SIF-018-1998

(SIF Document #SIF-AR-9701-001)

SIF APPROVED DOCUMENT

WORK GROUP: Architecture

TITLE: Guidelines for SONET DCN Architecture Engineering

DATE: August 13, 1998

EDITOR: **Name:** Corey Sharman
Voice: (972) 479-3720
email: corey.sharman@fnc.fujitsu.com

ABSTRACT: This document contains a compilation of the contributions for the SIF guidelines for SONET DCN Architecture Engineering document as of June 17, 1998.

NOTICE: This document has been prepared to assist SIF. It is offered to SIF as a basis for discussion and is not a binding proposal on Fujitsu Network Communications, Inc.. The requirements presented in this document are subject to change in form and numerical value after more study. Fujitsu Network Communications, Inc. specifically reserves the right to add to or amend the statements contained herein. Nothing contained herein shall be construed as conferring by implication, estoppel, waiver, or otherwise any license or right under any patent, whether or not the use of information herein necessarily employs an invention or any existing or later issued patent.

This document has received the approval of the SONET Interoperability Forum (SIF).

August 13, 1998

This document has received the approval of the SONET Interoperability Forum (SIF).

August 13, 1998

TABLE OF CONTENTS

TABLE OF FIGURES

This document has received the approval of the SONET Interoperability Forum (SIF).

August 13, 1998

TABLE OF TABLES

1

Introduction

This SIF report is intended to provide guidelines to SONET network providers to use in engineering their SONET DCN architectures. It focuses primarily on OSI routing issues, however, it considers the use of other protocols (e.g. TCP/IP) as well.

1.1 Background

The SONET DCN was originally envisioned to consist of the SONET DCC interconnected to CO LANs and X.25 WANs via GNEs. However, as SONET networks continue to grow, network providers are realizing the need to migrate their X.25 WANs to higher speed data networks using subnetwork technologies such as Frame Relay and/or ATM and using WAN routing protocols such as OSI or IP. The consequence of this need is that the WAN portion of the DCN becomes more and more complex to design and administer. The future WAN may consist of protocol conversion gateways and/or OSI routers. If OSI routers are present, they will need to interact with the OSI routers in the DCC and LAN portions of the DCN, thus impacting the performance and design constraints across the entire DCN.

1.2 Purpose

The purpose of this report is to meet the following objectives:

1. Describe DCN components and routing methods.
2. Identify reference architectures, including points of interface with the LAN and DCC portions of the DCN.
3. Provide guidelines for DCN design and router placement, as well as describe the interaction between WAN OSI routers, the DCC, and LAN routers.
4. Provide guidelines for routing area management, including guidelines for assigning OSI routing areas (e.g., maximum size of an area), OSI routing domains (e.g., maximum size of a domain), splitting areas, and level 1partition repair.
5. Assess the need for OSI inter-domain routing; if this is determined to be needed, then SIF should define requirements for the use of the OSI inter-domain routing protocol in SONET NEs and/or stand-alone routers.
6. Provide guidelines for DCN parameter provisioning.
7. Provide guidelines for name to address translation.

1.3 Scope

The entire SONET DCN (DCC, LAN and WAN) is within the scope of this report. The scope of protocols to be considered for the DCC and LAN is limited to the OSI protocols. The scope of protocols to be considered for the LAN/WAN is at a minimum the OSI protocols, however, other protocols may be considered as well (e.g. TCP/IP). When other protocols are considered, however, the point (and mechanism) of interworking with the OSI protocols must be clearly described.

1.4 Acronyms

ACSE	Association Control Service Element
AE	Application Entity
AP	Application Process
ATM	Asynchronous Transfer Mode
BIS	Boundary Intermediate System
BLSR	Bidirectional Line Switched Ring
CIT	Craft Interface Terminal
CLNP	Connectionless Network Protocol
CMISE	Common Management Information Service Element
CO	Central Office
DAP	Directory Access Protocol
DCC	Data Communications Channel
DCE	Data Communications Equipment
DCN	Data Communications Network
DIB	Directory Information Base
DIT	Directory Information Tree
DN	Distinguished Name
DSA	Directory System Agent
DS-n	Digital Signal
DUA	Directory User Agent
ENE	End-system Network Element
EMS	Element Management System
ES	End System
ES-IS	End System to Intermediate System protocol
ESH	End System Hello PDU
FTAM	File Transfer Access and Management
FTP	File Transfer Protocol
GNE	Gateway Network Element
IDRP	Inter-Domain Routing Protocol
IEEE	Institute of Electrical and Electronic Engineers
IIH	Intermediate-system to Intermediate-system Hello PDU
INE	Intermediate-system Network Element
IP	Internet Protocol
IS	Intermediate System
IS-IS	Intermediate System to Intermediate System protocol
ISH	Intermediate System Hello PDU
ISO	International Organization for Standardization
Kbps	Kilo-bits per second

L1	Level 1 Routing
L2	Level 2 Routing
LAN	Local Area Network
LAPB	Link Access Procedure Balanced
LAPD	Link Access Procedure for the D channel
LDAP	Lightweight Directory Access Protocol
LLC	Logical Link Control
LNC	Local Naming Context
LSDB	Link State Data Base
LSP	Link State PDU
Mbps	Mega-bits per second
MD	Mediation Device
MAC	Media Access Control
NARSE	Network Address Resolution Service Element
NE	Network Element
NEM	Network Element Manager
NET	Network Entity Title
NIST	National Institute of Standards and Technology
NLR	Network Layer Relay
NMS	Network Management System
NPDU	Network Protocol Data Unit
NSAP	Network Service Access Point
OAM	Operations, Administration and Maintenance
OS	Operations System
OSI	Open Systems Interconnect
OSIE	Open Systems Interconnect Environment
OSPF	Open Shortest Path First
PS	Packet Switch
PVC	Permanent Virtual Circuit
PDIS	Partition Designated Intermediate System
PDU	Protocol Data Unit
RA	Registration Agent
RBOC	Regional Bell Operating Company
RDN	Relative Distinguished Name
RM	Registration Manager
SDCC	Section Data Communications Channel
SDH	Synchronous Digital Hierarchy
SDS	SONET Directory Service
SIF	SONET Interoperability Forum
SNDCF	SubNetwork Dependent Convergence Functions
SONET	Synchronous Optical NETwork
SPF	Shortest Path First
STS-n	Synchronous Transfer Signal

SVC	Switched Virtual Circuit
T5	NE with a TARP processor and a DUA+RA function reserved to its own use
T5GW	System having a TARP processor and a DUA+RA function and capable of acting on behalf of TARP NEs
TARP	TID Address Resolution Protocol
TCP	Transmission Control Protocol
TID	Target IDentifier
TPn	Transport Protocol type n=0..4
TPDU	Transport Protocol Data Unit
UL	Upper Layers
UPSR	Unidirectional Path Switched Ring
WAN	Wide Area Network

1.5 References

ANSI T1.204-1997, *Telecommunications - Operations, Administration, Maintenance, and Provisioning (OAM&P) - Lower Layer Protocols for Telecommunications Management Network (TMN) Interfaces between Operations Systems and Network Elements*

ANSI T1.245-1995, *Directory Service for TMN and SONET*

ANSI X.3.216-1992, *Information Processing Systems - Data Communications - Structures and Semantics of the Domain Specific Part (DSP) of the OSI Network Service Access Point (NSAP) Address*

GR-199-CORE, *Operations Application Messages - Memory Administration Messages*, Issue 2 (Bellcore, November 1996)

GR-253-CORE, *Synchronous Optical Network (SONET): Common Generic Criteria, (A Module of TSGR, FR-440),* Issue 2, Revision 1 (Bellcore, December 1995), Section 8.

ISO/IEC 7498-1:1994: *Information Technology – OSI Reference Model*

ISO/IEC 8473:1988: *Information processing systems - Data communications - Protocol for providing the connectionless-mode network layer service*

ISO/IEC DIS 8473 1:1993: *Information technology - Protocol for providing the connectionless-mode network service*

ISO/IEC 9542:1988: *Information processing systems - Telecommunications and information exchange between systems - End system to Intermediate system routing exchange protocol for use in conjunction with protocol for providing the connectionless-mode Network Service (ISO 8473)*

ISO/IEC 9594-1: 1993:*Information Technology – The Directory: Overview of Concepts, Models and Services*

ISO/IEC 10589:1992: *Information technology - Telecommunications and information exchange between systems - Intermediate system to Intermediate system intra-domain routing information exchange protocol for use in conjunction with the protocol for providing the connectionless-mode Network Service (ISO 8473)*

ISO 10747:1994, *Information Technology - Telecommunications and Information Exchange between Systems - Protocol for Exchange of Inter-Domain Routing Information among Intermediate Systems to support forwarding of ISO 8473 PDUs*

ITU-T Recommendation G114, *Transmission Systems and Media – General Characteristics of International Telephone Connections and International Telephone Circuits,* February 1996

ITU-T Recommendation M3010, *Maintenance: Telecommunications Management Network, Principles for a Telecommunications Management Network*, October 1992

RFC1006, *ISO Transport Services on Top of the TCP*, Version 3, May 1987

SIF-DN-9706-045: *CIT/OS Software Platform Architecture Specification (approved)*

SIF-013-1997: *Directory Services and Network Address Resolution for CIT/OS Systems*

SIF-015-1997: *Lightweight Directory Access Protocol: ANSI T1.245*
SIF-004-1996:TARP/500: *The TARP and X.500 Directory Services Interworking Specification*

Tanenbaum, Andrew S., *Computer Networks – Third Edition*, 1996 Prentice Hall Inc., ISBN 0-13-349945-6

2 The Data Communications Network (DCN)

The term “Data Communications Network”, DCN, is one which has two definitions, a common usage one and an ITU-T M.3010 definition. These definitions are very similar, differing only in the architectural boundaries of DCN, and the extent of the subnetworks involved. However, even though the differences in the definitions are small, it is important to understand them as they may be significant in detailed technical discussions or documents.

In general, a data communications network provides management communications capabilities between elements of a TMN. That is, the DCN is a communications network between OSs, NEs, MDs, etc. The DCN may consist of a single or multiple subnetworks; this aspect of its architecture is determined by the agency deploying the network. The subnetworks need not be homogenous; a variety of different technology-specific subnetworks may be included in the DCN.

2.1 DCN Architectural Boundary Definition Differences

With respect to architectural boundaries, the common and loosely defined definition of DCN is that it includes all protocol stack layers.

The more restrictive and precisely defined definition of the term “DCN” is that of ITU-T M.3010, which restricts the DCN boundary to the network layer of the communications stack and places the layers above the network layer in a functional component called the Management Communications Function. The M.3010 definition of DCN states:

> “The DCN represents an implementation of OSI layers 1 to 3, which may include any relevant ITU-T or ISO standards for layers 1 to 3. The DCN provides no functionality at layers 4 to 7.”

This first release of the DCN guidelines uses the M.3010 definition of DCN architectural boundaries rather than the common definition.

2.2 DCN Subnetwork Definition Differences

The common definition of DCN with respect to the SONET TMN is often to exclude the Local Communications Networks (LCN) and the DCC subnetworks from the DCN.

However, from the M.3010 point of view, the LCN and DCC are merely subnetworks within the DCN, so in comparison, the M.3010 definition is more expansive and encompassing than the common definition of a DCN.

In this first release of the DCN guidelines the M.3010 definition of DCN subnetworks is used.

2.3 DCN Definition Comparison

The following table illustrates the differences between the common usage and M.3010 definition of DCN.

Table 1 DCN Definition Comparison

Difference	Common Usage	M.3010 Usage
Architectural Boundary	all stack layers included	DCN includes layers 1 to 3
Subnetwork inclusion	LCN and DCC excluded	LCN and DCC included

3 Overview of OSI Routing

This section provides a brief overview of OSI routing including terminology, routing hierarchy, addressing, PDU types, and routing levels.

3.1 Routing Terminology

OSI Networks consist of two types of systems:

- End Systems (ES)
- Intermediate Systems (IS)

ISs perform the relay function whereas ESs do not.

There are two types of ISs:

- Level 1 IS

  An OSI IS that performs IS-IS Level 1 routing, routes PDUs to a destination within the area or to the nearest Level 2 IS for destinations outside the area.

- Level 2 IS

  An OSI IS that performs IS-IS Level 2 routing, routes inter-area traffic. Level 2 ISs also perform Level 1 routing for the area they are in.

3.2 Topological Hierarchy

3.2.1 Routing Area

A routing area (often called a “Level 1 Area”) is the smallest grouping of OSI systems for routing purposes, and is a neighborhood of interconnected ESs and ISs. Area boundaries are somewhat arbitrary, but should be drawn so as to group systems that communicate with one another, as much as possible (reasons for this appear below).

3.2.2 Routing Domain

A routing domain (also referred to as a “level 2 area”) is a collection of routing areas, all of which must follow the same routing policies. All systems in the domain must have the same System ID length. All ISs in the domain must follow identical policies, and must not treat any systems in the domain preferentially to others (some notable SONET exceptions to this rule are discussed below).

3.2.3 Administrative Domain

An administrative domain is a set of routing domains whose addresses and routing policies are assigned and administered by the same authority. It exists only "on paper"; none of the routing protocols have any specific knowledge of it (in contrast to routing areas and domains.) A set of routing domains that are run by the same interexchange carrier, for example, would be considered an administrative domain (even if they are not interconnected via OSI.)

3.2.4 Global OSI Environment

The set of all routing domains that are interconnected form a Global OSI Environment (OSIE). The original intent was for there to be a single OSIE, but so far, this has been more of a pipe-dream than a reality.

3.3 Addressing

3.3.1 NSAP / NET

All routing is based on the Network Service Access Point (NSAP) Address or Network Entity Title (NET). The distinction between an NSAP and an NET is academic (the Network services user uses one, and the Network layer uses the other for its own purposes), and there is no difference in format or interpretation. In the following text, the term "NSAP" is used to mean either "NSAP Address" or "NET". For the purposes of IS-IS routing, an NSAP is divided into an Area Address, a System ID, and an N-Selector.

3.3.2 Area Address

The area address is the entire initial part of the NSAP up to but not including the System ID. In GR-253-CORE, a two-byte "Area ID" field is defined, but this only exists for address administration purposes and is treated by IS-IS as simply the last two bytes of the area address.

Each routing area has one or more unique area addresses. IS-IS implementations are required by ISO 10589 to support at least 3. The ability to have multiple area addresses for the same area is very important; it permits joining, splitting, or changing of area boundaries while the network is in use.

3.3.3 System ID

The System ID identifies the system within the area.

3.3.4 N-Selector

The N-Selector is used to select the user of the network layer, for example, Transport vs. TARP. It is ignored by IS-IS in routing decisions.

3.4 Link State Routing

3.4.1 Overview

Through the use of controlled flooding techniques, each IS spreads knowledge of its direct connectivity (e.g., its neighbors) to all other participating ISs. Level 1 routers advertise all of their neighbors to all other ISs in the area, and Level 2 routers advertise all their Level 2 neighbors to all Level 2 routers in the domain, along with their own area IDs. When advertising a link to a neighbor, the IS includes a **metric** value that indicate the link's capacity (and optionally, other metrics indicating expense, delay, or error probability).

As a result, each Level 1 IS has a complete picture of the area, in the form of a set of links. Likewise, each Level 2 IS has a complete picture of the set of links that can be used to join the areas together into a routing domain. Importantly, every IS in an area or domain has precisely the same picture of the area or domain, once the topology has stabilized and the routing information has been disseminated.

Using this picture, each IS traverses the entire graph using Dijkstra's Shortest Path First (SPF) algorithm. It finds, for each destination, the next hop to use on the best path to that destination. Level 1 ISs also find the shortest path to the nearest Level 2 IS. For level 1 routing, the "destination" is a System ID; for Level 2 routing, the "destination" is an area address.

It is necessary for all ISs in the area (for Level 1) or domain (for level 2) to have the same picture of the area or domain. Otherwise, routing loops can occur, because two ISs may each think that the other is on the shortest path to a destination, and will send PDUs for that destination back and forth between themselves until their NPDU lifetimes expire. This can occur after topology changes, but should be resolved when the updated information has been disseminated.

3.4.2 Hello PDUs

Hello PDUs are sent when a link first comes up, and periodically thereafter. Three different types of Hello PDUs are defined (i.e., ES sends ESH, IS sends ISH or IIH). A routing node uses the received Hello PDUs to determine its OSI adjacencies (i.e., routing neighbors). This information is also used to construct the LSP PDUs (see Section 0 for description of LSP PDUs).

3.4.3 Link State PDUs

Link State PDUs (LSPs) are the IS-IS PDUs that are used to advertise the links. Depending on the number of links to be advertised, one or more LSPs are required. Each IS generates and transmits all of its LSPs periodically. Therefore, background LSP traffic must be accounted for as a function of area size when engineering the traffic.

Whenever a topology change occurs (a link goes up or down), the IS or ISs affected regenerate and retransmit their LSPs, which get propagated throughout the area or domain.

Information from the LSP PDUs is stored inside the Link State Data Base (LSDB). The LSDB can be used to derive a topology map of all OSI routing entities (including NEs and routers) within the particular OSI routing area.

3.4.4 Level 1 vs. Level 2 Routing

Level 1 routing (when source and destination are in the same area) is optimal – data always travels on the shortest path from source to destination, provided that the topology is stable.

Level 2 routing (when source and destination are in different areas) is:

1. optimal from source to nearest level 2 IS in source area
2. optimal from Level 2 IS in source area to nearest Level 2 IS in destination area
3. optimal from Level 2 IS in destination area to destination
4. not necessarily optimal overall

As a result, it is usually desirable to engineer areas to place source and destination in the same area for the bulk of the traffic. This is less of a consideration in topologies where there are few alternate paths between areas.

3.5 Inter-Domain Routing

IS-IS supports inter-domain routing by permitting the Level 2 ISs at a domain boundary (called Boundary Intermediate Systems, or BISs) to be manually provisioned with a set of "reachable address prefixes" and the link to use for these prefixes. The BIS is only provisioned with the prefixes for routes to other domains that should be reached through it; the IS-IS protocol disseminates this information to all Level 2 ISs in the domain.

4 General Description of DCN Functions and Components

The following subsections describe the major functions and components of the DCN. Use of each of these components is illustrated in the reference DCN architectures of Section 0.

4.1 Local Area Networks

Local Area Networks (LANs) consist of shared communications media within a building. For most telco installations, LANs consist of twisted pairs (i.e., 10BaseT) connected to bridges, repeaters, and hubs. Thin-wire ethernet (i.e., 10Base2) or ethernet cables can also be used. IEEE 802.3 is the most common layer two protocol that is run over the shared LAN physical media. The ETHERNET protocol or token ring can also be run as the layer two protocol. As per GR-253-CORE, SONET NEs are required to support 10BaseT and IEEE 802.3 as the data link layer protocol.

4.1.1 Bridges

Bridges operate at the physical and data-link layers and are typically used to connect networks that run multiple network-layer protocols over the same data-link layer protocol.

4.1.2 Repeaters

Repeaters operate at the physical layer and simply boost signals from the received LAN segment and pass the boosted signal on to the next segment.

4.1.3 Hubs

The term “hub” is a generic term that refers to a device that incorporates a combination of bridge and repeater functions. Hubs are typically at the center of star topologies (i.e. an ethernet hub is a logical bus, physical star topology).

4.2 Mediation Devices

The term "mediation device" is one which has several usages, and accordingly, network administrators must be careful when using the term and in discussion and specifications, since confusion and misinterpretation can result from these dual usages.

4.2.1 Common Usage of Mediation Device

With respect to management communications, the common use of mediation device includes the interworking of different communications functions. In this sense, the mediation devices are protocol converters. These devices terminate protocol connections of one type and originate protocol connections of another type. For example, a protocol converter can terminate an X.25 connection, extract the user data, and insert this in the user data field of an OSI or a TCP/IP connection. This type of mediation device may be deployed either in central office locations, or within operations centers. It may be a stand alone piece of equipment or it may be integrated into operations systems or network elements.

Additionally, a second common usage of mediation device is that it functionally encompasses both the translation of information models or the translation of legacy NEs to a Q-interface compliant information model. This second type of mediation device may also be deployed in both central office locations and operations centers, and may also be a stand alone piece of equipment or integrated with other devices.

4.2.2 M.3010 Definition of Mediation Device

Like the common usage, ITU-T M.3010 also allows the use of the term "mediation" to describe protocol conversion but limits this use only to the protocol conversion necessary when DCN/DCN interworking occurs at the upper layers. M. 3010 uses the term network layer relay to identify a function that connects two different DCN technologies.

With respect to TMN information architecture, the M.3010 definition of "mediation device" is more restrictive and precisely defined than the common usage of the term.
M.3010 restricts "mediation device" to the translation of information models, only.
M.3010 uses the term "adaptation" (or q-adaptation) for the translation of legacy (non-TMN) NEs to a Q-interface compliant information model; accordingly, A Q-adaptor is the physical device which performs this adaptation function.

In M.3010 terms, when integrated into another physical building block, such as an NE or OS, the mediation or adaptation entity is called a Mediation Function (MF) or a Q-Adaptor Function (QAF).

4.2.3 Comparison of Common Usage and M.3010 Definition

The following table illustrates the differences between the common usage and M.3010 definition of the term mediation device, for both the information architecture and management communications cases.

Table 2 Mediation Device Usage

FUNCTION	COMMON USAGE	M.3010 USAGE
Translation between IMs	Mediation Device	Mediation Device
Translation from Legacy NE to IM	Mediation Device	Q-Adaptor
Protocol Interworking (UL)	Mediation Device	Mediation Device

4.2.4 Usage Recommendation

The SIF Architecture Work Group strongly recommends that all participants use the ITU-T M.3010 definition of mediation (and adaptation), as M.3010 is the architectural standard that SIF has adopted, the M.3010 definitions are more precise, and the M.3010 definition of mediation is the one that is used in SIF documents.

4.3 Router Function in SONET NEs

Standards specify that SONET/SDH equipment support OSI protocols. As such, SONET equipment must support OSI router functionality. This includes routing OSI traffic between DCCs, as well as OSI traffic between the operations interfaces and the DCCs. Please see Section 0 for a further discussion of router functionality.

4.4 Stand-alone Routers

Routers are network-layer devices (although routers can often perform bridging functions as well). Routers provide the capability to send traffic between multiple LAN/WAN segments. Stand-alone routers perform high-speed forwarding of network-layer Protocol Data Units (PDUs). Routers typically can support multiple network-layer protocols (e.g., CLNP and IP). Some stand-alone routers can also perform a protocol tunneling function (refer to Section 0 for a description of this function).

4.5 Transmission Facilities

Transmission facilities represent the communication pipes between DCN equipment. Transmission facilities are used to transport bits through the DCN. Transmission facilities are typically monitored for transmission errors; in many cases a poor performing or failed transmission facility will be protection switched to help prevent the loss of DCN traffic, or to prevent DCN congestion.

4.6 X.25 Packet Switches

X.25 packet switching technology dates back to the mid 1960s, and was designed to operate over low-speed links that have high transmission error rates (e.g., compared with fiber optics). X.25 networks perform hop-by-hop (where a hop is a connection between X.25 packet switches) error detection and correction. X.25 networks represent a large portion of the RBOC legacy DCN network. X.25 links operate at speeds of 1.2 Kbps up to DS3.

4.7 Gateway Network Element (GNE)

A Gateway Network Element is the NE in a collection of interconnected NEs (such as a UPSR or BLSR ring) that connects the collection of NEs to the management systems. For the most part three types of gateways exist: application gateways, transport gateways, and network layer relays.

The term as defined in GR-253-CORE, states that SONET GNEs can provide two distinct functions that support this role: the “CLNS LAN/DCC/WAN router” function (refer to Section 0), a required function; and an optional “TL1 X.25 to OSI Gateway” function (an example of an application gateway, refer to Section 0).

4.8 Application Gateway

Application gateways terminate the underlying protocol stack and the data or application is then delivered over a completely different protocol stack. The “TL1 X.25 to OSI Gateway” function described in GR-253-CORE is an excellent example of such a gateway. This application gateway terminates the X.25 protocol and “maps” the TL1 messages over a full OSI 7-layer stack. Since the application level is where the translation between the two protocol stacks exist, the result is a conversion at the application level. Notice that the underlying protocol stacks on either side of the application gateway do not require the same structure to be present, i.e. 3-layer stack versus 7-layer stack. Figure 1 depicts the application gateway described above.

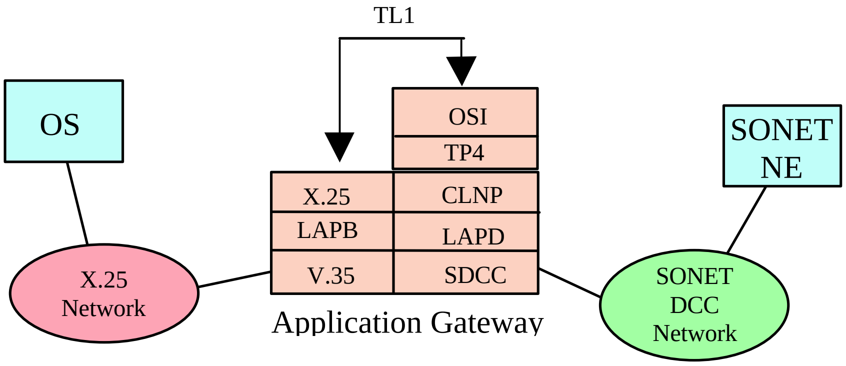


Figure 1 Application Gateway

4.9 Transport Gateway

Transport gateways terminate the underlying transport protocols and performs a translation. The best example of a transport gateway would be an OSI TP4 to TCP transport gateway. While it may seem at first glance that such a translation mechanism would be possible, it is not straight forward and trivial. Other problems arise with this approach such as supporting applications which are native to the specific transport service provided. An example would be running FTAM over OSI TP4 versus running FTP over TCP. If an encapsulation mechanism were used to support FTAM over TCP, the end systems would also need to support such a mechanism. Figure 2 depicts a transport gateway between TCP and OSI TP4.

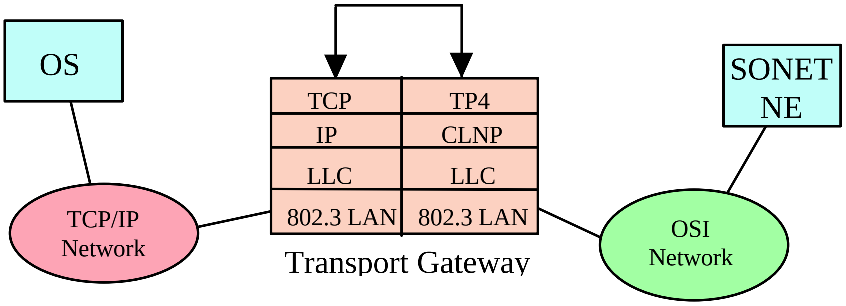


Figure 2 Transport Gateway

4.10 Network Layer Relay

With a network layer relay, the supported upper layers of the protocol stack remain intact but the underlying network service or subnetwork is different. The network layer relay is a more "pure" mechanism for providing gateway functionality between subnetworks which support different protocols. Network layer relays require that a common network layer sub-function exist between the two underlying protocols. An example of this would be running CLNP over X.25 and running CLNP over a 802.3 based LAN. The NLR is more of a routing function than a gateway function between the networks, although routing between points on different subnetworks which provide different services (e.g. connectionless subnetwork vs. connection oriented subnetwork) may require the use of sub-network convergence functions.

Figure 3 depicts an example of the Network Layer Relay between an OSI based X.25 network and the OSI based SONET DCC network. Note that the protocol stacks which support the two different subnetworks provide the same network layer.

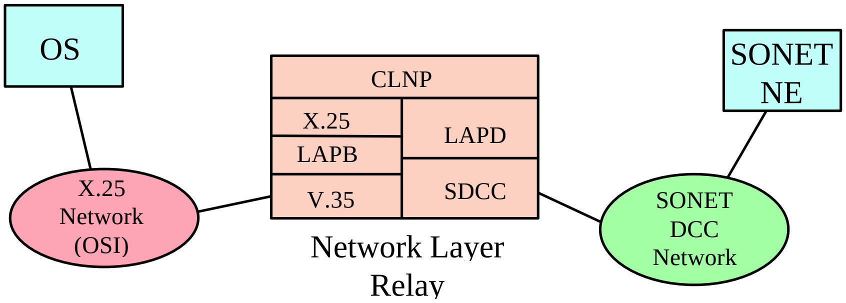


Figure 3 Network Layer Relay

4.11 Transport Service Bridge

Transport service bridges provide protocol conversion from one protocol suite to another, while supporting the same service over both protocols. Transport service bridges differ from transport gateways in focus, in that the applications need not worry about the underlying service changing as the protocol changes due to the interconnection of different protocols. The transport service bridge is an interworking function that maps the transport service primitives of one protocol to that of another protocol. Some of the more common transport service bridges are OSI TP4 to TCP/IP using RFC1006, OSI TP0 over X.25 to TCP/IP using RFC1006, and TP4 over CLNP to TP0 over X.25. The transport service bridge offers the advantage of supporting common applications, such as OSI, over protocol stacks which do not support the same native applications (this is the case using RFC1006 over TCP/IP), and offers a simpler solution to the transport gateway for cases where interworking between connection-oriented and connectionless protocol stacks is required (such as the case for TP4/CLNP to TP0/X.25).

The RFC1006 specification defines a thin transport layer service which when combined with TCP provides identical functionality to the OSI transport service class 4, TP4. The major difference between the TCP and TP4 transport protocols is the notion of a segment/stream oriented data transfer and a octet/packet oriented data transfer. RFC1006, which actually provide the TP0 service over TCP, provides a octet/packet oriented approach for the transport user, or applications.

The transport service bridge mechanism actually requires two transport connections to be established between the end systems which are communicating. In the OSI TP4 to RFC1006/TCP case, a transport connection must be established between the end system in the OSI TP4 network to the transport service bridge, and a second transport connection must be established between the transport service bridge and the end system in the TCP/IP network. The transport service bridge must copy the data packets, or TPDUs, between the two connections, once the two connections have been established, as well as map between primitives for maintaining or aborting the transport connections. In addition to the mapping required between the service primitives of the two protocol stacks, the transport service bridge must maintain a table for address mapping between the two protocols. For the TP4 to RFC1006/TCP case, this may include a TCP port/IP address to the NSAP address in the OSI side of the network. Figure 4 depicts the transport-service bridge between an RFC1006/TCP/IP stack and TP4 OSI based stack for interworking between TCP/IP based subnets and OSI based subnets.

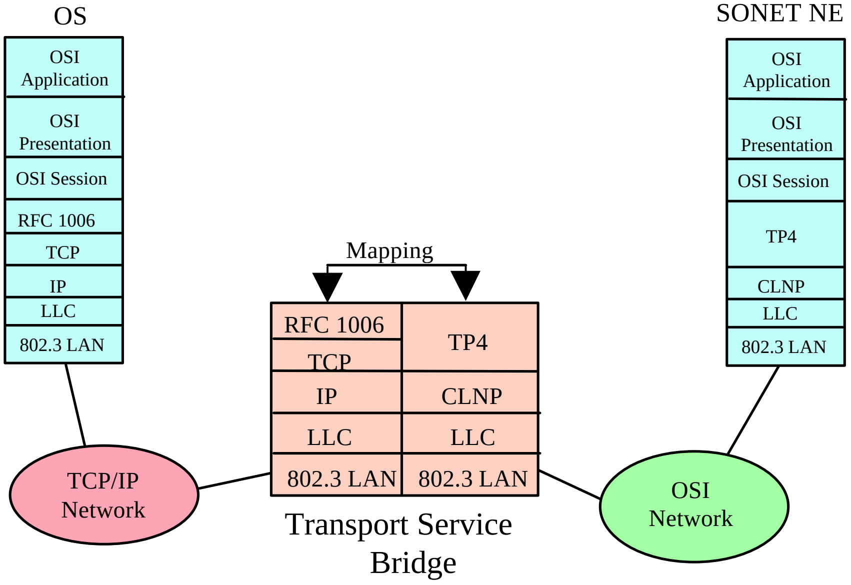


Figure 4 Transport Service Bridge

The transport service bridge offers advantages over the application gateways in that it accommodates multiple applications over multiple subnetwork technologies. For instance, OSI applications such as FTAM or CMISE can span an OSI based network as well as a TCP/IP based network while preserving all semantics of the applications between the communicating end systems. However, the transport service bridge is not without its drawbacks. Since the transport service bridge actually terminates the transport protocol on each side of the bridge, the end-to-end transport connection is not end system to end system. The transport service bridge also serves as a single point of failure, which could be resolved by providing multiple transport service bridges in the network.

5 DCN Routing

In addition to the physical transmission of data, DCNs support various combinations of routing. Routing functions which may appear in a network where multiple protocols and subnetworks are deployed, include application-gateways, transport-gateways, network layer relays, transport-service bridges and network-service tunnels, or encapsulation. While each function aims to solve the problem of multi-protocol interworking, each has certain advantages and disadvantages.

5.1 Native Routing

The function of Network routing is to forward packets based on the layer 3 destination address. In Figure 5 an operational support system is used to monitor and provision the SONET NE. A Network is needed to route the monitoring and provisioning packets to the SONET NEs. In addition, the SONET NE must be able to forward packets across the DCC to downstream SONET NEs. In Figure 5, the network is made up of standalone routers and SONET NE’s with routing functionality. Note: The standalone routers and the SONET NE’s can function as Level 1 IS (Intermediate Systems) or a Level 2 IS. Note that Figure 5 is for explanation purposes only.

An End System (ES) is generally an end user device on a network. The device is a non routing host or node on an OSI network. In other words the device does not have the ability to forward packets. In the context of SIF, the ES is generally an Operational support system that is used to monitor and provision SONET equipment. The end system will establish an adjacency to the Level 1 IS. The ES and IS will discover each other through End System Hello and an Intermediate System Hello. When an end system wants to forward a packet to another end system or a SONET NE , the end system forwards the packets to it’s adjacent Level 1 IS. The IS then looks up the packet destination address and forwards the packet along the best route. If the destination address is a SONET NE or End system in another area the packet is forwarded to the nearest Level 2 IS.

OSI allows for the network to be split up into multiple groups of contiguous End Systems and Intermediate Systems. A grouping is called a area. Within an area all Level 1 ISs communicate with other Level 1 ISs within the area. So a Level 1 IS will know the most efficient route for a packet within the area. A packet destined outside an area will be forwarded by the Level 1 IS to the nearest Level 2 IS. The Level 2 IS communicates with all of the other Level 2 ISs. The level 2 ISs will forward the packet until it reaches the Level 2 IS within the destination area. The packet will then be forwarded along the most efficient path to the ES by the Level 1 IS within the destination area.

Standalone routers are often utilized in large networks because they typically scale better than a SONET NE with IS functionality. The scaling refers to the number of Level 1 devices that can be placed within an area or the number of Level 2 devices within a domain. The number of Level 1 routers within an area is determined by the limitations of the worst performing Level 1 IS. The number of Level 2 ISs within a domain is determined by the performance of the worst performing Level 2 IS. So for example, often in large networks the level 2 IS functionality is left for dedicated routers. The standalone routers generally scale to the largest domain and are the most cost effective solution.

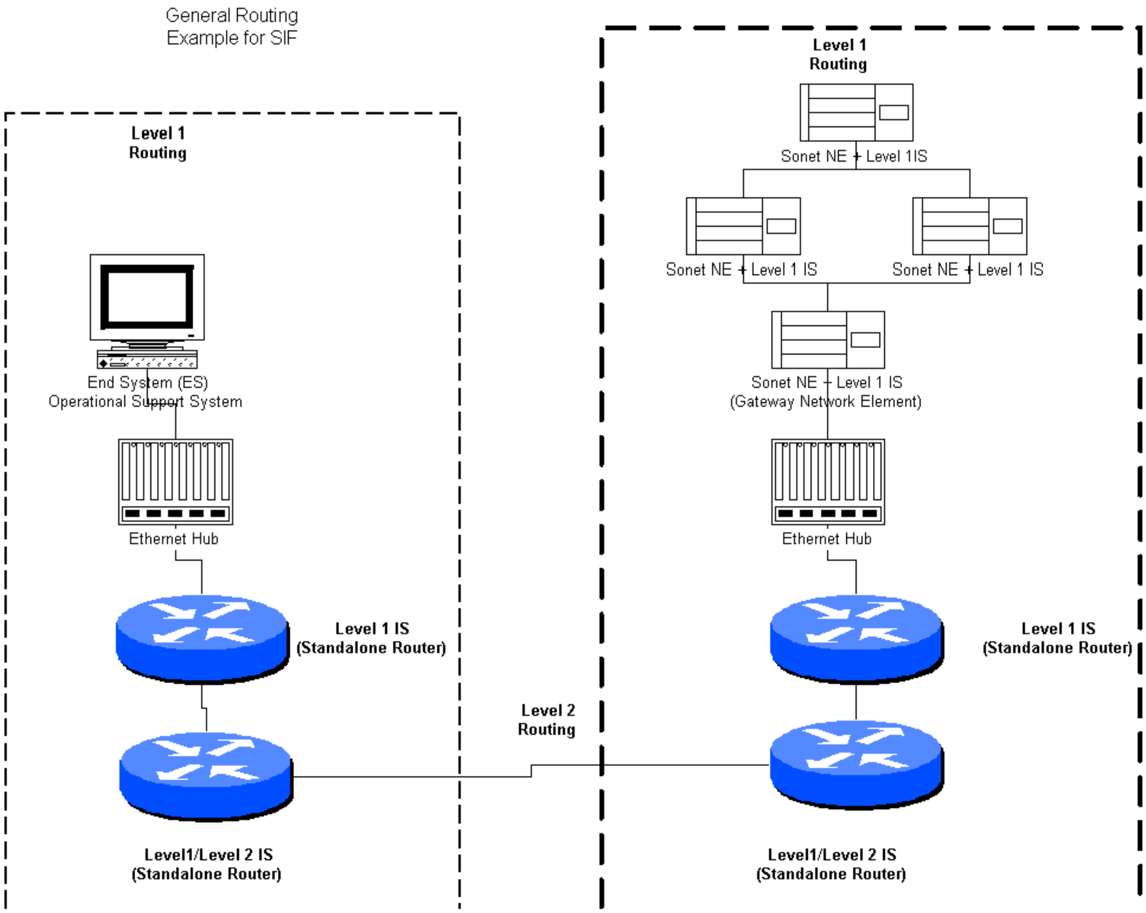


Figure 5 Network Routing

5.2 Dual Stack Approach

An alternative to the protocol conversion methods, is the use of dual stacks within every node in the system. This solution is simple since each protocol stack runs independently. The protocols and applications in the dual stack approach do not interoperate. The dual protocols may traverse the same underlying network facilities, such as LANs, bridges, or routers, while still maintaining separate identities within their respective networks.

The primary reason for using the dual stack approach is that the DCN is often shared with applications other than SONET surveillance. For example, a Central Office LAN may contain SNMP devices, PCs and UNIX workstations as well as SONET GNEs. In this case the commercial routers are often shared for routing all of this traffic onto a common "backbone" network.

Although the answer is simple, the implementation of this type of approach requires the use of additional routing related resources. Each system would require two of everything, including protocol stacks, addresses, and network management, to name a few. In addition, the cost to the each system with regards to CPU power, and memory resources would rise.

Note that most commercial routers that are available today can run dual stacks.

5.3 Encapsulation and Tunneling

Encapsulation and tunneling are methods for conveying protocol data units across a network that uses an incompatible protocol, permitting two otherwise disjoint networks to be joined using the services of the otherwise incompatible network. For this discussion, we will call the first the “original” network, and the second the “transit” network. At one end of the transit network, a tunneling or encapsulating router sends the original PDUs as user data of the transit protocol to a similar router at the other end of the transit network. This second router extracts the original PDUs and forwards them using the original protocol. When this is done using a connection-oriented service of the transit network, it is typically called “tunneling”, whereas when it is done using a connectionless service of the transit network, it is usually called “encapsulation”. Examples that could be used in a SONET context are:

- CLNP / IP: CLNP Encapsulation within IP packets
- CLNP / TCP: CLNP Tunneling using TCP connections
- IP / CLNP: IP Encapsulation within CLNP packets
- IP / TP4: IP Tunneling using TP4 connections
- CLNP / X.25: CLNP tunneling using X.25 connections (also called the CLNP X.25 Subnetwork Dependent Convergence Functions (SNDCF), ISO 8473-2.)

There are several potential advantages to tunneling and encapsulation. They can be used to permit the use of existing IP networks to join physically separate OSI networks, or vice versa.

No special routing or addressing considerations apply to the transit network, however, special care and attention to routing details is necessary for the original protocol. The communicating systems (e.g., OS and NE) would operate normally, with no need to be aware of the existence of the transit network.

CLNP / IP encapsulation is typically less efficient than CLNP / TCP, because each CLNP PDU incurs additional IP protocol overhead, and may also require additional (IP or CLNP) segmentation. ISIS does not have a subnetwork dependent convergence function (SNDCF) that would be appropriate for use across the transit network, therefore, ISIS “reachable address prefixes” would need to be provisioned on each encapsulating router, in addition to the IP address of each tunneling router that can be used to reach each prefix. CLNP / TCP Tunneling has the disadvantage over native routing that each tunneling router would need to be manually provisioned with the IP address of the other tunneling routers that it would communicate with.

CLNP / IP, CLNP / TCP, IP / CLNP, and IP / TP4 have the disadvantage that there are currently no applicable standards for interoperability.

In general, protocol tunneling is a useful mechanism for the initial deployment of a new protocol in the DCN. However, for large-scale deployment of a new protocol (such as OSI), it is recommended that the new protocol be deployed in its native mode, because of the configuration and interoperability issues discussed above.

6 Reference DCN Architectures

This section of the document illustrates various DCN architectures and the interworking of the protocols required to support the particular DCNs and the SONET subnetworks. Various hybrid DCN architectures consisting of multiple network technologies are also illustrated.

The reference DCN architectures herein are meant to illustrate the topologies in general, as well as the protocols required to support such architectures. They are broken up into Standard, Interim, and Non-Standard architectures, based on the following definitions:

Standard DCN: A standard DCN is one which does not violate the ITU-TM.3010 description of a DCN, i.e. “The DCN represents an implementation of the OSI layers 1 to 3, which include any relevant ITU-T or ISO standards for layers 1 to 3. A standard DCN provides no functionality at layers 4 to 7”.

Standard DCN Reference Architecture: A standard DCN reference architecture is one in which physical equipment is shown connected to a standard DCN (defined above) via the Q3 interface protocol suite as specified by ITU-T Q.811 and Q.812.

Interim DCN Reference Architecture: An interim DCN reference architecture is one which has been designated as interim by Bellcore GR-253-CORE (e.g. the TL1 over X.25 scenario).

Non-Standard DCN Reference Architecture: Any DCN reference architecture which does not fit into the definition of Standard or Interim as noted above shall be considered non-standard.

6.1 Standard Reference DCN Architectures

6.1.1 X.25 WAN Based

An X.25 WAN based DCN consists of an interconnected packet switch network between Packet Switches (PS) and Data Communication Equipment (DCE). It allows for communication between an OS and GNEs which provide connectivity to the SONET OSI based subnetworks. Figure 6 depicts such a DCN architecture. The architecture does not specify the protocol(s) which exist over the X.25 network other than what is required to transport data over the packet switched network. A standard option is called out in the protocol stacks shown in Figure 7. There is also an area which makes use of dual GNEs for OS-NE access.

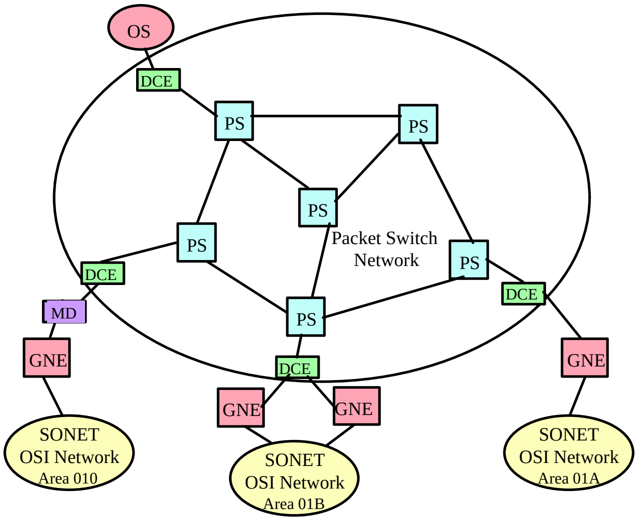

OS
DCE
PS
PS
PS
PS
Packet Switch Network
PS
DCE
MD
GNE
PS
DCE
PS
DCE
GNE
GNE
GNE
SONET OSI Network Area 010
SONET OSI Network Area 01B
SONET OSI Network Area 01A

Figure 6 X.25 WAN Based DCN

Figure 7 shows an OSI Application (i.e. CMISE) running over a full 7-layer OSI stack using an X.25 interface to the X.25 network to the GNE. The GNE, in this case, provides the convergence function from the X.25 network to the SONET subnetwork and makes use of the network layer relay function. Here, the GNEs operate as L2 IS nodes since they are connecting multiple areas over an OSI/X.25 network.

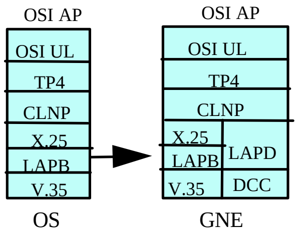


Figure 7 Standard X.25 WAN Based Protocol Stacks

6.1.2 OSI WAN Based

An OSI WAN based DCN consists of a collection of interconnected OSI routers which provide connectivity to SONET OSI subnetworks via routers in various central offices. The routers within the OSI network perform L2 IS functionality. The architecture shown in Figure 8 for the OSI WAN case depict the use of an Element Management System (EMS) for managing portions of the SONET network. Figure 8 also depicts the use of dual GNEs within a central office as well as GNEs performing L2 IS functionality. The GNEs labeled as L2 IS GNEs, provide L2 routing functionality for the OSI subnets they attach to (i.e. Area 01C and Area 01F) back to the DCN via the L2 IS OSI router. In this architecture, the end-to-end network is OSI with router and DCC subnets providing data connectivity.

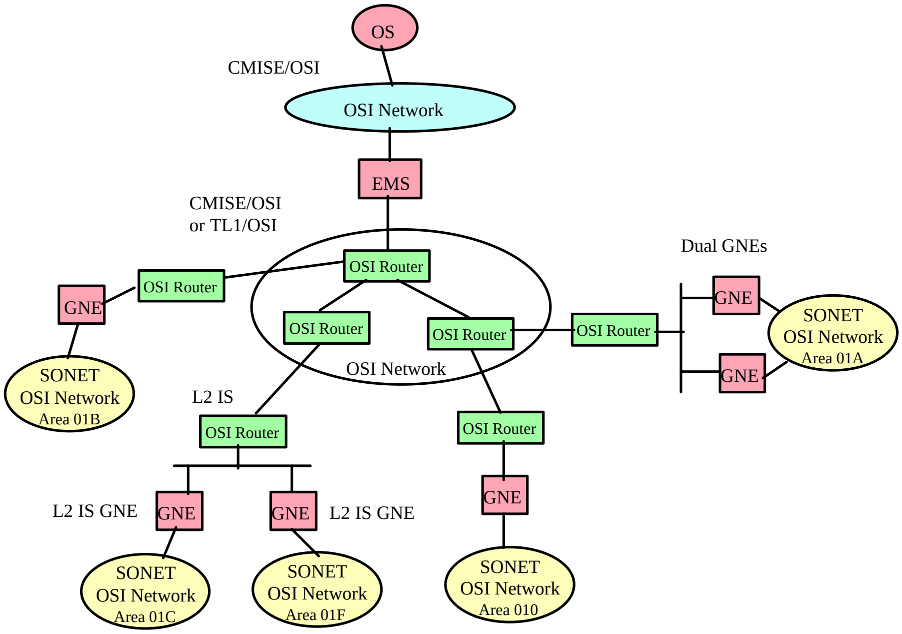


Figure 8 OSI WAN Based DCN

The OSI WAN based DCN architecture shown in Figure 8 depicts the EMS managing the subtending network, using an OSI application (i.e. CMISE) from the OS as well as from the EMS to the NEs.   The EMS may be providing subsequent messages to manage the NEs, or may be providing a direct “cut-through” for the OS.

The protocol stacks shown in Figure 9 depict the protocols active on the OS, EMS and GNEs. Note that no protocol or message translation is performed at the EMS in this case. However, the EMS may in fact generate subsequent OSI application (i.e. CMISE) messages to the NEs on behalf of the OS.   Again, the GNE in this case only provides the network layer relay between a LAN interface and the SONET DCC interface. This illustrates the target architecture for the DCN.

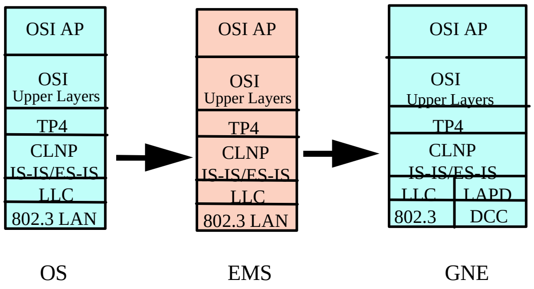


Figure 9 Standard OSI WAN Based Protocol Stacks

6.2 Interim Reference DCN Architectures

6.2.1 X.25 WAN Based

Referring to Figure 6, an interim option is called out in the protocol stacks shown in Figure 10. This option shows TL1 running directly over an X.25 interface from the OS through the X.25 network to the GNE. The GNE would terminate the X.25 interface and provide a full 7-layer OSI stack to carry the TL1 messages to the SONET subnetwork. In this case, the GNEs depicted only need to operate as L1 IS nodes, since there is no OSI routing between GNEs in the different areas.

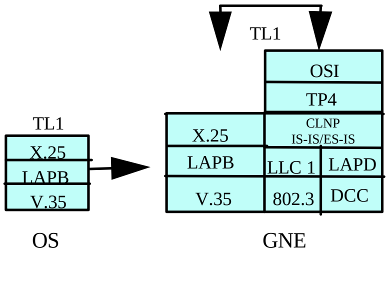


Figure 10 Interim X.25 WAN Based Protocol Stacks

6.3 Non-Standard Reference DCN Architectures

6.3.1 X.25 WAN Based

Referring to Figure 6, a non-standard option is called out in the protocol stacks shown in Figure 11. This option shows TL-1 (non-OSI application) running over a full 7-layer OSI stack using an X.25 interface to the X.25 network to the GNE. The GNE, in this case, provides the convergence function from the X.25 network to the SONET subnetwork and makes use of the network layer relay function. Here, the GNEs operate as L2 IS nodes since they are connecting multiple areas over an OSI/X.25 network.

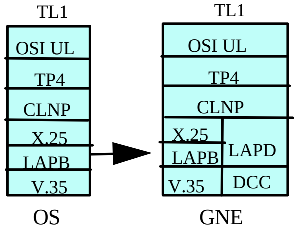


Figure 11 Non-Standard X.25 WAN Based Protocol Stacks

6.3.2 OSI WAN Based

Referring to Figure 8, another non-standard option has the EMS not only managing the subtending network, but also performing message conversion from CMISE to TL1 southbound from the OS; and TL1 to CMISE northbound from the NEs to the OS.

The protocol stacks shown in Figure 12 depict the protocol translation functions necessary at the EMS to provide the CMISE to TL1 translation from the OS to the subtending NEs. The GNE in this case only provides the network layer relay between a LAN interface and the SONET DCC interface.

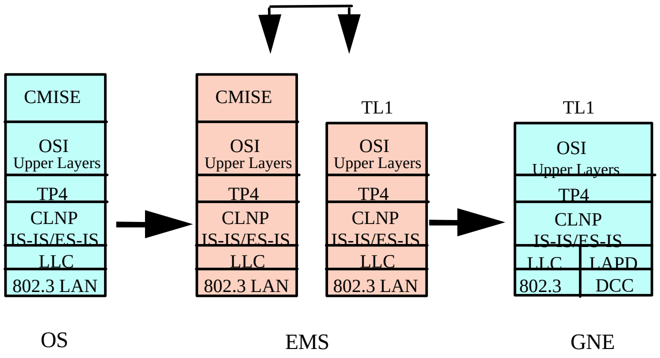


Figure 12 Non-Standard OSI WAN Based Protocol Stacks

6.3.3 IP WAN Based

An IP WAN based DCN consists of a collection of interconnected IP routers providing connectivity to multiple physical subnets. Some of these subnets would include SONET OSI subnetworks. The DCN architecture depicted in Figure 13 shows an internet consisting of a number of IP routers connecting multiple subnets, as well as OSs and GNEs which connect to SONET OSI networks in central offices. The one CO depicts dual GNEs within the office, while the other depicts GNEs connected to different routing areas. In the latter case, the GNEs function as L2 IS. The architecture configuration does not specify which protocols are active on the GNEs or OS. A number of choices are possible including the two shown in the protocol stacks indicated in Figure 14.

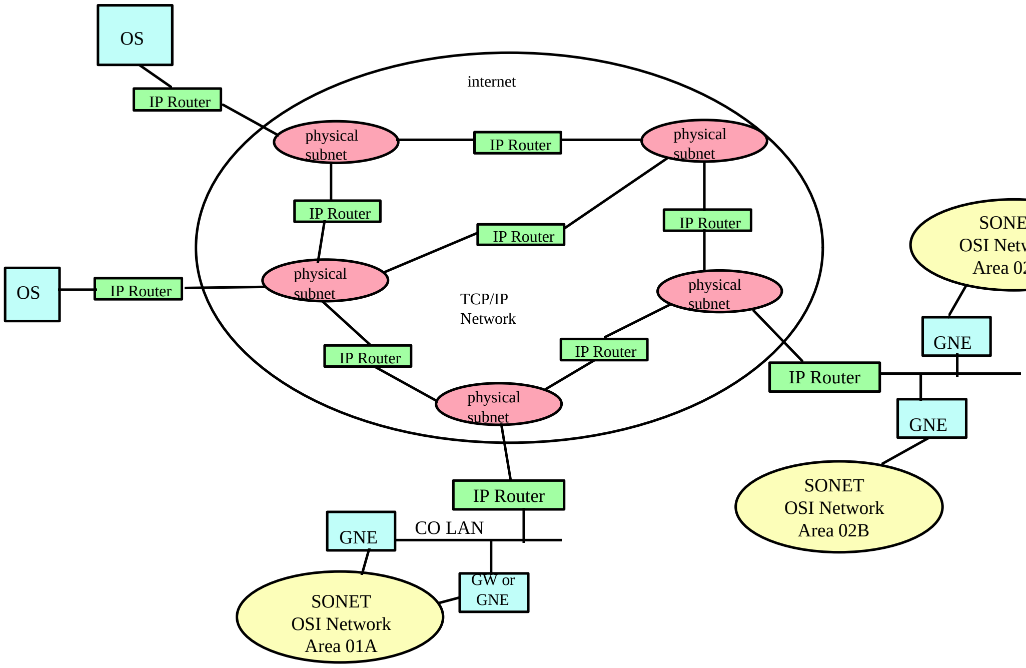


Figure 13 IP WAN Based DCN

The protocol stacks in Figure 14 show two of the possibilities for protocol support given the above architecture. The first interface assumes that TL1 is being run directly from the OS over a TCP/IP interface, over the IP network to the central office, where the GNE(s) perform a gateway function by terminating the TCP/IP stack and encapsulating the TL1 messages over a full 7-layer stack to the SONET OSI subnetwork. The second interface assumes that tunneling is used between the OS and the GNEs. The OS provides a full 7-layer OSI interface supporting TL1 to the TCP/IP network. Routers within the TCP/IP network would provide a tunnel entrance and exit point, encapsulating CLNP over TCP/IP. The GNEs, like the OS would provide a full 7-layer OSI interface and perform a network layer relay function between the CO LAN and the SONET DCC.

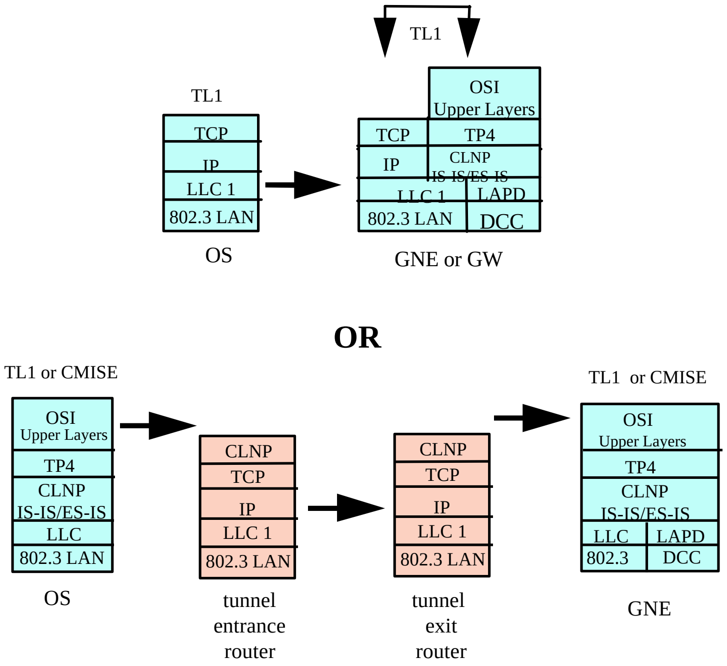


Figure 14 IP WAN Based Protocol Stacks

6.3.4 X.25/TCP/IP Hybrid

The hybrid DCN architecture depicted in Figure 15 illustrates an OS connected to an X.25 packet switched network, sending TL1 commands to an intervening EMS. The EMS performs a subsequent gateway function by terminating the X.25 network on one end, and sending the TL1 messages, both pass-through from the OS as well as any self generated, onto an IP network. The messages are subsequently sent through the IP network to a GNE in the CO. The GNE within the CO performs another gateway function from the TCP/IP protocol stack to a full 7-layer OSI protocol stack for the TL1 messages being sent to the NEs within the SONET OSI subnetwork. The GNEs shown in the CO require only L1 IS functionality, since they do not connect to more than one routing area. Also depicted is a SONET subnetwork utilizing dual GNEs to the DCN.

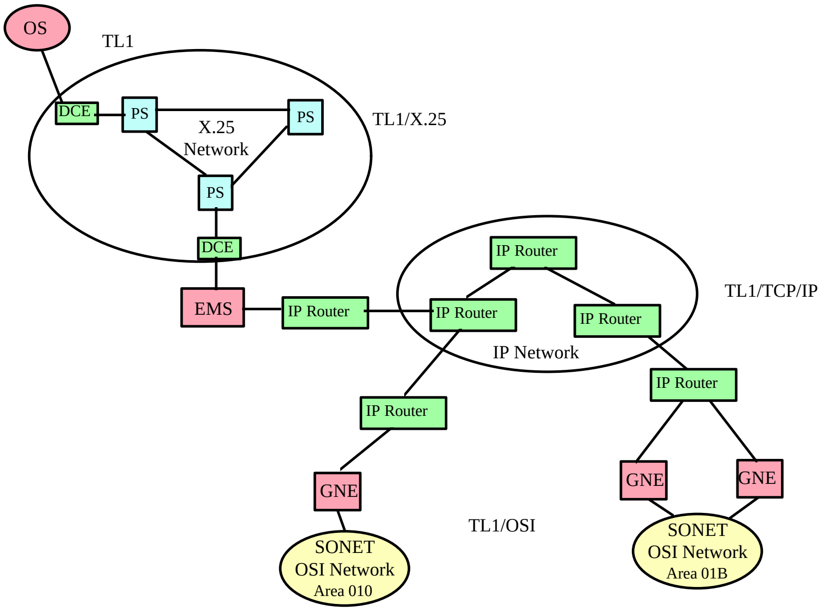


Figure 15 X.25/TCP/IP Hybrid DCN

The protocol stacks Figure 16 indicate the network interfaces which exist at the entry and exit points of the various DCNs, as well as the gateway functionality provided at the EMS and GNE for protocol conversion for the different subnetworks.

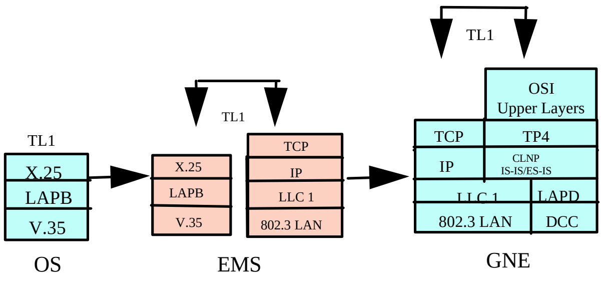


Figure 16 X.25/TCP/IP Hybrid Protocol Stacks

6.3.5 X.25/TCP/IP/X.25 Hybrid

The hybrid DCN architecture depicted in Figure 17 illustrates an OS connected to an X.25 packet switched network, sending TL1 commands to an intervening EMS. The EMS performs a subsequent gateway function by terminating the X.25 network on one end, and sending the TL1 messages, both pass-through from the OS as well as any self generated, onto an IP network. The messages are subsequently sent through the IP network to an IP-to-X.25 gateway. The second X.25 network terminates at a GNE in the CO. The GNE within the CO performs a gateway function from the X.25 protocol stack to a full 7-layer OSI protocol stack for the TL1 messages being sent to the NEs within the SONET OSI subnetwork. The GNEs shown in the CO require only L1 IS functionality, since they do not connect to more than one routing area. Also depicted is a SONET subnetwork utilizing dual GNEs to the DCN.

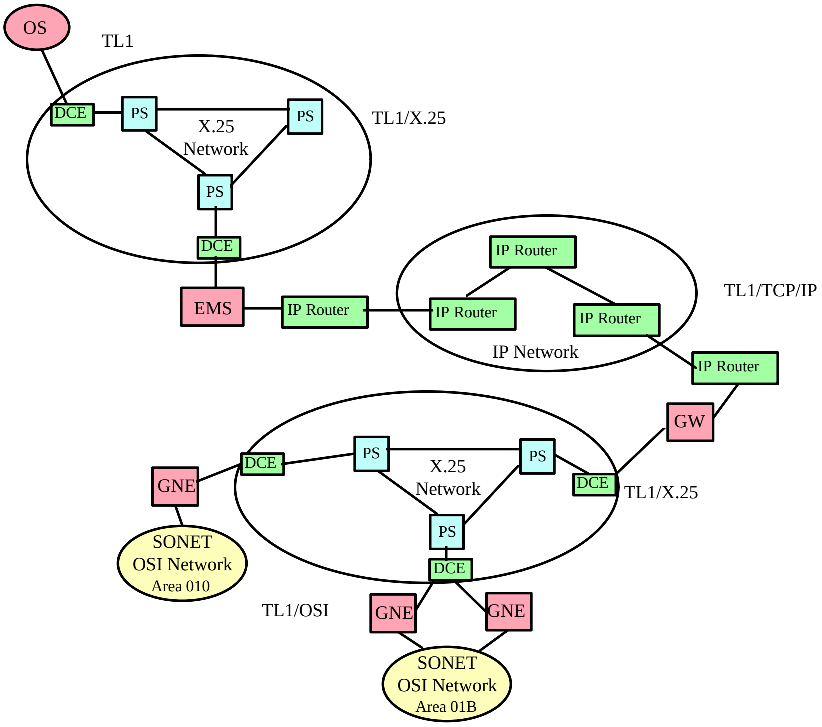


Figure 17 X.25/TCP/IP/X.25 Hybrid DCN

The protocol stacks in Figure 18 indicate the network interfaces which exist at the entry and exit points of the various DCNs, as well as the gateway functionality provided at the EMS, IP-to-X.25 Gateway and the GNE for protocol conversion at the different subnetwork interfaces.

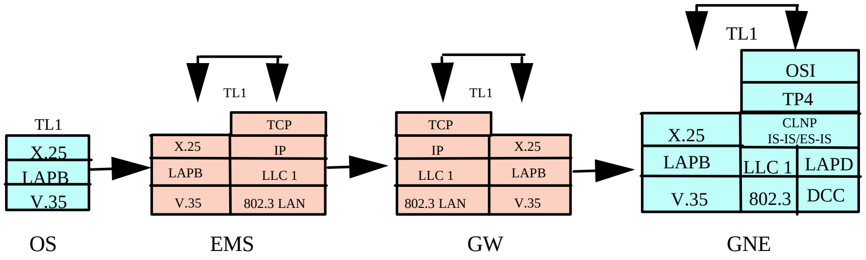


Figure 18 X.25/TCP/IP/X.25 Hybrid Protocol Stacks

6.3.6 RFC1006/TCP/IP/OSI Hybrid

The hybrid DCN architecture depicted in Figure 19 illustrates an OS connected to an IP network, sending TL1 or CMISE messages to an intervening EMS over a 7-layer protocol, using RFC1006. The EMS performs no protocol conversion in this case, but may perform message conversion or "pass-through". The messages from the EMS are sent through a second IP network to a transport service bridge at the edge of an OSI network. The messages pass over the bridge from the TCP/IP network to the OSI TP4/CLNP network, mapping RFC1006 PDUs to TP4 PDUs. The GNE within the CO performs a network layer relay function from the OSI LAN subnetwork to the SONET DCC subnetwork. The OSI routers within the OSI network perform L2 IS functionality, while the GNEs shown require only L1 IS functionality. Also depicted is a SONET subnetwork utilizing dual GNEs. (Note that the bridging function may not exist in a external box, but may in fact be a function of an NE (or GNE), itself.

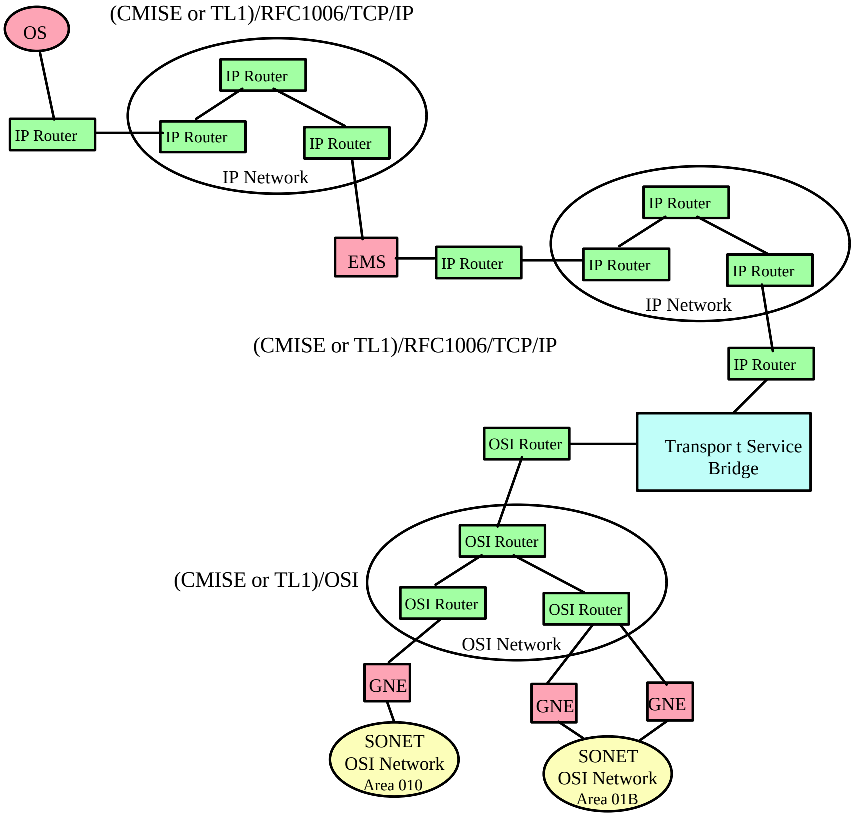


Figure 19 RFC1006/TCP/IP/OSI Hybrid DCN with TL1

The protocol stacks in Figure 20 indicate the network interfaces which exist at the entry and exit points of the various DCNs, as well as the functionality provided at the transport service bridge and the GNE for protocol conversion at the different subnetwork interfaces.

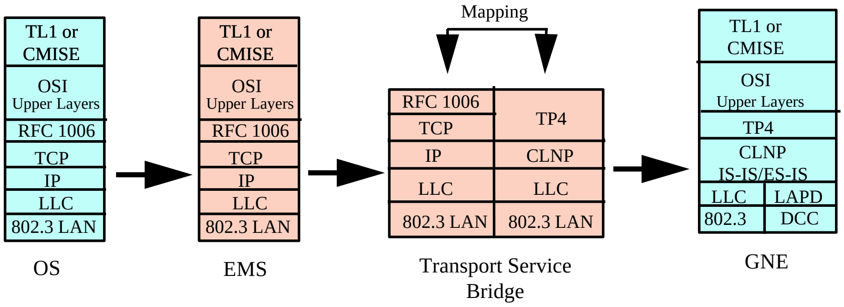
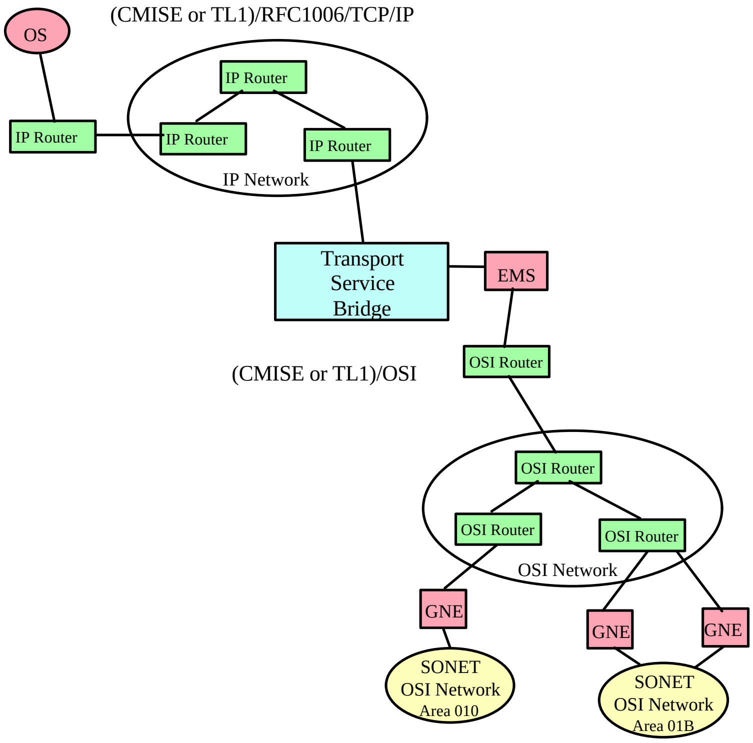


Figure 20 RFC1006/TCP/IP/OSI Hybrid with TL1 - Protocol Stacks

A second case of this hybrid DCN architecture depicted in Figure 21 illustrates an OS connected to an IP network, sending TL1 or CMISE messages to an intervening EMS over a 7-layer protocol, using RFC1006. In this case however, the messages pass through the transport service bridge prior to reaching the EMS. The EMS resides at the edge of the OSI network and is strictly an OSI node as shown. (Note that the transport bridge function could just as well reside in the EMS itself). The EMS performs no protocol conversion in this case, but may perform message conversion or “pass-through”. The GNE within the CO performs a network layer relay function from the OSI LAN subnetwork to the SONET DCC subnetwork. The OSI routers within the OSI network perform L2 IS functionality, while the GNEs shown require only L1 IS functionality.



This document has received the approval of the SONET Interoperability Forum (SIF).

August 13, 1998

Figure 21 RFC1006/TCP/IP/OSI Hybrid DCN with TL1 or CMISE

The protocol stacks in Figure 22 indicate the network interfaces which exist at the entry and exit points of the various DCNs, as well as the functionality provided at the transport service bridge and the GNE for protocol conversion at the different subnetwork interfaces. The difference here is that the EMS interfaces on the OSI side of the network.

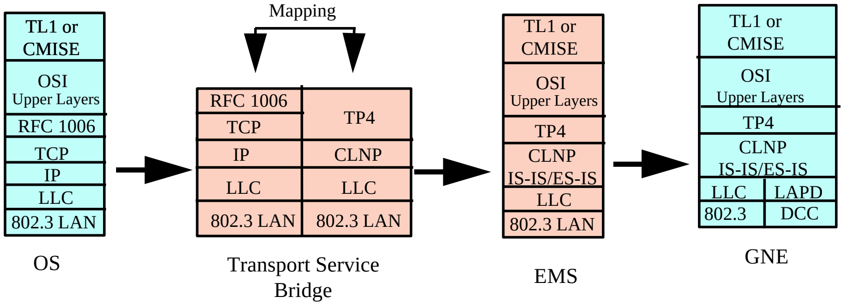


Figure 22 RFC1006/TCP/IP/OSI Hybrid with TL1 or CMISE - Protocol Stacks

7 DCN Design Guidelines

7.1 DCN Locations for Protocol Functions

OSI networks must have an architecture that enables an easily controlled and administered numbering plan that scales well as the number of NEs in an administrative domain grow. The architecture must provide adequate performance as well as assured availability, driving device and link redundancy.

Building on a solid architecture will allow for the network to be manageable. It will give the implementor the ability to add equipment, features, and services as necessary. And will generally allow for an equal distribution of access to the network, as well as, give a symmetric distribution of packet flow across the backbone transport.

When attempting to build a large network there are several guidelines which the designer must take into consideration. Large inter-domains are best designed in a modular or tiered structure. The overall architecture must embrace the concept of a “backbone” in which the routers form a core or transport utility. Secondly, when there are many remote locations dispersed geographically the designer should allow for router “switching centers” or “distribution” routers to be located around the backbone so as to have a symmetric connectivity to the central offices. Each central office will need routers called “access” routers for connectivity into their respective switching/distribution center (See Figure 23).

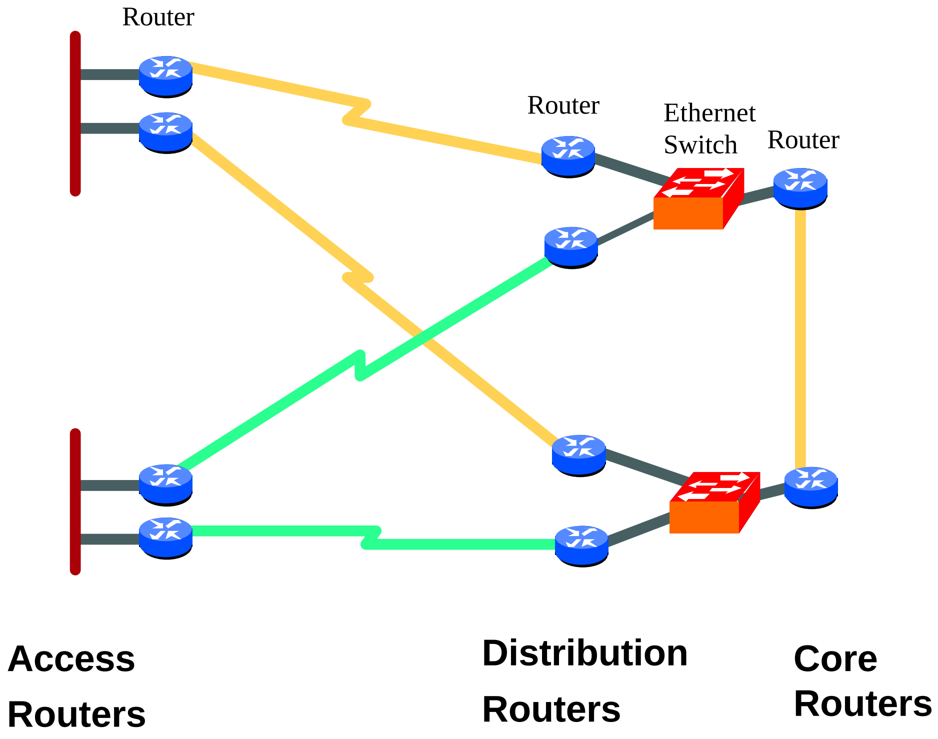


Figure 23 DCN Tiered Architecture

For a single routing domain a hierarchical system composed of SONET GNEs running level 1 routing, with local office (CO) access routers running level 1 and level 2 routing, and distribution routers at a gateway center running level 2 routing only, may be used to meet the stated requirements (See Figure 24).

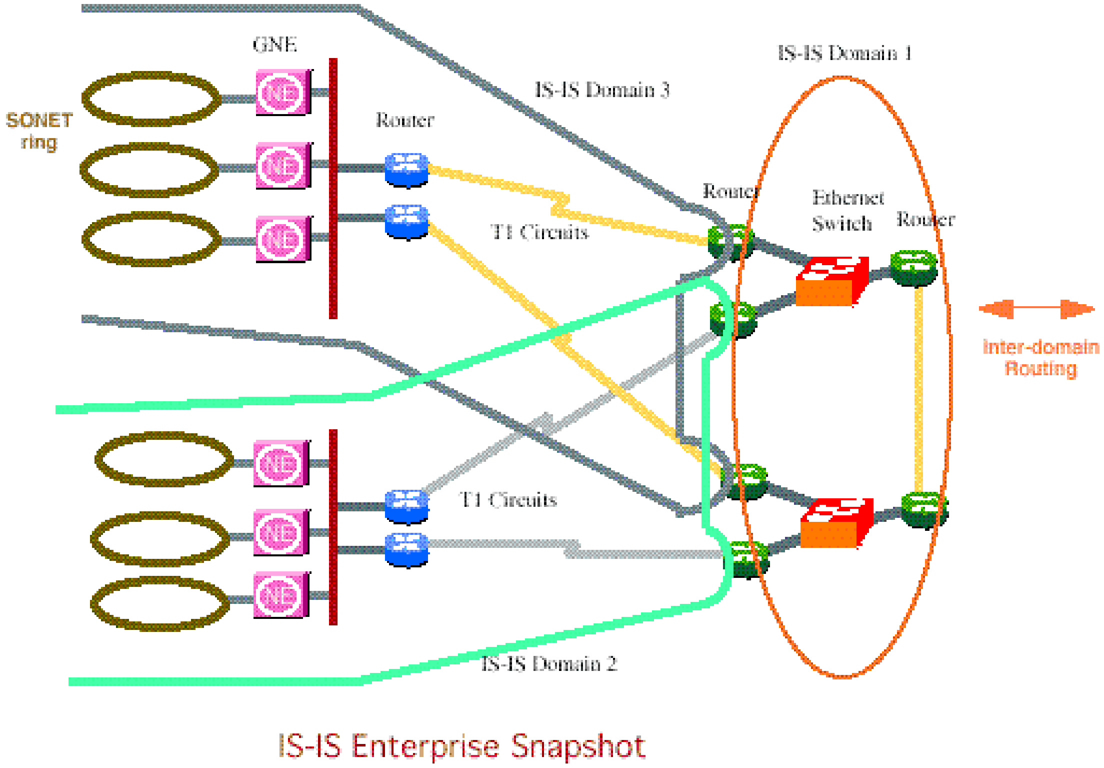


Figure 24 OSI DCN Architecture

Since, ISO 10589 (IS-IS) is a hierarchical layer three routing protocol, networks designed to carry CLNS traffic will work best when architected in a tiered design. With the inclusion of OSI routing at the SONET Section and Line DCCs it is important to use the DCC bandwidth efficiently. Keeping this in mind it is the SIF's desire to allow the Level 1 areas to grow in size as needed and also to not impact the size of the Level 2 routing tables. Where the CO’s consist of several SONET rings this design allows for the NE’s in the respective rings to be aggregated under the same OSI area. The routers function not only as a Level 1 area router but also as the necessary Level 2 router for communication outside the CO or between areas within the CO.

No two networks are alike. They may have similar topologies, but more likely than not will have protocol or bandwidth and media differences. So in trying to implement a network it is important to first decide on an architecture which scales and is manageable. During the design and implementation phases do not waver from the architecture.

Figure 25 shows an example where the GNEs are performing only level 1 routing.

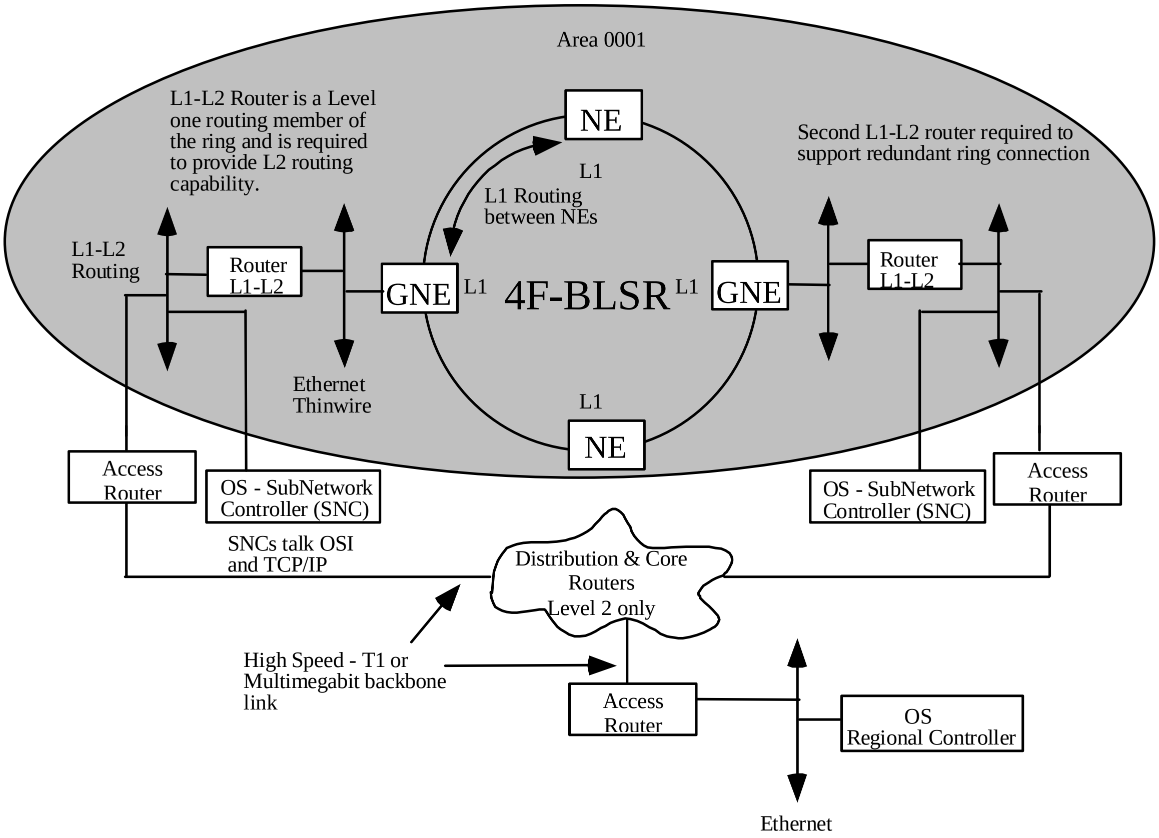

Area 0001
L1-L2 Router is a Level one routing member of the ring and is required to provide L2 routing capability.
NE
L1
L1 Routing between NEs
Second L1-L2 router required to support redundant ring connection
L1-L2 Routing
Router L1-L2
GNE
L1
4F-BLSR
L1
GNE
Router L1-L2
Ethernet Thinwire
L1
NE
Access Router
OS - SubNetwork Controller (SNC)
OS - SubNetwork Controller (SNC)
Access Router
SNCs talk OSI and TCP/IP
Distribution & Core Routers Level 2 only
High Speed - T1 or Multimegabit backbone link
Access Router
OS Regional Controller
Ethernet

Figure 25 GNEs Perform L1 Routing Only

Figure 26 and Figure 27 show some examples where the GNEs are performing level 1 and level 2 routing.

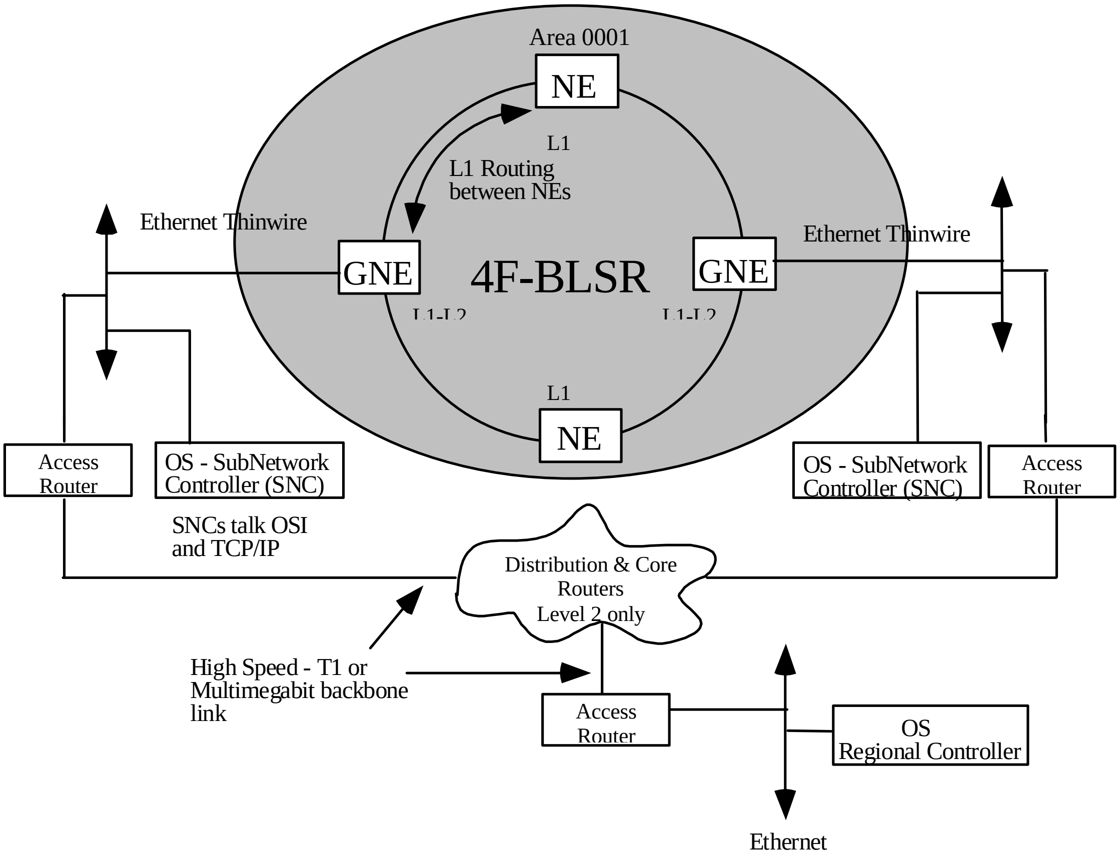


Figure 26 GNEs Perform L1 And L2 Routing

Notice that when the GNEs perform L1/L2 routing as shown in Figure 26, external routers are no longer required and the GNE can connect directly to the access router. Not having to have an external router is a distinct advantage when the GNEs are performing L1/L2 routing.

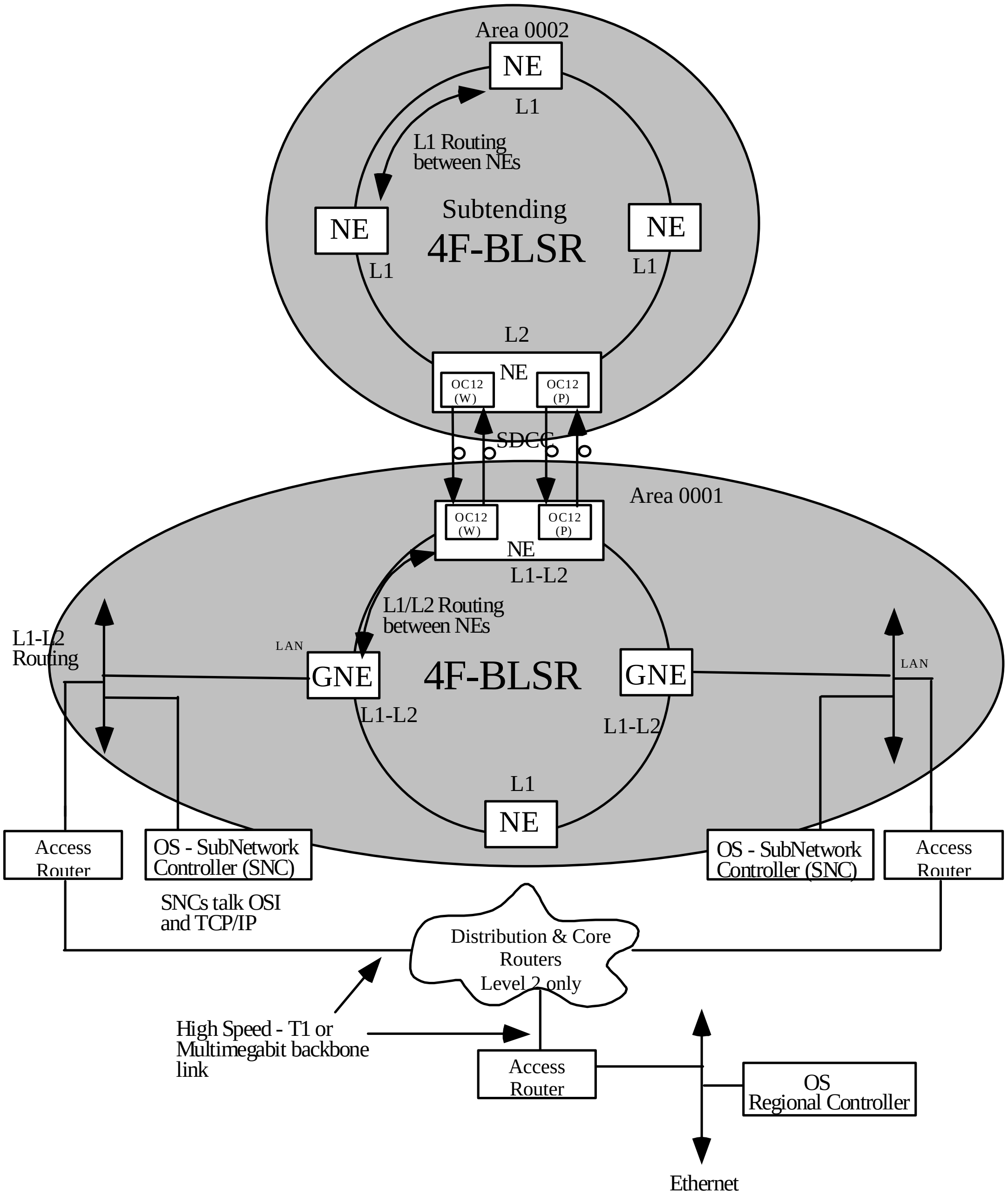


Figure 27 GNEs Perform L1 And L2 Routing To Manage Subtending Rings

Figure 27 shows how the GNE’s capability of L1 and L2 routing is used for the management of subtending ring/s with different routing areas.

7.1.1 Management System Considerations

Management systems should support the ES role of the ES-IS protocol (ISO 9542), including the ability to process Redirect PDUs for performance and cost matters.

In addition, a management system may act as an IS. For example, the reason why a management system may act as an IS could be dependent on the implementation architecture (e.g., a distributed management system architecture). Another reason why a management system may act as an IS would be to allow the management system to have access to information stored in the LSDB (see Section 0 for information on the LSDB).

Note that if the management system acts as an IS, this will result in additional routing related traffic on the LAN interface connected to the management system.

7.2 DCN Architecture Guidelines

Guideline 1: All SONET NEs in the same configuration with connected DCCs (e.g., a ring network) should reside in the same OSI routing area.

Guideline 2: A GNE and at least one CO router should reside in the same routing area.

Guideline 3: Data communications switching routers should not share a routing area with a CO router.

Guideline 4: Redundancy should be provided through multiple switching routers in separate physical locations with different routing areas, and multiple CO routers and WAN links to every GNE, so no single point of failure exists in the data communications network.

Guideline 5: SONET NEs deployed without a direct WAN interface should be accessible from a management system via more than one GNE (i.e., primary and secondary GNEs should be provided) to protect against single GNE point of failure.

Guideline 6: As a starting point, OSI routing areas should be limited to 50 routing devices (all SONET NEs and all data communications routers), until experience validates the actual numbers (Note: some vendors equipment may support a greater number of devices within an OSI routing area).

Guideline 7: Per GR-253-CORE, a GNE shall support L2 routing. However, the use of this L2 routing capability is optional.

7.3 IS-IS Deployment Requirements

This section identifies requirements imposed by ISO 10589 that must be understood by network planners or administrators.

7.3.1 System ID must be unique in OSI Area

No two systems in an OSI Area are permitted to have the same System ID. Since NEs are factory provisioned with globally unique System IDs, this criteria is easily met.

7.3.2 System ID for Level 2 IS must be unique in OSI Domain

No two Level 2 ISs in an OSI Domain are permitted to have the same System ID. Again, since NEs are factory provisioned with globally unique System IDs, this criteria is easily met.

7.3.3 Area Address must be unique in Global OSIE

No two OSI areas are permitted to have the same area address. There are exceptions for this for areas that never communicate outside the domain, but the SONET address format makes it easy for area addresses to be unique. Each administration obtains an Organization ID from ANSI, assigns arbitrary routing domain IDs and area IDs, and builds it's area address as:

39 840F 80 [org ID] 00 00 [domain ID] [area ID] [system ID] [N-Sel]

Refer to references ANSI T1.204 and ANSI X.3.216 for NSAP details.

7.3.4 Level 2 ISs must be interconnected

All Level 2 ISs in a domain must be directly interconnected to form a "Level 2 backbone". This does not mean that each Level 2 IS must have a link to all of the others, but rather, each must be able to reach all of the others without relaying through any Level 1 router or any router in another OSI Domain.

This is because Level 1 ISs always forward traffic that is destined for another area to the nearest Level 2 IS. As a result, Level 2 ISs cannot forward such traffic through a Level 1 IS, because the Level 1 IS will simply send the PDU back to the Level 2 IS (as it is likely to be the nearest). Also, a Level 2 IS will never send data destined for its own domain via another domain.

7.3.5 There must be a different Level 2 IS on each side of an Area boundary

There must be two Level 2 ISs at each area boundary, one on each side of the boundary. This is a requirement of the ISIS protocol.

7.3.6 LSPs must be small enough to reach every system

The originatingL1LSPBufferSize must be no larger than the smallest value for the Maximum NPDU Size for all Level 1 links in the OSI Area. Likewise, the originatingL2LSPBufferSize must be no larger than the smallest value for Maximum NPDU size for all Level 2 links in the OSI area.

When an IS needs to send multiple LSPs in order to advertise its links, it must divide them up itself; no other IS is permitted to segment the LSP when forwarding it on a link with a smaller maximum frame size. The IS-IS protocol does not specify what to do when an IS receives an LSP and cannot forward it on a link because it exceeds the maximum NPDU size (SN-SDU size) on the link. Some systems will attempt to forward the LSP and fail, others will drop the PDU, and others may treat the LSP is if it had a corrupt checksum and purge the LSP from the network. Inability to route deliverable PDUs, routing loops, or both may occur.

If an IS receives an LSP which is too large to propagate on a subnetwork, the correct behavior is not specified by ISO 10589. The problem can be avoided by consistent provisioning of the maximum generated LSP size on all systems.

7.3.7 IS-IS Deployment Example

Figure 28 shows an example where rings of ADMs (with subtending subnetworks) are interconnected using redundant DCSs.

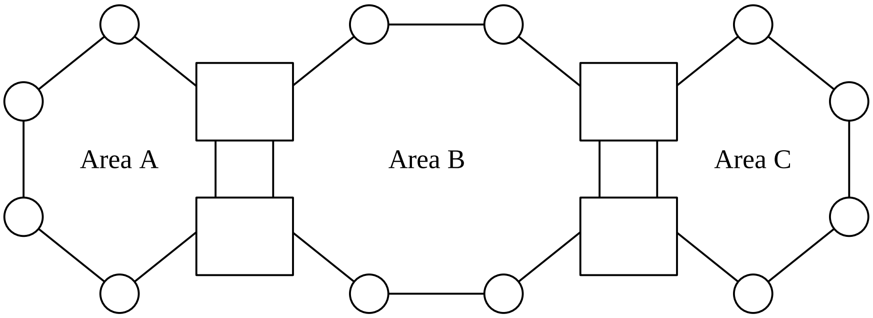


Figure 28 Dual Interconnected Ring Example

A) Unless a direct path is created between the DCSs, all NEs in Area B must be Level 2 ISs in order to survive a cable cut; otherwise, only half of the ring would need to be Level 2 ISs.

B) Each DCS must appear to be two ISs, one in each area. It must have two System IDs, and must operate the IS-IS protocol independently for each System ID. The most straightforward way to do this is for there to actually be two OSI stacks, one on each ring.

The easiest way to avoid these problems is to collapse all three areas into a single area, but there are disadvantages to very large areas.

C) If links in Area B are configured with a maximum NPDU size of 512, but systems in Area A or C are configured with a maximum NPDU size of 1497, then care must be taken to configure the originatingL2LSPBufferSize for routers in the latter areas to 512 rather than 1492, in order to permit propagation of the largest Level 2 LSP or SNP they might need to send (refer to Section 0).

7.4 Possible Need for Inter-Domain Routing

There are practical limits to the size of a routing domain. Since every Level 2 IS in the domain must keep information from every other Level 2 IS in the domain, the size of the domain is limited by the Level 2 IS with the least capability, either due to lack of memory or lack of processing power.

For this reason, it may be necessary to separate a network into multiple routing domains. If any communications is required between the domains, then inter-domain routing is required. Note that this does not necessarily mean that an inter-domain routing protocol is required, because ISIS supports a mechanism for provisioning inter-domain routes at the Level 2 ISs on the domain boundary and distributing these routes to the other Level 2 ISs in the domain. This is briefly described in Section 0, Inter-Domain Routing.

The only inter-domain routing protocol that is specified for use with CLNP is ISO 10747, “IDRP”. IDRP has not been widely implemented.

8 General Factors that Impact DCN Performance

8.1 Facility Transmission Rates

The major DCN functions and components are described in Section 0 of this document. Concerning facility transmission rates, the following functions and components are under consideration:

- Local Area Networks
  - Bridges
  - Repeaters
  - Hubs
- Mediation Devices
- Stand-Alone Routers
- Routing functions in SONET NEs
- X.25 Packet Switches
- Gateway Network Elements (GNEs)

The table below gives interconnection facility rates for the DCN components identified above. Local Area Networks are not used in the matrix, since they are comprised of Bridges, Repeaters , Hubs and other components already identified.

Each cell in the table shows recommended facility types and rates which can be used to connect the 2 DCN components forming the intersection of the cells. The technologies presented in Section 0 of this document have provided a basis for the facility rates shown in Table 3.

Table 3 Facility Transmission Rates

	Bridge	Repeater	Hub	Medtn. Device	Stand Alone Router	NE Routing Funcs.	X.25 Packet Switch	GNE Routing Func.
Bridge	Enet	Enet	Enet	Enet, DS1	Enet	Enet	DS1, T1	Enet
Repeater	Enet	Enet	Enet		Enet	Enet		Enet
Hub	Enet	Enet	Enet	Enet	Enet	Enet	Enet	Enet
Medtn. Device	Enet	Enet	Enet	X.25, Enet	Enet, X.25, Serial	Enet, X.25	9.6… 56K	DS3, Enet, 9.6… 56K
Stand Alone Router	Enet, Serial	Enet	Enet	Enet, X.25	All	Enet, X.25	X.25	Enet, X.25
NE Routing Funcs.	Enet	Enet	Enet	DS3, Enet, 9.6… 56K	Enet, X.25	SDCC, LDCC	9.6… 56K	SDCC, LDCC
X.25 Packet Switch			X.25	9.6… 56K	X.25	9.6… 56K	T1, FracT1	9.6… 56K
GNE Routing Funcs.	Enet	Enet	Enet	DS3, Enet, 9.6… 56K	Enet, X.25	SDCC, LDCC	9.6… 56K	9.6… 56K

Legend for above table:

Enet = 10Mbits/sec to 100Mbits/sec LAN interface
9.6…56K = 9.6Kbits/sec to 56Kbits/sec RS232 or V.35 or V.10 or V.11 interface
SDCC = Section Data Communication Channel in SONET overhead
LDCC = Line Data Communication Channel in SONET overheadSerial = Serial
FracT1 = Fractional T1

8.2 Number of Message Hops

The nodes making up a network are interconnected by data links and local area networks. Nodes and data links together form subnetworks where a group of nodes are attached so that each node is one hop from the other. A hop is defined as a path from one node to an adjacent node across a single data link. Subnetworks deployed on a broadcast media like CSMA/CD local area network or an ethernet can contain two or more adjacent nodes while nodes connected through physical links form a single point-to-point data link.

SONET NE’s form point-to-point data links using the Section DCC as the physical medium. Ethernet LANs and X.25 packet switched networks are examples of subnetworks which can contain multiple single hop nodes.

Packets are forwarded router by router. For example, a packet sent from a Network Management station destined for a remote NE is forwarded router by router, or hop by hop, until it reaches the destination NE.

A significant cost of traversing a network is associated with the number of hops a packet must follow. This is the cost of crossing a data link between two nodes or routers and is additive based on how many nodes/routers the packet must cross to reach it’s destination.

As a general rule when building a network, it is advisable to plan the topology of the management enterprise so that the diameter or longest paths across the network between any two far end nodes is constant. This balances out the response time across the network and guarantees fairness with respect to distance or number of hops a packet must cross to get from one side of the network to the other.

The farther the receiver is from the sender the more hops the packet must cross. In building routed networks, the best topologies turn out to be the ones having an architecture with the highest bandwidth at the core, a limited number of hops and a consistent diameter across the end-to-end topology.

When forwarding packets through a multi-hop network there are hidden costs associated with the end-to-end path. Relaying performed by stations, repeaters, bridges, routers, data links, and line propagation inject delays in the end-to-end response time. By limiting the number of hops, the network architect can minimize the impact of latency caused by these hidden costs.

8.3 Cross-node Delay

8.3.1 Node Delay Values

Cross-node delays are comprised of queuing delays, packet processing time, table lookup and a variety of other routing processes. Currently, values for these processing times are not generally available publically from equipment vendors. In the foreseeable future, vendors may come forth with performance values as more demands are placed on them by their customers.

8.3.2 Node Delay Optimization

The following is a brief list of components of DCN devices which affect end-to-end message delay. These items have been identified in Tanenbaum (pp. 561-565) (refer to Section 0) as good system design guidelines.

- CPU speed of routing elements – increase as much as possible

  The CPU speed of routing elements should be maximized to process packets.

- TPDU size – maximize

  If many small TPDUs are transferred, processing overhead is used for each packet. If larger TPDU sizes are used, the processing overhead is minimized for each packet.

- Context switches to process 1 packet – minimize

  Context switches between user code and operating system code further delays processing of packets. Packets should be allowed to arrive and be passed directly to a receiving process without intermediate processes such as buffering management, network layer handling, etc.

- Copying between operating system kernel, network layer processing, transport layer processing and application layer processing -- minimize

  Copying of buffers between various processes must be minimized to avoid delays in processing message streams at each device in the DCN. An example given by Tanenbaum (refer to Section 0 for the full reference) indicates even a 50 MIPS machine may only have the ability to handle a 16 Mbits/sec message stream because of excessive copying.

- Congestion avoidance algorithms -- use proven algorithms

  There are many algorithms which have been developed during the evolution of networks and can be found in society publications, journals, textbooks and other sources. An algorithm which have been utilized in OSPF networks is the NIST Sender Congestion Avoidance Algorithm.

- Timeout timers – use conservative setting (longer than could be expected)

Timers should be used sparingly. Every timeout causes an action to take place which does not advance packets through the network. Setting timers to higher values than would normally be used may reduce the overhead associated with timeout processing.

8.4 End-to-end Transport Delay

In order to maintain a typical SONET network it is assumed there will be a Network Management System. Figure 29 shows an NMS which communicates with NEs through a GNE. The topology shows a LAN/WAN hierarchical network which will allow packets to be forwarded to the proper Network Element.

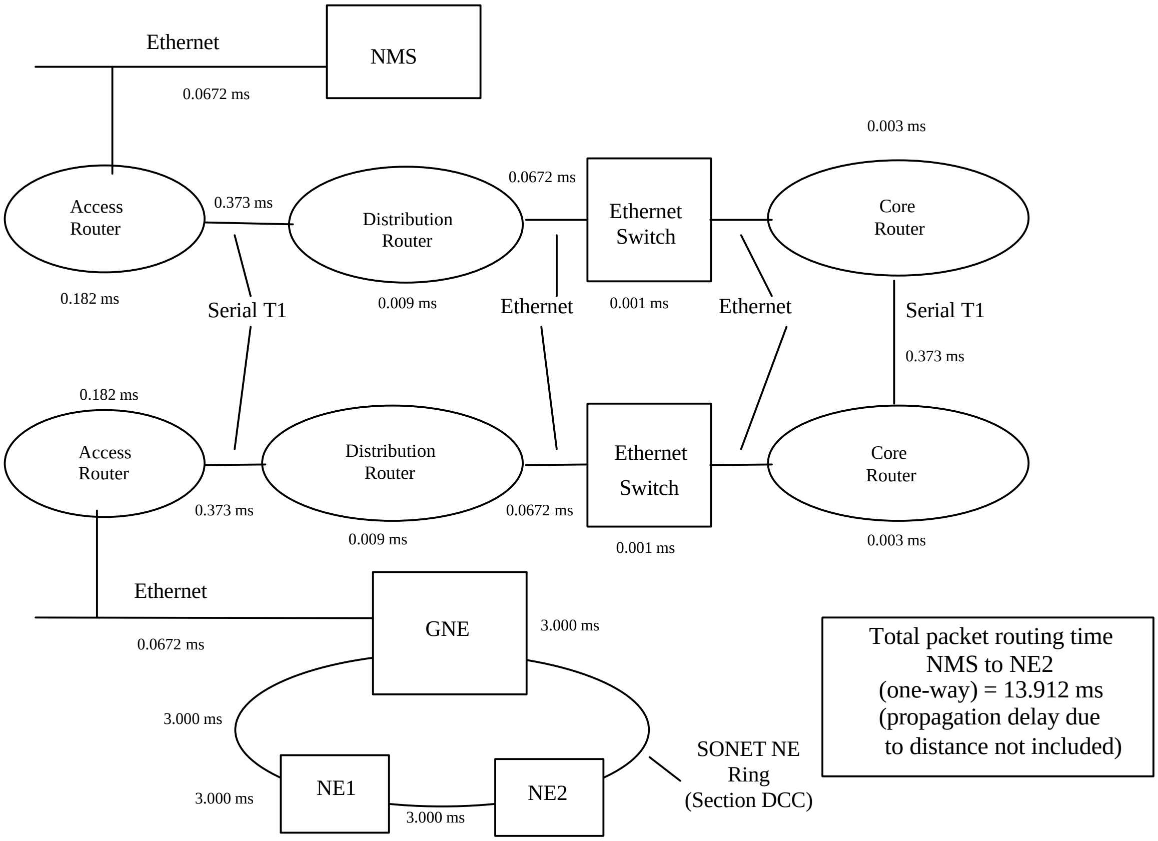


Figure 29 Typical OSI DCN End-to-End Delay

The time to transport a 64 byte packet one-way is shown next to each connection between elements. The time to process the packet is shown next to each element. The total time to transport and route the packet from the NMS to NE2 is the sum of all the transmission times and packet processing times (i.e. 13.912 ms).

Based on Figure 29 the typical node delays are shown in Table 4.

Table 4 Typical Node Delays for 64 byte Packets

Node or Element Type	***Delay for 64 Byte Packet (One Way) In Milliseconds***
10 Mbit/s Ethernet LAN	0.0672
Serial T1	0.373
Section DCC	3.000
Access Router	0.182
Distribution Router	0.009
Ethernet Switch	0.001
Core Router	0.003
Transport Network Element	3.000
Propagation Delay	7.000 us/mile (refer to ITU G.114 - Section 0)

8.4.1 Delay Analysis

The values cited in Figure 29 and Table 4 can be derived from the packet size, the transmission bandwidth and the protocol delay as follows:

media_delay = packet_size / bandwidth + protocol_delay

10 Mbit/s Ethernet delay: media_delay = ((pkt_size_in_bits) / media_BW) + (protocol_delay)

For an Ethernet connection, the protocol delay is 0.0096 ms of inter-packet gap required by the protocol.

64 byte packet = 576 bits ----> 576 / 10,000,000 + 0.0096 = 0.0672 ms delay

Serial T1 delay: media_delay = packet_size / bandwidth + protocol_delay

The protocol delay for a Serial T1 line is zero.

64 byte packet = 576 bits ----> 576/1,544,000 + 0 = 0.373 ms delay

Section DCC: media_delay = packet_size / bandwidth + protocol_delay

The protocol delay for Section DCC connections is zero.

64 byte packet = 576 bits ----> 576/192,000 + 0 = 3.0 ms delay

Access Router:

Router perfomance is measured in packets per second (PPS) for 64 byte packets. This number usually describes how fast a router can input a packet, do a routing table lookup, modify the headers for the new destination and send the packet out the interface. The size of a router is dependent on the latency of a packet to be forwarded. Three common router speeds are shown here for reference:

1. 5500 PPS router: latency delay = 1/5500 = 0.182 ms
2. 15,000 PPS router: latency delay = 1/15,000 = 0.067 ms
3. 85,000 PPS router: latency delay = 1/85,000 = 0.012 ms

Distribution Router:

115,000 PPS router: latency delay = 1/115,000 = 0.009 ms

Core Router: Make note here that today’s core routers are architecturally very complex. They not only have several hardware switching levels but will also have several software/microcode switching paths. It is important to understand the hardware makeup of the interfaces and the hardware path a packet will take

350,000 PPS router = .000003 seconds

Ethernet Switch: An Ethernet switch today is usually made up of highly complex ASICs. A packet arrives off a 10BaseT interface, the packet header is analyzed and forwarded out the appropriate interface. The forwarding engines, because they are finely tuned can achieve rates in excess of one million packets per second. Again testing is done with 64 byte packets. These switches perform line rate transitions between ingress and egress ports. Switches usually have been tested by an independent test lab.

Ethernet switch: delay = 1/1,000,000 = .000001 seconds

Transport network element: The performance numbers here are vendor dependent. Currently, values for NE routing function processing delay have not been made generally accessible.

Estimates from industry experts range from 0.5 ms to 20 ms. 3.0 ms (the Section DCC transport delay) will be used as a nominal value.

8.4.2 7.3.2 Delay as a Function of Packet Size

In the routing environment it is important to see how fast the router can forward packets of the minimum size. As it turns out 25-30% of packets seen in an average Internet are of minimum size. Performance of the forwarding process coupled with the hardware rates is important not only because of the number of packets seen at this size but also because at wire rates the minimum size packet consumes the least percentage of the overall media bandwidth. Hence, it is more important to understand how small packets impact the overall performance of the routing architecture then for maximum sized ones.

The CPUs in a router work harder forwarding smaller packets then they do when routing large packets. The CPU does the same amount of work to lookup a path and to manipulate the packet headers, but for small packets the CPU does it more often per second then for larger packets. Large packet forwarding allows other processes to gain control of the CPU because it takes longer to move the packet in/out of the router.

What follows is theoretical performance of 3 major transmission facilities for variable packet sizes:

Table 5 10 Mbit/s Ethernet Transfer Delays

packet size (bytes)	**64**	**128**	**512**	**1024**	**1500**	**1518**
# bits	576	1088	4160	8256	12064	12208
Line delay (ms.)	0.0576	0.1088	0.416	0.8256	1.2064	1.2208
Total transfer delay (ms.)	0.0672	0.118	0.426	0.835	1.217	1.232
One way rate (PPS)	14880	8445	2349	1197	822	812
bits / sec	7618560	8647680	9621504	9805824	9864000	9860928
% of bw used	76%	86%	96%	98%	99%	99%

Table 6 Serial T1 Transfer Delays

packet size (bytes)	64	128	512	1024	1500	1518
# bits	512	1024	4096	8192	12000	12144
Line delay (ms.)	0.331	0.663	2.653	5.306	7.772	7.865
Total transfer delay (ms.)	0.331	0.663	2.653	5.306	7.772	7.865
One way rate (PPS)	3015	1507	376	188	128	127
bits / sec	1543680	1543168	1540096	1540096	1536000	1542288
% of bw used	99.98%	99.95%	99.75%	99.75%	99.48%	99.89%

Table 7 Section DCC Transfer Delays

packet size (bytes)	64	128	512	1024	1500	1518
# bits	512	1024	4096	8192	12000	12144
Line delay (ms.)	2.666	5.333	21.333	42.666	62.500	63.250
Total transfer delay (ms.)	2.666	5.333	21.333	42.666	62.500	63.250
One way rate (PPS)	375	187	46	23	16	15
bits / sec	192000	191488	188416	188416	192000	182160
% of bw used	100.00%	99.73%	98.13%	98.13%	100.00%	94.88%

8.5 Protocol Parameter Settings

8.5.1 Default TPDU Size

The default TPDU size specified by GR-199-CORE is 128. This specification is based on a misunderstanding of other standards, and should be ignored. If equipment is deployed that has a default TPDU size of 128, the TPDU size should be provisioned to 1024 or 2048 to improve OSI Transport Protocol performance dramatically. The value does not need to be set consistently on different systems, because the value used in any connection is negotiated between communicating transport entities at connection initation time.

8.5.2 Default NPDU Size and Range

There are some deployment requirements concerning the maximum NPDU size. Sometimes, the maximum NPDU size is implicitly set by setting the maximum Link Service Data Unit size.

The maximum NPDU size must be the same at each end of a SONET link. Otherwise, ISIS systems will not interoperate over the link. While it is not necessary to use the same NPDU size on all links in a network, doing this may reduce segmentation and reassembly overhead. 512 bytes is a common default maximum NPDU size, because it is the default value for L2SIZE in GR-199-CORE and therefore is the default for most TL1 managed equipment. Use of larger sizes reduces OSI overhead for bulk data transfer applications like program download, at the cost of additional buffer space requirements on NEs. While GR-199-CORE only specifies sizes that are even powers of two, best results are usually obtained when the maximum NPDU size is the size of the typical CLNP header (57 or 60 bytes) plus the size of the maximum TPDU size used by the applications.

The maximum NPDU size must be the same for all Level 1 ISs on a LAN that are in the same OSI area. The maximum NPDU size must be the same for all Level 2 ISs on a LAN that are in the same routing domain. The standard size is 1497 bytes, which is derived from a maximum Ethernet user packet size of 1500 bytes, minus three bytes for the ISO 8802-2 (Logical Link Control) header.

The default NPDU size on Section DCC is 512, but should be set to a larger value such as 1497 or 2108 when possible. The value must be the same at each end of an SDCC link, or IS-IS will not operate over the link. Since SONET NEs are already deployed with a default (and possibly a maximum) NPDU size of 512, the default should not be changed, even though it leads to inefficient use of the DCC. The valid range for NPDU size should not be constrained to powers of two, as it is in GR-199-CORE. It should not be permitted to be lower than 512 (as it is in GR-199-CORE) because lower values are not permitted by the CLNP specification (ISO 8473).

8.5.3 LAPD Information Field Size

The value for LAPD N201 (Information field size) must be at least as large as the value for the maximum NPDU size. NE implementations should enforce this requirement.

8.5.4 IS-IS Maximum LSP Size

The originatingL1LSPBufferSize/ originatingL2LSPBufferSize provisioning parameters should default to 512, because this is the default maximum frame size on DCC subnetworks for SONET NEs that are already deployed. LSPs cannot be segmented by ISs which propagate them. This parameter must be consistent throughout the area for L1 and the domain for L2.

8.5.5 Transport Persistence Time

The Transport persistence time should be greater than the IS-IS LSP Regeneration Timer, which defaults to 30 seconds. This permits a connection to survive temporary IS-IS routing problems that can happen when multiple topology changes happen on the same router during a short interval.

A long persistence time is not important for easily retried, short-lived connections. It is most important for connections that are not easily retried or (most significantly) long-lived connections which must be completely redone on failure (i.e. FTAM downloads, for most FTAM applications).

The Transport persistence timer is often provisioned as a retry time and a retry count. If, however, exponential retry times are used or if the NIST Transport adaptive retransmission algorithm is used, the value may need to be specifically configured.

For SONET networks, it may be wise to either increase the recommended persistence time or reduce the minimum LSP regeneration interval.

8.5.6 IS-IS Metric Values

IS-IS metric values are important configuration parameters that affect routing decisions. On each IS, each link connected to that IS is provisioned with metric values describing the attributes of the link for routing purposes. Each link has from one to four metric values that are provisionable. The metric types are given the following names and intended uses:

default	indicates the capacity or maximum throughput of the link
delay	indicates the transit delay experienced on the link
expense	indicates the cost for transmitting data on the link
error	indicates the probability of error on the link

Support of the *default* metric is required on all IS-IS systems. Support of other metrics is optional, and actual use of other metrics increases memory requirements on the IS. Because all routers may not support or may not be provisioned to support the other metrics, their use should be avoided unless there is a strong reason to employ them and a clear understanding of how they work, which is outside the scope of this document. The text that follows assumes that only the *default* metric is used.

Smaller values for the metric indicate that the link is better with respect to the criteria; larger values indicate that the link is worse. When choosing a route, an IS essentially adds up the metric values for all links on all potential paths and chooses the path with the smallest total. For this reason, all Level 1 ISs in an area should be provisioned using the same conventions for choosing metric values. Likewise, all Level 2 ISs in a routing domain should be provisioned using the same conventions (although not necessarily the same as are used for Level 1).

Unfortunately, there are no standard conventions for metric values, and as a result, different vendors' equipment will choose different metric values for the same kind of link.

The following values are suggested for Level IS-IS *default* routing metrics. Other metrics, if supported, should be disabled. This should provide reasonable IS-IS operation in a multi-vendor network.

Link Speed	**value for *default* metric**
100 Mb/s	2
10 Mb/s	6
1.5 Mb/sec T1	10
576 Kb/s (Line DCC)	18
192 Kb/s (Section DCC)	34
56 Kb/s	63

NEs that route at speeds substantially lower than the link speeds should use metric values greater than those listed above, interpolating between the value for the link and the next higher value, based on the proportion of the link bandwidth that the NE can fill.

Service Providers are to be free to provision the values differently when they have consistent policies for metrics that differ from these suggestions.

8.5.7 NIST Adaptive Retransmission Policy

NEs should support and enable the NIST adaptive retransmission policy for default. This avoids the necessity for topology-dependent configuration and improves performance during network congestion.

9 OSI Network Administration

9.1 Splitting Existing Routing Areas

The objective of splitting a level 1 routing area is to take a large routing area and partition it into smaller, disjoint routing areas. There may be a number of reasons for wanting to do this. Changes in the administrative responsibility of a set of NEs may require the existing routing area to be re-partitioned according to these new administrative changes. A second reason may be that the routing area has grown too large. NEs within the routing area may not have the memory required to store a large routing table. Additionally, the NEs may not have the CPU cycles required to process the routing traffic associated with a large routing area (exchanging LSPs, Hellos etc.). Splitting level 1 routing areas will involve making changes to an NE’s manual area address set, as well as, configuring level 2 routers within a network.

To illustrate how a level 1 routing area would be split, consider an example where an existing level 1 routing area consisting of 150 nodes is to be split into three areas, A, B and C, each of 50 nodes. The routing area currently supports only nodes with area address A. Routing area A before the split is shown in Figure 30 and after the split in Figure 31.

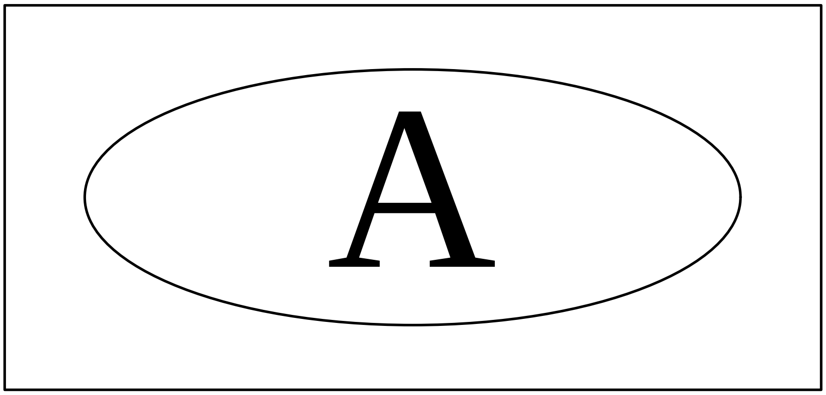


Figure 30 Routing Area A to be split into three routing areas

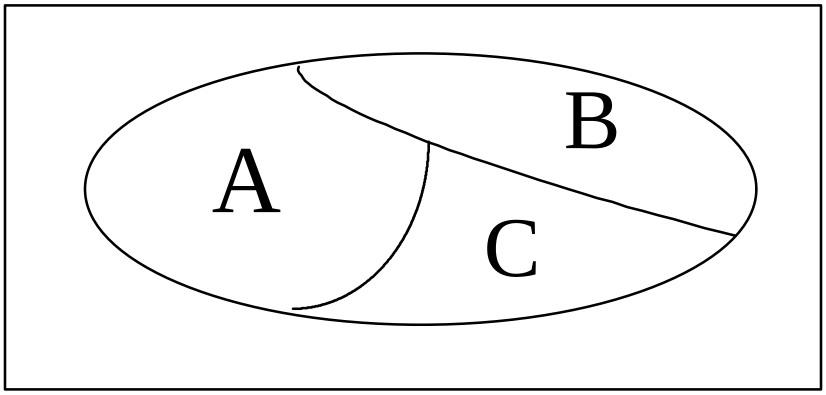


Figure 31 Area A split into 3 smaller areas, A, B and C

In order to consider splitting an area, all the NEs to be split into new areas must support provisioning new area addresses into their manual area address set. How many manual area addresses will an NE support? The number of manual area addresses supported by an NE must be the same as for it's neighbors or else an adjacency with those neighbors cannot be established. Each NE advertises the number of manual addresses it supports in its Link State PDUs (LSPs) and Intermediate system to Intermediate system Hellos (IIHs), if they don't match the adjacency cannot be brought up. NEs which support older implementations of ISO/IEC 10589 will support, by default, 3 entries in their manual area address set. NEs which support ISO/IEC 10589:1992 will have the number of entries supported in their manual area address set governed by the system parameter 'maximumAreaAddresses'. It is safest to engineer guidelines based on the lowest common denominator so for this paper it will be assumed that all nodes within the network support provisioning at most 3 manual area addresses. Splitting this routing area into three new areas, A, B and C could be accomplished as follows. First, visit each NE in new area B and provision the area address for B in their manual area address set, in addition to that of A. Next, visit each of the NEs in new area C and provision these NEs manual area address set with the area address C in addition to that of A.

If we were to proceed as above there is the potential for problems to arise. The problem may occur when each NE generates its set of computed area addresses. Each of the NEs in area B will generate LSPs and IIHs stating that they support the set of manual area addresses {A, B}. Each of the NEs in area C will generate LSPs and IIHs stating that they support the set of manual area addresses {A, C}. Each of the NEs in area A will send LSPs and IIHs stating that they support the set of manual area addresses given by {A}. When each NE puts these LSPs together they will generate a computed set of area addresses which will be given by the set {A, B, C}. This computed set of area addresses now consists of the maximum number of area addresses supported on each node.

The problem is what if an error is made adding area address B or C to any node ? Assume a finger problem creeps in when adding area address B to some node so that the area address is added not as B but as some other area address, say D. Each node can support a maximum of 3 area addresses in its computed set of area addresses but now there are actually 4 area addresses being advertised in the sum of the LSPs. Each NE will generate a computed set of area address equal to {A, B, C, D}. According to ISO/IEC 10589:1992, the three lowest area addresses are to be retained and the highest area address is to be dropped causing a manual area address dropped event. The NEs whose area address was dropped will now lose its network layer connectivity causing any active transport connections to time out. If OAM information had been flowing over that transport connection then the NE would be isolated from an OAM perspective when that connection timed out. If synonymous addressing is not supported by the NEs in this network then the problems experienced due to the area address being dropped will be much worse.

From the preceding example it can be seen that it’s easiest to split an existing area into two areas and then further split those resulting areas in two again as required. Thus the way to split routing area A from the example is first to split it as Figure 32.

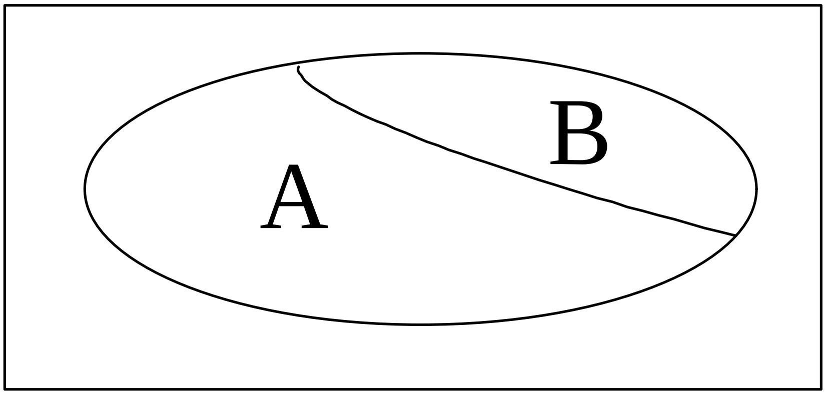


Figure 32 First split of Area A

Once area A is split as shown in Figure 32, the NEs in area B will support 2 manual area addresses, A and B. The NEs in area A will support area address A. The last step in separating the areas is to visit the NEs in area B and remove area address A from the list of supported manual area addresses. By doing this OAM connectivity will be lost between the two areas. If OAM connectivity to these two areas is to be maintained then an extra step must be followed before removing area address A from the list of manual area addresses supported by the NEs in area B. The extra step is to connect a pair of NEs, 1 in area A and 1 in area B by configuring them as level 2 routers. Once the level 2 routers are connected then the NEs in area B can have area address A removed from their list of supported area addresses. OAM connectivity will be maintained through the level 2 routing connection between the two areas. This is as shown in Figure 33.

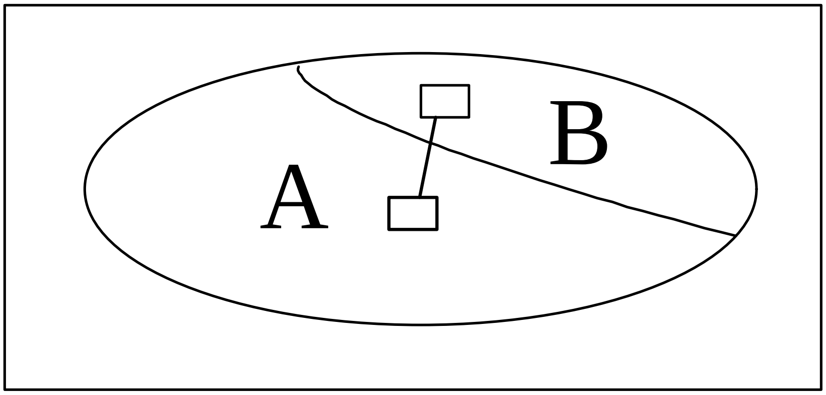


Figure 33 Areas A and B connected via level 2 routers

With the area now setup as shown in Figure 33, area A can now be further subdivided to create area C. This is the next iteration of applying the rules learned from earlier. The result is as shown in Figure 34.

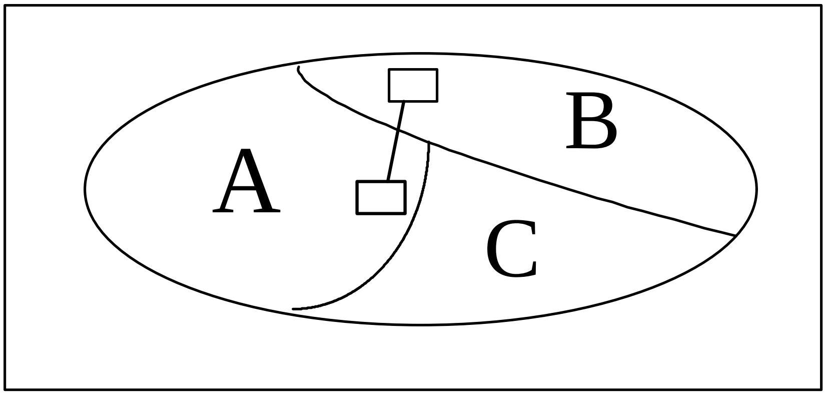


Figure 34 Area further split into Area C

It can be seen from the previous example that applying the same methodology can be used if any of areas A, B or C need to be further split. When partitioning the new areas it is advisable to partition (if possible) such that one NE per area can be used to interconnect the new areas via level 2 routing. The best case scenario in this example would be to have one level 2 router in each of domains A, B and C interconnected via one link between each pair as shown in Figure 35.

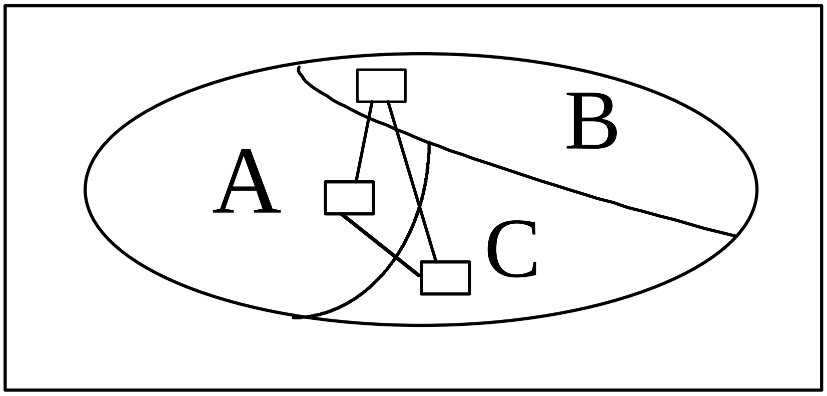


Figure 35 Areas A, B and C connected via level 2 routers

The worst case scenario would be to have one level 2 router in A connected to a level 2 router in B and another level 2 router in A connected to a level 2 router in C. Similarly B would have two level 2 routers in its area, one connected to A and one connected to C. Lastly C would also have two level 2 routers, one connected to the level 2 router in B and one connected to the level 2 router in A. Since ISO/IEC 10589:1992 states that each level 2 router must be connected to another level 2 router we now have to ensure that the pair of level 2 routers within each routing area must also be physically connected. If a physical connection did not exist before then one must be added. This is as shown in Figure 36.

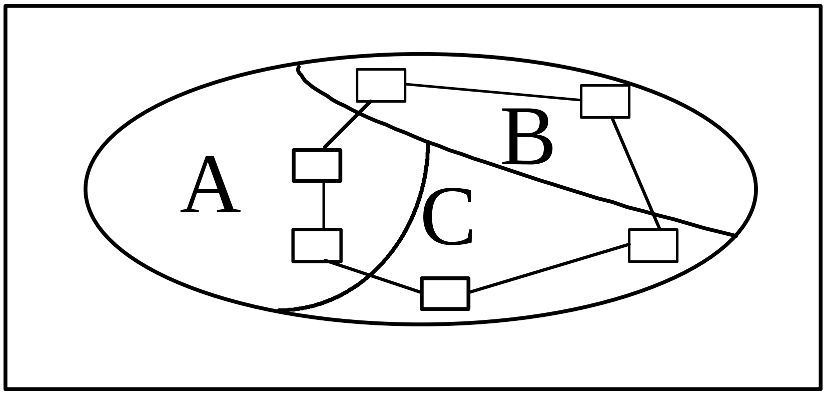


Figure 36 1 pair of level 2 routers per area pair

It becomes apparent from the preceding example that as the number of areas to be created increases so does the complexity of the interconnected level 2 network. As the interconnected level 2 network piggy backs on the underlying SONET network it may not always be possible to maintain the connectivity of the level 2 routers. Some other protocol would have to be found to get around this problem (perhaps IDRP ?).
If the areas to be created are to be kept isolated then the level 2 routers are not required and the complexity of connecting the level 2 routers is eliminated.

9.1.1 Guidelines for splitting areas

An area should not be split into more than 2 areas at any one point in time. If more areas are required then perform the splitting in an iterative fashion i.e. split the area first into 2 smaller areas, then split one of the newly created areas etc.

If the newly created areas are to maintain OAM connectivity then join one NE in one area to an NE in the other area via level 2 routing (adding a level 2 link if required).

9.2 Merging Existing Routing Areas

Joining areas into one or more areas is a much simpler process than splitting level 1 routing areas. In this case the areas which are to be joined need to have an area address in common. In fact, only the nodes which have a link across the area boundary need to have an area address in common. The rest of the NEs within the area will pick up the new area address when they calculate the computed area address set.

By visiting the NEs which have a link which crosses the boundary between the areas to be joined and provisioning an area address to be common to both, a level 1 routing adjacency will be created. Once the adjacency is created the NEs in each area will be able to route to each other and exchange LSPs. The level 2 routers which would have been used previously to route between these areas can now be removed.
The preceding assumes there are two level 1 routers, 1 in each area to be joined which are connected. If these two areas were only connected via level 2 routing, connectivity can be maintained by provisioning the circuit connecting the level 2 routers to be a level 1&2 circuit, or a level 1 circuit only.

Since it may be desirable to have only one area address in use across these newly joined areas, the following should be done. Provision the common area address to be used on all NEs in the newly formed area. Once all NEs have been provisioned with this information all other area addresses in use in the area other than the one in common may now be deleted from the NEs.

The number of areas joined as an operation should also be equal to 2 for simplicity. This way when joining the areas, if a problem arises (typing error etc.) there will be a free area address available to be used to rectify the problem.

9.3 Level 1 Partition Repair

The IS-IS standard offers a means of repairing a Level 1 area that was physically split in two or more partitions. An area may become partitioned as a result of failure of one or more Level 1 physical adjacencies in the area (i.e. a line breakdown).

It is possible to mend the area if the Level 2 sub-domain is not partitioned itself. The repair procedure involves electing Level 2 ISs (so-called Partition Designated Level 2 ISs) that stand for their partition. A partitioned area is then healed through so-called virtual links established across the Level 2 sub-domain between PDISs. Such systems are capable of tunneling CLNP and Level 1 IS-IS traffic for the purpose of conveying it through the Level 2 sub-domain.

First, it is worth giving an accurate definition of an area partition : two systems belong to the same partition if and only of they belong to the same routing area and they can communicate with each other by **only** traversing **non-virtual** links located within the area boundary. The largest partition of an area is the area itself, when it is not partitioned ! The term “virtual” refers to the characteristics of tunneling paths established by Partition Designated ISs.

Within each partition, a Partition Designated Level 2 Intermediate System is elected among the set of Level 2 ISs that :

1. implement the partition repair option,

2. are attached to the Level 2 sub-domain via the **default** metric,

3. act as Level 1 IS inside their partition (at least one of their circuits is not labeled Level 2-only),

4. are reachable without traversing **any virtual** links.

There is exactly one elected Partition Designated IS per partition. Obviously, only the partitions served by an PDIS can be mended !

When applied to the SONET field, relatively few systems might head a partition : this would typically be the role of Gateway Network Elements, or some DCN routers. In any case, a PDIS shall be located at the area boundary.

A Level 1 repair path is established between each PDISs. At each end of the path, the Partition Designated IS creates a new Level 1 adjacency, which is tagged as “virtual” and advertises other Level 1 ISs of its partition by inserting the new Level 1 neighbor in one of its Level 1 LSPs.

A Virtual Network Entity Title is formed by prepending the first listed area address from its Level 2 LSPs to the system ID of the Partition Designated IS represents each partition and then appending a ‘00’ Network Selector. A Partition Designated IS shall forward Level 1 IS-IS PDUs and CLNP PDUs addressed to the other partition by encapsulating them into CLNP DT PDU of which source address is its Virtual NET and the sink address is the Virtual NET of the virtual adjacency. In other words, intra-area information are tunneled through the Level 2 sub-domain.

Conversely, the receiving Partition Designated IS will decapsulate the user data of CLNP DT PDUs addressed to its Virtual NET and then apply the usual Level 1 IS-IS or CLNP procedures, after the actual content of the payload.

There are three general requirements to be fulfilled for putting partition repair at work :

1. The Level 2 sub-domain shall be fully connected. That is, any two Level 2 ISs of the domain must be able to communicate without leaving the domain and without traversing an intra-area link.

2. Each potential partition is fitted with at least one Level 2 system implementing the repair option and connected to the Level 2 sub-domain through the default metric.

3. The Level 2 sub-domain is able to route CLNP DT PDUs from and to Virtual NETs of systems that belong to the same routing area.

As a matter of fact, DCNs whose Level 2 subdomain is not entirely connected shall be regarded as ill-configured. Strictly speaking, this first requirement is NOT brought by partition repair. It is a topological constraint of Level 2 IS-IS routing, which is merely assumed by the Partition Repair procedures.

The precise locations of Repair-capable systems can not be predicted and might evolve with the growth of the area. For instance, in a linear topology where NEs form a collection of points along a segment, turning on partition repair at the two ends of the segment will make sense since this kind of topology is highly prone to partition.

Unfortunately, the third requirement is not accurately covered by ISO/IEC 10589 text. IS-IS standard omitted to describe a way of establishing a Level 1 repair path between two Partition Designated Intermediate Systems. By following ISO/IEC 10589 procedures, a Partition Designated IS can create a virtual adjacency with another Partition Designated IS without prior knowledge or warranty that they will be able to exchange CLNP PDUs from a Virtual NET to another. As a matter of fact, a blind application of ISO/IEC 10589 forwarding rules could even prevent such exchanges! Have a glance at this sample of routing domain :

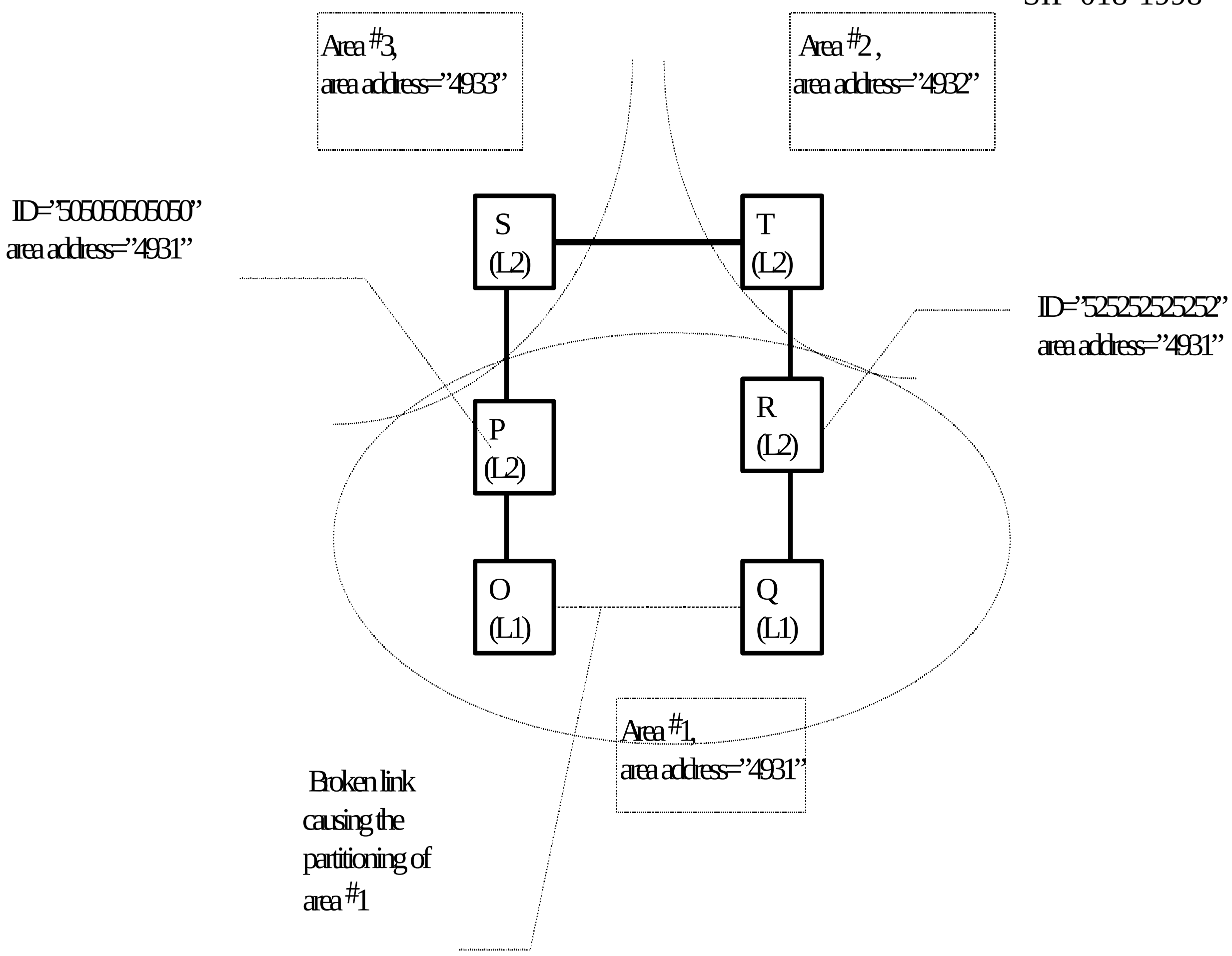


Figure 37 Non-repairable area with fully connected L2 sub-domain

Node P will create a virtual adjacency with node R and conversely node R will discover that node P heads another partition of area #1. Node P's virtual NET is "4931505050505000", Node R's virtual NET is "493152525252525200". Let node P transmit a Level 1 PDU to node R. At first, it encapsulates the IS-IS PDU into a CLNP DT PDU addressed to node R's Virtual NET. Then it routes the outer PDU. When the node S receives the PDU, it forwards it by following the regular IS-IS rules : it tries to match the sink address with one of its area addresses. If there is no match, it performs Level 2 routing : "consult the Level 2 forwarding database to determine the adjacency which is the next hop on the path to the *destination area*". Here comes the trouble : since the PDU is addressed to node R, the destination area is area #1 ! Thus, the next hop to area #1 could be either node T or node P, depending of the cost of the Level 2 links.

At the time of issue, the only workaround is the manual provisioning of routes on Level 2 ISs which encapsulated PDUs might go through (*à la OSPF* repair path setting). This solution works fine but puts much burden on the shoulders of the DCN managers. [1]

10 Name to Address Translation Resolution

10.1 TARP Solution Overview

The TID Address Resolution Protocol (TARP) is used to resolve the TID of a TL1 message to a CLNP address (NSAP). It is designed to prevent the user from manually mapping TIDs to NSAPs.

TARP uses a limited flooding technique to distribute TARP PDUs throughout the network. Note that there is concern that TARP may not scale to very large networks.

10.1.1 Provisioning of TARP Manual Adjacencies

TARP Manual Adjacencies are used for TARP PDUs propagation and origination in addition to the real adjacencies provided by network layer Routing Information Base.
There is a requirement to be satisfied for SONET management network architecture, when there are non-SONET NEs without TARP capability. The TARP protocol uses TIDs for addressing but non-SONET NEs such as commercial routers may not support TARP. This issue also rises when SONET NEs in a multi-vendor environment would try to pass traffic and some of the NEs do not support TARP. In order to make all SONET NEs (i.e., the ones that support TARP and the ones without TARP capability) interoperable, the NEs with the TARP capability should be provisionable to skip the NEs that do not support TARP. Therefore, it is necessary to support the manual TARP adjacency provisioning from a SONET NE. This capability allows the TARP request to hop through the non-SONET NEs.

The conclusion is to make sure the aforementioned capability is implemented in the SONET-NE.

10.2 Directory Services Solution Overview

10.2.1 The SONET Directory Service

Network Address resolution is one particular use of the SONET Directory Service (SDS) that is set up by ANSI T1.245 standard. T1.245 relies on X.500 models and services introduced by ISO/IEC 9594-1 and developed in other standards of the ISO/IEC 9594-x series.

[1] A defect report (ISO/IEC 10589 DR35) was submitted to ISO/JTC1 Secretariat with a proposed amendment to the standard that would allow an automated computation of routes to Virtual NETs by all Level 2 ISs of the domain. ISO/JTC1 next Plenary is due to vote upon the Defect Report 35 (June '98).

X.500 is a versatile directory service where entries are organized in a hierarchical manner to form a Directory Information Tree (DIT). Each entry (but the top one) is one of the subordinate entry of its immediate superior. A leaf entry has no subordinate.

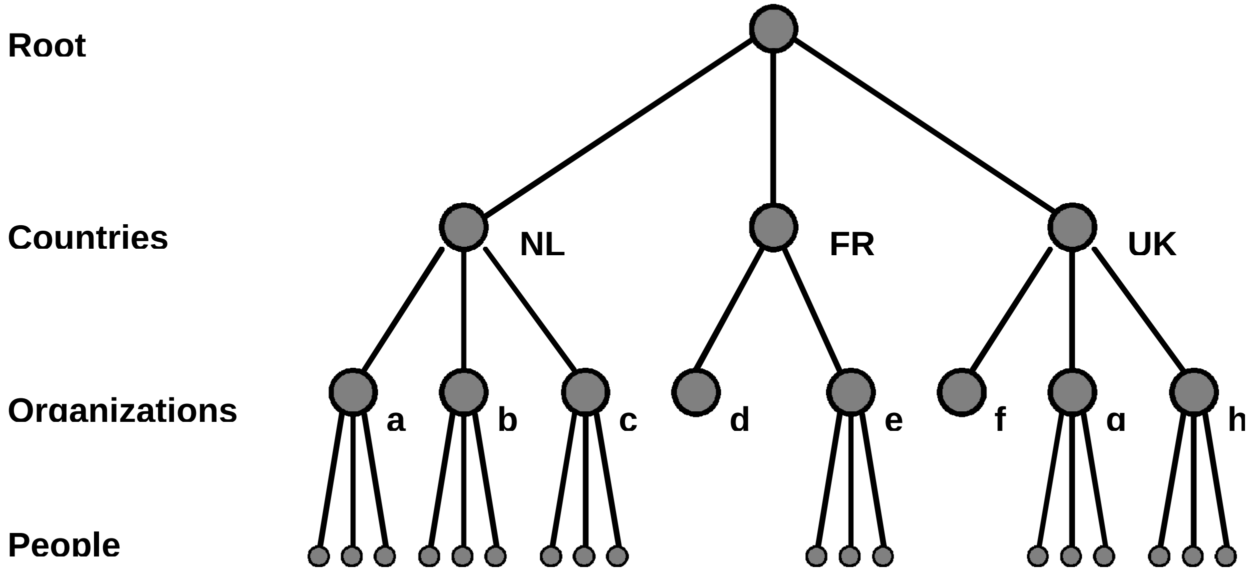


Figure 38 X.500 Directory Information Tree

An entry is an instance of a specific class that owns one or several attributes, one of which is called the naming attribute of the entry. The naming attribute defines the entry Relative Distinguished Name (RDN). **No two sibling entries may have the same RDN**. The sequence of all RDN from the top entry to the considered entry builds the entry's Distinguished Name (DN). An essential X.500 paradigm states that a DN uniquely identifies an entry, if the DN is valid within the DIT.

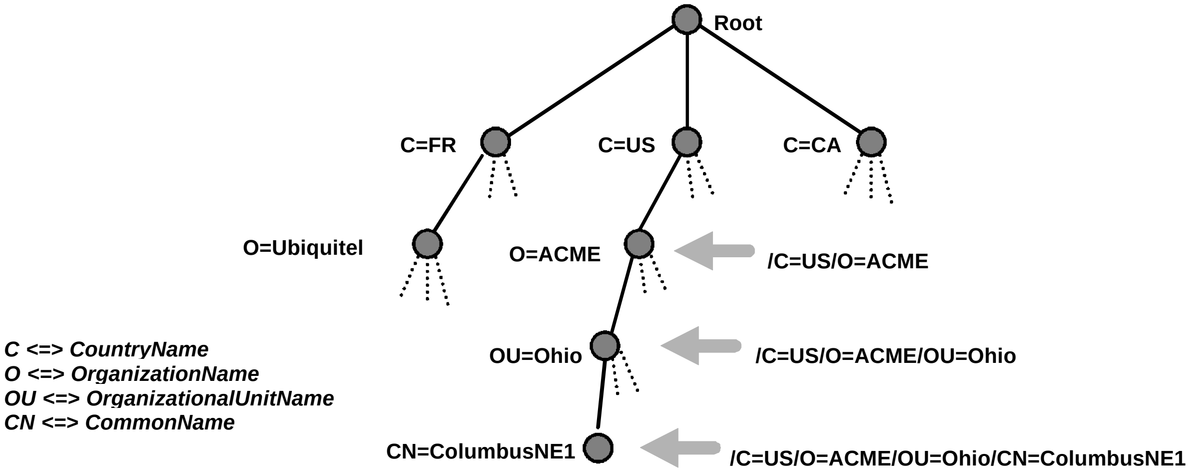


Figure 39 Naming of Directory objects

T1.245 defines a schema to insert TMN information describing SONET NEs from a management perspective into the DIT. The schema is loose enough to reflect the differences amongst NE architectures and accommodate the existence of mediation devices or miscellaneous proxy agents but still impose a naming hierarchy so that all items of information about a given NE are located below a stem ***tmnNE*** object, which stands for the NE.

tmnNE

Must contain	May contain
commonName	vendorName
entityAddress	localityName
managedElementId	neType
	proprietaryAddress

Figure 40 the *tmnNE* object class

An application related to a NE is represented by an ***applicationEntity*** entry (or an alias entry pointing to a ***applicationEntity*** entry of a proxy) that is subordinate to the root ***tmnNE*** entry. For convenience, one may insert an intermediate entry between ***tmnNE*** and ***applicationEntity***. This supplementary level is embodied by an ***applicationProcess*** object.

Information regarding the NE are to be automatically created and maintained, either by the NEs themselves or by proxy agents. The actual means of Directory population is thoroughly explained in T1.245 and (grossly) summarized in the following sketch where the operations are numbered according to their relative timing:

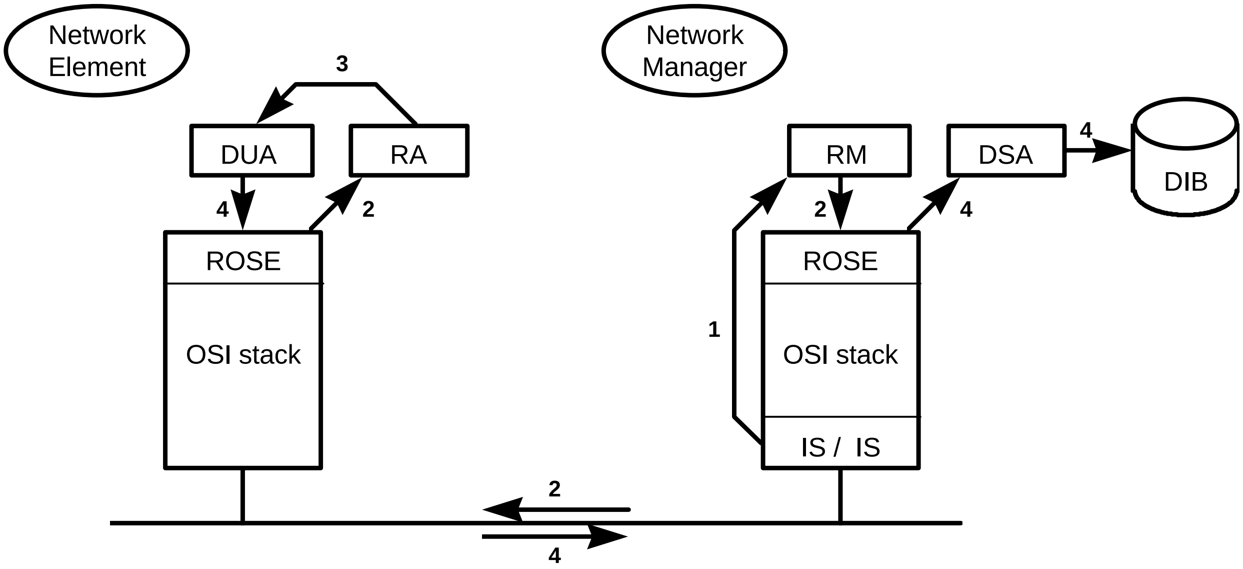


Figure 41 NE Discovery thanks to T1.245 procedures

A Registration Manager (RM) looks up[2] the set of routing information collected by IS-IS protocol to discover the systems within the same routing area. It establishes an association with each Registration Agent (RA) of the area and then transmits the DSA address(es) and an optional Name Prefix into a Registration Request Protocol (RRP) PDU conveyed by ROSE. A Name Prefix is syntactically equivalent to a DN. The knowledge of a Name Prefix alleviates the creation of NE-related objects when T1.245 registration is performed by the DUA.

All NEs of a given area shall share the same Local Naming Context (LNC). A LNC defines a partial branch in the DIT, beneath which NE entries reside. Use of LNC in the context of SONET is illustrated below: `/c=us/o=ACME/ou=NY-Metro` is the LNC of NE0 and NE1 entries. A RM would typically send out the watched area’s LNC as the value of the Name Prefix field of its RRP PDU.

[2] The means of communication between the Network layer and the RM is an implementation choice. The phrase “looks up” does not necessary imply that the RM IS TO periodically poll the local forwarding database.

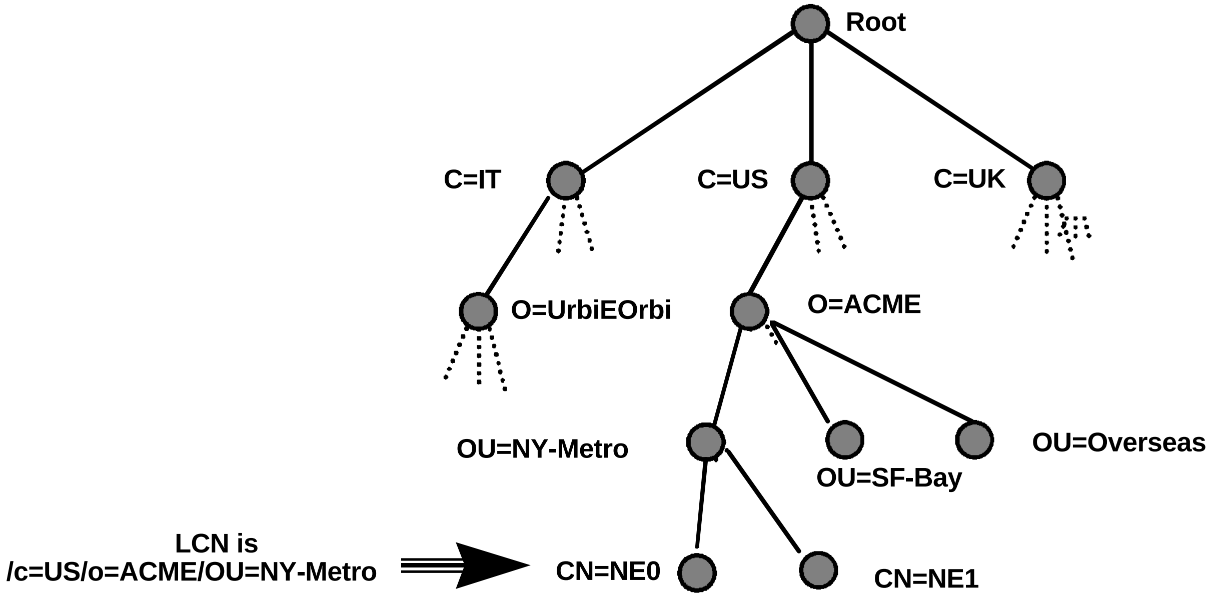


Figure 42 LNC of two sibling *tmnNE* entries

10.2.2 SIF NARSE documents as of today

Network Address Resolution Service Element (NARSE) is one of the mandatory services expected on Craft Terminals and Operation Systems. The aim of NARSE and the relationship with other ASEs are described in the on-going SIF-DN-9706-045 umbrella document. Though, programmatic access to NARSE is already well defined by SIF-RL-9704-043. This latter document profiled the IETF LDAP **API** to offer a unified interface for NARSE and proposes different mappings for the service provision, according to the desired profiles.

Although NARSE core profile can be rendered by TARP or by other means (local directory lookup), DAP or LDAP are the natural matches for the full featured NARSE, since the selected API names objects by enforcing T1.245 rules.

SIF-RL-9706-46 defines a LDAP version of ANSI T1.245 schema. It is the natural companion of SIF-RL-9704-043. Though, its scope is broader than NARSE: the document presents a comprehensive mapping of ANSI T1.245 schema and structure rules to the LDAP syntax.

10.2.3 Naming of OSI Applications

Within an OSI environment, an application is identified by its AE-Title. An AE-Title names an Application Entity which is part of an Application Process. An AE-Title is formed by concatenating an AP-Title (identifying the Application Process) with a AE-Qualifier that specifically points the entity. It is not uncommon that an Application Process owns only one Application Entity; though the AE-Qualifier can not be empty.

Knowing the AE-Title of a peer application is a good step to start with when seeking the peer address.

AE-Titles can take two possible forms. The first one uses registered Object Identifiers (still enforcing the Process/Qualifier hierarchy). The second syntax is equivalent to a Distinguished Name: the AP-Title is a sequence of RDNs and the AE-Qualifier itself is a RDN. In other words, a Name form AE-Title is a path in the DIT that leads down to the entry representing the application.

10.2.4 Retrieval of OSI Application Addresses

The information on a NE application is recorded in a ***applicationEntity*** object within the DIT. The following items are stored therein:

applicationEntity	
Must contain	**May contain**
commonName presentationAddress	description localityName organizationName organizationalUnitName supportedApplicationContext seeAlso

Figure 43 the *applicationEntity* object class

The OSI address to be called for associating with the application is contained in the *presentationAddress* attribute. Although the *supportedApplicationContext* attribute may be omitted according to ISO/IEC standards, T1.245 mandates its presence for the purpose of Network Address resolution.

10.2.4.1 Searching the OSI Address of an application having a Name form AE-Title

The retrieval of the address of a peer application is a trivial operation if the remote entity was assigned a Name form AE-Title. Once the calling application was given the peer's AE-Title, it only has to request its local NARSE provider to search for the entry whose DN value is the tendered AE-Title, with an automatic dereferencing of aliases if any are encountered. As a matter of fact, the AE-Title is actually a path to an ***applicationEntity*** object or an alias pointing such an object.

In the search argument, the user can ask either to fetch the entire entry content or to only select the *presentationAddress* attribute.

10.2.4.2 Searching the OSI Address of a TL1 Agent

Deployed TL1 agents usually have no AE-Title. One shall find other ways of going through the DIT.

The idea is to scan all the vertices of the DIT located below the NE root object to find an ***applicationEntity*** entry whose *supportedApplicationContext* attribute value matches the TL1 Application Context.

10.2.4.2.1TARP/X.500 Directory schema

According to SIF-006-1996, a TMN system containing a TL1 application entity shall have exactly one subordinate ***applicationEntity*** entry or ***apaeAlias*** entry (the latter option being handy to accommodate TL1 mediation devices for instance) below the ***tmnNE*** entry that represents the NE.

In both cases, the entry RDN value is merely the TID of the TL1 Agent. If an *applicationEntity* object is used, the *supportedApplicationContext* attribute shall be assigned the `tl1PeerComm` value defined in GR-253 while the *presentationAddress* takes the same value as a TID to Network Address TARP resolution and the use of GR-253 default selectors that TL1 would have yielded.

10.2.4.2.2Fitting legacy NEs into the DIT

Legacy SONET NEs are unable to register into the DIT on their own, either because they do not have sufficient Management Information for that purpose or simply because they can not access the X.500 DIB, in the case of TARP-only systems.

SIF-006-1996 addresses this concern by specifying a fashion of creating and filling up a ***tmnNE*** entry with the sole knowledge of the NE’s TID and its Network Address. The entry can be added by the NE itself or by a other system, called a TARP/X.500 Registration Proxy, on the behalf of the NE. The TID value is used as the RDN of the ***tmnNE***.

A Registration Proxy waits on NEs located within a given routing area. SIF-006-1996 mandates that ***tmeNE*** entries created by such a proxy be immediately (i.e. at the very next level) anchored under the Local Naming Context of the area. The area LNC can wisely be obtained by the proxy from the area’s RM.

This configuration is depicted in Figure 42 on page 74, if one assumes that “NE0” and “NE1” are respectively the TIDs of NE0 and NE1 and the area’s LNC is “/c=US/o=ACME/ou=NY-Metro”.

10.2.4.2.3Scenario for remote login to TARP-only systems

We have undertaken a stepwise journey from the broad scope of T1.245 to a view closer to the SONET world. It is worth highlighting that the combination of SIF-013-1997 and SIF-006-1996 permit a simple solution for the most common operation performed by deployed management systems: getting connected to the TL1 agent of a NE by only knowing its TID.

An area-wide TID to Network Address resolution consists of searching the subtree rooted by ./cn=<Target_Identifier> vertex and selecting the *presentationAddress* attribute of the ***applicationEntity*** object whose *supportedApplicationContext* attribute value is tl1PeerComm.

A domain-wide TID to Network Address resolution consists of starting a search from the domain’s root vertex (e.g. /c=US/o=ACME) and selecting the *presentationAddress* attribute of the ***applicationEntity*** object whose *supportedApplicationContext* attribute value is tl1PeerComm and whose *commonName* attribute value is equal to <Target_Identifier>.

In both cases, the requests can be derived from a well-defined pattern that takes only one variable parameter: the TID of the sought NE.

10.2.4.3 Searching other kinds of applications

Any OSI Application other than TL1 must be assigned a Name form AE-Title and must provision their application context in the *supportedApplicationContext* attribute of the corresponding entry.

Such applications may therefore be searched in the DIT either on the basis of their AE-Title or on the basis of the NE’s Distinguished Name and the application context.

10.3 Deployment considerations

10.3.1 Location of T1.245 Registration Manager

A RM uses Level 1 routing information to identify the systems intervening in the routing area. At least one RM must reside in a area where ANSI T1.245 registration takes place.

10.3.2 Location of T5GW(s)

As explained in SIF-006-1996, TARP to X.500 Interworking has two folds: being able to insert the legacy TARP NEs in to the DIT and rendering a TID to NET resolution that mimics TARP connectivity properties, from the standpoint of TARP systems initiating the resolution.

Chiefly, the first T5GW function can be described as a Registration Agent or Proxy, which will autonomously populate the DIB.

Conversely, the second T5GW function can be seen as a TID to NET server that acts in response to stimuli of remote TARP processors or of local TARP users.

10.3.2.1 T5GW as a X.500 Registration Proxy

The Registration Proxy mode is subjected to the same set of constraints as the RM because it has to extract Level 1 information from the local routing engine. Thus, at least one X.500 Registration Proxy is needed per watched area.

The proxy has to add new entries in the DIT and update their contents whenever a change of a TID/NET pairing is detected. Activating only one Registration Proxy at a time within a given area is a efficient method to avoid race conditions onto the DIB.

10.3.2.2 T5GW as a TID to NET Server Proxy

One Server Proxy should reside in an area to handle TARP requests from TARP-only NEs.

By installing one Server Proxy and one Registration Proxy per area, one can eliminate the domain-wide TARP propagation.

10.3.2.3 T5GW and T1.245 RM

The existence of a RM function within the local area greatly simplifies the deployment of a T5GW, because the T5GW can benefit from the RM information by playing the RA role.

10.3.2.4 Location of DSA(s)

On the contrary of the preceding types of systems, no constraint applies to the embodiment of the Directory Information Base and the exact location of the Directory System Agent(s).

ANSI T1.245 and SIF-006-1996 require a DAP-connector for the X.500-capable systems to talk to at least one DSA that participates to the Directory System.

LDAP front-ends can be placed anywhere in the Directory System, provided they are able to deal with ANSI T1.245 schema.

10.3.3 Possible use of T5 NEs

The term “T5” qualifies systems that have a TARP processor as well as direct connection to the Directory System. These hybrid systems have a strong property: they can be put anywhere in a area without disturbing the pre-existing TARP-only NEs. A system that is unable of propagating TARP PDUs can indeed jeopardize the correct operation of TARP-based managers, because its insertion reduces the overall TARP connectivity and may create propagation black holes.

When expanding a area where TARP is the dominant means of Network Addresses retrieval, the insertion of T5 NEs offers a stepwise extension of the area that remains upward compatible with legacy OSs (i.e. those relying on TARP for performing TID to NET translation). The planner may decide to only add a RM if TARP-only NEs need not to be recorded in the DIT or he/she may choose to install a RM and a T5GW so that the X.500 directory shows all NEs.

The utility of T5 NEs is more debatable in the case of the settlement of a brand new area. In such cases, T5 systems mainly lessen the scheduling constraints during the deployment, for the installed NEs can be managed by legacy TARP-based OSs until the new (or updated) managers gain a plain access to X.500 DIT.

Appendix A Document Contributors

Arnston, Bob	Tellabs Operations, Inc.
Dayton, Dean	Applied Innovation, Inc.
Degieux, Marc	ATOS
Ginsberg, Les	Vertel
Harvey, Jim	Ameritech
Henderson, Jeff	NORTEL
Hunt, Christopher	Lucent Technologies
Jamal, Rashid	Sprint
Klish, II, Cypryan (Kip)	NORTEL
Learman, Jeff	Open Networks Engineering
Pelosi, Steve	Fujitsu Network Communications, Inc.
Razzaghi, Heidi	NEC America, Inc.
Roman, Ron	Bellcore
Ross, Jeffrey	Fujitsu Network Communications, Inc.
Schriner, Don	CISCO Systems
Sharma, Sandeep	Alcatel Network Systems, Inc.
Sharman, Corey	Fujitsu Network Communications, Inc.
Tanabe, Junji	NEC America, Inc.
Thorp, Don	Ameritech
Truskowski, Mike	CISCO Systems
Wu, Li-Ran	Hitachi Telecom (USA), Inc.

This document has received the approval of the SONET Interoperability Forum (SIF).

August 13, 1998