AN ARCHITECTURE FOR NETWORK-LAYER ROUTING IN OSI

Paul F. Tsuchiya
The MITRE Corporation

ABSTRACT

Work on the standardization of routing protocols for OSI is in progress. The envisioned set of routing protocols is expected to work in nearly all of the environments which constitute OSI networks. Behind these routing protocols is an architecture which outlines problems and goals, establishes a framework upon which to base the development of protocols, and provides a conceptual baseline for continued work on unsolved problems. This architecture defines routing in the OSI network layer, functionally partitions the problem into its components, defines a routing hierarchy and an address hierarchy and discusses their relationship, and discusses arms-length routing relationships between differently administered networks. This paper presents that architecture, and discusses problems which remain to be solved as work progresses towards a global OSI network.

1. INTRODUCTION

In the International Organization for Standardization (ISO) Open Systems Interconnection (OSI) suite of data communications protocols, mature point-to-point communications standards exist for all seven layers of the reference model. Missing, however, are the various management standards (including routing standards) required to facilitate the operation of open systems. This paper concerns itself with the problem of routing in OSI, and the ongoing effort to define network layer routing standards for OSI.

In ISO, the primary motivation for network layer routing work has come from the American National Standards Institute (ANSI) Accredited Standards Committee (ASC) X3S3.3, Network and Transport Layer. In X3S3.3, work has been going on for several years to develop routing protocols for the network layer. This work has spawned the relatively mature DIS 9542, "End System to Intermediate System Routing Exchange Protocol for use in conjunction with the Protocol for the Provision of the Connectionless-mode Network Service (ISO 8473)" (ES-IS). A parallel effort for an ES-IS for use with ISO 8208 (ISO's X.25) is also underway. Further, solid and complete proposals for Intermediate System to Intermediate System (IS-IS) Routing within a single administrative domain have been submitted to X3S3.3 for consideration. (I apologize for this early barrage of alphabet soup. These concepts will be explained later in the paper.)

Behind these efforts is a set of architectural principles which are guiding the protocol developments. These architectural principles are continuously evolving as the problem of global routing in OSI is contemplated and solutions are put forth. They are based on X3S3.3's understanding of 1) the OSI environment, 2) the problems associated with that environment, 3) routing architectures, and 4) routing techniques.

This paper, then, describes the set of architectural principles which is guiding the development of routing protocols in ISO. The purpose of this paper is both to distribute information and to solicit comments. It represents not the final word on routing in the global OSI environment—rather, it represents the current position in an evolving process. Input to this process from as many communities as possible is requested.

2. DEFINITION OF ROUTING

When we say "routing", we refer to all of the procedures involved in building a routing table and relaying individual Protocol Data Units (PDUs). Routing consists of several component parts, which are illustrated in Figure 1.

In Figure 1 we see two kinds of flow—data flow and information flow. Data flow is flow that enters and leaves a System through PDUs. Information flow is internal to the System. (System is short for End System and/or Intermediate System.)

There are two types of data flows which invoke routing procedures: Routing PDUs and Data PDUs. Routing PDUs are those PDUs which are exchanged between systems to collect or distribute routing information such as link status, reachable Systems, and so on. Data PDUs are those PDUs, such as ISO 8473 packets, upon which a routing action must be made. We call this relaying, and it is decomposed into three functions: the Locate Function, the Route Function, and the Forward Function.

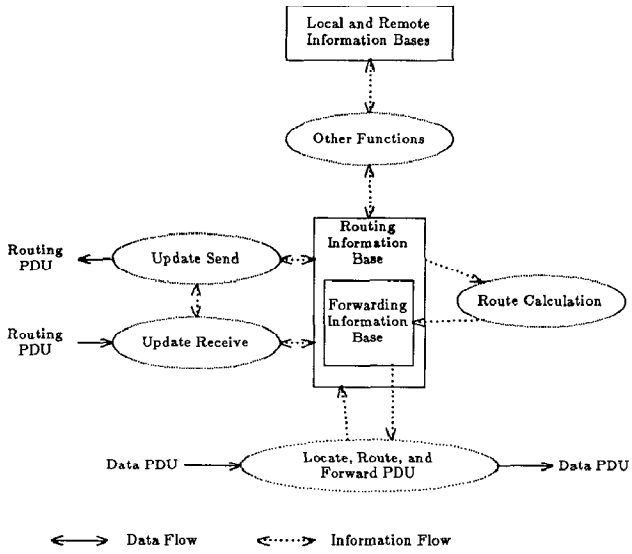

Figure 1

The Locate Function examines the destination Network Address—specifically, the Network Service Access Point (NSAP) Address in the case of ESs, and the Network Entity Title (NET) in the case of ISs—of the PDU and determines where (in the routing hierarchy) the destination ES is. This may be a simple masking operation on the address itself to extract the so called "routing information", or it may be a table or directory lookup function if the address contains no routing information. Relaying data PDUs is much simpler when the hierarchy location information is embedded in the Network Address in the data PDU. However, getting the hierarchy location information into the Network Address in the first place is not trivial.

The Route Function is what is often thought of as "routing". It is the function which returns the Network Address of the next hop when given the location of the destination ES derived from the Locate Function.

The Forwarding Function returns the Subnetwork Point of Attachment (SNPA) when given the Network Address of the next hop. An Ethernet address or an X.121 address are examples of SNPAs.

In addition, we define a data base called the Routing Information Base (RIB). This data base contains all information pertinent to routing, such as local link status, a network topology map, and so on. A subset of the RIB is the Forwarding Information Base (FIB). This data base contains the information which is directly accessed when the relaying functions are performed. It is what is often thought of as a "routing table"—that is, a list of destinations and the next hop to those destinations. The Route Calculation Function derives the FIB from information in the RIB.

Finally, we show Other Functions and the Local and Remote Information Bases in Figure 1. This is "everything else" routing procedures might require, such as directory service access, authentication information, and so on.

2.1. OSI Environment

In this following paragraphs, we describe the environment within which routing is expected to take place.

There will be a virtually unlimited number of ESs and ISs. An ES is the source or destination of data traffic. An IS relays traffic between ESs and/or other ISs. In particular, there will be many more ESs than ISs.

There will be collections of interconnected Systems which are administered by a single entity (a corporation, for instance). These collections of Systems, called Administrative Domains, will be routing traffic to and through other Administrative Domains. These Administrative Domains may not trust each other, however, and will require autonomy from each other and protection from each other's failures.

There will be very diverse means of transmission between Systems. These include direct links ranging in capacity from dial-up modems to fiber optic, Local Area Networks (LANs), satellites, packet switched data networks, and so on. These transmissions may be interconnected in any arbitrary fashion.

Systems in the OSI environment fall under a global addressing structure (ISO 8348/AD 2, "Network Service Definition - Addendum 2: Network Layer Addressing") which by definition allows addresses to be structured differently by different communities. In particular, it frees the assignment of addresses from necessarily being constrained by routing considerations.

The routing should be as automatic as possible. In particular, the routing protocols should to the extent possible automatically discover the appearance or disappearance of Systems and links, and to modify routes accordingly.

3. NETWORK LAYER ROUTING ARCHITECTURE

In this section, we develop the different but related components of the routing architecture—the routing hierarchy and its relation to the addressing hierarchy; the four functional tiers of routing; and the nature of routing and routing agreements between different Administrative Domains.

3.1. The Routing Hierarchy

As soon as one recognizes that a set of Systems (this Administrative Domain) wishes to distinguish itself from another set of Systems (that Administrative Domain) *for the purposes of routing*, one has embraced the notion of a routing hierarchy (as opposed to an addressing hierarchy, which is an entirely different thing). Any time a set of entities have been grouped such that they can be treated externally as a single entity, a hierarchy has been formed. If only one such level of grouping exists, then a two-level hierarchy results. If groups are recursively grouped, then a multiple level hierarchy results. (We avoid the temptation to provide a figure here. A drawing of the hierarchy must necessarily be simple, and invariably seems to mislead more than it enlightens.)

The address of a System may or may not correspond to the System's group memberships—that is, the address hierarchy may or may not correspond to the routing hierarchy. We say that an address corresponds to a grouping when all of the members of the group can be identified as belonging to the

group, and all non-members of the group can be identified as not belonging to the group, by merely examining the address of the member or non-member. In addition, there must be paths between all group members—that is, it must be possible to get from one group member to another group member without going through a non-member (this is true for a routing hierarchy, not for an address hierarchy).

In this architecture, we have three reasons for grouping Systems, or for grouping groups of Systems. Each grouping has an impact on how routing is done. We therefore have three different names for groups.

The first reason for grouping is simply to reduce the amount of routing information which must be spread around, thus achieving greater efficiency. This sort of grouping has been thoroughly studied in many contexts, but in particular by Kleinrock and Kamoun, and by Hagouel, in the context of data networks. We call this type of group a Cluster. Within Clusters and between Clusters, there is full trust, full agreement on routing procedures (algorithms, metrics, and so on)—in other words, no autonomy—and a high correspondence between clustering and addressing. This is not to say that there isn't a difference between intra-Cluster routing procedures and inter-Cluster routing procedures. For instance, handling cluster partitions (the situation where a message between two members of a Cluster cannot, in fact, be delivered without leaving the cluster) requires different intra-cluster and inter-cluster procedures. It is worth noting that many of the efficiencies achieved by clustering are lost if there is not a high correspondence between the clusters and the addresses.

The second reason for grouping Systems is because one group of Systems has different routing procedures than another group of Systems—in other words, the two groups are largely autonomous. We call this group a Routing Domain. We assume full trust between Routing Domains, and a maximum of autonomy. We say maximum (as opposed to full) here because full autonomy is not achievable in practice. For instance, if the efficiency benefits of consistent addressing are desired, then the Routing Domains do not have complete autonomy for choosing addresses. A better example, however, is that the two Routing Domains must use common routing procedures to talk to each other. These *exterior* routing procedures may have an impact on how the *interior* routing procedures must behave, thus limiting autonomy.

For instance, assume three Routing Domains A, B, and C, with A and B connected, and B and C connected, but with A and C not connected. Assume further that the border ISs (those which share a link with ISs in other Routing Domains) participate in the exterior routing procedures, and that all ISs participate in their respective interior routing procedures. Now, for traffic to pass from A to C, traffic must be forwarded from B's border IS with A to B's border IS with C. For this to happen, either 1) the interior ISs of B must have an awareness of the world outside of Routing Domain B, 2) or the border ISs of B must have some way of getting traffic to each other without requiring that the interior ISs have a notion of the world outside—the border ISs must either be directly connected, or they must tunnel through the interior ISs, either by enveloping the original packet, or by partial source routing (a feature of ISO 8473). In any event, the internal operation of Routing Domain B is effected by B's relationship with other Routing Domains, thus limiting true autonomy.

The third reason for grouping Systems is administrative commonality among those Systems—or, more to the point, a lack of administrative commonality between different groups of Systems. We call these groups Administrative Domains. In particular, there is a maximum of autonomy between Administrative Domains, there is a lack of trust, and there are very specific notions about where traffic may and may not go. Because of this, there is the notion of contractual agreements between Administrative Domains which determine whether the Administrative Domains will accept traffic from each other, and determine for which (if any) other Administrative Domains they will forward traffic. In particular, only routing information which is valid according to a set of agreements should be exchanged. This allows an Administrative Domain to protect itself from learning about routes which it does not wish to use (for security or legal reasons, for instance). It also limits the amount of potentially incorrect routing information an Administrative Domain can receive by limiting the set of other Administrative Domains which can pass third party information. However, it does so at the expense of dynamic and automatic configuration.

3.1.1. Structure of the Routing Hierarchy

In general, Clusters can consist of Systems and/or lower level Clusters. Routing Domains can consist of Systems and/or Clusters. Administrative Domains can consist of Routing Domains and lower level Administrative Domains. Beyond this, the architecture does not specify how groups may be arranged. For instance, they may overlap in arbitrary ways. The complexity of the arrangement of the hierarchies will be determined by the routing protocols, not by the architecture.

There will not be one global monolithic routing hierarchy. Instead, there will many separate routing hierarchies. Each of these separate routing hierarchies can be viewed as an Administrative Domain, and the global routing structure will consist of a flat (non-hierarchical) super-network of Administrative Domains, tightly or loosely coupled according to agreements made between Administrative Domains.

3.2. Addressing Hierarchy

The NSAP addressing hierarchy, on the other hand, is a global monolithic structure in that it is defined by ISO (ISO 8348/AD 2, "Network Service Definition - Addendum 2: Network Layer Addressing"), which specifies portions of the NSAP address space. The maximum NSAP address is 20 octets long. ISO 8348/AD 2 is designed to facilitate the assignment of globally unique NSAP addresses while accommodating (or more appropriately, subsuming) all existing standardized ISO and CCITT addresses. The technique for guaranteeing globally unique addresses is to recursively assign portions of the address to lower addressing authorities who are instructed to further parse the address as necessary. For instance, ISO has assigned a certain range of addresses to ANSI (specifically, all addresses where the first octet is hex 38 or 39, and the next octet and a half are hex 840). ANSI, as an address authority, then will break that space up into smaller chunks to

be handed to different organizations. These organizations may take their space and further divide it among sub-organizations, and so on. This recursive assignment essentially defines the address hierarchy.

After some number of recursions (hopefully not too many), some address authority will receive an address space which it hopes to parse in accordance with a routing hierarchy. It is not necessarily the case that at this point the remaining portion of the address must be parsed the same way by all Systems within this routing hierarchy. For instance, the address/routing authority may further assign the address by giving ranges to each of several regional networks, which may operate independently but communicate with each other extensively. Each of the regional networks may be quite different in composition, and may wish to parse their address spaces differently. This will work, however, because it is not necessary for Systems in Region A to understand how Systems in Region B parse their addresses. It is only necessary for Systems in Region A to recognize that an address is, in fact, in Region B, and to route it to a Region B System. The Region B System will then further parse the address for routing within its region.

It is not even necessary for the address/routing authority to assign the same size ranges to the regions. As long as one region (group) has a method to tell the other region which range (or ranges) of addresses are in its region, routing can take place. A typical method for doing this is to use a mask to indicate where the significant bits of address are, and an address to indicate the value of those bits. This technique is used both in the Internet Control Message Protocol (ICMP) Subnet Mask, and in ISO 9542.

3.3. Four Functional Tiers of Routing

Having defined a hierarchy for routing, we must consider the development of one or more routing protocols for use in that hierarchy. One option might be to define one standard which accomplished routing at all hierarchical levels. This, however, is not an appropriate approach for several reasons. First, we know how to do routing at only some levels of the hierarchy. If we try to delay any routing standards until we know how to do all routing, the development of useful routing protocols which can be of use in a limited environment will be unnecessarily delayed. Second, we can partition routing into several component parts because of functional differences between those parts. By partitioning the problem, we can work on individual parts without necessarily sacrificing the whole.

The first component part is routing between ESs and ISs, called ES-IS routing. The second is routing between ISs which fall under a single routing authority, share a common set of routing procedures, and which fully trust each other. This is called Intra-Domain IS-IS routing. The third is routing between ISs which fall under a single routing authority and which fully trust each other, but which have different routing procedures. This is called Inter-Domain IS-IS routing. The fourth is routing between ISs which fall under different routing authorities and which don't necessarily trust each other. This is called Inter-Administration IS-IS routing.

The creation of the first component, ES-IS routing, results from the observation that routing functions involving Systems which cannot relay traffic (ESs) are fundamentally different than those involving Systems which do relay traffic (ISs). In particular, since ESs cannot relay traffic, they have no reason to propagate routing updates *about other Systems*. This greatly simplifies the tasks that an ES must perform. An ES need only discover one or more ISs which will then perform relaying for the ES. The ISs can tell the ES which IS is the best choice for a given set of destinations. An ES, however, does not need to involve itself in a complex n-party routing algorithm. A side benefit of separating ES-IS routing from any IS-IS routing is that ESs can be as simple as possible—saving the complexity, and therefore expense, for the less numerous ISs.

In the second component, Intra-Domain IS-IS routing, a great deal of coordination between ISs is possible. Within the bounds of correct operation, an IS can accept information another IS gives it without question. Therefore, any discovered routes can be fully advertised (within the boundaries of the Domain). ISs can agree on routing metrics (such as delay, bandwidth, hops, etc.), a necessary requirement for the discovery and prevention of loops. The assignment of addresses can be controlled to reflect the hierarchical clustering which may exist within a Routing Domain.

The first two components are clear-cut and well understood. This is supported by an ES-IS protocol in ISO (DIS 9542), and two proposed Intra-Domain routing protocols in X3S3.3.

The last two components, Inter-Domain and Inter-Administration routing, are less clear-cut. It is not obvious what should be standardized with respect to these two components of routing. For example, for Inter-Domain routing, what can be expected from the Domains? By asking Domains to provide some kind of external behavior, we limit their autonomy. If we expect nothing of their external behavior, then routing functionality will be minimal.

Across administrations, it is not known how much trust there will be. In fact, the definition of trust itself can only be determined by the two or more administrations involved.

Fundamentally, the problem with Inter-Domain and Inter-Administration routing is that autonomy and mistrust are both antithetical to routing. Accomplishing either will involve a number of trade-offs which will require more knowledge about the environments within which they will operate.

4. EXAMPLE

The following example/problem gives perspective to some of the above discussion. Consider two corporations, A and B, each of which has research centers in several continents around the world. The research centers of their respective corporations communicate with each other through their corporate network—that is, the research centers do not have direct links with each other. Assume that the two corporations embark on a joint research project and Corporation A contributes one of its research centers in Continent 1 (called joint research center A), and Corporation B contributes one of its research centers in Continent 2 (called joint research center B). Let us assume that both Corporations parse their addresses in the format *Corporation.Continent.Region*, and that the address

of joint research center A is *A.1.1* and the address of joint research center B is *B.2.1*. It is necessary that the two joint research centers exchange data, but because the two corporations are competitors, they do wish to exchange data otherwise.

Assume it is determined that the most cost-effective way to allow the two joint research centers to communicate is by adding a link between the two corporate networks which connects locations which are not near either research center. This is depicted in Figure 2. If this link goes down, we assume that the two joint research centers wish to communicate alternatively via the more expensive public data networks. In Figure 2 we see a connection between an IS in Corporation A which is in a different region from Joint Research Center A to an IS in Corporation B which is in a different continent from Joint Research Center B. Because only communications between *A.1.1* and *B.2.1* can cross this link, it is necessary to constrain routing updates which cross that link to only mention *A.1.1* or *B.2.1*. It is also necessary to impose access control mechanisms on both ends of that link to prevent other data from crossing. While this may appear straight-forward enough, it is not a typical function of existing routing protocols.

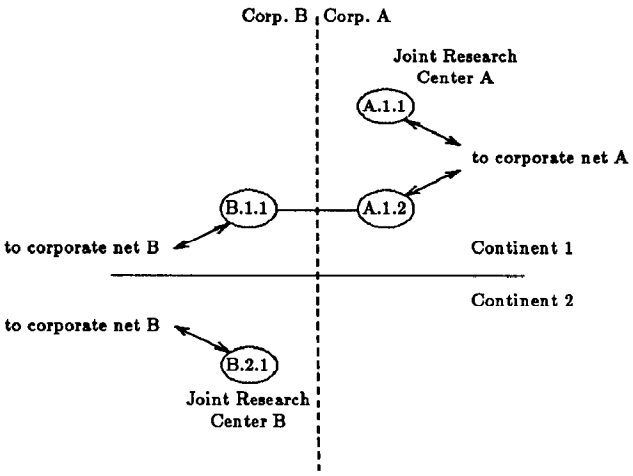

Figure 2

This kind of scenario can be made arbitrarily complex. For instance, what if a third corporation C had also contributed a research center, but this center was only connected to Corporation B, not Corporation A? Then the restrictions on the links between A and B and between B and C become still more complex. The amount of information required to describe complex restrictions on links and Systems, and the techniques for exchanging (or not exchanging) that information has not been studied.

5. DISCUSSION

We have outlined some approaches to the problems of routing in the global OSI data communications network. In spite of the extensive research in dynamic routing for data networks over the last twenty years (a modest bibliography is included), it seems that we still have more questions than answers. This is largely because most of the work has centered on routing in a non-hierarchical, singly-administered network. Some of the recent work has focused on hierarchical, singly-administered networks, but there is little work on hierarchical, multiply-administered networks.

With regard to multiply-administered networks, there are questions about what kinds of routing agreements can be made, how those agreements can be modeled, how they translate into restrictions on routing information and data traffic, and how they can be enforced.

There are questions concerning the nature of global networks. How much of the global connectivity will be provided by public domain networks, how much by private? How will routing algorithms in public domain networks interact with those in private networks? In many respects, good routing translates into power—power to circumvent public data networks, power to construct global networks without a central point of administration or control. How will this power be handled, legally and economically?

There are questions concerning faulty routing. How will faulty routers be isolated and fixed? This is hard enough to do in a singly administered network. How will it be done when routing takes place across arbitrarily complex topologies and agreements? How can the collapse of the global network due to faulty routing be prevented?

What is the relationship between routing at the network layer and routing at the application layer, for instance, mail handling systems? Should they be completely separate, or can information from one make the other more efficient?

6. SUMMARY

This paper gives a progress report on network layer routing architecture work in ANSI X3S3.3 for ISO. Some of the problems encountered in network layer routing are discussed, and an architectural basis for modeling and solving some of those problems is presented. In particular, a functional partition of the routing problem into four tiers (ES-IS routing, IS-IS routing within a single Routing Domain, IS-IS routing between different Domains but within an Administrative Domain, and IS-IS routing between Administrative Domains) is presented. A routing hierarchy, an address hierarchy, and the relationship between the two, are discussed. An example of a routing agreement between different administrations, and the constraints that are placed on routing as a result, is given. Finally, questions concerning outstanding problems in routing in the global OSI network are posed. It is hoped that these questions will help motivate discussion and research.

BIBLIOGRAPHY

American National Standards Institute ANSI X3S3.3, 3.3/86-215R2, *Draft Routing Architecture*, April, 1987.

Callon, R. and Lauer, G., *Hierarchical Routing for Packet Radio Networks*, Report No. 5945, SRNTN No. 31, BBN Laboratories Incorporated, June, 1985.

T. Cegrell, *A Routing Procedure for the Tidas Message-Switching Network*, IEEE Trans. on Communications, Vol. COM-23, No. 6, June 1975, pp. 575-585.

M. Gerla, *Controlling Routes, Traffic Rates, and Buffer Allocation in Packet Networks* IEEE Communications Magazine, Vol. 22, No.11, November 1984, pp. 11-23.

S. Gruchevsky, D. Piscitello, *The Burroughs Integrated Adaptive Routing System (BIAS),* ACM Computer Communication Review, Volume 17, Nos. 1 and 2, January/April 1987

J. Hagouel, *Issues in Routing for Large and Dynamic Networks,* PhD. Thesis, Columbia University, 1983.

J. M. Jaffe, F. H. Moss, *A Responsive Distributed Routing Algorithm for Computer Networks,* IEEE Trans. on Communications, COM-30, No. 7, July 1982, pp. 1758-1762.

Kamoun, F. and Kleinrock, L., *Hierarchical Routing for Large Networks: Performance Evaluation and Optimization, Computer Networks,* Vol. 1, pp. 155-174, 1977.

Kamoun, F. and Kleinrock, L., *Stochastic Performance Evaluation of Hierarchical Routing for Large Networks, Computer Networks,* Vol. 3, No. 5, pp. 387-353, November, 1979.

Kamoun, F. and Kleinrock, L., *Optimal Clustering Structures for Hierarchical Topological Design of Large Computer Networks, Computer Networks,* Vol. 10, No. 3, pp. 221-248, 1980.

Khanna, A. and Seeger, J., *Large Network Routing Study Design Document,* Report No. 6119, BBN Communications Corporation, January, 1986.

J. M. McQuillan, I. Richer, E. C. Rosen, *The New Routing Algorithm for the ARPANET,* IEEE Trans. on Communications, Vol. COM-28, No. 5, May 1980, pp. 711-719.

J. Mogul, J. Postel, *Internet Standard Subnetting Procedure,* RFC950, SRI International, Network Information Center, Menlo Park, CA, August 1985.

R. Perlman, *Hierarchical Networks and the Subnetwork Partition Problem,* Computer Networks and ISDN Systems 9, North-Holland, 1985, pp. 297-303.

R. Perlman, *Fault Tolerant Broadcast of Routing Information,* Computer Networks, Vol. 7, 1983, pp. 395-405

Saltzer, J., Reed, D., and Clark, D., *Source Routing for Campus-wide Internet Transport, Proceedings of the IFIP WG 6.4 Workshop on Local Networks,* August, 1980.

M. Schwartz and T.E. Stern, *Routing Techniques used in Computer Communication Networks,* IEEE Trans. on Communications, Vol. COM-28, No. 4, April 1980, 539-552

Shacham, N., *Hierarchical Routing in Large, Dynamic Ground Radio Networks,* SRI International, November, 1985.

Sparta Incorporated, *Design and Analysis for Area Routing in Large Networks,* McLean, VA, April, 1986.

Sunshine, C., *Addressing Problems in Multi-network Systems, Internet Engineering Note (IEN) 178,* April, 1981.

Zakon, S., *An Architecture for Routing in the ISO Connectionless Internet, Computer Communication Review,* Vol. 15, No. 5, pp. 10-39, October, 1985.