Internet Draft
Expiration Date: May, 24, 2001

Dimitris Pendarakis
Bala Rajagopalan
Debanjan Saha
Tellium Inc.

Routing Information Exchange in Optical Networks

draft-prs-optical-routing-01.ps

1 Status of this Memo

This document is an Internet-Draft and is in full conformance with all provisions of Section 10 of RFC2026.

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 and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as ‘‘work in progress.”

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html.

2 Abstract

The objective of this draft is to describe mechanisms for exchanging routing information between IP clients and optical networks, and between optical sub-networks. Such information exchange would allow automated establishment of end-to-end paths in an optical network comprising of multiple sub-networks. Furthermore, it would allow optical network clients, specifically IP routers, to discover their remote peers dynamically. Determination of reachability could be the first step in dynamic provisioning of optical paths using signaling.

3 Introduction

Consider the optical networking model as shown in Figure 1. Here, clients (e.g., IP routers) are attached to an optical core network, and connected to their peers over dynamically established switched optical paths. The interaction between theclients and the optical core is over a well-defined signaling and routing interface, shown as the User-Network Interface (UNI).

The optical network shown in Figure 1 consists of multiple Optical Cross-Connects (OXCs) interconnected by optical links in a general topology refered to as an "optical mesh network". This network may be multi-vendor, where individual vendor OXCs constitute sub-networks. Each sub-network itself is assumed to be mesh-connected. The interaction between the sub-networks is over a well-defined signaling and routing interface, shown as the Network-Network Interface (NNI).

The optical network essentially provides point-to-point connectivity between clients in the form of fixed bandwidth optical paths. The collection of optical paths therefore defines the topology of the virtual network interconnecting the clients. This topology may be static by design. In this case, the optical paths may be provisioned "manually", i.e., without any need for signaling protocols at the UNI. The more interesting case is when the interconnection topology can change dynamically. In this case, control protocols at the UNI, as well as at the NNI, are necessary for dynamic provisioning of paths.

Figures 2a and 2b illustrate some applications of dynamic provisioning. In Figure 2a, a Virtual Private Optical Network (VPON) for IP routers is shown. Here, optical paths are provisioned between routers belonging to the same VPON, but the interconnection topology may change with time depending on traffic demands between the VPON nodes. Figure 2b illustrates the case where two logically separate VPONs are interconnected dynamically based on policy decisions. Both these cases require routing interaction between the optical network and the IP networks to facilitate the automatic configuration of the VPON topology.

In the next section, the model for routing across the UNI is considered. Routing information exchange across the NNI is considered in Section 5. Finally, summary and conclusion are presented in Section 6.

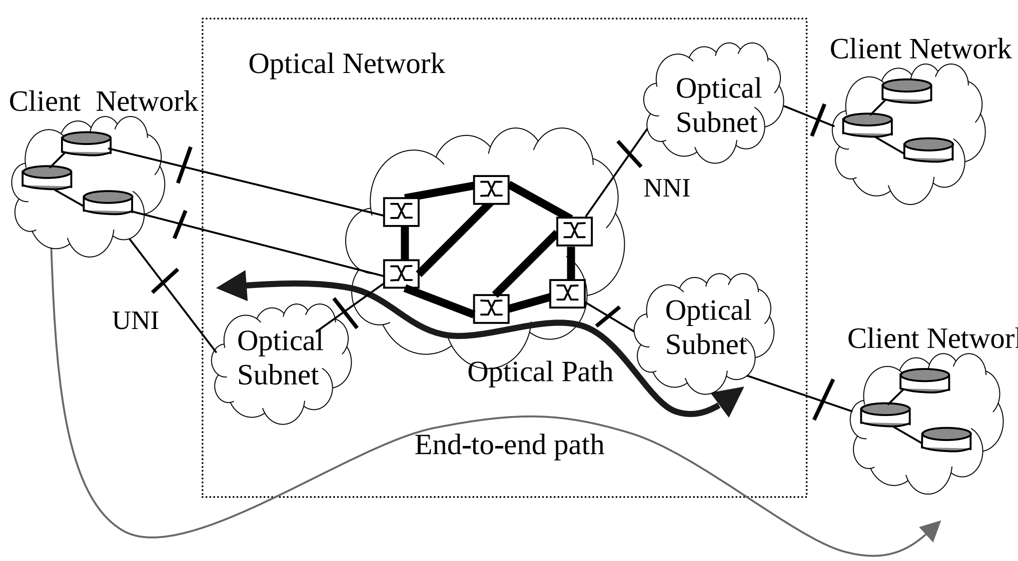


Figure 1: Optical Networking Model

4 Routing Information Exchange over the UNI

Let us consider the network model shown in Figure 1 and assume that the dynamic provisioning of an optical path is initiated by a client device using UNI signaling. Such a provisioning request must specify the destination endpoint in the optical network. This information can be made available to the client device in a number of ways. To describe these methods, let us consider the case where the client devices are IP routers. Let us designate the routers directly connected to the optical networks as *border* routers. The OXCs they are connected to are referred to as *border* OXCs.

1. The endpoint information may be configured in the client device. For example, each border router can be configured with optical endpoint addresses corresponding to IP destinations in other sites. This configured information, together with configured rules on dynamic provisioning, may be used to establish dynamic VPON topologies.
2. The endpoint information is obtained using a limited reachability protocol across the UNI. In this case, each border router runs the reachability protocol with the corresponding border OXC and obtains the address of every other border router belonging to the same VPON. Using this information, an initial set of IP routing adjacencies are established between border routers. The border routers then run an IP routing protocol amongst themselves to determine all the IP destinations reachable over the optical network.
3. The endpoint information is obtained using a routing protocol running across the UNI. In this case too, each border router runs the routing protocol (e.g., BGP) with the corresponding border OXC. Unlike

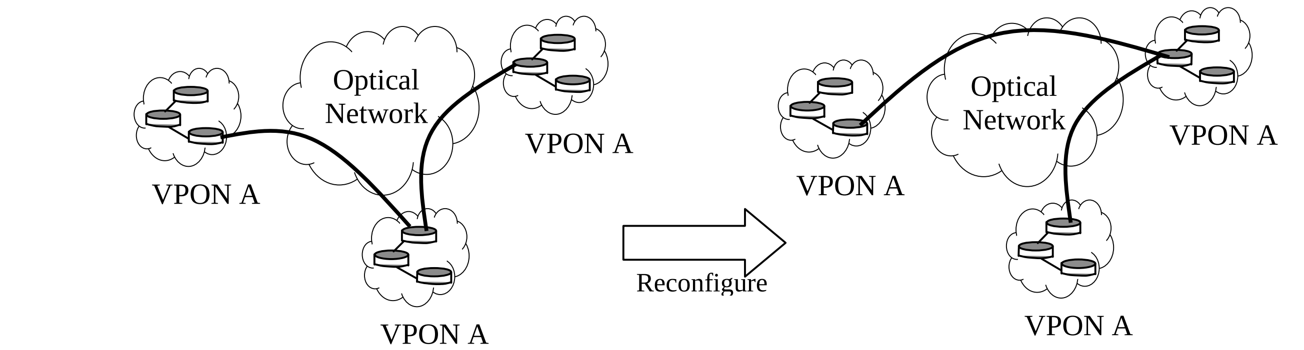


Figure 2a: Dynamic VPON

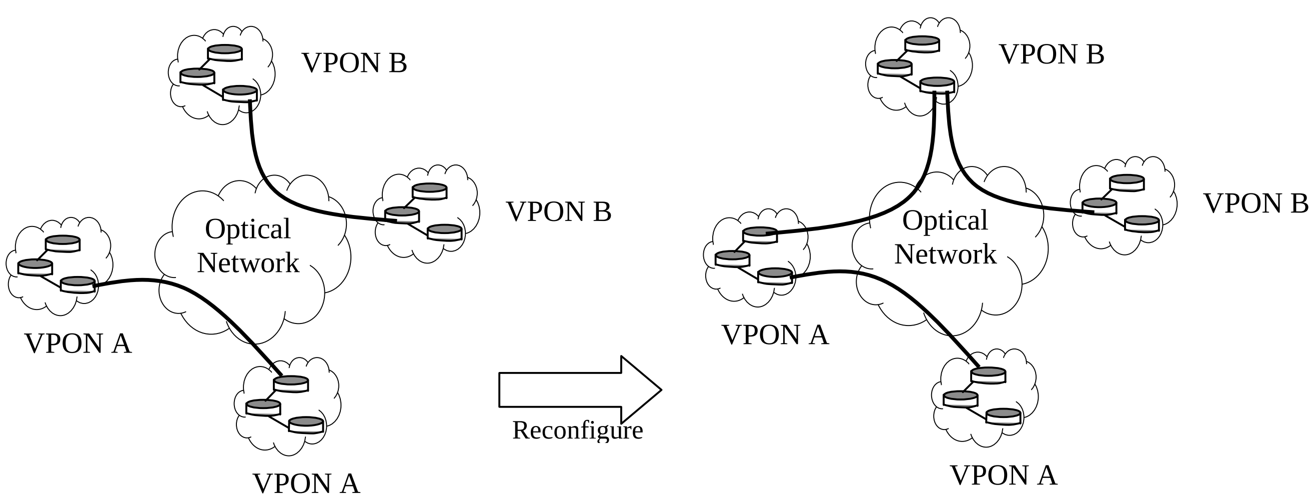


Figure 2b: Dynamic Interconnection of VPONs

(2) above, this protocol allows border routers to advertise all destinations reachable in their site to the corresponding border OXCs. The full reachability information is then conveyed to other border routers over the UNI. There is no need for border OXCs to establish IP routing adjacencies among themselves.

The last two options are referred to as *peer* routing organizations to reflect the fact that the client devices and OXCs are routing peers running a common routing protocol. We may, however, distinguish *partial* peering (option 2) which requires additional routing exchanges between clients across the optical network from *full* peering (option 3) which does not require these exchanges. (Here it is important to distinguish between the data and control planes over the UNI. The optical network provides a service to external entities in the form of fixed bandwidth transport pipes (optical paths). IP routers at the edge of the optical networks must necessarily establish such paths before communication at the IP layer can begin. Thus, the IP data plane over optical networks is realized over an *overlay* network of optical paths. On the other hand, IP routers and OXCs can have a *peer* relation on the control plane, especially for the implementation of a routing protocol that allows dynamic discovery of IP endpoints attached to the optical network.)

Clearly, not all of these three options may make sense in different networking scenarios. Specifically, with non-IP client devices configuration may be more convenient than other options. If the other options are considered in such settings, the reachability or routing protocols must convey non-IP client addresses across the optical network. For concreteness, however, we limit the following descriptions to the case where the clients are IP devices. It is easy to generalize the descriptions to non-IP clients. In the next two sub-sections, we consider full and partial peer routing models, respectively.

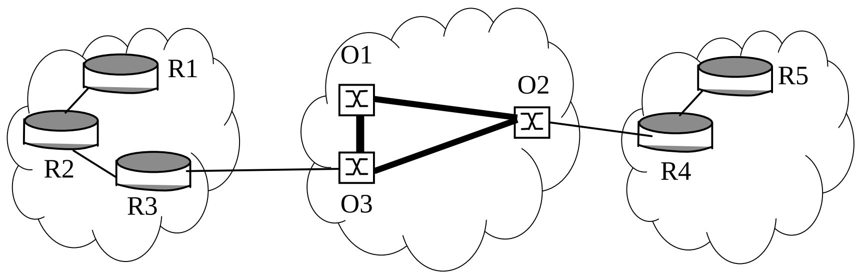
4.1 Full Peer Routing Models

We describe two different full peer routing organizations. First, a “flat” routing organization may be developed that allows a single routing protocol instance to run in both client and optical networks. Given that optical networks use IP-centric protocols, this organization is possible only with client networks that run the same IP routing protocols internally. Furthermore, this type of routing may be practical only when both the optical and client networks are administered by the same entity. Second, client and optical networks can be functionally separated, each running its own routing protocol, but exchanging full reachability information across the UNI using a standard protocol. This is a convenient solution, since it allows the development of provisioning and restoration procedures for optical sub-networks independent of client network routing. Also, this approach supports the common scenario where the optical network and client networks are administered by different entities. Additionally, there is a practical aspect to following this approach: it allows the same standard routing protocol to be used across the NNI for routing across disparate optical sub-networks (Section 3). In the following, these approaches are described in some detail.

4.1.1 Flat Routing Organization

Since the optical network implements IP-centric control protocols, it should be possible to export a representation of its internal topology to routers at the edge of the network. For example, an IP routing protocol like the Open Shortest-Path First (OSPF) [1] may be used across the IP/optical domains.

Figure 3 depicts flat routing organization using OSPF, where IP routers maintain a single topology database for the joint network consisting of IP and optical nodes. Here, the sample network topology has five IP routers and three optical switches. The OSPF Link State Advertisements (LSAs) at router R1 is represented abstractly in the table. Here, all the nodes and (unidirectional) links in the combined network are listed. Optical links are distinguished from other links, since the characterization of these links may include optical-link-specific information such as link type, composition of bundled links, etc. Thus, with flat routing, a router must maintain information that potentially has no meaning in the router network, but meaningful only within the optical network.



	LSA Type	ID	Advertising Node	Content of the LSA
1.	Nodal LSA	R1 IP addr.	R1	All links on R1
2.	Nodal-LSA	R2 IP addr.	R2	All links on R2
3.	Nodal-LSA	R3 IP addr.	R3	All links on R3
4.	Nodal-LSA	R4 IP addr.	R4	All links on R4
5.	Nodal-LSA	R5 IP addr.	R5	All links on R5
6.	Nodal-LSA	O1 IP addr.	O1	All links on O1
7.	Nodal-LSA	O2 IP addr.	O2	All links on O2
8.	Nodal-LSA	O3 IP addr.	O3	All links on O3
9.	Optical-LSA	R3-O3	R3	Optical link between R3and O3
10.	Optical-LSA	O3-R3	O3	Optical link between O3and R3
11.	Optical-LSA	R4-O2	R4	Optical link between R4and O2
12.	Optical-LSA	O2-R4	O2	Optical link between O2 and R4
13.	Optical-LSA	O1-O2	O1	Optical link between O1and O2
14.	Optical-LSA	O2-O1	O2	Optical link between O2and O1
15.	Optical-LSA	O1-O3	O1	Optical link between O1and O3
16.	Optical-LSA	O3-O1	O3	Optical link between O3and O1
17.	Optical-LSA	O2-O3	O2	Optical link between O2and O3
18.	Optical-LSA	O3-O2	O3	Optical link between O3 and O2

Figure 3: Flat Routing: Network Topology and Link State Information at R1

Assuming that routers are programmed to apply the correct semantics for the optical network information, IP routers can compute full paths to other IP destinations across the network. For example, router R1 can compute the path R1-R2-R3-O2-O3-R4-R5. This path may be signaled hop-by-hop from R1 to R5, using the appropriate protocols across the UNI and the NNI, and within router networks and optical sub-networks. Once the path is established, however, the segment R3-O2-O3-R4 must be treated as a single virtual link R3-R4 of fixed capacity (e.g., OC-48) and perhaps advertised as such in further OSPF updates. The restoration of the optical path within the optical network may be visible to all nodes in the network, thereby complicating the process.

4.1.2 Domain-Specific Routing Organization

Routing within the optical and IP domains may be separated, with a standard routing protocol running between domains. This is similar to the IP interdomain routing model. In this section, the focus is on the routing information exchange at the IP-optical UNI. There are two possibilities for this. We first consider the interdomain IP routing protocol, BGP [2], which may be adapted for exchanging routing information between IP and optical domains. We then consider the use of OSPF areas to facilitate routing across the UNI.

4.1.2.1 Routing Information Exchange using BGP

This would allow the routers to advertise IP address prefixes within their network to the optical network and to receive external IP address prefixes from the optical network. The optical network transports the reachability information from one IP network to others, as shown in Figure 4. Here, networks N1-N3 are assigned the address spaces indicated by the network prefixs *x.y.c.**, *a.b.c.**, and {*x.y.a.**, *x.y.b.**}. The propagation of the address prefixes from R4 to R3 through the optical network is shown. Exterior BGP (EBGP) is assumed to run between the IP routers and OXCs over the UNI (between *border* routers and *border* OXCs). Within the optical network, it is assumed that interior BGP (IBGP) is used between border OXCs within the same sub-network and EBGP is used over the optical NNI (in Figure 4, however, only a single optical sub-network is shown. Here, IBGP runs between border routers O1-O3). The IP address prefixes within the optical network are not advertised to routers using BGP. A border OXC receiving external IP prefixes from a router includes its own IP address as the egress point before propagating these prefixes to other border OXCs or border routers. For instance, in the example illustrated in Figure 4, the address of OXC O2 will be advertised along with the prefixes {*x.y.a.**, *x.y.b.**}. A border router receiving this information need not propagate the OXC address further, but it must keep the association between external IP addresses and egress OXC addresses. When a specific external IP address is to be reached, the border router can determine if an optical path has already been established to the appropriate egress OXC or a path must be newly established. Specific BGP mechanisms for propagating egress OXC addresses are to be determined, considering prior examples described in [3,4]. When VPONs are implemented, the address prefixes advertised by the border OXCs must be accompanied by some VPON identification (for example, VPN IPv4 addresses, as defined in [3], may be used). Border OXCs can then filter external addresses based on VPON identifiers before propagating them to routers, i.e., a router would only receive external IP addresses belonging to its own VPON. Once a router has determined reachability to external destinations, the dynamic provisioning of optical paths to reach these destinations may be based on traffic engineering mechanism implemented in the router.

4.1.2.2 Routing Information Exchange using OSPF

When the optical network and all client networks belong to a single routing domain, the routing information exchanged across the IP-optical UNI could be summarized using a hierarchical routing protocol such as OSPF.

OSPF supports a two-level hierarchical routing scheme through the use of OSPF areas. Routing within each area is flat, while detailed knowledge of an area's topology is hidden from all other areas. Routers attached to two or more areas are called Area Border Routers (ABRs). ABRs propagate IP addressing information from one area to another using summary LSAs. Within an OSPF routing domain, all areas are attached directly to a special area called the OSPF backbone area. The exchange of information between areas in some way is similar to BGP method of propagating reachability.

The use of a single OSPF routing domain with multiple areas is beneficial from the point of view of ease of

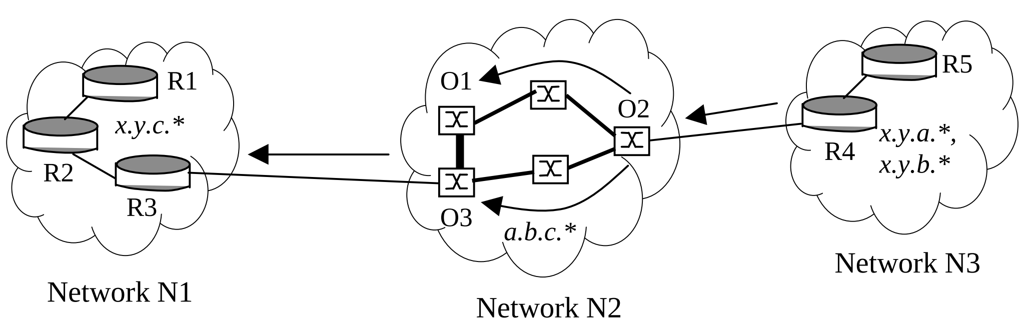


Figure 4: Domain-Specific Routing: BGP Example

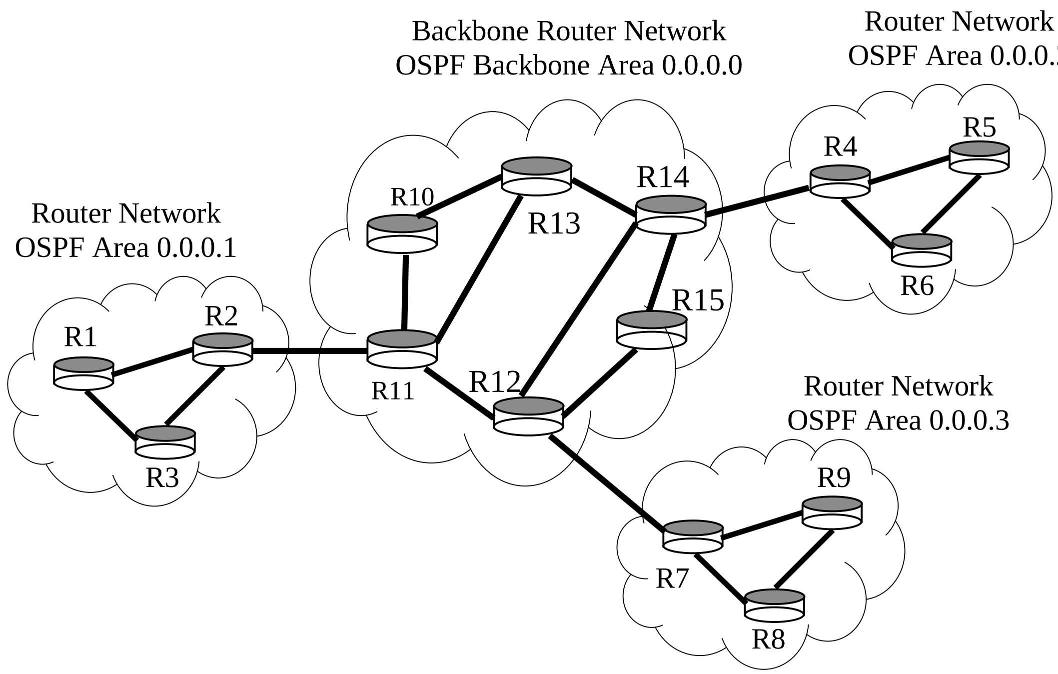


Figure 5: OSPF Routing Domain Organized into OSPF Areas

migration, as providers migrate to optically switched backbones. Figure 5 shows a single autonomous system organized into multiple OSPF areas. This sample network consists of 3 OSPF areas (0.0.0.1, 0.0.0.2 and 0.0.0.3) connected to the OSPF backbone area, denoted by 0.0.0.0. As shown in this figure, all areas are constructed using IP routers. Routers R2, R4, R7, R11, R12, R14 are ABRs and R2, R4, R7, R10-15 are backbone routers.

In Figure 6, the physical backbone router network of Figure 5 has been replaced with an optical network; this is simply achieved by replacing each backbone router with an optical switch. While the data plane characteristics of the optical network are completely different than those of the router backbone, the control plane remains essentially the same. As long as optical switches participate in the OSPF protocol, the optical network can serve as the OSPF backbone area, flooding summary LSAs between different areas. The optical network advertises external addresses into each area, along with the address of the ABR corresponding to each address and a cost metric. For example, switch O11 advertises addresses in an external area 0.0.0.3 into area 0.0.0.1. For those destinations advertised, it also indicates the address of R7, the ABR in area 0.0.0.3 and the cost of the path from O11 to O12 (to reach R7). The information about R7 is needed by R2 to determine where to establish a lightpath when communicating with destinations in area 0.0.0.3. The cost information may be used to select among multiple alternatives when a client network is multiply homed.

4.2 Partial Peer Routing Model

Running a protocol like BGP across the UNI may be considered too involved, at least for initial implementations of the UNI. A simpler approach would be to limit the reachability information passed through the optical network as follows:

1. Each border router belonging to a VPON registers a set of <IP address, VPON identifier> pairs (or a set of VPN IPv4 addresses) to a border OXC across the UNI.

2. The IP addresses of all border routers belonging to a VPON are propagated across the optical network (see Section 4)
3. These addresses are conveyed to each router that registers as a border router in the VPON.

The propagation of reachability information is illustrated in Figure 7. Here, three IP networks that are part of the same VPON are shown. The border routers have assigned addresses as shown. The flow of registration messages from border routers to border OXCs and the flow of reachability information (i.e., <IP address, VPON id> pair) in the reverse direction are shown. Within the optical network, a border OXC is assumed to originate routing advertisements for external IP addresses registered with it. This would allow interior OXCs to route optical paths destined to external IP addresses to the correct destination OXCs.

Now, once border routers in a VPON receive the address of other border routers within their own VPON, they may construct a VPON topology dynamically through UNI signaling. Assuming that each router has at least two interfaces to the optical network, a linear topology may be built automatically, as shown in Figure 8 (the initial topology may also be built based on configured information about routing adjacencies). Over this topology, the border routers may run their own IP routing protocol, for example, OSPF. In this case, the optical paths between the border routers will be represented as virtual links in the OSPF link state database. The initial topology may be modified dynamically, based on traffic engineering algorithms that are implemented in the border routers, as shown in Figure 2. Thus, the simple reachability protocol described above provides a mechanism for bootstrapping end-to-end IP routing within the VPONs across the optical network.

5 Routing Information Exchange over the NNI

Provisioning an end-to-end optical path across multiple sub-networks involves the establishment of path segments in each sub-network sequentially. Thus, a path segment is established from the source OXC to a

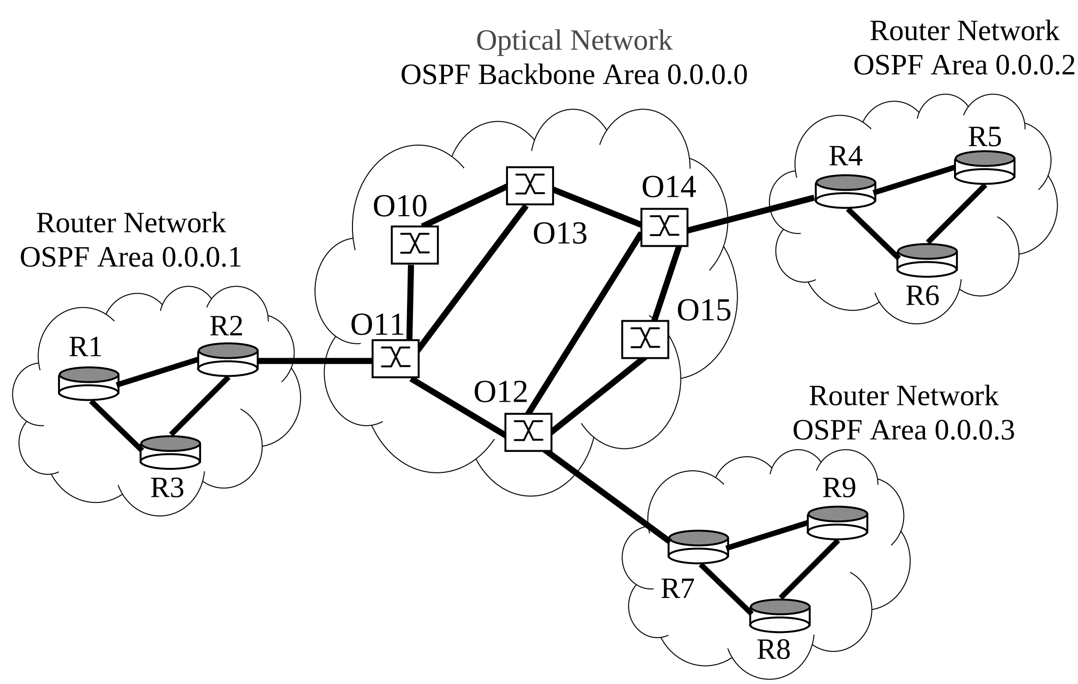


Figure 6: Optical OSPF Backbone Area

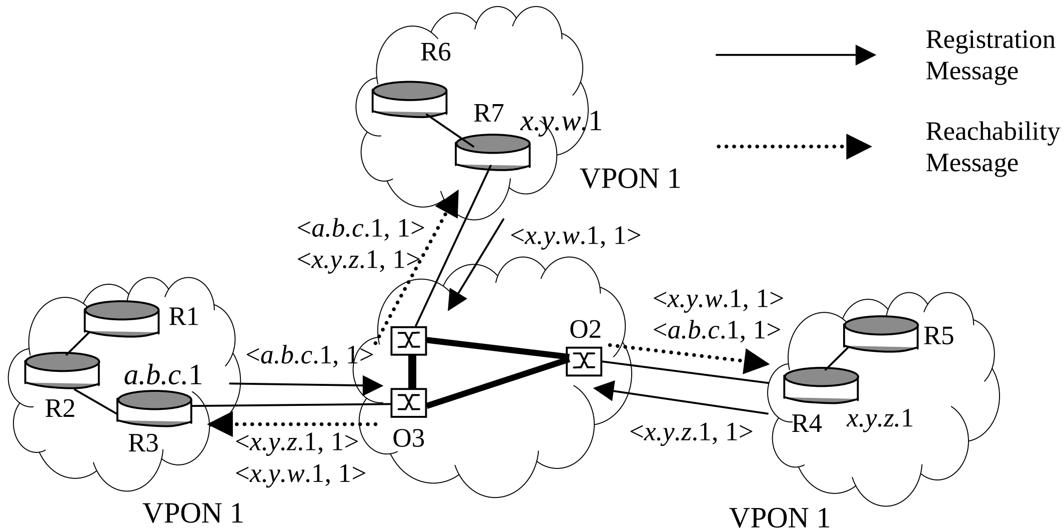


Figure 7: Domain-Specific Routing: Reachability Protocol

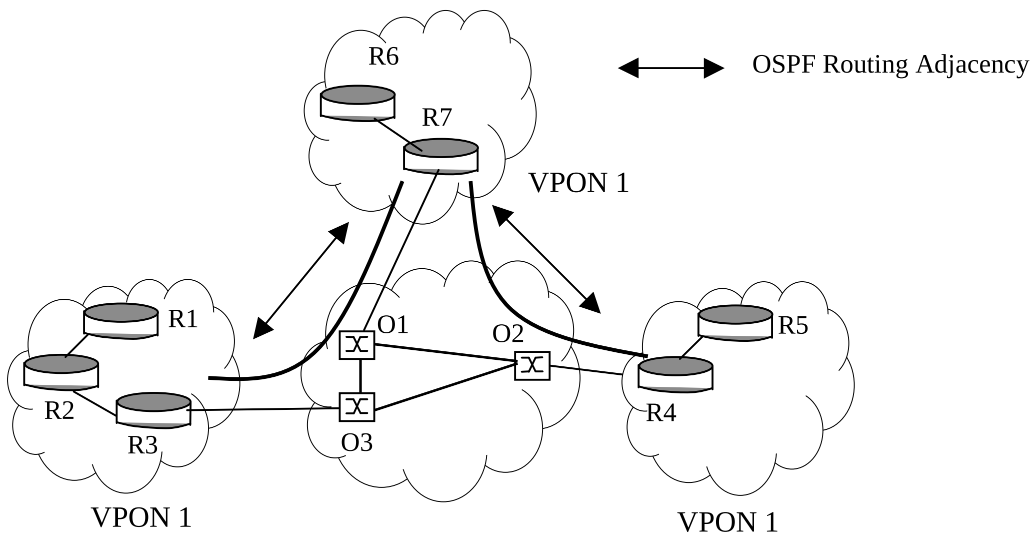


Figure 8: Domain-Specific Routing: Bootstrapping the VPON Topology

border OXC in the source sub-network. From this border OXC, signaling across the NNI is to establish a path segment to a border OXC in the next sub-network. Provisioning then continues in the next sub-network and so on until the destination OXC is reached. This procedure is shown in Figure 9, assuming CR-LDP signaling within sub-networks and a standardized NNI signaling for path set-up. To automate this process, it must be possible to determine the route to the destination OXC from within the source sub-network. A routing protocol must therefore run across the NNI between sub-networks. It is desirable that such a protocol allows the separation of routing between sub-networks. This allows proprietary provisioning schemes to be implemented within sub-networks while end-to-end provisioning is performed.

These objectives may be satisfied by running a version of BGP between border OXCs. For this, it is essential that the OXC IP addresses are unique network-wide. Using exterior BGP, adjacent border OXCs in different sub-networks can exchange reachability of OXCs and other external IP endpoints (border routers). Using interior BGP, the same information is propagated from one border OXC to others in the same sub-network. Thus, every border OXC eventually learns of all IP addresses reachable across different neighboring sub-networks. These addresses may be propagated to other OXCs within the sub-network thereby allowing them to select appropriate border OXCs as exit points for external destinations. To support the routing model described in Section 2.1, the external reachability information should include VPON identifiers.

It is clear that border OXCs must keep track of many IP addresses corresponding to different remote OXCs. The overhead for storage and propagating these addresses can be reduced if OXC addresses within a sub-network can be aggregated to a relatively few IP network prefixes. This is indeed possible if OXC addresses within a sub-network are derived from a small set of IP network prefixes.

Once border OXCs acquire reachability information regarding remote destinations, this information may be shared with other OXCs within the sub-network to enable end-to-end path provisioning. In short, a source OXC within a sub-network must determine the border OXC through which the ultimate destination can be reached. Also, if there is more than one such border OXC, a procedure must be available to select one of them. Finally, policy decisions may be involved in selecting a particular route. These issues are similar to interdomain routing in the Internet as discussed in [2].

5.1 Dynamic Provisioning Model

The model for provisioning an optical path across sub-networks is as follows. A provisioning request may be received by a source OXC from a management system, specifying the source and destination end-points. Such a request must explicitly specify the <OXC IP address, Index> pair identifying each end-point. On the other hand, a provisioning request may be received from an IP border router. In this case, the source end-point is implicit and the destination endpoint is identified by the IP address. In both cases, the routing of an optical path is done as follows:

1. The source OXC looks up its routing information corresponding to the specified destination IP address:
   a. If the destination is an OXC in the source sub-network, a path maybe directly computed to it.
   b. If the destination is an external address, the routing information will indicate a border OXC that would terminate the path in the source sub-network. A path is computed to the border OXC.
2. The computed path is signaled from the source to the destination OXC within the source sub-network. The complete destination endpoint address specified in the provisioning request (either <OXC IP address, Index> or <IP address>) is carried in the signaling message.
3. The destination OXC in the source sub-network determines if it is the ultimate destination of the path.
   a. If so, it checks if the destination endpoint identifier specified in the message includes a port index. In this case, it completes the optical path set-up using the port index. If the port index is not included, the address corresponds to a border router. In this case, the port through which the border router performed registration is used to complete the path set-up.

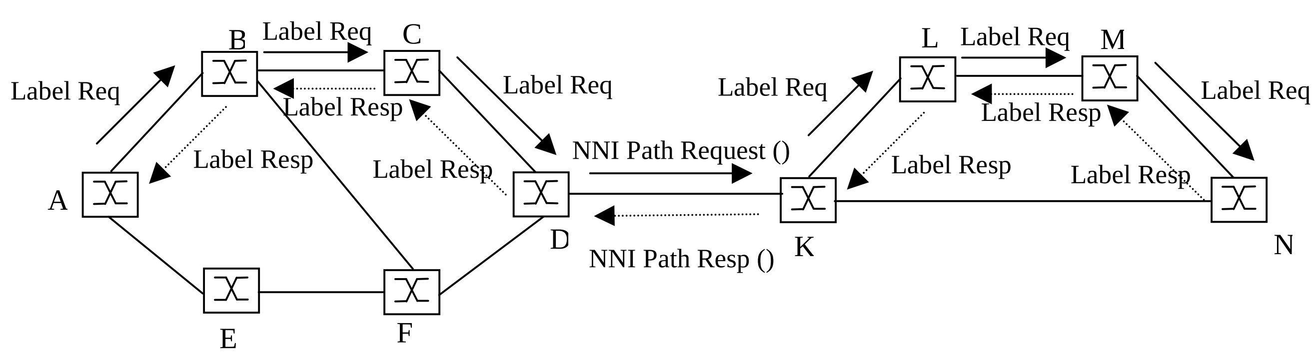


Figure 9: Provisioning Across Sub-Networks

b. If the OXC is not the ultimate destination, it determines the address of a border OXC in an adjacent sub-network that leads to the final destination. The path set-up is signaled to this OXC using NNI signaling. The next OXC then acts as the source for the path and steps 1-3 are followed.

6 Summary and Conclusions

This contribution considered several models for routing information exchange across the UNI and NNI. These models were classified into configuration-based, partial peering and full peering routing models. In summary, for UNI routing:

- Configuration-based: Requires optical end-point information to be configured in client systems. No routing protocol exchange across the UNI.
- Partial peering: Limited routing information is exchanged across the UNI. This information may be used for bootstrapping client specific routing exchange over the optical network.
- Full peering: Full routing information is exchanged across the UNI.

It was pointed out that not all of the UNI routing exchange models may be appropriate for all types of client networks. The descriptions in this contributions were mostly based on IP client networks. With other type of client networks, it is possible to carry non-IP addresses in routing protocols. The partial peering model can also accommodate other address types. For routing across the NNI, the full-peering model using BGP was described.

The routing models were described in abstract in this contribution. Further details of their operation are required. Also, the issue of how client devices determine which endpoints to establish an optical path is an issue. That was not considered in this contribution.

7 References

1. J. Moy, “OSPF Version 2,” RFC 1247, July, 1991.
2. Y. Rekhter and T. Li, “A Border Gateway Protocol 4 (BGP-4),” RFC 1771, March, 1995.
3. T. Bates, R. Chandra, D. Katz, and Y. Rekhter, “Multiprotocol Extensions for BGP-4,” RFC 2283, Feb., 1998.
4. E. Rosen and Y. Rekhter, “BGP/MPLS VPNs”, RFC 2547, March, 1999.

8 Author Information

Dimitris Pendarakis
Bala Rajagopalan
Debanjan Saha
Tellium Inc.
2 Crescent Place
P.O Box 901
Ocean Port, NJ 07757
Email: {dpendarakis, braja, dsaha}@tellium.com

****** This draft expires on May, 24, 2001 *******