ROUTING PROTOCOLS FOR OPTICAL NETWORKS

GHASSEN BEN BRAHIM* BILAL KHAN† ABDELLA BATTOU‡ MOHSEN GUIZANI §
GHULAM CHAUDHRY ¶

ABSTRACT

In wavelength division multiplexing (WDM) networks, communication between optical cross-connect (OXC) switches takes place along all-optical WDM channels which are commonly referred to as *lightpaths*. This paper (i) describes the central algorithmic problems whose solution an optical network routing protocol must facilitate, (ii) evaluates several candidate routing protocols that could be extended to operate in a WDM environment, and (iii) surveys currently ongoing efforts to standardize extensions to one of these, namely OSPF.

Keywords: Routing, WDM, Traffic Engineering, Extended OSPF.

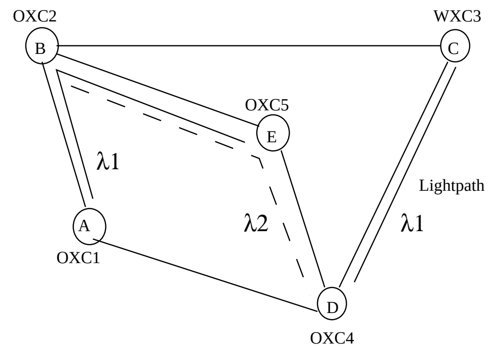

Figure 1: A WDM Routing Network with 5 OXCs

1 CENTRAL ALGORITHMIC PROBLEM

The central algorithmic problem in WDM routing is: Given a WDM network's physical topology (including the characteristics of its links) and a set of source-destination pairs, (i) *route assignment:* compute a route for a lightpath between each source-destination pair, and (ii) *wavelength assignment:* for each link traversed by this lightpath, determine the wavelength to be allocated for the lightpath on the given link. Together, these two assignment problems are often referred to as the *Routing and Wavelength Assignment (RWA)* problem. Typically, the RWA problem is solved by considering each of its two constituent subproblems in turn [20].

The RWA problem is complicated by the fact that an OXC switch may be (optionally) equipped with *wavelength conversion* hardware which permits lightpaths transient through the switch to enter and leave the switch on different wavelengths. Consider figure 1, which illustrates a WDM network consisting of five OXC switches interconnected by optical fiber links. If no switches are equipped with wavelength converters, then a lightpath in this network must always occupy the same wavelength on every fiber link it traverses. This restriction is commonly known as the *wavelength continuity constraint* for non-wavelength converting switches. An OXC switch that is equipped with wavelength conversion hardware is exempt from this constraint. The reader is referred to the book by Ramaswami and Sivarajan [18] for a comprehensive introduction to WDM network technologies.

A solution to an instance of the RWA problem must necessarily [1], [2] respect the following constraints: First, two lightpaths traveling on a given link must be assigned different wavelengths. Second, any lightpath that transits through an OXC switch that is not wavelength conversion capable, must use the same ingress and egress wavelengths. The central responsibility of the WDM routing protocol is to ensure that the information required for route assignment and wavelength assignment is maintained by the network in a scalable manner.

1.1 ROUTING ASSIGNMENT

There are three broad classes of strategies being used presently to address the routing assignment problem [20]. These strategies are referred to as *fixed routing, fixed alternate routing, and adaptive routing*.

Fixed routing is a simple technique which involves maintaining a fixed routing table at each candidate source node. The routing table consists of one entry for each candidate destination node, where the entry specifies the path from the source to the destination. Fixed routing is simple to implement, but is subject to unacceptably high blocking probabilities as wavelength availability on links becomes scarce, and when link failures occur.

In contrast, *fixed alternate routing* attempts to address the shortcomings of fixed routing by augmenting each entry in

*Electrical & Computer Engineering Department, University of Missouri Columbia.
†Advanced Engineering & Sciences ITT contracted to The Center for Computational Sciences, Naval Research Laboratory, Washington D.C.
‡Princeton Networks Inc, Greenbelt, Maryland.
§Computer Science Department, University of West Florida.
¶Electrical & Computer Engineering Department, University of Missouri Columbia/Kansas City.

the routing table to be a prioritized set of paths from source to destination, rather than just a single one. By doing so, fixed alternate routing is less sensitive to link failures and wavelength and thus offers lower connection blocking probabilities.

Finally, *adaptive routing* schemes attempt to select a path between a source-destination pair based on dynamically collected information concerning the network's state. This technique is much more resilient to link failures and less sensitive to wavelength scarcity, thus offers lower connection blocking probability than fixed adaptive routing approaches.

1.1.1 WAVELENGTH ASSIGNMENT

The second half of the RWA problem, wavelength assignment, is addressed algorithmically using strategies based either on heuristics, or on *graph coloring* algorithms. Presently considered heuristics include Random Wavelength Assignment, First-Fit, Least-Used (SPREAD), Most-Used (PACK). Min-product, Least Loaded, MAX-SUM, Relative Capacity Loss, Distributed Relative Capacity Loss, Wavelength Reservation, and Protecting Threshold. For a detailed comparison between the performance of these heuristics, the reader is referred to the review by Zang et al. in [20].

1.1.2 RWA AND TRAFFIC ENGINEERING

One popular extension of adaptive routing schemes are the so-called *least congested path* (LCP) schemes, which attempt to choose the path between source-destination pairs based on current traffic patterns in the network. By considering each wavelength on a link to be a fraction of the link's available bandwidth, LCP schemes are easily adapted to operate in a WDM environment. Unfortunately, LCP schemes introduce significant complexity, both in terms of the protocols that must operate to maintain information about network state, and the algorithms responsible for carrying out the route computation. Together, these concerns and their resolution are referred to as *Traffic Engineering (TE)*.

TE is responsible for mapping traffic flows to the managed physical network resources. Specifically, TE provides the ability to move traffic flows away from the shortest path and into less congested paths in order to satisfy the *Quality of Service (QoS)* requirements of higher layers. The main purpose of TE is to load balance the network traffic across various links, routers, and switches, thereby minimizing congestion and ensuring that the network resources are not over or under-utilized. Regrettably, many existing routing protocols do provide QoS and TE adequately to model WDM networks, and so require augmenting the protocol and packet specifications.

Major classes of TE attributes are described by Awduche et al. in [6]. These include: (1) *traffic parameter attributes* used to capture the characteristics of the traffic streams, (2) *generic path selection and maintenance attributes* specifying rules for selecting the route taken by an incoming traffic as well as the rules for maintenance of paths that have already been established, (3) *priority attributes* specifying relative importance of different traffic flows, (4) *preemption attributes* stating rules for preemption of resources, (5) *resilience attributes* specifying the behavior of certain traffic under fault conditions, (6) *policing attributes* defining the actions that should be taken by the underlying protocols when incoming traffic becomes non-compliant. In [11], Fedyk and Ghanwani present additional metrics and resource information for links tailored specifically for traffic engineering, including hop count metrics, bandwidth based metrics, delay based metrics, economic cost or expense based metrics, and administrative weight.

To summarize, the responsibility of an optical routing protocol is to facilitate lightpath establishment between OXC switches, i.e. it must provide access to sufficient information to solve the problems of Routing Assignment and Wavelength Assignment, while permitting TE information concerning the current network traffic and the QoS requirements of higher layers to be taken into consideration. In the next section, we present an overview of routing protocols, and evaluate specific candidates which might be extended to serve as a routing protocol for WDM networks.

2 CANDIDATE ADAPTIVE ROUTING PROTOCOLS FOR WDM NETWORKS

Routing protocols can be divided into two broad classes, *Interior Gateway Protocols (IGPs)*, and *Exterior Gateway Protocols (EGPs)*. IGPs are intended for use within an *Autonomous System*, i.e. a set of networks that are under a single or closely coordinated management organizations. Routing between different autonomous systems is maintained by the EGPs, of which the *Border Gateway protocols (BGPs)* are one example. While IGPs are designed to dynamically maintain information about network topology, produce optimal routes, and respond to changes quickly, EGPs emphasize stability and administrative control above route optimality. As the networking community moves towards further decentralization of control, EGPs are viewed as serving to protect one system of networks against errors or intentional misrepresentation by other systems.

Along a different axis, routing protocols can also be classified into: *distance vector routing* protocols which implement distributed algorithm to compute all-pairs of shortest paths, and *link state routing* protocols which propagate information about network topology to all routers.

In distance vector routing protocols, such as the *Routing Information Protocol (RIP)*, each router maintains the distance from itself to each possible destination, and the distance vector routing protocols relaxes these values down to their shortest value. In doing so, RIP is essentially a distributed implementation of the Bellman-Ford all-pairs shortest path algorithm.

In link state routing protocols, such as the *Open Shortest Path First (OSPF)* protocol, every router maintains a representation of the network's topology, and this representation is updated regularly on the basis of the information being exchanged between the network routers. In particular, each router is responsible for discovering its neighboring routers, and then advertising the identities of its adjacent neighbors and the "cost" to reach each. This information is advertised between routers by periodic exchange of link state packets. A link state protocol arms each router with a dynamic map of the network's topology; using this, the router may compute paths to any destination.

In the subsections that follow, we evaluate several candidate routing protocols and assess their suitability for operation in the WDM environment.

2.1 ROUTING INFORMATION PROTOCOL (RIP)

The Routing Information Protocol (see [17] [13] [12]) is a distance vector routing protocol for small networks. Each RIP-enabled router is classified as either active or passive: active routers advertise their routes to other routers; passive routers listen and update their routes based on the advertisements they receive but do not advertise any information themselves. Typically, routers run RIP in active mode, while hosts use passive mode. RIP is attractive because of its simplicity and elegance but suffers from several undesirable features, which we list below:

- *Network size limitations:* RIP was designed as a routing protocol for small networks of diameter $\leqslant 15$. The convergence time of RIP limits its applicability for larger networks.
- *Slow convergence:* The essential reasons behind the slow convergence are (1) the information that RIP needs to propagate is itself routed using the same routing tables that RIP strives to build, and (2) the accuracy of these routing tables increases only gradually because of the nature of distributed Bellman-Ford.
- *Count to infinity:* The failure of a link or node failure may result in two routers continually exchanging update vectors because each is claiming to have the right information. Many solutions (e.g. "Split Horizon", "Flush Updates", "Poison Reverse Update", etc.) have been proposed to address this issue. But count to infinity may still occur even when using these techniques.
- *Looping:* Incorrect routing information available in some nodes may result in forwarding loops within the network.
- *Restricted view of the network:* A RIP router does not have a view of the topology of the entire network, and so cannot detect problems such as looping or count to infinity.
- *Limitations on the metrics that can be used for shortest path computation:* RIP attempts to determine the shortest path by relaxing the count of intermediate hops between each source-destination pair. It is unclear how to extend RIP to support computation of the "best path" using multiple interdependent (particularly non-additive) link metrics, such as those required by TE.
- *Inability to support load balancing:* Occasionally the need arises to split traffic load between two parallel paths between two nodes, or to cache alternate paths for use in case of failures. RIP presently supports only the computation of the unique shortest path between nodes.

The undesirable characteristics described above make RIP a poor candidate to serve as the foundation for an optical routing protocol.

The next 3 protocols we consider are all link state routing protocols. By sequencing successive updates of link state information, these protocols sidestep the count-to-infinity problem exhibited by RIP. In addition, by enabling each switch to acquire a complete view of the network topology, they implicitly avoid RIP's routing loop problem.

2.2 INTERMEDIATE SYSTEM TO INTERMEDIATE SYSTEM (IS-IS)

In ISO terminology, a router is referred to as an *Intermediate System* (IS). The IS-IS protocol (see [17] [8] [14]) is a link state IGP that was originally developed for routing ISO/CLNP[1] packets. IS-IS has many desirable features, including:

- *Scalability to large networks:* IS-IS intra-domain routing is organized into a 2-levels of hierarchy, allowing large domains to be administratively subdivided.
- *Multi-path forwarding:* IS-IS uses a "shortest path first" algorithm to determine routes, and the new IS-IS protocol supports multi-path forwarding.
- *Genericity:* IS-IS can support multiple communication protocols and is therefore generic.

IS-IS does not flood any routing information across level 1 areas. Thus, internal IS-IS routers always send traffic destined to other level 1 areas to the nearest level 2 router Area Border Router (ABR). If there is more than one ABR, IS-IS

[1] International Organization for Standardization/Connectionless Network Protocol

may not pick the optimal route. In this manner, IS-IS trades off route optimality in exchange for lower control traffic, this allows IS-IS to scale to larger networks. Unfortunately, IS-IS is not an IETF standard and has only one RFC, which is not up to date with commercial implementations. As a consequence, IS-IS is no longer an open IGP routing protocol, and so is a risky choice for extension to the optical domain.

2.3 PRIVATE NETWORK TO PRIVATE NETWORK INTERFACE (PNNI)

PNNI is a link state routing protocol standardized by the ATM Forum [5]. Intended for use in ATM networks, PNNI offers many features that would make it a good starting point for developing a WDM routing protocol. These include:

* *Scalability to large networks:* PNNI implements hierarchic partitioning of networks into Peer Groups (upto 104 levels), and supports aggregation of information from within each Peer Group. Together, hierarchy and aggregation provide a flexible tradeoff curve between route optimality and protocol traffic. The PNNI hierarchy is configured automatically based on switch addresses using dynamic election protocols.

* *Support for traffic engineering and QoS:* PNNI distributes link state information in PNNI Topology State Elements (PTSEs); these come in many flavors, including Horizontal Link InfoGroup (HLinkIG) packets for advertising link characteristics, and Nodal InfoGroup (NodalIG) packets for advertising switch state information. HLinkIG packets include resource availabilities for a link on a per service-class basis; it would be quite easy to extend this description to include wavelength information. Similarly, it would be fairly straightforward to extend the NodalIG to include information about the wavelngth conversion capabilities of a switch.

* *Standardization:* PNNI was approved by the ATM Forum and has since been adopted by most ATM vendors and shown to work in the field.

In terms of the technical advantages outlined above, PNNI is a strong candidate to serve as the foundation for an optical routing protocol. PNNI supports policy-based QoS aware routing by by permitting the source switch (and each border switch) to compute a aprtial route through its peer-group twards the destination. This makes it difficult for PNNI to support multi-path forwarding. PNNI is, relatively a complex protocol, and perhaps as a consequence, the optical switch industry has turned away from considering PNNI-based WDM routing solutions. Interoperability with existing optical switch technology makes PNNI a problematic choice.

2.4 OPEN SHORTEST PATH FIRST (OSPF)

Open Shortest Path First [16] [13] [14] is a link-state IGP that distributes routing information between routers within a single autonomous system. OSPF has many desirable features, including:

* *Easily extended for TE and wavelength information.* OSPF distributes link state information in Link State Advertisement (LSA) packets. In the RFC [10], Coltun proposes Opaque LSAs to extended the OSPF protocol. Opaque LSAs have a standard LSA header followed by an application specific information, which may be used directly by OSPF or by other applications. Interoperability was acheived by introducing 3 new Link State (LS) types, each with fixed flooding regimes: (i) LS type 9 packets are not flooded beyond the local subnetwork, (ii) LS type 10 packets are not flooded beyond the borders of their associated area, and (iii) LS type 11 packets are flooded throughout the Autonomous System. Routers that are not opaque capable may receive Opaque LSAs because they cohabit broadcast networks with opaque capable routers; in this case the LSAs are discarded, because they have unknown LS types. Opaque LSAs make it possible to extend OSPF to incorporate new domain-specific TE information, as needed; a lot of effort has been expended to specify the exact format of these packets for TE in the IP domain.
* *Scalability:* OSPF supports 2-levels of hierarchy and groups routers into areas. Because OSPF advertises a summary of the information desribing the areas inside the Autonomous System, the protocol is more likely (compared to IS-IS) to compute an optimal path when routing between different areas. OSPF converges reasonably quickly even for large networks.
* *Simplicity and Acceptance:* OSPF is a relatively simple protocol (compared to PNNI), and enjoys widespread commercial adoption.
* *Standardization:* OSPF is a full IETF Standard (unlike IS-IS).

3 OSPF-BASED INITIATIVES FOR WDM ROUTING

Currently, most of the routing protocols being proposed for WDM are based on OSPF. We consider two of the most promising proposed extensions now.

3.1 THE QOS EXTENSION TO OSPF

In the RFC [4], Apostolopoulos et al. propose an extension to OSPF to support QoS. In the same RFC, the authors propose that the Type of Service (TOS) field within the protocol packet options field should be used to specify whether the originating router is QoS capable. The bandwidth and delay of each link will be encoded into the metric field of the LSA packet by QoS capable routers. By using three bits for the exponent part and the remaining thirteen bits for the mantissa, the scheme can be used to represent bandwidths ranging into Gbits/second.

The main advantage of this proposal is that it does not require the addition of any additional LSAs. This makes it simple to integrate QoS-capable routers running the extended protocol with routers running older versions of OSPF (i.e. non QoS-capable routers). However, the proposed scheme suffers from the drawback of not being able to advertise additional information such as a complete list of available wavelengths for routing inside WDM networks.

3.2 OSPF EXTENSION USING OPAQUE LSAS

Here we give a brief outline of ongoing efforts of proposals using Opaque LSAs to extend OSPF for routing in the optical domain.

Recent internet drafts by Chaudhuri et al. [9] and Basak et al. [7] present requirements for optical bandwidth management and restoration in a dynamically reconfigurable network. Both papers expand on the list of optical network characteristics that must be maintained in the OXC routing database. The information is classified into two categories: First there is the information that must be advertised using OSPF, e.g. the total number of active channels, the number of allocated channels, the number of preemptable channels, the risk groups throughout the network, etc. The second category consists of information that is kept locally and is not advertised for protocol scalability reasons. Examples of such information are the available capacity, the preemptable capacity, the association between fibers and wavelengths, etc. The proposals choose to put information about the wavelength availabilities in the second category, arguing that available wavelength changes occur frequently, so advertising the changes in the available wavelength does not yield a performance increase proportionate to the costs in terms of control traffic.

In the internet draft [15], Kompella et al. present an extension to OSPF for optical routing. In particular, they specify some of the parameters needed to be advertised by the Opaque OSPF LSAs. Like the author of [9], [7], Kompella et al propose that OSPF should not advertise any information pertaining to wavelength availability, and postpone wavelength assignment to the time when lightpaths are signaled.

In the internet draft[19], Wang et al. propose to use the OSPF protocols to advertise both the number of available wavelength per fiber and the available bandwidth. To address the objection that the advertisement of the available bandwidth is impractical because of rapid changes in network link usage, the authors also introduce the notion of “significant change”. In their extension, OSPF readvertises only if the available bandwidth changes by a factor which is higher than a certain tunable threshold.

3.2.1 UPDATE LSA FLOODING POLICIES

Some of the TE and QoS parameters change very frequently, raising the issue of when to advertise the changes of the network characteristics throughout the whole network. The original OSPF standard mandates a variety of tunable parameters controlling the flooding of LSAs, including the “MinLSInterval” which specifies the time between any two consecutive LSA originations, and the “MinLSArrival” which limits the frequency of accepting a newer instances of LSAs.

In [3], Apostolopoulos et al. present other policies dealing with the issue of when a router should flood a new LSA to advertise changes in its link metric. Some of the proposed policies include:

- *Threshold based policies* that trigger updates when the difference between the previously flooded and the current value of available link bandwidth is larger than a configurable threshold.
- Class based policies which partition the capacity of a link into a number of classes and re-advertise when a class boundary is crossed.
- *Timer based policies* which generates update at fixed, or enforce a minimum spacing between two consecutive updates. time intervals.

4 CONCLUSION

We have presented several candidate IGPs and evaluated their potential of each to serve as a foundation for a WDM routing protocol. In addition to the extensions needed to support RWA, we also considered the facility with which TE and QoS support could be added. RIP, IS-IS, PNNI were shown that they are not ideal for such extension. Being an IETF standard, with widewspread acceptance, scalable design, and support for multi-path forwarding, OSPF is a promising starting point for the development of a WDM routing protocol. We presented a survey of current efforts to standardize optical extensions to OSPF.

References

[1] M. Ali, B. Ramamurthy, and J. Deogun. Routing and Wavelength Assignment (RWA) with power considera-

tions in Optical Networks. *IEEE Globecom '99 Symposium on High-Speed Networks*, December, 1999.

[2] J. Anderson, J. Manchester, A. Rordiguez-Moral, and M. Veeraraghvan. Protocols and Architectures for IP Optical Networking. Bell Labs Technical Jounal, January, 1999.

[3] G. Apostolopoulos, R. Guerin, S. Kamat, and S. Tripathi. Quality of Service Based Routing: A performance Perspective. *Proceedings of SIGCOMM*, September, 1998.

[4] G. Apostolopoulos, D. Williams, S. Kamat, A. O. R. Guerin, and T. Przygienda. QoS Routing Mechanisms and OSPF Extensions. Request for Comments No. 2676, August, 1999.

[5] ATM-Forum. *Private Network-Network Interface Specification, Version 1.0*. ATM-Forum, 1996.

[6] D. Awduche, J. Malcolm, J. Agogbua, and J. M. M. O'Dell. Requirements for Traffic Engineering over MPLS. Request for Comments No. 2702, September, 1999.

[7] D. Basak, D. Awduche, J. Drake, and Y. Rekhter. Multi-Protocol Lambda Switching: Issues in Combining MPLS Traffic Engineering Control with Optical Corssconnects. Internet Draft, January, 2000.

[8] R. Callon. Use of OSI IS-IS for routing in TCP/IP and Dual Environments. Request for Comments No. 1195, December, 1990.

[9] S. Chaudhuri, G. Hjalmtysson, and J. Yates. Control of Lightpaths in an Optical Network. Internet Draft, February, 2000.

[10] R. Coltun. The OSPF Opaque LSA Option. Request for Comments No. 2370, July, 1998.

[11] D. Fedyk and A. Ghanwani. Metrics and Resource Classes for Traffic Engineering. Internet Draft, October, 1999.

[12] C. Hedrick. Routing Information Protocol. Request for Comments No. 1058, June, 1988.

[13] C. Huitema. *Routing in The Internet*. Prentice Hall.

[14] IETF. The MPLS Operations Mailing List. IETF's MPLS Working Group.

[15] K. Kompella, Y. Rekhter, D. Awduche, J. L. G. Hjalmtysson, S. Okamoto, D. Basak, G. Bernstein, J. Drake, N. Margalit, and E. Stern. Extensions to IS-IS/OSPF and RSVP in support of MPL(ambda)S. Internet Draft, October, 1999.

[16] J. Moy. OSPF Version 2. Request for Comments No. 2178, July, 1997.

[17] R. Perlman. *Interconnections Bridges and Routers*. Addison Wesley, June, 1993.

[18] R. Ramaswami and K. Sivarajan. *Optical Networks: A Practical Perspective*. Morgan Kaufmann Publishers, Inc, 1998.

[19] G. Wang, D. Fdyk, V. Sharma, K. Owens, G. Ash, M. Krishnaswamy, Y. Cao, M. Girish, H. Ruck, S. Bernstein, P. Nquyen, S. Ahluwalia, L. Wang, A. Doria, and H. Hummel. Extensions to OSPF/IS-IS for Optical Routing. Internet Draft, March, 2000.

[20] H. Zang, J. Jue, and B. Mukherjee. A Review of Routing and Wavelength Assignment Approaches for Wavelength-Routed Optical WDM Network. *Optical Networks Magazine*, August, 1999.