Interdomain optical routing

Greg M. Bernstein
Ciena Corporation
10480 Ridgeview Court, Cupertino, California 94014
gregb@ciena.com

Vishal Sharma
Metanoia, Incorporated
305 Elan Village Lane, San Jose, California 95134
v.sharma@ieee.org

Lyndon Ong
Ciena Corporation
10480 Ridgeview Court, Cupertino, California 94014
lyong@ciena.com

Received 5 January 2002; revised manuscript received 29 January 2002

As optical network deployment gains momentum, the efficient routing of optical connections across national and global networks assumes fundamental importance. When networks are partitioned into domains, for vendor interoperability, protocol scaling, or administrative ease, the nature and degree of topology and resource abstraction in routing protocols must be balanced by the need for intelligent information sharing to enable effective path computation. Here we focus on interdomain optical routing. We first highlight some major differences between optical circuit routing and Internet Protocol (IP) datagram routing and examine neighbor discovery and diverse routing in the optical case. We then develop a taxonomy of the information that can be shared between domains, and we discuss several applications of interdomain optical routing for the single-carrier case. Finally, we highlight some key issues for the multicarrier case.

OCIS codes: 060.0060, 060.4250.

1. Introduction

As carriers build global and national networks, the requirement to promote interworking and manageability creates a natural need for systematic partitioning of these networks into subnetworks. The control interfaces between these network partitions and the nature and extent of information shared across them then become important, since they influence *path selection*, which is critical for efficient traffic management. An important next step in the evolution of the control plane of optical networks therefore is *interdomain optical routing*.

Networks are typically partitioned to promote scaling of the intradomain routing protocols or to demarcate administrative or vendor boundaries. Thus the internal topological details of a domain (characteristics of its links, nodes, and subnetworks) are usually not shared externally (or shared to a limited extent). However, computing optimal and diverse paths across multiple domains in an optical network requires intelligent topology abstraction by means of routing protocols, which in turn requires careful analysis of the types of information that can be shared between domains in an optical network and of the applications that such information sharing enables. This is our focus in the current paper. [Interdomain signaling, which is the other piece of the puzzle necessary to make end-to-end provisioning work, is not discussed in this paper. The

interested reader is referred to documents from the Internet Engineering Task Force (IETF)[1] and the International Telecommunication Union (ITU)[2] for details.]

In Section 2 we define the concept of domains and network partitioning more formally and compare and contrast optical circuit routing with Internet Protocol (IP) datagram routing. We then focus on the issue of path computation and consider different models of route computation to motivate the need for sharing topology information. Finally, we look at the processes of neighbor discovery and the discovery of diverse routes in optical networks. In Section 3 we provide a decomposition of the information that may be shared (or hidden) between domains by use of the routing protocol. In Section 4 we consider various applications of interdomain optical routing within a single carrier's network. We discuss some of the most important cases and see how they relate to the taxonomy given in Section 3. In Section 5 we briefly discuss the main issues in interdomain optical routing across multiple carrier networks. Our concluding remarks appear in Section 6.

2. Background

2.A. Basic Concept of Domains and Network Partitioning

Domains are a partitioning of a network into subnetworks (or groups of subnetworks) with defined relationships between them. For example, the network in Fig.1 has been partitioned into three domains. In this paper we focus on how partitioning affects the control interfaces between domains. From a network control perspective, the motivations for dividing a network into domains may be as follows:

- To define administrative boundaries between network operators.
- To allow for scalability of routing and/or signaling.
- To enable isolation of partitions for security or reliability.
- To accommodate technology differences between systems, for example, by means of partitioning a single carrier's network into separate single-vendor subnetworks.

An *interdomain interface* is likely to have different characteristics than an *intradomain interface*, since the domain boundary exists for the purpose of hiding certain aspects of the domain from the outside world.

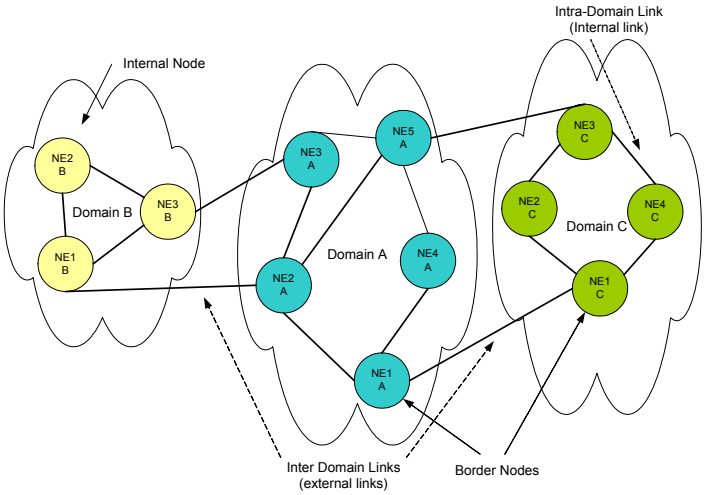

Fig.1. Partitioning a network into domains.

Examples of the use of domains include Border Gateway Protocol (BGP) *Autonomous Systems* (AS) and Open Shortest Path First (OSPF) Areas. An AS is a

subnetwork under the control of a single administration, as viewed from outside the domain. In the BGP case, an AS-to-AS interface may be an interface between two separate carriers or Internet service providers (ISPs), or an *intercarrier domain interface*. An OSPF Area, however, is commonly a subnetwork within an administrative domain (OSPF specific) belonging to the same carrier. An Area-to-Area interface is then between portions of the same carrier's network, or an *intracarrier domain interface*.

The domain concept as used here is orthogonal to the transport-network concept of layering. In the transport network, layers are technology specific and are used for multiplexing, performance monitoring, and fault management—with different layers providing different capabilities in each of these areas. An optical routing protocol must include information particular to the technology layer for which it is being used. For more information on layering and domain concepts, see ITU recommendation G.805.[3]

2.B. Optical Circuit and Internet Protocol Datagram Routing

In IP datagram networks, data (packet) forwarding is done on a hop-by-hop basis, with no connection established in advance. In circuit-switched optical networks, data forwarding is done on an end-to-end basis and requires a connection to be explicitly established ahead of time. Also, in IP networks each router makes an independent routing decision based on the destination address in a datagram. Thus, to prevent misrouting or loops, each node must use the same underlying topology database and use the same algorithm for route computation. By contrast, in optical networks the routing decision(s) are given in the setup request for the circuit, and existing traffic is unaffected, since the setup request is routed through the network. This is because distributed routing intelligence for circuit services does not affect existing traffic. In fact, in the absence of preemption, neither does distributed signaling.

Since the routing protocols in optical networks are not explicitly involved in data plane forwarding decisions, they do not affect service. For example, topology and resource status inaccuracies may affect whether a new connection (or a restoration connection) is established but will not (and should not) cause an existing connection to be torn down. This allows for greater flexibility in incorporating new information fields in optical routing protocols. Indeed, any information that can aid route computation or be used for service differentiation may be incorporated into the routing protocol, either as a standard element or as a vendor-specific extension. Since optical connections are explicitly (although perhaps loosely) routed, and since the basic route calculation is an atomic service that occurs, for a given connection, in a single network element (NE), NEs from different vendors need not use the same algorithm. Indeed, two different route computation algorithms may use the same information in different ways or not at all.

The bandwidth accounting needed in optical circuit-switched networks is also different than in packet networks. In packet networks that use either asynchronous-transfer-mode quality of service (ATM QoS) or multiprotocol-label-switching traffic engineering (MPLS-TE), complex statistical measures are used to characterize the load on a link, often with varying degrees of success. The inexactness of such measures and the compressibility of statistically multiplexed traffic imply that a small percentage change in link use can usually be absorbed by the network. By contrast, if an OC-192 link (OC, optical carrier) has just one STS-1 path (STS, synchronous transfer signal) occupied (less than 1% of the link bandwidth), it cannot accommodate an STS-192c path. In view of the relatively simple finite multiplex structures currently use in optical networks, tracking bandwidth resources is much easier than packet-switched networks; however, much stricter bandwidth accounting is required on circuit-switched links. In particular, with an individual optical-circuit switched link it is expected that this link can be fully utilized, although because of queuing effects a packet-switched link on average can never be run at full capacity and is typically run at less then 80% of capacity. This also affects how the protection (or restoration) bandwidth can be committed. In a packet-based network, although the protection path can be preconfigured, resources along it are not

used until a failure occurs. In circuit-based networks a protection path generally implies a committed resource. Such basic differences restrict the direct applicability of some of the traffic-engineering mechanisms used in packet-switched networks to circuit-switched networks.

2.C. Intradomain Routing for Packets and Circuits

We should point out that a routing protocol may also be used between nodes within a domain. Much effort has gone into extending existing IP-based intradomain routing protocols, also known as *Interior Gateway* Protocols (IGPs), for use in optical networks. For a tutorial overview of this work applied to synchronous optical network and synchronous digital hierarchy (SONET/SDH) optical networks see the paper by Bernstein *et al.*[4] The routing discussed in this paper builds from this existing conceptual base but applies these notions to the interdomain context.

The popularly used OSPF IGP cannot directly be applied in the interdomain case: first, because it can run only across areas that are under the same administrative entity, and second, because it assumes that each area, in turn, uses OSPF inside. The private network-to-network interface (PNNI) routing protocol from ATM, which has been the subject of considerable technical research, while supporting a notion of hierarchy still does not satisfy the requirements of a true domain-to-domain routing protocol. It also requires each level of the hierarchy to run a version of PNNI. Indeed, the only true domain-to-domain protocol in use today is the BGP protocol used in the Internet. BGP does not care or dictate which, if any, routing protocol is used within a domain. For example, BGP is frequently used between domains running different and incompatible internal routing protocols [typically, OSPF or IS–IS (Intermediate System to Intermediate System Protocol)]. Unfortunately, BGP is distance-vector based and so does not have mechanisms to deliver the topology and link status information that is so important in computing optical routes. Hence part of the motivation here is to examine closely what information an optical interdomain route protocol should disseminate.

2.D. Discovery Processes and Routing

The process of *neighbor discovery* is fundamental to many intradomain IP routing protocols. In the optical case this process involves discovery of link characteristics as well. In essence, neighbor discovery is used for automatically discovering a neighbor node *and* for understanding the fiber connectivity between the two nodes. Although most of the other information flows associated with the optical control plane can run either *in-fiber* or *out-of-fiber*, a portion of the discovery process must run in-fiber.

To facilitate interdomain optical routing, information concerning domain membership must be shared during discovery. For example, in the IP routing protocol OSPF, a subprocess, sometimes referred to as the Hello protocol, is used to discover IP layer neighbors. In addition, the Hello messages contain information used to identify an OSPF area (a subdomain within an OSPF routing domain). In Table 1 we show the domain, node, and port mappings obtained by the discovery process for the interdomain links in the network of Fig. 5 (below). A general theoretical model for the discovery function can be found in G.7714 (formerly G.disc),[5] one possible general mechanism is defined in the Link Management Protocol (LMP) draft of the IETF,[6] and an extension and specialization of LMP for SONET/SDH appears in the Optical Internetworking Forum's User Network Interface (UNI 1.0; Ref. 7) specification.

Table 1. Discovered Interdomain Neighbor Information for the Network of Fig. 5.

Link Endpoint			Link Endpoint		
Domain	Node	Port	Domain	Node	Port
A	NE2	5	B	NE3	3
A	NE2	7	B	NE1	1
A	NE3	9	B	NE3	4

2.E. Diversity in Optical Routing

One of the main uses of optical routing protocols is the discovery of diverse routes for optical connections. Multiple optical connections are often set up between the same end points, for example, the primary and backup connections for the same service. An important constraint on these connections is that they must be diversely routed.[8] In particular, they could be routed over paths that are *link diverse*, i.e., which do not share any common link, or *node diverse*, i.e., which do not share any common node.

When multiple physical and electronic levels are considered, ensuring link diversity is complicated by the possibility of two paths being diverse at one layer but failure correlated at another. The Shared Risk Link Group (SRLG) concept[9] generalizes the notion of link diversity to take into account diversity at multiple levels in addition to time division or wavelength multiplexing. These include the following:

1. *Cable (conduit) diversity.* Tells which fibers are in the same cable (conduit) and helps to avoid sending signals over routes that are most vulnerable to ordinary cable cuts (technically known as backhoe fades).
2. *Right of way (ROW) diversity.* Helps to avoid sending signals over routes that are subject to larger-scale disasters such as ship anchor drags or train derailments.
3. *Geographic route diversity.* Helps to avoid sending signals over routes that are subject to various larger-scale disasters such as earthquakes, floods, tornadoes, or hurricanes. A route could be approximately described by a piecewise set of latitude–longitude or universal transverse mercator coordinate pairs.

The concept of node diversity can also be generalized to abstract nodes. From a node's point of view, the diverse-routing specific concepts of interest include the following:

1. Nodes, i.e., individual switching elements.
2. Switching centers, i.e., a central office or exchange site.
3. Cities, or towns that contain more that one switching center.
4. Metro areas, or counties.
5. States.
6. Countries.
7. Geographic regions.

3. Routing Information Categorization

Different applications of interdomain optical routing call for different types of information to be shared or hidden between domains. In the following we decompose the information that can be transferred by means of a routing protocol broadly into link–topology information and node–domain information. We also further subdivide these categories and will use this taxonomy of routing information when discussing the routing applications.

3.A. Link and Topology Related Information

Internal topology information is information concerning the nodes and links and their connectivity within a domain. It is traditionally shared within a domain by means of an intradomain (interior gateway) routing protocol such as OSPF, IS–IS, or PNNI. For example, the existence of nodes that have links only to other nodes within the domain, i.e., that do not have links to other domains, is strictly internal topology information. These nodes are known as *internal nodes*, whereas nodes with links to other domains are known as *border nodes*. Also included in this information is link–port property information, such as whether the link is protected and, if protected, what type of protection is being used on it, e.g., linear 1 + 1, linear 1:*N*, or some type of ring such as a 4F-BLSR[10] (BLSR, bidirectional line-switched ring).

Internal resource information is concerned with the bandwidth available on links within a domain and possibly other resource-related information. It plays an important role in path selection within a domain.

Interdomain topology information is concerned with the interconnections between domains. This information can be key in interdomain path selection, for example, in determining diverse routes. For the network in Fig. 1 this information would let us know that domain A has two distinct links to domain B, and that domain A has two distinct links to domain C, but that domains B and C are not directly connected.

Interdomain resource information is concerned with the available bandwidth on interdomain links. It also is important for interdomain path selection and interdomain traffic engineering. For example, in Fig.1 this information would give us a bandwidth or capacity measure on the links between domains A and B, and on the links between domains A and C. The exact nature of this information may be application or context dependent.

3.B. Domain and Node Related Information

Reachability information tells us which addresses are directly reachable by means of a particular domain. These systems can be end systems (clients) to the network or nodes within the network, depending upon the application or context. Assume that in domain B of Fig.1 each of the network elements, NE1–NE3, has subtending end systems and that NE1–NE3 do not represent a valid final destination for a path. Under this assumption, the collection of the addresses of all the subtending end systems would form the reachability information for domain B.

Subnetwork capability information describes the capabilities or features offered by the domain as a whole. This information is used in applications in which sharing the internal topology and resource information is inappropriate. This information can include, for instance, (a) switching capabilities, (b) protection capabilities, (c) an overall available-capacity measure, and (d) reliability measures. For example:

(i) In SONET, one subnetwork may switch down to an STS-3c granularity while another switches down to an STS-1 granularity. Understanding which types of signals within a SDH/SONET multiplex structure can be switched by a subnetwork is important. Similar examples of granularity in switching also apply to the waveband case.

(ii) Some networking technologies, particularly SONET/SDH, provide a wide range of standardized protection techniques, but not all domains offer all protection options. For example, a 2/4-F BLSR-based subnetwork could offer extra data traffic, ring-protected traffic, and non-preemptable unprotected traffic (NUT),[10] while a mesh network might offer shared, SONET line layer linear protection and mesh protection.

(iii) Some domains may be in locations that have lower incidences of link failure. Such information could be helpful in computing routes to statistically "share the pain."

End-system capabilities information: Whereas properties of the subnetwork are important for deciding which domain to use to access a system (in the case of multihoming), end systems also posses a wide variety of capabilities. Allowing end-system capabilities (such as a system's ability to support SONET/SDH virtual concatenation) to be distributed into a routing protocol may not be that advantageous though, since it may counter the ability to summarize reachability information. Detailed end-system information may alternatively be obtained means of a directory service or some type of direct query between the end systems.

3.C. Diversity Services and Domain Representation

Although it is not necessary and, frequently, not desired, there are a number of situations in which revealing something about the internal topology and resources of a domain can

be advantageous. Three particular situations warrant revealing some internal domain information in the optical transport network:

(a) To promote better overall network efficiency by means of traffic engineering.
(b) When a domain wants to offer some form of diverse transit service.
(c) When a domain offers diverse connection origination and termination services for end systems.

Better overall network efficiency was an original driver behind the complex-node-with-exception representation of a domain in ATM's PNNI routing protocol.[11] In ATM terminology a domain is called a *peer group*; the entity representing it, a *logical group node*.

In Fig. 2 we show the network of Fig.1, with each domain represented by an *abstract node.* Note that complete connectivity within a domain is typically assumed, not including resource-allocation issues. To allow for better traffic engineering across an abstract node (domain), some abstraction (approximation) of internal link connectivity can be useful. Such internal *abstract links* are shown for domain A (abstract node A) in Fig. 2. Note that these abstract links do not necessarily correspond to internal physical links within the domain. In addition to being helpful with traffic engineering, such an internal abstract link representation is also useful if a domain desires to provide a *diverse transit service.* An example of such a diverse transit service is also shown in Fig. 2. In this case the internal abstract links of domain A are advertised, along with their SRLGs, to the rest of the network so diverse paths can be computed. Two connections that originate at location A in domain B and terminate at location Z in domain C make use of this internal information to use the diverse transit service offered by domain A.

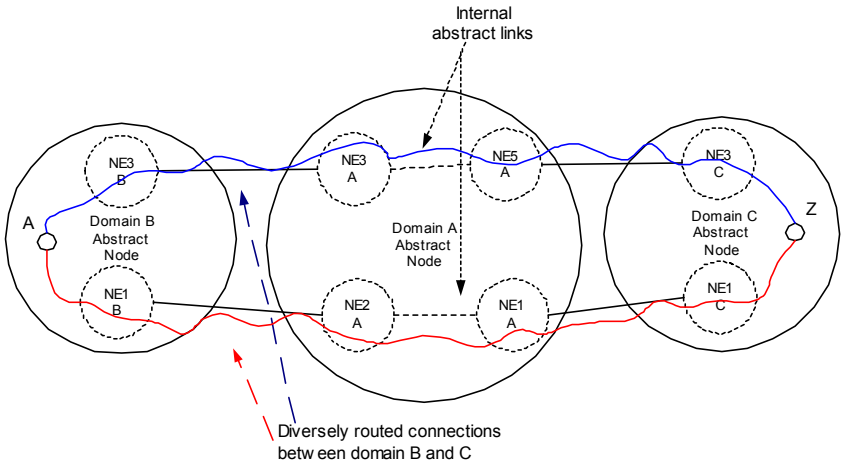

Fig. 2. Abstract node domain representation with abstract internal links to support a diverse transit service.

Figure 2 showed that modeling a domain by means of internal abstract links between border nodes maybe sufficient when a domain offers diverse transit services. However, it also seemed to leave some doubt concerning diversity at the origination and termination domains. In particular, this representation does not seem sufficient where a domain offers a *diverse origination and termination service* to end systems. Consider the control domain shown in Fig. 3. Here we show border nodes, nodes that terminate end systems, internal tandem nodes, internal links, and external links. To approximate this network to provide diversity connection services to end systems, we would need to reveal something about how those end systems are homed within the network.

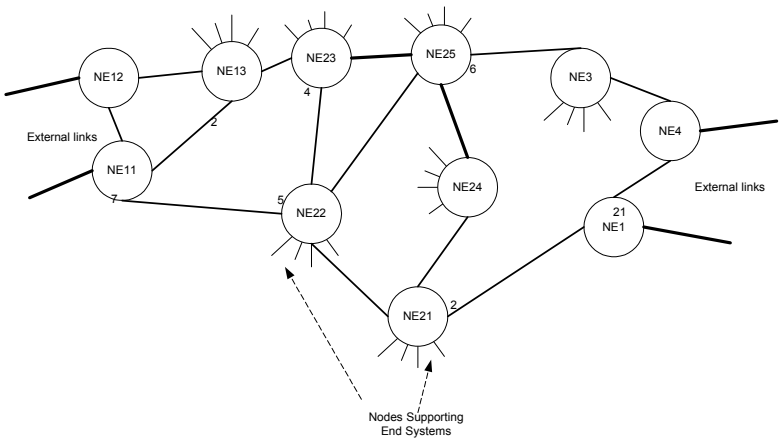

Fig. 3. Example control domain to be approximated.

In Fig. 4, we give an approximate control domain structure of the network shown in Fig. 3. In particular, we give a finer breakdown of internal structure in the form of an approximate subdomain view. For each domain, part of the basic information given is the reachability list for that domain. To do so, we use a subdomain abstract node model, where each abstract node for the subdomain has a subset of that domain's reachability list associated with it. Then we show subdomain abstract links connecting these subdomain abstract nodes to on another and, most importantly, to the border nodes. For example, in Fig. 4 abstract node 1 has associated with it reachability information for end systems attached to nodes 13, 23, and 25 in Fig. 3.

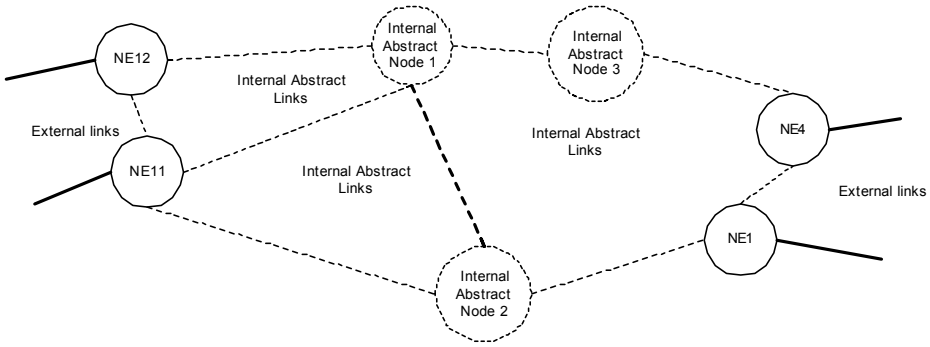

Fig. 4. Approximate representation for the control domain of Fig. 3

4. Intracarrier Interdomain Applications

Intracarrier interdomain routing arises when the optical network to be partitioned into areas is under the control of a single administrative entity. The main reasons for this partitioning may be, for example, scalability, intervendor interoperability, legacy equipment interoperability, and interlayer partitioning.

4.A. Intracarrier Scalability

As networks grow, it is useful to partition a carrier's network into separate optical routing domains that share limited or summarized information. This reduces the overhead of information exchange across the network and reduces the convergence time of routing protocols within a domain. Thus in the intercarrier scalability application we will hide or summarize internal topology and resource information while completely sharing interdomain topology and resource information to allow diverse path computation. Note

that general domain capabilities and capacity as well as reachability would tend to be shared as completely as possible.

Current IP routing protocols provide for this intracarrier scalability, by means of sharing only reachability information and not interdomain topology information. The main example is OSPF's multiple area capability. As discussed in Subsection 2.C, using multiarea OSPF or even ATM's PNNI protocol that supports hierarchical routing, sufficient information is not available at the source for diverse route calculation. When only an approximation of a domain's topology is made available to remote nodes, the signaling and call-processing functions at the domain border will receive an approximated (loose) route and must translate this into a precise route through the domain. Of course, the partitioning or grouping into domains can be recursive; i.e., we can have multiple levels of routing hierarchy to permit larger and larger networks. Such was the motivation behind the extensive hierarchy capability within ATM's PNNI routing protocol.

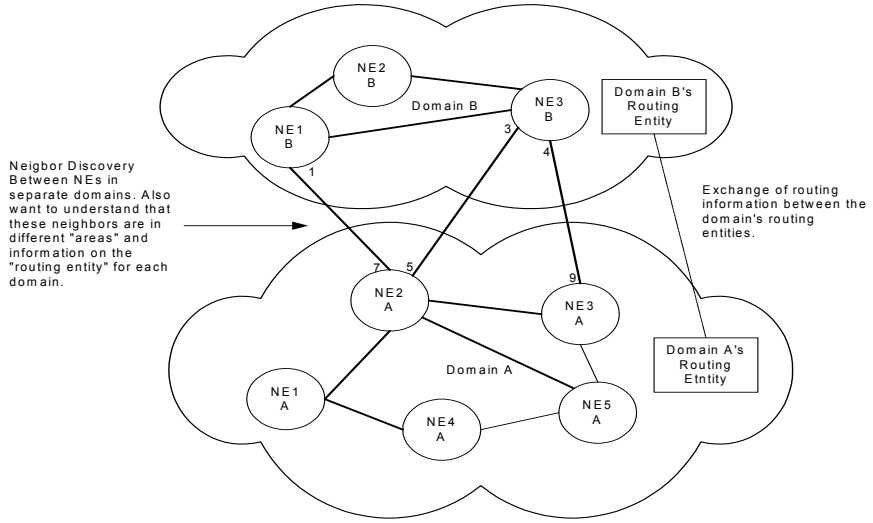

Fig. 5. Example network for interdomain routing.

Routing in optical networks involves calculating the entire route of the connection (albeit a loose route in some cases) through the network. As the request crosses different domains, a signaling or call-processing entity for the domain can take the loose routing information, and, from detailed information about its internal topology, compute a specific intradomain path through the domain. This process can be repeated within each domain, with the routing entities within each domain using different algorithms to calculate the portion of the path passing through their domain. For example, MPLS's loose routing capability allows one to specify a route for an optical connection in terms of a sequence of optical AS numbers, which can be handled by means of RSVP-TE's abstract node concept.

For example, in the network of Fig.1, a route between NE4 is domain C and NE2 in domain B could be specified simply as a sequence of domains {C, A, B} to be traversed by the setup request (and the circuit to be established between these nodes). This route would be locally expanded in domains A and B, either by a border node or by an appropriate entity in the domain responsible for resolving the route, to obtain the exact sequence of hops through that domain. For instance, within domain C, the route may be resolved as NE4 → NE1. The border node NE1 in domain A may resolve the local route to be NE1 → NE2 → NE3. Finally, border node NE3 in domain B may resolve the route (because of its local constraints) to be NE3 → NE2 → NE1. Once the corresponding acknowledgement is received by means of the RSVP-TE Resv message, the circuit would

have been established and become available to carry data (possibly after other end-to-end connectivity tests conducted by the carrier).

4.B. Intracarrier Intervendor

An important application of intracarrier optical routing is the intracarrier intervendor situation. From a carrier's perspective, the use of domains provides a clean way to isolate clouds of equipment belonging to different vendors while at the same time allowing for interoperability between those vendors.

An advantage of this method is that it allows the vendors complete freedom to use any combination of routing protocols or traditional management-based methods to propagate topology and resources internal to their domains. In other words, the *routing entity* in each domain (see Fig. 6) could obtain this information either by participating in a routing protocol such as OSPF, or by querying each NE by means of an EMS, or by simply having the required information manually configured into it.

In the intervendor case it can be particularly advantageous to centralize this system so that the flow of information can be monitored. A centralized routing entity could apply flooding and summarization mechanisms as if it is a switching system. The functions of the routing entity include (a) direct reachability exchange (that is, which NEs can be directly reached from this domain), (b) verification of area connectedness (that is, understanding how the two domains are interconnected), (c) area and domain topology (possibly summarized) exchange and updates, and (d) topology updates concerning other domains and areas.

The setup described above takes into account the three most important subcases of intercarrier intervendor partitioning:

(a) The first is where both vendor domains run proprietary distributed routing protocols within their domains.

(b) The second is where the optical subnetworks or domains (which includes a large number of existing installations) do not run any internal routing protocol (because the NEs are not capable of doing so), relying instead on EMS-based topology discovery and resource management.

(c) The third is where one domain has a centralized routing entity while the other runs a distributed routing protocol.

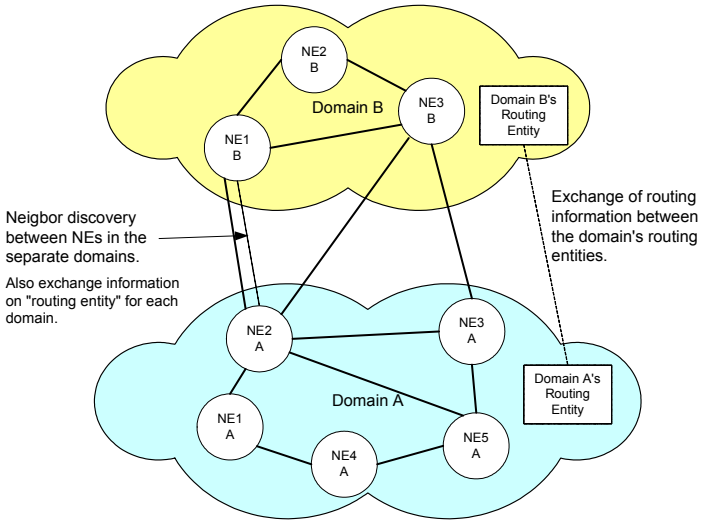

Fig. 6. Intracarrier intervendor routing domains.

4.C. Interbusiness Unit

A slightly different but interesting application of intracarrier optical routing occurs in the intracarrier interbusiness unit situation. This arises because a carrier often has multiple administrative domains, with groups of administrative domains being under the purview of independent BUs.

Different BUs represent independent cost centers with their own profit objectives and sales targets. As a result, although the BUs can profitably share topology information and would like to do so, they may not be so inclined to advertise the details of their resource usage into domains belonging to other BUs. Since each BU has its own revenue targets, advertising detailed resource availability information to other, potentially competing, BUs can have a negative effect on a BU's revenue generation. This is because detailed knowledge of available resources in one BU may enable other BUs within the carrier to requisition capacity from that BU. This would force the BU in question to yield to their request, possibly at the expense of selling capacity to more profitable, external, revenue-generating customers. Thus this situation has an additional dimension of information sharing, namely, policy-based information sharing, which does not apply to the other cases discussed so far.

Two examples in which this situation could become important are in metro-core-metro networks within a given carrier and in regional networks within a carrier. In the metro-core-metro situation in Fig. 7, the metro network domains are under the jurisdiction of one BU, while the core network domains belong to a different BU. In this case it is possible that the metro BU, armed with resource availability information about the core BU's domains, could claim capacity from the core network when needed. This harms the core network's profit goals, because the carrier may not be able to charge an internal customer the same rates that they could charge for the same capacity from an external customer. This motivates the need for selective, policy-based resource sharing between the BUs.

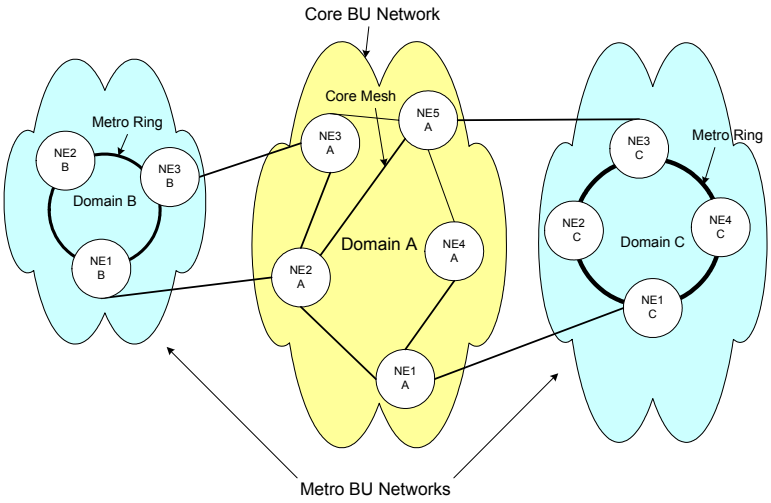

Fig. 7. Intracarrier interbusiness unit routing domains.

5. Intercarrier Interdomain Considerations

Interdomain optical routing across multiple carriers is considerably more complex than the single-carrier case. In the following, we highlight some of the major issues that effect the extent of sharing possible between carrier domains.

First, dealings with outside entities (different service providers) bring into play issues of trust and security not seen earlier. For instance, whereas revealing too much information about one's network is to be avoided, not revealing enough about certain

aspects, such as diversity capabilities, could drive customers elsewhere. Similarly, during route distribution a carrier may not wish to pass routing information that could point the way to a competing carrier's networks, thus affecting how route updates are performed!

The recent interest in bandwidth trading could motivate information sharing about a network's capability and connectivity but without revealing sensitive information such as reliance on other carriers for fibers or wavelengths, or information about unused network capacity, which could be financially sensitive. The issue is further complicated in that bandwidth trading could be realized with one of a number of business models, each having its own information requirements. In a *bandwidth broker model*, for example, the providers might post their asking prices for certain bandwidth services with the broker, and the broker could match a buyer with the lowest offer.

Another important issue is discovering diverse paths through domains belonging to multiple carriers. A simple extension of SRLGs to links between ASs would require that different ASs agree on the meaning of the 32-bit numbers composing a SRLG (which was previously local to an AS), which is difficult in itself. In addition, since path vectors that are AS diverse need not be fiber diverse, diverse routes may not be distinguishable based on differing path vectors alone! (Different carriers may have rented fiber from the same third party.) Therefore geographic link information may be the least contentious to get carriers to disclose.

Finally, in an interdomain intercarrier network, each carrier must be willing to advertise the destinations reachable via its network at each entry point and the specific formats of signals that each destination can accept. However, advertising this requires some protocol engineering, since the requirements of the optical case are at odds with the design goals of IP routing protocols such as BGP. With datagram routing the goal is simply to pick one route to a destination and make this choice consistent throughout the AS. Indeed, BGP specifically *reduces* the number of choices available, as per the following rule: "Fundamental to BGP is the rule that an AS advertises to its neighboring AS's only those routes that it uses. This rule reflects the 'hop-by-hop' routing paradigm generally used by the current Internet" (Ref. 12, p. 8).

Thus, whereas in the IP case we are merely interested in a consistent set of routes for hop-by-hop forwarding, in the optical case, we are interested in information that allows an optical path to be computed with whatever criteria are desired for that connection. Thus link-state-like routing protocols that do not obscure network topology (but perhaps provide an adequate abstraction of it) are required for the optical case, especially in the intercarrier situation. This is the subject of ongoing standards research.[13]

6. Concluding Remarks

In this paper we have given a detailed discussion of several issues in interdomain optical routing. We provided a categorization of routing information that can be shared between domains, depending on the relationship between those domains, and we discussed several intracarrier interdomain applications that highlight how the information in our taxonomy is used. In addition, we surveyed some of the issues in the intercarrier interdomain case, which has a number of economic, security, and trust considerations not present in the intracarrier case. We believe that our discussions point to some interesting possibilities for future research. One area is clearly the design of an optical interdomain routing protocol,[13] which needs to be non-topology-obscuring yet scalable. The other is a detailed consideration of intercarrier application situations. Finally, the connection between topology representation and loose routing is worth exploring, especially from a graph theoretic angle, where one may ask: To what extent can one abstract the physical network topology and still enable efficient traffic engineering and diversity services?

References

1. P. Ashwood-Smith, A. Banerjee, L. Berger, G. Bernstein, J. Drake, Y. Fan, K. Kompella, E. Mannie, J. P. Lang, B. Rajagopalan, Y. Rekhter, D. Saha, V. Sharma, G. Swallow, and Z. Bo Tang, "Generalized MPLS—signaling functional description," Work in Progress, draft-ietf-

mpls-generalized-signaling-07.txt, November 2001, http://search.ietf.org/internet-drafts/draft-ietf-mpls-generalized-signaling-07.txt.
2. ITU-T Draft Recommendation G.7713, “Distributed call and connection management,” October 2001, ITU Telecommunication Standardization Sector, http://www.itu.int/ITU-T/.
3. ITU-T Recommendation G.805, "Generic functional architecture of transport networks,” March 2000, ITU Telecommunication Standardization Sector, http://www.itu.int/ITU-T/.
4. G. Bernstein, E. Mannie, and V. Sharma, "Framework for MPLS-based control of optical SDH/SONET networks,” IEEE Netw. (July/August 2001), pp. 20–26.
5. ITU-T Draft Recommendation G.7714, “Generalized automatic discovery techniques, October 2001, ITU Telecommunication Standardization Sector, http://www.itu.int/ITU-T/.
6. J. P. Lang, K. Mitra, J. Drake, K. Kompella, Y. Rekhter, L. Berger, D. Saha, D. Basak, H. Sandick, A. Zinin, and B. Rajagopalan, “Link Management Protocol (LMP),” draft-ietf-ccamp-lmp-02.txt, Internet Draft, November 2001, http://search.ietf.org/internet-drafts/draft-ietf-ccamp-lmp-02.txt.
7. User Network Interface (UNI) 1.0 Signaling Specification, OIF2000.125.7, OIF, October 2001, Optical Internetworking Forum, http://www.oiforum.com/.
8. R. Bhandari, Survivable Networks: Algorithms for Diverse Routing (Kluwer, Boston, Mass., 1999).
9. K. Kompella, Y. Rekhter, A. Banerjee, J. Drake, G. Bernstein, D. Fedyk, E. Mannie, D. Saha, and V. Sharma, “IS-IS extensions in support of generalized MPLS,” Work in Progress, draft-ietf-isis-gmpls-extensions-01.txt, November 2000, http://search.ietf.org/internet-drafts/draft-ietf-isis-gmpls-extensions-07.txt.
10. ANSI T1.105.01-1995, “Synchronous optical network (SONET) automatic protection switching,” American National Standards Institute, http://www.ansi.org/.
11. Private Network-Network Interface Specification Version 1.0 (PNNI 1.0) af-pnni-0055.000, The ATM Forum, March 1996, The Asynchronous Transfer Mode Forum, http://www.atmforum.com/.
12. Y. Rekhter and P. Gross, "Application of the Border Gateway Protocol in the Internet,” RFC 1772, T. J. Watson Research Center, IBM Corporation, MCI, March 1995.
13. “Domain to Domain Routing Protocol (DDRP),” Work in Progress, OIF, December 2001, Optical Internetworking Forum, http://www.oiforum.com/.