAddress.

7.2.8 Robustness Checks

The Decision Process shall not utilise a link between two Intermediate Systems unless both ISs report the link.

NOTE - the check is not applicable to links to an End System.

Reporting the link indicates that it has a defined value for at least the default routeing metric. It is permissible for two endpoints to report different defined values of the same metric for the same link. In this case, routes may be asymmetric.

7.2.9 Construction of a Forwarding Database

The information that is needed in the forwarding database for routeing metric *k* is the set of adjacencies for each system *N*.

7.2.9.1 Identification of Nearest Level 2 IS by a Level 1 IS

Level 1 Intermediate systems need one additional piece of information per routeing metric: the next hop to the nearest level 2 Intermediate system according to that routeing metric. A level 1 IS shall ascertain the set, *R*, of “attached” level 2 Intermediate system(s) for metric *k* such that the total cost to *R* for metric *k* is minimal.

If there are more adjacencies in this set than maximumPathSplits, then the IS shall remove excess adjacencies as described in 7.2.7.

7.2.9.2 Setting the Attached Flag in Level 2 Intermediate Systems

If a level 2 Intermediate system discovers, after computing the level 2 routes for metric *k*, that it cannot reach any other areas using that metric, it shall:

- set AttachedFlag for metric *k* to False;
- regenerate its Level 1 LSP with LSP number zero; and
- compute the “nearest level 2 Intermediate system” for metric *k* for insertion in the appropriate forwarding database, according to the algorithm described in 7.2.9.1 for level 1 Intermediate systems.

NOTE - AttachedFlag for each metric *k* is examined by the Update Process, so that it will report the value in the ATT field of its Link State PDUs.

If a level 2 Intermediate system discovers, after computing the level 2 routes for metric *k*, that it can reach at least one other area using that metric, it shall

- set AttachedFlag for metric *k* to True;
- regenerate its Level 1 LSP with LSP number zero; and

- set the level 1 forwarding database entry for metric *k* which corresponds to "nearest level 2 Intermediate system" to SELF.

7.2.10 Information for Repairing Partitioned Areas

An area may become partitioned as a result of failure of one or more links in the area. However, if each of the partitions has a connection to the level 2 subdomain, it is possible to repair the partition via the level 2 subdomain, provided that the level 2 subdomain itself is not partitioned. This is illustrated in Figure 4.

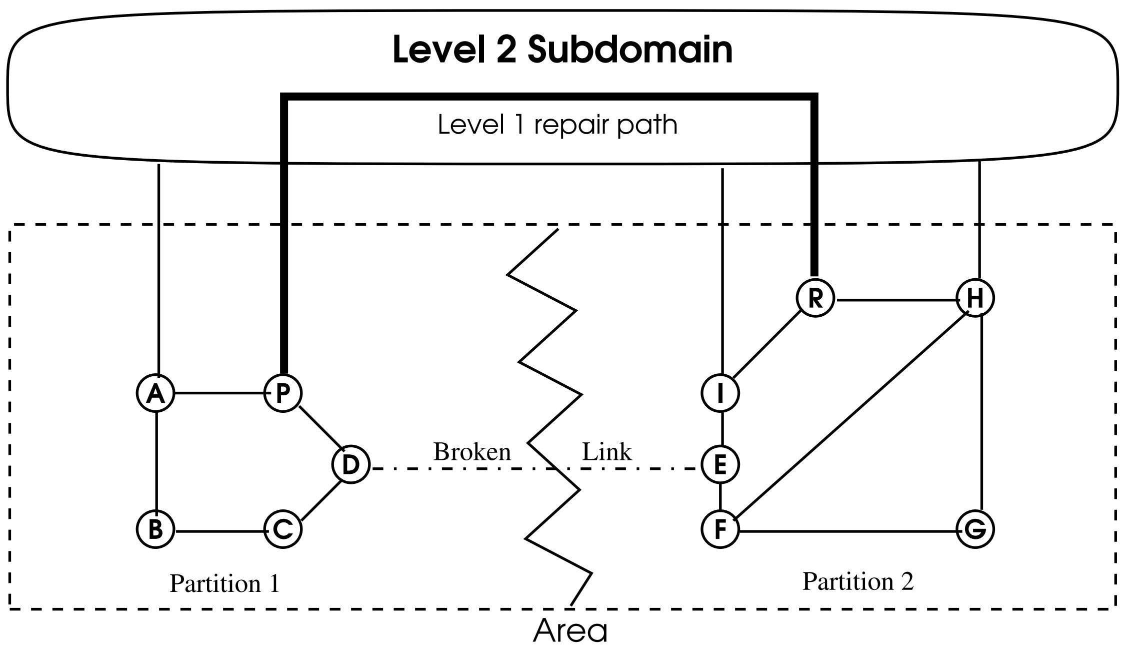


Figure 4 - Repair of partitioned level 1 area

All the systems A – I, R and P are in the same area *n*. When the link between D and E is broken, the area becomes partitioned. Within each of the partitions the *Partition Designated Level 2 Intermediate system* is selected from among the level 2 Intermediate systems in that partition. In the case of partition 1 this is P, and in the case of partition 2 this is R. The level 1 repair path is then established between between these two level 2 Intermediate systems. Note that the repaired link is now between P and R, not between D and E.

The Partition Designated Level 2 Intermediate Systems repair the partition by forwarding NPDUs destined for other partitions of the area through the level 2 subdomain. They do this by acting in their capacity as Level 1 Intermediate Systems and advertising in their Level 1 LSPs adjacencies to each Partition Designated Level 2 Intermediate System in the area. This adjacency is known as a "Virtual Adjacency" or "Virtual Link". Thus other Level 1 Intermediate Systems in a partition calculate paths to the other partitions through the Partition Designated Level 2 Intermediate System. A Partition Designated Level 2 Intermediate System forwards the Level 1 NPDUs through the level 2 subdomain by encapsulating them in 8473 Data NPDUs with its Virtual Network Entity Title as the source NSAP and the "adjacent" Partition Designated Level 2 Intermediate System's Virtual Network Entity Title as the destination NSAP. The following subclauses describe this in more detail.

7.2.10.1 Partition Detection and Virtual Level 1 Link Creation

Partitions of a Level 1 area are detected by the Level 2 Intermediate System(s) operating within the area. In order to participate in the partition repair process, these Level 2 Intermediate systems must also act as Level 1 Intermediate systems in the area. A partition of a given area exists whenever two or more Level 2 ISs located in that area are reported in the L2 LSPs as being a 'Partition Designated Level 2 IS'. Conversely, when only one Level 2 IS in an area is reported as being the 'Partition Designated Level 2 IS', then that area is not partitioned. Partition repair is accomplished by the Partition Designated Level 2 IS. The election of the Partition Designated Level 2 IS as described in the next subsection must be done before the detection and repair process can begin.

In order to repair a partition of a Level 1 area, the Partition designated Level 2 IS creates a "Virtual Network Entity" to represent the partition. The Network Entity Title for this virtual network entity shall be constructed from the first listed Area Address from its Level 2 Link State PDU, and the ID of the Partition Designated Level 2 IS. The IS shall also construct a virtual link (represented by a new Virtual Adjacency managed object) to each Partition Designated Level 2 IS in the area, with the NET of the partition recorded in the Identifier attribute. The virtual links are the repair paths for the partition. They are reported by the Partition Designated Level 2 IS into the entire Level 1 area by adding the ID of each "adjacent" Partition Designated Level 2 IS to the "Intermediate System Neighbors" part of its Level 1 Link State PDU. The "Virtual Flag" shall be set TRUE for these Intermediate System neighbors. The metric value for this virtual link shall be the default metric value d(N) obtained from this system's Level 2 PATHS database, where N is the "adjacent" Partition Designated Level 2 IS via the Level 2 subdomain.

An Intermediate System which operates as the Partition Designated Level 2 Intermediate System shall perform the following steps after completing the Level 2 shortest path computation in order to detect partitions in the Level 1 area and create repair paths:

a) Examine Level 2 Link State PDUs of all Level 2 Intermediate systems. Search Area Addresses for any address that matches any of the addresses in partitionAreaAddresses. If a match is found, and the Partition Designated Level 2 Intermediate system's ID does not equal this system's ID, then inform the level 1 update process at this system of the identity of the Partition Designated Level 2 Intermediate system, together with the path cost for the default routeing metric to that Intermediate system.

b) Continue examining Level 2 LSPs until all Partition Designated Level 2 Intermediate systems in other partitions of this area are found, and inform the Level 1 Update Process of all of the other Partition Designated Level 2 Intermediate systems in other partitions of this area, so that

   1) Level 1 Link State PDUs can be propagated to all other Partition designated level 2 Intermediate systems for this area (via the level 2 subdomain).

   2) All the Partition Designated Level 2 Intermediate systems for other partitions of this area can be reported as adjacencies in this system's Level 1 Link State PDUs.

If a partition has healed, the IS shall destroy the associated virtual network entity and virtual link by deleting the Virtual Adjacency. The Partition Designated Level 2 IS detects a healed partition when another Partition Designated Level 2 IS listed as a virtual link in its Level 1 Link State PDU was not found after running the partition detection and virtual link creation algorithm described above.

If such a Virtual Adjacency is created or destroyed, the IS shall generate a "PartitionVirtualLinkChange" event.

7.2.10.2 Election of Partition Designated Level 2 Intermediate System

The Partition Designated Level 2 IS is a Level 2 IS which:

- reports itself as “attached” by the default metric in its LSPs;
- reports itself as implementing the partiton repair option;
- operates as a Level 1 IS in the area;
- is reachable via Level 1 routeing without traversing any virtual links; and
- has the lowest ID

The election of the Partition Designated Level 2 IS is performed by running the decision process algorithm after the Level 1 decision process has finished, and before the Level 2 decision process to determine Level 2 paths is executed.

In order to guarantee that the correct Partition Designated Level 2 IS is elected, the decision process is run using only the Level 1 LSPs for the area, and by examining only the Intermediate System Neighbors whose Virtual Flag is FALSE. The results of this decision process is a set of all the Level 1 Intermediate Systems in the area that can be reached via Level 1, non-virtual link routeing. From this set, the Partition Designated Level 2 IS is selected by choosing the IS for which

- IS TYPE (as reported in the Level 1 LSP) is “Level 2 Intermediate System”;

- ATT indicates “attatched” by the default metric;

- P indicates support for the partiton repair option; and

- ID is the lowest among the subset of attached Level 2 Intermediate Systems.

7.2.10.3 Computation of Partition Area Addresses

A Level 2 Intermediate System shall compute the set of partitionAreaAddresses, which is the union of all manualAreaAddresses as reported in the Level 1 Link State PDUs of all Level 2 Intermediate systems reachable in the partition by the traversal of non-virtual links. If more than MaximumAreaAddresses are present, the Intermediate system shall retain only those areas with numerically lowest Area Address (as described in 7.2.11.1). If one of the local system’s ManualAreaAddresses is so rejected the event “Manual Address Dropped From Area” shall be generated.

7.2.10.4 Encapsulation of NPDUs Across the Virtual Link

All NPDUs sent over virtual links shall be encapsulated as ISO 8473 Data NPDUs. The encapsulating Data NPDU shall contain the Virtual Network Entity Title of the Partition Designated Level 2 IS that is forwarding the NPDU over the virtual link in the Source Address field, and the Virtual NET of the “adjacent” Partition Designated Level 2 IS in the Destination Address field. The SELECTOR field in both NSAPs shall contain the “IS-IS routeing” selector value. The QoS Maintenence field of the outer PDU shall be set to indicate forwarding via the default routeing metric (see table 1 on page 44).

For Data and Error Report NPDUs the Segmentation Permitted and Error Report flags and the Lifetime parameter of the outer NPDU shall be copied from the inner NPDU. When the inner NPDU is decapsulated, its lifetime parameter shall be decremented by the difference between the existing value of the parameter and the value of the Lifetime parameter in the outer NPDU.

For LSPs and SNPs the Segmentation Permitted flag shall be set to “True” and the Error Report flag shall be set to “False”. The Lifetime Parameter shall be set to 255. When an inner LSP is decapsulated, its remaining lifetime shall be decremented by half the difference between 255 and the value of the Lifetime parameter in the outer NPDU.

Data NPDUs shall not be fragmented before encapsulation, unless the total length of the Data NPDU (including header) exceeds 65535 octets. In that case, the original Data NPDU shall first be fragmented, then encapsulated. In all cases, the encapsulated Data NPDU may need to be fragmented by ISO 8473 before transmission in which case it must be reassembled and decapsulated by the destination Partition Designated Level 2 IS. The encapsulation is further described as part of the forwarding process in clause 7.4.3.2. The decapsulation is described as part of the Receive process in clause 7.4.4.

7.2.11 Computation of Area Addresses

A Level 1 or Level 2 Intermediate System shall compute the values of areaAddresses (the set of Area Addresses for this Level 1 area), by forming the union of the sets of manualAreaAddresses reported in the Area Addresses field of all Level 1 LSPs with LSP number zero in the local Intermediate system’s link state database.

NOTE - This includes all source systems, whether currently reachable or not. It also includes the local Intermediate system’s own Level 1 LSP with LSP number zero.

NOTE - There is no requirement for this set to be updated immediately on each change to the database contents. It is permitted to defer the computation until the next running of the Decision Process.

If more than MaximumAreaAddresses are present, the Intermediate system shall retain only those areas with numerically lowest Area Address (as described in 7.2.11.1). If one of the local system's ManualAreaAddresses is rejected the event "Manual Address Dropped From Area" shall be generated.

7.2.11.1 Comparison of Area Addresses

Numerical comparison of Area Addresses shall be performed on the encoded form of the address (i.e. with any padding of the IDI having been performed) by padding the shorter Area Address with trailing zeros to the length of the longer Area Address and then performing a numerical comparison in which the AFI is most significant.

7.3 The Update Process

The Update Process is responsible for generating and propagating Link State information reliably throughout the routeing domain.

The Link State information is used by the Decision Process to calculate routes.

7.3.1 Input and Output

INPUT

- Adjacency Database – maintained by the Subnetwork Dependent Functions
- Reachable Address managed objects - maintained by Network Management
- Notification of Adjacency Database Change – notification by the Subnetwork Dependent Functions that an adjacency has come up, gone down, or changed cost. (Circuit up, Circuit down, Adjacency Up, Adjacency Down, and Cost change events)
- AttachedFlag – (level 2 Intermediate systems only), a flag computed by the Level 2 Decision Process indicating whether this system can reach (via level 2 routeing) other areas
- Link State PDUs – The Receive Process passes Link State PDUs to the Update Process, along with an indication of which adjacency it was received on.
- Sequence Numbers PDUs – The Receive Process passes Sequence Numbers PDUs to the Update Process, along with an indication of which adjacency it was received on.
- Other Partitions – The Level 2 Decision Process makes available (to the Level 1 Update Process on a Level 2 Intermediate system) a list of ⟨Partition Designated Level 2 Intermediate system, Level 2 default metric value⟩ pairs, for other partitions of this area.

OUTPUT

- Link State Database
- Signal to the Decision Process of an event, which is either the receipt of a Link State PDU with different information from the stored one, or the purging of a Link State PDU from the database. The reception of a Link State PDU which has a different sequence number or Remaining Lifetime from one already stored in the database, but has an identical variable length portion, shall not cause such an event.

NOTE - An implementation may compare the checksum of the stored Link State PDU, modified according to the change in sequence number, with the checksum of the received Link State PDU. If they differ, it may assume that the variable length portions are different and an event signalled to the Decision Process. However, if the checksums are the same, an octet for octet comparison must be made in order to determine whether or not to signal the event.

7.3.2 Generation of Local Link State Information

The Update Process is responsible for constructing a set of Link State PDUs. The purpose of these Link State PDUs is to inform all the other Intermediate systems (in the area, in the case of Level 1, or in the Level 2 subdomain, in the case of Level 2), of the state of the links between the Intermediate system that generated the PDUs and its neighbours.

The Update Process in an Intermediate system shall generate one or more new Link State PDUs under the following circumstances:

a) upon timer expiration;

b) when notified by the Subnetwork Dependent Functions of an Adjacency Database Change;

c) when a change to some Network Management characteristic would cause the information in the LSP to change (for example, a change in manualAreaAddresses).

7.3.3 Use of Manual Routeing Information

Manual routeing information is routeing information entered by system management. It may be specified in two forms.

a) *Manual Adjacencies*

b) *Reachable Addresses*

These are described in the following sub-clauses.

7.3.3.1 Manual Adjacencies

An End system adjacency may be created by System Management. Such an adjacency is termed a *manual End system adjacency.* In order to create a manual End system adjacency, system managements shall specify:

a) the (set of) system IDs reachable over that adjacency; and

b) the corresponding SNPA Address.

These adjacencies shall appear as adjacencies with type “Manual”, adjacencyType “End system” and state “Up”. Such adjacencies provide input to the Update Process in a similar way to adjacencies created through the operation of ISO 9542. When the state changes to “Up” the adjacency information is included in the Intermediate System’s own Level 1 LSPs.

NOTE - Manual End system adjacencies shall not be included in a Level 1 LSPs issued on behalf of a pseudonode, since that would presuppose that all Intermediate systems on a broadcast subnetwork had the same set of manual adjacencies as defined for this circuit.

7.3.3.2 Reachable Addresses

A Level 2 Intermediate system may have a number of Reachable Address managed objects created by System management. When a Reachable Address is in state “On” and its parent Circuit is also in state “On”, the name and each of its defined routeing metrics shall be included in Level 2 LSPs generated by this system.

7.3.4 Multiple LSPs

Because a Link State PDU is limited in size to ReceiveLSPBufferSize, it may not be possible to include information about all of a system’s neighbours in a single LSP. In such cases, a system may use multiple LSPs to convey this information. Each LSP in the set carries the same sourceID field (see clause 9), but sets its own LSP Number field individually. Each of the several LSPs is handled independently by the Update Process, thus allowing distribution of topology updates to be pipelined. However, the Decision Process recognises that they all pertain to a common originating system because they all use the same sourceID.

NOTE - Even if the amount of information is small enough to fit in a single LSP, a system may optionally choose to use several LSPs to convey it; use of a single LSP in this situation is not mandatory.

The maximum sized Level 1 or Level 2 LSP which may be generated by a system is controlled by the values of the management parameters originatingL1LSPBufferSize or originatingL2LSPBufferSize respectively.

NOTE - These parameters should be set consistently by system management. If this is not done, some adjacencies will fail to initialise.

The IS shall treat the LSP with LSP Number zero in a special way, as follows:

a) The following fields are meaningful to the decision process only when they are present in the LSP with LSP Number zero:

   1) The setting of the “Infinite Hippity Cost” bit.
   2) The value of the Intermediate System Type field.
   3) The Area Addresses option. (This is only present in the LSP with LSP Number zero, see below).

b) When the values of any of the above items are changed, an Intermediate System shall re-issue the LSP with LSP Number zero, to inform other Intermediate Systems of the change. Other LSPs need not be reissued.

Once a particular adjacency has been assigned to a particular LSP Number, it is desirable that it not be moved to another LSP Number. This is because moving an adjacency from one LSP to another can cause temporary loss of connectivity to that system. This can occur if the new version of the LSP which originally contained information about the adjacency (which now does not contain that information) is propagated before the new version of the other LSP (which now contains the information about the adjacency). In order to minimise the impact of this, the following restrictions are placed on the assignment of information to LSPs.

a) The Area Addresses option field shall occur only in the LSP with LSP Number zero.

b) Intermediate System Neighbours options shall occur after the Area Addresses option and before any End system (or in the case of Level 2, Prefix) Neighbours options.

c) End system (or Prefix) Neighbour options (if any) shall occur after any Area Address or Intermediate System Neighbour options.

   NOTE — In this context, “after” means at a higher octet number from the start of the same LSP or in an LSP with a higher LSP Number.

   NOTE — An implementation is recommended to ensure that the number of LSPs generated for a particular system is within approximately 10% of the optimal number which would be required if all LSPs were densely packed with neighbour options. Where possible this should be accomplished by re-using space in LSPs with a lower LSP Number for new adjacencies. If it is necessary to move an adjacency from one LSP to another, the SRMflags (see clause 7.3.15) for the two new LSPs shall be set as an atomic action.[9]

When some event requires changing the LSP information for a system, the system shall reissue that (or those) LSPs which would have different contents. It is not required to reissue the unchanged LSPs. Thus a single End system adjacency change only requires the reissuing of the LSP containing the End system Neighbours option referring to that adjacency. The parameters maximumLSPGenerationInterval and minimumLSPGenerationInterval shall apply to each LSP individually.

[9]If the two SRMflags are not set atomically, a race condition will exist in which one of the two LSPs may be propagated quickly, while the other waits for an entire propagation cycle. If this occurs, adjacencies will be falsely eliminated from the topology and routes may become unstable for period of time potentially as large as maximumLSPGeneratonInterval.

7.3.5 Periodic LSP Generation

The Update Process shall periodically re-generate and propagate on every circuit with an IS adjacency of the appropriate level (by setting SRMflag on each circuit), all the LSPs (Level 1 and/or Level 2) for the local system and any pseudonodes for which it is responsible. The Intermediate system shall re-generate each LSP at intervals of at most MaximumLSPGenerationInterval seconds, with jitter applied as described in 10.1.

These LSPs may all be generated on expiration of a single timer or alternatively separate timers may be kept for each LSP Number and the individual LSP generated on expiration of this timer.

7.3.6 Event Driven LSP Generation

In addition to the periodic generation of LSPs, an Intermediate system shall generate an LSP when an event occurs which would cause the information content to change. The following events may cause such a change.

- an Adjacency or Circuit Up/Down event
- a change in Circuit metric
- a change in Reachable Address metric
- a change in ManualAreaAddresses
- a change in SystemID
- a change in Designated Intermediate System status
- a change in the waiting status

When such an event occurs the IS shall re-generate changed LSP(s) with a new sequence number. If the event necessitated the generation of an LSP which had not previously been generated (for example, an adjacency "Up" event for an adjacency which could not be accommodated in an existing LSP), the sequence number shall be set to one. The IS shall then propagate the LSP(s) on every circuit by setting SRMflag for each circuit. The timer maximumLSPGenerationInterval shall **not** be reset.

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

7.3.7 Generation of Level 1 LSPs (non-pseudonode)

The Level 1 Link State PDU not generated on behalf of a pseudonode contains the following information in its variable length fields.

- In the *Area Addresses* option— the set of manualAreaAddresses for this Intermediate System.
- In the *Intermediate System Neighbours* option — the set of 7-octet Intermediate system IDs of neighbouring Intermediate systems formed from:
  - The set of neighbourSystemIDs with an appended zero octet (indicating non-pseudonode) from adjacencies in the state "Up", on circuits of type "Point-Point", "In" or "Out", with
    - x adjacencyType "L1 Intermediate System"
    - x adjacencyType "L2 Intermediate System" and L2OnlyMode "False".

    The metrics shall be set to the values of Level 1 metric_k of the circuit for each supported routeing metric.
  - The set of L1CircuitIDs for all circuits of type "Broadcast". (i.e. the neighbouring pseudonode IDs) .

    The metrics shall be set to the values of Level 1 metric_k of the circuit for each supported routeing metric.

- The set of IDs with an appended zero octet derived from the Network Entity Titles of all Virtual Adjacencies of this IS. (Note that the Virtual Flag is set when encoding these entries in the LSP — see 7.2.10.)

  The default metric shall be set to the total cost to the virtual NET for the default routeing metric. The remaining metircs shall be set to the value indicating *unsupported*.

- In the *End System Neighbours* option — the set of 6-octet IDs of neighbouring End systems formed from:

  - The SystemID of the Intermediate System itself, with a value of zero for all supported metrics.

  - The set of endSystemIDs from all adjacencies with type “Auto-configured”, in state “Up”, on circuits of type “Point-to-Point”, “In” or “Out”, with adjacencyType “End system”.

    The metrics shall be set to the values of Level 1 metric_k of the circuit for each supported routeing metric.

  - The set of endSystemIDs from all adjacencies with type “Manual” in state “Up”, on all circuits.

    The metrics shall be set to the values of Level 1 metric_k of the circuit for each supported routeing metric.

7.3.8 Generation of Level 1 pseudonode LSPs

An IS shall generate a Level 1 pseudonode Link State PDU for each circuit for which this Intermediate System is the Level 1 LAN Designated Intermediate System. The LSP shall specify the following information in its variable length fields. In all cases a value of zero shall be used for all supported routeing metrics

- The Area Addresses option is not present.

  Note - This information is not required since the set of Area Addresses for the node issueing the pseudonode LSP will already have been made available via its own non-pseudonode LSP.

- In the Intermediate System Neighbours option — the set of 7-octet Intermediate System IDs of neighbouring Intermediate Systems on the circuit for which this pseudonode LSP is being generated formed from:

  - The Designated Intermediate System’s own SystemID with an appended zero octet (indicating non-pseudonode).

  - The set of neighbourSystemIDs with an appended zero octet (indicating non-pseudonode) from adjacencies on this circuit in the state “Up”, with adjacencyType “L1 Intermediate System”.

- In the End system Neighbours option — the set of 6 octet IDs of neighbouring End systems formed from:

  - The set of endsystemIDs from all adjacencies with type “Auto-configured”, in state “Up”, on the circuit for which this pseudonode is being generated, with adjacencyType “End system”.

7.3.9 Generation of Level 2 LSPs (non-pseudonode)

The Level 2 Link State PDU not generated on behalf of a pseudonode contains the following information in its variable length fields:

- In the Area Addresses option — the set of areaAddresses for this Intermediate system computed as described in 7.2.11.

- In the Partition Designated Level 2 IS option — the ID of the Partition Designated Level 2 Intermediate System for the partition.

- In the Intermediate System Neighbours option — the set of 7-octet Intermediate system IDs of neighbouring Intermediate systems formed from:

  - The set of neighbourSystemIDs with an appended zero octet (indicating non-pseudonode) from adjacencies in the state “Up”, on circuits of type “Point-to-Point”, “In” or “Out”, with adjacencyType “L2 Intermediate System”.

- The set of L2CircuitIDs for all circuits of type "Broadcast". (i.e. the neighbouring pseudonode IDs)

  The metrics shall be set to the values of Level 2 $metric_k$ of the circuit for each supported routeing metric.

- In the Prefix Neighbours option — the set of variable length prefixes formed from—

  - The set of names of all Reachable Address managed objects in state "On", on all circuits in state "On".

    The metrics shall be set to the values of Level 2 $metric_k$ for the reachable address.

7.3.10 Generation of Level 2 pseudonode LSPs

A Level 2 pseudonode Link State PDU is generated for each circuit for which this Intermediate System is the Level 2 LAN Designated Intermediate System and contains the following information in its variable length fields. In all cases a value of zero shall be used for all supported routeing metrics.

- The Area Addresses option is not present.

  Note - This information is not required since the set of Area Addresses for the node issueing the pseudonode LSP will already have been made available via its own non-pseudonode LSP.

- In the Intermediate System Neighbours option — the set of 7 octet Intermediate System IDs of neighbouring Intermediate Systems on the circuit for which this pseudonode LSP is being generated formed from:

  - The Designated Intermediate System's own SystemID with an appended zero octet (indicating non-pseudonode).

  - The set of neighbourSystemIDs with an appended zero octet (indicating non-pseudonode) from adjacencies on this circuit in the state "Up" with adjacencyType "L2 Intermediate System".

- The prefix neighbours option is not present.

7.3.11 Generation of the Checksum

This International Standard makes use of the checksum function defined in ISO 8473.

The source IS shall compute the LSP Checksum when the LSP is generated. The checksum shall never be modified by any other system. The checksum allows the detection of memory corruptions and thus prevents both the use of incorrect routeing information and its further propagation by the Update Process.

The checksum shall be computed over all fields in the LSP which appear after the Remaining Lifetime field. This field (and those appearing before it) are excluded so that the LSP may be aged by systems without requiring re-computation.

As an additional precaution against hardware failure, when the source computes the Checksum, it shall start with the two checksum variables (C0 and C1) initialised to what they would be after computing for the SystemID portion (i.e. the first 6 octets) of its SOURCE ID. (This value is computed and stored when the Network entity is enabled and whenever SystemID changes.) The IS shall then resume Checksum computation on the contents of the PDU after the first 6 octets of the SOURCE ID field.

NOTE - All Checksum calculations on the LSP are performed treating the SOURCE ID field as the first octet. This procedure prevents the source from accidentally sending out Link State PDUs with some other system's ID as source.

7.3.12 Initiating Transmission

The IS shall store the generated Link State PDU in the Link State Database, overwriting any previous Link State PDU with the same LSP Number generated by this system. The IS shall then set all SRMFlags for that Link State PDU, indicating it is to be propagated on all circuits with Intermediate System adjacencies.

An Intermediate System shall ensure (by reserving resources or otherwise) that it will always be able to store and internalise its own (and its own pseudonode) LSPs.

7.3.13 Preservation of order

When an existing Link State PDU is re-transmitted (with the same or a different sequence number), but with the same information content (i.e. the variable length part) as a result of there having been no changes in the local topology databases, the order of the information in the variable length part shall be the same as that in the previously transmitted LSP.

NOTE - If a sequence of changes result in the state of the database returning to some previous value, there is no requirement to preserve the ordering. It is only required when there have been no changes whatever. This allows the receiver to detect that there has been no change in the information content by performing an octet for octet comparison of the variable length part, and hence not re-run the decision process.

7.3.14 Propagation of LSPs

The update process is responsible for propagating Link State PDUs throughout the domain (or in the case of Level 1, throughout the area).

The basic mechanism is flooding, in which each Intermediate system propagates to all its neighbour Intermediate systems except that neighbour from which it received the PDU. Duplicates are detected and dropped.

Link state PDUs are received from the Receive Process. The maximum size control PDU (Link State PDU or Sequence Numbers PDU) which a system expects to receive shall be ReceiveLSPBufferSize octets. (i.e. the Update process must provide buffers of at least this size for the reception, storage and forwarding of received Link State PDUs and Sequence Numbers PDUs.) If a control PDU larger than this size is received, it shall be treated as if it had an invalid checksum (i.e. ignored by the Update Process and a “Corrupted LSP Received” event generated).

Upon receipt of a Link State PDU the Update Process shall perform the following functions:

a) Level 2 Link State PDUs shall be propagated on circuits which have at least one Level 2 adjacency.

b) Level 1 Link State PDUs shall be propagated on circuits which have at least one Level 1 adjacency or at least one Level 2 adjacency not marked “Level 2 only”.

c) When propagating a Level 1 Link State PDU on a broadcast subnetwork, the IS shall transmit to the multi-destination subnetwork address “All-L1-Intermediate systems”.

d) When propagating a Level 2 Link State PDU on a broadcast subnetwork, the IS shall transmit to the multi-destination subnetwork address “All-L2-Intermediate systems”.

NOTE — When propagating a Link State PDU on a general topology subnetwork the Data Link Address is unambiguous (because Link State PDUs are not propagated across Dynamically Assigned circuits).

e) An Intermediate system receiving a Link State PDU with an incorrect LSP Checksum or with an invalid PDU syntax shall

   1) log a circuit event, “Corrupted Link State PDU Received”,

   2) overwrite the Checksum and Remaining Lifetime with 0, and

   3) treat the Link State PDU as though its Remaining Lifetime had expired (see 7.3.16.4.)

f) A Intermediate system receiving a Link State PDU which is new (as identified in 7.3.16) shall

   1) store the Link State PDU into Link State database, and

   2) mark it as needing to be propagated upon all circuits except that upon which it was received.

g) When a Intermediate system receives a Link State PDU from source *S*, which it considers older than the one stored in the database for *S*, it shall set the SRM flag for *S*'s Link State PDU associated with the circuit from which the older Link State PDU was received. This indicates that the stored Link State PDU needs to be sent on the link from which the older one was received.

h) When a system receives a Link State PDU which is the same (not newer or older) as the one stored, the Intermediate system shall

1) acknowledge it if necessary, as described in 7.3.17, "Making the Update Reliable", and

2) clear the SRM flag for that circuit for that Link State PDU.

i) A Link State PDU received with a zero checksum shall be treated as if the Remaining Lifetime were 0. The age, if not 0, shall be overwritten with 0.

The Update Process scans the Link State Database for Link State PDUs with SRM Flags set. When one is found, provided the timestamp lastSent indicates that it was propagated no more recently than minimumLSPTransmissionInterval, the IS shall

a) transmit it on all circuits with SRM Flags set, and

b) update lastSent.

7.3.15 Manipulation of SRM and SSN Flags

For each Link State PDU, and for each circuit over which routeing messages are to be exchanged (i.e. not on DA circuits), there are two flags:

Send Routeing Message (SRM) – if set, indicates that Link State PDU should be transmitted on that circuit. On broadcast circuits SRM is cleared as soon as the LSP has been transmitted, but on non-broadcast circuits SRM is only cleared on reception of a Link State PDU or Sequence Numbers PDU as described below.

SRM shall never be set for an LSP with sequence number zero. (See 7.3.15.2)

Send Sequence Numbers (SSN) – if set, indicates that information about that Link State PDU should be included in a Partial Sequence Numbers PDU transmitted on that circuit. When the Sequence Numbers PDU has been transmitted SSN is cleared. Note that the Partial Sequence Numbers PDU serves as an acknowledgement that a Link State PDU was received.

7.3.15.1 Action on Receipt of a Link State PDU

When a Link State PDU is received on a circuit *C*, the IS shall perform the following functions:

a) If the LSP has zero Remaining Lifetime, perform the actions described in 7.3.16.4.

b) If the source *S* of the LSP is an IS or pseudonode for which the first 6 octets are equal to the SystemID of the receiving Intermediate System, and the receiving Intermediate System does not have that LSP in its database, or has that LSP, but no longer considers it to be in the set of LSPs generated by this system (e.g. it was generated by a previous incarnation of the system), then initiate a network wide purge of that LSP as described in 7.3.16.4.

c) If the source *S* of the LSP is a system (pseudonode or otherwise) for which the first 6 octets are equal to the SystemID of the receiving Intermediate system, and the receiving Intermediate system has an LSP in the set of currently generated LSPs from that source in its database (i.e. it is an LSP generated by this Intermediate system), perform the actions described in 7.3.16.1.

d) Otherwise, (the source *S* is some other system),

1) If the LSP is newer than the one in the database, or if an LSP from that source does not yet exist in the database:

   (i) Store the new LSP in the database, overwriting the existing database LSP for that source (if any) with the received LSP.

   (ii) Set SRMflag for that LSP for all circuits other than *C*.

   (iii) Clear SRMflag for *C*.

   (iv) If *C* is a non-broadcast circuit, set SSNflag for that LSP for *C*.

   (v) Clear SSNflag for that LSP for the circuits other than *C*.

2) If the LSP is equal to the one in the database (same Sequence Number, Remaining Lifetimes both zero or both non-zero, same checksums):

   (i) Clear SRMflag for *C*.

   (ii) If *C* is a non-broadcast circuit, set SSNflag for that LSP for *C*.

3) If the LSP is older than the one in the database:

   (i) Set SRMflag for *C*.

   (ii) Clear SSNflag for *C*.

When storing a new LSP, the Intermediate system shall first ensure that it has sufficient memory resources to both store the LSP and generate whatever internal data structures will be required to process the LSP by the Update Process. If these resources are not available the LSP shall be ignored. It shall neither be stored nor acknowledged. When an LSP is ignored for this reason the IS shall enter the "Waiting State". (See 7.3.19).

7.3.15.2 Action on Receipt of a Sequence Numbers PDU

When a Sequence Numbers PDU (Complete or Partial, see clause 7.3.17) is received on circuit *C* the IS shall perform the following functions:

a) For each LSP reported in the Sequence Numbers PDU:

   1) If the reported value equals the database value and *C* is a non-broadcast circuit, Clear SRMflag for *C* for that LSP.

   2) If the reported value is older than the database value, Clear SSNflag, and Set SRMflag.

   3) If the reported value is newer than the database value, Set SSNflag, and if *C* is a non-broadcast circuit Clear SRMflag.

   4) If no database entry exists for the LSP, and the reported Remaining Lifetime, checksum and sequence number of the LSP are all non-zero, create an entry with sequence number *0* (see 7.3.16.1.), and set SSNflag for that entry and circuit *C*. Under no circumstances shall SRMflag be set for such an LSP with zero sequence number.

   NOTE - This is because possessing a zero sequence number LSP is semantically equivalent to having no information about that LSP. If such LSPs were propagated by setting SRMflag it would result in an unnecessary consumption of both bandwidth and memory resources.

b) If the Sequence Numbers PDU is a Complete Sequence Numbers PDU, Set SRMflags for *C* for all LSPs in the database (except those with zero sequence number or zero remaining lifetime) with LSPIDs within the range specified for the CSNP by the START LSP ID and END LSP ID fields, which were not mentioned in the Complete Sequence Numbers PDU (i.e. LSPs this system has, which the neighbour does not claim to have).

7.3.15.3 Action on expiration of Complete SNP Interval

The IS shall perform the following actions every completeSNPInterval seconds for circuit *C*:

a) If *C* is a broadcast circuit, then

1) If this Intermediate system is a Level 1 Designated Intermediate System on circuit *C*, transmit a complete set of Level 1 Complete Sequence Numbers PDUs on circuit *C*. Ignore the setting of SSNflag on Level 1 Link State PDUs.

2) If this Intermediate system is a Level 2 Designated Intermediate System on circuit *C*, transmit a complete set of Level 2 Complete Sequence Numbers PDUs on circuit *C*. Ignore the setting of SSNflag on Level 2 Link State PDUs.

A complete set of CSNPs is a set whose startLSPID and endLSPID ranges cover the complete possible range of LSPIDs. (i.e. there is no possible LSPID value which does not appear within the range of one of the CSNPs in the set). Where more than one CSNP is transmitted on a broadcast circuit, they shall be separated by an interval of at least minimumBroadcastLSPTransmissionInterval seconds.

NOTE — An IS is permitted to transmit a small number of CSNPs (no more than 10) with a shorter separation interval, (or even "back to back"), provided that no more than *1000*/minimumBroadcastLSPTransmissionInterval CSNPs are transmitted in any one second period.

b) Otherwise (*C* is a point to point circuit, including non-DA DED circuits and virtual links), do nothing. CSNPs are only transmitted on point to point circuits at initialisation.

7.3.15.4 Action on expiration of Partial SNP Interval

The maximum sized Level 1 or Level 2 PSNP which may be generated by a system is controlled by the values of originatingL1LSPBufferSize or originatingL2LSPBufferSize respectively. An Intermediate system shall perform the following actions every PartialSNPInterval seconds for circuit *C* with jitter applied as described in 10.1:

a) If *C* is a broadcast circuit, then

1) If this Intermediate system is a Level 1 Intermediate System or a Level 2 Intermediate System with ManualL2OnlyMode "False", but is **not** a Level 1 Designated Intermediate System on circuit *C*, transmit a Level 1 Partial Sequence Numbers PDU on circuit *C*, containing entries for as many Level 1 Link State PDUs with SSNflag set as will fit in the PDU, and then clear SSNflag for these entries. To avoid the possibility of starvation, the scan of the LSP database for those with SSNflag set shall commence with the next LSP which was not included in the previous scan. If there were no Level 1 Link State PDUs with SSNflag set, do not transmit a Level 1 Partial Sequence Numbers PDU.

2) If this Intermediate system is a Level 2 Intermediate System, but is **not** a Level 2 Designated Intermediate System on circuit *C*, transmit a Level 2 Partial Sequence Numbers PDU on circuit *C*, containing entries for all as many Level 2 Link State PDUs with SSNflag set as will fit in the PDU, and then clear SSNflag for these entries. To avoid the possibility of starvation, the scan of the LSP database for those with SSNflag set shall commence with the next LSP which was not included in the previous scan. If there were no Level 2 Link State PDUs with SSNflag set, do not transmit a Level 2 Partial Sequence Numbers PDU.

b) Otherwise (*C* is a point to point circuit, including non-DA DED circuits and virtual links), for Level 1 and Level 2 (according to Intermediate system type) transmit a Partial Sequence Numbers PDU on circuit *C*, containing entries for as many Link State PDUs with SSNflag set as will fit in the PDU, and then clear SSNflag for these entries. To avoid the possibility of starvation, the scan of the LSP database for those with SSNflag set shall commence with the next LSP which was not included in the previous scan. If there were no Link State PDUs with SSNflag set, do not transmit a Partial Sequence Numbers PDU.

7.3.15.5 Action on expiration of Minimum LSP Transmission Interval

An IS shall perform the following actions every minimumLSPTransmissionInterval seconds with jitter applied as described in 10.1.

a) For all Point to Point circuits *C* transmit all LSPs that have SRMflag set on circuit *C*, but **do not** clear the SRMflag. The SRMflag will subsequently be cleared by receipt of a Complete or Partial Sequence Numbers PDU.

The interval between two consecutive transmissions of the same LSP shall be at least MinimumLSPTransmissionInterval. Clearly, this can only be achieved precisely by keeping a separate timer for each LSP. This would be an unwarranted overhead. Any technique which ensures the interval will be between MinimumLSPTransmissionInterval and *2 ** MinimumLSPTransmissionInterval is acceptable.

7.3.15.6 Controlling the Rate of Transmission on Broadcast Circuits

The Routeing characteristic minimumBroadcastLSPTransmissionInterval indicates the minimum interval between PDU arrivals which can be processed by the slowest Intermediate System on the LAN.

Setting SRMFlags on an LSP for a broadcast circuit does not cause the LSP to be transmitted immediately. Instead the Intermediate system shall scan the LSP database every minimumBroadcastLSPTransmissionInterval (with jitter applied as described in 10.1), and from the set of LSPs which have SRMFlags set for this circuit, one LSP shall be chosen at random. This LSP shall be multicast on the circuit, and SRMFlags cleared.

NOTE - In practice it would be very inefficient to scan the whole database at this rate, particularly when only a few LSPs had SRMFlags set. Implementations may require additional data structures in order to reduce this overhead.

NOTE - An IS is permitted to transmit a small number of LSPs (no more than 10) with a shorter separation interval, (or even “back to back”), provided that no more than *1000/*minimumBroadcastLSPTransmissionInterval LSPs are transmitted in any one second period.

In addition, the presence of any LSPs which have been received on a particular circuit and are queued awaiting processing shall inhibit transmission of LSPs on that circuit. However, LSPs may be transmitted at a minimum rate of one per second even in the presence of such a queue.

7.3.16 Determining the Latest Information

The Update Process is responsible for determining, given a received link state PDU, whether that received PDU represents new, old, or duplicate information with respect to what is stored in the database.

It is also responsible for generating the information upon which this determination is based, for assigning a sequence number to its own Link State PDUs upon generation, and for correctly adjusting the Remaining Lifetime field upon broadcast of a link state PDU generated originally by any system in the domain.

7.3.16.1 Sequence Numbers

The sequence number is a 4 octet unsigned value. Sequence numbers shall increase from zero to SequenceModulus. When a system initialises, it shall start with sequence number 1 for its own Link State PDUs.[10]

The sequence numbers the Intermediate system generates for its Link State PDUs with different values for LSP number are independent. The algorithm for choosing the numbers is the same, but operationally the numbers will not be synchronised.

If an Intermediate system *R* somewhere in the domain has information that the current sequence number for source *S* is greater than that held by *S*, *R* will return to *S* a Link State PDU for *S* with *R*’s value for the sequence number. When *S* receives this LSP it shall change its sequence number to be the next number greater than the new one received, and shall generate a link state PDU.

If an Intermediate system needs to increment its sequence number, but the sequence number is already equal to SequenceModulus, the event “Attempt to Exceed Maximum Sequence Number” shall be generated and the Routeing Module shall be disabled for a period of at least MaxAge + ZeroAgeLifetime, in order to be sure that any versions of this LSP with the high sequence number have expired. When it is re-enabled the IS shall start again with sequence number 1.

[10] It starts with 1 rather than 0 so that the value 0 can be reserved to be guaranteed to be less than the sequence number of any actually generated Link State PDU. This is a useful property for Sequence Numbers PDUs.

7.3.16.2 LSP Confusion

It is possible for an LSP generated by a system in a previous incarnation to be alive in the domain and have the same sequence number as the current LSP.

To ensure database consistency among the Intermediate Systems, it is essential to distinguish two such PDUs. This is done efficiently by comparing the checksum on a received LSP with the one stored in memory.

If the sequence numbers match, but the checksums do not and the LSP is not in the current set of LSPs generated by the local system, then the system that notices the mismatch shall treat the LSP as if its Remaining Lifetime had expired. It shall store one of the copies of the LSP, with zero written as the Remaining Lifetime, and flood the LSP.

If the LSP is in the current set of LSPs generated by the local system then the IS shall change the LSP's sequence number to be the next number greater than that of the received LSP and regenerate the LSP.

7.3.16.3 Remaining Lifetime Field

When the source generates a link state PDU, it shall set the Remaining Lifetime to MaxAge.

When a system holds the information for some time before successfully transmitting it to a neighbour, that system shall decrement the Remaining Lifetime field according to the holding time. Before transmitting a link state PDU to a neighbour, a system shall decrement the Remaining Lifetime in the PDU being transmitted by at least 1, or more than 1 if the transit time to that neighbour is estimated to be greater than one second. When the Remaining Lifetime field reaches 0, the system shall purge that Link State PDU from its database. In order to keep the Intermediate Systems' databases synchronised, the purging of an LSP due to Remaining Lifetime expiration is synchronised by flooding an expired LSP. See 7.3.16.4.

If the RemainingLifetime of the received LSP is zero it shall be processed as described in 7.3.16.4. If the Remaining Lifetime of the received LSP is non-zero, but there is an LSP in the database with the same sequence number and zero Remaining Lifetime, the LSP in the database shall be considered most recent. Otherwise, the PDU with the larger sequence number shall be considered the most recent.

If the value of Remaining Lifetime is greater than MaxAge, the LSP shall be processed as if there were a checksum error.

7.3.16.4 LSP Expiration Synchronisation

When the Remaining Lifetime on an LSP in memory becomes zero, the IS shall

a) set all SRMFlags for that LSP, and

b) retain only the LSP header.

c) record the time at which the Remaining Lifetime for this LSP became zero. When ZeroAgeLifetime has elapsed since the LSP Remaining Lifetime became zero, the LSP header shall be purged from the database.

NOTE - A check of the checksum of a zero Remaining Lifetime LSP succeeds even though the data portion is not present

When a purge of an LSP with non-zero Remaining Lifetime is initiated, the header shall be retained for MaxAge.

If an LSP from source *S* with zero Remaining Lifetime is received on circuit *C* :

a) If no LSP from *S* is in memory, then the IS shall

   1) send an acknowledgement of the LSP on circuit *C*, but

   2) shall not retain the LSP after the acknowledgement has been sent.

b) If an LSP from *S* is in the database, then

1) If the received LSP is newer than the one in the database (i.e. received LSP has higher sequence number, or same sequence number and database LSP has non-zero Remaining Lifetime) the IS shall:

   (i) overwrite the database LSP with the received LSP, and note the time at which the zero Remaining Lifetime LSP was received, so that after ZeroAgeLifetime has elapsed, that LSP can be purged from the database,

   (ii) set SRMflag for that LSP for all circuits other than *C*,

   (iii) clear SRMflag for *C*,

   (iv) if *C* is a non-broadcast circuit, set SSNflag for that LSP for *C*, and

   (v) clear SSNflag for that LSP for the circuits other than *C*.

2) If the received LSP is equal to the one in the database (i.e. same Sequence Number, Remaining Lifetimes both zero) the IS shall:

   (i) clear SRMflag for *C*, and

   (ii) if *C* is a non-broadcast circuit, set SSNflag for that LSP for *C*.

3) If the received LSP is older than the one in the database (i.e. received LSP has lower sequence number) the IS shall:

   (i) set SRMflag for *C*, and

   (ii) clear SSNflag for *C*.

c) If this system (or pseudonode) is *S* and there is an un-expired LSP from *S* (i.e. its own LSP) in memory, then the IS:

   1) shall not overwrite with the received LSP, but

   2) shall change the sequence number of the unexpired LSP from *S* as described in Clause 7.3.16.1,

   3) generate a new LSP; and

   4) set SRMflag on all circuits.

7.3.17 Making the Update Reliable

The update process is responsible for making sure the latest link state PDUs reach every reachable Intermediate System in the domain.

a) On point-to-point links the Intermediate system shall send an explicit acknowledgement encoded as a Partial Sequence Numbers PDU (PSNP)containing the following information:

   1) source's ID
   2) PDU type (Level 1 or 2)
   3) sequence number
   4) Remaining Lifetime
   5) checksum

   This shall be done for all received link state PDUs which are newer than the one in the database, or duplicates of the one in the database. Link state PDUs which are older than that stored in the database are answered instead by a newer link state PDU, as specified in 7.3.14 above.

b) On broadcast links, instead of explicit acknowledgements for each link state PDU by each Intermediate system, a special PDU known as a Complete Sequence Numbers PDU (CSNP), shall be multicast periodically by the Designated Intermediate System. The PDU shall contain a list of all LSPs in the database, together with enough information so that Intermediate systems receiving the CSNP can compare with their LSP database to determine whether they and the CSNP transmitter have synchronised LSP databases. The maximum sized Level 1 or Level 2 Sequence Numbers PDU which may be generated by a system is controlled by the values of originatingL1LSPBufferSize or originatingL2LSPBufferSize respectively. In practice, the information required to be transmitted in a single CSNP may be greater than will fit in a single PDU. Therefore each CSNP carries an inclusive range of LSPIDs to which it refers. The complete set of information shall be conveyed by transmitting a series of individual CSNPs, each referring to a subset of the complete range. The ranges of the complete set of CSNPs shall be contiguous (though not necessarily transmitted in order) and shall cover the entire range of possible LSPIDs.

The LAN Level 1 Designated Intermediate System shall periodically multicast complete sets of Level 1 CSNPs to the multi-destination address AllL1ISs. The LAN Level 2 Designated Intermediate System shall periodically multicast complete sets of Level 2 CSNPs to the multi-destination address AllL2ISs.

Absence of an LSPID from a Complete Sequence Numbers PDU whose range includes that LSPID indicates total lack of information about that LSPID.

If an Intermediate system, upon receipt of a Complete Sequence Numbers PDU, detects that the transmitter was out of date, the receiver shall multicast the missing information.

NOTE — Receipt of a link state PDU on a link is the same as successfully transmitting the Link State PDU on that link, so once the first Intermediate system responds, no others will, unless they have already transmitted replies.

If an Intermediate system detects that the transmitter had more up to date information, the receiving Intermediate system shall multicast a Partial Sequence Numbers PDU (PSNP), containing information about LSPs for which it has older information. This serves as an implicit request for the missing information. Although the PSNP is multicast, only the Designated Intermediate System of the appropriate level shall respond to the PSNP.

NOTE — This is equivalent to the PSNP being transmitted directly to the Designated Intermediate System, in that it avoids each Intermediate System unnecessarily sending the same LSP(s) in response. However, it has the advantage of preserving the property that all routeing messages can be received on the multi-destination addresses, and hence by a LAN adapter dedicated to the multi-destination address.

c) When a non-broadcast circuit (re)starts, the IS shall:

   1) set SRMflag for that circuit on all LSPs, and

   2) send a Complete set of Complete Sequence Numbers PDUs on that circuit.

7.3.18 Validation of Databases

An Intermediate System shall not continue to operate for an extended period with corrupted routeing information. The IS shall therefore operate in a *fail-stop* manner. If a failure is detected, the Intermediate system Network entity shall be disabled until the failure is corrected. In the absence of an implementation-specific method for ensuring this, the IS shall perform the following checks at least every maximumLSPGenerationInterval seconds.

a) On expiration of this timer the IS shall re-check the checksum of every LSP in the LSP database (except those with a Remaining Lifetime of zero) in order to detect corruption of the LSP while in memory. If the checksum of any LSP is incorrect, the event “Corrupted LSP Detected” shall be logged, and as a minimum the entire Link State Database shall be deleted and action taken to cause it to be re-acquired. One way to achieve this is to disable and re-enable the IS Network entity.

   NOTE — On point to point links, this requires at least that a CSNP be transmitted.

b) On completion of these checks the decision process shall be notified of an event (even if any newly generated LSPs have identical contents to the previous ones). This causes the decision process to be run and the forwarding databases re-computed, thus protecting against possible corruption of the forwarding databases in memory, which would not otherwise be detected in a stable topology.

c) The IS shall reset the timer for a period of MaximumLSPGenerationInterval with jitter applied as described in 10.1.

7.3.19 LSP Database Overload

As a result of network mis-configuration, or certain transitory conditions, it is possible that there may be insufficient memory resources available to store a received Link State PDU. When this occurs, an IS needs to take certain steps to ensure that if its LSP database becomes inconsistent with the other ISs', that these ISs do not rely on forwarding paths through the overloaded IS.

7.3.19.1 Entering the Waiting State

When an LSP cannot be stored, the LSP shall be ignored and Waiting State shall be entered. A timer shall be started for WaitingTime seconds, and the Intermediate System shall generate and flood its own LSP with zero LSP number with the "Infinite Hippity Cost" Bit set. This prevents this Intermediate system from being considered as a forwarding path by other Intermediate Systems.

It is possible that although there are sufficient resources to store an LSP and permit the operation of the Update Process on that LSP, the Decision Process may subsequently require further resources in order to complete. If these resources are not available, the Intermediate system shall then (i.e. during the attempt to run the Decision Process) enter Waiting State until such time as they are available and WaitingTime seconds have elapsed since the last LSP was ignored by the Update Process.

An implementation shall partition the available memory resources between the Level 1 and Level 2 databases. An overload condition can therefore exist independently for Level 1 or Level 2 (or both). The status attributes L1State and L2State indicate the condition for the Level 1 and Level 2 databases respectively. On entering Level 1 "Waiting State" the IS shall generate the "LSP L1 Database Overload" event, and on entering Level 2 "Waiting State" the IS shall generate the "LSP L2 Database Overload" event.

7.3.19.2 Actions in Level 1 Waiting State

While in Level 1 "waiting" state

a) If a Link State PDU cannot be stored, the IS shall ignore it and restart the timer for WaitingTime seconds.

b) The IS shall continue to run the Decision and Forwarding processes as normal.

c) When the WaitingTime timer expires, the IS shall:

   1) Generate an "LSP L1 Database Overload (recovered)" event.

   2) Clear the "Infinite Hippity Cost" bit in its own Level 1 LSP with zero LSP number and re-issue it.

   3) Set the L1State to "On".

   4) Resume normal operation.

7.3.19.3 Actions in Level 2 Waiting State

While in Level 2 "waiting" state

a) If a Link State PDU cannot be stored, the IS shall ignore it and restart the timer for WaitingTime seconds.

b) The IS shall continue to run the Decision and Forwarding processes as normal.

c) When the WaitingTime timer expires, the IS shall:

   1) Generate an "LSP L2 Database Overload (recovered)" event.
   2) Clear the "Infinite Hippity Cost" bit in its own Level 2 LSP with zero LSP number and re-issue it.
   3) Set the L2State to "On".
   4) Resume normal operation.

7.3.20 Use of the Link State Database

The only portion of the database relevant to the Decision Process is the data portion of the Link State PDUs.

The Update Process additionally uses the fields sequence number, Remaining Lifetime, and SRMFlag. The Remaining Lifetimes in the stored link state PDUs can either be periodically decremented, or converted upon receipt into an internal timestamp, and converted back into a Remaining Lifetime upon transmission.

7.3.20.1 Synchronisation with the Decision Process

Since the Update Process and the Decision Process share the Link State Database, care must be taken that the Update Process does not modify the Link State Database while the Decision Process is running.

There are two approaches to this. In one approach, the Decision Process signals when it is running. During this time, the Update Process queues incoming Link State PDUs, and does not write them into the Link State Database. If more Link State PDUs arrive than can fit into the queue allotted while the Decision Process is running, the Update Process drops them and does not acknowledge them.

Another approach is to have two copies of the Link State Database — one in which the Decision Process is computing, and the other in which the Update Process initially copies over the first database, and in which all new Link State PDUs are written. Additionally, depending on the hashing scheme, it is likely that a second copy of the address hash table will be required, so that the Update Process can do a rehash occasionally for efficiency.

When the Decision Process is ready to run again, it locks the new copy of the Link State Database, leaving the Update Process to copy over the information into the first area, and write new updates while the Decision Process runs again.

The advantage of the first approach is that it takes less memory. The advantage of the second approach is that Link State PDUs will never need to be dropped.

NOTE - If the decision process is implemented according to the specification in C.2, a finer level of parallelism is possible, as described below.

Arrival of a Link State PDU for a system before that system has been put into TENT is permitted. The new Link State PDU is used when that system is eventually put into TENT. Similarly, arrival of a new Link State PDU for a system after that system has been put into PATHS is permitted. That system has already been completely processed. The arrival of the new Link State PDU is noted and the decision process re-executed when the current execution has completed. An in-progress execution of the decision process shall not be abandoned, since this could prevent the decision process from ever completing.

Arrival of a Link State PDU for a system between that system being put on TENT and being transferred to PATHS shall be treated as equivalent to one of the previous two cases (for example, by buffering, or taking some corrective action).

7.3.20.2 Use of Buffers and Link Bandwidth

Implementations shall have a buffer management strategy that does not prevent other clients of the buffering service from acquiring buffers due to excessive use by the Update Process. They shall also ensure that the Update Process does not consume all the available bandwidth of links. In particular no type of traffic should experience starvation for longer than its acceptable latency. Acceptable latencies are approximately as follows:

- Hello traffic – Hello timer × 0.5
- Data Traffic – 10 seconds.

NOTE - The first of these requirements can be met by restricting the Update process to the use of a single buffer on each circuit for transmission. This may also cause the second requirement to be met, depending on the processor speed.

7.3.21 Parameters

MaxAge – This is the amount of time that may elapse since the estimated origination of the stored Link State PDU by the source before the LSP is considered expired. The expired LSP can be deleted from the database after a further ZeroAgeLifetime has expired. MaxAge shall be larger than maximumLSPGenerationInterval, so that a system is not purged merely because of lack of events for reporting Link State PDUs.

MaxAge is an architectural constant equal to 20 minutes.

ZeroAgeLifetime - This is the minimum amount of time for which the header of an expired LSP shall be retained after it has been flooded with zero Remaining Lifetime. A very safe value for this would be 2 × MaxAge. However all that is required is that the header be retained until the zero Remaining Lifetime LSP has been safely propagated to all the neighbours.

ZeroAgeLifetime is an architectural constant with a value of 1 minute.

maximumLSPGenerationInterval – This is the maximum amount of time allowed to elapse between generation of Link State PDUs by a source. It shall be less than MaxAge.

Setting this parameter too fast adds overhead to the algorithms (a lot of Link State PDUs). Setting this parameter too slow (and not violating constraints) causes the algorithm to wait a long time to recover in the unlikely event that incorrect Link State information exists somewhere in the domain about the system.

A reasonable setting is 15 minutes.

minimumLSPGenerationInterval – This is the minimum time interval between generation of Link State PDUs. A source Intermediate system shall wait at least this long before re-generating one of its own Link State PDUs.

Setting this too large causes a delay in reporting new information. Setting this too small allows too much overhead.

A reasonable setting is 30 seconds.

minimumLSPTransmissionInterval – This is the amount of time an Intermediate system shall wait before further propagating another Link State PDU from the same source system.

Setting this too large causes a delay in propagation of routeing information and stabilisation of the routeing algorithm. Setting this too small allows the possibility that the routeing algorithm, under low probability circumstances, will use too many resources (CPU and bandwidth).

Setting minimumLSPTransmissionInterval greater than minimumLSPGenerationInterval makes no sense, because the source would be allowed to generate LSPs more quickly than they'd be allowed to be broadcast. Setting minimumLSPTransmissionInterval smaller than minimumLSPGenerationInterval is desirable to recover from lost LSPs.

A reasonable value is 5 seconds.

CompleteSNPInterval – This is the amount of time between periodic transmissions of a complete set of Sequence Number PDUs by the Designated Intermediate system on a broadcast link. Setting this too low slows down the convergence of the routeing algorithm when Link State PDUs are lost due to the datagram environment of the Data Link layer on the broadcast link.

Setting this too high results in extra control traffic overhead.

A reasonable value is 10 seconds.

7.4 The Forwarding Process

The forwarding process is responsible for transmitting NPDUs originated by this system, and performing NPDU forwarding.

7.4.1 Input and Output

INPUT

- NPDUs from the ISO 8473 protocol machine
- PDUs from Update Process
- PDUs from Receive Process
- Forwarding Databases (Level 1 and 2) — one for each routeing metric

OUTPUT

- PDUs to Data Link Layer

7.4.2 Routeing Metric Selection

The Forwarding process selects a forwarding database for each NPDU to be relayed based on:

- the level at which the forwarding is to occur: level 1 or level 2; and
- a mapping of the ISO 8473 QoS Maintenance field onto one of the Intermediate system's supported routeing metrics.

The former selection is made by examining the Destination Address field of the NPDU.

The latter selection is made as follows:

a) If the QoS Maintenance field is not present in the NPDU, then the IS shall select the forwarding database calculated for the *default metric*.

b) If the QoS Maintenance field is present, the IS shall examine bits 7 and 8 of the parameter value octet. If these two bits specify any combination other than "1 1" (meaning globally unique QoS), then the IS shall select the forwarding database calculated for the *default metric*, otherwise

c) The IS shall select a forwarding database by mapping the values of bits 3, 2 and 1 of the parameter value as shown below in table 1 and shall proceed as follows:

   1) If the IS does not support the selected routeing metric, the IS shall forward based upon the *default metric;*

   2) If the forwarding database for one of the optional routeing metrics is selected and the database either does not contain an entry for the Destination Address in the NPDU being relayed, or contains an entry indicating that the destination is unreachable using that metric, then the IS shall attempt to forward based upon the *default metric*;

   3) Otherwise, forward based on the selected optional metric.

Table 1 - QoS Maintenance bits to routeing metric mappings

bit 1	*bit 2*	*bit 3*	*Selected routeing metric*
0	0	0	expense metric
1	0	0	default metric
0	1	1	default metric
0	1	0	expense metric
0	0	1	delay metric
1	1	0	error metric
1	0	1	delay metric
1	1	1	error metric

7.4.3 Forwarding Decision

7.4.3.1 Basic Operation

Let DEST = the Network Layer destination address of the PDU to be forwarded, or the next entry in the source routeing field, if present. It consists of sub-fields AREA ADDRESS, ID, and SELECTOR.

NOTE - The SELECTOR field in the destination address is not examined by Intermediate Systems. It is used by End Systems to select the proper Transport entity to which to deliver NSDUs.

This system's (the one examining this PDU for proper forwarding decision) address consists of sub-fields AREA ADDRESS and ID.

a) If the local system type is a level 1 Intermediate system, or the local system type is a level 2 Intermediate system and $AttachedFlag_k$ = False, then:

1) If the area address matches then

CALL Forward(Level 1,$metric_k$, DEST; ADJ),

and forward on adjacency ADJ.

2) ELSE

CALL Forward(Level 1,$metric_k$, 0; ADJ)

and forward on adjacency ADJ. (i.e. pass to nearest Level 2 Intermediate system.)

b) If the local system type is Level 2, and $AttachedFlag_k$ = True then:

1) If the area address matches then

CALL Forward(Level 1,$metric_k$, DEST; ADJ), and forward on adjacency ADJ.

2) ELSE

CALL Forward(Level 2, $metric_k$,DEST; ADJ)

and forward on adjacency ADJ. i.e. Forward according to normal Level 2 routeing

7.4.3.2 Encapsulation for Partition Repair

If this Intermediate system is the Partition Designated Level 2 IS for this partition, and the PDU is being forwarded onto the special adjacency to a Partition Designated Level 2 Intermediate system in a different partition of this area, encapsulate the complete PDU as the data field of a data NPDU (i.e., with an additional layer of header), making this system the Source address and the other Partition Designated Level 2 Intermediate system (obtained from the identifier attribute of the virtual Adjacency managed object) the Destination Address in the outer PDU header. Set the QoS Maintenance field of the outer PDU to indicate forwarding via the default routeing metric (see table 1). Then forward the encapsulated PDU onto an adjacency ADJ, obtained by calling Forward(Level 2, *default metric,* outer PDU DEST; ADJ).

7.4.3.3 The Procedure Forward

```
PROCEDURE Forward(
        level: (level1, level2),
        metric: (default,delay,expense,error);
        dest: NetworkLayerAddress,
        VAR adj: POINTER TO adjacency
        ) : BOOLEAN
```

This procedure chooses, from a Level 1 forwarding database – if level is level1, or from a Level 2 forwarding database – if level is level2, an adjacency on which to forward NPDUs for destination dest. A pointer to the adjacency is returned in adj, and the procedure returns the value "True". A destination of "0" at level 1 selects the adjacency for the nearest level 2 IS computed as described in 7.2.9.1.

If there are multiple possible adjacencies, as a result of multiple minimum cost paths, then one of those adjacencies shall be chosen. An implementation may chose the adjacency at random, or may use the possible adjacencies in "round robin" fashion.

If there is no entry in the selected forwarding database for the address dest, and the NPDU originated from the a local Transport entity and the system has one or more Intermediate System adjacencies, then one of those is chosen at random (or in "round robin" fashion) and the procedure returns the value "True". Otherwise the procedure returns the value "False".[11]

NOTE - Since the local adjacency database is pre-loaded into the decision process, there will always be an entry in the forwarding database for destinations to which an adjacency exists.

NOTE - The PDU to be forwarded may require fragmentation, depending on which circuit it is to be forwarded over. First the circuit is chosen for the PDU then, if necessary, the PDU is fragmented into multiple NPDUs which are all transmitted over that circuit.

7.4.3.4 Generating Redirect PDUs

In addition to forwarding an NPDU, the IS shall inform the local ISO 9542 protocol machine to generate a *Redirect PDU* if the PDU is being forwarded onto the same circuit from which it came, and if the source SNPA address of the NPDU indicates that the NPDU was received from an End System.

7.4.4 The Receive Process

The Receive Process is passed information from any of the following sources.

- received PDUs with the NPID of Intra-Domain routeing,

[11] This is done so that a system in the overloaded state will still be able to originate or forward NPDUs. If a system with a partial routeing information base were prohibited from attempting to forward to an unknown destination, system management would be unable to either communicate with this system, or route through it, for the purpose of diagnosing and/or correcting the underlying fault.

- configuration information from the ISO 9542 protocol machine,
- ISO 8473 data PDUs handed to the routeing function by the ISO 8473 protocol machine.

When an area is partitioned, a level 2 path is used as a level 1 link to repair the partitioned area. When this occurs, all PDUs (between the neighbours which must utilise a multi-hop path for communication) shall be encapsulated in a data NPDU, addressed to the “Intra-Domain routeing” SELECTOR. Control traffic (LSPs, Sequence Numbers PDUs) shall also be encapsulated, as well as data NPDUs that are to be passed between the “neighbours”.

NOTE - It is not necessary to transmit encapsulated IIH PDUs over a virtual link, since virtual adjacencies are established and monitored by the operation of the Decision Process and not the Subnetwork Dependent functions

The Receive Process shall perform the following functions:

- If it is a data NPDU, addressed to *this system* with SELECTOR = “Intra-Domain routeing”, then
  - decapsulate the NPDU (remove the outer NPDU header).
  - If the decapsulated PDU is a data NPDU, move the “congestion” indications to the decapsulated NPDU, and pass it to the ISO 8473 protocol machine.
  - Otherwise, if the decapsulated PDU is not an ISO 8473 PDU, perform the following steps on the decapsulated PDU:
- If it is a Link State PDU, pass it to the Update Process
- If it is a Sequence Numbers PDU, pass it to the Update Process
- If it is a IIH PDU, pass it to the appropriate Subnetwork Dependent Function
- If it is a data NPDU or Error Report for another destination, pass it to the Forwarding Process
- Otherwise, ignore the PDU

7.5 Routeing Parameters

The routeing Parameters set-able by System Management are listed for each managed object in clause 11.

7.5.1 Architectural Constants

The architectural constants are described in Table 2.

Table 2 - Routeing architectural constants

Name	*Value*	*Description*
MaxLinkMetric	63.	Maximum value of a routeing metric assignable to a circuit
MaxPathMetric	1023.	Maximum total metric value for a complete path
AllL1ISs	TBS by Digital	The multi-destination address “All Level 1 Intermediate Systems”
AllL2ISs	09-00-2B-02-00-00	The multi-destination address “All Level 2 Intermediate Systems”
AllIntermediateSystems	09-00-2B-00-00-05	The multi-destination address “All Intermediate Systems” used by ISO 9542
AllEndSystems	09-00-2B-00-00-04	The multi-destination address “All End Systems” used by ISO 9542
ISO-SAP	FE	The SAP for ISO Network Layer on ISO 8802-3 LANs
IntradomainRouteing-PD	TBS by TR9577	The Network Layer Protocol Discriminator assigned by ISO/TR 9577 for this Protocol
IntradomainRouteing-Selector	0.	The NSAP selector for the Intermediate System Network entity
SequenceModulus	2^{32}	Size of the sequence number space used by the Update Process
ReceiveLSPBuffer-Size	1492.	The size of LSP which all Intermediate systems must be capable of receiving.
MaxAge	1200.	Number of seconds before LSP considered expired.
ZeroAgeLifetime	60.	Number of seconds that an LSP with zero Remaining Lifetime shall be retained after propagating a purge.
MaximumAreaAddresses	3.	The maximum number of Area Addresses which may exist for a single area.
HoldingMultiplier	3.	The number by which to multiply helloTimer to obtain Holding Timer for ISH PDUs and for Point to Point IIH PDUs.
ISISHoldingMultiplier	10.	The number by which to multiply ISISHelloTimer to obtain Holding Timer for Level 1 and Level 2 LAN IIH PDUs.
Jitter	25.	The percentage of jitter which is applied to the generation of periodic PDUs.

8 Subnetwork Dependent Functions

The *Subnetwork Dependent Functions* mask the characteristics of the different kinds of Subnetworks from the *Subnetwork Independent Routeing Functions*. The only two types of circuits the Subnetwork Independent Functions recognise are *broadcast* and *general topology*.

The Subnetwork Dependent Functions include:

- The use of the ISO 8473 *Subnetwork Dependent Convergence Functions (SNDCF)* so that this protocol may transmit and receive PDUs over the same subnetwork types, using the same techniques, as does ISO 8473.

- Co-ordination with the operation of the ES–IS protocol (ISO 9542) in order to determine the Network layer addresses (and on Broadcast subnetworks, the subnetwork points of attachment) and identities (End System or Intermediate System) of all adjacent neighbours. This information is held in the *Adjacency* data base. It is used to construct Link State PDUs.

- The exchange of IIH PDUs. While it is possible for an Intermediate System to identify that it has an Intermediate System neighbour by the receipt of an ISO 9542 ISH PDU, there is no provision within ISO 9542 to indicate whether the neighbour is a Level 1 or a Level 2 Intermediate System. Specific PDUs (*LAN Level 1, LAN Level 2* and *Point to point IIH PDUs*) are defined to convey this information.

8.1 Multi-destination Circuits on ISs at a Domain Boundary

Routeing information (Link State PDUs) is not exchanged across a routeing domain boundary. All routeing information relating to a circuit connected to another routeing domain is therefore entered via the Reachable Address managed objects. This information is disseminated to the rest of the routeing domain via Link State PDUs as described in 7.3.3.2. This has the effect of causing NPDUs destined for NSAPs which are included in the addressPrefixes of the Reachable Addresses to be relayed to that Intermediate System at the domain boundary. On receipt of such an NPDU the Intermediate system shall forward it onto the appropriate circuit, based on its own Link State information. However in the case of multi-destination subnetworks (such as an ISO 8208 subnetwork using Dynamic Assignment, a broadcast subnetwork, or a connectionless subnetwork) it is necessary to ascertain additional subnetwork dependent addressing information in order to forward the NPDU to a suitable SNPA. (This may be the target End system or an Intermediate system within the other domain.)

In general the SNPA address to which an NPDU is to be forwarded can be derived from the destination NSAP of the NPDU. It may be possible to perform some algorithmic manipulation of the NSAP address in order to derive the SNPA address. However there may be some NSAPs where this is not possible. In these cases it is necessary to have pre-configured information relating an address prefix to a particular SNPA address.

This is achieved by additional information contained in the Reachable Address subordinate managed object. The mappingType characteristic may be specified as “Manual”, in which case a particular SNPA address or set of SNPA addresses is specified in the SNPA Address characteristic. Alternatively the name of an SNPA address extraction algorithm may be specified.

8.2 Point to Point Subnetworks

This clause describes the identification of neighbours on point to point links and on both Static and DCM DED circuits.

The IS shall operate the ISO 9542 protocol, shall be able to receive ISO 9542 ISH PDUs from other ISs, and shall store the information so obtained in the adjacency database.

8.2.1 Receipt of ESH PDUs – Database of End Systems

An IS shall enter an End system into the adjacency database when an ESH PDU is received on a circuit. If an ESH PDU is received on the same circuit, but with a different NSAP address, the new address shall be added to the adjacency, with a separate timer. A single ESH PDU may contain more than one NSAP address. When a new data link address or NSAP address is added to the adjacency database, the IS shall generate an "Adjacency State Change (Up)" event on that adjacency.

The IS shall set a timer for the value of "Holding Time" in the received ESH PDU. If another ESH PDU is not received from the ES before that timer expires, the ES shall be purged from the database, provided that the Subnetwork Independent Functions associated with initialising the adjacency have been completed. Otherwise the IS shall clear the adjacency as soon as those functions are completed.

When the adjacency is cleared, the Subnetwork Independent Functions shall be informed of an "Adjacency Change (Down)" event, and the adjacency can be re-used after the Subnetwork Independent Functions associated with bringing down the adjacency have been completed.

8.2.2 Receiving ISH PDUs by an Intermediate System

On receipt of an ISH PDU by an Intermediate System, the IS shall create an adjacency (with state "Initialising" and adjacencyType "Unknown"), if one does not already exist, and then perform the following actions:.

a) If the Adjacency state is "Up" and the ID portion of the NET field in the ISH PDU does not match the neighbourID of the adjacency then the IS shall:

   1) generate an "Adjacency State Change (Down)" event;

   2) delete the adjacency; and

   3) create a new adjacency with:

      (i) state set to "Initialising", and

      (ii) adjacencyType set to "Unknown".

   4) perform the following actions..

b) If the Adjacency state is "Initialising", and the adjacencyType status is "Intermediate System", the ISH PDU shall be ignored.

c) If the Adjacency state is "Initialising" and the adjacencyType status is not "Intermediate System", a point to point IIH PDU shall be transmitted as described in Clause 8.2.3.

d) The adjacencyType status shall be set to "Intermediate System" indicating that the neighbour is an Intermediate system, but the type (L1 or L2) is, as yet, unknown.

8.2.3 Sending Point to Point IIH PDUs

When required, an IS shall construct and transmit a Point to Point IIH PDU as follows.

a) The Circuit Type field shall be set according to Table 3.

b) The Local Circuit ID field shall be set to a value assigned by this Intermediate system when the circuit is created. This value shall be unique among all the circuits of this Intermediate system.

c) The first Point to Point IIH PDU (i.e. that transmitted as a result of receiving an ISH PDU, rather than as a result of timer expiration) shall be padded (with trailing PAD options containing arbitrary valued octets) so that the SNSDU containing the IIH PDU has a length of at least *maxsize - 1* octets[12] where *maxsize* is the maximum of

   - dataLinkBlocksize

[12]The minimum length of PAD which may be added is 2 octets, since that is the size of the option header. Where possible the PDU should be padded to maxsize, but if the PDU length is *maxsize- 1* octets no padding is possible (or required).

- originatingL1LSPBufferSize
- originatingL2LSPBufferSize

This is done to ensure that an adjacency will only be formed between systems which are capable of exchanging PDUs of length up to *maxsize* octets. In the absence of this check, it would be possible for an adjacency to exist with a lower maximum block size, with the result that some LSPs and SNPs (i.e. those longer than this maximum, but less than *maxsize*) would not be exchanged.

NOTE - It is necessary for the manager to ensure that the value of dataLinkBlocksize on a circuit which will be used to form an Intermediate system to Intermediate system adjacency is set to a value greater than or equal to the maximum of the LSPBufferSize characteristics listed above. If this is not done, the adjacency will fail to initialise. It is not possible to enforce this requirement, since it is not known until initialisation time whether or not the neighbour on the circuit will be an End system or an Intermediate system. An End system adjacency may operate with a lower value for dataLinkBlocksize.

Table 3 - Setting the value of the Circuit Type field

Routeing Type	Circuit manualL2OnlyMode	Circuit Type Field
Level 1	—	Level 1 only (1)
Level 2	“True”	Level 2 only (2)
Level 2	“False”	Level 1 and 2 (3)

8.2.4 Receiving Point to Point IIH PDUs

When a Point to Point IIH PDU is received by an Intermediate system, the area addresses of the two Intermediate Systems shall be compared to ascertain the validity of the adjacency. If the two Intermediate systems have an area address in common, the adjacency is valid for all combinations of Intermediate system types (except where a Level 1 Intermediate system is connected to a Level 2 Intermediate system with manualL2OnlyMode set “True”). However, if they have no area address in common, the adjacency is only valid if both Intermediate systems are Level 2, and the IS shall mark the adjacency as Level 2 Only. This is described in more detail below.

On receipt of a Point to Point IIH PDU, each of the Area Addresses from the PDU shall be compared with the set of Area Addresses in the manualAreaAddresses characteristic.

a) If a match is detected between any pair the following actions are taken.

   1) If the local system is of Routeing Type “L1IntermediateSystem” the IS shall perform the action indicated by Table 4.

   2) If the local system is of Routeing Type “L2IntermediateSystem” and the Circuit manualL2OnlyMode has the value “False”, the IS shall perform the action indicated by Table 5.

   3) If the local system is of Routeing Type “L2IntermediateSystem” and the Circuit manualL2OnlyMode has the value “True”, the IS shall perform the action indicated by Table 6.

b) If a no match is detected between any pair, the following actions shall be performed.

   1) If the local system is of Routeing Type “L1IntermediateSystem” and the adjacency is not in state “Up”, the IS shall delete the adjacency (if any) and generate an “Initialisation Failure (Area Mismatch)” Event.

   2) If the local system is of Routeing Type “L1IntermediateSystem” and the adjacency is in state “Up”, the IS shall delete the adjacency and generate an “Adjacency State Change (Down – Area Mismatch)” Event .

   3) If the local system is of Routeing Type “L2IntermediateSystem” the IS shall perform the action indicated by Table 7 (irrespective of the value of manualL2OnlyMode for this circuit).

c) If the action taken is “Up”, as detailed in the tables referenced above, the IS shall compare the Source ID field of the PDU with the local SystemID.

1) If the local Intermediate system has the higher Source ID, the IS shall set the Circuit CircuitID status to the concatenation of the local SystemID and the Local Circuit ID (as sent in the Local Circuit ID field of point to point IIH PDUs from this Intermediate System) of this circuit.

2) If the remote Intermediate system has the higher Source ID, the IS shall set the Circuit CircuitID status to the concatenation of the remote system's Source ID (from the Source ID field of the PDU), and the remote system's Local Circuit ID (from the Local Circuit ID field of the PDU).

3) If the two source IDs are the same (i.e. the system is initialising to itself), the local SystemID is used.

NOTE — The circuitID status is not used to generate the Local Circuit ID to be sent in the Local Circuit ID field of IIH PDUs transmitted by this Intermediate system. The Local Circuit ID value is assigned once, when the circuit is created and is not subsequently changed.

d) If the action taken is "Accept" and the new value computed for the circuitID is different from that in the existing adjacency, the IS shall

1) generate an "Adjacency State Change(Down)" event, and

2) delete the adjacency.

e) If the action taken is "Up" or "Accept" the IS shall

1) copy the Adjacency neighbourAreas entries from the PDU,

2) set the holdingTimer to the value of the "Holding Time" from the PDU, and

3) set the neighbourSystemID to the value of the "Source ID" from the PDU.

Table 4 - Level 1 State table for matching areas

Circuit Type[1]	Adjacency Type		
	none[2]	Level 1[3]	Level 2[4]
Level 1 only	Up[5] L1[6]	Accept	Down[7] (Wrong system)
Level 2 only	Reject[8] (Wrong system)	Down[7] (Wrong system)	Down[7] (Wrong system)
Level 1 & 2	Up[5] L2[9]	Down[7] (Wrong system)	Accept

[1]The value of the "Circuit Type" field in the received PDU.

[2]The adjacency is not in state "Up"

[3]The adjacency is in state "Up" and the Adjacency adjacencyType is "L1 Intermediate System".

[4]The adjacency is in state "Up" and the Adjacency adjacencyType is "L2 Intermediate System".

[5]The adjacency is accepted and an "Adjacency State Change (Up)" event is generated. If the Adjacency adjacencyType was "Unknown" (i.e. no ISH PDU has yet been received), a point to point IIH PDU is also transmitted.

[6]The adjacencyType status is set to "L1 Intermediate System".

[7]An "Adjacency State Change (Down)" event is generated, with the specified reason, and the adjacency deleted.

[8]An "Initialization Failure" event is generated with the specified reason.

[9]The adjacencyType status is set to "L2 Intermediate System".

Table 5 - Level 2 State table for matching areas

Circuit Type[1]	Adjacency Type			
	none[2]	Level 1[3]	Level 2[4]	Level 2 Only[5]
Level 1 only	Up[6] L1[7]	Accept	Down[8] (Wrong system)	Down[8] (Wrong system)
Level 2 only	Up[6] L2O[9]	Down[8] (Wrong system)	Down[8] (Wrong system)	Accept
Level 1 & 2	Up[6] L2[10]	Down[8] (Wrong system)	Accept	Down[8] (Wrong system)

[1]The value of the “Circuit Type” field in the received PDU.

[2]The adjacency is not in state “Up”

[3]The adjacency is in state “Up” and the Adjacency adjacencyType is “L1 Intermediate System”.

[4]The adjacency is in state “Up” and the Adjacency adjacencyType is “L2 Intermediate System, with L2OnlyMode “False”.

[5]The adjacency is in state “Up” and the Adjacency adjacencyType is “L2 Intermediate System, with L2OnlyMode “True”.

[6]The adjacency is accepted and an “Adjacency State Change (Up)” event is generated. If the Adjacency adjacencyType was “Unknown” (i.e. no ISH PDU has yet been received), a point to point IIH PDU is also transmitted.

[7]The adjacencyType status is set to “L1 Intermediate System”.

[8]An “Adjacency State Change (Down)” event is generated, with the specified reason, and the adjacency deleted.

[9].The adjacencyType status is set to “L2 Intermediate System”, and the L2OnlyMode is set to “True”.

[10]The adjacencyType status is set to “L2 Intermediate System”, and the L2OnlyMode is set to “False”.

Table 6 - Level 2 Only State table for matching areas

Circuit Type[1]	Adjacency Type		
	none[2]	Level 2[3]	Level 2 Only[4]
Level 1 only	Reject[5] (Wrong system)	Down[6] (Wrong system)	Down[6] (Wrong system)
Level 2 only	Up[7] L2O[8]	Down[6] (Wrong system)	Accept
Level 1 & 2	Up[7] L2O[8]	Down[6] (Wrong system)	Accept

[1]The value of the "Circuit Type" field in the received PDU.

[2]The adjacency is not in state "Up"

[3]The adjacency is in state "Up" and the Adjacency adjacencyType is "L2 Intermediate System", with L2OnlyMode "False".

[4]The adjacency is in state "Up" or and the Adjacency adjacencyType is "L2 Intermediate System" with L2OnlyMode "True".

[5]An "Initialization Failure" event is generated with the specified reason.

[6]An "Adjacency State Change (Down)" event is generated, with the specified reason, and the adjacency deleted.

[7]The adjacency is accepted and an "Adjacency State Change (Up)" event is generated. If the Adjacency adjacencyType was "Unknown" (i.e. no ISH PDU has yet been received), a point to point IIH PDU is also transmitted.

[8]The adjacencyType status is set to "L2 Intermediate System" with L2OnlyMode "True".