Network Working Group
Request for Comments: 1xxx
Obsoletes RFC 1195

Ross W. Callon
Digital Equipment Corporation
January, 1993

Use of OSI IS-IS for Routing in TCP/IP and Multi-Protocol Environments

Status of this Memo

This Internet Draft specifies a protocol on the IAB Standards Track for the internet community, and requests discussion and suggestions for improvements. Please refer to the current edition of the "IAB Official Protocol Standards" for the standardization state and status of this protocol. Distribution of this memo is unlimited.

The document is an Internet Draft. Internet Drafts are working documents of the Internet Engineering Task Force (IETF), its Areas, and its Working Groups. Note that other groups may also distribute working documents as Internet Drafts.

Internet Drafts are draft documents valid for a maximum of six months. Internet Drafts may be updated, replaced, or obsoleted by other documents at any time. It is not appropriate to use Internet Drafts as reference material or to cite them other than as a "working draft" or "work in progress". Please check the 1id-abstracts.txt listing contained in the internet-drafts Shadow Directories on nic.ddn.mil, nnsc.nsf.net, nic.nordu.net, ftp.nisc.sri.com, or munnari.oz.au to learn the current status of any Internet Draft.

Abstract

This RFC specifies an integrated routing protocol, based on the OSI Intra-Domain IS-IS Routing Protocol, which may be used as an interior gateway protocol (IGP) to support TCP/IP as well as OSI. This allows a single routing protocol to be used to support pure IP environments, pure OSI environments, and multi-protocol environments. This specification was developed by the IS-IS working group of the Internet Engineering Task Force. The final international standard version of the OSI IS-IS protocol is contained in ISO 10589 [1].

Comments should be sent to “isis@merit.edu”.

Contents

Figures

1 Introduction: Overview of the Protocol

The TCP/IP protocol suite has been growing in importance as a multi-vendor communications architecture. With the anticipated emergence of OSI, we expect coexistence of TCP/IP and OSI to continue for an extended period of time. There is a critical need for routers to support both IP traffic and OSI traffic in parallel.

There are two main methods that are available for routing protocols to support dual OSI and IP routers. One method, known as “Ships in the Night”, makes use of completely independent routing protocols for each of the two protocol suites. This specification presents an alternate approach, which makes use of a single integrated protocol for interior routing (i.e., for calculating routes within a routing domain) for both protocol suites.

This integrated protocol design is based on the OSI Intra-domain IS-IS routing protocol [1], with IP-specific functions added. This RFC is considered a companion to the OSI IS-IS Routing spec, and will only describe the required additional features. OSI IS-IS supports routing for the OSI ConnectionLess Network Protocol (CLNP [4]).

By supporting both IP and OSI CLNP traffic, this integrated protocol design supports traffic to IP hosts, OSI end systems, and multi-protocol end systems. This approach is “integrated” in the sense that the IS-IS protocol can be used to support pure-IP environments, pure-OSI environments, and multi-protocol environments. In addition, this approach allows interconnection of multi-protocol (IP and OSI) routing domains with other multi-protocol domains, with IP-only domains, and with OSI-only domains.

IS-IS may be extended to support other protocol suites in addition to CLNP and IP. For example, IS-IS has been extended to support DECnet™ [9]. The approach specified in this RFC is compatible with these additional extensions. However, the details of how other protocols are supported using IS-IS are outside of the scope of this specification.

This specification obsoletes RFC 1195. The differences between this specification and RFC 1195 are discussed in section 6.

The protocol specified here is based on the work of the IETF IS-IS working group.

1.1 What the Integrated IS-IS offers

The Integrated IS-IS provides a single routing protocol which will simultaneously provide an efficient routing protocol for TCP/IP, and for OSI CLNP. This design makes use of the OSI IS-IS routing protocol, augmented with IP-specific information. This design provides explicit support for IP subnetting, variable subnet masks, TOS-based routing, and external routing.

Both OSI and IP packets are forwarded “as is” — i.e., they are transmitted directly over the underlying link layer services without the need for mutual encapsulation. The Integrated IS-IS is a dynamic routing protocol, based on the SPF (Dijkstra) routing algorithm.

The protocol described in this specification allows for mixing of IP-only, OSI-only, and dual (IP and OSI) routers, as defined below.

An IP-only IS-IS router (or "IP-only" router) is defined to be a router which: (i) Uses IS-IS as the routing protocol for IP, as specified in this report; and (ii) Does not otherwise support OSI protocols. For example, such routers would not be able to forward OSI CLNP packets.

An OSI-only router is defined to be a router which uses IS-IS as the routing protocol for OSI, as specified in [1]. Generally, OSI-only routers may be expected to conform to OSI standards, and may be implemented independent of this specification.

A dual IS-IS router (or "dual" router) is defined to be a router which uses IS-IS as a single integrated routing protocol for both IP and OSI, as specified in this report.

This approach does not change the way that IP packets are handled. IP-only and dual routers are required to conform to the Requirements for IP Routers [10]. The Integrated IS-IS protocol described in this report outlines an Interior Gateway Protocol (IGP) which will provide routing within a TCP/IP routing domain (i.e., autonomous system). Other aspects of router functionality (e.g., operation of ICMP, ARP, EGP, etc.) are not affected by this proposal.

Similarly, this approach does not change the way that OSI packets are handled. There will be no change at all to the contents nor to the handling of ISO 8473 Data packets and Error Reports, nor to ISO 9542 Redirects and ES Hellos. ISO 9542 IS Hellos transmitted on LANs are similarly unchanged. ISO 9542 IS Hellos transmitted on point-to-point links are unchanged except for the addition of IP-related information. Similarly, other OSI packets (specifically those involved in the IS-IS intra-domain routing protocol) remain unchanged except for the addition of IP-related information.

This approach makes use of the existing IS-IS packets, with IP-specific fields added. Specifically: (i) the protocols supported by each router, as well as each router's IP addresses, are specified in ISO 9542 IS Hello, IS-IS Hello and Link State Packets; (ii) internally reachable IP addresses are specified in all Link State Packets; and (iii) externally reachable IP addresses, and external routing protocol information, are specified in level 2 Link State Packets. The detailed encoding and interpretation of this information is specified in sections 3, 4, and 5 of this RFC.

The protocol described in this report may be used to provide routing in an IP-only routing domain, in which all routers are IP-only. Similarly, this protocol may be used to provide routing in a pure dual domain, in which all routers are dual. Finally, this protocol may be used to provide routing in a mixed domain, in which some routers are IP-only, some routers are OSI-only, and some routers are dual. The specific topological restrictions which apply in this latter case are described in detail in section 1.4 ("Support of Mixed Routing Domains"). The use of IS-IS for support of pure OSI domains is specified in [1].

This protocol specification does not constrain which network management protocol(s) may be used to manage IS-IS-based routers. An SNMP-based management information base (MIB) for managing IP-only, OSI-only, and dual routers, is the subject of a separate companion document [14].

1.2 Overview of the ISO IS-IS Protocol

The IS-IS Routing Protocol [1] has been developed in ISO to provide routing for pure OSI environments. In particular, IS-IS is designed to work in conjunction with ISO 8473 (The ISO Connectionless Network Layer Protocol [4]), and ISO 9542 (The ISO End System to Intermediate System Protocol [5]). This section briefly describes the manner in which IS-IS is used to support pure OSI environments. Enhancements for support of IP and multi-protocol environments are specified elsewhere in this report.

In IS-IS, the network is partitioned into “routing domains”. The boundaries of routing domains are defined by network management, by setting some links to be “external domain”. If a link is marked as “external domain”, no IS-IS routing messages are sent on that link.

ISO is currently working on a standard for inter-domain routing (i.e., for routing between separate autonomous routing domains) [6]. This is currently at the "committee draft" stage. Until this protocol is available, manual configuration is used. The link is statically configured with the set of address prefixes reachable via that link, and with the method by which they can be reached (such as the DTE address to be dialed to reach that address, or the fact that the DTE address should be extracted from the IDP portion of the ISO address).

OSI IS-IS routing makes use of two-level hierarchical routing. A routing domain is partitioned into “areas”. Level 1 routers know the topology in their area, including all routers and end systems in their area. However, level 1 routers do not know the identity of routers or destinations outside of their area. Level 1 routers forward all traffic for destinations outside of their area to a level 2 router in their area. Similarly, level 2 routers know the level 2 topology, and know which addresses are reachable via each level 2 router. However, level 2 routers do not need to know the topology within any level 1 area, except to the extent that a level 2 router may also be a level 1 router within a single area. Only level 2 routers can exchange data packets or routing information directly with external routers located outside of the routing domains.

ISO provides for flexible, variable length addresses (known as Network Service Access Point Addresses, or NSAPs [7]), which allow for multi-level hierarchical address assignment. These addresses provide the flexibility needed to simultaneously solve two critical problems: (i) How to administer a worldwide address space; and (ii) How to assign addresses in a manner which makes routing feasible in a worldwide Internet. However, assignment of addresses needs to be done with great care, if the potential advantages of this addressing flexibility are to be realized in actual networks [8]. The ISO NSAP address format is illustrated in figure 1.

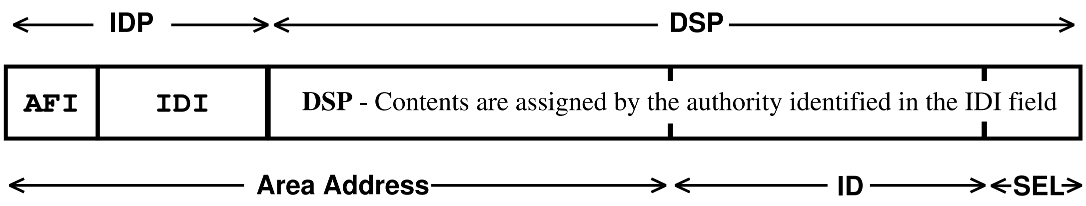


Figure 1 - ISO Hierarchical Address Structure

For administrative purposes, ISO addresses are subdivided into the Initial Domain Part (IDP), and the Domain Specific Part (DSP). The IDP is the part which is standardized by ISO, and specifies the format and authority responsible for assigning the rest of the address. The IDP is further subdivided into the Authority and Format Identifier, and the Initial Domain Identifier (IDI). The DSP is assigned by whatever addressing authority is specified by the IDP.

For routing purposes, ISO addresses are subdivided into the area address, the system identifier (ID), and the NSAP selector (SEL). The area address identifies both the routing domain and the area within the routing domain. Generally, the area address corresponds to the IDP plus a high-order part of the DSP. The ID field may be from 1 to 8 octets in length, but uses a single fixed length for any one routing domain. The selector is always one octet in length.

Usually, all nodes in an area have the same area address. However, sometimes an area might have multiple addresses. Motivations for allowing this are:

- It might be desirable to change the address of an area. The most graceful way of changing an area from having address A to having address B is to first allow it to have both addresses A and B, and then after all nodes in the area have been modified to recognize both addresses, then one by one the nodes can be modified to “forget” address A.

- It might be desirable to merge areas A and B into one area. The method for accomplishing this is to, one by one, add knowledge of address B into the A partition, and similarly add knowledge of address A into the B partition.

- It might be desirable to partition an area C into two areas, A and B (where “A” might equal “C”, in which case this example becomes one of removing a portion of an area). This would be accomplished by first introducing knowledge of address A into the appropriate nodes (those destined to become area A), and knowledge of address B into the appropriate nodes, and then one by one removing knowledge of address C.

Since OSI addressing explicitly identifies the area, it is very easy for level 1 routers to identify packets going to destinations outside of their area, which need to be forwarded to level 2 routers.

In IS-IS, there are two types of routers:

- Level 1 intermediate systems — these nodes route based on the ID portion of the ISO address. They route within an area. They recognize, based on the destination address in a packet, whether the destination is within the area. If so, they route towards the destination. If not, they route to the nearest level 2 router.

- Level 2 intermediate systems — these nodes route based on the area address (i.e., on the combination of [IDP, HO-DSP]). They route towards areas, without regard to the internal structure of an area. A level 2 IS may also be a level 1 IS in one area.

A level 1 router will have the area portion of its address manually configured. It will refuse to become a neighbor with a node whose area addresses do not overlap its area addresses. However, if level 1 router has area addresses A, B, and C, and a neighbor has area addresses B and D, then the level 1 router will accept the other node as a neighbor.

A level 2 router will accept another level 2 router as a neighbor, regardless of area address. However, if the area addresses do not overlap, the link would be considered by both routers to be “level 2 only”, and only level 2 LSPs would flow on the link. External links (to other routing domains) must be from level 2 routers.

IS-IS provides an optional partition repair function. In the unlikely case that a level 1 area become partitioned, this function (if implemented) allows the partition to be repaired via use of level 2 routes.

IS-IS requires that the set of level 2 routers be connected. Should the level 2 backbone become partitioned, there is no provision for use of level 1 links to repair a level 2 partition.

In unusual cases, a single level 2 router may lose connectivity to the level 2 backbone. In this case the level 2 router will indicate in its level 1 LSPs that it is not “attached”, thereby allowing level 1 routers in the area to route traffic for outside of the domain to a different level 2 router. Level 1 routers therefore route traffic to destinations outside of their area only to level 2 routers which indicate in their level 1 LSPs that they are “attached”.

An end system may autoconfigure the area portion of its address by extracting the area portion of a neighboring router’s address. If this is the case, then an endnode will always accept a router as a neighbor. Since the standard does not specify that the end system *must* autoconfigure its area address, an end system may be configured with an area address. In this case the end system would ignore router neighbors with non-matching area addresses.

Special treatment is necessary for broadcast subnetworks, such as LANs. This solves two sets of issues: (i) In the absence of special treatment, each router on the subnetwork would announce a link to every other router on the subnetwork, resulting in n-squared links reported; (ii) Again, in the absence of special treatment, each router on the LAN would report the same identical list of end systems on the LAN, resulting in substantial duplication.

These problems are avoided by use of a “pseudonode”, which represents the LAN. Each router on the LAN reports that it has a link to the pseudonode (rather than reporting a link to every other router on the LAN). One of the routers on the LAN is elected “designated router”. The designated router then sends out an LSP on behalf of the pseudonode, reporting links to all of the routers on the LAN. This reduces the potential n-squared links to n links. In addition, only the pseudonode LSP includes the list of end systems on the LAN, thereby eliminating the potential duplication (for further information on designated routers and pseudonodes, see [1]).

The IS-IS provides for optional Quality of Service (QOS) routing, based on throughput (the default metric), delay, expense, or residual error probability. This is described in greater detail in section 3.8, and in [1].

IS-IS has a provision for authentication information to be carried in all IS-IS PDUs. Currently the only form of authentication which is defined is simple passwords. A password may be associated with each link, each area, and with the level 2 subdomain. A router not in possession of the appropriate password(s) is prohibited from participating in the corresponding function (i.e., may not initialize a link, be a member of the area, or a member of the level 2 subdomain, respectively).

Procedures are provided to allow graceful migration of passwords without disrupting operation of the routing protocol. The authentication functions are extensible so that a stronger, cryptographically-based security scheme may be added in an upwardly compatible fashion at a future date.

1.3 Overview of the Integrated IS-IS

The Integrated IS-IS allows a single routing protocol to be used to route both IP and OSI packets. This implies that the same two-level hierarchy will be used for both IP and OSI routing. Each area will be specified to be either IP-only (only IP traffic can be routed in that particular area), OSI-only (only OSI traffic can be routed in that area), or dual (both IP and OSI traffic can be routed in the area).

This proposal does not allow for partial overlap of OSI and IP areas. For example, if one area is OSI-only, and another area is IP-only, then it is not permissible to have some routers be in both areas. Similarly, a single backbone is used for the routing domain. There is no provision for independent OSI and IP backbones.

Similarly, within an IP-only or dual area, the amount of knowledge maintained by routers about specific IP destinations will be as similar as possible as for OSI. For example, IP-capable level 1 routers will maintain the topology within the area, and will be able to route directly to IP destinations within the area. However, IP-capable level 1 routers will not maintain information about destinations outside of the area. Just as in normal OSI routing, traffic to destinations outside of the area will be forwarded to the nearest level 2 router. Since IP routes to subnets, rather than to specific end systems, IP routers will not need to keep nor distribute lists of IP host identifiers (note that routes to hosts can be announced by using a subnet mask of all ones).

The IP address structure allows networks to be partitioned into subnets, and allows subnets to be recursively subdivided into smaller subnets. However, it is undesireable to require any specific relationship between IP subnet addresses and IS-IS areas. For example, in many cases, the dual routers may be installed into existing environments, which already have assigned IP and/or OSI addresses. In addition, even if IP addresses are not already pre-assigned, the address limitations of IP constrain what addresses may be assigned. We therefore will not require any specific relationship between IP addresses and the area structure. The IP addresses can be assigned completely independently of the OSI addresses and IS-IS area structure. As will be described in section 3.2 (“Hierarchical Abbreviation of IP Reachability Information”), greater efficiency and scaling of the routing algorithm can be achieved if there is some correspondence between the IP address assignment structure and the area structure.

Within an area, level 1 routers exchange link state packets which identify the IP addresses reachable by each router. Specifically, zero or more [IP address, subnet mask, metric] combinations may be included in each Link State Packet. Each level 1 router is manually configured with the [IP address, subnet mask, metric] combinations which are reachable on each interface. A level 1 router routes as follows:

- If a specified destination address matches an [IP address, subnet mask, metric] reachable within the area, the packet is routed via level 1 routing.

- If a specified destination address does not match any [IP address, subnet mask, metric] combination listed as reachable within the area, the packet is routed towards the nearest level 2 router.

Flexible use of the limited IP address space is important in order to cope with the anticipated growth of IP environments. Thus an area (and by implication a routing domain) may simultaneously make use of a variety of different address masks for different subnets in the area (or domain). Generally, if a specified destination address matches more than one [IP address, subnet mask] pair, the more specific address is the one routed towards (the one with more "1" bits in the mask — this is known as "best match" routing).

Level 2 routers include in their level 2 LSPs a complete list of [IP address, subnet mask, metric] specifying all IP addresses reachable in their area. As described in section 3, this information may be obtained from a combination of the level 1 LSPs (obtained from level 1 routers in the same area), and/or by manual configuration. In addition, Level 2 routers may report external reachability information, corresponding to addresses which can be reached via routers in other routing domains (autonomous systems).

Default routes may be announced by use of a subnet mask containing all zeroes. Default routes should be used with great care, since they can result in “black holes”. For example, a default route announced using an internal metric at level 2 will prevent use of routes with external metrics at level 2 (since routes with internal metrics are always preferred to routes with external metrics). Similarly, a default route at level 1 will prevent use of level 2 routes within that area (see sections 3.5 through 3.7 for a further discussion of how this may be used).

The Integrated IS-IS provides optional Type of Service (TOS) routing, through use of the QOS feature from IS-IS.

1.4 Support of Mixed Routing Domains

The Integrated IS-IS proposal specifically allows for three types of routing domains:

- Pure IP
- Pure OSI
- Dual

In a pure IP routing domain, all routers must be IP-capable. IP-only routers may be freely mixed with dual routers. Some fields specifically related to OSI operation may be included by dual routers, and will be ignored by IP-only routers. Only IP traffic will be routed in an pure IP domain. Any OSI traffic may be discarded (except for the IS-IS packets necessary for operation of the routing protocol).

In a pure OSI routing domain, all routers must be OSI-capable. OSI-only routers may be freely mixed with dual routers. Some fields specifically related to IP operation may be included by dual routers, and will be ignored by OSI-only routers. Only OSI traffic will be routed in a pure OSI domain. Any IP traffic may be discarded.

In a dual routing domain, IP-only, OSI-only, and dual routers may be mixed on a per-area basis. Specifically, each area may itself be defined to be pure IP, pure OSI, or dual.

In a pure IP area within a dual domain, IP-only and dual routers may be freely mixed. Only IP traffic can be routed by level 1 routing within a pure-IP area.

In a pure-OSI area within a dual domain, OSI-only and dual routers may be freely mixed. Only OSI traffic can be routed by level 1 routing within a pure OSI area.

In a dual area within a dual routing domain only dual routers may be used. Both IP and OSI traffic can be routed within a dual area.

Within a dual domain, if both IP and OSI traffic are to be routed between areas then all level 2 routers must be dual.

1.5 Advantages of Using Integrated IS-IS

Use of the Integrated IS-IS protocol, as a single protocol for routing both IP and OSI packets in a dual environment, has significant advantages over using separate protocols for independently routing IP and OSI traffic.

An alternative approach is known as “Ships In the Night” (S.I.N.). With the S.I.N. approach, completely separate routing protocols are used for IP and for OSI. For example, OSPF [11] may be used for routing IP traffic, and IS-IS [1] may be used for routing OSI traffic. With S.I.N., the two routing protocols operate more or less independently. However, dual routers will need to implement both routing protocols, and therefore there will be some degree of competition for resources.

Note that S.I.N. and the Integrated IS-IS approach are not really completely separate options. In particular, if the Integrated IS-IS is used within a routing domain for routing of IP and OSI traffic, it is still possible to use other independent routing protocols for routing other protocol suites.

Optional extensions to IS-IS may be defined for routing other common protocol suites. However, such extensions are outside of the scope of this document. This section will compare Integrated IS-IS and S.I.N. for routing of IP and OSI only.

A primary advantage of the Integrated IS-IS relates to the network management effort required. Since the Integrated IS-IS provides a single routing protocol, within a single coordinated routing domain using a single backbone, this implies that there is less information to configure. This combined with a single coordinated MIB simplifies network management.

Note that the operation of two routing protocols with the S.I.N. approach is not really independent, since the routing protocols must share common resources. However, with the Integrated IS-IS, the interactions are explicit, whereas with S.I.N., the interactions are implicit. Since the interactions are explicit, again it may be easier to manage and debug dual routers.

Another advantage of the Integrated IS-IS is that, since it requires only one routing protocol, it uses fewer resources. In particular, less implementation resources are needed (since only one protocol needs to be implemented), less CPU and memory resources are used in the router (since only one protocol needs to be run), and less network resources are used (since only one set of routing packets need to be transmitted). Primarily this translates into a financial savings, since each of these three types of resources cost money. This implies that dual routers based on the Integrated IS-IS should be less expensive to purchase and operate than dual routers based on S.I.N.

Note that the operation of two routing protocols with the S.I.N. approach are not really independent, since they must share common resources. For example, if one routing protocol becomes unstable and starts to use excessive resources, the other protocol is likely to suffer. A bug in one protocol could crash the other. However, with the Integrated IS-IS, the interactions are explicit and are defined into the protocol and software interactions. With S.I.N., the interactions are implicit.

The use of a single integrated routing protocol similarly reduces the likely frequency of software upgrades. Specifically, if you have two different routing protocols in your router, then you have to upgrade the software any time EITHER of the protocols change. If you make use of a single integrated routing protocol, then software changes are still likely to be needed, but less frequently.

Finally, routing protocols have significant real time requirements. In IS-IS, these real time requirements have been explicitly specified. In other routing protocols, these requirements are implicit. However, in all routing protocols, there are real time guarantees which must be met in order to ensure correct operation. In general, it is difficult enough to ensure compliance with real time requirements in the implementation of a single real time system. With S.I.N., implementation of two semi-independent real-time protocols in a single device makes this more difficult.

Note that both Integrated IS-IS and S.I.N. allow for independence of external routes (for traffic from/to outside of the routing domain), and allow for independent assignment of OSI and TCP/IP addresses.

2 Symbols and Abbreviations

AA	Administrative Authority (a three octet field in the GOSIP version 2.0 NSAP address format)
AFI	Authority and Format Identifier (first octet of all OSI NSAP addresses — identifies format of the rest of the address)
CLNP	Connection-Less Network Protocol (ISO 8473, the OSI connectionless network layer protocol — very similar to IP)
DFI	DSP Format Identifier (a one octet field in the GOSIP version 2.0 NSAP address format)
ES	End System (The OSI term for a host)
ES-IS	End System to Intermediate System Routeing Exchange Protocol (ISO 9542 — OSI protocol between routers and end systems)

ICD	International Code Designator (ISO standard for identifying organizations)
IP	Internetwork Protocol (an Internet Standard Network Layer Protocol)
IS	Intermediate System (The OSI term for a router)
IS-IS	Intermediate System to Intermediate System Routeing Exchange Protocol (the ISO protocol for routing within a single routing domain)
IS-IS Hello	An Hello packet defined by the IS-IS protocol (a type of packet used by the IS-IS protocol)
ISH	An Hello packet defined by ISO 9542 (ES-IS protocol). (not the same as IS-IS Hello)
ISO	International Organization for Standardization (an international body which is authorized to write standards of many kinds)
LSP	Link State Packet (a type of packet used by the IS-IS protocol)
NLPID	Network Layer Protocol ID (A one-octet field identifying a network layer protocol)
NSAP	Network Service Access Point (a conceptual interface point at which the network service is made available)
SEL	NSAP Selector (the last octet of NSAP addresses, also called NSEL)
OSI	Open Systems Interconnection (an international standard protocol architecture)
RD	Routing Domain (the set of routers and end systems using a single instance of a routing protocol such as IS-IS)
SNPA	Subnetwork Point of Attachment (a conceptual interface at which a subnetwork service is provided)
TCP	Transmission Control Protocol (an Internet Standard Transport Layer Protocol)
TCP/IP	The protocol suite based on TCP, IP, and related protocols (the Internet standard protocol architecture)

3 Subnetwork Independent Functions

3.1 Exchange of Routing Information

The exchange of routing information between routers makes use of the normal routing packet exchange as defined in the OSI IS-IS routing spec, with additional IP-specific information added to the IS-IS routing packets.

The IS-IS protocol provides for the inclusion of variable length fields in all IS-IS packets. These fields are encoded using a “Code, Length, Value” triplet, where the code and length are encoded in one octet each, and the value has the length specified (from 0 to 254 octets). IS-IS requires that: “Any codes in a received PDU that are not recognised are ignored and passed through unchanged”. This requirement applies to all routers implementing IS-IS, including OSI-only, IP-only, and dual routers. This allows IP-specific information to be encoded in a manner which OSI-only routers will ignore, and also allows OSI-specific information to be encoded in a manner which IP-only routers will ignore.

IP-capable (i.e., all IP-only and dual) routers need to know what network layer protocols are supported by other routers in their area. This information is made available by inclusion of a “protocols supported” field in all IS-IS Hello and Link State Packets. This field makes use of the NLPID (Network Layer Protocol Identifier), which is a one-octet value assigned by ISO to identify network level protocols. NLPID values have been assigned to ISO 8473 and to IP.

IP-capable routers need to know the IP address of the adjacent interface of neighboring routers. This is required for sending ICMP redirects (when an IP-capable router sends an ICMP redirect to a host, it must include the IP address of the appropriate interface of the correct next-hop router). This information is made available by inclusion of the IP interface address in the IS-IS Hello packets. Specifically, each IS-IS Hello packet contains the IP address(es) of the interface over which the Hello is transmitted. The IS-IS allows multiple IP addresses to be assigned to each physical interface.

IP-capable routers in some cases may have interfaces to point-to-point links which do not have any IP address assigned (known as "unnumbered" links). According to the requirements of IP routers [10], a router which has unnumbered links must also have a special IP address, called a "router-ID". The router-ID is one of the router’s IP addresses, and is essentially used as if it were the IP address of all unnumbered interfaces. Where unnumbered links are used, routers need to know a "router-ID" of the neighbor router over that link. This is accomplished by inclusion of the "router-ID" field in IS-IS Hello packets transmitted over unnumbered point to point links.

In some cases, it will be useful for IP-capable routers to be able to determine an IP address(es) of all other routers at their level (i.e., for level 1 routers: all other routers in their area; for level 2 routers: all other level 2 routers in the routing domain). This is useful whenever an IP packet is to be sent to a router, such as for encapsulation or for transmission of network management packets. This information is made available by inclusion of IP address in LSPs. Specifically, each IS-IS LSP includes one or more IP addresses of the router which transmits the LSP. An IP-capable router is required to include at least one of its IP addresses in its LSPs, and may optionally include several or all of its IP addresses. Where a single router operates as both a level 1 and a level 2 router, it is required to include the same IP address(es) in its level 1 and level 2 LSPs.

IP-capable routers need to know, for any given IP destination address, the correct route to that destination. Specifically, level 1 routers need to know what IP addresses are reachable from each level 1 router in their area. In addition, level 1 routers need to find level 2 routers (for traffic to IP addresses outside of their area). Level 2 routers need to know what IP addresses are reachable internally (either directly, or via level 1 routing) from other level 2 routers, and what addresses are reachable externally from other level 2 routers. All of this information is made available by inclusion of IP reachable address information in the Link State Packets.

Internal (within the routing domain) and external (outside the domain) reachability information is announced separately in level 2 LSPs. Reachable IP addresses include a default metric, and may include multiple TOS-specific metrics. In general, for external routes, metrics may be of type “internal” (i.e., directly comparable with internal metrics) or of type “external” (i.e not comparable with the internal metric). A route using internal metrics (i.e., either announced as “IP internal reachability information”, or announced as “IP external reachability information” with an internal metric) is always preferred to a route using external metrics (i.e., announced as “IP external reachability information”, with an external metric).

In some rare cases, a router may have **only** unnumbered point to point links (i.e., a router might not have **any** operational interface with an assigned IP address). In such cases, the router is still required to have at least one IP address assigned to it, which can be used as its router-ID [10]. In this case, the router is required to announce a host route to itself in its LSPs (announced in the "IP internal reachability information" field, with the address specifying its router-ID and a mask of all ones). This is announced in level 1 LSPs, and is also announced in level 2 LSPs unless superceded by summary information (see section 3.2 below). This allows packets destined for that router to be delivered correctly. For example, this allows other routers to forward network management packets to this router.

The detailed encoding of the IP-specific information included in routing packets is provided in section 5 (“Structure and Encoding of PDUs”).

3.2 Hierarchical Abbreviation of IP Reachability Information

Level 2 routers include in their level 2 LSPs a list of all [IP address, subnet mask, metric] combinations reachable in their area. In general, this information may be determined from the level 1 LSPs from all routers in the area. If we ignore resource constraints, then it would be permissible for a level 2 router to simply duplicate all [IP address, subnet mask, metric] entries from all level 1 routers in its area (with appropriate metric adjustment), for inclusion in its level 2 LSP. However, in order for hierarchical routing to scale to large routing domain sizes, it is highly desired to abbreviate the reachable address information.

This is accomplished by manual configuration of summary addresses. Each level 2 router may be configured with one or more [IP address, subnet mask, metric] entries for announcement in their level 2 LSPs.

The set of reachable addresses obtained from level 1 LSPs is compared with the configured reachable addresses. Redundant information obtained from level 1 LSPs is not included in level 2 LSPs. Generally it is expected that the level 2 configured information will specify more inclusive addresses (corresponding to a subnet mask with fewer bits set to “1”). This will therefore allow one configured address/submask pair (or a small number of such pairs) to hierarchically supercede the information corresponding to multiple entries in level 1 LSPs.

The manually configured addresses are included in level 2 LSPs only if they correspond to at least one address which is reachable in the area. For manually configured level 2 addresses, the associated metric values to announce in level 2 LSPs are also manually configured. The configured addresses will supercede reachable address entries from level 1 LSPs based only on the IP address and subnet mask — metric values are not considered when determining if a given configured address supercedes an address obtained from a level 1 LSP.

If multiple summary addresses are configured, then each may be included in the level 2 LSP (provided that each summary address corresponds to at least one address which is reachable within the area). In some unusual cases a summary address may be configured which is itself hierarchically superceded by another manually configured summary address. In this case, both summary addresses are included in the level 2 LSP (i.e., a manually configured summary address will not be omitted from the level 2 LSP solely because it is itself superceded by a different manually configured summary address).

Any address obtained from a level 1 LSP which is not superceded by the manually configured information is included in the level 2 LSPs. In this case, the metric value announced in the level 2 LSPs is calculated from the sum of the metric value announced in the corresponding level 1 LSP, plus the distance from the level 2 router to the appropriate level 1 router. Note: If this sum results in a metric value greater than 63 (the maximum value that can be reported in level 2 LSPs), then the value 63 must be used. Delay, expense, and error metrics (i.e., those TOS metrics other than the default metric) will be included only if (i) the level 2 router supports the specific TOS; (ii) the path from the level 2 router to the appropropriate level 1 router is made up of links which support the specific TOS; and (iii) the level 1 router which can reach the address directly also supports the specific TOS for this route, as indicated in its level 1 LSP.

In general, the same [IP address, subnet mask] pair may be announced in level 1 LSPs sent by multiple level 1 routers in the same area. In this case (assuming the entry is not superceded by a manually configured entry), then only one such entry shall be included in the level 2 LSP. The metric value(s) announced in level 2 LSPs correspond to the minimum of the metric value(s) that would be calculated for each of the level 1 LSP entries.

A level 2 router will have IP addresses which are directly reachable via its own interfaces. For purposes of inclusion of IP reachable address information in level 2 LSPs, these “directly reachable” addresses are treated exactly the same as addresses received in level 1 LSPs.

Manually configured addresses may hierarchically supercede multiple level 1 reachable address entries. However, there may be some IP addresses which match the manually configured addresses, but which are not reachable via level 1 routing. If a level 2 router receives an IP packet whose IP address matches a manually configured address which it is including in its level 2 LSP, but which is not reachable via level 1 routing in the area, then the packet must be discarded. In this case, an error report may be returned (as specified in [10]), with the reason for discard specifying destination unreachable.

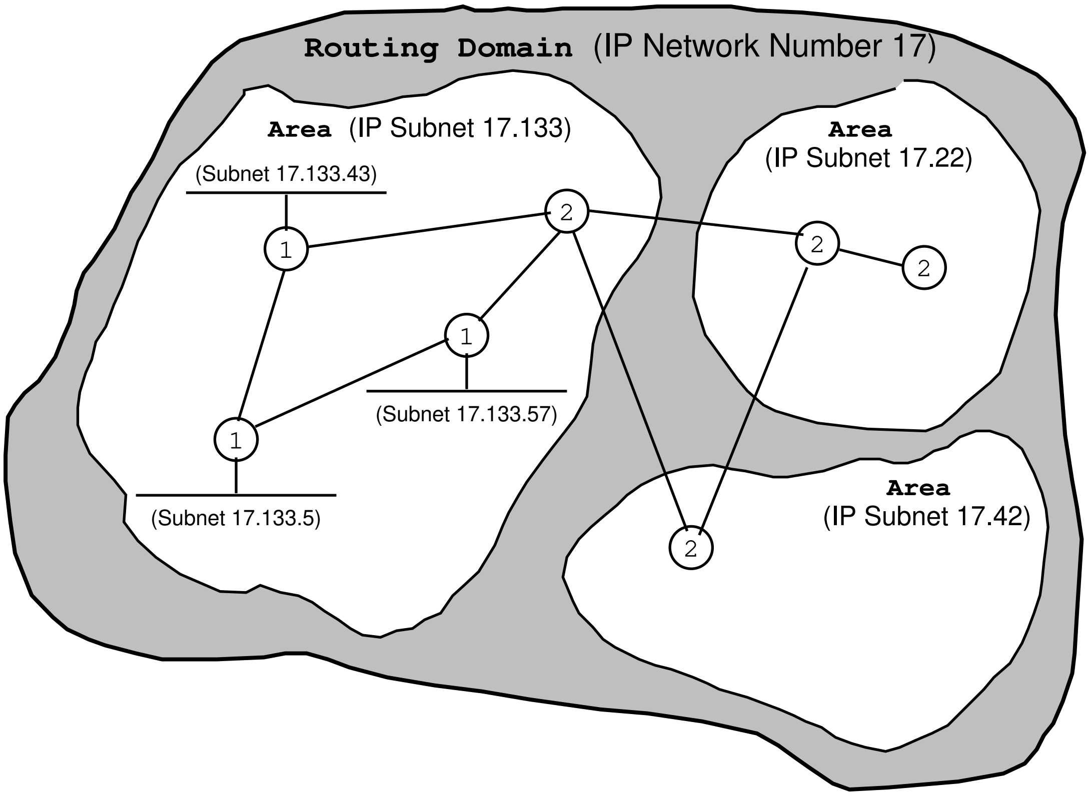


Figure 2 - An Example

An example is illustrated in figure 2. Suppose that the network number for the entire routing domain is 17 (a class A network). Suppose each area is assigned a subnet number consisting of the next 8 bits. The area may be further subdivided by assigning the next eight bits to each LAN in the area, giving each a 24 bit subnet mask (counting the network and subnet fields). Finally 8 bits are left for the host field. Suppose that for a particular area (given subnet number 17.133) there are a number of IP capable level 1 routers announcing (in the special IP entry in their level 1 LSPs) subnets 17.133.5, 17.133.43, and 17.133.57.

Suppose that in this example, in order to save space in level 2 LSPs, the level 2 routers in this area are configured to announce subnet 17.133. Only this one address needs to be announced in level 2 LSPs. Thus if an IP packet comes along for an address in subnet 17.133.5, 17.133.43 or 17.133.57, then other level 2 routers, in other areas, will know to pass the traffic to this area.

The inclusion of 17.133 in level 2 LSPs means that the three subnet addresses starting with 17.133 do not all have to be listed separately in level 2 LSPs.

If any traffic comes along that is for an unreachable address such as 17.133.124.7, then level 2 routers in other areas in this particular domain will think that this area can handle this traffic, will forward traffic to level 2 routers in this area, which will have to discard this traffic.

Suppose that subnet number 17.133.125 was actually reachable via some other area, such as the lower right hand area. In this case, the level 2 router in the left area would be announcing (in its level 2 LSPs — according to manually configured information) reachability to subnet 17.133. However, the level 2 router in the lower right area would be announcing (in its level 2 LSPs — according to information taken from its received level 1 LSPs), reachability to subnet 17.133.125.

Due to the use of “best match” routing, this works correctly. For example, suppose that a packet were to come from a source host located on subnet 17.133.5, and has a destination address for a host on subnet 17.133.125. Since subnet 17.133.125 is not located in the source (leftmost) area, the level 1 routers in that area would forward the packet to the nearest level 2 router. When the packet arrives at the level 2 router (still in the leftmost area), since the destination is not reachable in the area, the packet would be routed according to level 2 information towards the level 2 router in the lower right area (i.e., the packet would be routed towards the level 2 router which is announcing subnet 17.133.125 in its LSP, since this is a better match than the router which is announcing subnet 17.133). Generally, all traffic from other areas destined to subnet 17.133.125 would be sent to the level 2 router in the lower right area, and all other traffic to subnet 17.133 (i.e., traffic to any IP address starting with 17.133, but not starting with 17.133.125) would be sent to the level 2 router in the leftmost area.

3.3 Optimization of Route Leaking from Level 1 to Level 2

Normally, all routers which are both level 1 and level 2 routers will "leak" information learned from level 1 LSPs into level 2 (i.e., for each entry learned from an IP reachability information included in a received level 1 LSP, the router will pass the appropriate information into level 2 LSPs). In general this means that for areas with many level 2 routers, the same information from level 1 LSPs is included in multiple level 2 LSPs. This section describes an optional optimization to reduce the amount of information that must be included in level 2 LSPs.

Routers which are both level 1 and level 2 routers need to run the level 1 Dijkstra algorithm before they can determine which routes are fed from level 1 into level 2. For example, this is necessary in order to determine what metric value is announced with each entry in level 2 LSPs. The optimization described here makes use of the "next hop" information which is calculated for each route. In particular:

For each reachability entry {IP address, mask} learned from level 1 LSPs:

1) If the entry is superceded by any manually configured summary route, then this entry is not leaked into level 2 LSPs (as discussed in section 3.2 above);

2) Else, if the next hop along the best path to the entry is a level 1 only router, then the entry **must** be leaked into the level 2 LSP;

3) Else, the entry **may** optionally be leaked into the level 2 LSP.

Reachability entries which are directly reachable from the router (i.e., which the router places in its own level 1 LSP) **must** be leaked into level 2 LSPs (unless superceded by a summary route).

Where path splitting is used, the entry **must** be leaked if **all** best paths to the destination are via level-1-only routers, and **may** be leaked if **any** best path to the destination is via a level 2 router.

There are three different possible implementations which have been proposed which conform to the requirements outlined above. Other implementations are also permitted (and will interoperate correctly), provided they conform to the above requirements. These three implementations are presented only to give general guidance as to possible implementations of this feature:

1) Implementations may **always** feed routes from level 1 to level 2 (if not superceded by summary routes). For example, this is the approach specified in RFC 1195.

2) Implementations may check to see if they have **any** neighboring router which is a level-1-only router. If they have no such neighbor, then they feed only those routes directly reachable from the router (i.e., the routes that the router itself announces in its own level 1 LSP). If they have **any** neighboring level 1 only router, then they feed all routes from level 1 to level 2.

3) Implementations may individually check each IP reachability entry to determine whether the path to the nearest level 1 router announcing that entry is via a next hop router which is level 1 only. This solution allows routers to only feed those routes which are required according to the criteria above. This solution minimizes the amount of information which is leaked into level 2 LSPs, but is also the most complex to implement.

3.4 Addressing Routers in IS-IS Packets

The IS-IS packet formats explicitly require that OSI-style addresses of routers appear in the IS-IS packets. For example, these addresses are used to determine area membership of routers. It is therefore necessary for all routers making use of the IS-IS protocol to have OSI style addresses assigned. For IP-only routers, these addresses will be used only in the operation of the IS-IS protocol, and are not used for any other purpose (such as the operation of EGP, ICMP, or other TCP/IP protocols).

For OSI-only and dual routers, assignment of NSAP addresses is straightforward, but is outside of the scope of this specification. Address assignment mechanisms are being set up by standards bodies which allow globally unique OSI NSAP addresses to be assigned. All OSI-only and dual routers may therefore make use of normal OSI addresses in the operation of the IS-IS protocol.

For IP-only routers, again NSAP addresses may be obtained for use with the IS-IS protocol via the normal assignment authorities for OSI NSAP addresses. However, in some cases, routers may have only TCP/IP addresses, and it may be undesireable to have to go through the normal mechanisms for assignment of NSAP addresses. Thus it is desireable to make use of existing IP address assignments in determination of a valid OSI NSAP address. An NSAP address format is being pursued independent of this specification which will allow representation of IP addresses in NSAPs. This will allow any system which has a valid IP address assigned to it to automatically have a valid NSAP address assigned to it.

The previous version of this specification (RFC 1195) provided an algorithmic method for NSAP addresses to be derived from IP addresses. This algorithmic method will be rendered obsolete and therefore should not be used. Until final definition of the algorithmic method for producing valid NSAP addresses from IP addresses, IP-only IS-IS routers will need to have a valid ISO NSAP address assigned to them, obtained through normal OSI-related addressing mechanisms.

Within an area all routers must have consistent area address(es). The area address is used in IS-IS Hello messages and is the basis by which routers recognize whether neighbor nodes are in or out of their area.

3.5 External Links at Level 2

External connectivity (i.e., communications with routers outside of the routing domain) is normally done only by level 2 routers. The ISO version of IS-IS allows external OSI routes to be reported as “reachable address prefixes” in level 2 LSPs. The Integrated IS-IS also allows external IP reachable addresses (i.e., IP addresses reachable via inter-domain routing) to be reported in level 2 LSPs in the “IP external reachability information” field. External OSI and external IP routes are handled independently.

In some cases, external routes may need to be advertised at level 1. This is particularly likely to occur during transition from use of one routing protocol to another routing protocol (such as transition from RIP to Integrated IS-IS), and is discussed in section 3.6 ("External Links at Level 1").

The routes announced in IP external reachability information entries include all routes to outside of the routing domain. This includes routes learned from OSPF, EGP, RIP, or any other external protocol.

External routes may make use of “internal” or “external” metrics. Internal metrics are comparable with the metrics used for internal routes. Thus in choosing between an internal route, and an external route using internal metrics, the metric values may be directly compared. In contrast, external metrics cannot be directly compared with internal metrics. Any route defined solely using internal metrics is always preferred to any route defined using external metrics. When an external route using external metrics must be used, the lowest value of the external metric is preferred regardless of the internal cost to reach the appropriate exit point.

It is useful, in the operation of external routing protocols, to provide a mechanism for border routers (i.e., routers in the same routing domain, which have the ability to route externally to other domains) to determine each other’s existence, and to exchange external information (in a form understood only by the border routers themselves). This is made possible by inclusion of “inter-domain routing protocol information” fields in level 2 LSPs. The inter-domain routing protocol information field is not included in pseudonode LSPs.

In general there may be multiple types of external inter-domain routing protocol information exchanged between border routers. The IS-IS therefore specifies that each occurance of the inter-domain routing protocol information field include a "type" field, which indicates the type of inter-domain routing protocol information enclosed. Values to be used in the type field will be specified in future versions of the "Assigned Numbers" RFC. Initial values for this field are specified in Annex A of this specification.

Information contained in the inter-domain routing protocol information field will be carried in level 2 LSPs, and will therefore need to be stored by all level 2 routers in the domain. However, only those level 2 routers which are directly involved in external routing will use this information. In designing the use of this field, it is important to carefully consider the implications that this may have on storage requirements in level 2 routers (including those level 2 routers which are not directly involved in external routing).

The protocols used to exchange routing information directly between border routers, and external routers (in other routing domains/autonomous systems) are outside of the scope of this specification.

3.6 External Links at Level 1

In most cases, external links should be attached to the routing domain only via level 2 routers. This is important because level-1-only routers do not know which destinations are reachable in the routing domain (level-1-only routers know which destinations are reachable in the area, but forward other traffic to level 2 routers). Thus if a level-1-only router has a connection to outside of the routing domain, in general for any destination which is reachable via the external link the level-1-only router does not know if there is a better route within the domain. In addition, a level-1-only router does not know which routes inside of the domain can be announced to the external router.

However, there are two situations in which external routes may need to be announced in level 1 LSPs, both of which will occur during transition from other routing protocols to IS-IS: (i) In some cases, an area may be IP-capable in its level 1 routing, even though level 2 is not IP-capable; (ii) In other cases, small islands of routers running older protocols (such as RIP) may continue to exist in a network long after the bulk of the network has switched over to running Integrated IS-IS.

Each of these cases can be handled by announcing external routes in level 1 LSPs. Such routes are announced as if they were internal routes, using the "IP Internal Reachability Information" field in the level 1 LSP. In this case, the metrics used must be comparable to the normal internal metrics used in level 1 LSPs within the area.

External links at level 1 should be used with extreme caution, since it is easy to mis-configure such links in ways that could cause black holes and/or routing loops. Announcement of external routes in level 1 LSPs should be strictly limited to the two cases mentioned above, where level 2 is not IP-capable; **or** where the external route is to a small isolated set of routers and the only path out of the small set of external routers is via the IS-IS routing domain. In other cases, it is preferable to increase the number of routers which operate as level 2 routers, rather than introducing external routes at level 1 (note that IS-IS allows routers to be both level 1 routers within one area, as well as level 2 routers, and that it is generally reasonable for a small to medium size routing domain with up to several hundred routers to have most or all routers be level 2 routers).

3.7 Mixed Operation

In most cases, Integrated IS-IS will be used in environments where all routers support both TCP/IP and OSI CLNP. However, there may be situations where level 1 is IP-capable and level 2 is not, or vice versa. This may occur because OSI-only IS-IS-based routers are used (for example, level 1 might be using Integrated IS-IS, while level 2 is using OSI IS-IS). Alternatively, this may also occur during transition from DECnet Phase 4 to DECnet Phase 5 routing (for example, level 1 might be using Integrated IS-IS to support OSI, TCP/IP, and DECnet, while level 2 is still using DECnet Phase 4 routing).

3.7.1 Level 1 is IP-Capable, Level 2 is Not IP-Capable

If level 1 is IP-capable within a particular area, and level 2 is not, then operation of the routing protocol within the area is straightforward. Each level 1 router in the area calculates routes to all other routers in the area, including correct IP routes. The one problem which occurs involves how to route IP traffic to destinations outside of the area.

In the normal operation of Integrated IS-IS, where both level 1 and level 2 are IP-capable, then traffic destined to outside of the area is routed to the nearest level 2 routers. Level 1 routers therefore will generally assume that a level 2 router represents a valid route to outside of the area. However, if level 2 is not IP-capable, then this will not work correctly (since level 2 routers will not have routes to IP destinations outside of the area). Also, in general, a level-1-only router will not know whether the level 2 backbone is IP-capable.

This problem is solved by use of IP default routes in level 1 LSPs. If any router in the area announces an IP default route in its level 1 LSP, then this takes precedence over forwarding traffic to the level 2 router. IP routers in the area which have routes to outside of the routing domain may include these external routes in their LSPs, either by including specific route entries (as discussed in section 3.6), or by including IP default routes. In general, IP default routes in level 1 LSPs may be included by any router which is manually configured to do so, regardless of whether or not the router happens to also be a level 2 router. Any level-1-only IP-capable router which receives an IP default route in an LSP must allow the IP default route to take precedence over forwarding traffic destined for outside of the area to a level 2 router.

As an example, one case of interest is where IS-IS is used for level 1 routing within an area, OSI IS-IS or DECnet Phase 4 is used for routing at level 2, and one or two routers within the area are also running some other IP routing protocol (such as RIP) in order to determine IP routes to outside of the area. In this case, traffic from within the area destined to outside of the area should go to the routers which are running RIP. They will accomplish this by announcing a default route in their level 1 LSP. This allows other level 1 routers within the area to know which routers have routes to outside. Those routers which are running RIP (with other routers outside of the IS-IS routing domain) announce, in their RIP packets, routes corresponding to those destinations which are reachable within the area. Which routes to announce in this manner can be determined by a combination of manual configuration and/or extracting routes from the level 1 LSPs (a complete description of the interactions between multiple IP routing protocols is outside of the scope of this document).

The use of IP default routes at level 1 should be done with care, and should only be used where the level 2 backbone is not IP-capable, or where there is good reason to restrict the forwarding of traffic for destinations outside of the area.

3.7.2 Level 1 is Not IP-Capable, Level 2 is IP-Capable

If level 1 is not IP-capable within a particular area, but level 2 is IP-capable, then the protocol operation is straightforward. In this case each level 2 router which happens to be in the non-IP-capable area will not receive any IP routes in level 1 LSPs. In this case, each level 2 router in the non-IP-capable area will include in its level 2 LSPs those IP addresses which are reachable directly from the router, but will not include additional IP reachable addresses from level 1 LSPs.

3.8 Type of Service Routing

The Integrated IS-IS protocol provides IP Type of Service (TOS) routing [15], through use of the Quality of Service (QOS) feature of IS-IS. This allows for routing on the basis of throughput (the default metric), delay, expense (monetary cost), or residual error probability. Note than any particular packet may be routed on the basis of any one of these four metrics. Routing on the basis of general combinations of metrics is not supported.

The support for TOS/QOS is optional. If a particular packet calls for a specific TOS, and the correct path from the source to destination is made up of routers all of which support that particular TOS, then the packet will be routed on the optimal path. However, if there is no path from the source to destination made up of routers which support that particular type of service, then the packet will be forwarded using the default metric instead. This allows for TOS service in those environments where it is needed, while still providing acceptable service in the case where an unsupported TOS is requested.

The IP TOS field is mapped onto the four available metrics as follows:

Bits 0-2 (Precedence): This field does not affect the route, but rather may affect other aspects of packet forwarding.

Bits 3-6 (TOS):

0000 (normal service)	Use default metric
1000 (minimize delay)	Use delay metric
0100 (maximize throughput)	Use default metric
0010 (maximize reliability)	Use reliability metric
0001 (minimize monetary cost)	Use cost metric
other	Use default metric

3.9 Multiple LSPs and SNPs

In some cases, IS-IS packets (specifically Link State Packets and Complete Sequence Number Packets) may be too large to fit into one packet. OSI IS-IS allows for LSPs and CSNPs to be split into multiple packets. This is independent of ISO 8473 segmentation, and is also independent of IP fragmentation. Use of independent multiple packets has the advantages (with respect to segmentation or fragmentation) that: (i) when information in the IS-IS changes, only those packets effected need to be re-issued; (ii) when a single packet is received, it can be processed without the need to receive all other packets of the same type from the same router before beginning processing.

The Integrated IS-IS makes use of the same multiple packet function, as defined in [1]. IP-specific fields in IS-IS packets may be split across multiple packets. As specified in section 5 ("Structure and Encoding of PDUs"), some of the IP-specific fields (those which may be fairly long) may be split into several occurences of the same field, thereby allowing splitting of the fields across different packets.

Multiple LSPs from the same router are distinguished by LSP number. Generally, most variable length fields may occur in an LSP with any LSP number. Some specific variable length fields may be required to occur in LSP number 0. Except where explicitly stated otherwise, when an IS-IS router issues multiple LSPs, the IP-specific fields may occur in an LSP with any LSP number.

Complete Sequence Number Packets may be split into multiple packets, with the range to which each packet applies explicitly reported in the packet. Partial Sequence Number Packets are inherently partial, and so can easily be split into multiple packets if this is necessary. Again, where applicable, IP-specific fields may occur in any SNP.

3.10 IP-Only Operation

For IP-only routers, the format for IS-IS packets remains unchanged. However, there are some "variable length" fields from the IS-IS packets that can be omitted. Specifically:

IS-IS Hello Packets:

- no change

IS-IS Link State Packets:

- the "End Systems Neighbours" entries are omitted
- the "Prefix Neighbours" entries are omitted

IS-IS Sequence Number Packets:

- no change

3.11 Encapsulation

Integrated IS-IS has the potential of allowing encapsulation (of IP packets in OSI packets, or vice versa). This would allow encapsulation to be done where necessary (to allow IP packets to traverse OSI-only routers, or vice versa). Since Integrated IS-IS allows routers to know which routers are IP-only, which are OSI-only, and which are dual, this encapsulation could be done automatically without the need for manual configuration. However, at the time that this specification was written, no standard exists for encapsulation of IP packets in OSI packets, or vice versa. Therefore, it is not currently feasible to specify the details of the encapsulation function.

A future addendum to the Integrated IS-IS may specify optional encapsulation mechanisms for partition repair, and for forwarding packets through incompatible routers (i.e., for forwarding OSI packets through IP-only routers, and forwarding IP packets through OSI-only routers). The details of encapsulation and decapsulation are for further study. Routers complying with the Integrated IS-IS are not required to implement encapsulation nor decapsulation.

3.12 Order of Preference of Routes / Dijkstra Computation

We define the term “IP reachability entry” to mean the combination of the [IP address, subnet mask]. The Dijkstra calculation must calculate routes to each distinct IP reachability entry. For the Dijkstra calculation, each IP reachability entry can be treated in much the same manner as an OSI end system. Naturally, each IP reachability entry is treated as distinct from any OSI end systems which may also be reachable in the same area or routing domain.

For any particular IP reachability entry, this is the same as another entry if and only if: (i) the subnet masks are identical; and (ii) for each bit in the subnet mask which has the value “1”, the IP address is identical. This can easily be tested by zeroing those bits in the IP address which correspond to a zero bit in the mask, and then treating the entry as a 64 bit quantity, and testing for equality between different 64 bit quantities. The actual calculation of routes to IP reachability entries is therefore no more complex than calculation of routes to OSI end systems (except for the replacement of a 48-bit test with a 64-bit test).

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

The Integrated IS-IS prefers routes within the area (via level 1 routing) whenever possible. If level 2 routes must be used, then routes within the routing domain (specifically, those routes using internal metrics) are prefered to routes outside of the routing domain (using external metrics).

The Integrated IS-IS protocol makes use of “best match” routing of IP packets. This implies that a particular destination address may match more than one entry in the forwarding database. If a particular IP packet has a destination address which matches two different IP reachability entries, then the entry who’s mask contains the most “1” bits is preferred.

IP packets whose destination is a router are routed the same way as any other IP packet, by forwarding first to the appropriate subnet, and then forwarding on that subnet to the destination host (which just happens to be a router in this case). In particular, the IP forwarding database does not contain explicit routes to the individual "IP interface addresses" listed by each router in its LSP.

However, host routes (routes with a subnet mask of all ones) may of course be included in the IP reachability entries, and will be handled in the same manner as other IP reachability entries.

In order to ensure correct interoperation of different router implementations, it is necessary to specify the order of preference of possible routes. For OSI destinations, this is outside of the scope of this report. For IP destinations, this is specified in sections 3.12.1 and 3.12.2 below. Annex B specifies a detailed Dijkstra calculation and forwarding algorithm which is compatible with the order of preference of routes specified here.

With IS-IS, if a route to a given destination is advertised, or a link between routers is advertised, then metric values associated with some or all of the specified TOS metric types may be associated with that destination or link. However, the default metric must always be available. Normally this ensures that if a route using any TOS metric is available, then a route using the default metric will also be available. The only exception to this is where the corresponding route using the default metric has a total cost (within the area, or within the level 2 backbone) greater than MaxPathMetric.

In determining the route to a particular destination for a specified TOS, only routes using either the requested TOS metric, or the default TOS metric, are considered.

3.12.1 Order of Preference of Routes In Level 1 Routing

If a given destination is reachable within an area via a route using either the requested TOS or the default TOS, then the IS-IS will always make use of a path within the area (via level 1 routing), regardless of whether an alternate path exists outside of the area (via level 2 routing). In this case, routes within the area are selected as follows:

1) Amongst routes in the area, if the specified destination address matches more than one [IP address, subnet mask] pair, then the more specific address match (the one with more “1” bits in the mask) is prefered.

2) Amongst routes in the area to equally specific address matches, routes on which the requested TOS (if any) is supported are always prefered to routes on which the requested TOS is not supported.

3) Amongst routes in the area of the same TOS to equally specific address matches, the shortest routes are prefered. For determination of the shortest path, if a route on which the specified TOS is supported is available, then the specified TOS metric is used, otherwise the default metric is used. Amongst routes of equal cost, loadsplitting may be performed as specified in [1].

For a level 1 only router (i.e., a router which does not take part in level 2 routing, or a level 2 router which is not "attached"), if a given destination is not reachable within an area, level 1 routing will always route to a level 2 router as follows:

1) Amongst routes in the area to attached level 2 routers, routes on which the requested TOS (if any) is supported are always prefered to routes on which the requested TOS is not supported.

2) Amongst routes in the area of the same TOS to attached level 2 routers, the shortest routes are prefered. For determination of the shortest path, if a route on which the specified TOS is supported is available, then the specified TOS metric is used, otherwise the default metric is used. Amongst routes of equal cost, loadsplitting may be performed as specified in [1].

NOTE: An IP default route will match any IP destination. Thus if an IP default route is announced at level 1 (for either the requested TOS or the default TOS), then for a destination that is not otherwise reachable in the area, level 1 routing will prefer a route to the router that announces an IP default route over routing the packet to the nearest level 2 router.

3.12.2 Order of Preference of Routes in Level 2 Routing

For those level 2 routers which also take part in level 1 routing, routes learned via level 1 routing, using either the requested TOS or the default TOS, are always prefered to routes learned through level 2 routing. For destinations which are not reachable via level 1 routing, or for level 2 only routers (routers which do not take part in level 1 routing), then level 2 routes are selected as follows:

1) Routes using internal metrics only are always preferred to routes using external metrics.

2) If a route using internal metrics only is available:

   a) If the specified destination address matches more than one [IP address, subnet mask] pair, then the more specific address match (i.e., the largest number of “1”s present in the subnet mask) is prefered.

   b) Amongst routes with equally specific address matches (i.e., an equal number of “1”s present in the subnet mask), routes on which the requested TOS (if any) is supported are always preferred to routes on which the requested TOS is not supported.

   c) Amongst routes of the same TOS with an equally specific address matches, the shortest path is prefered. For determination of the shortest path, if a route on which the specified TOS is supported is available, then the specified TOS metric is used, otherwise the default metric is used. Amongst routes of equal cost, loadsplitting may be performed as specified in [1].

   NOTE: Internal routes (routes to destinations announced in the “IP Internal Reachability Information” field), and external routes using internal metrics (routes to destinations announced in the “IP External Reachability Information” field, with a metric of type “internal”) are treated identically for the purpose of the order of preference of routes, and the Dijkstra calculation.

3) If a route using internal metrics only is not available, but a route using external metrics is available:

   a) If the specified destination address matches more than one [IP address, subnet mask] pair, then the more specific address match is prefered.

      NOTE: For external routes, the subnet mask will normally correspond precisely to the network number. This implies that this test will usually discover equal length matching strings. However, this test is included to allow more general handling of external addresses. There are cases (particularly during transition from one routing protocol to another) in which the network corresponding to a single IP network number is connected by routers using several different routing protocols (implying that an external route may be to part of a network, implying a subnet mask which is longer than the network number). Similarly, Classless Inter-Domain Routing (CIDR) requires grouping IP network numbers, implying a subnet mask which is shorter than the network number. In these situations this test is necessary in order to ensure correct operation.

   b) Amongst routes with equally specific matches, routes on which the requested TOS (if any) is supported are always preferred to routes on which the requested TOS is not supported. NOTE: for external routes, the route is considered to support the requested TOS only if the internal route to the appropriate border router supports the requested TOS, and the external route reported by the border router also supports the requested TOS

   c) Amongst routes of the same TOS with an equal length matching address string, the shortest path is prefered. For determination of the shortest path:

      (i) Routes with a smaller announced external metric are always prefered.

      (ii) Amongst routes with an equal external metric, routes with a shorter internal metric are prefered. Amongst routes of equal cost, loadsplitting may be performed as specified in [1].

For level 2 routers which are announcing manually configured summary addresses in their level 2 LSPs, in some cases there will exist IP addresses which match the manually configured addresses, but which do not match any addresses which are actually reachable via level 1 routing in the area. Generally, packets to such addresses are handled according to the following rules:

1) If the specified destination is reachable via level 1 routing, then according to the order of preference of routes specified above, the packet will be delivered via level 1 routing.

2) If the specified destination is not reachable via level 1 routing, but is reachable via 2 routing, and there are other level 2 routers which offer more desireable routes according to the rules specified above (for example a route with a more specific match, or a route with an equally specific match which supports the correct TOS), then level 2 routing will forward the packet according to the more desireable route.

3) If the specified destination is not reachable via level 1 routing, and the manually configured summary address advertised by this router (the router which has received the packet and is

trying to forward it) represents the most desireable route, then the destination is unreachable and the packet must be discarded.

3.13 Use of Two Instances of IS-IS in S.I.N. Mode

A single instance of IS-IS may be used to support routing for either: (i) OSI CLNP only; (ii) IP-only; or (iii) both OSI CLNP and IP. In unusual cases it may be desireable to use one instance of IS-IS to support routing for OSI CLNP, while a different independent instance of IS-IS may simultaneously be used to support routing for IP. In such cases, it is possible to have a single physical router behave as if it were two routers, one for each protocol suite. This is known as operating IS-IS in "Ships in the Night" (S.I.N.) mode, because the two instances of IS-IS will operate independently, much as two ships passing in the night operate independently. Annex D of this specification specifies how two instances of IS-IS can be operated in this manner.

It is expected that such operation will be unusual. Routers are not required to implement the ability to operate two instances of Integrated IS-IS in S.I.N. mode. However, if a router does implement the capability to do so, then it **must** do so in the manner specified in Annex D. Also, if a router implements the capability of operating two instances of IS-IS in S.I.N. mode, then it **must** also implement the capability of migrating between two instances of IS-IS (one supporting IP-only operation, and the other supporting OSI-only operation) and a single instance of Integrated IS-IS (supporting IP and OSI), again as specified in Annex D.

4 Subnetwork Dependent Functions

4.1 Link Demultiplexing

Multiprotocol routers may receive a combination of OSI packets, and IP packets. It is necessary for the multiprotocol routers to be able to clearly and unambiguously distinguish the two protocol suites.

This problem is not unique to the Integrated IS-IS routing protocol. In fact, this problem will occur in any multi-protocol environment. This problem is currently being worked on independently, and is outside of the scope of this specification.

In general, the link type is a configuration parameter. For example, whether to use PPP, HDLC, or some other point-to-point protocol over a point-to-point link would be configured. For any particular link type, a method must be defined for encapsulation of both OSI and IP packets. Definition of such methods for common link types is outside of the scope of this specification.

IP packets are encapsulated directly over the underlying link layer service, using the normal method for transmssion of IP packets over each type of link. Similarly OSI packets are encapsulated directly over the underlying link layer service, using the normal method for transmission of OSI packets over each type of link. Finally, note that IS-IS packets are encapsulated using the normal method for transmission of OSI packets over any particular link type. This implies that all

IS-IS routers, including IP-only routers, must be able to receive IS-IS packets using the normal encapsulation for OSI packets.

4.2 Multiple IP Addresses per Interface

The Integrated IS-IS allows each router to have multiple IP addresses for each physical interface, up to the maximum number which may be contained in a single "IP Interface Address" field (i.e., up to a maximum of 63 addresses per interface). For example, where there are two logical subnets on the same LAN, the interface may have two IP addresses, one corresponding to each logical subnet. Each IS-IS Hello packet contains a list of IP addresses associated with the physical interface over which the Hello is transmitted.

It is permissible to implement routers which conform to the Integrated IS-IS specification which restrict the number of IP addresses per interface. However, IP-capable routers must be able to interact correctly with other routers which assign multiple IP addresses per physical interface (up to the maximum of 63 addresses per interface).

Where appropriate (for example, in some cases on point-to-point links), some interfaces may have no IP addresses assigned. In this case, the IS-IS Hello transmitted on that interface may omit the IP Interface Address field, or may include the IP Interface Address field with zero entries.

4.3 LANs, Designated Routers, and Pseudonodes

The maintenance of designated routers and pseudonodes is specified in [1], and is not changed by this proposal. In the case that IP-only and dual routers (or OSI-only and dual routers) are mixed on the same LAN in a pure IP area (or a pure OSI area, respectively), any router on the LAN may be elected designated router.

However, there is a fundamental difference in the way that OSI and TCP/IP deal with LANs, and other broadcast subnetworks.

With OSI, the use of the ES-IS protocol (ISO 9542) allows the end systems and routers to automatically determine their connectivity, thereby allowing all end systems on the LAN to potentially route via any of the routers on the LAN.

In contract, TCP/IP explictly assigns subnet identifiers to each local area network. In some cases, a single physical LAN could have multiple subnet identifiers assigned to it. In this case, end systems (hosts) which have an address on one logical subnet are explicitly precluded from sending IP packets directly to a router whose address places it on a different logical subnet. Each router is manually configured to know which subnets it can reach on each interface. In the case that there are multiple logical subnets on the same LAN, each router can only exchange IP packets with those end systems which are on the same logical subnet. This implies that it is not sufficient for the pseudonode LSP to announce all subnets on the LAN (i.e., all [IP address, subnet mask] pairs reachable on the LAN).

It is therefore necessary for each router to announce in its LSPs those subnets which it can reach on each interface, including interfaces to broadcast subnetworks such as LANs. The pseudonode

LSP does not specify the IP addresses which are reachable on the LAN (i.e., does not contain the the IP reachability field).

As specified elsewhere (see “Requirements for IP Routers” [10]), routers may send ICMP redirects only if: (i) the IP packet is being forwarded over the same physical interface over which it arrived; and (ii) the source address of the forwarded IP packet, the IP address of this router’s interface (as indicated by the source address of the ICMP redirect), and the IP address of the router to which the packet is being redirected (again, as indicated in the ICMP redirect) are all on the same IP subnet.

4.4 Maintaining Router Adjacencies

The IS-IS determines whether an adjacency is to be established between two routers using means which are independent of the IP interface addresses of the routers. Where multiple logical subnets occur on the same physical LAN, this potentially allows adjacencies to be brought up between two routers which share physical connectivity to each other, but which don’t have a logical subnet in common. IP-capable IS-IS routers therefore must be able to forward IP packets over existing adjacencies to routers with which they share physical connectivity, even when the IP address of the adjacent interface of the neighboring router is on a different logical IP subnet.

For point-to-point links, IS-IS requires exchange of ISO 9542 ISHs, as the first step in establishing the link between routers. All IS-IS routers are therefore required to transmit and receive ISO 9542 ISH packets on point-to-point links.

The “protocols supported” field (defined in section 5 below) must be present in all IS-IS Hello packets sent by dual and IP-only routers. If this field is missing, then it is assumed that the packet was transmitted by an OSI-only router. Similarly, those 9542 ISHs sent over point-to-point links, where there is (or may be) another IS-IS router at the other end of the point-to-point link, must also contain the "protocols supported" field. Note that if this field is mistakenly sent in a 9542 ISH where there is an ordinary OSI-only End System at the other end of the link, then (in accordance to ISO 9542) the End System is required to ignore the field and interpret the ISH correctly. It is therefore safe to always include this field in ISHs sent over point-to-point links.

Dual routers must operate in a dual fashion on every link in the routing domain over which they are running IS-IS. Thus, the value of the "protocols supported" field must be identical on every link (i.e., for any one router running IS-IS, all of the Hellos and LSPs transmitted by it must contain the same "protocols supported" values).

4.5 Forwarding to Incompatible Routers

There may be times when a dual router has to forward an IP packet to an OSI-only router, or forward an OSI packet to an IP-only router. In this case the packet must be discarded. An error report may be transmitted, in accordance with the IP or ISO 8473 specification (respectively). The reason for discard specified in the error report should specify “destination host unreachable” (for IP), or “destination unreachable” (for OSI).

Similarly, due to errors, in some cases an IP-only router may have to forward an IP packet to an OSI-only router. Again, the packet must be discarded, as specified above. This may only occur if IP-only and OSI-only routers occur in the same area, which is a configuration error.

4.6 Router-IDs on Unnumbered Point to Point Links

As specified in router requirements [10], when an IP router inserts its address into an IP record route, strict source and record route, loose source and record route, or timestamp option, it must use the IP address of the logical interface on which the packet is being sent. When this rule cannot be obeyed because the output interface has no IP address (i.e., it is an unnumbered interface -- an interface which does not have an IP address) the router must instead insert its "router-ID". For routers which have one or more unnumbered interfaces, the router must chose one of its IP addresses as a router-ID, and use that one unique router-ID for all of its unnumbered interfaces.

For strict source route packets sent over unnumbered links, it is necessary for a router to know the router-ID of its neighbor router. This information is also useful for network management purposes (so that a router can list the IP interface address of each of its neighboring routers).

The router-ID is reported to the neighboring router by inclusion of the router-ID field in IS-IS Hello PDUs sent over unnumbered links. In particular, when a router transmits an IS-IS Hello PDU over an interface:

- If the interface has one or more IP addresses associated with it, then the IS-IS Hello PDU must contain the "IP Interface Address" field indicating the IP address(es) associated with that interface.

- If the interface has no IP addresses associated with it, then the IS-IS Hello PDU must contain the "router-ID" field indicating the router-ID associated with the router. In this case, the "IP Interface Address" field may be omitted, or may be included with zero entries.

As specified in section 3.1, if a router has no operational interface with an associated IP address (i.e., if every active interface of a router is attached to an unnumbered point to point link), then the router must announce a host route to itself (using its router-ID and a mask of all ones) in the "IP internal reachability information" field in its level 1 LSP, and must also announce the same host route to itself (using the same router-ID value) in its level 2 LSP unless superceded by a manually configured summary route.

5 Structure and Encoding of PDUs

This clause describes the additional packet fields for use of the ISO IS-IS Intra-Domain Routing protocol in pure IP and dual environments. Specifically, the same packet types are used as in IS-IS [1], and all fixed fields remain the same. Additional variable length fields are defined in this section.

5.1 Overview of IS-IS PDUs

The packets used in IS-IS routing protocol fall into three main classes: (i) Hello Packets; (ii) Link State Packets (LSPs); and (iii) Sequence Number Packets (SNPs).

Hello packets are used to initialize and maintain adjacencies between neighboring routers. There are three types of IS-IS Hello packets: (i) “Level 1 LAN IS to IS Hello PDUs” are used by level 1 routers on broadcast LANs. (ii) “Level 2 LAN IS to IS Hello PDUs” are used by level 2 routers on broadcast LANs. (iii) “Point-to-Point IS to IS Hello PDUs” are used on non-broadcast media, such as point-to-point links, or general topology subnetworks.

On point-to-point links, the exchange of ISO 9542 ISHs (intermediate system Hellos) is used to initialize the link, and to allow each router to know if there is a router on the other end of the link, before IS-IS Hellos are exchanged. All routers implementing IS-IS (whether IP-only, OSI-only, or dual), if they have any interfaces on point-to-point links, must therefore be able to transmit ISO 9542 ISHs on their point-to-point links.

Link State Packets (LSPs) are used to exchange link state information. There are two types of LSPs: (i) “Level 1 Link State PDUs” are transmitted by level 1 routers. (ii) “Level 2 Link State PDUs” are transmitted by level 2 routers. Note that level 2 routers will, in most cases, also be level 1 routers, and will therefore transmit both sorts of LSPs.

Sequence number PDUs are used to ensure that neighboring routers have the same notion of what is the most recent LSP from each other router. The sequence number PDUs therefore serve a similar function to acknowledgement packets, but allow more efficient operation. There are four types of sequence number packets: (i) “Level 1 Complete Sequence Numbers PDU”; (ii) “Level 2 Complete Sequence Numbers PDU”; (iii) “Level 1 Partial Sequence Numbers PDU”; and (iv) “Level 2 Partial Sequence Numbers PDU”. A partial sequence number packet lists the most recent sequence number of one or more LSPs, and operates much like an acknowlegement. A partial sequence number packet differs from an conventional acknowledgement in the sense that it may acknowlege multiple LSPs at once, and in the sense that it may act as a request for information. A complete sequence number packet contains the most recent sequence number of all LSPs in the database. A complete sequence number packet may therefore be used to ensure synchronization of the database between adjacent routers either periodically, or when a link first comes up.

5.2 Overview of IP-Specific Information for IS-IS

There are six new fields defined for the Integrated IS-IS: (i) “Protocols Supported”; (ii) “IP Interface Address”; (iii) “Router-ID”; (iv) “IP Internal Reachability Information”; (v) “IP External Reachability Information”; and (vi) “Inter-Domain Routing Protocol Information”.

The “Protocols Supported” field identifies the protocols which are supported by each router. This field must be included in all IS-IS Hello packets and all LSPs with LSP number 0 transmitted by IP-capable routers. If this field is not included in an IS-IS Hello packet or an LSP with LSP number 0, it may be assumed that the packet was transmitted by an OSI-only router. The "Protocols Supported" field must also be included in ISO 9542 ISHs send by IP-capable routers over point-to-point links to other IS-IS routers.

The “IP Interface Address” is included in all IS-IS Hello packets and LSPs transmitted by IP-only and dual routers. In the Hello packets, this field occurs once only, and contains the IP address(es) of the interface on which the Hello packet is transmitted (up to a maximum of 63 IP addresses on each interface). If an IS-IS Hello is transmitted over an interface which does not have an IP address assigned, then this field may be omitted, or may be included with zero entries. In Link State Packets, this field contains a list of one or more IP addresses corresponding to one or more interfaces of the router which originates the LSP. Each IP-capable router must include this field in its LSPs. This field may occur multiple times in an LSP, and may occur in an LSP with any LSP number.

The "Router-ID" field is included in those IS-IS Hello packets transmitted by IP-only and dual routers over interfaces which do not have an IP address assigned to the interface. If an IS-IS Hello is transmitted over an interface which does not have an IP address assigned, then this field must be present. If an IS-IS Hello is transmitted over an interface which does have an IP address assigned, then this field must be omitted.

The “IP Internal Reachability Information” field may be present in all LSPs transmitted by IP-capable routers. If present, it identifies a list of zero or more [IP address, subnet mask, metrics] reachable by the router which originates the LSP. Each entry must contain a default metric, and may contain delay, expense, and error metrics. If an IP-capable router does not directly reach any IP addresses, then it may omit this field, or may include the field with zero [IP address, subnet mask, metrics] entries. If included in level 1 LSPs, this field includes only entries directly reachable by the router which originates the LSP, via one of its interfaces. If included in level 2 LSPs, this field includes only entries reachable by the router which originates the LSP, either via one of its interfaces, or indirectly via level 1 routing. This field may occur multiple times in an LSP, and may occur in an LSP with any LSP number.

The “IP External Reachability Information” field may be present in level 2 LSPs transmitted by level 2 IP-capable routers. If present, it identifies a list of zero or more [IP address, subnet mask, metrics] entries reachable by the router which originates the level 2 LSP. Each entry must contain a default metric, and may contain delay, expense, and error metrics. Each entry may contain metrics of type “internal”, or of type “external”. If a level 2 router does not have any external routes (via neighboring routers in other routing domains), when it may omit this field, or may include the field with zero entries. This field includes only entries reachable by the router which originates the LSP, via a direct link to an external router. This field may occur multiple times in a level 2 LSP, and may occur in an LSP with any LSP number.

The “Inter-Domain Routing Protocol Information” field may be present in level 2 LSPs transmitted by level 2 IP-capable routers. This field is transmitted for the convenience of the external routing protocol, and is not used by the IS-IS. For example, this may be used to allow border routers to find each other. This field may occur multiple times in a level 2 LSP, and may occur in an LSP with any LSP number.

All IP-specific information is encoded in IS-IS packets as variable length fields. All variable length fields in IS-IS are encoded as follows:

Field	No. of Octets
CODE	1
LENGTH	1
VALUE	LENGTH

Figure 3 - Encoding of Variable Length Fields

Any codes in a received PDU that are not recognised shall be ignored and, for those packets which are forwarded (specifically Link State Packets), passed on unchanged.

In general, an IS-IS PDU may contain multiple variable length fields, some of which contain OSI-specific information (specified in [1]) and some of which contain IP-specific information (specified below). Except where explicitly stated otherwise, these variable length fields may occur in any order.

5.3 Encoding of IP-Specific Fields in IS-IS PDUs

This section specifies the detailed encoding of all IP-specific fields in IS-IS PDUs. Where a particular field may be present in more than one type of PDU, the field is repeated for each type of PDU to which it applies.

Bit and octet numbering is the same as in [1]. In particular, octets in a PDU are numbered starting from 1, in increasing order. Bits in an octet are numbered from 1 to 8, where bit 1 is the least significant bit and is pictured on the right. When consecutive octets are used to represent a number, the lower octet number has the most significant value.

5.3.1 Level 1 LAN IS to IS Hello PDU

- Additional codes for IP support are:
  - Protocols Supported – the set Network Layer Protocol Identifiers for Network Layer protocols that this Intermediate System is capable of relaying
    - x CODE – 129
    - x LENGTH – total length of the value field (one octet per protocol supported).
    - x VALUE – one octet NLPID (as assigned by ISO/TR 9577) for each supported data protocol.

Field	No. of Octets
NLPID	1
NLPID	1

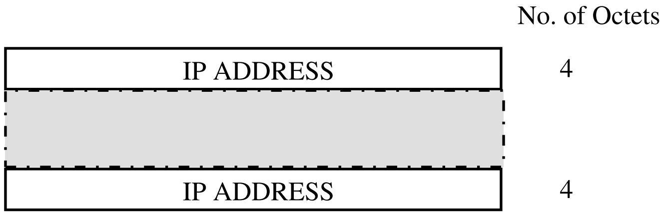
· NLPID – ISO/TR 9577 registered Network Layer Protocol Identifier.

- IP Interface Address – the IP address(es) of the interface corresponding to the SNPA over which this PDU is to be transmitted.

  x CODE – 132

  x LENGTH – total length of the value field (four octets per address).

  x VALUE –



  · IP ADDRESS – 4 octet IP Address of the Interface.

5.3.2 Level 2 LAN IS to IS Hello PDU

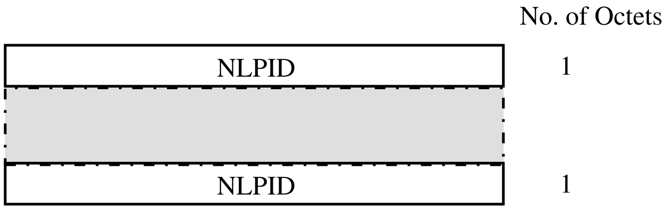
- Additional codes for IP support are:

- Protocols Supported – the set Network Layer Protocol Identifiers for Network Layer protocols that this Intermediate System is capable of relaying

  x CODE – 129

  x LENGTH – total length of the value field (one octet per protocol supported).

  x VALUE – one octet NLPID (as assigned by ISO/TR 9577) for each supported data protocol.



  · NLPID – ISO/TR 9577 registered Network Layer Protocol Identifier.

- IP Interface Address – The IP address(es) of the interface corresponding to the SNPA over which this PDU is to be transmitted.

  x CODE – 132

  x LENGTH – total length of the value field (four octets per address).

x VALUE –

	No. of Octets
IP ADDRESS	4
IP ADDRESS	4

· IP ADDRESS – 4 octet IP Address of the Interface.

5.3.3 Point-to-Point IS to IS Hello PDU

- Additional codes for IP support are:
  - Protocols Supported – the set Network Layer Protocol Identifiers for Network Layer protocols that this Intermediate System is capable of relaying

    x CODE – 129

    x LENGTH – total length of the value field (one octet per protocol supported).

    x VALUE – one octet NLPID (as assigned by ISO/TR 9577) for each supported data protocol.

	No. of Octets
NLPID	1
NLPID	1

· NLPID – ISO/TR 9577 registered Network Layer Protocol Identifier.

  - IP Interface Address – The IP address(es) of the interface corresponding to the SNPA over which this PDU is to be transmitted.

    x CODE – 132

    x LENGTH – total length of the value field (four octets per address).

    x VALUE –

	No. of Octets
IP ADDRESS	4
IP ADDRESS	4

· IP ADDRESS – 4 octet IP Address of the Interface.

- Router-ID – an IP address of the router transmitting the PDU. This field may be used only when the Hello PDU is transmitted over an interface which does not have an IP address assigned, and thus the IP interface address field is either not present, or contains zero entries.

  x CODE – 134

  x LENGTH – 4.

  x VALUE –

	No. of Octets
IP ADDRESS	4

· IP ADDRESS – 4 octet IP Address of the Interface.

5.3.4 Level 1 Link State PDU

- Additional codes for IP support are:

- Protocols Supported – the set Network Layer Protocol Identifiers for Network Layer protocols that this Intermediate System is capable of relaying.

  This must appear once in LSP number 0.

  x CODE – 129

  x LENGTH – total length of the value field (one octet per protocol supported).

  x VALUE – one octet NLPID (as assigned by ISO/TR 9577) for each supported data protocol.

	No. of Octets
NLPID	1
NLPID	1

· NLPID – ISO/TR 9577 registered Network Layer Protocol Identifier.

- IP Interface Addresses – The IP addresss of one or more interfaces corresponding to the SNPAs enabled on this Intermediate system (i.e., one or more IP addresses of this router).

  This is permitted to appear multiple times, and in an LSP with any LSP number.

  x CODE – 132

x LENGTH – total length of the value field (four octets per address).

x VALUE –

Field	No. of Octets
IP ADDRESS	4
…	
IP ADDRESS	4

· IP ADDRESS – 4 octet IP Address

- IP Internal Reachability Information – IP addresses within the routing domain reachable directly via one or more interfaces on this Intermediate system.

  This is permitted to appear multiple times, and in an LSP with any LSP number. However, this field must not appear in pseudonode LSPs.

  x CODE – 128.

  x LENGTH – a multiple of 12.

  x VALUE –

			No. of Octets
0	I/E	DEFAULT METRIC	1
S	R	DELAY METRIC	1
S	R	EXPENSE METRIC	1
S	R	ERROR METRIC	1
IP ADDRESS			4
SUBNET MASK			4
…			
0	I/E	DEFAULT METRIC	1
S	R	DELAY METRIC	1
S	R	EXPENSE METRIC	1
S	R	ERROR METRIC	1
IP ADDRESS			4
SUBNET MASK			4

· DEFAULT METRIC is the value of the default metric for the link to the listed neighbor. Bit 8 of this field is reserved, and must be set to zero on transmission and ignored on reception. Bit 7 of this field (marked I/E) indicates the metric type (internal or external) for all four TOS metrics, and must be set to zero indicating internal metrics.

- DELAY METRIC is the value of the delay metric for the link to the listed neighbor. If this IS does not support this metric it shall set the bit “S” to 1 to indicate that the metric is unsupported. Bit 7 of this field is reserved, and must be set to zero on transmission and ignored on reception.
- EXPENSE METRIC is the value of the expense metric for the link to the listed neighbor. If this IS does not support this metric it shall set the bit “S” to 1 to indicate that the metric is unsupported. Bit 7 of this field is reserved, and must be set to zero on transmission and ignored on reception.
- ERROR METRIC is the value of the error metric for the link to the listed neighbor. If this IS does not support this metric it shall set the bit “S” to 1 to indicate that the metric is unsupported. Bit 7 of this field is reserved, and must be set to zero on transmission and ignored on reception.
- IP ADDRESS is a 4-octet Internet address
- SUBNET MASK is a 4 octet IP subnet mask.

5.3.5 Level 2 Link State PDU

- Additional codes for IP support are:
  - Protocols Supported – the set Network Layer Protocol Identifiers for Network Layer protocols that this Intermediate System is capable of relaying.

    This must appear once in LSP number 0.

    x CODE – 129

    x LENGTH – total length of the value field (one octet per protocol supported).

    x VALUE – one octet NLPID (as assigned by ISO/TR 9577) for each supported data protocol.

	No. of Octets
NLPID	1
NLPID	1

- NLPID – ISO/TR 9577 registered Network Layer Protocol Identifier.

  - IP Interface Addresses – The IP addresss of one or more interfaces corresponding to the SNPAs enabled on this Intermediate system (i.e., one or more IP addresses of this router).

    This is permitted to appear multiple times, and in an LSP with any LSP number. Where a router is both a level 1 and level 2 router, it must include the same IP addresses in its level 1 and level 2 LSPs.

    x CODE – 132

x LENGTH – total length of the value field (four octets per address).

x VALUE –

	No. of Octets
IP ADDRESS	4
IP ADDRESS	4

· IP ADDRESS – 4 octet IP Address

- IP Internal Reachability Information – IP addresses within the routing domain reachable directly via one or more interfaces on this Intermediate system.

  This is permitted to appear multiple times, and in an LSP with any LSP number. However, this field must not appear in pseudonode LSPs.

  x CODE – 128.

  x LENGTH – a multiple of 12.

  x VALUE –

			No. of Octets
0	I/E	DEFAULT METRIC	1
S	R	DELAY METRIC	1
S	R	EXPENSE METRIC	1
S	R	ERROR METRIC	1
IP ADDRESS			4
SUBNET MASK			4
0	I/E	DEFAULT METRIC	1
S	R	DELAY METRIC	1
S	R	EXPENSE METRIC	1
S	R	ERROR METRIC	1
IP ADDRESS			4
SUBNET MASK			4

· DEFAULT METRIC is the value of the default metric for the link to the listed neighbor. Bit 8 of this field is reserved, and must be set to zero on transmission and ignored on reception. Bit 7 of this field indicates the metric type (internal or external) for all four TOS metrics, and must be set to zero indicating internal metrics.

- DELAY METRIC is the value of the delay metric for the link to the listed neighbor. If this IS does not support this metric it shall set the bit “S” to 1 to indicate that the metric is unsupported. Bit 7 of this field is reserved, and must be set to zero on transmission and ignored on reception.
- EXPENSE METRIC is the value of the expense metric for the link to the listed neighbor. If this IS does not support this metric it shall set the bit “S” to 1 to indicate that the metric is unsupported. Bit 7 of this field is reserved, and must be set to zero on transmission and ignored on reception.
- ERROR METRIC is the value of the error metric for the link to the listed neighbor. If this IS does not support this metric it shall set the bit “S” to 1 to indicate that the metric is unsupported. Bit 7 of this field is reserved, and must be set to zero on transmission and ignored on reception.
- IP ADDRESS is a 4-octet Internet address
- SUBNET MASK is a 4 octet IP subnet mask.

• IP External Reachability Information – IP addresses outside the routing domain reachable via interfaces on this Intermediate system.

This is permitted to appear multiple times, and in an LSP with any LSP number. However, this field must not appear in pseudonode LSPs.

x CODE – 130.

x LENGTH – a multiple of 12.

x VALUE –

			No. of Octets
0	I/E	DEFAULT METRIC	1
S	R	DELAY METRIC	1
S	R	EXPENSE METRIC	1
S	R	ERROR METRIC	1
IP ADDRESS			4
SUBNET MASK			4
0	I/E	DEFAULT METRIC	1
S	R	DELAY METRIC	1
S	R	EXPENSE METRIC	1
S	R	ERROR METRIC	1
IP ADDRESS			4
SUBNET MASK			4

· DEFAULT METRIC is the value of the default metric for the path to the listed IP addresses. Bit 8 of this field is reserved, and must be set to zero on transmission and ignored on reception. Bit 7 of this field indicates the metric type (internal or external) for all four TOS metrics, and may be set to zero indicating internal metrics, or may be set to 1 indicating external metrics.

· DELAY METRIC is the value of the delay metric for the path to the listed IP addresses. If this IS does not support this metric it shall set the bit “S” to 1 to indicate that the metric is unsupported. Bit 7 of this field is reserved, and must be set to zero on transmission and ignored on reception.

· EXPENSE METRIC is the value of the expense metric for the link to the listed IP addresses. If this IS does not support this metric it shall set the bit “S” to 1 to indicate that the metric is unsupported. Bit 7 of this field is reserved, and must be set to zero on transmission and ignored on reception.

· ERROR METRIC is the value of the error metric for the link to the listed IP addresses. If this IS does not support this metric it shall set the bit “S” to 1 to indicate that the metric is unsupported. Bit 7 of this field is reserved, and must be set to zero on transmission and ignored on reception.

· IP ADDRESS is a 4-octet Internet address

· SUBNET MASK is a 4 octet IP subnet mask

- Inter-Domain Routing Protocol Information – Inter-domain routing protocol information carried transparently through level 2 for the convenience of any Inter-Domain protocol that may be running in the boundary ISs.

  This is permitted to appear multiple times, and in an LSP with any LSP number.

  x CODE – 131.

  x LENGTH – total length of the value field

  x VALUE –

	No. of Octets
Inter-Domain Information Type	1
External Information	VARIABLE

· INTER-DOMAIN INFORMATION TYPE indicates the type of the external information which is encoded in the field.

· EXTERNAL INFORMATION contains inter-domain routing protocol information, and is passed transparently by the IS-IS protocol.

5.3.6 Level 1 Complete Sequence Numbers PDU

- No additional codes are needed for IP support.

5.3.7 Level 2 Complete Sequence Numbers PDU

- No additional codes are needed for IP support.

5.3.8 Level 1 Partial Sequence Numbers PDU

- No additional codes are needed for IP support.

5.3.9 Level 2 Partial Sequence Numbers PDU

- No additional codes are needed for IP support.

5.3.10 ISO 9542 ISH PDU

- Additional codes for IP support are:
  - Protocols Supported – the set Network Layer Protocol Identifiers for Network Layer protocols that this Intermediate System is capable of relaying.

    This appears in ISO 9542 ISH PDUs transmitted on point-to-point links.

    x CODE – 129

    x LENGTH – total length of the value field (one octet per protocol supported).

    x VALUE – one octet NLPID (as assigned by ISO/TR 9577) for each supported data protocol.

	No. of Octets
NLPID	1
NLPID	1

· NLPID – ISO/TR 9577 registered Network Layer Protocol Identifier.

6 Compatibility with Earlier Versions

This specification supercedes RFC 1195. However, at the time that this is written, routers based on RFC 1195 have been deployed, and are in use in operational networks. This section describes the changes between the current version of Integrated IS-IS and the version specified in RFC 1195, and discusses interoperation between routers based on these specifications.

6.1 Version of IS-IS

RFC 1195 makes use of the draft proposal DP 10589 [2]. This version makes use of the final International Standard version of ISO 10589 [1]. The changes between the DP and final International Standard versions of ISO 10589 that effect interoperation are discussed in this section.

Other changes in ISO 10589 (those that do not effect interoperation, or which are purely editorial in nature) are not discussed here.

6.1.1 Length of ID field in NSAP Addresses

The DP version of IS-IS requires that the ID field in the NSAP address be 6 octets long. The final version of IS-IS requires that the ID field be fixed for any given routing domain, and that for any particular domain the ID field may be from 1 to 8 octets in length.

In order for older routers (based on RFC 1195 and DP 10589) to interoperate with current routers within the same routing domain, the routing domain must use an ID length of 6 octets.

6.1.2 Number of Area Addresses

DP 10589 allowed a maximum of 3 area addresses for any area. ISO 10589 allows a larger number of area addresses to be assigned to a single area.

In order for older routers conforming to RFC 1195 and DP 10589 to interoperate with newer routers in the same area within the same routing domain, the area must have a maximum of 3 area addresses assigned to it.

6.1.3 Encoding of Variable Length Fields in Sequence Number Packets

DP 10589 made use of an encoding of the variable length fields which did not follow the normal {type, length, value} triplet format used in other IS-IS PDUs. However, RFC 1195 requires that the newer triplet format be used. The encoding specified in Annex B of RFC 1195 is compatible with the encoding used in the DIS and final International Standard versions of 10589. Therefore, any router what conforms with RFC 1195 will use an encoding of the variable length fields which is compatible with ISO 10589.

6.1.4 Network Management

The GDMO (ISO Management Information Base) has changed substantially between the DP and International Standard versions of 10589. This does not directly effect interoperation. However, the network management center will need to distinguish between routers conforming to DP 10589 and routers conforming to ISO 10589 for management purposes.

6.2 Encoding of the Authentication Field

RFC 1195 defines an authentication field, which optionally allows use of a password, and provides flexibility for possible future extension to other authentication mechanisms.

Subsequent to the publication of RFC 1195, the OSI IS-IS protocol was enhanced by addition of a functionally identical authentication mechanism. Due to the identical functionality of the two authentication fields, this has been removed from the current specification (thus routers conforming to this specification will use the authentication mechanisms specified in [1]).

Therefore, if routers conforming to RFC 1195 are intermixed with routers conforming to this specification, it should be understood that RFC-1195-conformant routers will ignore the authentication field of newer routers, and vice versa.

6.3 Router-IDs in IS-IS Hello PDUs

This specification allows router-IDs to be specified in IS-IS Hello packets. This is useful for operation of Integrated IS-IS over unnumbered point to point links.

If a router conforming to RFC 1195 is connected directly to a router conforming to this specification over a point to point link, then each router's interface to that point to point link must have an IP address assigned.

6.4 External Routes at Level 1

This specification allows external routes to be input at level 1. However, this is done in a way which does not require external metrics, and which is compatible with older routers.

6.5 Mixed Operation

This RFC specifies the details of operation when level 1 is IP-capable, and level 2 is not IP-capable, or vice versa (level 2 is IP capable, and level 1 is not IP-capable).

Operation when level 1 is IP-capable (at least within one area) and level 2 is not IP-capable requires that, if IP packets are to be routed to destinations outside of the area, then either: (i) all level 2 routers in the area must be IP-capable and have consistent understanding of external routes; or (ii) default routes be announced in level 1 LSPs by those routers capable of routing to outside of the area. However, older routers conformaing to RFC 1195 are not capable of responding properly to an IP default route in the level 1 LSP. The second solution therefore cannot be used unless all routers in the area conform to this specification, or other means are used to ensure that all routers in the area can deal correctly with default routes at level 1.

Operation when level 2 is IP-capable, but level 1 is not (at least within one area), requires that no IP traffic be routed by level-1-only routers within the non-IP-capable area. This does not involve any change from RFC 1195.

6.6 Addresses for IP-Only Routers

Integrated IS-IS allows IP-only routers to use IS-IS as their routing protocol. However, the operation of IS-IS requires valid OSI NSAP addresses. RFC 1195 provided a tentative method for algorithmically deriving a valid OSI address from an existing IP address using the OSI-ICD NSAP format.

Subsequent to the publication of RFC 1195, independent work was started in other standards bodies to allow IP addresses to be directly represented as NSAP addresses. The alternate encoding suggested in RFC 1195 is therefore obsolete and should not be used.

6.7 Optimization of Route Leaking from Level 1 to Level 2

This specification provides an optimization for reducing the amount of information required in level 2 LSPs resulting from level 1 routing information. This optimization is optional, and is compatible with earlier implementations.

6.8 Running Two Instances of IS-IS in S.I.N. Mode

This specification proposes an optional method for running two instances of IS-IS, one for routing IP only, and the other for routing OSI CLNP only. Routers conforming to RFC 1195 will not know how to operate in this manner. Thus, a router conforming to RFC 1195 may operate a single instance of IS-IS for support of either IP-only operation, OSI-only operation, or multi-protocol operation. However, a router conforming to RFC 1195 will not be able to simultaneously run multiple independent instances of IS-IS.

7 Security Considerations

ISO 10589 has a provision for carrying authentication information in all IS-IS packets. This is extensible to multiple authentication mechanisms. However, currently the only defined mechanism is a simple password, transmitted in the clear without encryption. The use of a simple password does not provide useful protection against intentional misbehavior. Rather, this should be thought of as a weak protection against accidental errors such as accidental mis-configuration. Definition of other authentication mechanisms is beyond the scope of this document.

Other aspects of security are not discussed in this document.

8 Author's Address

Ross Callon
Digital Equipment Corporation
550 King Street, LKG 1-2/A19
Littleton, MA 01460-1289
508-486-5009

9 References

[1] "Intermediate System to Intermediate System Intra-Domain Routeing Exchange Protocol for use in Conjunction with the Protocol for Providing the Connectionless-Mode Network Service (ISO 8473)", International Standard ISO 10589, 1992.

[2] “Intermediate System to Intermediate System Intra-Domain Routeing Exchange Protocol for use in Conjunction with the Protocol for Providing the Connectionless-mode Network Service (ISO 8473)”, Draft Proposal ISO DP 10589, February 1990.

[3] Callon,R., "Use of OSI IS-IS for Routing in TCP/IP and Dual Environments, RFC 1195, December 1990.

[4] “Protocol for Providing the Connectionless-Mode Network Service”, ISO 8473, March 1987.

[5] “End System to Intermediate System Routeing Exchange Protocol for Use in Conjunction with the Protocol for Providing the Connectionless-Mode Network Service (ISO 8473)”, ISO 9542, March 1988.

[6] "Protocol for Exchange of Inter-domain Routeing Information among Intermediate Systems to Support Forwarding of ISO 8473 PDUs", ISO/CD 10747, 1991.

[7] "Network Service Definition, Addendum 2: Network Layer Addressing", ISO 8348/Addendum 2, 1988.

[8] Colella,R., Gardner,E., Callon,R., "Guidelines for OSI NSAP Allocation in the Internet", RFC 1237, July 1991.

[9] DECnet™ Phase 5 routing spec

[10] Almquist,P., “Requirements for IP Routers”, Internet Draft, October 1, 1991.

[11] Moy,J., “OSPF Version 2”, RFC 1247, July 1991.

[12] Postel,J., “Internetwork Protocol”, RFC 791, September 1981.

[13] Postel,J., “Internet Control Message Protocol”, RFC 792, September 1981.

[14] Gunner, C., “Integrated IS-IS Management Information Base”, Internet Draft, November 1991.

[15] Almquist, P., “Type of Service in the Internet Protocol Suite", Internet Draft, January 1992.

Annex A
Inter-Domain Routing Protocol Information

This annex specifies the contents and encoding of the Inter-Domain Routing Protocol Information (IDRPI) field. This annex is an integral part of the Integrated IS-IS specification. However, it is expected that this annex will be augmented or superceded by future efforts outside of the scope of the IS-IS specification. For example, at the time this specification is being written, there are efforts underway to specify the manner that Integrated IS-IS will interact with BGP. This will result in additional "type" codes being assigned and additional uses of the IDRPI field being specified.

A.1 Inter-Domain Information Type

As specified in sections 3.5 and 5.3, the IDRPI field consists of a one-octet inter-domain information type field, plus a variable external information field. This section specifies initial values for the inter-domain information type field. Other values for inter-domain information type will be assigned and maintained in future versions of the "Assigned Numbers" RFC. Uses of the IDRPI field for carrying other types of inter-domain information will be specified in future RFCs.

The following types have been assigned:

Type = 0 reserved

Type = 1 local (uses routing-domain specific format)

Type = 2 AS Number Tag

Type = 1 indicates that the inter-domain routing protocol information uses a format which is local to the routing domain.

Type = 2 indicates that the inter-domain routing protocol information includes autonomous system information used to tag IP external reachability information. In this case the inter-domain routing protocol information entry must include a single AS number, which is used to tag all subsequent External IP Reachability entries until the end of the LSP, or until the next occurence of the Inter-Domain Routing Protocol Information field.

A.2 Encoding

As specified in section 5.3.5, the IDPRI entry is encoded as a variable length field, as follows:

x CODE — 131

x LENGTH — total length of the value field x VALUE –

Field	No. of Octets
Inter-Domain Information Type	1
External Information	VARIABLE

- INTER-DOMAIN INFORMATION TYPE indicates the type of the external information which is encoded in the field.
- EXTERNAL INFORMATION contains inter-domain routing protocol information, and is passed transparently by the IS-IS protocol.

The Inter-domain information type field indicates the type of information which is contained in the external information field, as follows:

Type = 0 is reserved (must not be sent, and must be ignored on receipt).

Type = 1 indicates that the external information field contains information which follows a locally specified format.

Type = 2 indicates that the external information field contains an autonomous system number tag, to be applied to subsequent IP external reachability information entries. In this case, this "inter-domain routing protocol information" entry must contain precisely one 2 octet AS number. The AS tag is associated with subsequent IP External Reachability entries, until the end of the LSP, or until the next occurence of the Inter-Domain Routing Protocol Information field. In this case, the VALUE contains the following:

x VALUE –

Field	No. of Octets
Inter-Domain Information Type = 2	1
Autonomous System Number	2

Annex B
Dijkstra Calculation and Forwarding

Annex C.2 of ISO 10589 [1] specifies the SPF (Dikskstra) algorithm for calculating routes with the IS-IS routing protocol. This annex specifies modifications to the SPF algorithm for supporting IP and dual routing, and specifies a compatible method for forwarding IP packets. This will result in an order of preference of routes which is compatible with that specified in section 3.12.

This annex is included for informational purposes.

B.1 SPF Algorithm for IP and Dual Use

This section specifies an SPF Algorithm for calculating routes with the IS-IS routing protocol, for support of both TCP/IP and OSI. This is based on an extention to the algorithm specified in annex C.2 of ISO 10589 [1].

An algorithm invented by Dijkstra known as shortest path first (SPF) is used as the basis for the route calculation. It has a computational complexity of the square of the number of nodes, which can be decreased to the number of links in the domain times the log of the number of nodes for sparse networks (networks which are not highly connected).

A number of additional optimizations are possible:

1) If the routing metric is defined over a small finite field (as in this standard), the factor of *log n* may be removed by using data structures which maintain a separate list of systems for each value of the metric rather than sorting the systems by logical distance.

2) Updates can be performed incrementally without requiring a complete recalculation. However, a full update must be done periodically to ensure recovery from data corruption, and studies suggest that with a very small number of link changes (perhaps 2) the expected computation complexity of the incremental update exceeds the complete recalculation. Thus, this annex specifies the algorithm only for the full update.

3) If only End System LSP information has changed, it is not necessary to re-compute the entire Dijkstra tree. If the proper data structures are used, End Systems (including IP reachability entries) may be attached and detached as leaves of the tree and their forwarding information base entries altered as appropriate.

The original SPF algorithm does not support load splitting over multiple paths. The algorithm in this annex does permit load splitting by identifying a set of equal cost paths to each destination rather than a single least cost path.

B.1.1 Databases

PATHS — This represents an acyclic directed graph of shortest paths from the system *S* performing the calculation. It is stored as a set of triples of the form *<N, d(N), {Adj(N)}>*, where:

N is a system identifier. In the level 1 algorithm, *N* is a 6 octet ID for OSI end systems, a 7 octet ID for routers, or an 8 octet IP Internal Reachability Information entry. For a router which is not a pseudonode, it is the 6 octet system ID, with a 0 appended octet. For a pseudonode it is a true 7 octet quantity, comprised of the 6 octet Designated Intermediate System ID and the extra octet assigned by the Destinated Router. The IP Internal Reachability Information entries consist of a 4 octet IP address plus a 4 octet subnet mask, and will always be a leaf, i.e., "End System", in PATHS.

In the level 2 algorithm, *N* is either a 7 octet router or pseudonode ID (as in the level 1 algorithm); a variable length OSI address prefix; an 8 octet IP Internal Reachability Information Entry, or an 8 octet IP External Reachability Information entry. The variable length OSI address prefixes, and 8 octet IP Reachability Information entries will always be a leaf, i.e., "End System" in PATHS. As above, the IP Reachability Information entries consist of an [IP address, subnet mask] combination.

d(N) is *N*'s distance from *S* (i.e., the total metric value from *N* to *S*).

{Adj(N)} is a set of valid adjacencies that *S* may use for forwarding to *N*.

When a system is placed on PATHS, the path(s) designated by its position in the graph is guaranteed to be a shortest path.

TENT — This is a list of triples of the form *<N, d(N), {Adj(N)}>*, where *N*, *d(N)*, and *{Adj(N)}* are as defined above for PATHS.

TENT can intuitively be thought of as a tentative placement of a system in PATHS. In other words, the triple *<N, x, {A}>* in TENT means that if *N* were placed in PATHS, *d(N)* would be *x*, but *N* cannot be placed on PATHS until is is guaranteed that no path shorter than *x* exists.

Similarly, the triple *<N, x, {A,B}>* in TENT means that if *N* were placed in PATHS, then *d(N)* would be *x* via either adjacency *A* or *B*.

Note: It is suggested that the implementation maintain the database TENT as a set of list of triples of the form *<*,Dist,*>*, sorted by distance *Dist*. In addition, it is necessary to be able to process those systems which are pseudonodes before any non-pseudonodes at the same distance *Dist*.

The 8 octet system identifiers which specify IP reachability entries must always be distinguishable from other system identifiers. As specified in section 3.12, two IP reachability entries which differ only in the subnet mask are still considered to be separate, and will therefore have distinct system identifiers *N*. The SPF algorithm will therefore calculate routes to each such entry, and the correct entry will be selected in the forwarding process.

B.1.2 Use of Metrics in the SPF Algorithm

Internal metrics are not comparable to external metrics. For external routes (routes to destinations outside of the routing domain), the cost *d(N)* of the path from *N* to *S* may include both internal and external metrics. *d(N)* may therefore be maintained as a two-dimensioned vector quantity (specifying internal and external metric values).

d(N) is initialized to [internal metric = 0, external metric = 0].

In incrementing *d(N)* by 1, if the internal metric value is less than the maximum value MaxPath-Metric, then the internal metric value is incremented by one and the external metric value left unchanged; if the internal metric value is equal to the maximum value MaxPathMetric, then the internal metric value is set to 0 and the external metric value is incremented by 1. Note that this can be implemented in a straightforward manner by maintaining the external metric as the high order bits of the distance.

In the code of the algorithm below, the current path length is held in the variable "tentlength". This variable is a two-dimensional quantity tentlength=[internal metric, external metric], and is used for comparing the current path length with *d(N)* as described above. Tentlength is incremented in the same manner as *d(N)*.

B.1.3 Overview of the Algorithm

The basic algorithm, which builds PATHS from scratch, starts out by putting the system doing the computation on PATHS (no shorter path to SELF can possibly exist). TENT is then pre-loaded from the local adjacency database.

Note that a system is not placed on PATHS unless no shorter path to that system exists. When a system *N* is placed on PATHS, the path to each neighbor *M* of *N*, through *N*, is examined, as the path to *N* plus the link from *N* to *M*. If $<M,*,*>$ is in PATHS, this new path will be longer, and thus ignored.

If $<M,*,*>$ is in TENT, and the new path is shorter, the old entry is removed from TENT and the new path is placed in TENT. If the new path is the same length as the one in TENT, then the set of potential adjacencies *{Adj(M)}* is set to the union of the old set (in TENT) and the new set *{Adj(N)}*. If *M* is not in TENT, then the path is added to TENT.

Next the algorithm finds the triple $<N,x,\{Adj(N)\}>$ in TENT, with minimal *x*. Note: This is done efficiently because of the optimization described above. When the list of triples for distance *Dist* is exhausted, the algorithm then increments *Dist* until it finds a list with a triple of the form $<*,Dist,*>$.

N is placed in PATHS. We know that no path to *N* can be shorter than *x* at this point because all paths through systems already in PATHS have already been considered, and paths through systems in TENT still have to be greater than *x* because *x* is minimal in TENT.

When TENT is empty, PATHS is complete.

Note that external metrics can only occur in "IP External Reachability Information" entries, which correspond to a leaf (i.e., End System in PATHS). Any route utilizing an entry with an external metric will always be considered to be less desireable than any entry which uses an internal metric. This implies that in the addition of systems to PATHS, all systems reachable via internal routes are always added before any system reachable via external routes.

B.1.4 The Algorithm

The Decision Process Algorithm must be run once for each supported routing metric (i.e., for each supported Type of Service). A level 1 router runs the algorithm using the level 1 LSP database to compute level 1 paths (for those level 1 routers which are not level 2 routers, this includes the path to the nearest attached level 2 router). Level 2 routers also separately run the algorithm using the level 2 LSP database to compute level 2 paths. IP-capable level 2 routers must keep level 2 internal IP routes separate from level 2 external IP routes.

Note that this implies that routers which are both level 1 and level 2 routers, and which support all four routing metrics, must run the SPF algorithm 8 times (assuming partition repair is not implemented).

If this system is a Level 2 Router which supports the partition repair optional function the Decision Process algorithm for computing Level 1 paths must be run twice for the default metric. This first execution is done to determine which of the area's manualAreaAddresses are reachable in this partition, and to elect a Partition Designated Level 2 Router for the partition. The partition Designated Level 2 Router will determine if the area is partitioned and will create virtual Level 1 links to the other Partition Designated Level 2 Routers in the area in order to repair the Level 1 partition. This is further described in section 7.2.10 of [1].

The SPF algorithm specified here will calculate routes for both OSI and IP. In particular, routes are calculated to all system identifiers *N*, where *N* may specify an OSI End System, the OSI address of a router, or an IP reachability entry. In computing the forwarding database, it is an implementation specific issue whether the IP forwarding database is kept separately from the OSI forwarding database. Where appropriate, this annex will refer separately to entries in these two forwarding databases. This is not meant to preclude any specific implementation method.

OSI and IP use separate mechanisms to determine whether a packet is in the area (in particular, OSI makes use of area addresses, and IP determines that a destination is not in an area by looking in the level 1 forwarding database and determining that no entry exists for that destination within the area). The route to the nearest level 2 router will result in separate entries in the forwarding database for OSI and IP. For IP, the route to the nearest attached level 2 router may be entered in the forwarding database as a default route (i.e., a route with a subnet mask of all 0).

One approach would be to put the results of each Dijkstra algorithm in a separate forwarding database. For a router which supports both level 1 and level 2 routing (including level 2 internal and level 2 external routes), and which supports all four types of service, this would result in twelve separate forwarding databases for IP. Implementations may choose to minimize the number of forwarding databases by combining the information from the multiple Dijkstra calculations into a single database per supported TOS. This is discussed in section B.2 below.

The SPF algorithm specified in annex C.2.3 of [1] is amended to appear as follows:

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

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

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

2) Now pre-load TENT with the local adjacency database (Each entry made to TENT must be marked as being either an End System or a router to enable the check at the end of Step 2 to be made correctly — Note that each local IP reachability entry is included as an adjacency, and is marked as being an End System). For each adjacency *Adj(N)* (including level 1 OSI Manual Adjacencies, or level 2 OSI enabled reachable addresses, and IP reachability entries) on enabled circuits, to system *N* of SELF in state "Up" compute:

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

   Adj(N) = the adjacency number of the adjacency to *N*

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

   If $x = d(N)$, then $\{Adj(M)\} \leftarrow \{Adj(M)\} \cup \{Adj(N)\}$.

4) If *N* is a router or an OSI End System entry, and there are now more adjacencies in *{Adj(M)}* than maximumPathSplits, then remove excess adjacencies as described in Clause 7.2.7 of [1]. If *N* is an IP Reachability Entry, then excess adjacencies may be removed as desired. This will not effect the correctness of routing, but may eliminate the determinism for IP routes (i.e., IP packets still follow optimal routes within an area, but where multiple equally good routes exist, will not necessarily follow precisely the route that any one particular router would have anticipated).

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

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

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

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

9) For all broadcast circuits in state "On", find the pseudonode LSP for that circuit (specifically, the LSP with number zero and with the first 7 octets of LSPID equal to L_nCircuitID for that circuit, where *n* is 1 (for level 1 routing) or 2 (level 2 routing)). If it is present, for all the

neighbors *N* reported in all the LSPs of this pseudonode which do not exist in TENT add an entry *<N, d(N), {Adj(N)}>* to TENT, where:

d(N) = metric_k of the circuit.

Adj(N) = the adjacency number of the adjacency to the DR.

10)Go to Step 2.

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

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

$$dist(P,N) = d(P) + metric_k(P,N)$$

for each neighbor *N* (both End System and router) of the system *P*. If the "Infinite Hippity Cost" bit is set, only consider the End System neighbors of the system *P*. Note that the End Systems neighbors of the system *P* includes IP reachable address entries included in the LSPs from system *P*. Here, *d(P)* is the second element of the triple

<P, d(P), {Adj(P)}>

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

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

3) If *<N,d(N),{Adj(N)}>* is in PATHS, then do nothing.

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

4) If a triple *<N,x,{Adj(N)}>* is in TENT, then:

a) If $x = dist(P,N)$, then $\{Adj(N)\} \leftarrow \{Adj(N)\} \cup \{Adj(P)\}$.

b) If *N* is a router or an OSI end system, and there are now more adjacencies in *{Adj(N)}* than maximumPathSplits, then remove excess adjacencies, as described in clause 7.2.7 of [1]. For IP Reachability Entries, excess adjacencies may be removed as desired. This will not effect the correctness of routing, but may eliminate the determinism for IP routes (i.e., IP packets will still follow optimal routes within an area, but where multiple equally good routes exist, will not necessarily follow precisely the route that any one particular router would have anticipated).

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

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

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

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

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

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

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

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

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

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

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

B.2 Forwarding of IP packets

The SPF algorithm specified in section B.1 may be used to calculate (logically) separate IP forwarding tables for each type of service, and for level 1, level 2 internal, and level 2 external routes. Section B.2.1 describes how to forward IP packets, based on these multiple forwarding databases. Section B.2.2 describes how the multiple forwarding databases can be combined into a single forwarding database per supported TOS.

B.2.1 Basic Method for Forwarding IP packets

For level 1-only routers:

- Determine if the IP destination address matches any entry in the level 1 forwarding table for the specified TOS.

- Determine if the IP destination address matches any entry in the level 1 forwarding table for the default TOS.

- If default TOS resulted in more specific entry, forward according to default TOS.

- If equally specific entries found, or specified TOS resulted in more specific entry, forward according to specified TOS

- If no entry was found (which includes no default route entry), then destination is unreachable.

  Note: For level 1 only routers, if no default route has been announced at level 1 (for either the default metric, or the specific TOS metric), then the route to the nearest attached level 2 router will be entered into the forwarding database as a default route (i.e., a route with a subnet mask which is all 0). Thus this last event (no entry found) can occur only if there is no attached level 2 router reachable in the area.

For routers which are both level 1 and level 2 routers:

- Determine if the IP destination address matches any entry in the level 1 forwarding table for the specified TOS.

- Determine if the IP destination address matches any entry in the level 1 forwarding table for the default TOS.

- If default TOS resulted in more specific entry (i.e., more bits in the subnet mask take the value 1), forward according to default TOS.

- If equally specific entries found, or specified TOS resulted in more specific entry, forward according to specified TOS

- If no entry found:

- Determine if the IP destination address matches any entry in the level 2 internal forwarding table for the specified TOS.

- Determine if the IP destination address matches any entry in the level 2 internal forwarding table for the default TOS.

- If default TOS resulted in more specific entry, forward according to default TOS.

- If equally specific entries found, or specified TOS resulted in more specific entry, forward according to specified TOS

- If no entry found:

- Determine if the IP destination address matches any entry in the level 2 external forwarding table for the specified TOS.

- Determine if the IP destination address matches any entry in the level 2 external forwarding table for the default TOS.

- If default TOS resulted in more specific entry, forward according to default TOS.

- If equally specific entries found, or specified TOS resulted in more specific entry, forward according to specified TOS

- If no entry is found, then destination is unreachable

For level 2-only routers, the above algorithm can be used, except since there is no level 1 forwarding database, the corresponding steps can be skipped.

As discussed in section 3.12.2, for level 2 routers which are announcing manually configured summary addresses in their level 2 LSPs, in some cases there will exist IP addresses which match the manually configured addresses, but which do not match any addresses which are reachable via level 1 routing in the area. Packets to such addresses are handled according to the rules specified in section 3.12.2. This may be accomplished by adding the manually configured [IP address, subnet mask] entry to the level 2 forwarding database (for the appropriate TOS), with a special "next hop" address which specifies that packets for which this entry is selected are to be discarded. This will work correctly because more desireable entries (such as delivering the packet via level 1 routing to the correct destination, or a more specific level 2 route) will automatically take precedence according to the forwarding rules specified above. Less desireable routes (such as using a level 2 external route to the "default route" entry) are not possible because other level 2 routers will believe the summary addresses advertised by this router.

B.2.2 Reduction of IP Forwarding Databases

The multiple forwarding databases used in the basic forwarding method in section B.2.1 can be reduced, by combining the multiple databases into one database for each supported TOS.

For reduction of IP forwarding databases, it is assumed that for any two overlapping address entries, either the entries are identical, or one range contains the other. In other words, for any two [IP address, subnet mask] entries *A* and *B*, if there is at least one IP address which matches both entries, then either: (i) the two entries are identical; or (ii) entry *A* contains entry *B* (i.e., any IP address which matches entry *B* also matches entry *A*); or (iii) entry *B* contains entry *A* (any IP address which matches entry *A* also matches entry *B*).

Non-contiguous subnet masks can be configured to violate this assumption. For example, consider the two entries:

- *A*=[address="01010101 00000101 00000000 00000000", mask="11111111 00001111 00000000 00000000"]

- *B*=[address="01010101 01010000 00000000 00000000", mask="11111111 11110000 00000000 00000000"]

In this case neither entry contains the other. Specifically;

- there are IP addresses which match both *A* and *B* (e.g., "01010101 01010101 xxxxxxxx xxxxxxxx"),

- there are IP addresses which match *A* but not *B* (e.g., "01010101 11110101 xxxxxxxx xxxxxxxx")

- there are IP addresses which match *B* but not *A* (e.g., "01010101 01011111 xxxxxxxx xxxxxxxx").

The reduction of the multiple forwarding databases for each TOS to a single database for each TOS is based on the use of "best match" routing, combined with reduction of the entries placed in the forwarding database in order to eliminate entries which are not to be selected (based on the order of preference of routes specified in section 3.12). The specific algorithm for creation of the IP forwarding database can be described as follows:

1) Make use of the the Dijkstra algorithm described in section B.1 to create separate forwarding databases for each supported TOS for level 1 routes, level 2 internal routes, and level 2 external routes. (Note that each entry in the forwarding database will specify an [IP address, subnet mask] combination, as well as the next hop router for IP packets which match that entry).

2) For each level 1 route entry which has been placed in the level 1 IP forwarding database for a specific TOS, copy that entry into the overall IP forwarding database for that TOS.

3) For each route entry *X* which has been placed in the level 2 internal IP forwarding database for a specific TOS, search for overlapping entries in the level 1 IP forwarding database for the specific TOS, and also for the default TOS:

   a) If there is any overlapping entry *Y* in the level 1 forwarding database (for the specfic TOS, or for the default TOS) such that either (i) *Y* contains X; or (ii) *Y* is identically specific to *X*; then ignore entry *X*.

   b) Otherwise, copy entry *X* into the overall IP forwarding database for the specific TOS.

4) For each route entry *X* which has been placed in the level 2 external IP forwarding database for a specific TOS, search for overlapping entries in the level 1 IP forwarding database for the specific TOS, and for the default TOS, and the level 2 internal IP forwarding database for the specific TOS, and for the default TOS.

   a) If there is an overlapping entry *Y* such that either (i) *Y* contains *X*; or (ii) *Y* is identically specific to *X*; then ignore entry *X*.

   b) Otherwise, copy entry *X* into the overall IP forwarding database for the specific TOS.

This method will result in one forwarding database for each supported TOS. The forwarding of packets can then be simplified to be as follows:

1) For IP packets which map to the default TOS metric (or to an unsupported TOS metric), search the default TOS forwarding database and select the entry which has the most specific match. Forward the packet accordingly.

2) For packets which map to a specific (non-default) TOS metric, search the specific TOS forwarding database and select the entry *j* which has the most specific match. Also search the default TOS forwarding database and select the entry *k* which has the most specific match. Forward the packet as follows:

   a) If *k* is more specific than *j*, forward according to entry *k*

   b) If *j* and *k* are equally specific, forward according to entry *j*

c) If *j* is more specific than *k*, forward according to entry *j*

Annex C
Interaction of the Integrated IS-IS with Brouters

A "brouter" is a device which operates an both a bridge and a router. One possible type of brouter acts as a router for IP traffic, and acts as a bridge for all other types of traffic.

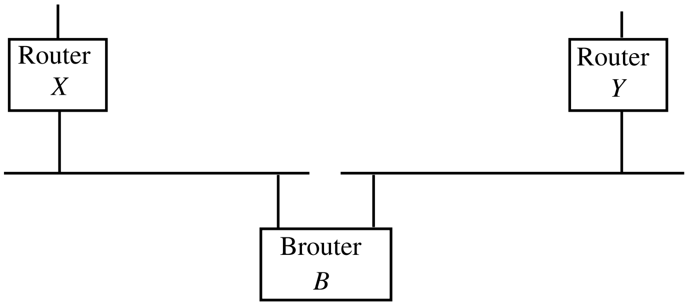
Depending upon the manner in which a brouter is implemented, and depending upon the network topology, there is an obscure bug which can result from the interaction of the Integrated IS-IS protocol, and brouters. This appendix gives an example of the bug, and proposes a simple correction to the operation of brouters to correct the problem.

This annex is included for informational purposes.

C.1 The Problem

Suppose that we have a brouter which treats IP packets as if it were a normal IP router, and which treats all other packets as if it is a bridge.

Suppose that two routers "*X*" and "*Y*" (which implement the Integrated IS-IS protocol), two Ethernets, and a brouter "*B*" are all connected as follows:



Here suppose that *X* and *Y* are running the Integrated IS-IS protocol, and are both level 1 routers in the same area. Thus *X* and *Y* send IS-IS Hello packets on the LAN. These Hello packets are received and forwarded by the brouter (using normal bridge functions). Similarly, *X* and *Y* receive each other's IS-IS LSP packets. In this way, it appears to the Brouter that *X* and *Y* are exchanging OSI packets, and so they are forwarded using normal bridge functions. It appears to *X* and *Y* as if they are on the same LAN, and so they learn each others 48-bit Ethernet addresses and exchange routing information.

Now, suppose that *X* receives an IP packet, which it needs to forward via *Y*. Since *X* thinks that it and *Y* are on the same Ethernet, it just forwards the IP packet directly, using normal Ethernet encapsulation and using the 48-bit Ethernet address of *Y* as the destination address in the Ethernet header. Brouter *B*, when thinking as a bridge says: "this is an IP packet, I don't forward this as a bridge”. Brouter *B*, when thinking like an IP router says: “this is an IP packet, I know how to forward IP packets. However, this is sent to an Ethernet address which is not me, thus I will ignore it”. The result is that the IP packet does not get forwarded.

This problem relates directly to the fact that *X* and *Y* are exchanging OSI packets to determine the connectivity of the path between them, but then are trying to send IP packets over the path. Also, there is a device between *X* and *Y* on the path which treats OSI and IP packets differently.

Also note that this problem can also occur in more complex topologies, whenever a brouter is treating OSI and IP packets in a fundamentally different manner.

C.2 Possible Solutions

C.2.1 More Sophisticated Brouter

One solution is that brouter *B* needs to be a little more sophisticated. In particular, it needs to use the following rules:

- For packets which are not IP packets, act as a bridge (this is the same as before).
- For IP packets sent to an Ethernet broadcast or multicast address, act as an IP router (this is also the same as before).
- For IP packets sent to my own Ethernet 48-bit address(es), act as an IP router (this is also the same as before).
- For IP packets sent to a single station 48-bit address which is not one of my addresses, act at a bridge (THIS IS NEW).

With this change, the IP packet transmitted from *X* to *Y* is forwarded by the brouter, acting as a bridge. This allows the Brouter and the multi-protocol routers to interoperate properly.

C.2.2 Dual Router / Brouter

An alternate solution would be for the Brouter to route both OSI and IP equally. If the Brouter used the Integrated IS-IS for this purpose, then it could be part of the same routing domain and interoperate like any other dual router (except for the ability to bridge other protocol suites). If it used other protocols for routing OSI and IP, then it would need to be part of another routing domain, and could interoperate with Integrated IS-IS routers like any other external router.

Annex D
Operation of IS-IS in "Ships in the Night" Mode

In some cases it may be desireable to use one instance of IS-IS to support routing for OSI CLNP, while a separate independent instance of IS-IS may simultaneously be used to support routing for IP. In such cases, it is possible to have a single physical router behave as it it were two routers, one for each protocol suite. Thus some routers would be running a single instance of IS-IS and only supporting CLNP. Other routers would be running a single instance of IS-IS and only suporting IP. Other routers would be running two instances of IS-IS, and supporting both CLNP and IP. This is known as operating IS-IS in "Ships in the Night" (S.I.N.) mode.

Routers are **not** required to implement the ability to operate two instances of Integrated IS-IS in S.I.N. mode. However, if a router does implement the capability to do so, then it must do so in the manner specified in section D.1. Also, if a router implements the capability of operating two instances of IS-IS in S.I.N. mode, then it must also implement the ability to migrate between S.I.N. and integrated operation, as specified in section D.2.

D.1 Operation in S.I.N. Mode

When two instances of Integrated IS-IS are being used, all IS-IS packets used for one instance are independent of all packets used for the other instance. Thus, for any one IS-IS packet, it is necessary to determine unambiguously whether the packet is related to one instance or the other instance. This separation is performed by use of the password capability provided by the authentication field from ISO IS-IS [1]. In particular, all IS-IS packets must use the password field from [1], and the set of passwords used for one instance must be different from the passwords used for the other instance.

Normally, when two independent instances of IS-IS are being used, one will support IP-only (and thus the protocols supported field of associated packets will contain the NLPID corresponding to IP) and the other will support OSI only (implying that the protocols supported field of associated packets will contain the NLPID corresponding to OSI). This would imply that in principle it would be possible to demultiplex the packets based on the protocols supported field. However, during migration from S.I.N. mode to integrated mode it will be necessary to have both instances support the same protocol suite, implying that it is not acceptable to demultiplex based on the protocols supported field.

When running in S.I.N. mode, the Dijkstra algorithm must be run independently for each instance of IS-IS.

D.2 Migration Between S.I.N. Mode and Integrated Mode

This section described how to make a smooth transition from an initial state of ships in the night operation to a final state of integrated operation. For migration in the reverse direction (from integrated to S.I.N. operation) the same transition steps can be followed in the reverse direction. The transition is constrained as follows:

- INITIAL STATE: Some routers are running IS-IS for support of CLNP. Some other routers (may be the same routers, an overlapping set, or a different set of routers) are running IS-IS for support of IP.

- FINAL STATE: One set of routers has been updated to run Integrated IS-IS for support of IP and CLNP. The other routers have been turned off. Some resources may have been reassigned in order to support the integrated operation (for example, some of the links and/or routers associated with the set of routers which have been turned off might have been added to the Integrated IS-IS set).

- ADDITIONAL CONSTRAINT: We must allow continuous correct operation with no period of network disruption. We must be able to update routers gradually on a one-at-a-time basis during the migration period.

Assuming a large network with a substantial number of routers, it may not be possible for all routers to be reconfigured in a short time frame. Rather, it may take days, weeks, or even months for all routers to be reconfigured.

The main problem in transitioning from S.I.N. to integrated routing is the need to maintain a consistent view of the network by all routers at all times (or at least all routers at the same level, such as all level 2 routers, or all level 1 routers within a single area). In particular, IS-IS (and all link state protocols) assumes that the routers at the same level all have an identical view of topology.

However, IS-IS routers which are configured to run IS-IS for support of OSI-only will ignore IP-specific information in IS-IS packets, and vice versa. Also, those routers running two instances of IS-IS in S.I.N. mode will run the Dijkstra algorithm twice, while those running integrated IS-IS will run the Dijkstra algorithm only once. Thus, if we simply start updating some routers to run IS-IS in integrated mode, while other routers are still running IS-IS in S.I.N. mode, then routers will have an inconsistent view of topology.

It is therefore necessary to select one instance of IS-IS (for example, the instance originally used for support of CLNP) and incrementally update it to support both protocol suites, while making sure that this instance is not used for support of the second protocol suite (for example, is not used for support of IP) until all routers have been updated. This is accomplished by configuring routers to announce in the appropriate IS-IS packets that they can support both protocol suites, but not actually announce that they can reach any destinations in the "to be added" protocol suite until all routers have been updated.

The description below assumes that the instance of IS-IS originally used for support of CLNP will be updated to be dual, and the instance of IS-IS originally used for support of IP will be phased out. A similar migration scenario may be done the other way, with the instance originally used for IP migrated to being dual and the instance originally used for CLNP being phased out. Routers which implement the ability to run two instances of IS-IS in S.I.N. mode must be able to do the migration either way. The migration scenario is as follows:

1) INITIAL STATE: Two instances of IS-IS are being used: One instance for support of CLNP only, one for support of IP only. Some routers route one protocol suite only (and therefore only take part in the operation of one instance of IS-IS). Some routers route both protocol suites, and therefore take part in the operation of both instances of IS-IS. All routers make use of passwords carried in the Authentication field from IS-IS, and the two instances of IS-IS use different passwords.

   Note: For ease of description the first instance of IS-IS (initially used for CLNP only) will be called "instance A". The set of routers which take part in the first instance of IS-IS (the instance initially used for support of CLNP only) will be referred to as "set A". Similarly, the second instance of IS-IS (used for IP only) will be referred to as "instance B". The set of routers which take part in the second instance of IS-IS will be referred to as "set B".

2) One at a time, for each router in set A, update the router to support both IP and CLNP with instance A of IS-IS as follows:

   a) No IP routes are announced in instance A of IS-IS (i.e., either the "IP reachability information" fields are empty, or not present).

   b) If any IP routes are found announced in LSPs received from other routers taking part in instance A of IS-IS, then these routes are used.

   c) The "protocols supported" field for instance A contains the NLPIDs for both CLNP and IP.

   d) If some of these routers are also part of set B (i.e., are also running the second instance of IS-IS for support of IP), then they continue to take part in instance B of IS-IS.

   e) For those routers which are also part of set B the routes for IP (if any) determined from instance A take precedence over those routes determined from instance B.

3) When this is complete, all routers in set A will be running Integrated IS-IS, and will be ready to (a) interpret IP-specific information in instance A IS-IS packets; and (b) correctly forward

IP packets on the basis of this information. However, at this time no IP routes are actually being announced into the IS-IS packets corresponding to instance A. This means that routing of IP packets is still using instance B, and therefore that no disruption of routing has yet occurred.

4) One at a time, tell the routers in set A to announce their own local routes into instance A of IS-IS. Similarly, tell those routers which are announcing external routes into instance B to also announce similar external routes into instance A.

5) When this is complete, IP packets are being forwarded using instance A. However, for those routers which are in instance B but not in instance A, they still are using the instance B routes. Also, such routes have routes which are being announced into instance B but not into instance B. Such routers may, one at a time, be migrated to participate in instance A.

6) When this is complete, instance B may be turned off. Associated resources may be retired or used to support instance A.

7) FINAL STATE: Instance A of IS-IS is supporting both IP and CLNP. Instance B of IS-IS is no longer in use.