ECMA

EUROPEAN COMPUTER MANUFACTURERS ASSOCIATION

END SYSTEM ROUTING

TR/38

January 1987

Free copies of this document are available from ECMA,
European Computer Manufacturers Association
114 Rue du Rhône – 1204 Geneva (Switzerland)

ECMA

EUROPEAN COMPUTER MANUFACTURERS ASSOCIATION

END SYSTEM ROUTING

TR/38

January 1987

BRIEF HISTORY

This Technical Report is concerned with the standardization of a routing protocol to be used between End Systems and suppliers of routing information. The protocol to be used by suppliers of routing information between themselves is expected to be the subject of a future ECMA publication.

The Technical Report has the following objectives:

(i) To state the ECMA position with regard to End System Routing,

(ii) To serve as a vehicle for influencing decisions in other standards arenas,

(iii) To formalise the work carried out by ECMA.

It is not the intention to define completely new approaches or solutions in this area, but rather to complement, improve, and ensure compatibility with existing work of standardization bodies such as ISO and ANSI.

The basis for this Technical Report is the ongoing activity in ISO and ANSI.

Adopted by the General Assembly of December 11, 1986.

TABLE OF CONTENTS

1. SCOPE AND FIELD OF APPLICATION

To facilitate compatible interconnection of data processing equipment conforming to the Open Systems Interconnection architecture, this Technical Report describes:

(i) Principles by which End Systems which do not have inbuilt knowledge and ability to identify for themselves the route which should be used for the transmission of data in the Network Layer are able to obtain the necessary information from other systems.

(ii) A protocol for use by such End Systems in obtaining routing information.

The following are outside the scope of this Technical Report:

(i) The methods by which the suppliers of routing information determine the routes which they notify by the means described in this Report.

(ii) The operation and contents of Network Layer directories.

(iii) Any form of determination of routes other than at the Network Layer (for example, in the Application Layer).

This Technical Report is concerned with routing information required for the provision of the Network Service in the following circumstances:

(i) When providing Connectionless-mode Network Service (CLNS) using the methods described in ISO 8473.

In this case the procedures described in this Report constitute further activities of the Network Entity which are additional to the procedures described in ISO 8473.

(ii) When providing Connection-mode Network Service (CONS) over X.25.

In this case the procedures described in this Report constitute further activities of the Network Entity in addition to those described in ISO 8878 and ISO 8881.

When provision of the Network Service in circumstances other than these is standardized further associated routing activities may be necessary.

2. CONFORMANCE

Systems claiming Conformance with the Protocol defined in this Technical Report shall:

(i) Implement the operations defined in 9.5 and 9.6 and state which if any of the further operations defined in 9.1, 9.2, 9.3 and 9.4 are implemented, and

(ii) Construct Protocol Data Units (PDUs) according to Clause 10, and

(iii) Discard PDUs relating to operations which are not implemented.

3. REFERENCES

ECMA-117	Domain-Specific Part of Network Layer Addresses.
ISO 7498	Information Processing Systems - Open Systems Interconnection - Basic Reference Model.
ISO 7498/DAD1	Information Processing Systems - Open Systems Interconnection - Addendum to ISO 7498 Covering Connectionless-mode Transmission.
ISO 8208	Information Processing Systems - X.25 Packet Level Protocol for Data Terminal Equipment.
ISO 8348	Information Processing Systems - Telecommunications and Information Exchange between Systems - Network Service Definition.
ISO 8348/DAD1	Information Processing Systems - Telecommunications and Information Exchange between Systems - Addendum to the Network Service Definition Covering Connectionless-mode Transmission.
ISO 8348/DAD2	Information Processing Systems - Telecommunications and Information Exchange between Systems - Addendum to the Network Service Definition Covering Network Layer Addressing.
ISO 8473	Information Processing Systems - Telecommunications and Information Exchange between Systems - Protocol for Providing the Connectionless Network Service.
ISO 8648	Information Processing Systems - Telecommunications and Information Exchange between Systems - Internal Organization of the Network Layer.
ISO 8802	Information Processing Systems - Local Area Networks.
ISO 8878	Information Processing Systems - Data Communications - Use of X.25 to Provide the OSI Connection-mode Network Service.
ISO 8881	Information Processing Systems - Data Communications - Use of the X.25 Packet Level Protocol in Local Area Networks.

4. DEFINITIONS

4.1 Reference Model Definitions

This Technical Report makes use of the following concepts from ISO 7498, Basic Reference Model:

- End System (ES)
- Network Entity
- Network Layer
- Network Protocol
- Network Protocol Data Unit (NPDU)

- Network Relay
- Network Service Access Point (NSAP)
- Network Service Access Point Address
- Routing

4.2 Network Layer Architecture Definitions

This Technical Report makes use of the following concepts from ISO 8648, Internal Organization of the Network Layer:

- Subnetwork (SN)
- Subnetwork Access Protocol (SNAcP)
- Subnetwork Service

4.3 Network Layer Addressing Definitions

This Technical Report makes use of the following concepts from ISO 8348/DAD2, Addendum to the Network Service Definition Covering Network Layer Addressing:

- Subnetwork Address
- Subnetwork Point of Attachment (SNPA)

and of the following concepts from ISO 8802, Local Area Networks:

- Broadcast Medium
- Multicast Address

4.4 Additional Definitions

For the purposes of this Technical Report the following additional definitions apply:

4.4.1 Configuration

The collection of systems attached to a single subnetwork. This is defined in terms of system types (i.e. End or Intermediate Systems), Network Service Access Point (NSAP) addresses present, Network Entities present, and the correspondence between systems and Subnetwork Point of Attachment (SNPA) addresses.

4.4.2 Intermediate System (IS)

A system which does not contain application processes, and is used only to enable interconnection of subnetworks.

4.4.3 Management Information Base

The conceptual repository for Management Information.

4.4.4 Network Entity Title (NET)

An identifier for a Network Entity which has the same abstract syntax as a Network Service Access Point (NSAP) address, and which can be used to unambiguously identify a Network Entity in an End or Intermediate System.

4.4.5 Routing Information Base

That part of the Management Information Base which is concerned with routing.

4.4.6 Routing Information Supplier (RIS)

A system which is responsible for determining routing information and supplying it to other systems.

5. GENERAL OVERVIEW

The Network Layer is responsible for transmission of data from one End System to another, via whatever routes through intervening subnetworks are appropriate. Therefore, when an End System wishes to transmit data, the Network Layer is responsible for determining the route which should be followed and the Intermediate Systems which should be used. In the case of Connection-mode Network Service (CONS), this determination is made when a connection is established, and all subsequent data for the connection follows the same route. In the case of Connectionless-mode Network Service (CLNS), the route of each Network Protocol Data Unit (NPDU) may be determined separately. This Technical Report is concerned with the actions to be taken by End Systems in order that the correct routes may be determined.

The process of determining what routes are to be used, taking into account changing conditions throughout all the interconnected subnetworks, may be a very complex task involving the exchange and processing of a large amount of information. There is therefore a special class of systems, Routing Information Suppliers (RIS), which are responsible for cooperatively determining routing information and supplying it to systems as required.

Intermediate Systems are responsible for passing data from one subnetwork to another. They therefore require a different level of routing information. Moreover they are necessarily involved in the overall process of determining routes, not only because the information about the status of network components must at the very least pass through them, but also because the status of Intermediate Systems is an important part of this information. The routing process can be optimized if the Routing Information Suppliers are the Intermediate Systems. The definitions in this Report assume that this is the case. The function of supplying routing information may be carried out by other systems either as well as, or instead of, the Intermediate Systems, but this is outside the scope of this Technical Report.

There are three basic requirements to enable this routing process to work:

(i) End Systems must be able to determine how to access operational suppliers of routing information.

(ii) Protocols must exist to enable End Systems to request specific information required, and to enable Routing Information Suppliers to transmit it to them.

(iii) Routing Information Suppliers, in order to carry out their work of identifying operational routes, must be able to determine which End Systems are operational.

Note 1:

Routing Information Suppliers also have a requirement for other information about the network, but the methods by which this information is obtained are outside the scope of this Technical Report.

The three basic requirements may be satisfied to a greater or lesser extent by the underlying subnetwork. The protocol defined in Clause 9 contains functions to enable all the requirements to be satisfied, but network service providers are required to use the functions only to the extent necessary to support the Quality of Service (QoS) they wish to provide. For example:

(a) A Connection-oriented subnetwork may give a very reliable indication of the status of remote systems whenever an attempt is made to transmit to them. In this case a network service provider in an End System may choose not to implement the Report Configuration function, if the QoS which it wants to provide can be achieved on routes which the information suppliers will be able to determine on the basis of information provided by the subnetwork itself. Similarly, information suppliers might not implement the Report Configuration function in such cases.

(b) The addressing system used on a subnetwork may make it possible to determine whether a destination is on the same subnetwork, and what its subnetwork address is. An End System might then choose not to use the functions described in Clause 9 for communication with systems on the same subnetwork, if the QoS supported by the subnetwork itself is adequate.

(c) The subnetwork may carry out redirection itself when necessary, making use of the Redirect function unnecessary.

6. PRINCIPLES

In order to satisfy the three basic requirements described in Clause 5, three basic functions have to be carried out by an End System. These are:

(i) Routing Information Supplier Identification:

Determining the Subnetwork Point of Attachment (SNPA) address of suppliers of routing information, and whether those information suppliers are operational.

(ii) Route Identification:

Determining the SNPA addresses of systems which could be used as the first, or only, step of routes to other NSAPs, and discovering whether those systems are operational.

(iii) Self Identification:

Ensuring that its own SNPA address, and the fact that it is operational, are known to other systems.

The methods used for each of these functions are essentially similar; they are:

(a) To enable the determination of an SNPA address, either:

- it is pre-configured, or
- special protocol messages are used to indicate where systems are located.

(b) To enable determination of the operability of systems, either:

- information is obtained from the normal operation of the subnetwork during the provision of the network service (for example, if attempts are made to establish an X.25 virtual circuit), or
- where the nature of the subnetwork is such that adequate information is not obtained in this way, use is made of an additional protocol designed to indicate systems' operability.

7. PROTOCOL OVERVIEW

7.1 Information Provided by the Protocol

This Protocol provides two types of information to Network Entities which support its operation:

- Configuration Information, and
- Route Redirection Information.

7.1.1 Configuration Information

Configuration Information permits End Systems (ESs) to discover the existence and reachability of Intermediate Systems (ISs) and permits ISs to discover the existence and reachability of ESs. This information allows ESs and ISs attached to the same subnetwork to dynamically discover each other's existence and availability, thus eliminating the need for manual intervention at ESs and ISs to establish the identity of Network Entities that can be used to route Network Protocol Data Units (NPDUs).

Configuration Information also permits End Systems to obtain information about each other in the absence of an available Intermediate System.

Note 2:

The term Configuration Information is not intended in the broad sense of configuration as used in the context of OSI system management. Rather, only functions specifically defined herein are intended.

7.1.2 Route Redirection Information

Route Redirection Information allows Intermediate Systems to inform End Systems of (potentially) better paths to use when forwarding NPDUs to a particular destination. A better path could either be another IS on the same subnetwork as the ES, or the destination ES itself, if it is on the same subnetwork as the source ES. Allowing the ISs to inform the ESs of routes minimizes the complexity of routing decisions in End Systems, and improves performance because the ESs may make use of the better IS or local subnetwork access for subsequent transmissions.

7.2 Types of Subnetwork

In order to evaluate the applicability of this protocol in particular configurations of End Systems, Intermediate Systems and subnetworks, three generic types of subnetwork are identified.

These are:

- the Point-to-Point subnetwork,
- the Broadcast subnetwork, and
- the General Topology subnetwork.

These subnetwork types are discussed below.

7.2.1 Point-to-Point Subnetworks

A Point-to-Point subnetwork supports exactly two systems. The two systems may be either two End Systems, or an End System and a single Intermediate System. A single point-to-point data link connecting two Network Entities is an example of a Point-to-Point subnetwork.

7.2.1.1 Configuration Information

On a Point-to-Point subnetwork the Configuration Information of this protocol informs the communicating Network Entities as to whether:

- the topology consists only of two End Systems, or
- one of the two systems is an Intermediate System.

Note 3:

On a Point-to-Point subnetwork, if both systems are Intermediate Systems then this protocol is inapplicable to the situation, since an IS-to-IS protocol should be employed instead. However, there is no reason why the Configuration Information could not be employed in an IS-to-IS environment to ascertain the topology and initiate operation of an IS-to-IS protocol.

The Intermediate System is informed of the NSAP address(es) supported by the Network Entity in the End System. This permits reachability information and routing metrics concerning these NSAPs to be disseminated to other Intermediate Systems for the purpose of calculating routes to and from this End System.

7.2.1.2 Route Redirection Information

Route Redirection Information is not employed on Point-to-Point subnetworks because there are never any alternate routes.

7.2.2 Broadcast Subnetworks

A Broadcast subnetwork supports an arbitrary number of End Systems and Intermediate Systems, and additionally is capable of transmitting a single Subnetwork Service Data Unit (SNSDU) to all or a subset of these systems in response to a single SN_UNITDATA Request. An example of a Broadcast subnetwork is a Local Area Network (LAN) conforming to ISO 8802/2, type 1 operation.

Note 4:

While not strictly necessary, it is also presumed that the cost of this operation is close to the cost of transmitting a single SNSDU rather than the cost of transmitting n SNSDUs, where n is the number of systems attached to the subnetwork.

7.2.2.1 Configuration Information

On a Broadcast subnetwork the Configuration Information of this protocol is employed to inform the communicating Network Entities of the following:

- End Systems are informed of the reachability, Network Entity Title (NET), and SNPA address(es) of each active Intermediate System on the subnetwork.
- Intermediate Systems are informed of the subnetwork address and NSAP address(es) of each End System. Once the Intermediate System obtains this information, reachability information and routing metrics concerning these NSAPs may be disseminated to other ISs for the purpose of calculating routes to and from each ES on the subnetwork.
- In the absence of an available Intermediate System, End Systems may query over a Broadcast subnetwork to discover whether a particular NSAP is reachable on the subnetwork, and if so, what SNPA address to use to reach that NSAP.

7.2.2.2 Route Redirection Information

Route Redirection Information may be employed on Broadcast subnetworks to permit Intermediate Systems to inform End Systems of superior routes to a destination NSAP. The superior route might be another IS on the same subnetwork as the ES, or it might be the destination ES itself, if it is directly reachable on the same subnetwork as the source ES.

7.2.3 General Topology Subnetworks

A General Topology subnetwork 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. An example of a General Topology subnetwork is a subnetwork employing ISO 8208.

Note 5:

The crucial distinguishing characteristic between the Broadcast subnetwork and the General Topology subnetwork is the «cost» of an n-way transmission to a potentially large subset of the systems on the subnetwork. On a General Topology subnetwork, the cost is assumed to be close to the cost of sending an individual PDU to each SNPA on the subnetwork. Conversely, on a Broadcast subnetwork the cost is assumed to be close to the cost of sending a single PDU to one SNPA on the subnetwork. Intermediate situations between these extremes are of course possible; in such cases it would be possible to treat the subnetwork as either in the Broadcast or General Topology category.

7.2.3.1 Configuration Information

On a General Topology subnetwork the Configuration Information is generally not employed because this protocol can be very costly in the utilization of, and charging for, subnetwork resources.

7.2.3.2 Route Redirection Information

Route Redirection Information may be employed on General Topology subnetworks to permit Intermediate Systems to inform End Systems of superior routes to a destination NSAP. The superior route might be another IS on the same subnetwork as the ES, or it might be the destination ES itself if it is directly reachable on the same subnetwork as the source ES.

7.3 Services Assumed by the Protocol

The type of facilities required to support the protocol depends on the mode of Network Service which is being provided.

7.3.1 Facilities Required for Connectionless-mode Network Service (CLNS)

The subnetwork service required to support this protocol is defined as comprising the following primitives:

Primitives		Parameters
SN_UNITDATA	Request Indication	SN_Destination_Address, SN_Source_Address, SN_Quality_of_Service, SN_Userdata

Figure 1: Subnetwork Service Primitives

7.3.1.1 Subnetwork Addresses

The source and destination addresses specify the points of attachment to a public or private subnetwork(s) involved in the transmission (Subnetwork Points of Attachment (SNPAs)). Subnetwork addresses are defined in the Service Definition of each individual subnetwork.

This protocol is designed to take advantage of subnetworks which support Broadcast, Multicast, or other forms of multi-destination addressing for n-way transmission. It is assumed that the SN_Destination_Address parameter may take on one of the following multi-destination addresses in addition to a normal single-destination address:

- «All End System Network Entities»,
- «All Intermediate System Network Entities».

Where a real subnetwork does not inherently support Broadcast or other forms of transmission to multi-destination addresses, a convergence function may be used to provide n-way transmission to these multi-destination addresses.

When the SN_Destination_Address on the SN_UNITDATA Request is a multi-destination address, the SN_Destination_Address parameter in the corresponding SN_UNITDATA Indication shall be the same multi-destination address.

The syntax and semantics of subnetwork addresses, except for the properties described above, are not defined in this Technical Report.

7.3.1.2 Subnetwork User Data

The SN_Userdata is an ordered multiple of octets, and is transferred transparently between the specified subnetwork service access points.

The subnetwork service is required to support an SNSDU size of at least that required to operate the Protocol for Providing the Connectionless Network Service (ISO 8473).

7.3.1.3 Service Assumed from Local Environment

A timer service must be provided to allow the protocol entity to schedule events. This is identical to the S-TIMER service defined in ISO 8473.

7.3.2 Facilities Required for Connection-mode Network Service (CONS)

Where the CONS is being provided, the routing protocol requires the facilities provided by the X.25 Packet Layer as described in ISO 8208. Where the facilities required for CLNS as described above are also available, further protocol elements may be implemented which make use of them.

7.4 Summary of Protocol Elements

The protocol is based on a number of PDUs which contain information relating to routing, together with a time limit restricting the validity of the information. The use of these PDUs is defined in Clauses 8 and 9, and their format is described in Clause 10. The PDUs are as follows:

Name	Information Contained
ESH	An NSAP address located at the SNPA from which the PDU was sent
ISH	The title of an Intermediate System entity located at the SNPA from which the PDU was sent
RD	An SNPA which constitutes a better route towards a specified NSAP

Figure 2: Summary of PDUs

Time limits on the validity of the information are specified by means of a Holding Time contained in the PDU.

The Holding Timer (HT) applies to both Configuration Information and Route Redirection Information. The value of the HT is set by the source of the information and transmitted in the appropriate PDU. The recipient of the information is expected to retain the information no longer than the HT. Old Configu-

ration or Route Redirection Information must be discarded after the HT expires to ensure the correct operation of the protocol.

Further discussion of the rationale for these timers, and guidelines for their use, is given in Appendix A.

In addition, the PDUs contain a checksum to allow validity to be confirmed. The operation of the checksum is as defined in clause 6.11 of ISO 8473.

The End System Hello (ESH) and Intermediate System Hello (ISH) PDUs are used only when the facilities described in 7.3.1 are available, and are transported in SN_UNITDATA primitives. The Redirect (RD) PDU is transmitted in the User Data field of an X.25 CLEAR packet when the facilities described in 7.3.2 are used, and otherwise in an SN_UNITDATA primitive.

8. ROUTING FUNCTIONS

8.1 Identifying Operational Intermediate Systems

8.1.1 Location of Intermediate Systems

The existence of an Intermediate System at a particular Subnetwork Point of Attachment (SNPA) is discovered either:

(a) by pre-configured information, or

(b) by receipt of an Intermediate System Hello (ISH) PDU.

Method (b) is only likely to be practical on Broadcast subnetworks because of the cost on General Topology subnetworks of ensuring that all End Systems receive ISH PDUs.

8.1.2 Operational Status of Intermediate Systems

The operability and reachability of an Intermediate System may be notified by the subnetwork when an attempt is made to communicate with it, for example in the case of success or failure of an attempt to establish an X.25 virtual circuit. Where such information is not available with sufficient reliability to enable provision of the required Quality of Network Service, the following procedure is used:

A Flush Old Configuration function is executed periodically to remove stale Configuration Information from the local Network Entity's Routing Information Base. All entries obtained from ESH and ISH PDUs previously received are scanned, and if the Holding Timer has elapsed, the information corresponding to that ES or IS is discarded.

8.2 Self Identification

The location of End Systems may be known to other systems by pre-configured information, or by algorithms based on address structures, for example those defined in Standard ECMA-117. Where it is not desired to rely on such methods, or where provision of the desired QoS requires that additional information about the reliability of the system must be made available to other systems, then the following procedure is followed:

A Configuration Timer (CT) is defined, which is a local timer (i.e. maintained independently by each system). The timer determines how often a system reports its availability to the other systems on the same subnetwork. The shorter the CT, the more quickly other systems on the subnetwork will become aware when the reporting system becomes available or unavailable. The increased responsiveness must be traded off against increased use of resources in the subnetwork and in the recipient systems.

Every time the local CT expires in an End System, it constructs and transmits one ESH PDU for each NSAP it serves, as described in 9.1. An ESH PDU is also transmitted as part of the operation of the Query Configuration procedure as described in 9.1.1.

8.3 Route Identification

An End System may be aware of routes to distant NSAPs by any of the following methods:

(a) Some routes may be pre-configured into the system, or obtained by pre-defined algorithms based on address structure,

(b) Routes may be identified by receipt of an RD PDU,

(c) Routes may be deduced from the fact that an ESH PDU has been received indicating an SNPA on a directly connected subnetwork where the NSAP is located. This only results in a useful route if the QoS available on the subnetwork is adequate.

A route which has been identified is no longer usable when:

(a) Attempting to use it results in a report from the subnetwork service that the required SNPA is no longer reachable, or

(b) Where the route was based on ESH or RD information, the Holding Timer has expired, or

(c) An RD PDU is received indicating a better route to the same destination with adequate QoS.

If no usable route is known to a required destination, the End System proceeds as follows:

(i) If an operational Intermediate System (IS) has been identified, the End System assumes that the destination can be reached by using that IS as the destination of the first hop. If more than one operational IS is available, the choice between them is a local matter not covered by this Technical Report. Details of subsequent routing activity which can arise from this procedure are given in 9.4.

(ii) If there is no operational IS, but the subnetwork to which the End System is attached is a Broadcast subnetwork, then the Configuration Request procedure is followed (see 9.3).

(iii) If there is no operational IS on a non-Broadcast subnetwork, then the destination NSAP is unreachable in this situation.

9. PROTOCOL PROCEDURE DETAILS

9.1 Operation of End System Hello (ESH)

Every time the local Configuration Timer expires in an End System, it constructs and transmits one ESH PDU for each NSAP it serves, and issues one SN_UNITDATA Request with the ESH PDU as the SNSDU on each subnetwork to which it is attached.

The Holding Time field is set to approximately twice the value of the ES's Configuration Timer parameter. This value is set large enough so that at least every other ESH PDU may be lost in the subnetwork or discarded due to lack of resources and the Configuration Information will still be maintained. The value must be set small enough so that Intermediate Systems can respond in a timely fashion to End Systems becoming available or unavailable.

The SN_Destination_Address parameter is set to the group address that indicates «All Intermediate System Network Entities». This ensures that the PDU will reach all of the active Intermediate Systems.

Note 6:

The actual value of the SN_Destination_Address used to mean «All Intermediate System Network Entities» is subnetwork-dependent and will most likely vary from subnetwork to subnetwork. It would of course be desirable on widely-used subnetwork types (such as those based on ISO 8802) that this value, and the value of the «All End System Network Entities» group address, be standardized.

9.1.1 Use of ESH in Query Configuration Procedure

An ESH PDU is also transmitted when an End System attached to a Broadcast subnetwork receives an NPDU addressed to one of its NSAPs, with the SN_Destination_Address from the SN_UNITDATA Indication set to the group address «All End System Network Entities». This occurs as a result of another End System having performed the Query Configuration procedure described in 9.3.1.

The End System constructs an ESH PDU identical in content to the ESH PDU constructed on expiry of the Configuration Timer (see 9.1) for the NSAP to which the received NPDU was addressed. It then transmits the ESH PDU to the source of the original NPDU by issuing an SN_UNITDATA Request with the SN_Destination_Address set to the value of the SN_Source_Address received in the SN_UNITDATA Indication with the original NPDU.

9.2 Operation of Configuration Information Recording

When an End System receives ESH or ISH PDUs, it extracts the Configuration Information, and adds or replaces the corresponding Configuration Information in the local Network Entity's Routing Information Base. If insufficient space is available to store new Configuration Information, the PDU is discarded. No error report is generated.

9.3 Operation of Configuration Request

9.3.1 CLNS - Use of Query Configuration Procedure

The Query Configuration procedure is performed under the following circumstances:

(i) The End System is attached to a Broadcast subnetwork,

(ii) There is no Intermediate System currently reachable on the subnetwork (i.e. no ISH PDUs have been received since the last information was flushed by the Flush Old Configuration function),

(iii) The Network Layer's Route PDU function needs to obtain the SNPA address to which to forward a PDU destined for a certain NSAP, and

(iv) The SNPA address cannot be obtained either by a local transformation or a local table lookup.

Note 7:

Despite appearances, this is actually a quite common case, since it is likely that there will be numerous isolated Local Area Networks without Intermediate Systems to rely upon for routing information (for example via the Request Redirect function of this protocol).

Further, if the Intermediate System(s) are temporarily unavailable, without this capability communication on the local subnetwork would suffer unless manually-entered tables were present in each End System or all NSAPs of the subnetwork had the subnetwork SNPA address embedded in them.

The End System, when needing to route an NPDU to a destination NSAP whose SNPA is unknown, issues an SN_UNITDATA Request with the NPDU as the SN_Userdata. The SN_Destination_Address parameter is set to the group address that indicates «All End System Network Entities».

Subsequently an ESH PDU may be received containing the NSAP address along with the corresponding SNPA address (see 9.1.1). In such a case the End System executes the Record Configuration procedure for that NSAP, and therefore will be able to route subsequent PDUs to that destination using the specified SNPA. If no ESH PDU is received, the End System may declare the destination NSAP not reachable. The length of time to wait for a response or the possibility of repeating the process some number of times before returning a failure are local matters and are not specified in this Technical Report.

9.3.2 CONS - Use of Broadcast CALL

The Broadcast CALL REQUEST procedure is used, as described in clause 6.7 of ISO 8881.

9.4 Operation of Redirect Requests

9.4.1 Connectionless-mode Network Service (CLNS)

9.4.1.1 Record Redirect Procedure

This procedure is invoked whenever a Redirect (RD) PDU is received. An RD PDU is generated by Intermediate Systems in order to inform End Systems of a better route. The Record Redirect function extracts the

Redirect Information and adds or replaces the corresponding Redirection Information in the local Network Entity's Routing Information Base. The essential information is the redirection mapping from a Destination Address (DA) to a subnetwork address, along with the corresponding QoS parameters and the Holding Time for which this mapping is to be considered valid.

Note 8:

If insufficient memory is available to store new Redirection Information, the RD PDU may be safely discarded since in any event the original Intermediate System will continue to forward PDUs on behalf of this Network Entity.

9.4.1.2 Flushing Old Redirect Information

A procedure is executed periodically to remove stale Redirection Information from the local Network Entity's Routing Information Base. All entries obtained from RD PDUs previously received are scanned, and if the Holding Timer (HT) has elapsed, the information corresponding to that Destination Address is discarded.

9.4.1.3 Refresh Redirect Procedure

This procedure is invoked whenever an NPDU is received by a destination ES. It is closely coupled with the function that processes received NPDUs at a destination Network Entity (this is the PDU Decomposition function in ISO 8473). The purpose of this function is to increase the longevity of a redirection without allowing an incorrect route to persist indefinitely.

The Source Address (SA) and QoS parameters associated with the NPDU are extracted and compared to any Destination Address (DA) and QoS parameters being maintained in the Routing Information Base (such information would have been stored by the Record Redirect procedure). If a corresponding entry is found, the previous hop of the PDU is obtained from the SN_Source_Address parameter of the SN_UNITDATA Indication primitive by which it was received. If this address matches the next hop address stored with the Redirection Information, the remaining Holding Time for the redirection is reset to the original Holding Time that was obtained from the RD PDU.

Note 9:

The purpose of this function is to avoid timing out redirection entries when the Network Entity is receiving return traffic from the destination via the same path over which it is currently sending traffic. This is particularly useful when the destination system is on the same subnetwork as the source, since after one redirect no IS need be involved in the ES-to-ES traffic.

This function must operate in a very conservative fashion however, to prevent the formation of «black holes». The remaining Holding Time should be refreshed only under the exact conditions specified above. For a discussion of the issues surrounding the refresh of Redirection Information, see Appendix A.

9.4.2 Connection-mode Network Service (CONS)

When an attempt is made to establish a connection via an Intermediate System which knows of a better route, then the call request will be rejected by means of a CLEAR

INDICATION packet containing the diagnostic code value of 120 (temporary routing problem). The User Data field of the CLEAR INDICATION will contain an RD PDU giving details of the preferred route. If the End System wishes to establish a connection to the destination NSAP, it establishes a new call to the SNPA specified in the RD PDU.

The Redirect Information is then extracted and stored in the Routing Information Base as described in 9.4.1, for possible use in establishing further connections to the same destination NSAP.

9.5 Operation of Protocol Error Processing

Any PDU which contains a valid Network Layer Protocol Identifier (NPID) as defined in 10.2.2 and a valid Version/ Protocol Identifier Extension (V/P) as defined in 10.2.4 shall be considered a protocol error if it deviates in any way from the structure and encoding of PDUs defined in the remainder of Clause 10. Any PDU which is a protocol error shall be discarded.

Note 10:

PDUs which do not contain both the correct NPID and the correct V/P are outside the scope of this Technical Report.

9.6 Operation of PDU Error Detection

The operation of PDU Error Detection is identical to that described in clause 6.11 of ISO 8473.

10. STRUCTURE AND ENCODING OF PROTOCOL DATA UNITS (PDUs)

10.1 Structure

All Protocol Data Units (PDUs) shall contain an integral number of octets. The octets in a PDU are numbered sequentially, the first octet being numbered one (1). The bits in an octet are numbered from one (1) to eight (8), where bit one (1) is the low-order bit.

When consecutive octets are used to represent a binary number, the lowest octet number has the most significant value.

Note 11:

The encoding of the PDUs for this protocol is compatible with that used in ISO 8473.

Note 12:

When the encoding of a PDU is illustrated using a diagram within this Clause, the following representation is used:

- octets are shown with the lowest-numbered octet to the left, higher-numbered octets being further to the right,
- within an octet, bits are shown with bit eight (8) to the left and bit one (1) to the right.

PDUs shall contain, in the following order:

(i) the Fixed part,

(ii) the Network Address part,

(iii) the Subnetwork Address part, if present, and

(iv) the Options part, if present.

10.2 Fixed Part

10.2.1 General

The Fixed part contains frequently occurring parameters including the Type (ESH, ISH, or RD) of the PDU. The length and the structure of the Fixed part are defined by the PDU code.

The Fixed part has the following format:

	Octet
Network Layer Protocol Identifier (NPID)	1
Length Indicator (LI)	2
Version/Protocol Identifier Extension (V/P)	3
reserved (must be set to ZEROs)	4
0 | 0 | 0 | Type (TP)	5
Holding Time	6,7
Checksum (CS)	8,9

Figure 3: PDU Header - Fixed Part

10.2.2 Network Layer Protocol Identifier (NPID)

This field shall be set to bit pattern 1000 0010.

10.2.3 Length Indicator (LI)

The length is indicated in binary notation, with a maximum value of 254 (1111 1110). The length indicated is the length of the entire PDU (which consists entirely of header, since this protocol does not carry user data) in octets, as described in Clause 10. The value 255 (1111 1111) is reserved for possible future extensions.

10.2.4 Version/Protocol Identifier Extension (V/P)

This field shall be set to bit pattern 0000 0000.

10.2.5 Type (TP)

The Type field identifies the type of the PDU. Allowed values are given in Figure 4; all other values are reserved.

Bits		5	4	3	2	1
ESH PDU		0	0	0	1	0
ISH PDU		0	0	1	0	0
RD PDU		0	0	1	1	0

Figure 4: Valid PDU Types

10.2.6 Holding Time

The Holding Time field specifies for how long the receiving Network Entity should retain the Configuration/Routing Information contained in this PDU. The receiving Network Entity should discard any information obtained from this PDU from its internal state when the Holding Time expires.

The Holding Time field is encoded as an integral number of seconds. A value of zero is to be interpreted as an infinite Holding Time.

10.2.7 PDU Checksum (CS)

The checksum is computed on the entire PDU.

A checksum value of zero is reserved to indicate that the checksum is to be ignored. The operation of the PDU Error Detection function ensures that the value zero does not represent a valid checksum. A non-zero value indicates that the checksum shall be processed. If the checksum calculation fails, the PDU shall be discarded.

10.3 Network Address Part

10.3.1 General

Address parameters are distinguished by their location. The different PDU Types carry different address parameters however. The ESH PDU carries a Source NSAP Address (SA), the ISH PDU carries an Intermediate System Network Entity Title (NET), and the RD PDU carries a Destination NSAP Address (DA).

10.3.2 Network Protocol Address Information Encoding (NPAI)

The Destination and Source Addresses are Network Service Access Point (NSAP) addresses as defined in ISO 8348/DAD2, Addendum to the Network Service Definition Covering Network Layer Addressing. The Network Entity Title address parameter in defined in 4.4.4 of this Technical Report. The Destination Address (DA), Source Address (SA), and Network Entity Title (NET) are encoded as NPAI using the binary syntax defined in clause 8.3.1 of ISO 8348/DAD2.

The address information is of variable length. Each address parameter is encoded as follows:

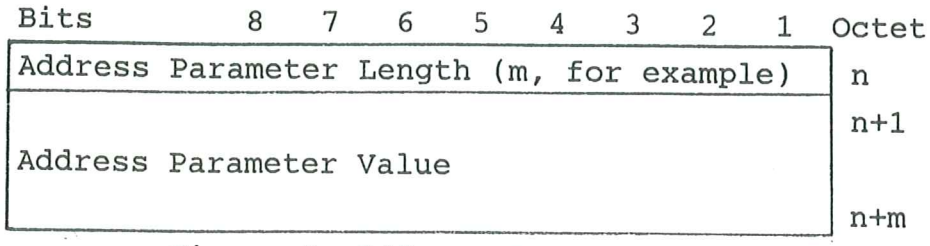


Figure 5: Address Parameters

10.3.3 Source Address Parameter for ESH PDU

The Source Address (SA) is the NSAP address of an NSAP served by the Network Entity sending the ESH PDU. It is encoded in the ESH PDU as follows:

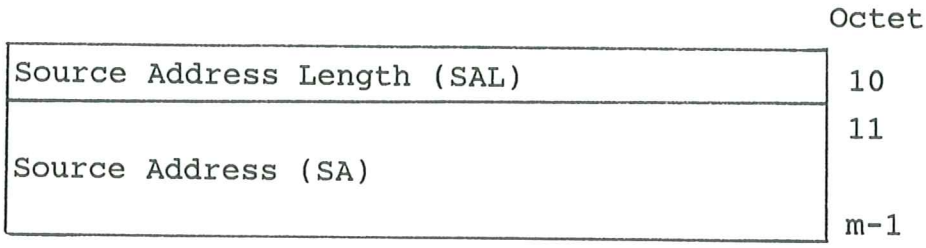


Figure 6: ESH PDU - Network Address Part

10.3.4 Network Entity Title Parameter for ISH PDU

The Network Entity Title parameter is the Network Entity Title (NET) of the Intermediate System sending the ISH PDU. It is encoded in the ISH PDU as follows:

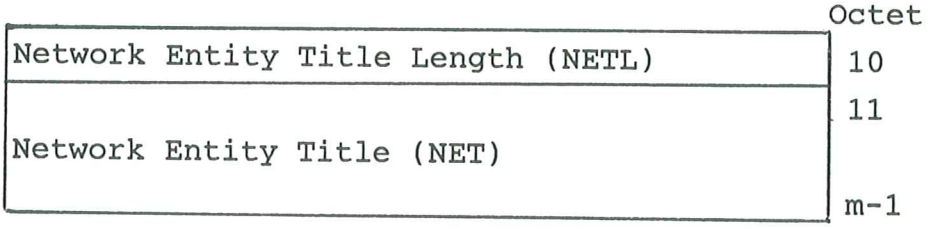


Figure 7: ISH PDU - Network Address Part

10.3.5 Destination Address Parameter for RD PDU

The Destination Address (DA) is the NSAP address of a destination associated with some NPDU being forwarded by the Intermediate System sending the RD PDU. It is encoded in the RD PDU as follows:

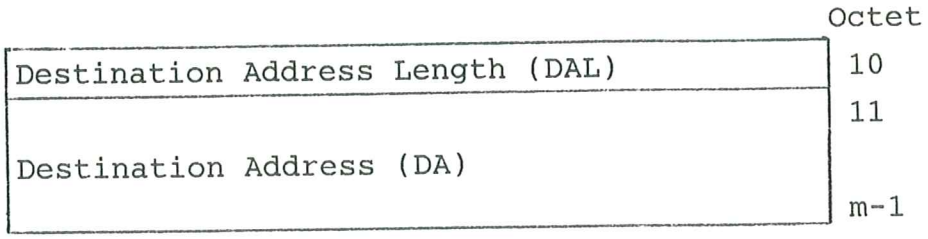


Figure 8: RD PDU - Network Address Part

10.4 Subnetwork Address Part

The Subnetwork Address part is present only in RD PDUs. It is used to indicate the subnetwork address of another Network Entity on the same subnetwork as the End System (and Intermediate System) which may be a better path to the destination specified in the Network Address part.

The Subnetwork Address parameter is encoded in the same manner as the Network Address parameters; see Figure 5.

10.4.1 Subnetwork Address Parameter for RD PDU

The Subnetwork Address parameter is encoded in the RD PDU as shown in Figure 9:

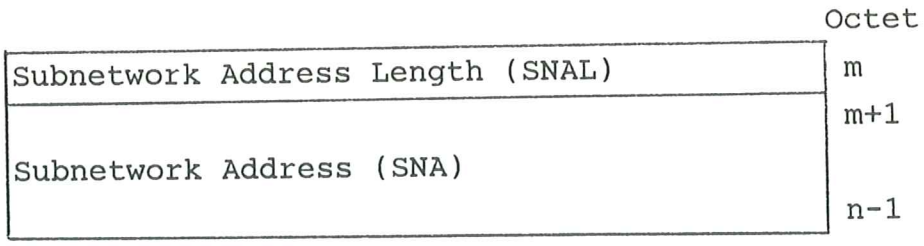


Figure 9: RD PDU - Address Part

10.5 Options Part

10.5.1 General

The Options part is used to convey optional parameters. The Options part of the PDU Header is illustrated below:

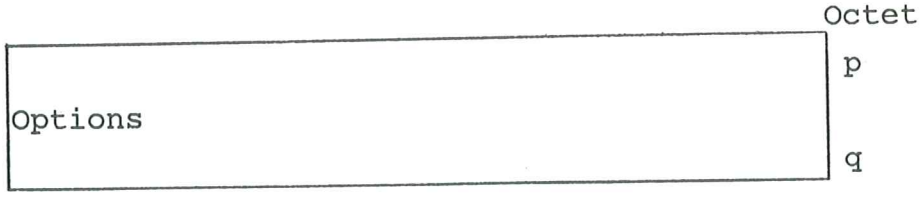


Figure 10: All PDUs - Options Part

If the Options part is present, it may contain one or more parameters. The number of parameters that may be contained in the Options part is constrained by the length of the Options part, which is determined by the following formula:

PDU length - (length of Fixed part + length of Address part),

and by the length of the individual optional parameters.

Parameters defined in the Options part may appear in any order. Duplication of options is not permitted. Receipt of a PDU with an option duplicated must be treated as a protocol error.

The encoding of parameters contained within the Options part of the PDU header is illustrated in Figure 11:

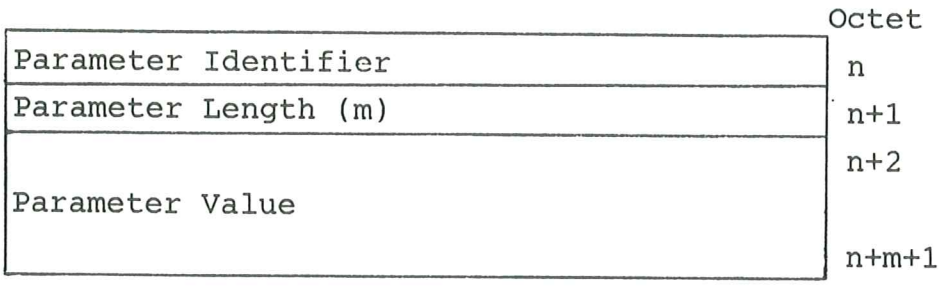


Figure 11: Encoding of Option Parameters

The Parameter Identifierfield is coded in binary and, without extensions, provides a maximum of 255 different parameter identifiers. No Parameter Identifiers shall use bits 8 and 7 with the value 00, so the actual maximum number of available parameters is lower. The Parameter Identifier of 255 (binary 1111 1111) is reserved for possible future extensions.

The Parameter Length field indicates the length, in octets, of the Parameter Value field. The length is indicated by a positive binary number, m, with a theoretical maximum value of 254. The practical maximum value of m is lower. For example, in the case of a single parameter contained within the Options part, two octets are required for the Parameter Identifier and the Parameter Length indicators. Thus the value of m is limited to:

m = 252 -(length of Fixed part + length of Address part).

For each subsequent parameter the maximum value of m decreases.

The Parameter Value field contains the value of the parameter identified in the Parameter Identifier field.

The following parameters are permitted in the Options part.

10.5.2 Security

The Security parameter conveys information about the security requested in the Data PDU that caused the RD PDU to be generated.

This parameter has the same encoding and semantics as the Security parameter in ISO 8473:

Parameter Identifier: 1100 0101

Parameter Length: variable

Parameter Value: see clause 7.5.3 of ISO 8473

10.5.3 Quality of Service Maintenance

The Quality of Service (QoS) parameter conveys information about the QoS requested in the Data PDU that caused the RD PDU to be generated.

This parameter has the same encoding and semantics as the QoS Maintenance parameter in ISO 8473:

Parameter Identifier: 1100 0011

Parameter Length: variable

Parameter Value: see clause 7.5.6 of ISO 8473

10.5.4 Priority

The Priority parameter conveys information about the priority requested in the Data PDU that caused the RD PDU to be generated.

This parameter has the same encoding and semantics as the Priority parameter in ISO 8473:

Parameter Identifier: 1100 1101

Parameter Length: variable

Parameter Value: see clause 7.5.7 of ISO 8473

10.5.5 Additional Domain-Specific Information

This parameter is present only in ISH PDUs, and conveys additional information which is defined by the authority responsible for the NSAP addressing domain to which the IS Network Entity belongs. Use by End Systems of the information contained in this parameter is optional.

Parameter Identifier: 1100 1110

Parameter Length: variable

Parameter Value: as determined by the authority responsible for the domain

Note 13:

Authorities which have adopted the ECMA Standard for the Domain-Specific Part of Network Layer Addresses (Standard ECMA-117) are recommended to encode this parameter value in the following manner:

The Parameter Value is itself comprised of a number of parameters, encoded in the form of Identifier, Length, and Value fields in the same way as shown in Figure 11 for Option parameters.

The values of the Identifier field are as follows:

0000 0000 to 0111 1111 : reserved for future allocation

1000 0000 to 1011 1111 : private encodings as required by the authority

1100 0000 : indicates that the Value field contains the Subnet-ID value allocated to the subnetwork on which this ISH is being transmitted

1100 0001 to 1111 1110 : reserved for future ECMA definition

1111 1111 : reserved for expansion

Duplication of parameters is treated as a protocol error.

10.6 End System Hello Protocol Data Unit (ESH PDU)

10.6.1 Structure

The ESH PDU has the format shown in Figure 12:

	Octet
Network Layer Protocol Identifier (NPID)	1
Length Indicator (LI)	2
Version/Protocol Identifier Extension (V/P)	3
reserved (must be set to ZEROs)	4
0 | 0 | 0 | Type (TP)	5
Holding Time	6,7
Checksum (CS)	8,9
Source Address Length (SAL)	10
Source Address (SA)	11 m-1
Options	m p-1

Figure 12: ESH PDU Format

10.7 Intermediate System Hello Protocol Data Unit (ISH PDU)

10.7.1 Structure

The ISH PDU has the format shown in Figure 13:

	Octet
Network Layer Protocol Identifier (NPID)	1
Length Indicator (LI)	2
Version/Protocol Identifier Extension (V/P)	3
reserved (must be set to ZEROs)	4
0 | 0 | 0 | Type (TP)	5
Holding Time	6,7
Checksum (CS)	8,9
Network Entity Title Length (NETL)	10
Network Entity Title (NET)	11 m-1
Options	m p-1

Figure 13: ISH PDU Format

10.8 Redirect Protocol Data Unit (RD PDU)

10.8.1 Structure

The RD PDU has the format shown in Figures 14 and 15:

	Octet
Network Layer Protocol Identifier (NPID)	1
Length Indicator (LI)	2
Version/Protocol Identifier Extension (V/P)	3
reserved (must be set to ZEROs)	4
0 | 0 | 0 | Type (TP)	5
Holding Time	6,7
Checksum (CS)	8,9
Destination Address Length (DAL)	10
Destination Address (DA)	11 ... m-1
Subnetwork Address Length (SNAL)	m
Subnetwork Address (SNA)	m+1 ... n-1
Network Entity Title Length (NETL)	n
Network Entity Title (NET)	n+1 ... p-1
Options	p ... q-1

Figure 14: RD PDU Format when Redirect is to an IS

	Octet
Network Layer Protocol Identifier (NPID)	1
Length Indicator (LI)	2
Version/Protocol Identifier Extension (V/P)	3
reserved (must be set to ZEROs)	4
0 | 0 | 0 | Type (TP)	5
Holding Time	6,7
Checksum (CS)	8,9
Destination Address Length (DAL)	10
Destination Address (DA)	11 m-1
Subnetwork Address Length (SNAL)	m
Subnetwork Address (SNA)	m+1 n-1
Network Entity Title Length (NETL) = zero	n
Options	n+1 p-1

Figure 15: RD PDU Format when Redirect is to an ES

11. SUBJECTS FOR FURTHER STUDY

(a) End System Routing functions for use with other methods of providing the Network Service (over ISDN, for example).

(b) Management operations related to the routing function, for example:

- Initial establishment of the Routing Information Base,
- Setting of Timers,
- Dynamic changing of entries in the Routing Information Base (due to configuration changes, breakdown, maintenance purposes, additional QoS, and so on),
- Traffic statistics.

APPENDIX A

SUPPORTING TECHNICAL MATERIAL

A.1 Use of Timers

This protocol makes extensive use of timers to ensure the timeliness and accuracy of information disseminated using the Configuration and Route Redirection functions. A discussion of the rationale for using these timers follows, together with some background on how they operate.

Systems using this protocol learn about other systems exclusively by receiving PDUs sent by those systems. In a connectionless environment, a system must periodically receive updated information to ensure that the information it previously received is still correct. For example, if a system on a subnetwork becomes unavailable (either it has ceased operating, or its SNPA becomes inoperative) the only way another system can detect this fact is by the absence of transmissions from that system. If information were retained in the absence of new PDUs being received, Configuration and/or Routing Information would inevitably become incorrect. The Holding Timers specified by this protocol guarantee that old information will not be retained indefinitely.

A useful way of thinking of the Configuration and Route Redirection Information is as a cache maintained by each system. The cache is periodically flushed to ensure that only up-to-date information is stored. Unlike most caches, however, the time to retain information is not a purely local matter. Rather, information is held for a period of time specified by the source of the information.

Some examples will help clarify this operation.

A.1.1 Example of Holding Timer for Route Redirection Information

Route Redirection Information is obtained by an End System through the Request Redirect procedure (see 9.4). It is quite possible that an Intermediate System might redirect an End System to another IS which has recently become unavailable (this might happen if the IS-to-IS routing algorithm is still converging following a configuration change). If the Holding Timer were not present, or was set very long by the sending IS, an End System would have been redirected into a «black hole» from which none of its Data PDUs would ever emerge. The length of the Holding Time on Redirects specifies, in essence, the length of time black holes are permitted to exist.

On the other hand, setting the Holding Timer on Route Redirects very short to minimize the effect of black holes has other, undesirable, consequences. First, for each PDU that causes a redirect, an additional PDU beside the original Data PDU must be composed and transmitted; this increases overhead. Second, each time a «working» redirect's Holding

Timer expires, the redirected End System will revert to a poorer route for at least one PDU.

A.1.2 Example of Holding Timer for Configuration Information

A similar type of problem can occur with respect to Configuration Information. If the Holding Time of an ISH PDU is set very long, and the only Intermediate System (which has been sending this Configuration Information) on the subnetwork becomes unavailable, a subnetwork-wide black hole can form. During this time, End Systems on the subnetwork may not be able to communicate with each other because they presume that an Intermediate System is operating which will forward their Data PDUs to destination ESs on the local subnetwork and return RD PDUs. Once the Holding Time expires, the ESs will realise that no IS is available and will take their only recourse, which is to send their traffic directly on the local subnetwork.

Given the types of problems that can occur, it is important that responsibility for incorrect information can be unambiguously assigned to the source of the information. For this reason all Holding Timers are calculated by the source of the Configuration or Route Redirection Information and communicated explicitly to each recipient in the appropriate PDU.

A.2 Refresh and Timeout of Route Redirection Information

The protocol allows End Systems to refresh Route Redirection Information without first allowing the Holding Time to expire and being redirected by an Intermediate System for a second (or subsequent) time. Such schemes are prevalent in connectionless subnetworks and are often called Reverse Path Information, Previous Hop Cache, or something similar.

Refreshing the Route Redirection Information has obvious performance benefits, but can be dangerous if not handled in a very conservative fashion. In order for a redirection to be safely refreshed, all of the following conditions must hold:

(i) The Source Address of the received PDU must be exactly the same as the Destination Address specified in a prior RD PDU (this defines a «match» on the Route Redirection Information). Making assumptions about the equivalence of abbreviated addresses, group addresses, or similar «special» addresses is dangerous since routing for these addresses cannot be assumed to be the same.

(ii) The Quality of Service (QoS) parameters of the received PDU must be exactly the same as the QoS parameters specified in the matching (by Destination Address) redirection entry. Again, there is no guarantee that PDUs with different QoS parameters will be routed the same way. It is quite possible that the redirected path is even a black hole for certain values of the QoS parameters (the Security field is a good example).

(iii) The «previous hop» of the received Data PDU must match the «next hop» stored in the Route Redirection Information. Specifically, the SN_Source_Address of the

SN_UNITDATA Indication which received the PDU must match exactly the SN_Destination_Address specified in the redirect to be used for sending traffic via the SN_UNITDATA Request primitive. This comparison ensures that redirects are refreshed only when the reverse traffic is being received from the same IS (or destination ES) as the forward traffic is being sent through (or to). This check ensures that redirects are not refreshed solely on the basis of traffic being received from the destination. It is quite possible that the traffic is simply indicating that the forward path in use is not working!

Note that these conditions still allow refresh in the most useful and common cases where either the destination is another ES on the same subnetwork as the source ES, or the redirection is to an IS which is passing traffic to and from the destination in both directions (i.e. the path is symmetric).

A.3 Configuration Information for CONS on Broadcast Subnetworks

The functions which provide Configuration Information, involving the use of broadcast messages such as ESH, can be applied equally well to disseminating Configuration Information for use in CONS. However, it can be expected that methods of supporting CONS, such as the use of X.25 as proposed in ISO 8881, will include mechanisms for detecting failure of the Network Entities to or through which a connection is established. In this case detection of failure by the absence of «hello» messages becomes less important, and it may be possible to use larger values for the Holding Timers than would be the case for CLNS (i.e. larger than twice the Configuration Timer, which is still constrained by the need to notify the presence of the system to other newly-operational systems).

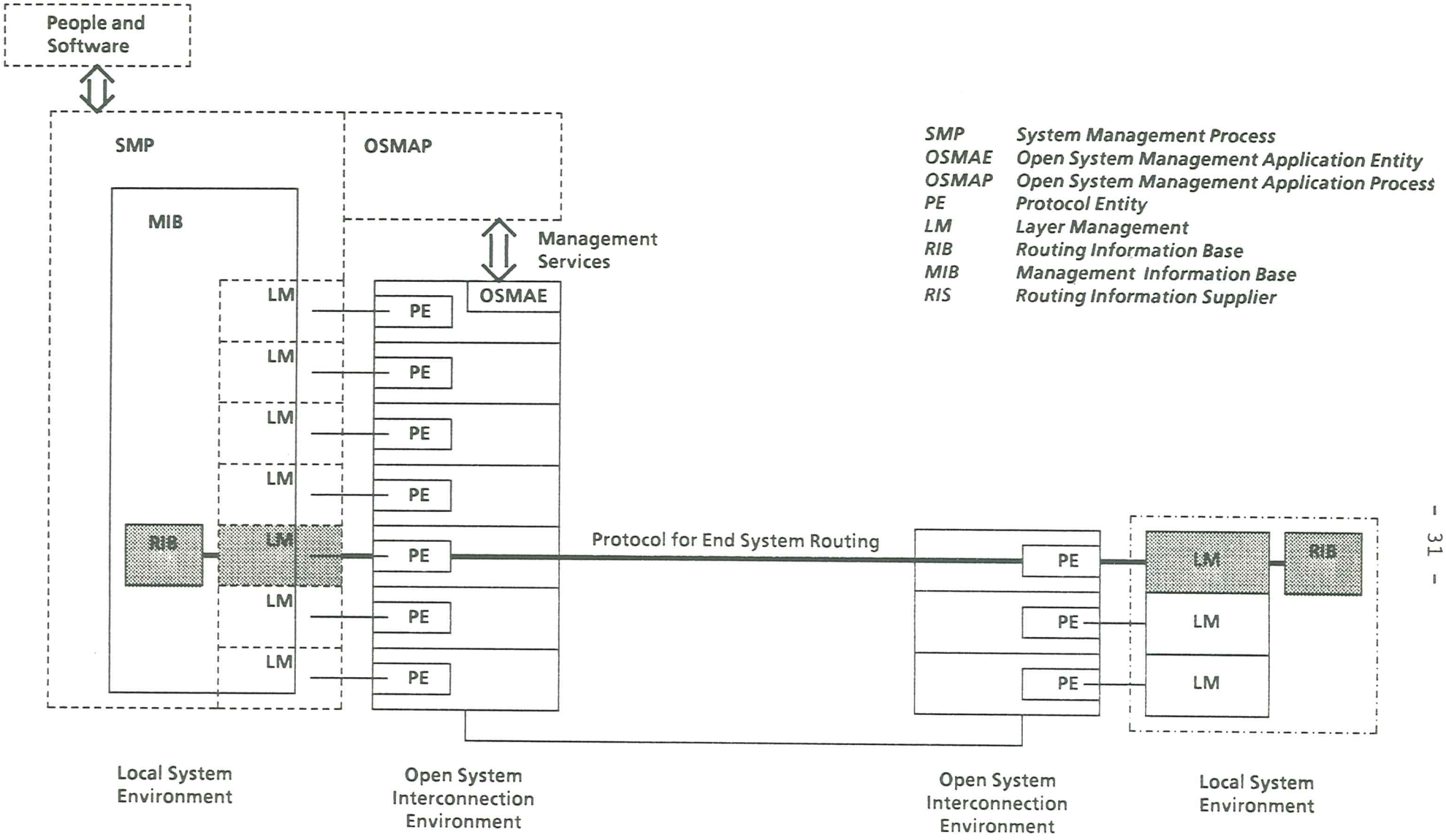
It is appropriate to define an infinite Holding Timer for systems using only this type of protocol in order to relieve communicating systems of the necessity to run a timer. The principle remains, however, that the system which originates a «hello» message is responsible for deciding the time for which it is valid.

APPENDIX B

RELATIONSHIP TO MANAGEMENT FRAMEWORK

The ECMA Technical Report on the subject of OSI Management Framework describes the position of this routing function as a layer-manager to layer-manager information exchange.

This is illustrated in the diagram on the next page.



SMP *System Management Process*
OSMAE *Open System Management Application Entity*
OSMAP *Open System Management Application Process*
PE *Protocol Entity*
LM *Layer Management*
RIB *Routing Information Base*
MIB *Management Information Base*
RIS *Routing Information Supplier*

Integration of ES-Routing into Management Framework

APPENDIX C

LIST OF ACRONYMS

CL	Connectionless-mode
CLNS	Connectionless-mode Network Service
CO	Connection-mode
CONS	Connection-mode Network Service
CS	Checksum
CT	Configuration Timer
DA	Destination Address
DAL	Destination Address Length
ES	End System
ESH	End System Hello
ESH PDU	End System Hello Protocol Data Unit
HT	Holding Timer
IS	Intermediate System
ISH	Intermediate System Hello
ISH PDU	Intermediate System Hello Protocol Data Unit
LAN	Local Area Network
LI	Length Indicator
NET	Network Entity Title
NETL	Network Entity Title Length
NPAI	Network Protocol Address Information
NPDU	Network Protocol Data Unit
NPID	Network Layer Protocol Identifier
NS	Network Service
NSAP	Network Service Access Point
PDU	Protocol Data Unit
QoS	Quality of Service
RD	Redirect
RD PDU	Redirect Protocol Data Unit
RIS	Routing Information Supplier
RT	Redirect Timer
SA	Source Address
SAL	Source Address Length
SN	Subnetwork
SNA	Subnetwork Address
SNAcP	Subnetwork Access Protocol
SNAL	Subnetwork Address Length
SNPA	Subnetwork Point of Attachment
SNSDU	Subnetwork Service Data Unit
TP	Type
V/P	Version/Protocol Identifier Extension