INTERNATIONAL TELECOMMUNICATION UNION

ITU-T G.7712/Y.1703

(03/2003)

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

SERIES G: TRANSMISSION SYSTEMS AND MEDIA, DIGITAL SYSTEMS AND NETWORKS

Digital terminal equipments – Operations, administration and maintenance features of transmission equipment

SERIES Y: GLOBAL INFORMATION INFRASTRUCTURE AND INTERNET PROTOCOL ASPECTS

Internet protocol aspects – Operation, administration and maintenance

Architecture and specification of data communication network

ITU-T Recommendation G.7712/Y.1703

ITU-T G-SERIES RECOMMENDATIONS

TRANSMISSION SYSTEMS AND MEDIA, DIGITAL SYSTEMS AND NETWORKS

INTERNATIONAL TELEPHONE CONNECTIONS AND CIRCUITS	G.100–G.199
GENERAL CHARACTERISTICS COMMON TO ALL ANALOGUE CARRIER-TRANSMISSION SYSTEMS	G.200–G.299
INDIVIDUAL CHARACTERISTICS OF INTERNATIONAL CARRIER TELEPHONE SYSTEMS ON METALLIC LINES	G.300–G.399
GENERAL CHARACTERISTICS OF INTERNATIONAL CARRIER TELEPHONE SYSTEMS ON RADIO-RELAY OR SATELLITE LINKS AND INTERCONNECTION WITH METALLIC LINES	G.400–G.449
COORDINATION OF RADIOTELEPHONY AND LINE TELEPHONY	G.450–G.499
TESTING EQUIPMENTS	G.500–G.599
TRANSMISSION MEDIA CHARACTERISTICS	G.600–G.699
DIGITAL TERMINAL EQUIPMENTS	G.700–G.799
DIGITAL NETWORKS	G.800–G.899
DIGITAL SECTIONS AND DIGITAL LINE SYSTEM	G.900–G.999
QUALITY OF SERVICE AND PERFORMANCE	G.1000–G.1999
TRANSMISSION MEDIA CHARACTERISTICS	G.6000–G.6999
DIGITAL TERMINAL EQUIPMENTS	G.7000–G.7999
General	G.7000–G.7099
Coding of analogue signals by pulse code modulation	G.7100–G.7199
Coding of analogue signals by methods other than PCM	G.7200–G.7299
Principal characteristics of primary multiplex equipment	G.7300–G.7399
Principal characteristics of second order multiplex equipment	G.7400–G.7499
Principal characteristics of higher order multiplex equipment	G.7500–G.7599
Principal characteristics of transcoder and digital multiplication equipment	G.7600–G.7699
Operations, administration and maintenance features of transmission equipment	**G.7700–G.7799**
Principal characteristics of multiplexing equipment for the synchronous digital hierarchy	G.7800–G.7899
Other terminal equipment	G.7900–G.7999
DIGITAL NETWORKS	G.8000–G.8999

For further details, please refer to the list of ITU-T Recommendations.

ITU-T Recommendation G.7712/Y.1703

Architecture and specification of data communication network

Summary

This Recommendation defines the architecture requirements for a Data Communication Network (DCN) which may support distributed management communications related to the Telecommunication Management Network (TMN), distributed signalling communications related to the Automatic Switched Transport Network (ASTN), and other distributed communications (e.g., Orderwire or Voice Communications, Software Download). The DCN architecture considers networks that are IP-only, OSI-only, and mixed (i.e., support both IP and OSI). The interworking between parts of the DCN supporting IP-only, parts supporting OSI-only, and parts supporting both IP and OSI are also specified.

Various applications (e.g., TMN, ASTN, etc.) require a packet-based communications network to transport information between various components. For example, the TMN requires a communications network, which is referred to as the Management Communication Network (MCN) to transport management messages between TMN components (e.g., NEF component and OSF component). ASTN requires a communication network which is referred to as the Signalling Communication Network (SCN) to transport signalling messages between ASTN components (e.g., CC components). This Recommendation specifies data communication functions that can be used to support one or more application's communication network.

The data communication functions provided in the 11/2001 version of this Recommendation support connection-less network services. This revision of the Recommendation adds the support of connection-oriented network SCN services by including specific MPLS-based mechanism.

This Recommendation forms part of a family of Recommendations covering transport networks.

Source

ITU-T Recommendation G.7712/Y.1703 was revised by ITU-T Study Group 15 (2001-2004) and approved under the WTSA Resolution 1 procedure on 16 March 2003.

Document history	
Issue	**Notes**
1.0	Output of Q14/15 October 2001 meeting
1.1	Output of Q14/15 April 2002 meeting
1.2	Cleanup of 1.1
1.3	Output of Q14/15 October 2002 meeting
1.4	Cleanup of version 1.3: Replace section headings 7.1.a, etc. with 7.1.13, etc., Remove editor comments
1.5	Output from the Q14/15 meeting and submitted for consent

Keywords

Data Communication Network, Internet Protocol (IP), Open System Interface (OSI).

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications. The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

Compliance with this Recommendation is voluntary. However, the Recommendation may contain certain mandatory provisions (to ensure e.g. interoperability or applicability) and compliance with the Recommendation is achieved when all of these mandatory provisions are met. The words "shall" or some other obligatory language such as "must" and the negative equivalents are used to express requirements. The use of such words does not suggest that compliance with the Recommendation is required of any party.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, ITU had not received notice of intellectual property, protected by patents, which may be required to implement this Recommendation. However, implementors are cautioned that this may not represent the latest information and are therefore strongly urged to consult the TSB patent database.

CONTENTS

		Page
1	Scope	1
2	References	1
3	Terms and definitions	3
4	Abbreviations	4
5	Conventions	6
6	DCN characteristics	7
	6.1 TMN application	9
	6.2 ASTN application	16
	6.3 Other applications requiring communication networks	23
	6.4 Separation of various applications	23
7	DCN functional architecture and requirements	23
	7.1 Specification of data communication functions	24
	7.2 Provisioning requirements	35
	7.3 Security requirements	35
Annex A – Requirements for three-way handshaking		35
	A.1 Point-to-Point three-way adjacency TLV	35
	A.2 Adjacency three-way state	36
Annex B – Requirements for automatic encapsulation		37
	B.1 Introduction	37
	B.2 Scope	37
	B.3 Description of the AE-DCF	37
	B.4 Requirements and limitations	39
Appendix I – Constraints of the interworking functions in DCN		49
	I.1 General assumptions	49
	I.2 Common to all scenarios	49
Appendix II – Example implementation of automatic encapsulation		51
	II.1 Introduction	51
	II.2 Updates to Dijkstra's Algorithm	52
Appendix III – Commissioning guide for SDH NEs in dual RFC 1195 environment and impact of automatic encapsulation option		55
	III.1 Introduction	55
	III.2 Integrated IS-IS without automatic encapsulation	55
	III.3 Integrated IS-IS with automatic encapsulation	59
Appendix IV – Example illustration of packet 1+1 protection		62
	IV.1 Packet 1+1 protection overview	62
	IV.2 Packet 1+1 protection illustration	63
	IV.3 Operation of selector algorithm under various failure scenarios	65
Appendix V – Bibliography		69

ITU-T Recommendation G.7712/Y.1703

Architecture and specification of data communication network

1 Scope

This Recommendation defines the architecture requirements for a Data Communication Network (DCN) which may support distributed management communications related to the Telecommunication Management Network (TMN), distributed signalling communications related to the Automatic Switched Transport Network (ASTN), and other distributed communications (e.g., Orderwire or Voice Communications, Software Download). The DCN architecture considers networks that are IP-only, OSI-only, and mixed (i.e., support both IP and OSI). The interworking between parts of the DCN supporting IP-only, parts supporting OSI-only, and parts supporting both IP and OSI are also specified.

The DCN provides Layer 1 (physical), Layer 2 (data-link) and Layer 3 (network) functionality and consists of routing/switching functionality interconnected via links. These links can be implemented over various interfaces, including Wide Area Network (WAN) interfaces, Local Area Network (LAN) interfaces, and Embedded Control Channels (ECCs).

Various applications (e.g., TMN, ASTN, etc.) require a packet-based communication network to transport information between various components. For example, the TMN requires a communication network, which is referred to as the Management Communication Network (MCN) to transport management messages between TMN components (e.g., NEF component and OSF component). ASTN requires a communication network, which is referred to as the Signalling Communication Network (SCN) to transport signalling messages between ASTN components (e.g., CC components). This Recommendation specifies data communication functions that can be used to support one or more application's communication network.

The data communication functions provided in this Recommendation support connection-less network services. Additional functions may be added in future versions of this Recommendation to support connection-oriented network services.

2 References

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published. The reference to a document within this Recommendation does not give it, as a stand-alone document, the status of a Recommendation.

– ITU-T Recommendation G.707/Y.1322 (2000), *Network node interface for the synchronous digital hierarchy (SDH)*.

– ITU-T Recommendation G.709/Y.1331 (2003), *Interfaces for the Optical Transport Network (OTN)*.

– ITU-T Recommendation G.783 (2000), *Characteristics of synchronous digital hierarchy (SDH) equipment functional blocks*.

– ITU-T Recommendation G.784 (1999), *Synchronous digital hierarchy (SDH) management*.

– ITU-T Recommendation G.798 (2002), *Characteristics of optical transport network hierarchy equipment functional blocks*.

- ITU-T Recommendation G.807/Y.1302 (2001), *Requirements for automatically switched transport networks (ASTN)*.
- ITU-T Recommendation G.872 (2001), *Architecture of optical transport networks*.
- ITU-T Recommendation G.874 (2001), *Management aspects of the optical transport network element*.
- ITU-T Recommendation G.7710/Y.1701 (2001), *Common equipment management function requirements*.
- ITU-T Recommendation G.8080/Y.1304 (2001), *Architecture for the automatically switched optical networks (ASON)*.
- ITU-T Recommendation M.3010 (2000), *Principles for a telecommunications management network*.
- ITU-T Recommendation M.3013 (2000), *Considerations for a telecommunications management network.*
- ITU-T Recommendation M.3016 (1998), *TMN security overview*.
- ITU-T Recommendation Q.811 (1997), *Lower layer protocol profiles for the Q3 and X interfaces*.
- ITU-T Recommendation X.263 (1998) | ISO/IEC TR 9577:1999, *Information technology – Protocol identification in the Network Layer*.
- ISO/IEC 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/IEC 10589:2002, *Information technology – Telecommunications and information exchange between systems – Intermediate System to Intermediate System intra-domain routeing information exchange protocol for use in conjunction with the protocol for providing the connectionless-mode network service (ISO 8473)*.
- IETF RFC 791 (1981), *Internet Protocol DARPA Internet Program Protocol Specification*.
- IETF RFC 792 (1981), *Internet Control Message Protocol*.
- IETF RFC 826 (1982), *An Ethernet Address Resolution Protocol*.
- IETF RFC 894 (1984), *A Standard for the Transmission of IP Datagrams over Ethernet Networks*.
- IETF RFC 1122 (1989), *Requirements for Internet Hosts – Communication Layers*.
- IETF RFC 1172 (1990), *The Point-to-Point Protocol (PPP) Initial Configuration Options*.
- IETF RFC 1195 (1990), *Use of OSI IS-IS for Routing in TCP/IP and Dual Environments*.
- IETF RFC 1332 (1992), *The PPP Internet Protocol Control Protocol (IPCP)*.
- IETF RFC 1377 (1992), *The PPP OSI Network Layer Control Protocol (OSINLCP)*.
- IETF RFC 1661 (1994), *The Point-to-Point Protocol (PPP)*.
- IETF RFC 1662 (1994), *PPP in HDLC-like Framing*.
- IETF RFC 1812 (1995), *Requirements for IP Version 4 Routers*.
- IETF RFC 2328 (1998), *OSPF Version 2*.
- IETF RFC 2460 (1998), *Internet Protocol, Version 6 (IPv6) Specification*.

– IETF RFC 2463 (1998), *Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification*.

– IETF RFC 2472 (1998), *IP Version 6 over PPP*.

– IETF RFC 2740 (1999), *OSPF for IPv6*.

– IETF RFC 2784 (2000), *Generic Routing Encapsulation (GRE)*.

3 Terms and definitions

3.1 Terms defined in ITU-T Rec. G.709/Y.1331:

a) Optical Channel Data Unit (ODUk)

b) Optical Channel Transport Unit (OTUk)

c) Optical Overhead Signal (OOS)

3.2 Term defined in ITU-T Rec. G.784:

a) Data Communications Channel (DCC)

3.3 Terms defined in ITU-T Rec. G.807/Y.1302:

a) Automatic Switched Transport Network (ASTN)

b) Network – Network Interface (NNI)

c) User – Network Interface (UNI)

3.4 Terms defined in ITU-T Rec. G.8080/Y.1304:

a) Call Controller (CallC)

b) Connection Controller (CC)

c) Connection Controller Interface (CCI)

d) Subnetwork Controller (SNCr)

3.5 Terms defined in ITU-T Rec. G.874:

a) General Communications Channel (GCC)

b) General Management Communications Overhead (COMMS OH)

3.6 Terms defined in ITU-T Rec. G.7710/Y.1701:

a) X Management Network

b) X Management Subnetwork

3.7 Term defined in ITU-T Rec. G.872:

a) Optical transport network (OTN)

3.8 Terms defined in ITU-T Rec. M.3010:

a) Adaptation Device (AD)

b) Data Communications Function (DCF)

c) Mediation Device (MD)

d) Network Element (NE)

e) Network Element Function (NEF)

f) Operations System (OS)

g) Operations System Function (OSF)

h) Q-interface

i) Translation Function

j) Workstation Function (WSF)

3.9 Term defined in ITU-T Rec. M.3013:

a) Message Communications Function (MCF)

3.10 This Recommendation defines the following terms:

3.10.1 Data Communication Network (DCN): The DCN is a network that supports Layer 1 (physical), Layer 2 (data-link), and Layer 3 (network) functionality. A DCN can be designed to support transport of distributed management communications related to the TMN, distributed signalling communications related to the ASTN, and other operations communications (e.g., orderwire/voice communications, software downloads, etc.).

3.10.2 Embedded Control Channel (ECC): An ECC provides a logical operations channel between NEs. The physical channel supporting the ECC is technology specific. Examples of physical channels supporting the ECC are; a DCC channel within SDH, GCC channel within OTN OTUk/ODUk, or the COMMS OH channel within the OTN OOS.

3.10.3 IP routing interworking function: An IP Routing InterWorking Function allows IP topology or routes to be passed from one IP routing protocol to a different incompatible IP routing protocol. For example, an IP Routing InterWorking Function may form a gateway between an Integrated IS-IS routed DCN and an OSPF routed DCN.

3.10.4 network-layer interworking function: A Network-Layer InterWorking Function provides interoperability between nodes that support incompatible network-layer protocols. An example of a Network-Layer InterWorking Function is static GRE tunnels, or an AE-DCF.

3.10.5 Automatic Encapsulating Data Communication Function (AE-DCF): An AE-DCF automatically encapsulates packets when necessary so that they may be routed by NEs that would otherwise be unable to forward them. An AE-DCF also features a matching de-encapsulation function to restore the packet back to its original form once it has traversed incompatible NEs.

4 Abbreviations

This Recommendation uses the following abbreviations:

AD	Adaptation Device
AE-DCF	Automatic Encapsulating Data Communication Function
ARP	Address Resolution Protocol
ASON	Automatic Switched Optical Network
ASTN	Automatic Switched Transport Network
ATM	Asynchronous Transfer Mode
CallC	Call Controller
CC	Connection Controller
CCI	Connection Controller Interface
CLNP	ConnectionLess Network layer Protocol
CLNS	ConnectionLess Network layer Service
COMMS OH	General Management Communications Overhead
DCC	Data Communication Channel
DCF	Data Communication Function

DCN	Data Communication Network
DF	Don't Fragment
ECC	Embedded Control Channel
EMF	Equipment Management Function
ES	End System
ESH	End System Hello (ISO 9542)
ES-IS	End System-to-Intermediate System
GCC	General Communication Channel
GNE	Gateway Network Element
GRE	Generic Routing Encapsulation
HDLC	High Level Data Link Control
ICMP	Internet Control Message Protocol
ID	Identifier
IIH	IS-IS Hello
IntISIS	Integrated Intermediate System-to-Intermediate System
IP	Internet Protocol
IPCP	Internet Protocol Control Protocol
IPv4	Internet Protocol Version 4
IPv6	Internet Protocol Version 6
IS	Intermediate System
ISDN	Integrated Services Digital Network
ISH	Intermediate System Hello (ISO 9542)
IS-IS	Intermediate System-to-Intermediate System
IWF	Interworking Function
LAN	Local Area Network
LAPD	Link-Access Procedure D-Channel
LCN	Local Communication Network
LSP	Link State Protocol Data Unit
MAC	Media Access Control
MCF	Message Communication Function
MCN	Management Communication Network
MD	Mediation Device
MTU	Maximum Transmission Unit
NE	Network Element
NEF	Network Element Function
NLPID	Network Layer Protocol Identifier
NNI	Network-to-Network interface

NSAP	Network Service Access Point
ODUk	Optical Channel Data Unit
OOS	OTM Overhead Signal
OS	Operations System
OSC	Optical Supervisory Channel
OSF	Operations System Function
OSI	Open System Interface
OSINLCP	OSI Network Layer Control Protocol
OSPF	Open Shortest Path First
OTM	Optical Transport Module
OTN	Optical Transport Network
OTUk	Optical Channel Transport Unit
PDU	Protocol Data Unit
PPP	Point-to-Point Protocol
RFC	Request For Comment
SCN	Signalling Communication Network
SDH	Synchronous Digital Hierarchy
SID	System Identifier
SNCr	SubNetwork Controller
SP	Segmentation Permitted
SPF	Shortest Path First
TCP	Transmission Control Protocol
TF	Translation Function
TLV	Type Length Value
TMN	Telecommunication Management Network
TNE	Transport Network Element
UNI	User-to-Network Interface
WAN	Wide Area Network
WS	Work Station
WSF	Work Station Function
xMS	X Management Subnetwork

5 Conventions

The following conventions are used throughout this Recommendation:

Mixed DCN: A mixed DCN supports multiple network layer protocols (e.g., OSI and IPv4). It is possible in a mixed DCN, that the path between two communicating entities (e.g., an OS and a managed NE) will traverse some parts that only support one network layer protocol (e.g., OSI) and other parts that only support another network layer protocol (e.g., IPv4). To provide communication

between such entities, one network layer protocol should be encapsulated into the other network layer protocol at the boundary of those parts supporting different network layer protocols.

OSI-only DCN: An OSI-only DCN supports only CLNP as the network layer protocol. Therefore, the end-to-end path between two communicating entities (e.g., an OS and a managed NE) will support CLNP and encapsulation of one network layer protocol into another network layer protocol is not required to support such communications.

IPv4-only DCN: An IPv4-only DCN supports only IPv4 as the network layer protocol. Therefore the end-to-end path between two communicating entities (e.g., an OS and a managed NE) will support IPv4 and encapsulation of one network layer protocol into another network layer protocol is not required to support such communications.

IPv6-only DCN: An IPv6-only DCN supports only IPv6 as the network layer protocol. Therefore the end-to-end path between two communicating entities (e.g., an OS and a managed NE) will support IPv6 and encapsulation of one network layer protocol into another network layer protocol is not required to support such communications.

6 DCN characteristics

Various applications (e.g., TMN, ASTN, etc.) require a packet-based communication network to transport information between various components. For example, the TMN requires a communication network, which is referred to as the Management Communication Network (MCN) to transport management messages between TMN components (e.g., NEF component and OSF component). ASTN requires a communication network, which is referred to as the Signalling Communication Network (SCN) to transport signalling messages between ASTN components (e.g., CC components). This Recommendation specifies data communication functions that can be used to support one or more application's communication network.

Figure 6-1 illustrates example applications that can be supported via the DCN. Each application can be supported on separate DCNs or on the same DCN depending on the network design.

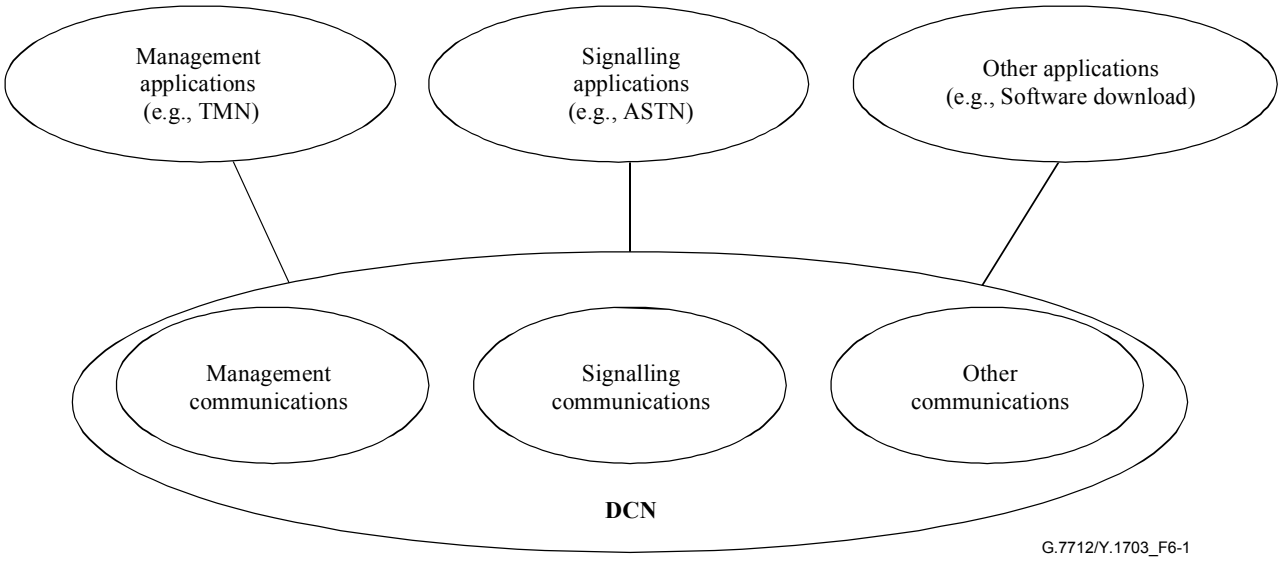


Figure 6-1/G.7712/Y.1703 – Example applications supported by a DCN

The conceptual DCN is a collection of resources to support the transfer of information among distributed components. As discussed above, examples of distributed communication that can be supported by the DCN are distributed management communications related to the TMN, and distributed signalling communications related to the ASTN. In the case of a DCN supporting

distributed management communications, the distributed components are TMN components (NEs, ADs, OSs, MDs, and WSs containing TMN functions such as OSF, TF, NEF, WSF). ITU-T Recs M.3010 and M.3013 provide further specifications for the TMN functions. In the case of a DCN supporting distributed signalling communications, the distributed components are ASTN components (NEs containing ASTN SNCr functions). ITU-T Recs G.807/Y.1302 and G.8080/Y.1304 provide further specifications for the ASTN functions.

A number of telecommunication technologies can support the DCN functions such as, circuit switching, packet switching, LAN, ATM, SDH, and the OTN. Important aspects of the DCN are the quality of service, information transfer rate, and diversity of routing to support specific operational requirements of the distributed communications supported across the DCN (e.g., distributed management communications, distributed signalling communications).

The goal of an interface specification is to ensure meaningful interchange of data between interconnected devices through a DCN to perform a given function (e.g., TMN function, ASTN function). An interface is designed to ensure independence of the type of device or of the supplier. This requires compatible communication protocols and compatible data representations for the messages, including compatible generic message definitions for TMN management functions and ASTN control functions.

The DCN is responsible for providing compatible communication at the network layer (Layer 3), data-link layer (Layer 2), and physical layer (Layer 1).

Consideration of interfaces should be given to compatibility with the most efficient data transport facilities available to each individual network element (e.g., leased circuits, circuit-switched connections, packet-switched connections, Signalling System No. 7, Embedded Communication Channels of the SDH, OTN, and ISDN access network D- and B-channels).

This Recommendation specifies the lower three layers for data communication and therefore any interworking between protocols within the lower three layers. Such interworking is provided by the Data Communication Function (DCF). Examples of such interworking are illustrated in Figure 6-2. Note that such interworking does not terminate the Layer 3 protocols. One example is interworking between different physical layers via a common Layer 2 protocol (e.g., bridging MAC frames from a LAN interface to an ECC). Another example is interworking between different data-link layer protocols via a common layer 3 protocol (e.g., routing IP packets from a LAN interface to an ECC). The third example illustrated in Figure 6-2, shows interworking between different network layer protocols via a Layer 3 tunnelling function (in this example OSI is encapsulated/tunnelled over IP, however, IP over OSI encapsulation/tunnelling is also possible).

The type of information transported between the distributed components depends on the type of interfaces supported between the components. A DCN supporting distributed management communications related to the TMN needs to support the transport of information associated with the TMN interfaces defined in ITU-T Rec. M.3010. A DCN supporting distributed signalling communications related to the ASTN needs to support the transport of information associated with the ASTN interfaces defined in ITU-T Rec. G.807/Y.1302.

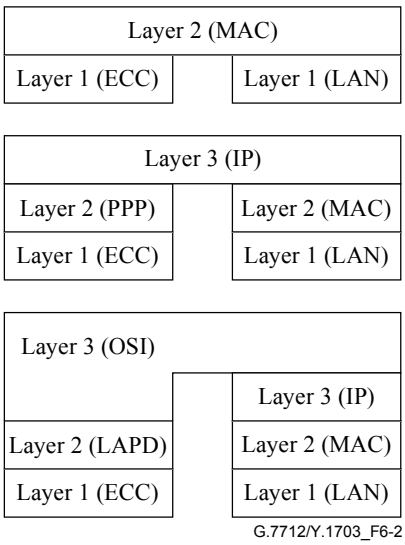


Figure 6-2/G.7712/Y.1703 – Examples of DCN Interworking

6.1 TMN application

The TMN requires a communications network, which is referred to as the Management Communication Network (MCN) to transport management messages between TMN components (e.g., NEF component and OSF component). Figure 6-3 illustrates an example relationship of the MCN and the TMN. The interfaces between the various elements (e.g., OS, WS, NE) and the MCN as illustrated in Figure 6-3 are logical and can be supported over a single physical MCN interface or multiple MCN interfaces.

Figure 6-4 illustrates an example of a physical implementation of a MCN supporting distributed management communications. Depending on the choice of implementation of the MCN, the physical elements may support any combination of ECC interfaces, LAN interfaces, and WAN interfaces. Figure 6-4 also illustrates the types of management plane functional blocks that can be supported in various physical elements. Refer to ITU-T Recs M.3010 and M.3013 for detailed specifications regarding these management functional blocks. A Data Communication Function (DCF) is part of each physical element and provides data communication functions.

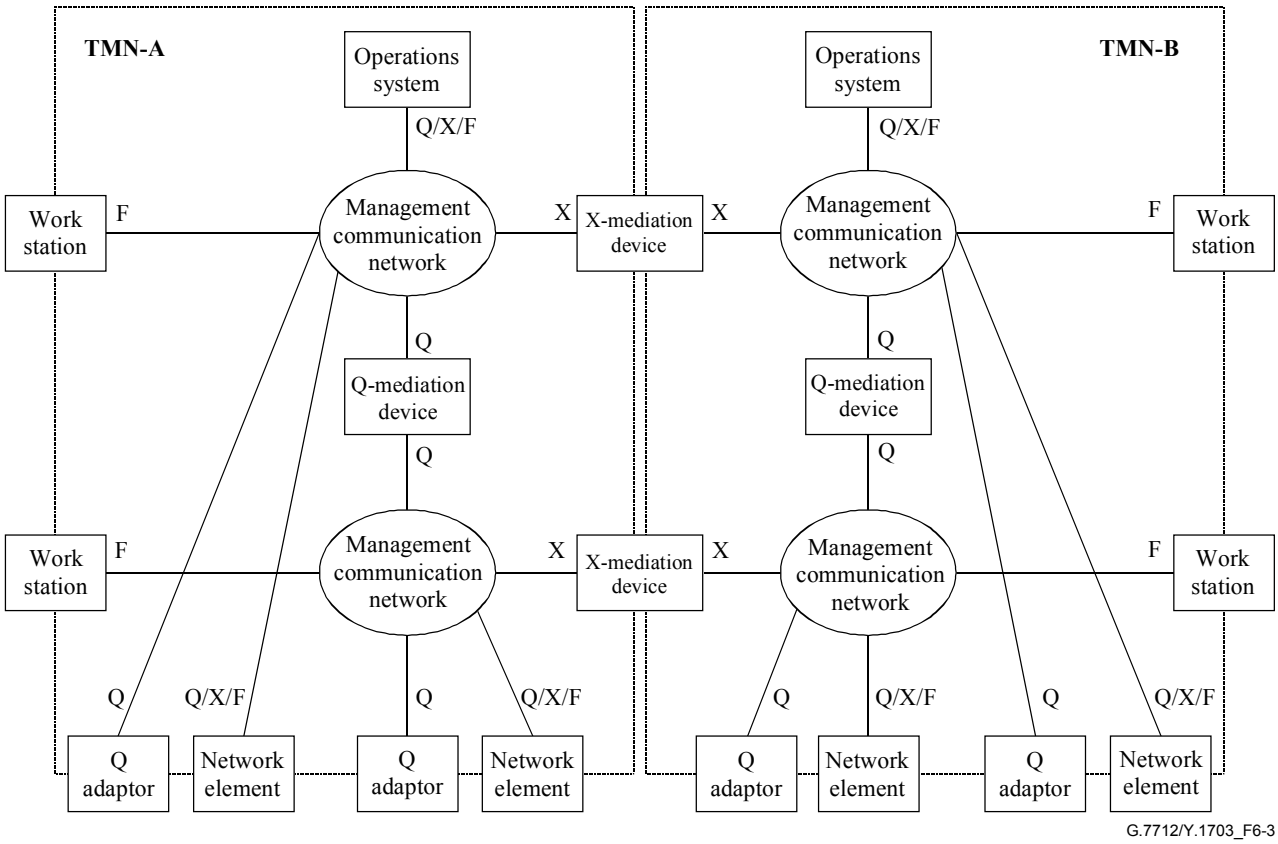


Figure 6-3/G.7712/Y.1703 – Example relationship of TMN interfaces and MCN

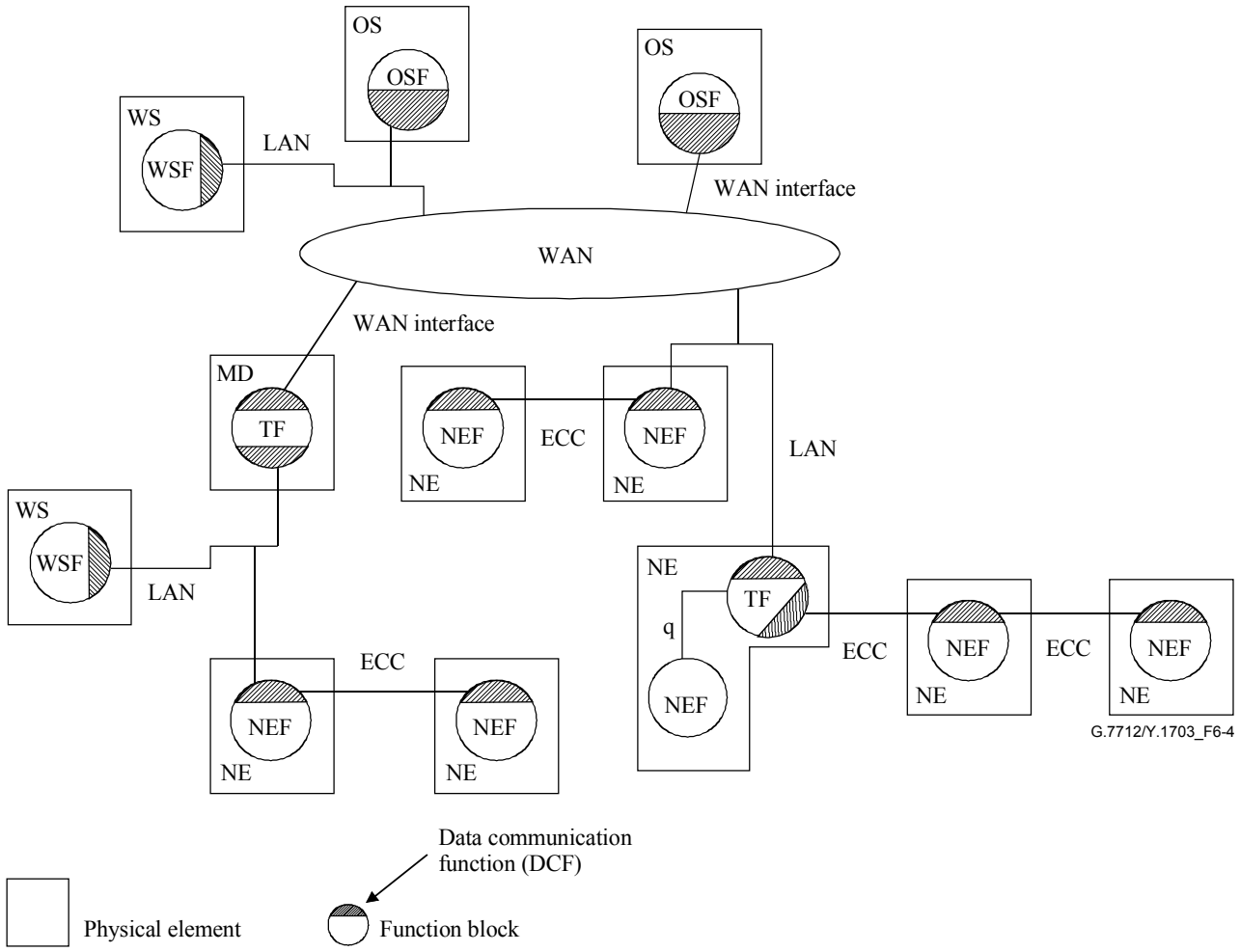


Figure 6-4/G.7712/Y.1703 – Example of physical implementation of MCN supporting TMN

6.1.1 X management subnetwork architecture

In Figure 6-5, a number of points should be noted concerning the architecture of a X Management Subnetwork (xMS):

– *Multiple NEs at a single site*

Multiple addressable SDH or OTN NEs may appear at a given site. For example, in Figure 6-5, NE_E and NE_G may be collocated at a single equipment site.

– *SDH/OTN NEs and their communication functions*

The message communication function of an SDH or OTN NE terminates (in the sense of the lower protocol layers) routes, or otherwise processes messages on the ECC or connected via an external interface.

i) All NEs are required to terminate the ECC. This means that each NE must be able to perform the functions of an OSI end system or IP host.

ii) NEs may also be required to route ECC messages between ports according to routing control information held in the NE. This means that an NE may also be required to perform the functions of an OSI intermediate system or IP router.

– *SDH/OTN inter-site communications*

The inter-site or inter-office communications link between SDH/OTN NEs may be formed from the SDH/OTN ECCs.

– *SDH/OTN intra-site communications*

Within a particular site, SDH/OTN NEs may communicate via an intra-site ECC or via a Local Communication Network (LCN). Figure 6-5 illustrates both instances of this interface.

NOTE – A standardized LCN for communicating between collocated network elements has been proposed as an alternative to the use of an ECC. The LCN would potentially be used as a general site communication network serving SDH, OTN, and non-SDH/OTN NEs (NNEs).

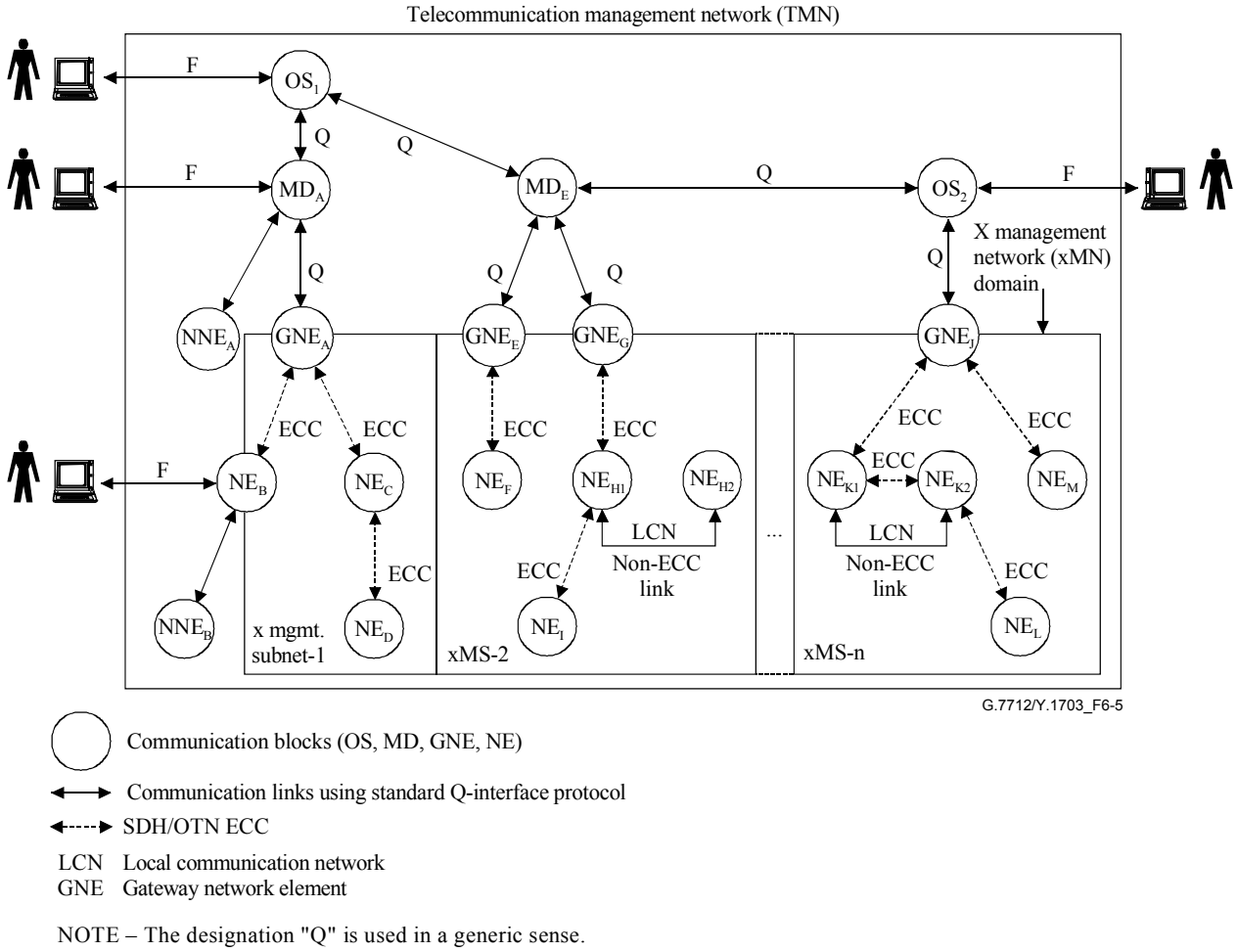


Figure 6-5/G.7712/Y.1703 – TMN, management network and management subnetwork model

6.1.1.1 Topology for management subnetwork

Figure 6-6 illustrates example MCN topologies such as linear, ring, mesh, and star utilizing ECCs and/or Local Communication Networks (LCN) (e.g., Ethernet LAN) as the physical links interconnecting the Network Elements. Figure 6-7 illustrates how a Management Subnetwork could be supported on each topology. Common to each topology are the dual Gateways (GNE_1 and GNE_2) which allow reliable access to the NEs within the Management Subnetwork. Another common aspect to each of the example topologies is that each topology allows multiple diverse paths between any NE within the Management Subnetwork and the Operations System (OS).

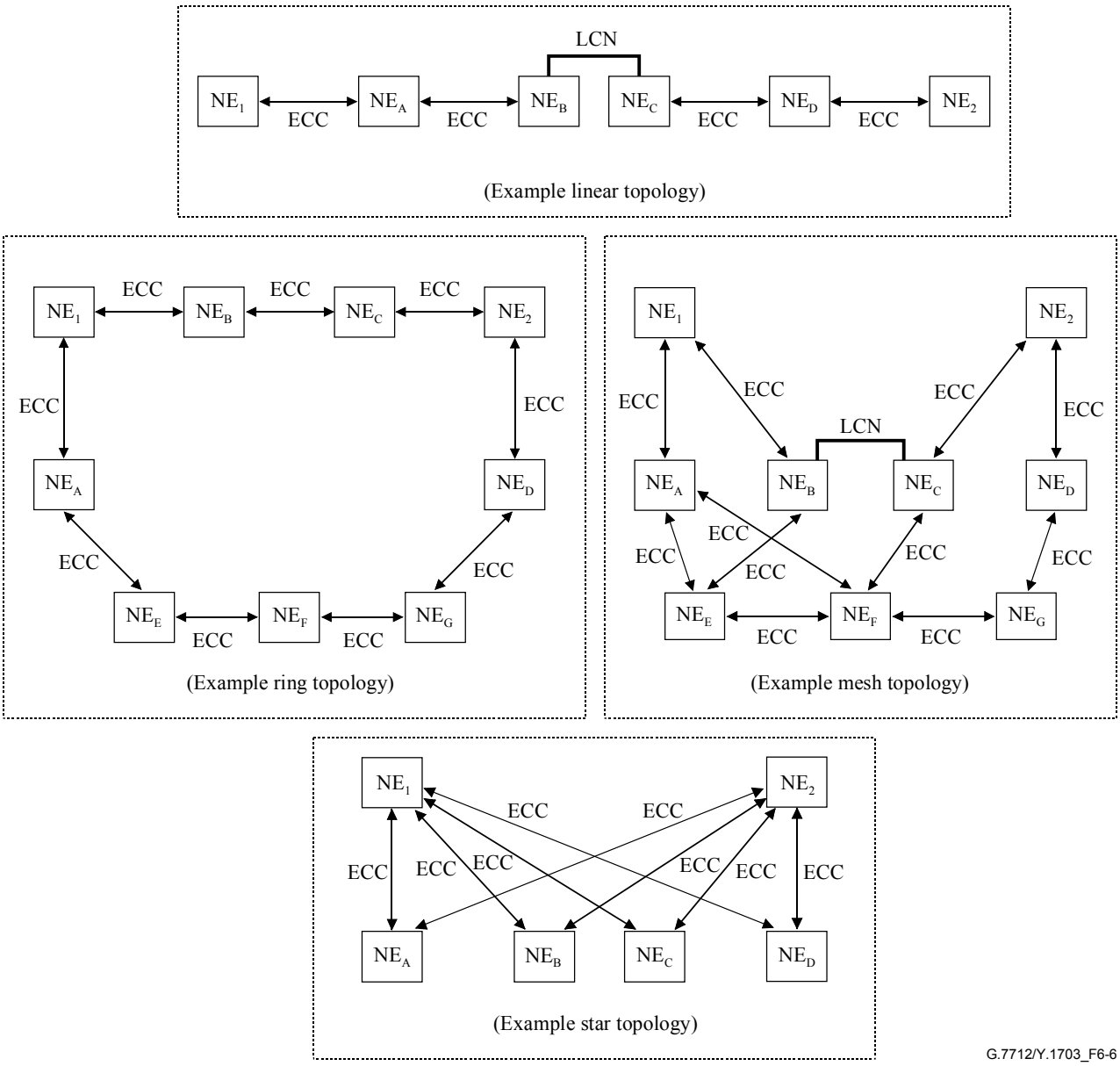


Figure 6-6/G.7712/Y.1703 – Example topologies

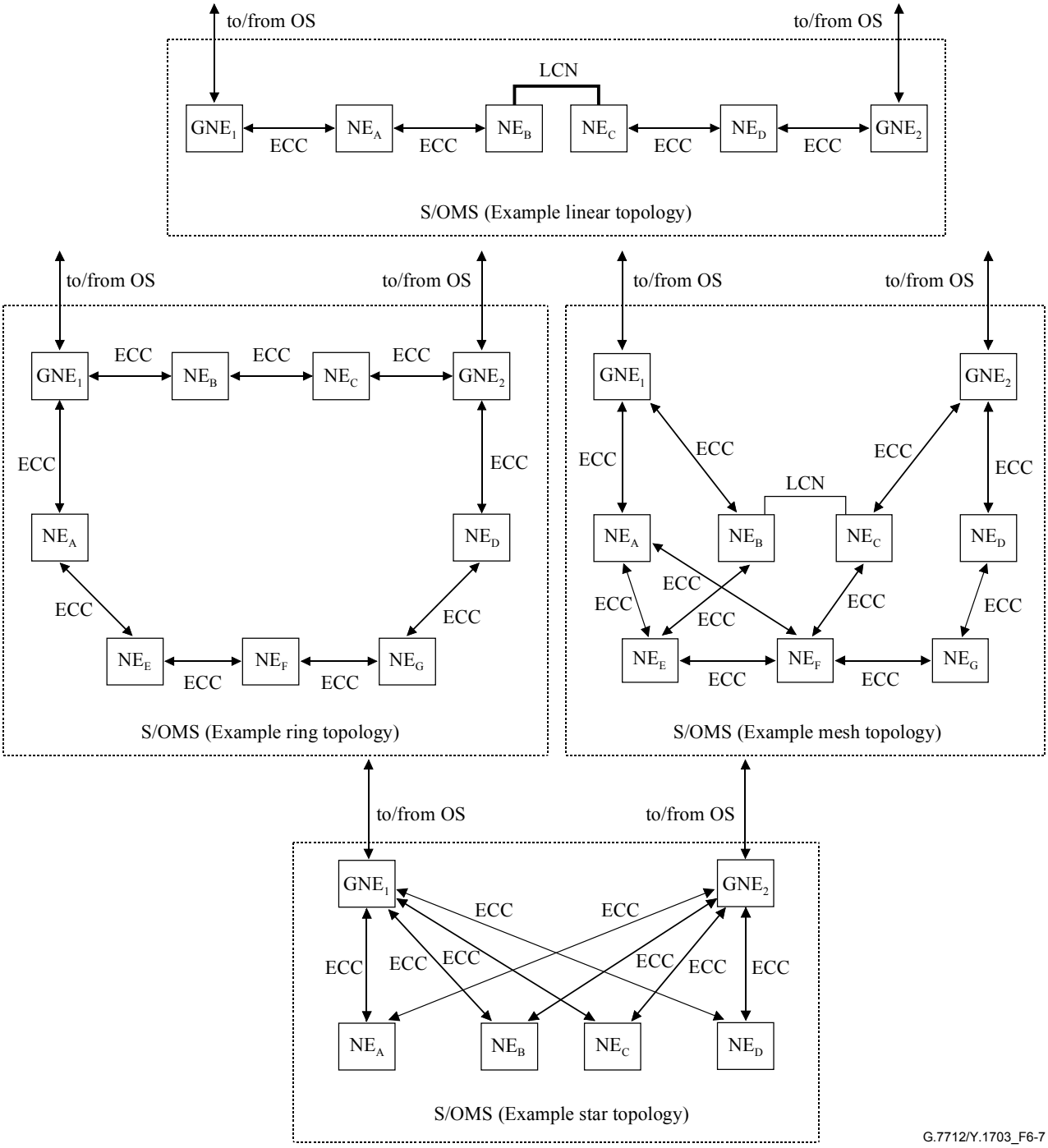


Figure 6-7/G.7712/Y.1703 – Supporting a management subnetwork on various topologies

6.1.2 Reliability of MCN

A MCN should be designed to prevent a single fault from making the transfer of critical management messages impossible.

A MCN should be designed to ensure that congestion in the MCN does not cause the blocking or excessive delay of network management messages that are intended to correct a failure or fault.

OSs and NEs that provide an emergency function may require alternate or duplicate access channels to the MCN for redundancy.

6.1.3 Security of MCN

See ITU-T Rec. M.3016 for MCN security requirements.

6.1.4 MCN data communication functions

The DCF within the TMN entities shall support End System (ES) (in OSI terms) or Host (in IP terms) functionality.

- When the DCF within the TMN entities support ECC interfaces, the following functions are required to be supported:
  - ECC Access Function (as specified in 7.1.1)
  - ECC Data-Link Termination Function (as specified in 7.1.2)
  - "Network Layer PDU into ECC Data-Link Layer" Encapsulation Function (as specified in 7.1.3)
- When the DCF within the TMN entities support Ethernet LAN interfaces, the following functions are required to be supported:
  - Ethernet LAN Physical Layer Termination Function (as specified in 7.1.4)
  - "Network Layer PDU into Ethernet Frame" Encapsulation Function (as specified in 7.1.5)

The DCF within the TMN entities may operate as an Intermediate System (IS) (in OSI terms) or as a Router (in IP terms). The DCF within TMN entities that operate as IS/Routers must be capable of routing within their Level-1 area and, therefore, must provide the functionality of a Level-1 IS/Router. Additionally, the DCF within a TMN entity may be provisioned as a Level-2 IS/Router, which provides the capability of routing from one area to another. The functionality of a Level-2 IS/Router is not needed in the DCF of all TMN entities. An example of a DCF supporting Level-2 IS/Router functionality might be the DCF within a gateway NE.

- When the DCF within the TMN entities operate as an IS/Router, the following functions are required to be supported:
  - Network Layer PDU Forwarding Function (as specified in 7.1.6)
  - Network Layer Routing Function (as specified in 7.1.10)

The DCF within a TMN entity that supports IP may be connected directly to a DCF in a neighbouring TMN entity that supports only OSI.

- When the DCF within a TMN entity that supports IP is connected directly to a DCF in a neighbouring TMN entity that supports only OSI, the following function is required to be supported in the DCF supporting IP:
  - Network Layer PDU Interworking Function (as specified in 7.1.7)

The DCF within a TMN entity may have to forward a Network Layer PDU across a network that does not support the same Network Layer type.

- When the DCF within a TMN entity must forward a Network Layer PDU across a network that does not support the same Network Layer type, the following functions are required to be supported:
  - Network Layer PDU Encapsulation Function (as specified in 7.1.8)
  - Network Layer PDU Tunnelling Function (as specified in 7.1.9)

The DCF within a TMN entity that supports IP using OSPF routing may be connected directly to a DCF in a neighbouring TMN entity that supports IP using IntISIS.

- When the DCF within a TMN entity that supports IP using OSPF routing is connected directly to a DCF in a neighbouring TMN entity that supports IP using IntISIS, the following function is required to be supported in the DCF supporting OSPF:
  - IP Routing Interworking Function (as specified in 7.1.11)

6.2 ASTN application

ASTN requires a communications network, which is referred to as the Signalling Communication Network (SCN) to transport signalling messages between ASTN components (e.g., CC components).

Figure 6-8 illustrates an example relationship of the SCN and the ASTN. The interfaces between the various elements and the SCN as illustrated in Figure 6-8 are logical and can be supported over a single physical SCN interface, or multiple SCN interfaces.

Figure 6-9 illustrates an example of a physical implementation of a SCN supporting distributed signalling communications. Depending on the choice of implementation of the SCN, the physical elements may support any combination of ECC interfaces, LAN interfaces, and WAN interfaces. Figure 6-9 also illustrates the types of control plane functional blocks that can be supported in various physical elements. Refer to ITU-T Recs G.807/Y.1302 and G.8080/Y.1304 for detailed specifications regarding these control functional blocks. A Data Communication Function (DCF) is part of each physical element and provides data communication functionality.

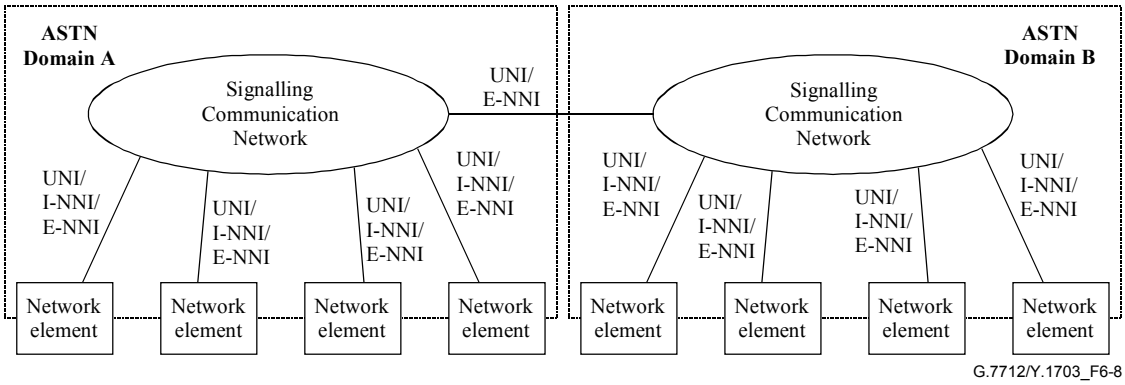


Figure 6-8/G.7712/Y.1703 – Example relationship of ASTN interfaces to SCN

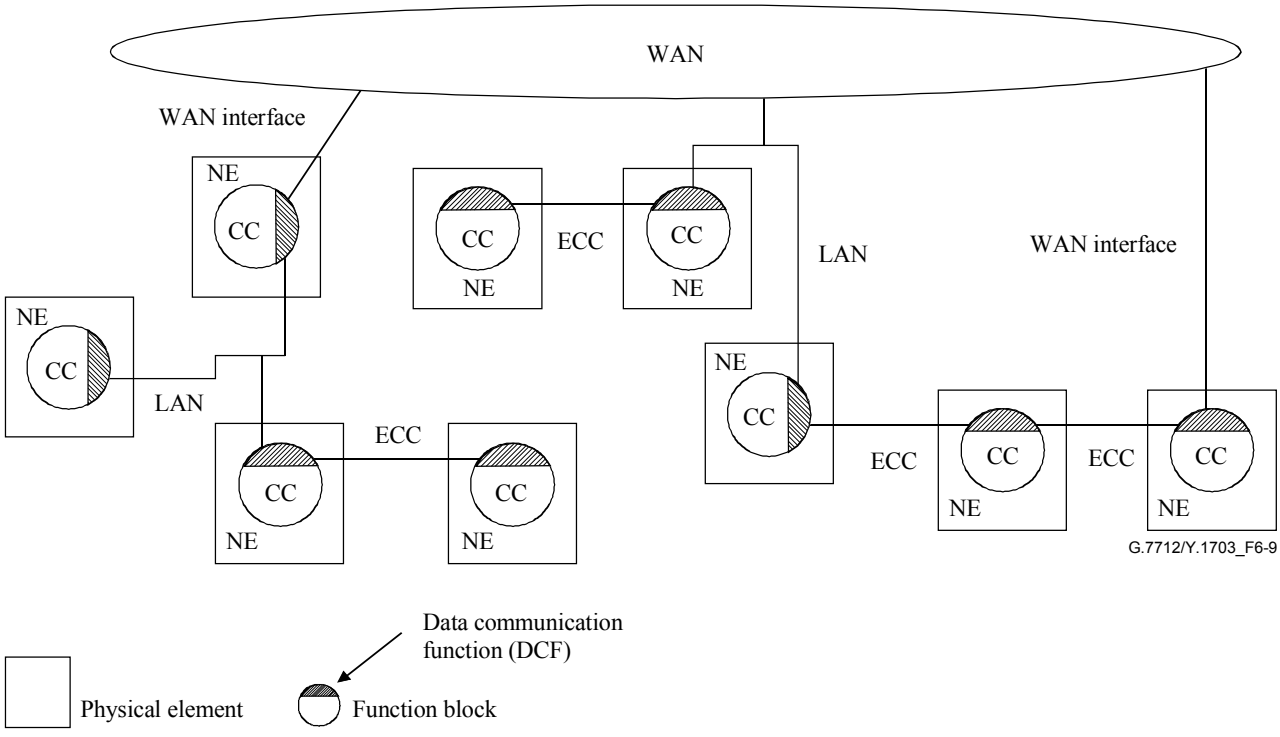


Figure 6-9/G.7712/Y.1703 – Example of physical implementation of SCN supporting ASTN

6.2.1 Topology of SCN

Figure 6-10 illustrates example topologies such as linear, ring, mesh, and star utilizing ECCs and/or Local Communication Networks (LCN) (e.g., Ethernet LAN) as the physical links interconnecting the Network Elements. Figure 6-11 illustrates how an ASTN Signalling Network could be supported on each topology. Common to each topology is that alternate diverse paths exist between the communicating entities (i.e., the ASTN capable NEs). Note that to support alternate diverse paths between communicating ASTN NEs under a linear topology, an external WAN link could be provided between the edge ASTN NEs.

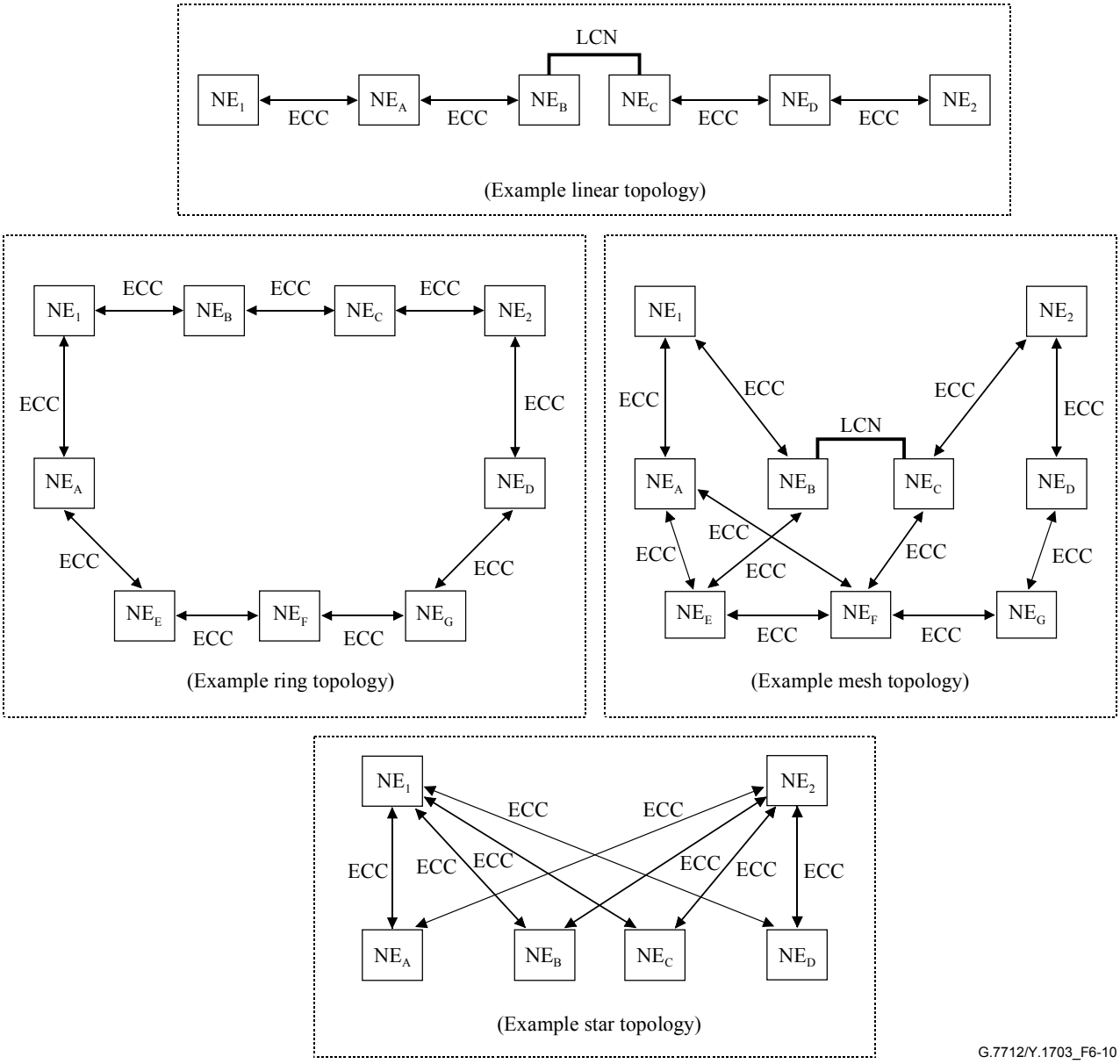


Figure 6-10/G.7712/Y.1703 – Example topologies

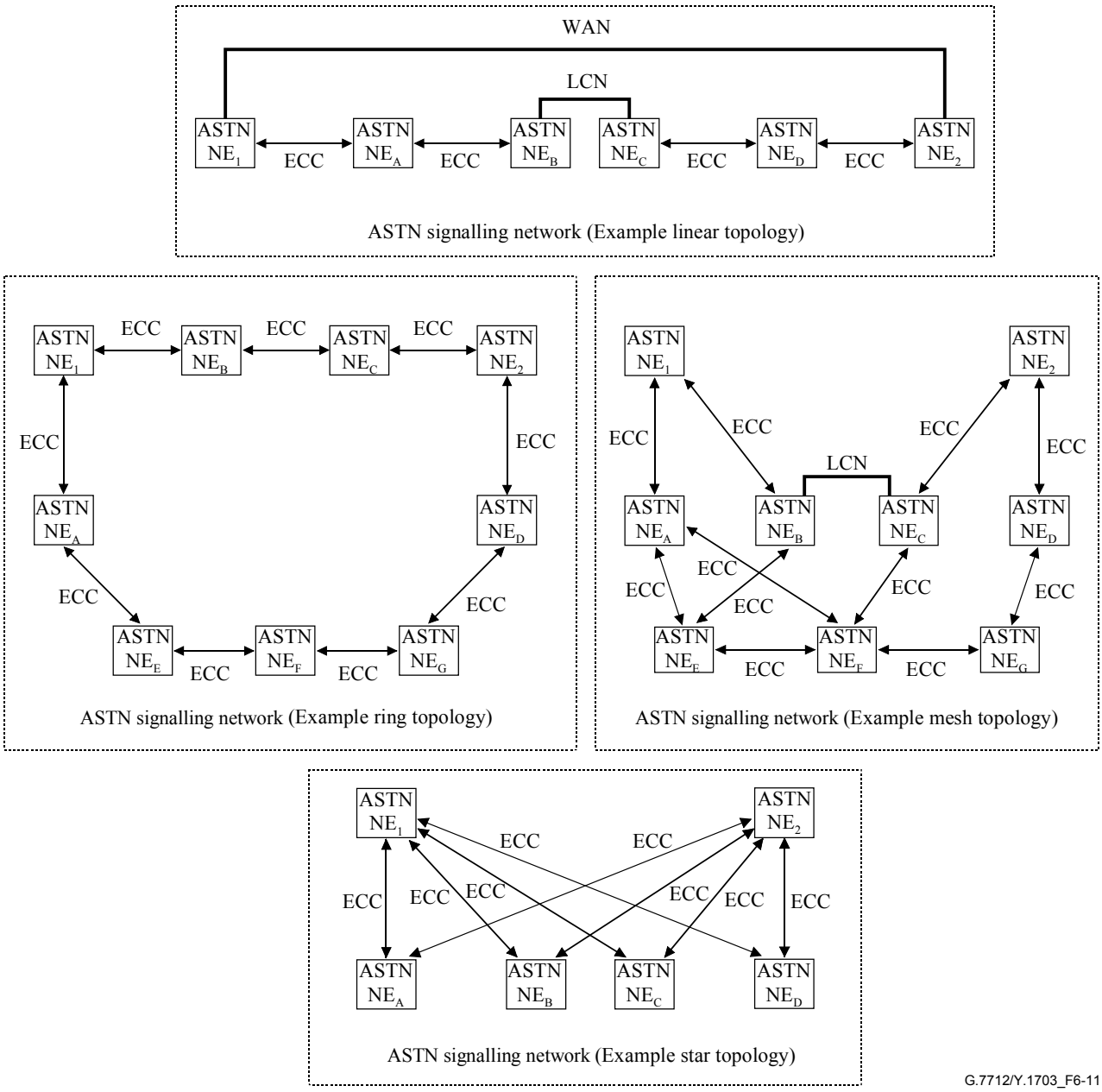


Figure 6-11/G.7712/Y.1703 – Supporting an ASTN signalling network on various topologies

Figure 6-12 illustrates how the ASTN Signalling Network could consist of three different portions; the customer-network portion, the intra-administrative domain portion, and the inter-administrative domain portion. This example shows a mesh topology utilizing ECCs, Local Communication Networks (e.g., Ethernet LAN), and Leased Lines (e.g., DS1/E1, VC-3/4) as the physical links interconnecting the ASTN NEs. The topology of the intra-administrative domain portion allows signalling to have alternate diverse paths between two communicating ASTN NEs. The topology of the inter-administrative domain portion depends on agreements between Administrative Domains A and B. This example illustrates dual access points between the Administrative Domains. The topology of the Customer-Network portion depends on agreements between the customer and service provider. This example illustrates a single access point between the customer and the network.

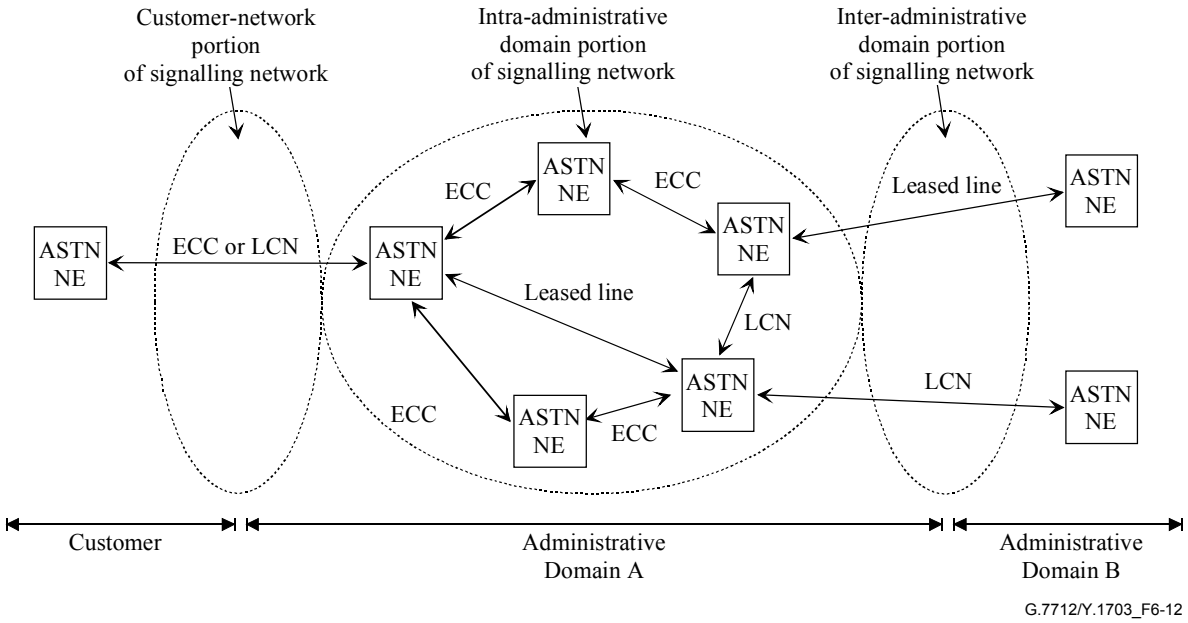


Figure 6-12/G.7712/Y.1703 – Example SCN

6.2.2 Reliability of SCN

Figure 6-13 illustrates ASTN control messages being transported over a SCN. It illustrates the following logical interfaces:

UNI User-to-Network Interface.

NNI Network-to-Network Interface.

CCI Connection Controller Interface.

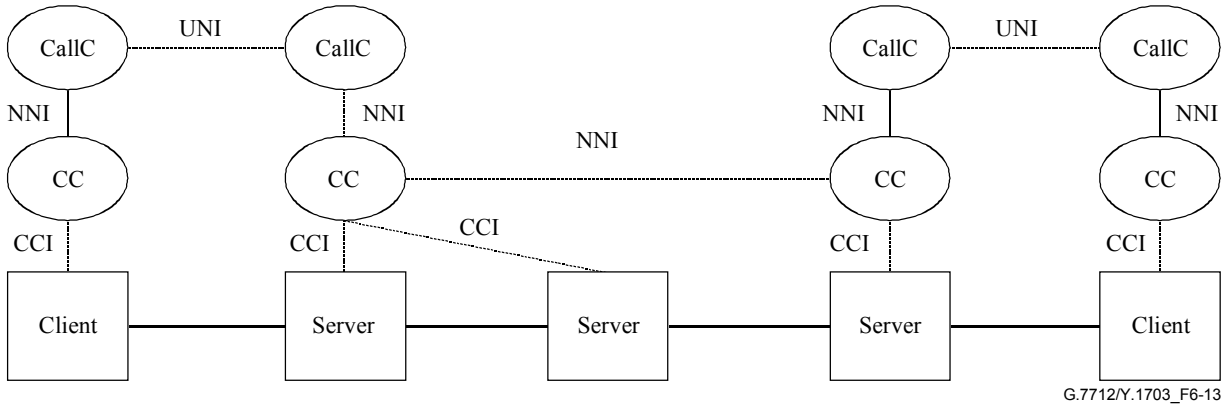


Figure 6-13/G.7712/Y.1703 – ASTN interfaces supported on SCN

In this example, the UNI, NNI, and CCI logical interfaces are carried via the SCN network. The SCN may consist of various subnetworks, where logical links in some subnetworks may share common physical routes with the transport network but such a configuration is neither required nor excluded.

It is possible for the SCN to experience an independent failure from the transport network. Such a scenario is illustrated in Figures 6-14 and 6-15. In this example, which focuses on ASTN messages transported over the SCN, an independent failure to the SCN would affect new connection set-up and connection tear-down requests.

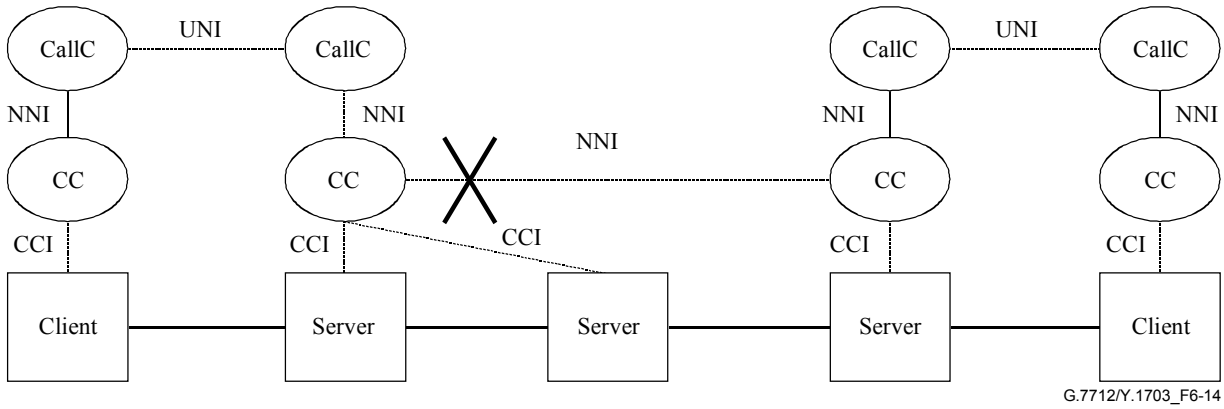


Figure 6-14/G.7712/Y.1703 – SCN failure impacting signalling interface

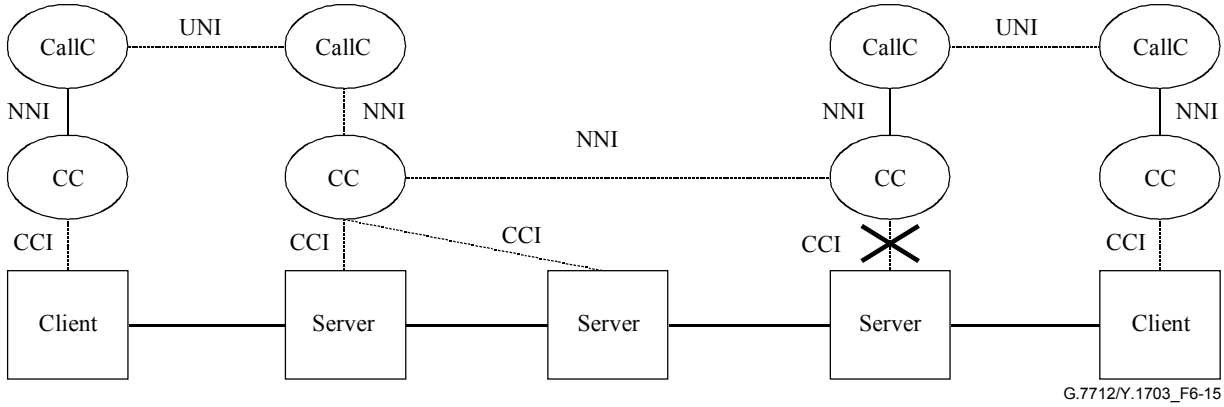


Figure 6-15/G.7712/Y.1703 – SCN failure impacting CCI interface

As indicated in Figure 6-15, it is also possible for some logical links within the SCN to share common physical routes with the transport network. In this case, it is possible for the SCN to experience a failure that is not independent from the transport network (i.e., failure interrupts both SCN traffic as well as transport traffic), as shown in Figure 6-16. In this example, which focuses on ASTN messages transported over the SCN, such a failure may impact restoration when ASTN is used to provide restoration of existing connections. It is, therefore, critical for the SCN to provide resiliency when transporting restoration messages.

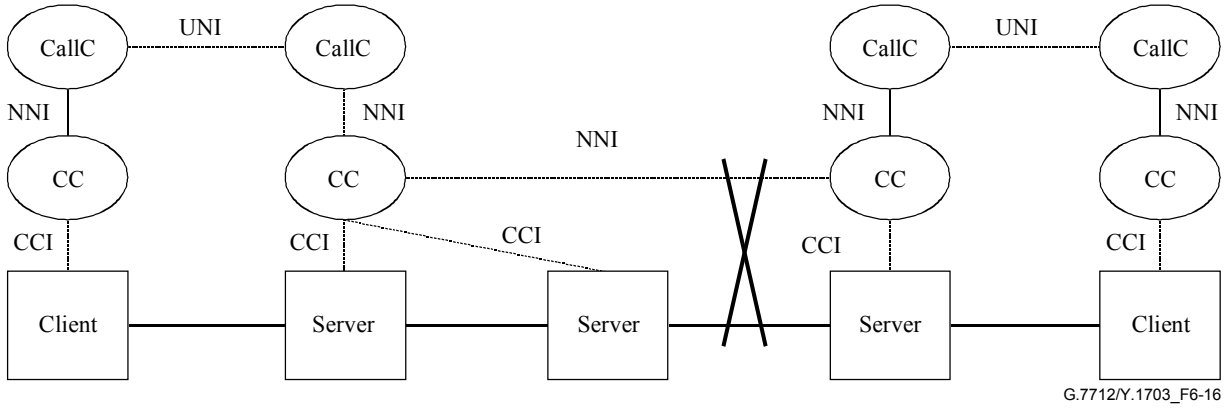


Figure 6-16/G.7712/Y.1703 – SCN failure impacting both signalling and data interfaces

If the ASTN application is only used to provide connection-setup and teardown, a connection-less SCN may be sufficient. However, if the ASTN application is also used to provide restoration, a connection-oriented SCN may be required. A connection-oriented SCN would require specification of additional functions to support connection-oriented network services.

The SCN reliability requirements are as follows:

The SCN shall support various levels of restoration depending on the reliability requirements of the communicating components for which it provides transport (i.e., restoration can be supported between those communicating components requiring highly reliable communications without requiring restoration to be supported among all communicating components).

One way of achieving reliable SCN is through use of Packet 1+1 protection for connection-oriented protocol such as MPLS as described in 6.2.4.

The SCN may provide transport for restoration messages. In such a case, the SCN shall provide restoration speeds which allow proper operation of the connections for which the restoration messages control.

6.2.3 Security of SCN

A SCN supporting ASTN messages may provide connectivity between different administrative domains. When a SCN provides connectivity between administrative boundaries, precautions must be taken such that only those messages that are allowed to pass between the two administrative domains are able to cross the interface, while other messages which are not allowed to pass between administrative domains are prevented from crossing the interface. The SCN needs to ensure that only a select set of messages, which are allowed by the administrative parties on either side of the interface, are actually able to pass across the interface.

6.2.4 SCN data communication functions

The DCF within the ASTN entities shall support End System (ES) (in OSI terms) or Host (in IP terms) functionality.

- When the DCF within the ASTN entities support ECC interfaces, the following functions are required to be supported:
  - ECC Access Function (as specified in 7.1.1)
  - ECC Data-Link Termination Function (as specified in 7.1.2)
  - "Network Layer PDU into ECC Data-Link Layer" Encapsulation Function (as specified in 7.1.3)
- When the DCF within the ASTN entities support Ethernet LAN interfaces, the following functions are required to be supported:
  - Ethernet LAN Physical Layer Termination Function (as specified in 7.1.4)
  - "Network Layer PDU into Ethernet Frame" Encapsulation Function (as specified in 7.1.5)

The DCF within the ASTN entities may operate as an Intermediate System (IS) (in OSI terms), or as a Router (in IP terms). The DCF within ASTN entities that operate as IS/Routers must be capable of routing within their Level-1 area and therefore must provide the functionality of a Level-1 IS/Router. Additionally, the DCF within an ASTN entity may be provisioned as a Level-2 IS/Router, which provides the capability of routing from one area to another. The functionality of a Level-2 IS/Router is not needed in the DCF of all ASTN entities.

- When the DCF within the ASTN entities operate as an IS/Router, the following functions are required to be supported:
  - Network Layer PDU Forwarding Function (as specified in 7.1.6)
  - Network Layer Routing Function (as specified in 7.1.10)

The DCF within a ASTN entity that supports IP may be connected directly to a DCF in a neighbouring ASTN entity that supports only OSI.

- When the DCF within an ASTN entity that supports IP is connected directly to a DCF in a neighbouring TMN entity that supports only OSI, the following function is required to be supported in the DCF supporting IP:
  - Network Layer PDU Interworking Function (as specified in 7.1.7)

The DCF within a ASTN entity may have to forward a Network Layer PDU across a network that does not support the same Network Layer type.

- When the DCF within a ASTN entity must forward a Network Layer PDU across a network that does not support the same Network Layer type, the following functions are required to be supported:
  - Network Layer PDU Encapsulation Function (as specified in 7.1.8)
  - Network Layer PDU Tunnelling Function (as specified in 7.1.9)

The DCF within a ASTN entity that supports IP using OSPF routing may be connected directly to a DCF in a neighbouring ASTN entity that supports IP using IntISIS.

- When the DCF within an ASTN entity that supports IP using OSPF routing is connected directly to a DCF in a neighbouring ASTN entity that supports IP using IntISIS, the following function is required to be supported in the DCF supporting OSPF:
  - IP Routing Interworking Function (as specified in 7.1.11)

The DCF within the ASTN entities may operate as an Label Edge Router (LER).

When the DCF within the ASTN entities operates as an LER, the following functions are required to be supported:

- If the DCF supports ECC interfaces, the "MPLS PDU into ECC Data-Link Layer" Encapsulation Function (as specified in 7.1.13).
- If the DCF supports LAN interfaces, the "MPLS PDU into Ethernet Frame" Encapsulation Function (as specified in 7.1.14).
- MPLS LSP Signalling Function (as specified in 7.1.15).
- MPLS LSP Forwarding Function (as specified in 7.1.16).
- MPLS LSP Path Computation Function (as specified in 7.1.17).
- "Network Layer PDU into MPLS" Encapsulation Function (as specified in 7.1.18).

The DCF within the ASTN entities may operate as a Label Switch Router (LSR).

When the DCF within the ASTN entities operate as an LSR, the following functions are required to be supported:

- If the DCF supports ECC interfaces, the "MPLS PDU into ECC Data-Link Layer" Encapsulation Function (as specified in 7.1.13).
- If the DCF supports LAN interfaces, the "MPLS PDU into Ethernet Frame" Encapsulation Function (as specified in 7.1.14).
- MPLS LSP Signalling Function (as specified in 7.1.15).
- MPLS LSP Forwarding Function (as specified in 7.1.16).

The DCF within the ASTN entities may provide packet 1+1 protection capability.

The minimum requirements to provide packet 1+1 protection service are as follows:

- There is no additional capability required on the interior nodes of the network;
- The network should support the establishment of diversely routed connections.
- *Ingress Node*
  - Must be able to associate the two connections that are used to provide packet level 1+1 protection between two end nodes;
  - Must support the carrying of an identifier in the packet which will be used to identify duplicate copies of a packet at the egress node;
  - Must be able to dual-feed each packet on these two mated connections.
- *Egress Node*
  - Must be able to associate the two connections that are used to provide packet level 1+1 protection between two end nodes;
  - Must be able to identify the duplicate copies of a dual-fed packet using the identifier;
  - Must be able to select and forward one and only one copy of a packet.

The mechanism to associate the two diverse connections as well as the format and location of the sequence identifier shall be as described in 7.1.19.

6.3 Other applications requiring communication networks

Besides TMN and ASTN applications, other applications such as voice communications (e.g., orderwire), software downloads and operator specific communications require a communication network to provide transport of information between components.

6.4 Separation of various applications

Depending on the network design, network size, link capacity, security requirements and performance requirements, various levels of separation between the multiple applications (e.g., TMN, ASTN) are possible. The level of separation that is provided is a choice that is made among operators and vendors when designing the network. The following are examples of various levels of separation.

Option A: The DCN can be designed such that the MCN, SCN, and other applications (e.g., operator-specific communications) are supported on the same layer 3 network (e.g., share the same IP network).

Option B: The DCN can be designed such that the MCN, SCN, and other applications (e.g., operator-specific communications) are supported on separate layer 3 networks, however, they may share some of the same physical links.

Option C: The DCN can be designed such that the MCN, SCN, and other applications (e.g., operator-specific communications) are supported on separate physical networks (i.e., separate layer 3 networks that do not share any of the same physical links).

7 DCN functional architecture and requirements

The DCN architecture requirements in this clause apply to IP-only Domains, OSI-only Domains, and mixed IP+OSI domains. The DCN architecture requirements are technology independent. Technology-specific Recommendations such as ITU-T Rec. G.784 for SDH and ITU-T Rec. G.874 for OTN will specify which requirements are applicable for that particular technology.

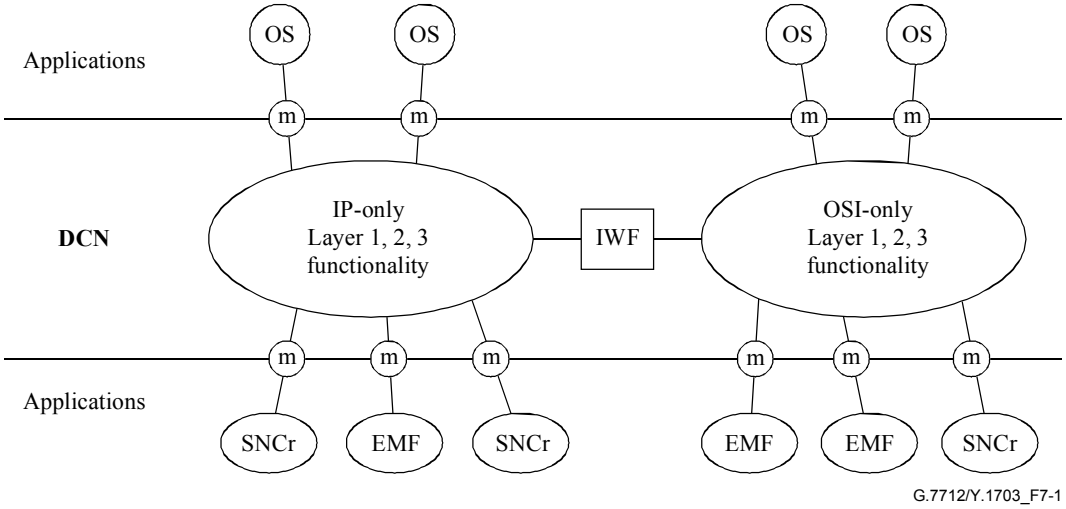
The DCN is aware of Layer 1, Layer 2, and Layer 3 protocols and is transparent to upper-layer protocols used by the applications for which it transports.

A DCN may be designed such that only IP is supported. A DCN supporting only IP may consist of various subnetworks using different physical and data link layer protocols, however, all subnetworks will support IP as the network layer protocol.

However, since embedded DCN networks support OSI, some DCNs may consist of parts that support IP-only, parts that support OSI-only, and parts that support both IP and OSI.

Those parts of the DCN supporting IP (i.e., either those parts supporting only IP or those parts supporting IP and OSI) may consist of DCFs that support IP-only (i.e., a single stack IP-only DCFs) and/or DCFs supporting IP and OSI (e.g., a dual-stack DCF which is capable of routing both IP and OSI packets). Those parts of the DCN supporting only OSI, would consist of DCFs that support OSI-only (i.e., a single stack OSI-only DCF).

Figure 7-1 illustrates the functional architecture of the DCN. As discussed above, the DCN may be composed of parts that only support IP, parts that only support OSI, and parts that support both IP and OSI. An Interworking Function (IWF) between those parts of the DCN supporting IP-only, OSI-only, and IP and OSI, and mapping functions which map applications to the IP layer are also specified. To provide such transport, the DCN supports Layer 1 (physical), Layer 2 (data-link), and Layer 3 (network) functionality. The architecture requirements for those parts of the DCN supporting IP only, OSI only as well as the requirements for interworking between those parts of the DCN supporting IP-only, OSI-only, and IP and OSI are specified. The cloud in Figure 7-1, representing the IP-only part of the DCN, is an abstract view of the DCN and therefore may also apply to a single IP NE interconnected to OSI NEs via an IWF.



IWF Interworking Function
SNCr Subnetwork Connection Controller
EMF Equipment Management Function
OS Operations System
m mapping between Application and DCN

Figure 7-1/G.7712/Y.1703 – Functional architecture of DCN

7.1 Specification of data communication functions

This clause provides specifications for various data communication functions related to ECC interfaces, Ethernet LAN interfaces, and network layer capabilities.

7.1.1 ECC access function

An ECC Access Function provides access to the ECC bit stream. This function is defined in technology specific equipment Recommendations (e.g., ITU-T Recs G.783 and G.798). The bit rates and definitions of the various ECCs (e.g., DCC, GCC, and COMMS OH in OSC) is provided in the technology specific Recommendations (e.g., ITU-T Recs G.784 and G.874).

7.1.2 ECC data-link layer termination function

An ECC Data-Link Layer Termination Function provides the common data-link layer processing regardless of the network layer PDU encapsulated within the Data-Link Layer Frame. The mapping of the Data-Link Layer Frame into the ECC is also provided by this function. This function is specified in the technology specific Recommendations. However, the specification for the SDH ECC Data-Link Layer Termination Function is provided below.

7.1.2.1 SDH ECC data-link layer termination function

7.1.2.1.1 Mapping the SDH data-link layer frame into the ECC

The HDLC framed signal is a serial bit stream containing stuffed frames surrounded by one or more flag sequences. The HDLC framed signal format is defined in ITU-T Rec. Q.921 for LAPD, and RFC 1662 for PPP in HDLC framing. A HDLC frame consists of N octets as presented in Figure 7-2. The HDLC frame is transmitted right to left and top to bottom. A 0 bit is inserted after all sequences of five consecutive 1 bits within the HDLC frame content (octets 2 to N-1) ensuring that a flag or abort sequence is not simulated within a frame.

The mapping of the HDLC framed signal into the DCC channel is bit-synchronous (rather than octet-synchronous) since the stuffed HDLC frame does not necessarily contain an integer number of octets as a consequence of the 0 insertion process. Therefore, there is no direct mapping of a stuffed HDLC frame into bytes within a DCC channel. The HDLC signal generator derives its timing from the ServerLayer/DCC_A function (i.e., the DCC_CI_CK signal) for SDH. The following ServerLayer/DCC_A functions are defined in ITU-T Rec. G.783; MSn/DCC_A function, MS256/DCC_A function, and RSn/DCC_A function.

The HDLC frame signal is a serial bit stream and will be inserted into the DCC channel such that the bits will be transmitted on the STM-N in the same order that they were received from the HDLC frame signal generator.

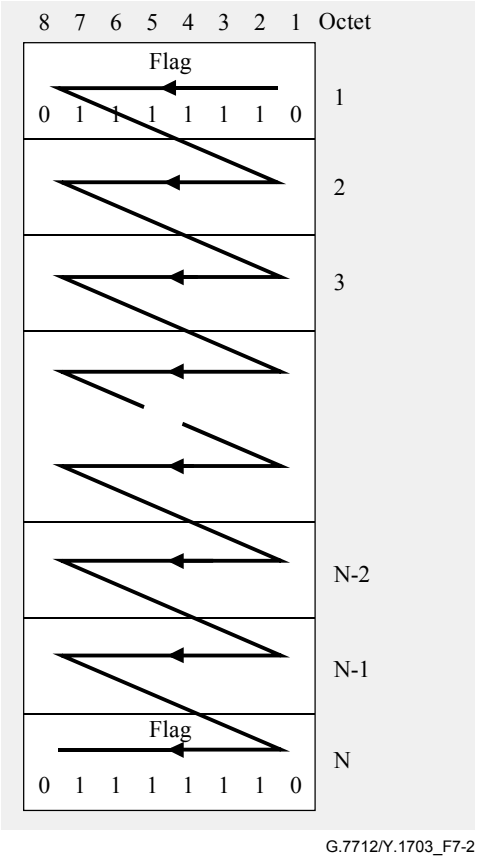


Figure 7-2/G.7712/Y.1703 – HDLC frame format

7.1.2.1.2 SDH ECC data-link layer protocol specification

The three types of interfaces identified are; IP-only interfaces, OSI-only interfaces, and Dual interfaces (Dual interfaces are interfaces that can carry both IP and OSI packets). When carrying only IP over the DCC, PPPinHDLC framing shall be used as the data-link layer protocol. Since Dual Interfaces can carry both IP and OSI, it is possible for a Dual Interface to be connected to either an IP-only interface, an OSI-only interface, or another Dual interface. OSI-only interfaces exist in networks today, and the data-link protocol used on such interfaces is LAPD as defined in ITU-T Rec. G.784. To allow Dual Interfaces to connect to either an IP-only interface or an OSI-only interface, the data-link layer protocol supported on a Dual Interface must be configurable to support either PPPinHDLC or LAPD. An exception is allowed for embedded SDH NEs supporting LAPD in hardware that are upgraded to support Dual Interfaces. To limit the amount of hardware upgrades, it is allowed for upgraded SDH NEs to support only LAPD.

7.1.2.1.2.1 IP-only interface

IP-only interfaces are illustrated in Figure 7-3.

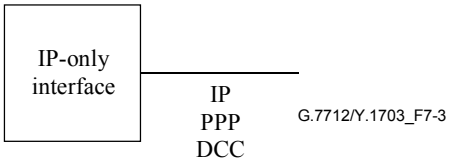


Figure 7-3/G.7712/Y.1703 – IP-only interface

IP-only interfaces shall use PPP as per RFC 1661.

7.1.2.1.2.2 OSI-only interface

OSI-only interfaces are illustrated in Figure 7-4.

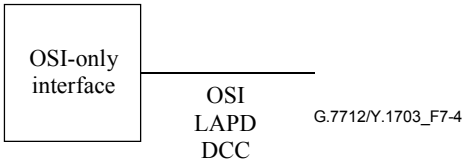


Figure 7-4/G.7712/Y.1703 – OSI-only interface

OSI-only interfaces shall use LAPD as per ITU-T Rec. G.784.

7.1.2.1.2.3 Dual interface (IP+OSI)

Dual interfaces (Dual interfaces are interfaces that can carry OSI and IP packets) can be connected to IP-only interfaces, OSI-only interfaces, or other Dual interfaces. To allow Dual interfaces to be connected to other IP-only interfaces or other OSI-only interfaces, the data-link protocol on the Dual interface must be configurable to switch between PPP in HDLC framing (as per RFC 1662) and LAPD (as per ITU-T Rec. G.784) as illustrated in Figure 7-5. Note that embedded SDH NEs supporting LAPD in hardware that are upgraded to support IP are not required to support PPP in HDLC framing on its dual interfaces. Therefore its dual interfaces are only required to support LAPD.

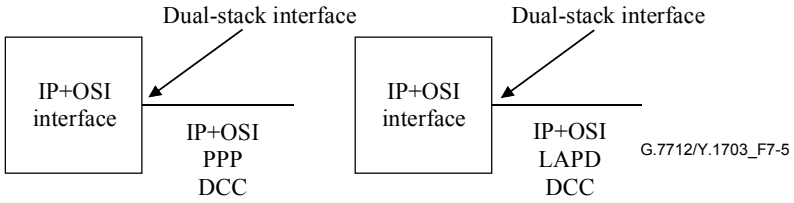


Figure 7-5/G.7712/Y.1703 – Dual interface

Dual interfaces supporting PPP shall use PPP as per RFC 1661.

Dual interfaces supporting LAPD shall use LAPD as per ITU-T Rec. G.784.

7.1.3 "Network layer PDU into ECC data-link frame" encapsulation function

A "Network Layer PDU into ECC Data-Link Frame" Encapsulation Function encapsulates and unencapsulates the Network Layer PDU into the Data-Link Frame. This function also processes the protocol identifier. This function is defined in the technology specific Recommendations. However, the specification for the "Network Layer PDU into SDH ECC Data-Link Frame" Encapsulation Function is provided below.

7.1.3.1 "Network layer PDU into SDH ECC data-link frame" encapsulation function

The specification of the "Network Layer PDU into SDH ECC Data-Link Frame" Encapsulation Function for IP-only interfaces, OSI-only interfaces, and Dual Interfaces is provided below.

7.1.3.1.1 IP-only interface

IP-only interfaces must use only PPPinHDLCframing/DCC as per RFC 1662.

An IP-only interface is defined as follows:

The Transmit End

– Shall put IS-IS packets directly into PPP Information Field as per RFC 1661 with the OSI protocol value as per RFC 1377 into the PPP Protocol Field.

– Shall put IPv4 packets directly into PPP Information Field as per RFC 1661 with the IPv4 protocol value as per RFC 1332 into the PPP Protocol Field.

– Shall put IPv6 packets directly into PPP Information Field as per RFC 1661 with the IPv6 protocol value as per RFC 2472 into the PPP Protocol Field.

The Receive End

– An IS-IS packet is identified if the PPP Protocol Field has the OSI protocol value as per RFC 1377, and if the packet has the NLPID for IS-IS as specified in ITU-T Rec. X.263 | ISO/IEC 9577.

– An IPv4 packet is identified if the PPP Protocol Field has the IPv4 protocol value as per RFC 1332.

– An IPv6 packet is identified if the PPP Protocol Field has the IPv6 protocol value as per RFC 2472.

7.1.3.1.2 OSI-only interface

OSI-only interfaces must use only LAPD/DCC as per ITU-T Rec. G.784.

An OSI-only interface is defined as follows:

The Transmit End

– Shall put CLNP, IS-IS, and ES-IS packets directly into LAPD payload as per ITU-T Rec. G.784.

The Receive End

– Shall inspect the protocol identifier located in the first octet of the LAPD payload. The value of this identifier is consistent with the values assigned in ITU-T Rec. X.263 | ISO/IEC 9577. If the PDU received is for a protocol not supported by the receiver, then the PDU shall be discarded.

7.1.3.1.3 Dual (IP+OSI) interface

A Dual interface supporting PPP as the data-link protocol is defined as follows:

The Transmit End

– Shall put CLNP, IS-IS, and ES-IS packets directly into PPP Information Field as per RFC 1661 with the OSI protocol value as per RFC 1377 into the PPP Protocol Field.

– Shall put IPv4 packets directly into PPP Information Field as per RFC 1661 with the IPv4 protocol value as per RFC 1332 into the PPP Protocol Field.

– Shall put IPv6 packets directly into PPP Information Field as per RFC 1661 with the IPv6 protocol value as per RFC 2472 into the PPP Protocol Field.

The Receive End

– An OSI packet is identified if the PPP Protocol Field has the OSI protocol value as per RFC 1377.

– An IPv4 packet is identified if the PPP Protocol Field has the IPv4 protocol value as per RFC 1332.

– An IPv6 packet is identified if the PPP Protocol Field has the IPv6 protocol value as per RFC 2472.

A Dual interface supporting LAPD as the data-link protocol is defined as follows:

The Transmit End

– Shall put CLNP, IS-IS, and ES-IS packets directly into LAPD payload as per ITU-T Rec. G.784.

– Shall put IP packets directly into LAPD payload, with a one-octet protocol identifier prepended. This identifier will be consistent with the ITU-T Rec. X.263 | ISO/IEC 9577 assigned values for IPv4 and IPv6.

The Receive End

– Shall inspect the protocol identifier located in the first octet of the LAPD payload. The value of this identifier is consistent with the values assigned in ITU-T Rec. X.263 | ISO/IEC 9577. If the PDU received is for a protocol not supported by the receiver, then the PDU shall be discarded.

7.1.4 Ethernet LAN physical termination function

An Ethernet LAN Physical Termination Function terminates the physical Ethernet interface.

One or more of the following rates shall be supported: 1 Mbit/s, 10 Mbit/s, 100 Mbit/s.

Access to terminated ECC channels is allowed by Network Elements supporting Ethernet LAN interfaces. Not all network elements supporting ECC channels need to support Ethernet LAN ports, as long as there is an ECC path from a Network Element terminating the ECC channel and another Network Element providing Ethernet LAN ports.

7.1.5 "Network layer PDU into Ethernet frame" encapsulation function

This function encapsulates and unencapsulates a Network Layer PDU into an 802.3 or Ethernet (version 2) frame.

It shall encapsulate Network Layer PDUs into 802.3 or Ethernet (version 2) frames according to the following rules.

- It shall encapsulate and unencapsulate CLNP, IS-IS, and ES-IS PDUs into 802.3 frames as per ITU-T Rec. Q.811.
- It shall encapsulate and unencapsulate IP packets into Ethernet (version 2) frames as per RFC 894.
- IP addresses shall be mapped to Ethernet MAC addresses utilizing the Address Resolution Protocol in RFC 826.

It shall determine the received frame type (802.3 or Ethernet version 2) as per Section 2.3.3 in RFC 1122.

7.1.6 Network layer PDU forwarding function

The Network Layer PDU Forwarding Function forwards network layer packets.

If this function forwards CLNP packets, it shall forward CLNP packets as per ITU-T Rec. Q.811.

If this function forwards IPv4 packets, it shall forward IPv4 packets as per RFC 791.

If this function forwards IPv6 packets, it shall forward IPv6 packets as per RFC 2460.

The preferred addressing format is IPv6. The IP routing protocol should be able to deal with IPv6 and IPv4 addressing.

7.1.7 Network layer PDU interworking function

The Network Layer PDU Interworking Function ensures that neighbouring DCF functions, running different network layer protocols can communicate. The DCF supporting IP is required to support OSI to allow communication to the neighbouring DCF supporting only OSI.

7.1.8 Network layer PDU encapsulation function

The Network Layer PDU Encapsulation Function encapsulates and unencapsulates one network layer PDU into another network layer PDU.

CLNP packets shall be encapsulated over IP using Generic Routing Encapsulation (GRE), as specified in RFC 2784, as payload in an IP packet using an IP protocol number of 47 (decimal) and with the *Don't Fragment* (DF) bit not set. As per RFC 2784, the GRE shall contain an Ethertype to indicate what network layer protocol is being encapsulated. The industry standard for OSI Ethertype, which is 00FE (hex) shall be used.

IP packets shall be encapsulated over CLNS using GRE, as specified in RFC 2784, as the data payload of a CLNP Data Type PDU as specified in ISO/IEC 8473-1, using an NSAP selector value of 47 (decimal) and with the SP (segmentation permitted) flag set. Further information is available in RFC 3147.

IP packets shall be encapsulated over IP using GRE, as specified in RFC 2784, as payload in an IP packet using an IP protocol number of 47 (decimal) and with the *Don't Fragment* (DF) bit not set.

As an option, the Network Layer PDU Encapsulation function may forward PDUs across incompatible nodes via the automatic encapsulation procedure described in Annex B. Note that a DCF supporting the automatic encapsulation procedure described in Annex B is compatible with and can be deployed in the same area as a DCF that does not support the automatic encapsulation procedure.

7.1.9 Network layer tunnelling function

The Network Layer PDU Tunnelling Function provides a static tunnel between two DCFs supporting the same network layer PDU. For a tunnel with a configured MTU size, any IP packet that cannot be forwarded over the tunnel because it is larger than the MTU size, and has its DF bit set, should be discarded, and an ICMP unreachable error message (in particular the "fragmentation needed and DF set" code) should be sent back to the originator of the packet.

7.1.10 Network layer routing function

The Network Layer Routing Function routes network layer packets.

A DCF supporting OSI routing shall support IS-IS as per ISO/IEC 10589.

A DCF supporting IP routing shall support Integrated IS-IS (see 7.1.10.1 for Integrated IS-IS requirements) and may also support OSPF as well as other IP routing protocols.

7.1.10.1 Integrated IS-IS requirements

A DCF supporting Integrated IS-IS shall support RFC 1195.

A DCF supporting Integrated IS-IS shall support Three-way Handshaking on all point-to-point links (see Annex A for Three-way Handshaking requirements). Three-way handshaking modifies the adjacency creation and maintenance behaviour specified in ISO/IEC 10589.

7.1.10.1.1 Network-layer protocol aware adjacency creation

The DCF shall include a "protocols supported" TLV in all IIH and ISH PDUs on all interfaces, and in all LSPs with LSP number 0, as per RFC 1195.

On receipt of an IS-IS ISH or IIH PDU, the DCF shall inspect the PDU to see if it contains a "protocols supported" TLV. This shall take place on all interfaces, whether LAN, DCC or other links. If an ISH or IIH PDU does not contain a "protocols supported" TLV, then it shall be treated as if it contains a "protocols supported" TLV containing only the NLPID for CLNP.

The DCF shall compare the NLPIDs listed in the "protocols supported" TLV (assuming only CLNP if none is present) with the network layer protocols that the DCF is itself capable of forwarding.

If no adjacency exists with the neighbour that sent the ISH or IIH, and if the DCF is not capable of forwarding any of the network layer protocols listed in the "protocols supported" TLV of the ISH or IIH received from the neighbour, then the DCF shall not form an adjacency with that neighbour.

If an adjacency does exist with the neighbour that sent the ISH or IIH, and if the DCF is not capable of forwarding any of the network layer protocols listed in the "protocols supported" TLV of the ISH or IIH received from the neighbour, then the DCF shall delete the adjacency with that neighbour and generate a ProtocolsSupportedMismatch Event.

If the DCF is itself capable of forwarding one or more of the network layer protocols listed in the "protocols supported" TLV of a received ISH or IIH, then the DCF shall process the ISH or IIH as normal.

The DCF shall not consider the value of the "protocols supported" TLV of LSPs during this process.

A DCF that cannot forward CLNP PDUs shall ignore ESH PDUs and consequently shall not advertise reachability to OSI End Systems.

7.1.10.1.2 IS-IS domain-wide IP prefix distribution

DCFs supporting Level-1, Level-2 Integrated IS-IS shall support the advertising of configured IP destination prefixes learned via Level-2 into Level-1 LSPs, as well as IP destination prefixes learned via Level-1 into Level-2 LSPs. The default behaviour, when no IP destination prefixes have been configured, shall be to not propagate any Level-2 prefixes into Level-1 LSPs, while all Level-1 learned prefixes shall be propagated into Level-2 LSPs.

7.1.10.1.2.1 Configuration prefixes

The operator shall provision two tables that control the propagation of prefixes. One table shall control propagation from Level-1 to Level-2, while the other controls propagation from Level-2 to Level-1.

7.1.10.1.2.2 Tagging of propagated prefixes

Since propagating prefixes from Level-2 into Level-1 and subsequently from Level-1 back into Level-2 can introduce routing loops, a tag is necessary to identify the source of the prefix. This tag, called the up/down bit, is stored in the previously unused high-order bit (bit 8) of the Default Metric field in IP Reachability TLVs and IP External Reachability TLVs. Existing implementations of IS-IS that support RFC 1195 will not be impacted by the redefinition of this bit as RFC 1195 requires it to be set to zero when originating LSPs, and ignored upon receipt. Further information is available in RFC 2966.

IP Reachability TLVs and IP External Reachability TLVs shall be processed in the same manor. The type of TLV received will be the same type used when the prefix is propagated from the Level-2 to a Level-1 area, as well as from a Level-1 area to the Level-2.

This is different than RFC 1195, which limited IP External Reachability TLVs to appearing only in Level-2 LSPs.

7.1.10.1.2.2.1 Transmission of LSPs with IP reachability TLVs and IP external reachability TLVs

As with normal RFC 1195, the value of the up/down bit shall be zero for all IP TLVs in Level-2 LSPs. The value of the up/down bit shall be zero for Level-1 LSPs originated within a Level-1 area.

The up/down bit shall be set to one in an IP TLVs in Level-1 LSP when a Level-1, Level-2 Integrated IS-IS NEs is propagating a configured prefix from Level-2 to Level-1.

7.1.10.1.2.2.2 Reception of LSPs with IP reachability TLVs and IP external reachability TLVs

A DCF supporting Integrated IS-IS shall ignore the value of the up/down bit when developing routes for use within a Level-1 area or for the Level-2.

A DCF supporting Level-1, Level-2 Integrated IS-IS that receives an LSP with an IP TLV for a prefix that matches an entry in the Level-1 to Level-2 Propagation table shall advertise the appropriate prefix from Level-1 to Level-2.

A DCF supporting Level-1, Level-2 Integrated IS-IS that receives an LSP with an IP TLV with the up/down bit set to one shall never use the prefix for propagation of information from Level-1 to Level-2.

7.1.10.1.2.2.3 Use the up/down bit in Level-2 LSPs

The use of up/down bit in Level-2 LSPs is for further study.

7.1.10.1.2.3 Route preference

Given that prefixes can now be propagated from Level-2 to Level-1, the Route Preferences specified in RFC 1195 must be updated to take into account this new source. The resulting Route Preference order is as follows:

1) L1 intra-area routes with internal metric;
   L1 external routes with internal metric.
2) L2 intra-area routes with internal metric;
   L2 external routes with internal metric;
   Inter-area routes propagated from L1 into the L2 with internal metric;
   Inter-area external routes propagated from L1 into the L2 with internal metric.
3) Inter-area routes propagated from L2 into an L1 area with internal metric;
   External routes propagated from L2 into an L1 area with internal metric.
4) L1 external routes with external metric.
5) L2 external routes with external metric;
   Inter-area external routes propagated from L1 into the L2 with external metric.
6) Inter-area external routes propagated from L2 into an L1 area with external metric.

7.1.11 IP routing interworking function

A DCF supporting the IP Routing Interworking Function shall support route-filtering mechanisms, per Sections 7.5 and 7.6 of RFC 1812, so that networks with two routing protocols can be connected via more than one exchange point.

7.1.12 "Applications to Network Layer" mapping function

OSI applications running over (a part of) the DCN that only supports IP may be mapped into IP as specified in 2.1.6/Q.811 dealing with RFC 1006/TCP/IP protocol profile. Such a mapping is a Layer 4 solution and is therefore outside the scope of this Recommendation. Another option for carrying OSI applications across (a part of) the DCN that only supports IP is to provide OSI over IP Layer 3 encapsulation as specified in 7.1.8.

The mapping of IP applications over (a part of) the DCN supporting IP shall be in accordance with IP suite specifications.

7.1.13 "MPLS PDU into ECC data-link layer" encapsulation function

This function encapsulates and unencapsulates a MPLS PDU into an ECC Data-Link Layer frame.

If PPP is the supported data link protocol on the ECC interface, the following is required:

– *At Transmit End*

  Shall put MPLS packets directly into PPP Information Field as per RFC 1661 with the MPLS protocol value of 0281 hex into the PPP Protocol Field as per RFC 3032, Section 4.3, for MPLS Unicast.

– *At Receive End*

  An MPLS packet is identified if the PPP Protocol Field has the MPLS protocol value of 0281 hex as per RFC 3032, Section 4.3, for MPLS Unicast.

7.1.14 "MPLS PDU into Ethernet frame" encapsulation function

This function encapsulates and unencapsulates a MPLS PDU into an Ethernet (version 2) frame.

It shall encapsulate MPLS PDUs into Ethernet (version 2) frames as per RFC 894 using an ethertype value of 8847 hex as per RFC 3032, Section 5, for MPLS Unicast.

7.1.15 MPLS LSP signalling function

The MPLS LSP Signalling Function provides the signalling necessary to set-up the MPLS LSP.

A DCF supporting the MPLS LSP Signalling Function shall support the following reservation model: Explicit Path with a strict route via simple nodes (32 bits IP-address), for point-to-point unicast LSP, via the Reservation Style "FF" over IPv4.

The Path message is forwarded to the destination along a path specified by a list of IP-addresses in the Explicit Route Object (ERO). Each node (LSR) in the path records the ERO. Via the Label Request object the nodes (LSR's) provide label binding for the session. See RFC 3209 – RSVP-TE, Sections 2.2, 3.1, 4.2 and 4.3.

The destination node responds with a Resv message, which is sent upstream towards the sender, in reverse order of the node-list in the ERO. The Label in the Label object of the Resv message is used in each intermediate LSR to associate outgoing traffic with this LSP. If the node is not the sender, it allocates a new Label and places that in the Label object of the Resv message, which it sends upstream to the PHOP. See RFC 3209 – RSVP-TE, Sections 2.2 and 3.2 and 4.1.

If the node cannot fulfil the request, it sends a PathErr or ResvErr message to the sender node. See RFC 3209 – RSVP-TE, Section 4.5.

The soft-state procedure of RSVP implies periodic sending of a full representation of the LSP state in Resv and Path messages to maintain the LSP. The Srefresh message is used in place of the periodic sending of standard Path and Resv messages. Each MessageID in the Srefresh message represents a full Path or Resv message, for which the state is not changed. See RFC 2961 – RSVP-ORE, Section 5.5.

A MESSAGE_ID_NACK object is used to indicate that a received MessageID does not match, and a full Path or Resv message is needed to restore the LSP. See RFC 2961 – RSVP-ORE, Section 5.4.

A MESSAGE_ID_ACK object is used to acknowledge the receipt of messages containing the MESSAGE_ID object and for which the ACK_Desired flag is set. It is part of the Srefresh re-transmission algorithm as described in RFC 2961 – RSVP-ORE, Section 6.3.

7.1.16 MPLS LSP forwarding function

The MPLS LSP Forwarding Function forwards the incoming MPLS packet to an outgoing interface based on its MPLS label and the Next Hop Label Forwarding Entry (NHLFE) as per RFC 3031.

The sequence of packets must be maintained within an LSP.

7.1.17 MPLS LSP path computation function

The MPLS LSP Path Computation Function calculates the path for a unidirectional LSP. This function shall be able to calculate paths for two unidirectional LSPs to the same destination such that their paths do not traverse the same node or subnetwork.

7.1.18 "Network layer packet into MPLS" encapsulation function

The "Network Layer Packet into MPLS" Encapsulation Function adds/removes the MPLS label stack entry to/from the network layer packet as per RFC 3032.

7.1.19 MPLS packet 1+1 protection function

7.1.19.1 Associating two LSPs

The ingress and egress nodes shall identify and associate the two LSPs providing packet 1+1 service. This association between two LSPs can be established using either network management interface or signalling.

For the case of signalling, an identifier shall be transferred across each of the diverse LSPs. The identifier shall be identical on each of the diverse LSPs and shall be unique amongst LSPs initiated by the ingress node and amongst LSPs terminated by the egress node.

The specific mechanism for assigning the identifier, as well as how the identifier is transported within the signalling protocol, is for further study. The mechanism will be similar to the one required for associating LSPs for other MPLS-based protection mechanisms such as 1+1 or 1:1.

In order to meet the requirement that there are no signalling extensions required at the intermediate nodes, the Identifier and the LSP Service Type (i.e., packet 1+1) shall be carried within opaque objects.

7.1.19.2 Sequence identifier format

The sequence number shall be used as the identifier for packet 1+1 protection. Each copy of the dual-fed packet is assigned the same unique sequence number by the ingress node. The sequence number of the next packet is generated by adding one to the current sequence number.

The egress node uses the sequence number to make sure that only the first received copy of the packet is selected, whereas the second received copy is discarded. The egress node strips off the sequence number from the packet right after its selection and before passing to the upper layer of the stack. Note that packet 1+1 recovery scheme is independent of the applications/protocols supported above MPLS.

The sequence number shall be carried in every packet as the first four bytes inside the shim header of each of the LSP providing packet 1+1 protection. The initial sequence number that is assigned to the first packet by the ingress node shall be agreed between the ingress and egress nodes. The default value of the initial sequence number is zero.

The sequence number is located after the 4-bytes MPLS encapsulation header as illustrated in Figure 7-6. Note that packet 1+1 can be provided at any level of hierarchy of a nested LSP.

4-bytes Shim header	4-bytes Sequence number
Encapsulation header	Sequence number

Figure 7-6/G.7712/Y.1703 – Sequence identifier format

7.2 Provisioning requirements

Every NE must support the creation of an interface that does not have any physical manifestation. This interface must be provisionable with an IP address.

The LSP size shall be configurable.

This allows the MTU size within the domain to be set.

Area ID provisioning per interface, including ECC channels and LAN, is required for OSPF.

7.3 Security requirements

Care must be taken to avoid unwanted interactions (addresses, etc.) between a public IP network and a DCN supporting IP.

Annex A[1]

Requirements for three-way handshaking

The three-way handshaking procedure is based upon and designed to be compatible with, the IETF IS-IS Working Group's Three-way Handshaking function (RFC 3373).

A.1 Point-to-Point three-way adjacency TLV

A DCF supporting Integrated IS-IS shall include a TLV in all point-to-point IIH PDUs. The structure of the TLV shall be:

Type = 0xF0 (decimal 240)

Length = 5 to 17 octets

Value:

Adjacency Three-way State (one octet):

0 = Up

1 = Initializing

2 = Down

Extended Local Circuit ID of four octets

Neighbour System ID of zero to eight octets if known

Neighbour Extended Local Circuit ID of four octets if known

[1] NOTE – This new Annex A replaces that of ITU-T Rec. G.7712/Y.1703 version 2001.

The Extended Local Circuit ID shall be assigned by the DCF when the circuit is created, and the DCF shall use a different value for each point-to-point circuit that it has.

The Adjacency Three-way State reported in the TLV shall be as specified in clause A.2.

A.2 Adjacency three-way state

A DCF supporting Integrated IS-IS shall have an adjacency three-way state for each point-to-point circuit. This state is different to the state specified in ISO/IEC 10589.

If no adjacency exists on a link, then the adjacency three-way state shall be set to "Down".

If a DCF receives an ISH on a point-to-point link and this results in a new adjacency being created with adjacency state "Initializing", then the adjacency three-way state shall be set to "Down".

If a DCF receives a point-to-point IIH that does not contain a three-way adjacency TLV, then the DCF shall behave as per ISO/IEC 10589, but shall include the TLV in IIH PDUs on that link reporting the adjacency three-way state as "Down".

If a DCF receives a point-to-point IIH PDU that contains a three-way adjacency TLV, then the DCF shall behave differently to ISO/IEC 10589 IIH PDU processing as follows:

- If the Neighbour System ID and the Neighbour Extended Local Circuit ID fields of the TLV are present and if either Neighbour System ID does not match the ID of the DCF, or the Neighbour Extended Local Circuit ID does not match the Extended ID of the DCF, then the IIH PDU shall be discarded and shall not be processed.
- If the IIH PDU results in the ISO/IEC 10589 state tables producing an "Up" or "Accept", and if the received Adjacency Three-way State is "Down", then the DCF shall set its adjacency three-way state to "Initializing".
- If the IIH PDU results in the ISO/IEC 10589 state tables producing an "Up" or "Accept", and if the received Adjacency Three-way State is "Initializing", then the DCF shall change its adjacency three-way state from "Down" or "Initializing" to "Up" and generate an "AdjacencyChangeState(Up)" event.
- If the IIH PDU results in the ISO/IEC 10589 state tables producing an "Up" or "Accept", and if the received Adjacency Three-way State is "Initializing", then if the DCF already has an adjacency three-way state of "Up", it shall maintain the adjacency three-way state of "Up".
- If the IIH PDU results in the ISO/IEC 10589 state tables producing an "Up" or "Accept", and if the received Adjacency Three-way State is "Up", then if the DCF already has an adjacency three-way state of "Down", it will generate an "AdjacencyStateChange(Down)" event with the reason "Neighbour restarted" and the adjacency shall be deleted with no further IIH PDU processing taking place.
- If the IIH PDU results in the ISO/IEC 10589 state tables producing an "Up" or "Accept", and if the received Adjacency Three-way State is "Up", then if the DCF already has an adjacency three-way state of "Initializing", then it will change its adjacency three-way state to "Up" and generate an "AdjacencyChangeState(Up)" event.
- If the IIH PDU results in the ISO/IEC 10589 state tables producing an "Up" or "Accept", and if the received Adjacency Three-way State is "Up", then if the DCF already has an adjacency three-way state of "Up", it shall maintain the adjacency three-way state of "Up".
- Following the comparison of source ID from the PDU with the local system, ID and manipulation of the Circuit ID shall not be performed.

If the IIH PDU results in the ISO/IEC 10589 state tables producing an "Up" or "Accept" then the DCF shall:

1) copy the adjacency areaAddressOfNeighbour entries from the Area Addresses field of the PDU;
2) set the holdingTimer value of the Holding Time field from the PDU; and
3) set the neighbourSystemID to the value of the Source ID field from the PDU as per ISO/IEC 10589.

Annex B

Requirements for automatic encapsulation

B.1 Introduction

This annex provides a specification for the optional AE-DCF that enables nodes that support routing of differing incompatible network layer protocols, such as CLNS, IPv4 or IPv6 to be present in a single IS-IS level-1 area or level-2 subdomain, and which automatically encapsulates one network layer protocol into another as required, provided that all of the nodes support IS-IS or Integrated IS-IS routing.

B.2 Scope

The AE-DCF is an optional function. When it is provided, it shall function as specified in this annex. The requirements in this annex apply only to DCFs that contain the additional functionality of an AE-DCF. The AE-DCF also requires certain behaviours from DCFs that do not include AE-DCF functionality, in order to interwork with them. Requirements for DCFs that do not include AE-DCF functionality are found in 7.1.10.1 for IP and dual nodes, and in ISO/IEC 10589 for OSI nodes.

B.3 Description of the AE-DCF

B.3.1 Introduction

Integrated IS-IS as specified in RFC 1195 was originally designed to be able to route IP and CLNS using a single routing protocol, and a single SPF algorithm. For this, it represents IPv4 addresses and subnet masks as a 64 bit number which is then treated by the SPF algorithm as if it were an OSI End System address. Integrated IS-IS nodes are required to have an IS-IS Area Address and a System Identifier, which is treated in the same way as an NSAP address is in an OSI-only node. Integrated IS-IS nodes then form adjacencies and flood System Identifiers and metrics throughout their Level-1 area (Level-1 routers) or their Level-2 subdomain (Level-2 routers) in the same way as OSI-only IS-IS nodes.

SIDs (System Identifiers) and metrics to other SIDs are flooded throughout a Level-1 area or Level-2 subdomain using LSPs (Link State PDUs) that are common to both IS-IS and Integrated IS-IS nodes. IP-specific information is then added to these LSPs using TLV extensions that are understood only by IP capable nodes. OSI-only routers cannot decode these TLVs but still flood them onwards to all of their adjacencies. In this way, an SPF tree can be built by any IS-IS or Integrated IS-IS node whether it can route CLNS, IPv4 or IPv6. OSI-capable nodes will calculate shortest paths to OSI End Systems, IPv4-capable nodes will calculate shortest paths to IPv4 addresses or prefixes and IPv6-capable nodes will calculate shortest paths to IPv6 addresses or prefixes.

One consequence of this is that an OSI-only node will calculate a shortest path to an OSI End System that goes through an IP-only node, even though that IP-only node cannot forward CLNS packets. Similarly, an IP-only node will calculate a shortest path to an IP destination that goes through an OSI-only node, even though the OSI-only node cannot forward IP packets. Thus an OSI-only capable node must not be placed in a part of a network where there is any possibility of it being on the shortest path to IP destinations, and an IP-only node must not be placed in a part of the network where there is any possibility of it being on the shortest path to an OSI End System.

The Integrated IS-IS algorithm can only use a single SPF algorithm for two or more network layer protocols due to an assumption that all network-layer protocols have access to the same resources, in other words, the same network with the same topology. Thus Integrated IS-IS requires any node in a Level-1 area or Level-2 subdomain to be able to route any network layer protocol that is present in the area or domain respectively.

For this reason, RFC 1195 places topological restrictions on networks that are routed by Integrated IS-IS, requiring that all of the nodes support both IP and CLNS in an area that have both CLNS traffic and IP traffic present in them.

Consequently, according to RFC 1195, if one node is upgraded and forwards IP packets, then all of the others in the Level-1 area or Level-2 subdomain must also be upgraded.

The solution proposed here allows this topological restriction to be removed, and it automatically encapsulates CLNS packets inside IP packets for forwarding across IP-only nodes and encapsulates IP packets inside CLNS packets for forwarding across OSI-only nodes. The solution proposed here is fully compatible with existing OSI-only nodes, which will not require any upgrade. It places one requirement upon IPv4-only or IPv6-only nodes above those in RFC 1195, specifically the Network-layer Protocol Aware Adjacency Creation Function specified in 7.1.10.1.1.

B.3.2 The basic concept

This feature takes advantage of the fact that all Integrated IS-IS and IS-IS nodes share basic topology information in the same way, and of the behaviour that OSI-only nodes will attempt to forward a packet across an IP-only node and vice versa, even though that node is incapable of actually forwarding the packet. Normally, this would result in packet loss, but an AE-DCF encapsulates packets before they are forwarded across incompatible nodes so that they are not lost.

When two islands of IP capable Integrated IS-IS nodes are connected using a central network that supports only OSI, and if all of the nodes participate in the same area (for Level-1 nodes), then the IP capable nodes will receive the LSPs from all of the other IP capable nodes, even those in the other island, as well as the LSPs from all of the OSI-only nodes in the centre. Thus they calculate shortest paths across the OSI-only nodes for all of the IP destinations in the island on the far side. It is only when an IP-capable node actually forwards an IP packet to an OSI-only node that things go wrong, and the packet is lost. Hence, the topological restrictions in RFC 1195.

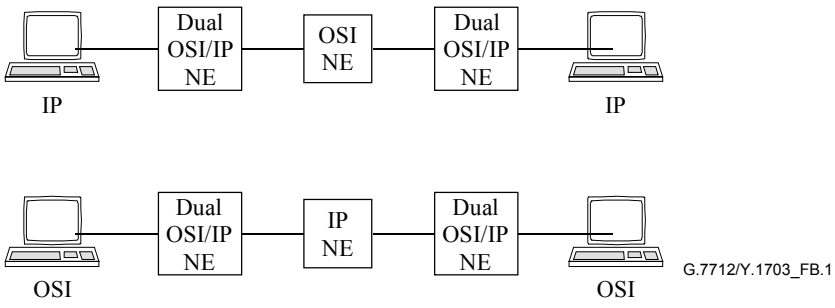


Figure B.1/G.7712/Y.1703 – Illegal topologies

The above simple networks illustrated in Figure B.1 are illegal topologies according to RFC 1195. In the top network IP packets will be routed from one side of the network to the other, but on arrival at the OSI-only node will be discarded. Similarly, on the bottom network CLNS packets will be routed from one side of the network to the other, but on arrival at the IP-only node will be discarded. An AE-DCF specified here corrects this behaviour.

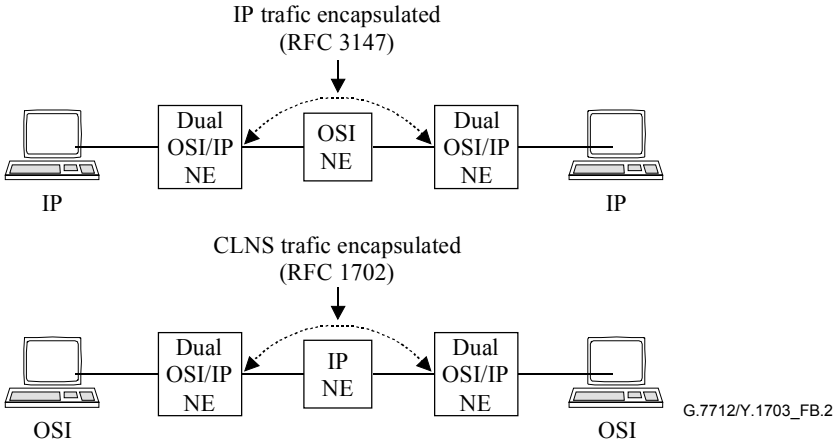


Figure B.2/G.7712/Y.1703 – Encapsulation "repair"

The AE-DCF resides in dual nodes, and enables them to recognize that a particular neighbour will discard certain traffic, and so to encapsulate it into a form that will not be discarded (see Figure B.2). This 'repairs' the network so that the part of network between the dual nodes acts as if it is comprised of all dual nodes when, in actual fact one, or more of the nodes are not dual.

An AE-DCF does not alter the path that a packet will take across the network; any individual packet will still cross the network using the shortest path as calculated by the normal IS-IS SPF algorithm.

The Network-layer Protocol Aware Adjacency Creation Function specified in 7.1.10.1.1 forces traffic to go through nodes that support both IP and OSI whenever the shortest path takes traffic across a boundary between IP-capable and OSI-capable parts of an area. The AE-DCF then enables those dual nodes to encapsulate a packet if necessary, so that it can be forwarded by nodes that do not support that network layer protocol. This encapsulation takes place only when necessary, and thus these tunnels are automatically created and are dynamic. The resulting tunnels are not maintained in any way and exist only as entries in forwarding tables. The tunnels do not appear as a circuit or interface as far as the routing protocol is concerned. Thus packets still cross the network along the shortest path that each node calculates normally, and there is no need for IS-IS packets to be encapsulated, only IP and CLNS traffic is encapsulated.

B.4 Requirements and limitations

B.4.1 Requirements for OSI-only nodes

In order to interwork with the AE-DCF OSI-only nodes are required to be conformant to ISO/IEC 10589.

B.4.2 Requirements for IP-capable nodes

In order to interwork with the AE-DCF, IP-only nodes are required to be conformant to RFC 1195.

In particular, IP-capable nodes are required to ignore the "protocol supported" TLV in LSPs of nodes that they are considering as candidates for shortest paths when running the SPF algorithm.

An IP-capable node that only includes IP-capable nodes in its SPF calculation would not conform to RFC 1195, where it states:

– From page 26 of RFC 1195: "The Dijkstra computation does not take into consideration whether a router is IP-only, OSI-only, or dual. The topological restrictions specified in section 1.4 ensure that IP packets will only be sent via IP-capable routers, and OSI packets will only be sent via OSI-capable routers."

The AE-DCF is compatible with RFC 1195 implementations that conform to the above statement. An implementation that only includes IP-capable nodes in its SPF calculation would not view paths through OSI-only nodes as being a suitable route, and so will not take advantage of the AE-DCF.

In order to interwork with the AE-DCF, IP nodes are required to conform to 7.1.10.1.1. The reason for this is stated below:

– This solution is dependant upon IP packets arriving at an OSI-only node, having first gone through an AE-DCF, and upon CLNS packets arriving at an IP-only node, having first gone through an AE-DCF. The AE-DCF is then responsible for encapsulating these packets so that they can be forwarded.

– Therefore an IP-only node must never have an adjacency with an OSI-only node.

– If this solution is used to mix IPv4 and IPv6 nodes in the same Level-1 area or Level-2 subdomain then similarly an IPv4-only node must never have an adjacency with an IPv6-only node.

– This requirement is met if all IP-capable nodes conforming to 7.1.10.1.1. Note that this requirement is not present in RFC 1195.

Alternatively, an operator may manually ensure that nodes that do not support a network layer protocol in common do not have adjacencies.

B.4.3 Requirements for automatically encapsulating dual or multi-lingual nodes

If this feature is to be used in a Level-1 area or Level-2 subdomain, then nodes that support more than one network layer protocol, but that do not support the AE-DCF, may be used with caution. A safer alternative is either to comply with the topological restrictions of RFC 1195, or to use only dual or multilingual nodes that contain the AE-DCF.

B.4.3.1 Encapsulation capability TLV

The AE-DCF will include a new TLV in LSPs with LSP number equal to zero. The new TLV will have the following structure:

Code: 16 (decimal)
Length: The length of the value
Value: A variable length part containing the following:
Sub-TLV type: 1
Sub-TLV length: 3 times the number of encapsulation modes in the sub-TLV
Sub-TLV value:
47 indicating that the next two bytes are a GRE encapsulation;
The NLPID of a packet that may be encapsulated (inner);
The NLPID of a packet that transports the encapsulated packet (outer);
Bytes 4,5,6: A second encapsulation mode (if needed);
Bytes 7,8,9: A third encapsulation mode (if needed);
Etc.

The NLPIDs that are used shall be those as specified in ITU-T Rec. X.263 | ISO/IEC 9577. Nodes that transmit this TLV shall indicate the formats that a node can both receive and transmit. Nodes must be able to both automatically encapsulate and automatically unencapsulate the formats that are described in the TLV, so that traffic may be received, and so that traffic may return in the reverse direction.

It is recommended that dual nodes supporting an AE-DCF are able to encapsulate/unencapsulate A over B, and B over A (where A and B are the two supported network layer protocols) making two encapsulation modes in a typical dual node.

For example, the contents of the TLV for a typical OSI and IPv4 AE-DCF will be:

16: the code;
8: the value length (in this example);
1: sub-TLV type 1;
6: sub-TLV length (in this example);
47: next two bytes are a supported GRE mode;
129: IPI for CLNP from ITU-T Rec. X.263 | ISO/IEC 9577;
204: IPI for IPv4 from ITU-T Rec. X.263 | ISO/IEC 9577;
47: next two bytes are a supported GRE mode;
204: IPI for IPv4 from ITU-T Rec. X.263 | ISO/IEC 9577;
129 IPI for CLNP from ITU-T Rec. X.263 | ISO/IEC 9577.

An OSI, IPv4, IPv6 AE-DCF will thus typically use six encapsulation modes to indicate CLNP over IPv4, CLNP over IPv6, IPv4 over CLNS, IPv4 over IPv6, IPv6 over CLNS, and IPv6 over IPv4, giving a value length of 20.

This TLV will not be included in pseudonode LSPs.

An AE-DCF that does not have any IPv4 addresses must not place any encapsulation formats in its TLV of type equal 16 that include IPv4 as an encapsulation transport (outer) NLPID until such time as an IPv4 address is provisioned and advertised.

An AE-DCF that does not have any IPv6 addresses must not place any encapsulation formats in its TLV of type equal 16 that include IPv6 as an encapsulation transport (outer) NLPID until such time as an IPv6 address is provisioned and advertised.

B.4.3.2 Forwarding process

As the AE-DCF does not modify the path that a packet follows, an AE-DCF may calculate a shortest path for an IP packet that results in the next hop being an OSI-only node.

When this happens the AE-DCF must not simply forward a packet to an adjacent node that does not support that type of network layer protocol. Instead, the AE-DCF must encapsulate the packet inside a new packet of a type that the next hop does support. The criteria for whether an adjacent node does or does not support a particular network layer protocol is whether that network layer protocol is listed in the "protocols supported" TLV in IS-IS Hello PDUs received from the node on the adjacency which is the next hop for that destination.

This new packet requires a network layer protocol, a destination address, and a source address to encapsulate the original packet:

– The network layer protocol of the new packet must be one that is supported by the next hop as defined by the "protocols supported" TLV of Hello PDUs received from the next hop.
– The destination address of the new packet must be equal to the identity of the next node along the shortest path to the original destination that has transmitted an encapsulation mode that has both the type of network layer protocol that the original packet is as the encapsulated (inner) NLPID, and a network layer protocol that is supported by the next hop

(as defined by the "protocols supported" TLV of Hello PDUs received from the next hop) as the encapsulation transport (outer) NLPID.

– This must be achieved by inspection of the new TLV of type equal to 16 from LSPs received from each node in the path to the destination, until the first is found that meets the above requirement.
– When inspecting TLVs of type equal to 16, an AE-DCF shall ignore any sub-TLVs that it does not understand, and shall jump to the next sub-TLV and shall inspect that, either until it finds all of the encapsulation modes that it is looking for, or until it reaches the end of the TLV.
– The source address of the new packet must be equal to the identity of the AE-DCF that constructs the new encapsulation packet.

If an AE-DCF can forward a packet without encapsulation because the next hop supports that type of packet, then the AE-DCF must forward the packet without encapsulating it.

An AE-DCF might send LSPs containing IP reachability from an IP-only node on to a split stack node, or vice versa, and consequently might then be required to encapsulate packets headed for a split stack node, or unencapsulate packets received from a split stack node.

Thus an automatically encapsulating split stack node must also follow the same process of inspecting LSPs of nodes between itself and the destination looking for a node that has a suitable encapsulation format.

Note that a split stack node might be capable of receiving an IPv4 packet only encapsulated inside CLNS for example. In this case, the split stack node will transmit only "CLNS" in the "protocols supported" field of its Hello packets, and will only include one encapsulation mode in its TLV of type equal 16 in its LSPs. This single encapsulation mode will specify IPv4 as the encapsulated (inner) packet NLPID and CLNS as the encapsulation transport (outer) packet NLPID.

B.4.3.3 Receipt process

When an AE-DCF receives a packet that is destined for itself, it must inspect that packet to see if it has another packet encapsulated inside it. The resultant unencapsulated CLNS, IPv4 or IPv6 packet must then be forwarded as normal. If the resultant unencapsulated packet then contains another packet destined for this node, the process repeats; this is because multiple layers of encapsulation may require unencapsulation at a single AE-DCF.

IS-IS packets are not compatible with IP packets and cannot be forwarded across the public Internet or other IP-only networks. This is a security advantage as it makes it difficult for a malicious entity to remotely launch IS-IS packets at IS-IS or Integrated IS-IS nodes across the public Internet. In order not to remove this advantage, then, if an IS-IS or ES-IS packet arrives encapsulated inside another packet destined for an AE-DCF, then the AE-DCF must discard it unless it came from a node with which the AE-DCF has a manually provisioned tunnel with IS-IS provisioned to run across it. Optionally, an error report may be raised informing the network manager of information such that a packet was received and dropped, where it came from, or that it is a potential malicious event.

All packets must be encapsulated using GRE encapsulation as specified in 7.1.8.

B.4.3.4 MTU size and fragmentation requirements

The encapsulation of one packet inside another may result in a new packet that is longer than the MTU size of the link over which this new packet must be forwarded. This new GRE packet must not be discarded, therefore, these packets must not have the Don't Fragment bit set if they are IPv4 packets and must have the Segmentation Permitted flag set if they are CLNS packets, as per 7.1.8.

The resultant encapsulation packets must then be fragmented before being forwarded if the packet is now longer than the MTU limit of the link.

It is not necessary to fragment a packet before encapsulating it, as the resultant encapsulation packet will be fragmented if necessary.

B.4.3.5 Requirements for AE-DCF with broadcast (LAN) interfaces

B.4.3.5.1 Pseudo-node election process

According to 7.1.10.1.1 IP-only nodes are not allowed to form an adjacency with OSI-only nodes, and IPv4-only nodes are not allowed to form an adjacency with IPv6-only nodes.

Therefore, when IP-only and OSI-only nodes are connected to the same LAN and in the same Level-1 area or Level-2 subdomain, then the IP-only nodes will form adjacencies with one another and will elect a pseudonode, whilst the OSI-only nodes will form separate adjacencies and will elect a different pseudonode. Therefore, there will be two separate pseudonodes on the LAN, one for the OSI-only nodes, and one for the IP-only nodes.

A similar thing may happen if IPv4-only and IPv6-only nodes are connected to the same LAN.

An AE-DCF must, therefore, take part in these separate pseudonode election processes independently for each network layer that it supports. A Level-1/Level-2 AE-DCF must take part in two pseudonode election processes for each network layer protocol that it supports (one for Level-1 and one for Level-2).

Each pseudonode on the LAN residing on a node of a network layer protocol compatible with the AE-DCF, will have an adjacency with the AE-DCF. Thus on an IP & OSI LAN the AE-DCF will correctly be the one that has valid adjacencies both with the IP pseudonode and with the OSI pseudonode (if multiple pseudonodes are present on the LAN). The AE-DCF will have an adjacency with the IP pseudonode and with the OSI pseudonode, but the IP pseudonode will not have a direct adjacency with the OSI pseudonode, and vice versa, but will instead gain connectivity only through the AE-DCF, thus guaranteeing that CLNS packets are encapsulated by the AE-DCF before being forwarded to IP-only nodes, and that IP packets are encapsulated by the AE-DCF before being forwarded to OSI-only nodes.

An IP- and OSI-capable AE-DCF may be elected as the Designated Router by the IP-capable nodes on the LAN, but not by the OSI-capable nodes; in this case, the AE-DCF must create a pseudonode, but the pseudonode must declare adjacencies in its LSPs only with the IP-capable nodes on the LAN.

Similarly, an IP- and OSI-capable AE-DCF may be elected as the Designated Router by the OSI-capable nodes on the LAN, but not by the IP-capable nodes; in this case, the AE-DCF must create a pseudonode, but the pseudonode must declare adjacencies in its LSPs only with the OSI-capable nodes on the LAN.

An IP- and OSI-capable AE-DCF may be elected as the Designated Router both by the IP-capable and by the OSI-capable nodes on the LAN; in this case, the AE-DCF must create a pseudonode that declares adjacencies in its LSPs to all of the nodes on the LAN.

In essence, an AE-DCF takes part in a separate election process for each network layer protocol that it supports, and if it wins any of the elections then it creates a pseudonode, but the pseudonode will declare adjacencies in its LSPs only with the set, or sets, of nodes that elected it.

Consequently, OSI-only or IP-only nodes may receive LSPs from a pseudonode that lists adjacencies to nodes on the LAN that they do not have adjacencies with. If a packet should need to be forwarded via such a node, then it should be sent to the Designated IS as per ISO/IEC 10589 section C.2.5 item "h", and as per RFC 1195 section C.1.4 step 0 clause 8 on page 73. Note that these clauses in ISO/IEC 10589 and RFC 1195 are non-normative. It is possible that there are

implementations that do not exhibit this behaviour. Such an implementation will drop packets rather than send traffic to an AE-DCF for automatic encapsulation, if the AE-DCF is the Designated Router, and if non-compatible nodes on the same LAN are on the shortest path.

Implementers and operators, therefore, have a choice to make, the choice is:

1) Set the priority of the AE-DCF to a high value. This results in a single pseudonode appearing on the LAN, supported by an AE-DCF. The disadvantage of this approach is that there is a small chance that a legacy implementation exists on the LAN that does not forward traffic to an AE-DCF if a non-compatible node on the LAN is on the shortest path.

or

2) Set the priority of the AE-DCF to a low value. This results in one pseudonode appearing on the LAN for every network-layer protocol supported, explicitly sending traffic for non-compatible nodes through an AE-DCF. This improves interoperability but doubles the amount of LSPs transmitted onto the LAN, possibly reducing scalability.

It is recommended that the priority of an AE-DCF is operator configurable.

B.4.3.5.2 LSP update process

ISO/IEC 10589 states in section 7.3.15.1 that an LSP received, that does not come from a valid adjacency, must be discarded. A strict OSI-only implementation will therefore reject LSPs that are transmitted onto a LAN interface by an IP-only node, as the IP-only node has rejected the adjacency as per 7.1.10.1.1. Thus the OSI-only node can receive such an LSP only from an AE-DCF. Without modified behaviour, a dual node would only forward such an LSP during periodic LSP database synchronization.

An AE-DCF is, therefore, required to have modified LSP flooding behaviour so that OSI-only or IP-only nodes do not need to wait for the next LSP database synchronization event.

An AE-DCF must check incoming LSPs that arrive on LAN interfaces to see if they come from a neighbour that supports all of the network layer protocols that the AE-DCF does. This must be achieved by inspection of the "protocols supported" TLV in Hello packets received from that neighbour.

If the LSP is received from a neighbour that does support all of the network layer protocols that the AE-DCF supports, then the AE-DCF shall behave as per ISO/IEC 10589 and unset the SRM flag for that LSP on that LAN interface if it already has the LSP, or shall flood it out of all other interfaces if it does not already have the LSP.

If the LSP is received from a neighbour that does not support all of the network layer protocols that the AE-DCF supports, and, if it does not already have the LSP, then the AE-DCF shall set the SRM flag for that LSP on the LAN interface over which the LSP was received, in addition to all other interfaces, resulting in the AE-DCF retransmitting the LSP onto the LAN.

In this way, if an LSP is transmitted onto the LAN by an IP-only node, then an AE-DCF will retransmit the LSP, so that it may be received on a valid adjacency by OSI-only nodes on the LAN and vice-versa.

B.4.3.5.3 Redirects

If an AE-DCF originates an ICMP redirect request, the request must not redirect IPv4 packets from an IPv4-capable node to a non-IPv4-capable node. Likewise, if an AE-DCF originates ISO/IEC 9542 Redirect PDUs, the redirect must not redirect CLNS packets from an OSI capable node to a non-OSI-capable node.

B.4.3.5.4 Mixing of dual RFC 1195 only and automatically encapsulating nodes on a LAN

A dual node that is conformant to RFC 1195, but that does not support an AE-DCF, must not reside on a LAN in the same Level-1 area or Level-2 subdomain as both IP-only and OSI-only nodes, as it may forward IP traffic to an OSI-only node, or CLNS traffic to an IP-only node, resulting in packet loss. This is a topological restriction of RFC 1195.

A dual node that is conformant to RFC 1195, but that does not support an AE-DCF, may reside on a LAN in the same Level-1 area or Level-2 subdomain as an AE-DCF.

Additionally, it may reside on a LAN with an OSI-only node if it can forward only CLNS traffic to that node, an IPv4-only node if it can forward only IPv4 traffic to that node, or an IPv6-only node if it can forward only IPv6 traffic to that node.

B.4.4 Requirements for automatically encapsulating split stack nodes

A split stack node initiates and terminates packets of a network-layer protocol type that it cannot forward natively in its DCC channels. Therefore, the only way that such a node may initiate or terminate such packets is if they are in an encapsulated form.

This solution is particularly useful for adding an IP card into a predominantly OSI node, or a node that will be installed into an existing OSI network, for example. It may also be easier to upgrade an OSI gateway NE to a split stack node, rather than to a dual AE-DCF, so that IP traffic can get in and out of the network for which the node is a gateway.

The split stack node must be able to internally route any packets that it receives that are of a network-layer protocol equal to one of those listed in the "protocols supports" TLVs of its IS-IS LSPs.

A split stack node must use the "protocols supported" TLV in IS-IS Hello PDUs to indicate only the network-layer protocols that it can receive and forward natively on any individual interface (or not support this TLV if it is an OSI-only interface).

That is, an IP-over-OSI node can route CLNS natively in its DCC channels, and can route IP traffic that arrives for it in IP-over-OSI GRE encapsulated packets, or possibly an Ethernet interface.

Thus, a split stack node may indicate one network layer protocol in the "protocols supported" TLV of Hello packets on one interface, and a different network layer protocol in the "protocols supported" TLV of Hello packets on another interface. Such a node would be able to route both network layer protocols internally, and so would advertise both in the "protocols supported" TLV of its LSPs.

A split stack node must use IP reachability TLVs in IS-IS LSPs to indicate the address range of encapsulated packets that it is able to terminate.

A split stack node might receive IP reachability extensions from an IP-only node, via a dual AE-DCF. Therefore the split stack node must be able to send traffic to a destination via an AE-DCF, which it will use to unencapsulate its packets. To achieve this, a split stack node must search for the next node along the path to each destination capable of unencapsulation, or for a split stack destination, in exactly the same way that an AE-DCF does.

An automatically encapsulating split stack node shall advertise the encapsulation modes that it supports using Encapsulation Capability TLV as per B.4.3.1.

When a split stack node receives a packet that is destined for itself, it must inspect that packet to ascertain whether it has another packet encapsulated inside it. If so, then the packet will be processed internally, unless it is an IS-IS or ES-IS packet, in which case it must be discarded (unless a manually provisioned tunnel exists with IS-IS provisioned to run across it) in the same way as it would be by a dual AE-DCF.

In the same way as a dual AE-DCF, a split stack node must support GRE encapsulation as specified in 7.1.8.

B.4.5 Use of IP nodes that do not conform to 7.1.10.1.1 with the AE-DCF

IPv4-only or IPv6-only nodes that are conformant to RFC 1195, but that do not support the Protocol Aware Adjacency Creation Function specified in 7.1.10.1.1, may be used in the same mixed Level-1 area or Level-2 subdomain as an AE-DCF, but the network manager must manually ensure that such a node does not have any adjacencies with other nodes that might forward packets to it that it does not support.

B.4.6 Use of dual nodes with no AE-DCF and dual nodes with AE-DCF in the same IS-IS area

Dual nodes that are conformant to RFC 1195, but that do not support an AE-DCF, may be used in mixed Level-1 areas or Level-2 subdomains with an AE-DCF with the restrictions below:

Integrated IS-IS nodes (or clusters of nodes) that support more than one network layer protocol but which do not support an AE-DCF are still subject to the topological restrictions of RFC 1195. This means that the network manager must ensure that such a node cannot pass packets to a neighbouring node that cannot forward that type of packet.

That is, dual signifies a dual Integrated IS-IS node that conforms to RFC 1195, but that does not contain an AE-DCF.

OSI-AEDCF-dual-AEDCF-IP is a safe combination;

OSI-AEDCF-dual-dual-dual-AEDCF-IP is a safe combination;

IPv4-AEDCF-dual IPv4&IPv6-AEDCF-IPv6 is a safe combination;

dual-AEDCF-OSI-AEDCF-dual is a safe combination;

OSI-IPv4&OSIAEDCF-dual IPv4&OSI-dual IPv4&IPv6-IPv4&IPv6 AEDCF-IPv6 is not a safe combination;

OSI-IPv4&OSIAEDCF-dual IPv4&OSI-IPv4&IPv6&OSI-dual IPv4&IPv6-IPv4&IPv6 AEDCF-IPv6 is not a safe combination.

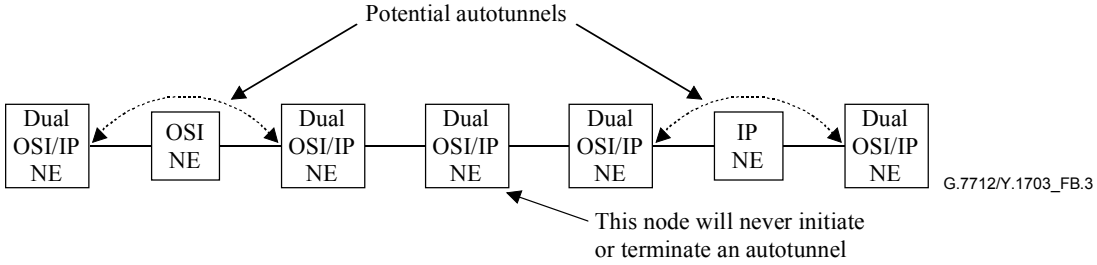


Figure B.3/G.7712/Y.1703 – Topological requirements for IS-IS area dual nodes

B.4.7 Requirements for Level-1 and Level-2 nodes

It is recommended that nodes that support both Level-1 and Level-2 routing, and that are present in an area in which this AE-DCFs are used either:

– Support all network layer protocols that are present in both the Level-1 and the Level-2 subdomain in which the node participates and support an AE-DCF.

or

– Support all network layer protocols that are present in both the Level-1 and the Level-2 subdomain in which the node participates and be either directly connected to, or connected through, continuous strings of other nodes that support all network layer protocols in the area, to a node that supports an AE-DCF and that supports all of the network layer protocols in the area.

That is, dual signifies an Integrated IS-IS node that conforms to RFC 1195, but that does not support an AE-DCF:

L2_subdomain-dual_L1/L2-non_dual is safe (as per RFC 1195);

L2_subdomain-dual_L1/L2-dual-dual-non_dual is safe (as per RFC 1195);

L2_subdomain-dual_L1/L2-AE-DCF-mixted_network is safe;

L2_subdomain-dual_L1/L2-dual-dual-AE-DCF-mixted_network is safe;

L2_subdomain-dual_L1/L2-non_dual-dual is not safe (unless RFC 1195 restrictions are applied);

L2_subdomain-dual_L1/L2-non_dual-AE-DCF is not safe (unless RFC 1195 restrictions are applied).

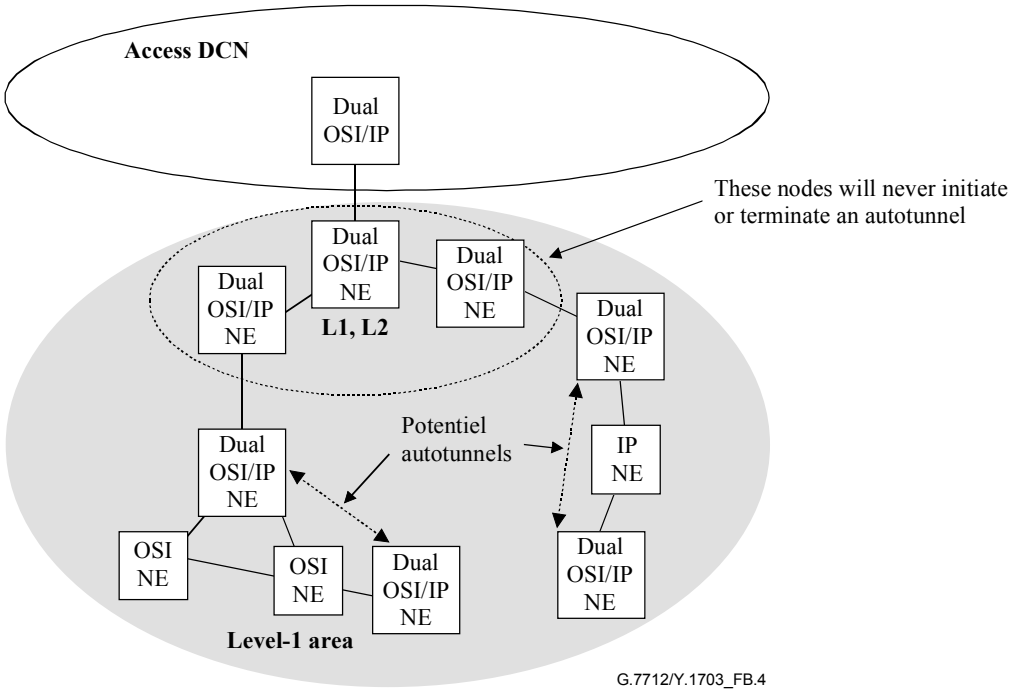


Figure B.4/G.7712/Y.1703 – Requirements for Level-1, Level-2 nodes

However, it is understood that a gateway NE, and therefore a L1, L2 router, may be an existing OSI-only device. In this case, it is possible to have IP and automatic encapsulation in the area by using the following method, with care:

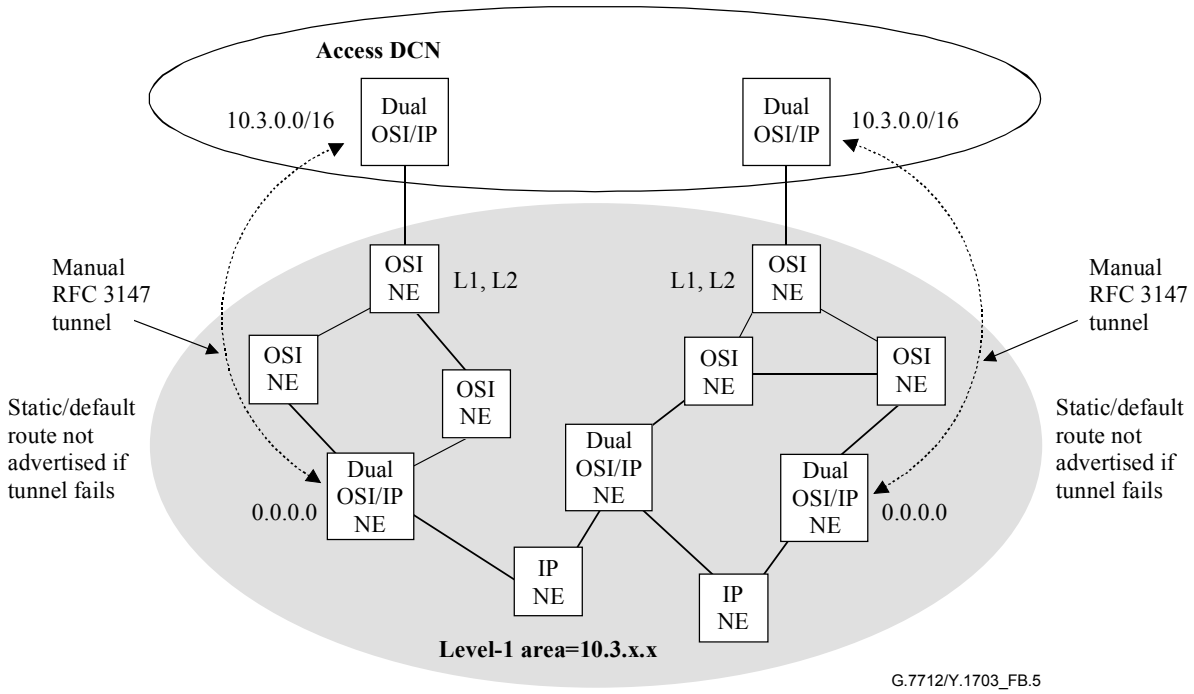


Figure B.5/G.7712/Y.1703 – Use of an OSI-only device as a gateway

One or more dual nodes in the area may be chosen as gateways for IP packets. These nodes will be configured to advertise a default route (0.0.0.0) into the area to attract all "out of area" IP traffic to them. These nodes will then forward all "out of area" traffic across a manually provisioned GRE tunnel, which passes through the Level-1, Level-2 OSI-only node to another dual node outside of the area.

The dual node that is outside of the area must have a prefix manually provisioned into it to attract all IP traffic bound for the area to it, and send it over the tunnel into the area. Optionally, a mechanism, such as an IP routing protocol, may be provisioned across the tunnel so that each end may see if the other is alive; however, if Integrated IS-IS is used, then it must be a different routing instance to that used generally in the area, as it is effectively a different routing domain.

If such a mechanism is used, then if the far end disappears, the dual node inside the area should stop advertising a default route, and the dual node outside of the area should stop advertising the prefix that represents the nodes in the area. In this way, redundant IP gateways can be provisioned.

Note that RFC 1195 states that default routes should not be advertised within Level-1 LSPs. This solution requires that this rule be broken. Normally a Level-1 RFC 1195 node would consider a Level-1, Level-2 node to be its default route. This solution requires that this behaviour be overwritten by receipt of a default route advertisement in a Level-1 LSP. If this is not possible, then a work-around is for the IP gateway nodes to be configured with a selection of static routes that cover all possible "out of area" destinations that an IP stack in the area is likely to try to reach.

B.4.8 Requirements for the Level 2 subdomain

It is acceptable to route all protocols present natively in the Level-2 subdomain, as per RFC 1195, in which case, none of the Level-2 nodes need to support an AE-DCF, but all of them must support all of the network layer protocols present.

Alternatively, it is acceptable to use Level-2 nodes that support less than all of the network layer protocols present in the domain, in which case, the Level-2 dual or multilingual nodes will be required to support an AE-DCF so that packets may be automatically encapsulated in order to pass through such nodes.

Appendix I[2]

Constraints of the interworking functions in DCN

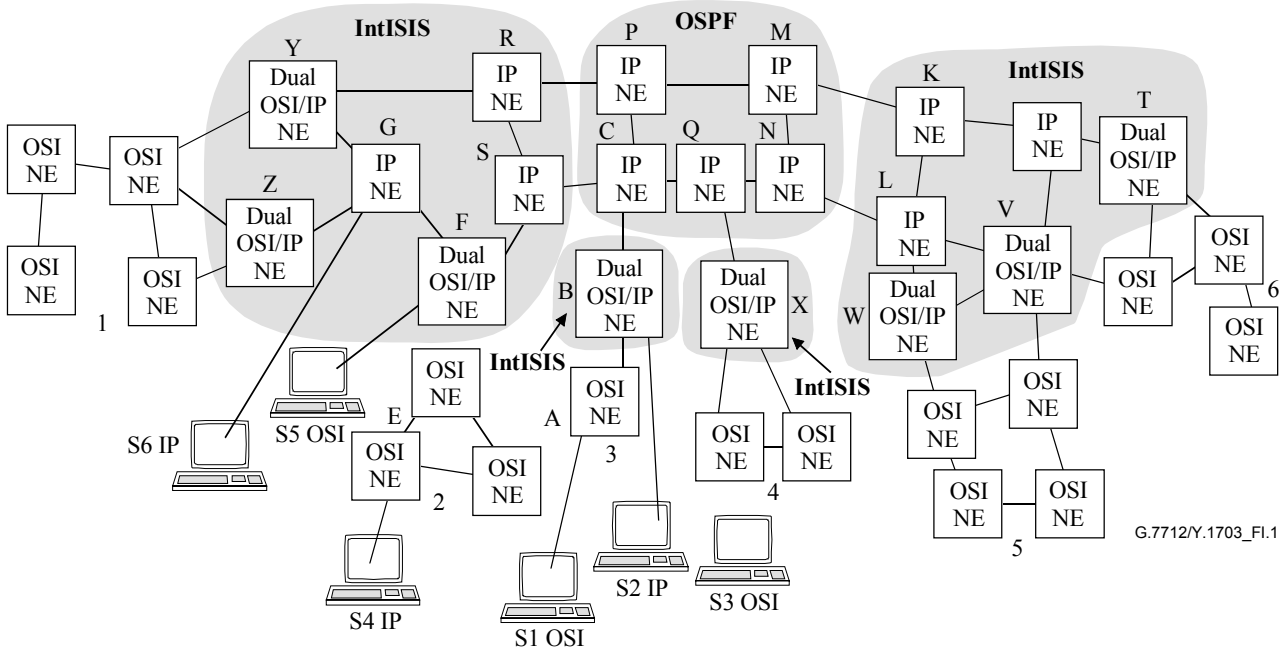


Figure I.1/G.7712/Y.1703 – Interworking scenarios

I.1 General assumptions

DCN covers the IWF for Layer 2-3 of the IP-OSI stacks. Interworking mechanisms that apply to other layers are out of the scope of this Recommendation, (i.e., mediation).

See 7.1.7 for a definition of interworking.

Tunnels are based on RFCs.

The IP-only NEs support IP routing and may contain redistribution between integrated IS-IS and OSPF.

I.2 Common to all scenarios

Dynamic routing is accomplished through the use of route redistribution of IP address information between OSPF and IS-IS NEs. Route redistribution is preformed on the OSPF nodes between the pairs; (R,P), (S,C), (M,K), (N,L).

I.2.1 Scenario 1: OSI-based management system connected to node A

There must be at least one tunnel configured from B to one or more of Y or Z.

There must be a tunnel configured from B to X.

There must be a tunnel configured from B to F.

There must be at least one tunnel configured from B to one or more of W, V or T.

---

2 NOTE – This new Appendix I replaces that of ITU-T Rec. G.7712/Y.1703 version 2001.

The above tunnels will probably have IS-IS running across them (inside the tunnel), however, inter-domain routing techniques is also a possibility. Under some conditions, some tunnels could become congested as a result of routing choices.

An OSI-based management system now has CLNS connectivity to any OSI-only or dual stack NE in the network, but does not have connectivity with IP-only NEs. Although an OSI-based manager will be able to send CLNS packets to a dual stack NE, it will not be able to manage it unless it is OSI manageable.

I.2.2 Scenario 2: IP-based management systems connected to node B

In this particular network, IP traffic can be forwarded from B to all IP NEs without requiring tunnels. OSPF NEs P, C, M, and N must support redistribution of IP routes into Integrated IS-IS. Filters will have to be configured on OSPF nodes P, C, M, and N in order to stop routing loops from forming.

An IP-based management system now has IP connectivity to any IP-only or dual stack NE in the network, but does not have connectivity with OSI-only NEs. Although an IP-based manager will be able to send IP packets to a dual stack NE, it will not be able to manage it unless it is IP manageable.

I.2.3 Scenario 3: OSI-based management systems connected to node C

NE C cannot provide OSI connectivity, and so CLNS packets cannot be forwarded; therefore, an OSI-based management system cannot function at this location.

I.2.4 Scenario 4: IP-based management systems connected to node E

NE E cannot provide IP connectivity, and so IP packets cannot be forwarded; therefore, an IP-based management system cannot function at this location.

I.2.5 Scenario 5: OSI-based management systems connected to node F

CLNS traffic can pass through NE F to OSI network 2 without requiring tunnels as NE F can forward CLNS packets natively.

There must be a tunnel configured from F to B.

There must be at least one tunnel configured from F to one or more of Z or Y.

There must be a tunnel configured from F to X.

There must be at least one tunnel configured from F to one or more of W, V or T.

The above tunnels will probably have IS-IS running across them (inside the tunnel), however, interdomain routing techniques are also a possibility. Under some conditions, some tunnels could become congested as a result of routing choices.

An OSI-based management system now has CLNS connectivity to any OSI-only or dual stack NE in the network, but does not have connectivity with IP-only NEs. Although an OSI-based manager will be able to send CLNS packets to a dual stack NE, it will not be able to manage it unless it is OSI manageable.

I.2.6 Scenario 6: IP-based management systems connected to node G

In this particular network, IP traffic can be forwarded from G to all IP NEs without requiring tunnels. OSPF NEs P, C, M, and N must support redistribution of IP routes into Integrated IS-IS. Filters will have to be configured on each OSPF nodes P, C, M, and N in order to stop routing loops from forming.

An IP-based management system now has IP connectivity to any IP-only or dual stack NE in the network, but does not have connectivity with OSI-only NEs. Although an IP based manager will be able to send IP packets to a dual stack NE, it will not be able to manage it unless it is IP manageable.

Appendix II

Example implementation of automatic encapsulation

II.1 Introduction

This appendix is not a requirement but gives brief example details on how a node may be implemented with respect to one aspect of the feature specified in this Recommendation.

The simplest way (but not the only way) for a node to calculate the next node along the shortest path to the final destination of a packet that can unencapsulate, is to modify the SPF algorithm to achieve this.

The algorithm can be modified to find the next node along the shortest path to the destination that can accept IP over OSI encapsulated traffic, and the next node along the shortest path to the destination that can accept OSI over IP encapsulated traffic. Note that these two may be the same node, or may be two separate nodes. A modified Dijkstra algorithm is provided below that achieves this.

This additional process only need happen when the next hop does not support the network layer protocol of the type that corresponds to the destination address for that path. If the next hop does support that type of network layer protocol (as specified in the "protocols supported" TLV present in IS-IS Hello PDUs received from that node), then packets to that destination may simply be forwarded natively and forgotten, and so the search for a node along the path that can unencapsulate is not necessary.

The algorithm must then identify an IP address for this next unencapsulation node if the destination of the path is an OSI End System, and must then identify an OSI address for this next unencapsulation node if the destination of the PATH is an IP address.

Failure to find an IP address for this next unencapsulation node indicates a configuration error in that node (no IP address); this may optionally result in an error message being sent to the network administrator. Packet loss will result if a CLNS packet requires tunnelling to that node over IP as, without an IP destination address, encapsulation may not be possible and the packet will be discarded instead.

Failure to find a node that can unencapsulate indicates a network design error, more specifically, a failure to conform to the topological restrictions stated in this Recommendation. This should result in a "destination unreachable" error report.

For each IP destination that requires encapsulation to get beyond the next hop, the node can then put a marker in the IP forwarding table indicating the OSI destination address that must be used to encapsulate all IP packets destined for that address.

For each OSI destination that requires encapsulation to get beyond the next hop, the node can then put a marker in the OSI forwarding table indicating the IP destination address that must be used to encapsulate all OSI packets destined for that address.

A node that supports IPv4, IPv6 and OSI may find two addresses (for example an IPv4 address and an IPv6 address) that could be used to encapsulate. In this case, it may choose either as long as it results in a packet that is of a network layer protocol type that the next hop supports (as specified in the "protocols supported" TLV present in IS-IS Hello PDUs received from that node).

II.2 Updates to Dijkstra's Algorithm

The following clauses contain the full Dijkstra's algorithm including extensions to support auto-tunnelling. It is based on the algorithm as specified in RFC 1195. The algorithm shown is suitable for a dual IPv4 and CLNS automatically-encapsulating node. Changes to this algorithm are shown in ***Bold Italic***.

The algorithm produces a PATHS database containing, for each destination, the identity of the first node from S to N capable of unencapsulating IP over OSI, and the identity of the first node from S to N capable of unencapsulating OSI over IP.

For each IP destination, the first node from S to N capable of unencapsulating IP over OSI may have its OSI address loaded into the IP forwarding table as the destination address to be used in any CLNP packet used to encapsulate IP over OSI, if the next hop does not support IP.

For each OSI End System, the first node from S to N capable of unencapsulating OSI over IP may have one of its IP addresses loaded into the OSI forwarding table as the destination address to be used in any IP packet used to encapsulate OSI over IP, if the next hop does not support OSI.

II.2.1 Changes to database

The PATHS and TENTS database should be updated to contain an extension to the {Adj(N)}, element of the triple. The adjacency N element will contain two corresponding Dual Protocol Support (IDP(N)-ODP(N)) entries which will represent the System ID of the first Dual router on the path from S to N capable of de-encapsulating IP over OSI tunnelled packets (IDP(N)) and the System ID of the first dual router on that path from S to N capable of de-encapsulating OSI over IP tunnelled packets (ODP(N). If no *DP(N) router exists on the PATH, then this value will be set to zero. If multiple Adj(N) entries exist in either the TENTS or the PATHS database, then each adjacency will have corresponding *DP(N) entries. Thus, each triple will take the format <N,d(N),{Adj(N)-IDP(N)-ODP(N)}>

If the value of IDP(N) is set to 0, then this means that no dual router exists on the path to the destination capable of de-encapsulating and encapsulating IP over OSI packets.

If the value of ODP(N) is set to 0, then this means that no dual router exists on the path to the destination capable of de-encapsulating and encapsulating OSI over IP packets.

II.2.2 Changes to Algorithm

The SPF algorithm specified in section C.1.4 of RFC 1195 of is amended to appear as follows:

```
Step 0: Initialize TENT and PATHS to empty. Initialize tentlength to
[internalmetric=0, externalmetric=0].

    (tentlength is the pathlength of elements in TENT that we are
    examining.)

    1) Add <SELF,0,W-0-0> to PATHS, where W is a special value indicating
       traffic to SELF is passed up to internal processes (rather than
       forwarded).

    2) Now pre-load TENT with the local adjacency database (each
       entry made to TENT must be marked as being either an End System,
       or a router, to enable the check at the end of Step 2 to be made
       correctly - Note that each local IP reachability entry is
       included as an adjacency, and is marked as being an End System).
       For each adjacency Adj(N) (including level 1 OSI Manual
```

Adjacencies, or Level 2 OSI enabled reachable addresses, and
IP reachability entries) on enabled circuits, to system N of
SELF in state "Up" compute:

d(N) = cost of the parent circuit of the adjacency (N),
obtained from metric.k , where k = one of {default metric,
delay metric, monetary metric, error metric}

Adj(N)***-IDP(N)-ODP(N)*** = the adjacency number of the adjacency to N***,***
the SID of the next-hop router along the path to the neighbour capable of de-encapsulating IP over OSI packets, and the SID of the next-hop router along the path to the neighbour capable of de-encapsulating OSI over IP packets . In this case, i.e., during initialization, both DP values will be set to 0

3) If a triple <N,x,{Adj(M)-***IDP(N)-ODP(N)***}> is in TENT, then:

If x = d(N), then {Adj(M)***-IDP(N)-ODP(N)***} <--- {Adj(M)***-IDP(M)-ODP(M)***}
U {Adj(N)***-IDP(N)-ODP(N)***}.

4) If N is a router or an OSI End System entry, and there are now more adjacencies in {Adj(M)} than maximumPathSplits, then remove excess adjacencies as described in Clause 7.2.7 of ISO/IEC 10589. If N is an IP Reachability Entry, then excess adjacencies may be removed as desired. This will not effect the correctness of routing, but may eliminate the determinism for IP routes (i.e., IP packets still follow optimal routes within an area, but where multiple equally good routes exist, will not necessarily follow precisely the route that any one particular router would have anticipated).

5) If x < d(N), do nothing.

6) If x > d(N), remove <N,x,{Adj(M)***-IDP(M)-ODP(M)***}> from TENT and add the triple
<N,d(N),{Adj(N) ***-IDP(N)-ODP(N)***}>.

7) If no triple <N,x,{Adj(M)***-IDP(M)-ODP(M)*** }> is in TENT, then add
<N,d(N),{Adj(N)***-IDP(N)-ODP(N)***}> to TENT.

8) Now add systems to which the local router does not have adjacencies, but which are mentioned in neighbouring pseudonode LSPs. The adjacency for such systems is set to that of the designated router. Note that this does not include IP reachability entries from neighbouring pseudonode LSPs. In particular, the pseudonode LSPs do not include IP reachability entries.

9) For all broadcast circuits in state "On", find the pseudonode LSP for that circuit (specifically, the LSP with number zero and with the first 7 octets of LSPID equal to LnCircuitID for that circuit, where n is 1 (for Level 1 routing) or 2 (Level 2 routing)). If it is present, for all the neighbours N reported in all the LSPs of this pseudonode which do not exist in TENT add an entry <N,d(N),{Adj(N)***-IDP(N)-ODP(N)***}> to TENT, where:

d(N) = metric.k of the circuit.
Adj(N) = the adjacency number of the adjacency to the DR.

10) Go to Step 2.

Step 1: Examine the zeroeth link state PDU of P, the system just placed on PATHS (i.e., the LSP with the same first 7 octets of LSPID as P, and LSP number zero).

1) If this LSP is present and the "Infinite Hippity Cost" bit is clear
For each Adj(*) – IDP(*) – ODP(*) pair in the PATHS database for P.
If this is not a pseudo-node LSP and if IDP(*) is equal to zero then check the unencapsulation capability field of the LSP, if it supports IP over OSI then set the IDP(P) value for this adjacency to be the system ID of P.
if ODP(*) is equal to zero then check the unencapsulation capability field of the LSP, if it supports OSI over IP then set the IDP(P)value for this adjacency to be the system ID of P

2) If this LSP is present, and the "Infinite Hippity Cost" bit is clear, then for each LSP of P (i.e., all LSPs with the same first 7 octets of LSPID and P, irrespective of the value of SP number) compute:

dist(P,N) = d(P) + metric.k(P,N)

for each neighbour N (both End System and router) of the system P. If the "Infinite Hippity Cost" bit is set, only consider the End System neighbours of the system P.

Note that the End Systems neighbours of the system P includes IP reachable address entries included in the LSPs from system P. Here, d(P) is the second element of the triple

<P,d(P),{Adj(P)***-IDP(P)-ODP(P)***}>

and metric.k(P,N) is the cost of the link from P to N as reported in P's link state PDU.

3) If dist(P,N) > MaxPathMetric, then do nothing.

4) If <N,d(N),{Adj(N)***-IDP(N)-ODP(N)***}> is in PATHS, then do nothing.

NOTE – d(N) must be less than dist(P,N), or else N would not have been put into PATHS. An additional sanity check may be done here to ensure that d(N) is in fact less than dist(P,N)

5) If a triple <N,x,{Adj(N)***-IDP(N)-ODP(N)***}> is in TENT, then:

a) If x = dist(P,N), then {Adj(N)***, IDP(N)-ODP(N)***} <-- {Adj(N)***-IDP(N)-ODP(N)***} U {Adj(P)***-IDP(P)-ODP(P)***}.
Note that even if the value of Adj(N) is equal to the value Adj(P) but the corresponding values of either IDP(P) or ODP(P) and IDP(N) or ODP(N) are different then this should be treated as a different adjacency and will represent a different path to the destination.

b) If N is a router or an OSI end system, and there are now more adjacencies in {Adj(N)} than maximumPath Splits, then remove excess adjacencies, as described in clause 7.2.7 of ISO/IEC 10589. For IP Reachability Entries, excess adjacencies may be removed as desired. This will not effect the correctness of routing, but may eliminate the determinism for IP routes (i.e., IP packets will still follow optimal routes within an area, but where multiple equally good routes exist, will not necessarily follow precisely the route that any one particular router would have anticipated).

c) if x < dist(P,N), do nothing.

d) if x > dist(P,N), remove <N,x,{Adj(N)***-IDP(N)-ODP(N)***}> from TENT, and add <N,dist(P,N),{Adj(P)***-IDP(P)-ODP(P)***}>

6) if no triple <N,x,{Adj(N)}> is in TENT, then add <N,dist(P,N),{Adj(P)}> to TENT.

Step 2: If TENT is empty, stop. Else:

1) Find the element <P,x,{Adj(P)***-IDP(P)-ODP(P)***}>, with minimal x as follows:

a) If an element <*,tentlength,*> remains in TENT in the list for tentlength, choose that element. If there are more than one elements in the list for tentlength, choose one of the elements (if any) for a system which is a pseudonode in preference to one for a non-pseudonode. If there are no more elements in the list for tentlength, increment tentlength and repeat Step 2.

b) Remove <P,tentlength,{Adj(P)***-IDP(P)-ODP(P)***}> from TENT.

c) Add <P,d(P),{Adj(P)***-IDP(P)-ODP(P)***}> to PATHS.

d) If this is the Level 2 Decision Process running, and the system just added to PATHS listed itself as Partition Designated Level 2 Intermediate system, then additionally add <AREA.P,d(P),{Adj(P)}> to PATHS, where AREA.P is the Network Entity Title of the other end of the Virtual Link, obtained by taking the first AREA listed in P's LSP and appending P's ID.

e) If the system just added to PATHS was an end system, go to step 2. Else go to Step 1.

NOTE – In the Level 2 context, the "End Systems" are the set of Reachable Address Prefixes (for OSI), the set of Area Addresses with zero cost (again, for OSI), plus the set of IP reachability entries (including both internal and external).

Appendix III

Commissioning guide for SDH NEs in dual RFC 1195 environment and impact of automatic encapsulation option

III.1 Introduction

This appendix provides guidance on installing Integrated IS-IS nodes in a dual IPv4 and OSI network, and on how to use the optional automatic encapsulation feature described in Annex B.

III.2 Integrated IS-IS without automatic encapsulation

III.2.1 Introduction and Rules of RFC 1195

Integrated IS-IS, as specified in RFC 1195, was originally written as a dual routing protocol. Specifically, it was written to be able to route both IPv4 and CLNP using a single SPF calculation, a single set of metrics for both IP and CLNP, and, a single set of Hellos and LSPs.

More specifically, Integrated IS-IS routers conforming to RFC 1195 calculate shortest paths across a Level-1 area or Level-2 subdomain without considering whether any candidate router can actually forward a specific type of packet.

This is clearly stated in RFC 1195 in section 3.10:

– "The Dijkstra computation does not take into consideration whether a router is IP-only, OSI-only, or dual. The topological restrictions specified in section 1.4 ensure that IP packets will only be sent via IP-capable routers, and OSI packets will only be sent via OSI-capable routers."

With Integrated IS-IS, a router is just a router. The assumption is that any router in the network can handle any type of packet that is thrown at it.

Therefore, Integrated IS-IS routers calculate routes, and forward packets based on this assumption, and it is the responsibility of an operator to make sure that the assumption is actually true.

Thus, there are the topological restrictions of RFC 1195. Failure to enforce the topological restrictions of RFC 1195 may result in packet loss, as packets disappear into the black-hole of a router that simply discards packets that it cannot forward, as it does not support them.

In a simple single Level-1 area network, the rules are quite simple. These are:

1) If IPv4 packets are to be forwarded in an area, then all of the routers in the area must be able to forward IPv4 packets.

2) If CLNP packets are to be forwarded in an area, then all of the routers in the area must be able to forward CLNP packets.

3) If both IPv4 and CLNP packets are to be forwarded in an area, then all of the routers in the area must be dual, i.e., able to forward both.

Thus, it is fairly easy to classify IS-IS Level-1 areas into the classes "OSI-only area", IP-only area", and "Dual area". This is shown in Figure III.1.

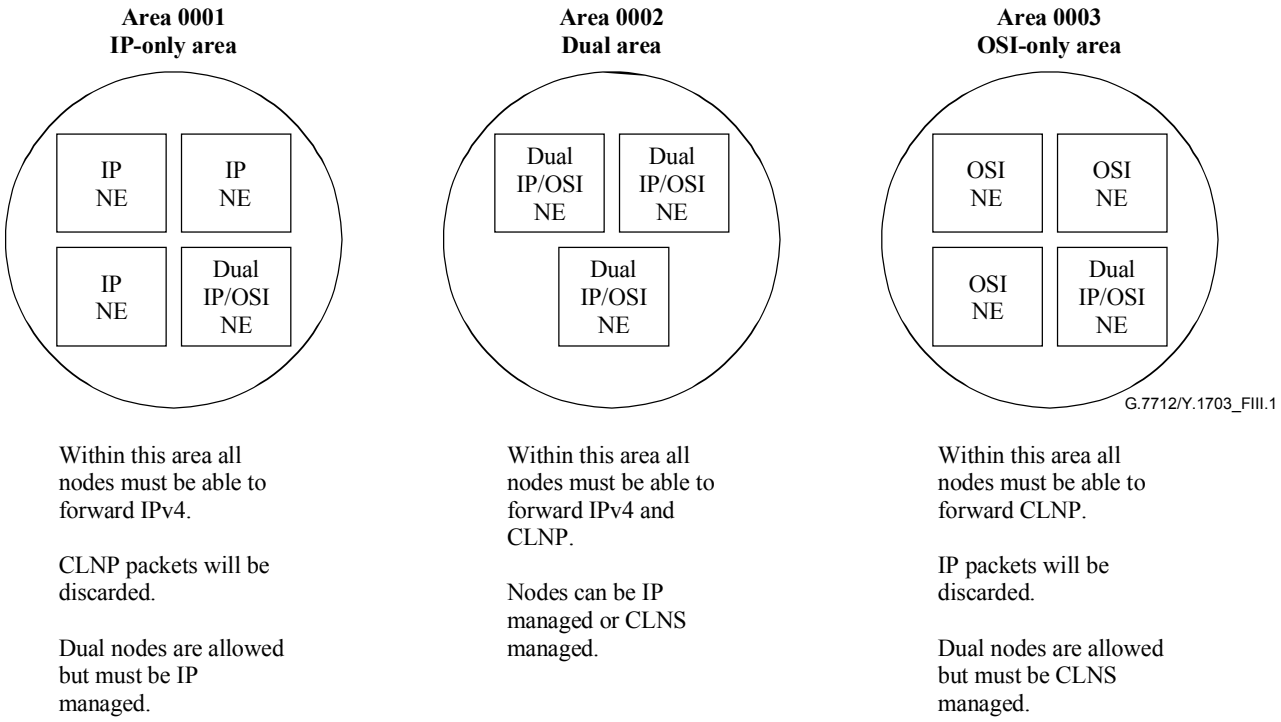


Figure III.1/G.7712/Y.1703 – Classification of IS-IS Level-1 areas

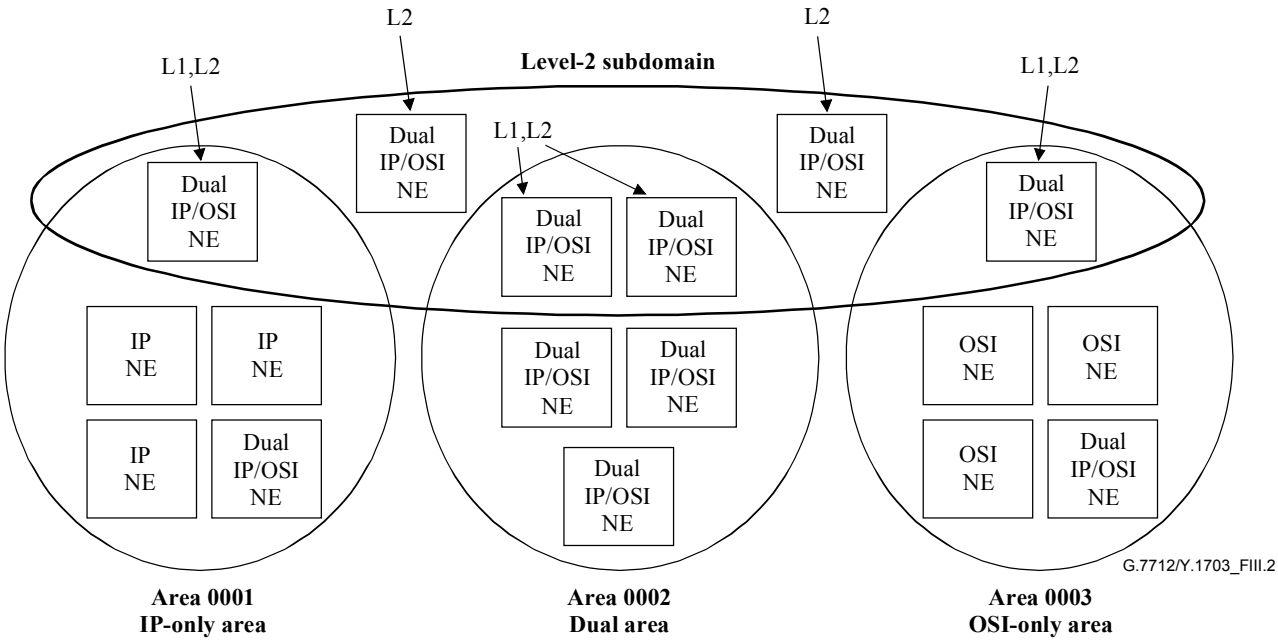
III.2.2 Level-2 subdomain

If a larger network is needed, requiring Level-2 routing, then the Level-2 subdomain forwards packets between the Level-1 areas and, thus, must support all of the types of packets present in all of those Level-1 area. The rules for the Level-2 subdomain are:

1) If IPv4 packets are forwarded in any of the areas (IP-only or Dual areas), then all of the routers in the Level-2 subdomain must be able to forward IPv4.

2) If CLNP packets are forwarded in any of the areas (OSI-only or Dual areas), then all of the routers in the Level-2 subdomain must be able to forward CLNP.

Therefore, if any of the areas are dual, or if both OSI-only and IP-only areas exist, then, the routers in the Level-2 subdomain must be dual. This is illustrated in Figure III.2.



As both IPv4 and CLNP are forwarded within the Level-1 areas, all of the nodes in the Level-2 subdomain must be dual, even those present in IP-only or OSI-only areas.

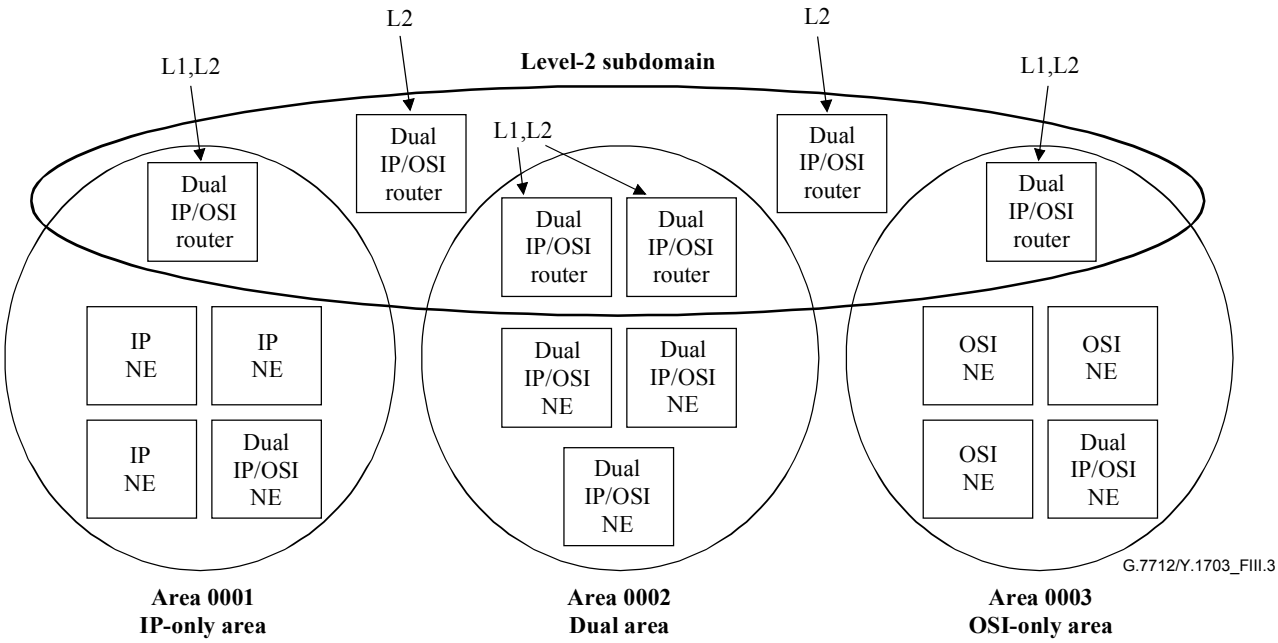
A node is in the Level-2 subdomain if it runs Level-2 routing.

Figure III.2/G.7712/Y.1703 – Level-2 subdomain

III.2.3 Level-2 subdomain with external routers running integrated IS-IS

Many operators currently run Level-1 IS-IS routing in their OSI-only SDH NEs, and then link up multiple areas using Level-2 IS-IS routing in an external router network.

If an operator wishes to use a similar model for a dual network, then they can run Level-1 Integrated IS-IS in each area, and Level-2 Integrated IS-IS in an external router network. This gives a very similar network to the previous one, as shown in Figure III.3



As both IPv4 and CLNP are forwarded within the Level-1 areas, all of the routers in the Level-2 subdomain must be dual, even those present in IP-only or OSI-only areas.

Figure III.3/G.7712/Y.1703 – Level-2 IS-IS routing in an external router network

III.2.4 External routers running OSPF or other IP routing protocols

Many operators currently run Level-2 IS-IS in their external routers, and OSPF, or other routing protocols, for IP. In this case, the external router must remain as the Level-2 router for the SDH NEs, and so, for a dual area, must be a dual Integrated IS-IS router. However, the router may be configured to route all IP packets using OSPF by configuring redistribution of IP routes between IS-IS and OSPF. In this way, all IP packets will be OSPF routed, whilst CLNP packets continue to be Level-2 IS-IS routed. This is shown in Figure III.4.

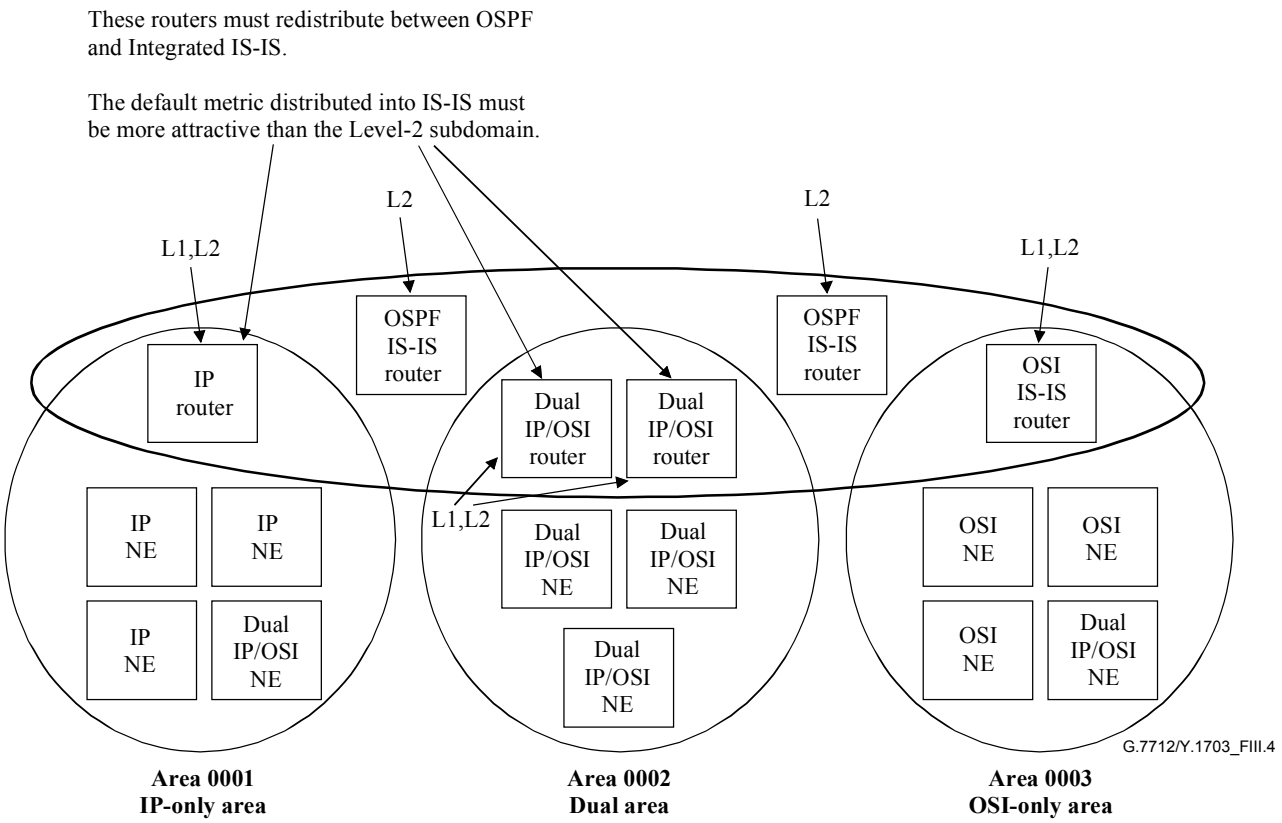


Figure III.4/G.7712/Y.1703 – External routers running OSPF

Note that the Integrated IS-IS stack in the external routers will not be aware that the Level-2 subdomain is meant only for CLNP packets. The OSPF learned routes must, therefore, be redistributed into Integrated IS-IS with a low default metric, to make them more attractive to IP packets than the Level-2 subdomain.

III.3 Integrated IS-IS with automatic encapsulation

III.3.1 Introduction and effect on topological restrictions

The automatic encapsulation option allows the topological rules of RFC 1195 to be broken. Automatic encapsulation effectively makes a node, or group of nodes, appear to be able to forward packets that, intact, they cannot.

This is shown in Figure III.5.

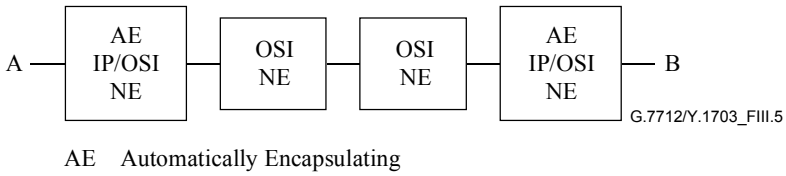


Figure III.5/G.7712/Y.1703 – Group of nodes with automatic encapsulation

This group of nodes will now forward both IPv4 and CLNP packets, as long as the packets enter at point A or B, through one of the automatically encapsulating nodes.

The group of nodes may now safely be put into a dual area, or a dual Level-2 subdomain, as the pair of automatically encapsulating nodes will forward IPv4 packets by encapsulating them inside CLNP packets, so that they will be forwarded by the OSI-only NEs rather than being discarded.

A valid dual area may now look something like that shown in Figure III.6.

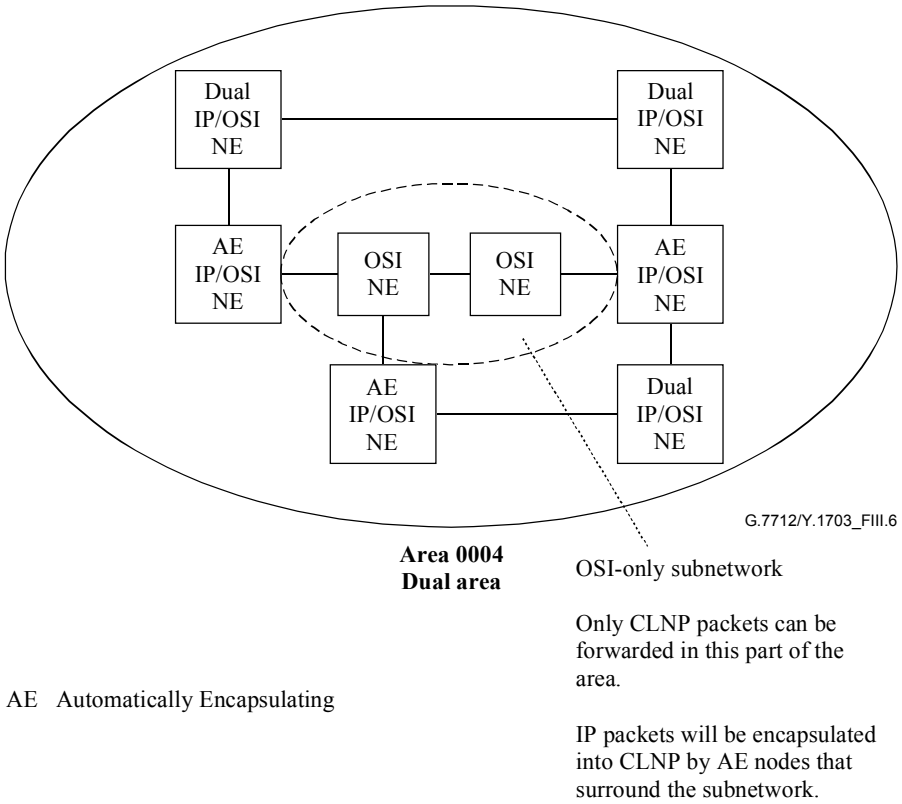


Figure III.6/G.7712/Y.1703 – Example of a valid dual area

Note that the OSI-only nodes must not be directly connected to one of the dual nodes that do not have the automatic encapsulation option. It is only the presence of the automatic encapsulating nodes that prevent IPv4 packets from being sent to an OSI-only node.

A dual node may be connected directly to an OSI-only node if it is also treated as an OSI-only node, as shown in Figure III.7.

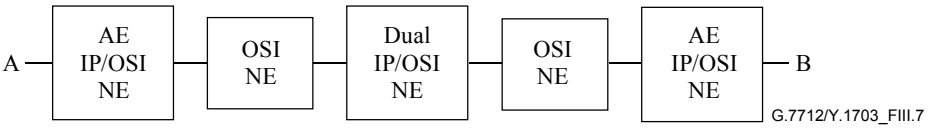


Figure III.7/G.7712/Y.1703 – Connection of a dual node to an OSI-only node

In this case, the network acts as a dual network for packets going from point A to B, but IPv4 packets cannot reach the central dual node. This dual node is inside an OSI-only subnetwork. This dual node will be able to forward CLNP packets only, and must be CLNS managed. There must be no other connections to the central dual node, as, if IPv4 packets were introduced at the central node, then they might be forwarded to an OSI-only node and be discarded.

III.3.2 Getting IP traffic in and out of the SDH embedded network

III.3.2.1 IP capable gateway NE

Both IP and CLNP packets must be able to enter and leave a dual area, whether or not automatic encapsulation is used. Normally traffic enters and leaves an IS-IS area via Level-1, Level-2 routers. These are routers that participate both in the Level-1 area and in the Level-2 subdomain.

The simplest way to build this is to ensure that any Level-1, Level-2 routers are dual, as shown in Figure III.8.

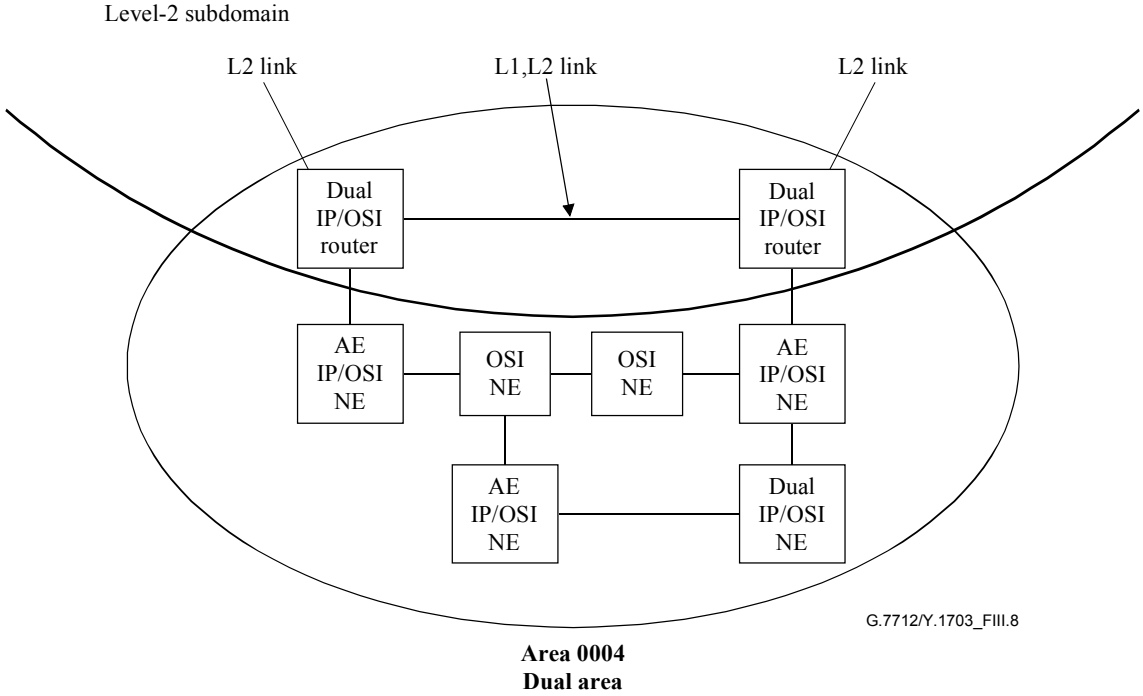
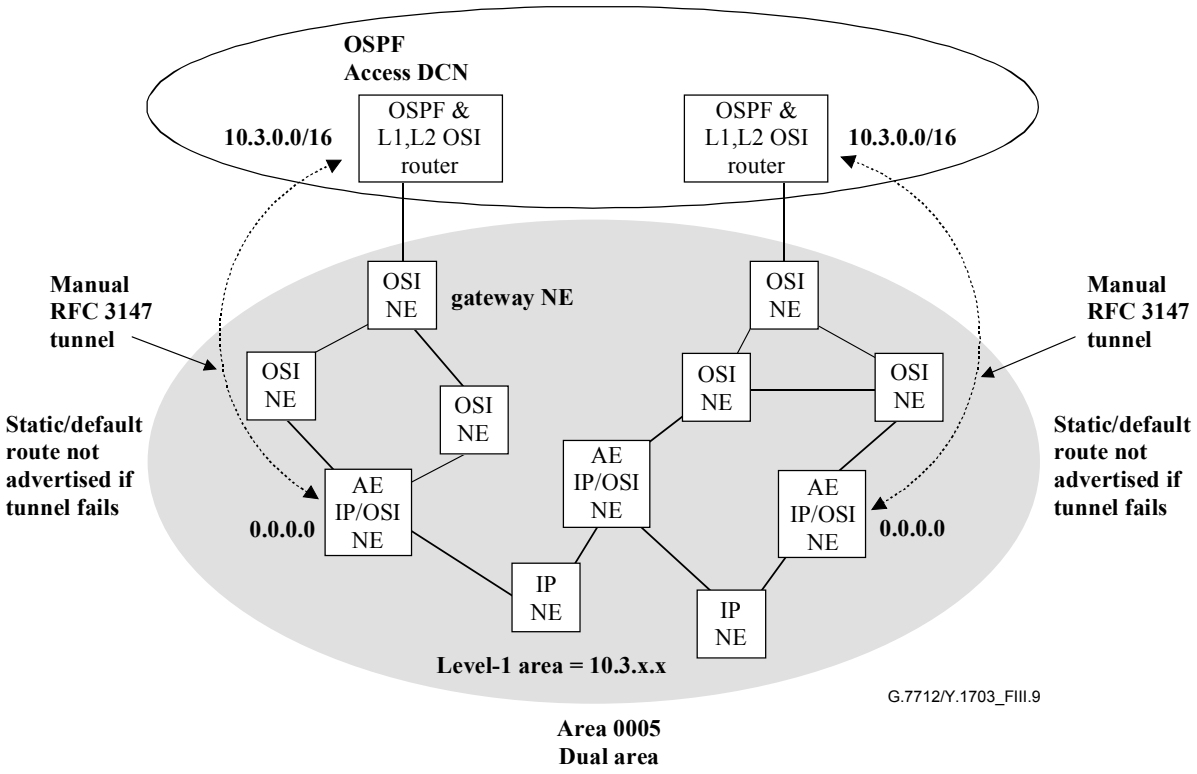


Figure III.8/G.7712/Y.1703 – Dual gateway

III.3.2.2 OSI-only gateway NE

Occasionally, automatically encapsulating nodes will be used to upgrade an existing OSI-only area to make it effectively into a dual area. In this case, the gateway nodes may have to remain as OSI-only nodes. In such a case, a network can be built as shown in Figure III.9.



AE Automatic Encapsulating

Figure III.9/G.7712/Y.1703 – OSI-only gateway

In this network, the CLNP packets that need to leave the Level-1 area continue to go to the OSI Level-1, Level-2 router. The nodes that have a manual tunnel leading out of the Level-1 area advertise this as a default route. Consequently, the IP-capable nodes will all add an entry to the bottom of their routing table telling them to send all IPv4 packets to one of the nodes that has the manual tunnel, unless they have a more specific route. In this way an IPv4 packet is never sent to a Level-1, Level-2 node, but is always sent across one of the manual tunnels.

The router in the Access DCN that terminates the manual tunnel does not need to run Integrated IS-IS. It may run any IP routing protocol that an operator wishes to use. In this way, an existing network that uses OSPF and Level-2 IS-IS in the Access DCN, and Level-1 IS-IS in the SDH NEs, may have the Level-1 areas upgraded to dual areas with little impact on the existing OSI-only SDH NEs, or on the Access DCN.

Appendix IV

Example illustration of packet 1+1 protection

IV.1 Packet 1+1 protection overview

Packet 1+1 path protection provides a packet level protection service similar in some respects to the conventional connection level 1+1 service, with several important distinctions. Packet level 1+1 allows selection of incoming packets from any connection, irrespective of the connection from which the last packet was selected. That is, packet 1+1 protection treats both connections as working connections, as opposed to designating one connection as working, and the other as the protection. In the latter, packets are selected from the working connection until a detection of failure on the working connection causes a switching to the protection connection. In contrast, packet 1+1

does not require explicit failure detection and protection switching. This allows the packet level 1+1 scheme to recover from any failure instantaneously and transparently. Similar to the connection level 1+1 protection, only edge nodes need to be service-aware, which makes interoperability easier.

To provide packet 1+1 protection service between two connection-oriented network edge nodes, a pair of connections is established along disjoint paths. Packets from an application flow subscribing to the service are dual-fed at the ingress node onto the two connections. Disjoint paths, in the simplest case, may be link or node disjoint but, in general, may involve more complicated notion such as shared risk groups. At the egress edge node, one of the two copies of the packets selected and forwarded from the two possible received copies, each traversing a disjoint path. Given this, any single failure in the network, other than the ingress or egress node itself, can affect at most one copy of each packet. This allows the service to withstand a single failure transparently. In terms of restoration time, this can be characterized as an instantaneous recovery from a failure since there is no need to detect, notify and switch to protection path explicitly. The scheme can be easily extended to protect against multiple failures by employing more than two disjoint paths.

IV.2 Packet 1+1 protection illustration

Figure IV.1 illustrates a realization of the service using sequence numbers as identifiers. After passing through the classifier, each packet that needs to be forwarded on the mated LSPs is assigned a distinct sequence number by the service-aware source edge node. This packet with the distinct identification is then duplicated and forwarded onto the two disjoint LSPs. The egress node shall only select one copy of the duplicated packet. For appropriately selecting the packet exactly once, the destination must be able to identify the duplicate packets and then select one, and handle all possible variations. This selection process at the packet level is non-trivial as the duplicate packets may not arrive at the same time (due to propagation delay and buffering) and also these packets may get lost (due to transmission errors and buffer overflows).

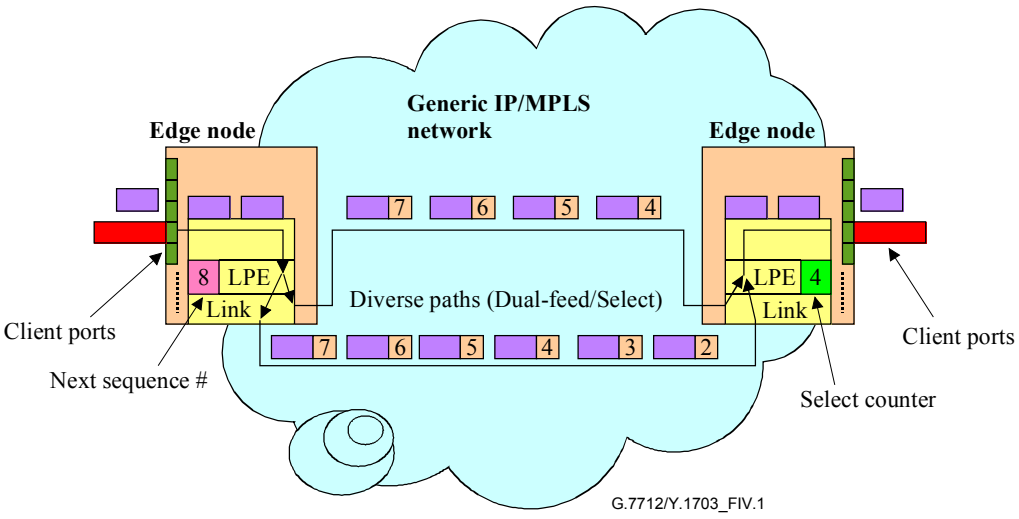


Figure IV.1/G.7712/Y.1703 – 1+1 protection

The ingress node inserts the sequence number as defined in 7.1.19.2. The packet is then duplicated and transported over diverse LSPs. Due to the diversity of the LSPs, there will be a leading LSP and a trailing LSP. The leading LSP will deliver the packets to the egress node faster than the trailing LSP. Therefore, under non-failure conditions, the egress node will select the packets from the leading LSP. The packets received on the trailing LSP will be duplicate packets and will therefore, be discarded.

The decision whether to accept or discard a received packet is based on the received packet's sequence number and a counter + sliding window at the egress node. The counter indicates the sequence number of the next packet it is expecting. The counter, plus sliding window, provides a window of acceptable sequence numbers. The sliding window is needed to properly accept and reject packets. If the received packet falls in the window, it is considered legitimate and can be accepted. Otherwise, it is rejected. The size of the window should be larger than the maximum number of consecutive packets a working (an alive) LSP can lose.

The sliding window is used to solve the problem of losing packets on the leading LSP when the leading LSP's sequence number is very close to the wrap around point. Figure IV.2 illustrates a leading LSP (LSP 1) that delivers a packet with sequence number 29. The packet is accepted and the counter is incremented to 30. If we assume that 2 consecutive packets are lost (i.e., packets with sequence numbers 30 and 31), the next received packet on LSP 1 will be 0. Without a sliding window, the egress node will reject the packet since 0 < 30. By implementing a sliding window that is larger than the maximum number of consecutive packets a working (an alive) LSP can lose, this problem can be solved. For example, let's say that the maximum number of consecutive packets that a working LSP can loose is 5, then a sliding window of 6 can be defined. Taking the same example as before, however, now using the sliding window, the egress node will accept packets in the range of {30, 31, 0, 1, 2, 3, 4}. Therefore, even if 5 packets are lost (i.e., the maximum number of consecutive packets that can be lost on a working LSP) the next packet received will have sequence number 3 and the packet will be accepted.

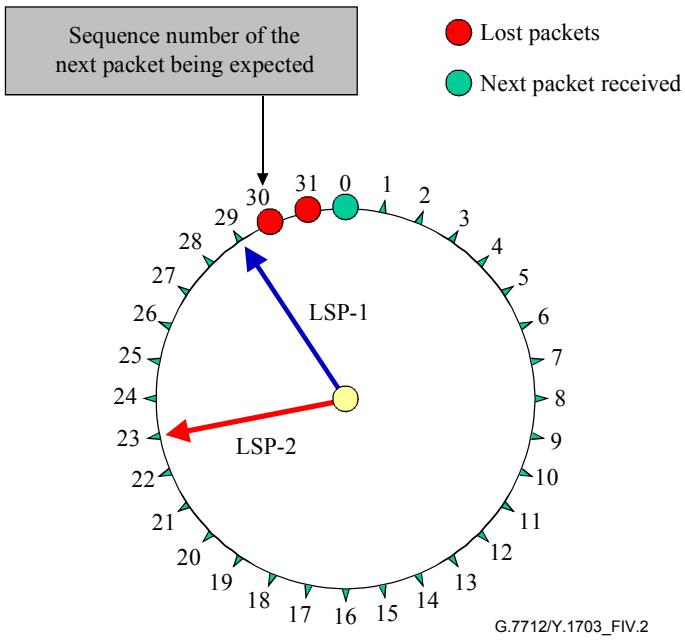


Figure IV.2/G.7712/Y.1703 – Sliding window mechanism

Note that this idea of sliding window only works if the falling behind LSP cannot fall back in the sliding window range. If a packet with a sequence number in the range of the sliding window is received from the falling behind LSP, then it will be mistakenly accepted. A falling behind LSP can only receive a packet with a sequence number in the range of the sliding window if it falls back by more than (2^N – size of sliding window). Therefore the number of bits "N" used for the sequence number must support the following equation:

$$2^N > \text{SlidingWindow} + \text{DelayWindow}$$

where;

SlidingWindow > maximum number of consecutive packets that can be lost on a LSP

and

DelayWindow = maximum number of packets the trailing LSP can fall behind the leading LSP

Note that 7.1.19.2 defines a 4-byte field for carrying the sequence number. The 4-byte field provides a sequence of more than 4 billion numbers which is large enough to accommodate worst-case consecutive packet losses and delay differentials.

One reasonable way of engineering the size of the sliding and delay windows is to make the size of the sliding window equal to the size of the delay window. (Note that it is assumed that the size of the delay window is generally larger than the size of the sliding window.) This guarantees selection of packets from the leading LSP in all scenarios after a failed LSP gets repaired. This point is further elaborated in the following clause which discusses various failure scenarios.

IV.3 Operation of selector algorithm under various failure scenarios

One way to view the operation of the selector algorithm is to picture a clock with 2^N intervals. Figure IV.3 illustrates an example where N = 4 (i.e., 4-bit sequence number) and, therefore, the sequence number ranges from 0 through 15.

In this example, the SlidingWindow is set equal to the DelayWindow, which is 5.

Figure IV.3 shows the Leading LSP ahead of the Trailing LSP by 3 sequence numbers. The Leading LSP delivers a packet with sequence number = 1 and the counter is now set to 2.

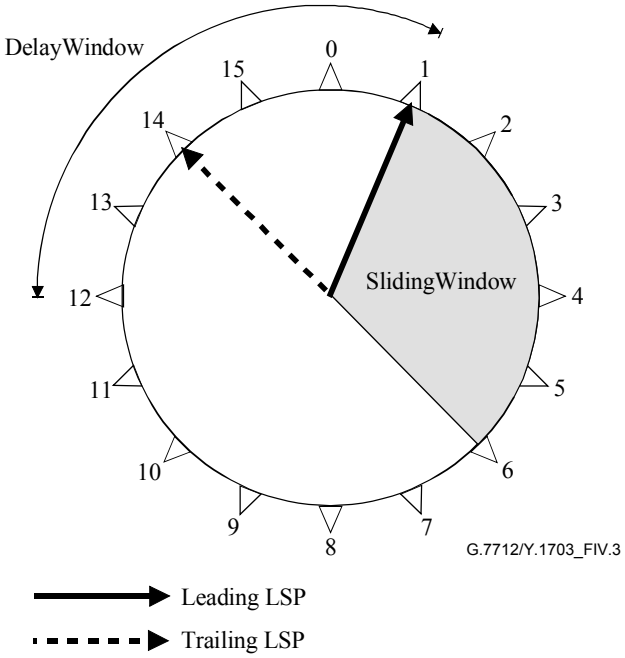


Figure IV.3/G.7712/Y.1703 – Selector algorithm operation

Figure IV.4 shows that, prior to receiving a packet with sequence number equal to 2 on the Leading LSP, the Leading LSP fails. Until the packet with sequence number equal to 2 is delivered from the Trailing LSP, the egress node will not select any packets and the counter will remain equal to 2.

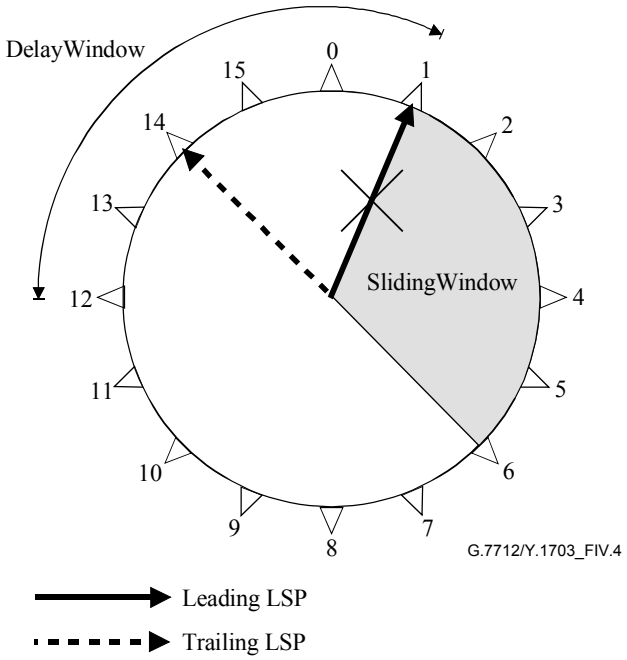


Figure IV.4/G.7712/Y.1703 – Leading LSP failure

Figure IV.5 illustrates that, when the packet with sequence number equal to 2 is received on the Trailing LSP, the egress node increments the counter to 3 and the sliding window shifts so that a packet with sequence number in the range of 3 through 7 can be accepted.

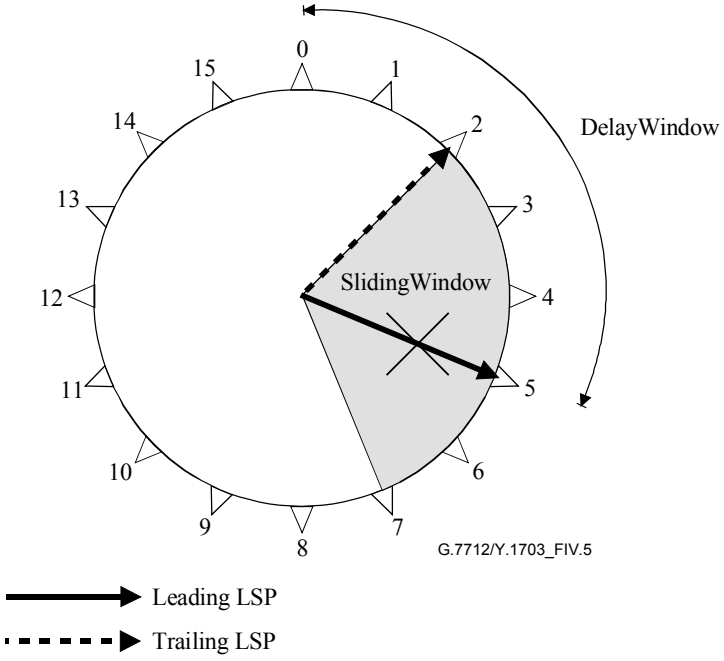


Figure IV.5/G.7712/Y.1703 – Reception of packet 2 by trailing LSP

Figure IV.6 illustrates that, prior to receiving a packet with sequence number equal to 3 from the Trailing LSP, the Leading LSP is repaired and a packet with sequence number equal to 6 is received from the Leading LSP. Since 6 is within the sliding window range, the packet is accepted. Note that it is important that, so long as the Leading LSP is working, packets are received from the Leading LSP. Therefore, to ensure that, when the Leading LSP is repaired, that it delivers a packet with a

sequence number value that is within the sliding window range, the SlidingWindow should be equal to or greater than the DelayWindow, which is the case for this example.

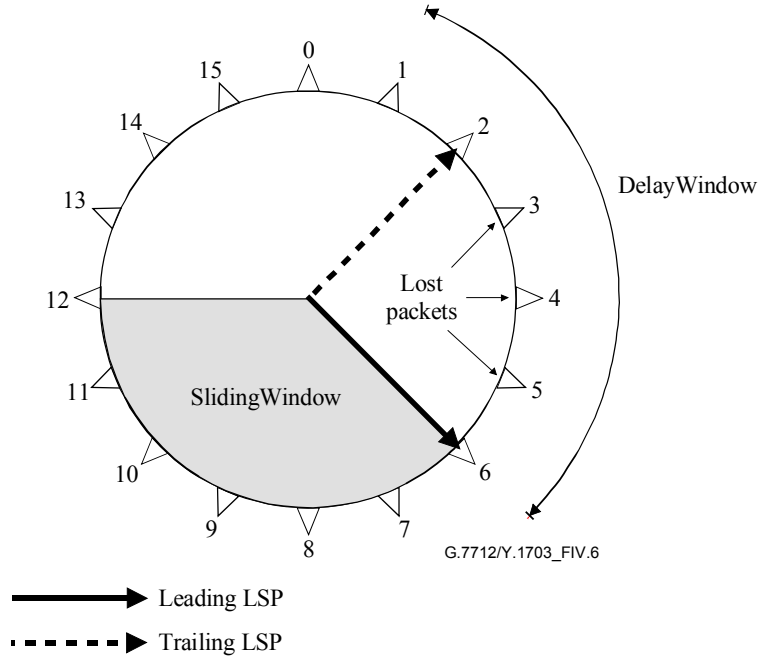


Figure IV.6/G.7712/Y.1703 – Leading LSP repaired

Figures IV.7, IV.8 and IV.9 illustrate a problem if the SlidingWindow is set smaller than the DelayWindow. In this case, it is possible that, when the Leading LSP is repaired, it delivers packets with sequence numbers that fall outside the SlidingWindow and, therefore, the egress node continues to accept packets from the Trail LSP. If, at a later time, the Trailing LSP fails, there is a potential to lose many packets (worst case would be 2^N – *size_of_sliding_window*, where N is the number of bits used for the sequence number).

Figure IV.7 shows an example where the SlidingWindow is set to 3, while the DelayWindow can be up to 7. In this example, the Trailing LSP trails the Leading LSP by 4 sequence numbers. Since the Leading LSP is failed, the packets are selected from the Trailing LSP.

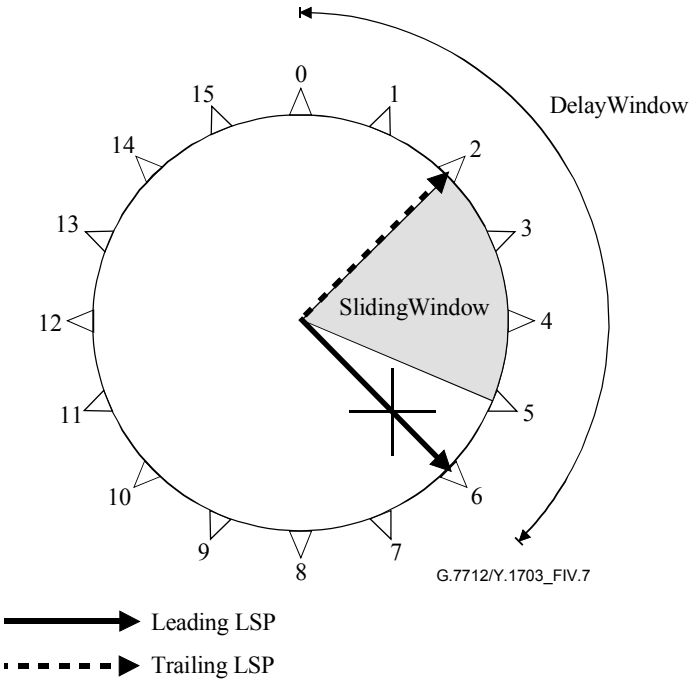


Figure IV.7/G.7712/Y.1703 – Sliding window too small: packets selected from the trailing LSP

Figure IV.8 illustrates that, at the time when the Leading LSP is repaired, it delivers a packet with sequence number equal to 7 which is outside the SlidingWindow and, therefore, rejected. The packets continue to be selected from the Trailing LSP.

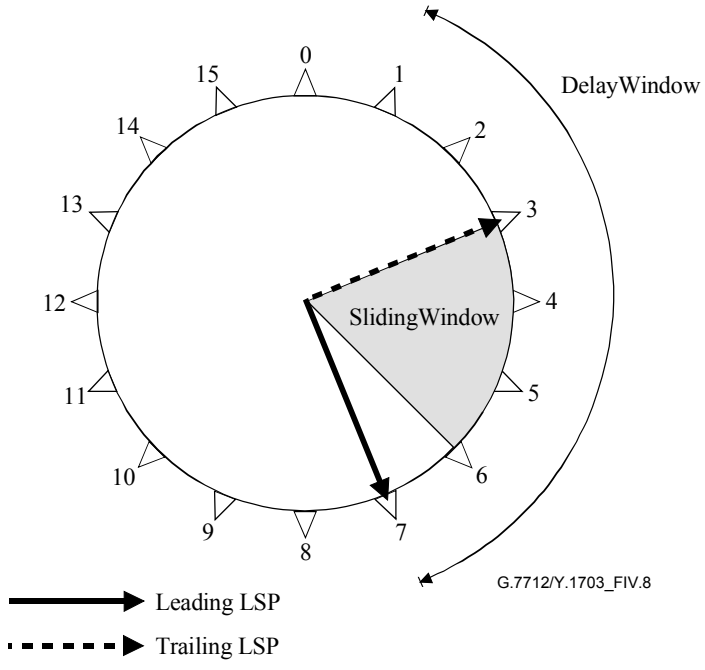


Figure IV.8/G.7712/Y.1703 – Sliding window too small: rejection of packets delivered by the repaired leading LSP

Figure IV.9 illustrates a failure to the Trailing LSP. Since the Leading LSP delivers packets outside the SlidingWindow and, therefore, those packets are rejected, the egress node will not start accepting packets until the Leading LSP comes all the way around and starts to deliver packets with a sequence number that falls within the SlidingWindow. This can result in a significant loss of packets. Therefore, to prevent such an occurrence, it is recommended that this type of selector algorithm set the SlidingWindow equal to the DelayWindow.

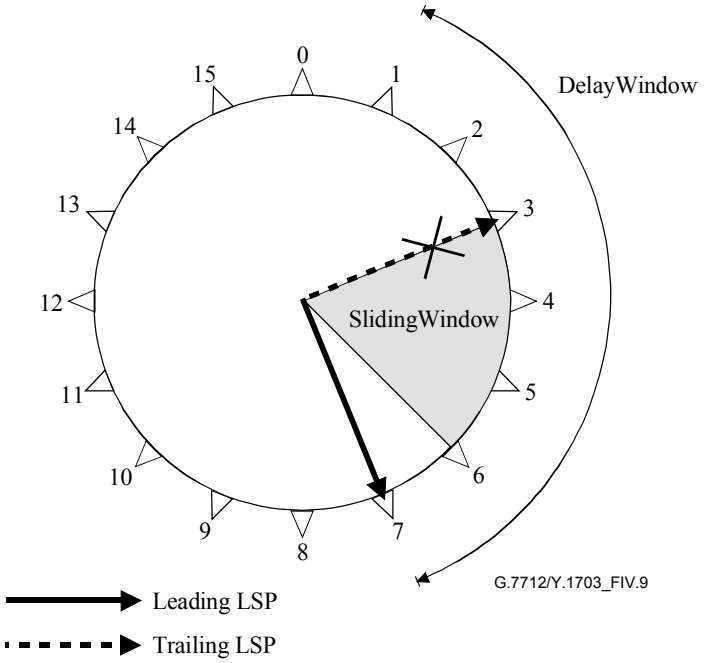


Figure IV.9/G.7712/Y.1703 – Sliding window too small: effect of a failure for trailing LSP

Appendix V

Bibliography

– IETF RFC 1006 (1997), *ISO Transport Service on top of the TCP Version 3*.

– IETF RFC 2966 (2000), *Domain-wide Prefix Distribution with Two-Level IS-IS*

– IETF RFC 3147 (2001), *Generic Routing Encapsulation of CLNS Networks*.

– IETF RFC 3373 (2002), *Three-Way Handshake for Intermediate System to Intermediate System (IS-IS) Point-to-Point Adjacencies*.

ITU-T Y-SERIES RECOMMENDATIONS

GLOBAL INFORMATION INFRASTRUCTURE AND INTERNET PROTOCOL ASPECTS

GLOBAL INFORMATION INFRASTRUCTURE	
General	Y.100–Y.199
Services, applications and middleware	Y.200–Y.299
Network aspects	Y.300–Y.399
Interfaces and protocols	Y.400–Y.499
Numbering, addressing and naming	Y.500–Y.599
Operation, administration and maintenance	Y.600–Y.699
Security	Y.700–Y.799
Performances	Y.800–Y.899
INTERNET PROTOCOL ASPECTS	
General	Y.1000–Y.1099
Services and applications	Y.1100–Y.1199
Architecture, access, network capabilities and resource management	Y.1200–Y.1299
Transport	Y.1300–Y.1399
Interworking	Y.1400–Y.1499
Quality of service and network performance	Y.1500–Y.1599
Signalling	Y.1600–Y.1699
Operation, administration and maintenance	**Y.1700–Y.1799**
Charging	Y.1800–Y.1899

For further details, please refer to the list of ITU-T Recommendations.

SERIES OF ITU-T RECOMMENDATIONS

Series A	Organization of the work of ITU-T
Series B	Means of expression: definitions, symbols, classification
Series C	General telecommunication statistics
Series D	General tariff principles
Series E	Overall network operation, telephone service, service operation and human factors
Series F	Non-telephone telecommunication services
Series G	**Transmission systems and media, digital systems and networks**
Series H	Audiovisual and multimedia systems
Series I	Integrated services digital network
Series J	Cable networks and transmission of television, sound programme and other multimedia signals
Series K	Protection against interference
Series L	Construction, installation and protection of cables and other elements of outside plant
Series M	TMN and network maintenance: international transmission systems, telephone circuits, telegraphy, facsimile and leased circuits
Series N	Maintenance: international sound programme and television transmission circuits
Series O	Specifications of measuring equipment
Series P	Telephone transmission quality, telephone installations, local line networks
Series Q	Switching and signalling
Series R	Telegraph transmission
Series S	Telegraph services terminal equipment
Series T	Terminals for telematic services
Series U	Telegraph switching
Series V	Data communication over the telephone network
Series X	Data networks and open system communications
Series Y	**Global information infrastructure and Internet protocol aspects**
Series Z	Languages and general software aspects for telecommunication systems

Printed in Switzerland
Geneva, 2003