

11-14-2006

On multicasting in wavelength-routing mesh networks

Ashraf Hamad
Iowa State University

Tao Wu
Microsoft Corporation

Ahmed Kamal
Iowa State University, kamal@iastate.edu

Arun K. Somani
Iowa State University, arun@iastate.edu

Follow this and additional works at: https://lib.dr.iastate.edu/ece_pubs



Part of the [Digital Communications and Networking Commons](#)

The complete bibliographic information for this item can be found at https://lib.dr.iastate.edu/ece_pubs/294. For information on how to cite this item, please visit <http://lib.dr.iastate.edu/howtocite.html>.

This Article is brought to you for free and open access by the Electrical and Computer Engineering at Iowa State University Digital Repository. It has been accepted for inclusion in Electrical and Computer Engineering Publications by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

On multicasting in wavelength-routing mesh networks

Abstract

This paper considers multicasting on wavelength-routing mesh optical networks. Although multicasting has been studied extensively in different network environments, multicasting in this environment is different, and more involved. The paper discusses the challenges of multicast support in optical wavelength routing networks, and reports on the advances made so far in this venue. The paper introduces a classification and a comparison of such techniques, and a study of their advantages and disadvantages.

Keywords

Optical Networks, Wavelength Division Multiplexing (WDM), Wavelength-Routing Networks, All-Optical Multicasting (AOM), All-Electronic Multicasting (AEM), Hybrid Multicasting (HM)

Disciplines

Digital Communications and Networking

Comments

This is a manuscript of an article published as Hamad, Ashraf, Tao Wu, Ahmed E. Kamal, and Arun K. Somani. "On multicasting in wavelength-routing mesh networks." *Computer Networks* 50, no. 16 (2006): 3105-3164. DOI: [10.1016/j.comnet.2005.12.012](https://doi.org/10.1016/j.comnet.2005.12.012). Posted with permission.

Creative Commons License



This work is licensed under a [Creative Commons Attribution-Noncommercial-No Derivative Works 4.0 License](https://creativecommons.org/licenses/by-nc-nd/4.0/).

On Multicasting in Wavelength-Routing Mesh Networks*

Ashraf Hamad

Tao Wu

Ahmed Kamal †

Arun Somani

Department of Electrical and Computer Engineering
Iowa State University
Ames, IA 50011 U.S.A.
{ashraf,wutao,kamal,arun}@iastate.edu

Abstract

This paper considers multicasting on wavelength-routing mesh optical networks. Although multicasting has been studied extensively in different network environments, multicasting in this environment is different, and more involved. The paper discusses the challenges of multicast support in optical wavelength routing networks, and reports on the advances made so far in this venue. The paper introduces a classification and a comparison of such techniques, and a study of their advantages and disadvantages.

Key words: Optical Networks, Wavelength Division Multiplexing (WDM), Wavelength-Routing Networks, All-Optical Multicasting (AOM), All-Electronic Multicasting (AEM), Hybrid Multicasting (HM).

I Introduction

The introduction of Wavelength Division Multiplexing (WDM) [1, 2] is considered an impetus to a new field of research. As such, the last decade has witnessed accelerated research towards the realization of the optical network, which is capable of providing different classes of service, to different applications, with an abundance of bandwidth. Such research efforts, were accompanied by rapid developments in the field, in terms of the introduction of new devices, and the advancement of the enabling technologies (for example, see [3], [4] and [5] for the state-of-the-art in optical technology). Although the number of commercially available channels per fiber is currently in the hundreds, thousands of channels per fiber are not uncommon in research laboratories [6].

Optical networks have gone through two generations of development. In the first generation, optical networks were similar in principle to their copper-based counterparts, but allow concurrent

*This research was supported in part by the National Science Foundation under grant ANI-0087746.

†Corresponding author: Telephone: (515) 294-3580; Fax: (515) 294-1152, E-mail: kamal@iastate.edu.

transmissions on different wavelengths. Several protocols were developed for packet switching on such networks using single and multihop delivery strategies [7, 8]. The design and operation of second generation optical networks, is entirely different. The basic principle was to exploit the inherent characteristics of optics, and realize end-to-end communication in a true, and transparent all optical domain. Second generation networks establish lightpaths which are end-to-end all optical channels. Such lightpaths are used for circuit-switched service only, and packet switched service must be carried on such channels. Connectivity is achieved by implementing a virtual topology onto the physical topology, using lightpaths. Second generation optical networks are usually referred to as *Wavelength Routing Networks*.

Unlike first generation networks, whose design is very similar to copper-based network design, the design of second generation optical networks takes a completely different approach. Network design must go through a number of phases, and may have to also undergo a number of iterations. These phases are shown in Figure 2, and will be described in detail in Section II.

One class of service which can make use of the huge amount of bandwidth provided by optical networks is the class of multicast service. It is expected that a sizable portion of the traffic in future high performance networks will be of this type [9]. High bandwidth applications of the multicast type include video distribution, network news distribution, database replication and search queries, storage area networks updates and backups, multi-party videoconferences, computer-supported collaborative work, etc. The topic of multicast service support in WDM-based networks has received significant interest in the last few years. The problem of multicast support in first generation optical networks boils down to the design of medium access control protocols capable of providing such service. As such, packet replication in this case is implemented in the electronic domain. Several protocols of this type have been developed, and a comprehensive review of such protocols is provided by the authors in [10].

Multicast service support in second generation optical networks assumes a completely different nature, and faces unique problems. Packet replication cannot be carried out in the electronic domain, since this means that either different lightpaths must be established between the source and all the destinations, or that the lightpaths must be terminated at the branching points. Both approaches have their own shortcomings. In the first approach, bandwidth will be unnecessarily wasted, while in the second approach signal transparency will be compromised due to the multiple terminations at the splitting points. Therefore, it is more advantageous to implement packet replication in the optical domain. Although this seems like an obvious solution, this introduces a set of design and implementation problems. Signal splitting entails power loss, and the split signal must be amplified more often. This will not only increase the cost of implementation, but will also

degrade the signal to noise ratio [4]. The splitting ratio should also be adjustable to accommodate different multicast trees [11]. Moreover, the construction of the multicast tree should take into account the power loss, and signal degradation. There are several other related problems which must be addressed in the course of implementing multicasting in optical networks.

This paper is concerned with multicasting in second generation optical networks. Although interest in this area is relatively recent, some advances have already been made. Reference [79] provided an introduction to this important area, however, this article is more comprehensive. In this context, we provide an account of the challenges faced in designing and implementing such networks, as well as a comprehensive survey and comparison of the most significant work which has appeared in the literature. The paper also addresses some of the open problems in this area.

The paper is organized as follows. We start by briefly discussing the planning, design and operation process of second generation optical networks in Section II. The significance of this discussion is twofold. First, it helps one realize the complexity involved in designing and operating those types of networks. It also provides the basis to our treatment of the challenges to multicasting. In Section III we provide a brief overview of the multicasting techniques used in traditional IP-based networks, and in which multicasting is done in the electronic domain. We then address the challenges faced in providing multicasting service in all-optical second generation networks in Section IV. Section V considers some of the technical problems in designing optical networks for multicasting. The design of optical elements for the support of multicasting is considered. Also, the limitations imposed by these optical components, power budget and network diameter on the multicast group size are outlined. In addition, network dimensioning and provisioning, as well as network operation for multicasting are discussed. Section VI surveys and compares the multicasting techniques introduced in the literature. Finally, Section VII concludes the paper with some remarks and elaborates on some of the open problems in this area. In order to develop an understanding of the various multicasting aspects and techniques in wavelength-routing WDM networks, our treatment in this paper is enriched by several examples that are all based on a common network topology, namely the popular NSF Network (NSFNET) backbone topology which is shown in Figure 1. In these example, the multicast session is represented as an ordered-pair (s,D) , where s represents the source node and D takes the form of a set of destinations.

II Network Planning and Operation

The goal of the network planning process in optical networks, and in networks at large, is to design the network, including the determination of the resources and their capacities [12]. The objective of

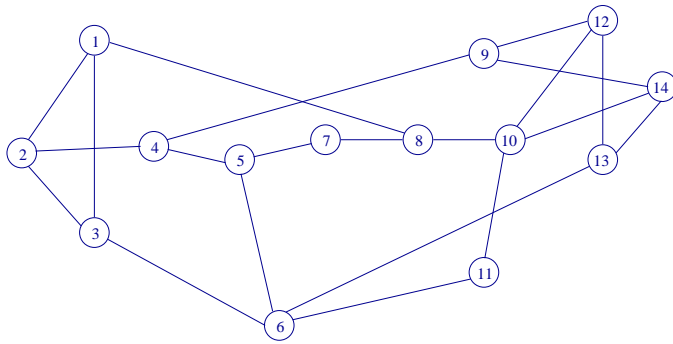


Figure 1: NSF Network (NSFNET)

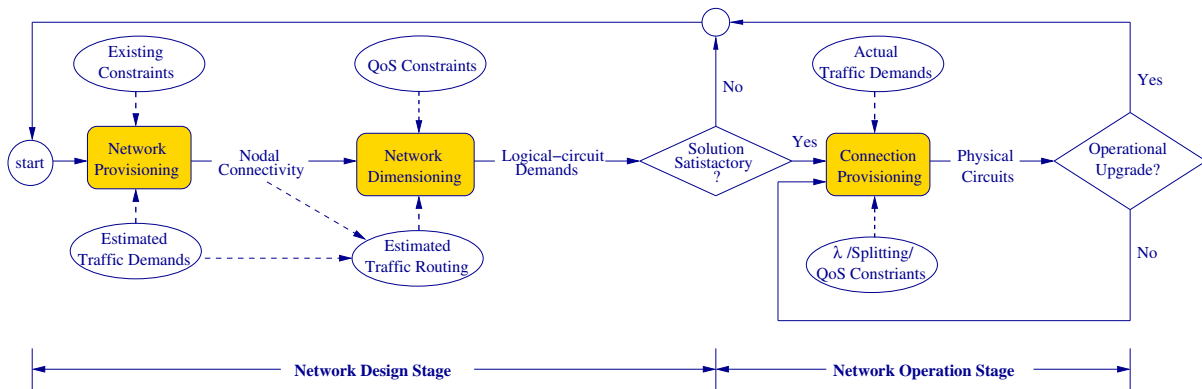


Figure 2: Optical Network Planning Stages under Multicast Communication

this design process is to be able to accommodate a certain set of traffic conditions, while satisfying certain levels of quality of service (QoS), and incurring minimal cost. Network planning is a very complex, and imprecise process. The imprecision is due to the lack of complete knowledge of the traffic requirements, and their future evolution. While this contributes to the complexity of the problem, the complexity usually stems from the fact that an optimal determination, and allocation of resources is usually an NP-Hard problem.

Network operation, on the other hand, has to deal with the allocation of existing resources to real traffic demands, which may deviate from the projected ones. The objective in this case is to minimize the loss in the QoS, and in the carried traffic, given the available network resources.

Network planning and operation can be viewed as consisting of three phases, which have to be executed iteratively, namely, Network Provisioning, Network Dimensioning, and Connection Provisioning. The relation, and interaction between the three phases is shown in Figure 2.

Network Provisioning Stage

The Network Provisioning and Network Dimensioning phases are performed during the network design and they are closely related, but we would like to make a distinction between them. Network Provisioning is the first task faced by the network designer/planner. In this stage, the planner

must carry out a careful assessment of the requirements of the users, in term of traffic demands, their characteristics, their duration, and their required QoS. Since this information will be used to plan and implement a network that will be used by different users and for several years, not only the current traffic demands should be taken into account, but also the projected traffic demands. Usually, such traffic demands are somehow overestimated in order to extend the lifetime of the network. In this phase, the network topology is determined. This includes:

- The allocation (which includes determining the locations and the quantities) of the various network resources, e.g., the optical fibers, the wavelength cross connects, the amplifiers, the wavelength converters, and the light splitters¹.
- The determination of the virtual topology which must be provisioned to serve the estimated traffic demands, given the network resources. This requires the determination of the different lightpaths/light-trees.

This phase is subject to a number of *existing* constraints that include peer-nodes connectivity constraints, geographical distribution of the nodes, the existence of other networks that may be used to carry part of the traffic, or at least participate in switching the traffic.

The above is usually formulated as a optimization problem, e.g., a constrained resource allocation problem, and the solution of which can be obtained using conventional optimization techniques. The objective function in this case, is usually a function of the cost of the components, which needs to be minimized.

The output of the Network Provisioning stage takes the form of connectivity matrix that specifies the connection pattern between the various network components along with their optimal (or near optimal) physical locations in the network. This output in conjunction with the estimated traffic demands and their QoS requirements form the input to the next phase of the network design stage, namely, the Network Dimensioning phase.

Network Dimensioning Stage

The main focus of the Network Dimensioning phase is to determine the optimal dimension of the various network resources such that QoS requirements are met, given the projected traffic demands, a specific routing strategy, and the resources determined during the Network Provisioning stage. In particular, the following must be determined:

- The link capacity, in terms of the number of fibers per link,

¹For the case of networks provisioned for multicasting.

- The fiber capacity, in terms of the number of wavelength channels per fiber, and
- The channel capacity, in terms of the transmission rate.

Up to this point, the network design in terms of its topology and resources capacities are computed. Such a design is examined in order to validate its effectiveness before physical implementation. A Design Effectiveness Metric (DEM) is used for this purpose, and it can be defined in a number of ways. The most widely used definition is the ratio of the number of accommodated (accepted) calls to the total cost of the designed network, or call per dollar. If the DEM of a certain solution for the network design stage fails to satisfy a specific threshold measure, the network design is revised in an iterative manner until the most appropriate network design candidate is found.

Connection Provisioning Stage

During network operation, traffic demands may, or may not meet the projected demands. This requires operating the network, with the current resources, the current topology, and the currently accepted connections, in a manner that maximizes the number of accepted connections, while satisfying the required QoS measures. This is known as Connection Provisioning. In particular, Connection Provisioning includes all of the followings:

- Route determination,
- Wavelength assignment and resource allocation along the computed routes,
- Call establishment,
- Error recovery due to nodes and/or links failures,
- Traffic rerouting and wavelength reassignment in order to accommodate more traffic sessions(especially in the case of dynamic routing and wavelength assignment), and
- Call termination and the deallocation of their associated network resources.

As such, the Connection Provisioning function is exercised before call establishment, during the lifetime of calls, and after the termination of calls.

It is to be noted that the fundamental objective that network operator attempts to achieve is the maximization of the effective utilization of the network, while guaranteeing the different connections QoS requirements. This directly translates into a greater revenue to the network operator. Network utilization can be maximized by maximizing (minimizing) the ratio of the actual accepted (rejected) calls with respect to the total number of arriving calls. These measures are usually referred to as the acceptance, or blocking probabilities, respectively.

III Multicasting

In multicasting, the network replicates data as it is being sent from a source to the different receivers. This design feature reduces the amount of bandwidth used for one-to-many communications. A tree consists of a root, that is the source node, branches, segments of the tree that connects a parent node with child nodes, receivers, a set of nodes that join a group, and leaf router, that is a node or router at the edge of the tree that connects a receiver to the tree. A multicast tree is a non-cyclic connected graph and is used by a parent node (source) to replicate data to be sent to one or more children.

The current mode of multicasting used in IP networks, known as the Deering-Cheriton model [13], is based on a receiver oriented architecture. The basis for group membership and the corresponding construction or deconstruction of the tree (and/or parts thereof) is based on the actions of individual receivers. Hence, sources never enumerate the receivers of a group and only send data to a single group address. This receiver driven approach allows multicasting or one-to-many communication to scale to a vast numbers of receivers. Control signaling in IP multicasting is separated between that of host-to-router and router-to-router. The Internet Group Management Protocol (IGMP) is used to support host-to-router communication [14]. Specifically, hosts use this protocol to indicate that a receiver (application) is joining or leaving a group. This action then triggers a subsequent grafting or pruning of branches on a tree by the multicast routing protocol.

Currently, there are three types of IP multicast routing protocols responsible for constructing three different types of trees, namely, Source-trees, Shared-trees, and Hybrid trees.

III.1 Source Trees

Source tree protocols are based on a data driven scheme. It involves broadcast, reverse path distribution, and explicit pruning to build a delivery tree. The scheme ensures that data is only forwarded on interfaces that have downstream group members. Initially, the distribution method was based on Reverse Path Forwarding (RPF) and was used as a means of preventing loops as information was broadcast through the network [15]. This can be accomplished by only allowing routers to forward traffic through its other interfaces if the input interface used to receive the multicast data is also the output interface when the source address is a unicast destination. This action forms a spanning tree from the source to all possible receivers, and it removes the need to retain per data packet state in order to prevent loops. The root of the spanning tree is rooted at the first-hop router attached to the source-host of the multicast group. The initial spanning tree used to reach all possible hosts thus evolves into a source-based tree with the minimal set of

branches used to reach all group members.

In order to distinguish one source tree from another, routers must retain source and group states for each constructed tree. This state is either in the form of forwarding state indicating that the node is a member of the tree, or prune state indicating that the node is not a member of the tree and should not receive nor forward multicast traffic for the specific group. The source-tree protocols are best suited for dense topologies in which the receiver sets are close to each other. The benefit of a source-based tree is that its branches exhibit minimal delay between the source and the receivers of the group. The downside is the amount of state information that needs to be maintained. Several algorithms have been developed based on the above concept.

The Distance Vector Multicast Routing Protocol (DVMRP): The DVMRP algorithm [16] was the first IP multicast protocol to follow the above multicast model and was used in the Multicast Backbone (Mbone) [17]. DVMRP forwards initial data on a truncated broadcast tree, which then gets pruned to form the multicast tree. The truncation is accomplished by its routing protocol, which forms child parent relationships before any data is sent. State is maintained in DVMRP for all potential sources and all prunes that have been received. Several modifications have been made to DVMRP to cope with the relatively large and ever increasing set of networks.

Protocol Independent Multicast (PIM): The PIM algorithm is a tree construction protocol that operates independently over any unicast routing protocol. It can be divided into two perspectives: sparse mode (PIM-SM) and dense mode (PIM-DM) [18, 19], corresponding to receiver sets that are relatively dense (and constructs source-based tree) and sparse (and build shared trees), respectively. PIM-DM is very similar to DVMRP. It operates over any unicast routing protocol, does not distribute or maintain a separate routing table of its own, and uses the routing table constructed by the underlying unicast routing protocol to determine if ingress multicast traffic conforms to RPF checks. PIM-SM is a soft state multicast protocol that is designed to operate over any unicast routing protocol. Explicit join messages are periodically sent from leaf routers towards the core, which installs state on a hop-by-hop basis along the path. If a 'better' path is available, the join messages are rerouted and a new state is installed, while the old state on the old path is timed out and removed. PIM-SM also uses a threshold heuristic to determine if an additional source tree is to be constructed from a source sending data to the group. If the threshold is surpassed, then the source tree is constructed.

Multicast Open Shortest Path First (OSPF): The OSPF algorithm is a link state intra-domain routing protocol that operates at two levels: 1) within an Area (an arbitrary collection of

subnets), and 2) among Areas of a domain. Each router floods Link State Advertisements (LSAs) to the other routers that are peers at the same level. From these advertisements, routers are able to construct their own routing tables using a Dijkstra shortest path algorithm [20]. MOSPF specifies extensions that allow LSA messages to contain information regarding routers that have neighbor hosts that are members of a multicast group. When a new receiver joins a group in a subnet, and no other receivers exist, then the designated router advertises this information to its peers via flooding. MOSPF builds source-based trees. When a new source sends its initial data to the group, downstream routers construct a shortest path tree from the perspective of each downstream node to the receivers (actually, subnets) of the group. This calculation identifies the downstream nodes and subsequently forwards data to those neighbor routers. The source address is used to distinguish one tree from other trees for the group.

III.2 Shared Trees

The second option is based on the concept of shared trees that involves the use of a single tree for all sources of the group. Specifically, a shared root, or core is used as a point of reference from which to graft branches onto the tree. Trees of this type are known as core based trees, and the maximum delay bound of a core-based tree is twice that of a shortest path tree.

Several extensions have been introduced to address the need for dynamic group membership and tree maintenance due to connectivity failure. In addition, the orientation of tree construction in these new extensions has been shifted from a source perspective to a receiver-initiated perspective. In using a receiver-initiated approach, shared trees are built with an explicit joining mechanism. Within the context of today's protocols, explicit unicast join messages are triggered by IGMP-join messages and are sent from the leaf router to the shared core. This action results in the installation of state and grafting a branch, along the path of the join message. This leaf router then sends the explicit join on a hop-by-hop basis along a path to the shared core installing state along each node of the path.

Both the Core Based Tree (CBT) version 1 and the PIM-sm, are used to build shared trees. They share a common feature of operating independently over any unicast routing protocol. It is also implicitly understood that both CBT and PIM construct trees based on the shortest hop metric, since most routing protocols use this metric to calculate destination reachability. In addition, both protocols install state on a hop-by-hop basis from the direction of the leaf router towards the shared root. This is of particular interest since the best' path from the leaf to the root may not be the same from the opposite direction. On the other hand, a core-Based Tree contains a number of distinctive features representing several design choices that separates it from the subsequent

PIM-sm protocol. One of the principle differences involves CBTs construction of bi-directional trees, which allows it to support any-point distribution. The fundamental rule in its distribution is that multicast data is sent out to all the interfaces that have state for the group except the interface that the data was received on. It also supports the notion of a primary and a set of secondary cores. This allows the protocol to support a measure of fault tolerance. The Ordered CBT (OCBT) is an augmented version of CBT which supports a more robust set of cores for a group. In addition, OCBT incorporates extensions to construct inter-domain branches within the architecture of the Hierarchical IP multicast routing scheme protocol (HIP). OCBT also fixes a problem in CBT in keeping the tree intact in case of a failure. The Simple Multicast Protocol is a derivative of CBT and PIM and has been developed with the goal of supporting bi-directional trees and obviating the need for a multicast address allocation mechanism.

III.3 Hybrid Trees

Two primary concerns associated with shared tree protocols are core placement and traffic concentrations. Most of todays shared tree protocols make a priori selections of which node is to be selected as the shared root for a multicast group. This selection is made before any group members join the group and thus is orthogonal to the topological location of receivers. If the selected node/core is close to the receiver set, then the end-to-end delay between source and receivers will most likely be reasonable in comparison to a tree placed at the source of the multicast data. However, the core could be selected at a location far away from the rest of the receiver set, which can adversely impact the end-to-end delay. The concern regarding traffic concentrations arises in cases where several sources are sending data to the same group. Using just a single shared tree for data distribution, all on-tree nodes are then candidates for experiencing congestion. The severity of the congestion is dependent on the bandwidth of a link, the cumulative traffic from group sources, and other flows transiting the on-tree node.

Hybrid tree protocols attempt to address these concerns by constructing a single shared tree for relatively low amounts of group data, and also construct source trees for high rates of data transmission from a given source. Hybrid tree protocols produce a combination of a shared tree and source tree(s) for the same group. Some of the source tree branches in a hybrid tree have disjoint paths from the shared tree. Also, the leaf portions of the shared tree share the same links as the individual source trees, but segments that are closer to the shared root do not have the same amount of state or traffic loads for all sources. Hybrid trees may see oscillation in tree construction/deconstruction, corresponding to fluctuations in state storage and signaling overhead due to the decision the algorithm needs to make regarding partition between shared and source

tree portions.

BGMP, with its extensions to BGP, is viewed as an exterior multicast protocol and operates amongst domains. The reason for this distinction is primarily associated with the ability of administrators to apply inter-domain policies in constructing trees with BGMP. It constructs inter-domain multicast trees using the M-IGP within a domain as well as the multicast address allocation suite of protocols being developed under the Multicast-Address Allocation (MALLOC) working group of the IETF. When the M-IGP establishes a branch that spans the domain and reaches a border router, the border router extends the branch on an inter-domain basis. The use of a multicast address allocation mechanism, and its assignment of addresses to specific domains, addresses the issue of core/group mappings. The MALLOC suite of protocols dynamically assigns blocks of addresses to domains, which in turn can subdivide the blocks into more granular blocks (e.g., an ISP allocating sub-blocks to its customer stub domains).

IV Challenges of Multicast Communication in Wavelength-Routing Networks

Multicast communication support in single channel networks, e.g., IP networks, presents a number of challenges, and has thus been the subject of extensive research (see Section III). Nonetheless, it is even more challenging in the multi-channel wavelength-routing networks. Wavelength-routing networks have a number of characteristics which contribute to these difficulties. Such characteristics include the mode of operation of such networks, which requires connection provisioning, and power constraints of optical networks, which are exacerbated by signal splitting. In addition, hardware requirements, in terms of splitters and converters, which may not be deployed at all nodes can also impose a number of limitations. This section addresses these challenges based on their sources, and discusses how they impact the implementation of multicasting in wavelength-routing networks.

IV.1 Challenges Due to High-Transmission Rates

With the WDM technology, each channel supports transmission rates on the order of 10 Gbit/sec to 40 Gbits/sec. The delay-bandwidth product in this case is very large, and increases further with an increasing network diameter. This feature adversely impacts modes of operation which depend on feedback from the network. For example, on-demand routing strategies which probe the network resources when a connection is to be established will result in bandwidth wastage, as well as an increased connection latency. In addition, optimal provisioning, which requires increased

computation, and node coordination, can also result in far from optimal resource utilization.

IV.2 Challenges Due to the Characteristics of Wavelength-Routing Networks

Wavelength-routing networks have a number of special characteristics that affect multicasting support in wavelength-routing networks. These limitations are basically due to the use of multiple channels, the circuit-switched connection mode, as well as the employment of optical components. In the remainder of this subsection, a brief description of each of these limitations is presented.

IV.2.i Multi-Channels Environment Limitations

First, the use of multiple wavelength channels on the mesh topology precludes the use of several conventional multicasting techniques. This is especially true if no, or limited wavelength conversion is used, and wavelength continuity is required. Notice here that wavelength continuity must be satisfied both in depth due to signal propagation, and in breadth, due to multicasting and signal branching. Channels on different fibers cannot therefore be treated independently, as it is the case with IP multicasting.

IV.2.ii Circuit-Switched Communications Limitations

Circuit-switching communication mode used in wavelength routing networks is characterized by the following [12]: (1) route determination, (2) resource reservation along the route, (3) long setup time, (4) connection request blocking due to lack of resources, and (5) the static and the long-duration natures of the connection. Therefore, performance metrics for multicasting in this environment must be chosen carefully. For example, the call acceptance probability may have to be defined differently, e.g., allowing delivery to a subset of the multicast group [21, 22], where members of the subset may have different weights. Also, reducing the connection setup time is a significant issue in circuit-switched networks [23] in addition to the desire to obtain a reasonable per receiver propagation delay [56] or a total multicast tree cost [56, 58, 59, 71].

These features do not only affect network operation, but also network provisioning and dimensioning, as they impact the cost of the network, and its ability to handle the projected traffic.

IV.2.iii Optical Hardware Limitations

The limitations of the state-of-the-art in optical components, and the high cost associated with their usage preclude full deployment of such components at all nodes. This results in a network structure that is unbalanced as far as the node functionalities are concerned. For example, not

all nodes may be multicasting capable, or having wavelength conversion capability, and if it is available, conversion can be done between different regions of the optical spectrum at different nodes. This asymmetric deployment has a significant effect on the multicast routing and tree maintenance, as will be described later in the paper.

Optical Splitting Impact

The wavelength splitting capability is a key enabling technology for multicast communication in wavelength-routing networks. Light splitting is equivalent to packet replication in the electronic domain. However, it is theoretically simpler since it can be achieved by using optical passive devices, which are inexpensive and have very little impact on the signal-to-noise ratio of the optical signal. It also eliminates the need for buffers usually needed for data duplication in the electronic domain. Moreover, they maintain the optical signal transparency. Any node that is equipped with the light-splitting capability is called a Multicast-Capable (MC) node, otherwise it is called a Multicast-Incapable (MI) node [25]. In addition, a simple version of the splitting capability, which is assumed to be available at all nodes², is the drop-and-continue (DaC) [56, 58] or sometimes called the tap-and-continue (TaC) [40, 71]. This refers to tapping a small amount of the power which is used for signal detection by the receiver(s) connected to the node.

The example shown in Figure 3 illustrates the significance of employing light-splitters for routing a multicast session, denoted by $m=(11,\{6,8,9,12,14\})$, in the NSFNET. All nodes are assumed to have the DaC capability and node 10 is equipped with a light-splitter³. Also, wavelength channels are sufficiently available; hence, no call blocking is encountered. The multicast tree, called the *light-tree*, is then constructed such that the number of hops between the source and each destination is minimal, as depicted in Figure 3-(a). Node 6 receives the signal transmitted by node 11, while nodes 8, 12 and 14 receive the signal after being split by node 10. Node 9 will receive the same signal received by node 12, but after tapping a small amount of power at node 12 for signal detection by using the DaC capability. Effectively, a single multicast light-tree is sufficient to support the multicast connection, m , using one wavelength.

However, with the same system setup, but without a splitter at the branching node 10, constructing a single multicast distribution tree for this session is not feasible because only a single offspring can be attached immediately to node 10 on any particular wavelength (i.e., the fanout of node 10 is 1 as will be explained later). Figure 3-(b) illustrates this case in which node 11 needs to transmit on three different wavelengths (λ_0 , λ_1 and λ_2) to node 10, which in turn relays

²Some multicast routing schemes assume that some nodes only are equipped with this capability, e.g. [24].

³Also, node 10 has a *complete* splitting capability. The concept of complete splitting will be discussed shortly.

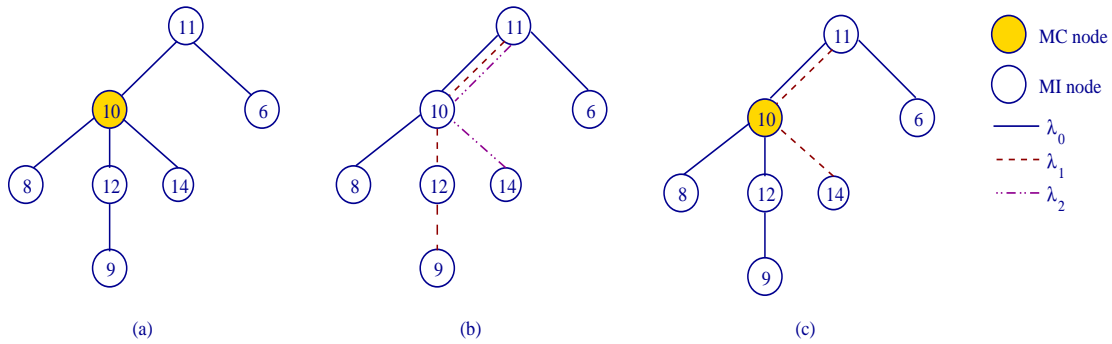


Figure 3: The multicast delivery structure (light-tree or light-forest) for $m=(11, \{6,8,9,12,14\})$ in NSFNET when f factor for node 10 equals (a) 3, (b) 1, and (c) 2

these transmissions to the different branches in order to reach destinations 8, 9 and 14. Multicast routing in this case consists of three distinct light-trees (called a *light-forest*) and employs two additional wavelengths.

The use of splitters reduces the signal strength and amplifiers may have to be used. However, optical amplifiers are still expensive, and their use degrades signal to noise ratio due to amplification to spontaneous emission [4]. Therefore, the benefits of using signal splitters are significantly reduced. Besides, integrating optical amplifiers within optical switches⁴ complicates the switch design [39].

On the other hand, not using optical amplifiers to compensate the power loss impacts many operational parameters, e.g., the number of splitters along each source-destination path. For example, assume that the receiver sensitivity is denoted by P_{sen} . Also, let splitters be implemented using 2×2 couplers which results in a $3dB$ signal loss. If the source power is limited by a maximum signal power, P_{max} , with value, say $XdBm$, the upper-bound on the number of splitters per source-destination path is $\frac{X-P_{sen}}{3}$. In practice, $XdBm$ is limited to $0.1mW$, and for the case of non-coherent detection method, P_{sen} should be no less than $-30dBm$. Therefore, no more than 7 splitters are allowed. However, with coherent detection, P_{sen} can be as low as $-50dBm$, which means that up to 14 splitters can be used⁵.

As a consequence of the above, a practical situation arises in which only some of the switching nodes are MC nodes while the rest are MI nodes. Networks in which both MC and MI nodes coexist are known in the literature as networks with *sparse-splitting* [27] and many schemes were introduced to deal with this situation, e.g., [56, 59, 71, 58]. Empirically, it was shown that around 50% of the nodes in the networks need to be MC nodes in order to achieve good performance level

⁴This integration is referred to as *on-site amplification scheme* and it will be discussed later in this subsection.

⁵The calculations for finding this upper-bound are made with the assumption that there is no attenuation on fiber.

[27, 41]. Such networks are known to implement Partial Packet Replication (PPR) [24]. In addition, two extreme strategies for MC node deployment were investigated in the literature in which all the nodes are either MI nodes, e.g., [40], or MC nodes, e.g., [21, 53, 22]. These strategies are called, *no-splitting* (or No Packet Replication-NPR), and *full-splitting* (or Full Packet Replication-FPR), respectively. The degree of splitting in the network directly influences the design of multicast routing algorithms as will be explained in Subsection VI.1.

Multicast support in wavelength-routing networks is also influenced by the splitting fanout, which is the maximum number of multicast tree branches supported per node. The fanout of an MC node is denoted by f , where $1 \leq f \leq d-1$ and d is the nodal degree. When f equals $d-1$, the switch is said to have a *complete* splitting capability; otherwise, the splitting capability of the node is *limited* since only a selected subset of the node's neighbors can receive the split signal. The splitting fanout is an important parameter in the design of multicast trees and it also impacts the choice of the number of amplifiers, their placement, and signal-to-noise ratio.

Consider the same system setup in Figure 3-(a) but with the assumption that the f factor for node 10 is 2. With this configuration, node 10 can forward the signal to only two of its offsprings, say nodes 8 and 12. A separate connection from the source node must then be established to reach the remaining node, viz., node 14, via node 10, but on another wavelength. The multicast delivery structure then consists of 2 separate trees on two different wavelengths, as shown in Figure 3-(c).

The actual splitting ratio also influences multicast tree construction. Signal splitting ratio can be either *fixed* or *adjustable*. With fixed splitting, the light-splitter divides the input signal equally into f output signals (ignoring the excess loss), where f is the fanout factor. This kind of splitter is a passive one, that is simple and inexpensive. However, it can be power inefficient, especially when the multicast tree is unbalanced. In this case, some signal power is wasted and the multicast tree size is affected. On the other hand, splitters with adjustable splitting ratios can significantly reduce the power wastage as well as maximize the multicast group size. An analysis of this issue will be presented in Section V. It is also worth mentioning that the use of splitters with adjustable splitting ratios becomes imperative in the context of dynamic multicasting where the multicast tree structure dynamically changes. However, splitters with adjustable-ratio are active devices that are more expensive than their fixed-ratio counterparts, may be unstable and may not be noise-immune.

Wavelength Conversion Effect

Wavelength Conversion, or in fact, the lack of it, also influences multicast routing. Wavelength-converters are active optical devices which shift the optical signal from one optical frequency to

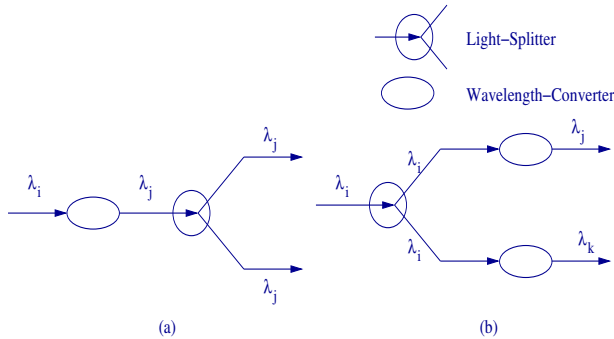


Figure 4: Splitter-converter relative locality in the switch (a) Pre-Conversion Scheme, and (b) Post-Conversion Scheme.

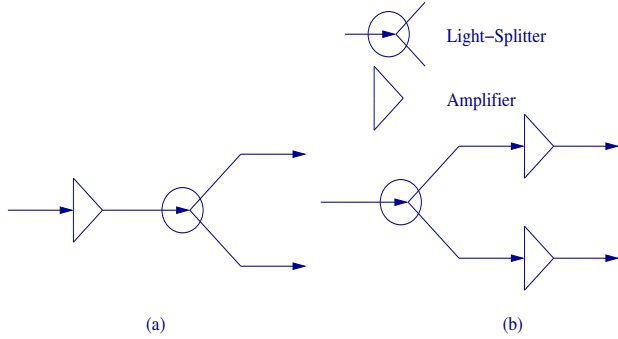


Figure 5: Splitter-Amplifier relative locality in the switch (a) Pre-Amplification Scheme, and (b) Post-Amplification Scheme.

another. The employment of the wavelength converters provides flexibility in the network operation and enhances its performance. The use of wavelength conversion can simplify multicast routing since wavelength continuity will not be a constraint.

All-optical wavelength converters [28, 29], however, are still very expensive and immature. Also, equipping the switch with wavelength converters complicates the switch design and increases its cost, which, similar to light-splitters, hinders *full* deployment of wavelength converters. Instead, *sparse* and *limited* conversion are usually more commonplace. Limited conversion means that conversion from a certain wavelength is limited to only a sub-band of the optical spectrum. Sparse conversion, on the other hand, refers to the situation in which only a subset of nodes are equipped with converters. Both configurations, i.e., limited and/or sparse conversion, requires some degree of wavelength continuity, which restricts the multicast tree construction.

Integrating wavelength converters with splitters can take two forms, namely, *pre-conversion* or *post-conversion*. As shown in Figure 4, these schemes correspond to placing the converters before or after the light splitter, respectively⁶. Since a single converter is needed, the cost associated with the Pre-Conversion scheme is much less than that of the Post-Conversion scheme. Both schemes provide the same degree of flexibility if converters always convert from λ_i to λ_j . Otherwise, the Post-Conversion scheme in general exhibits more flexibility and results in higher system performance.

It is known that the *SNR* can be improved by using all-optical wavelength converters [28, 29], and such improvement is enhanced by cascading the converters [30]. Moreover, very little power loss will be suffered when a signal passes through a wavelength converter because of the amplification

⁶Because each wavelength requires a separate converter, Figure 4 depicts the requirements for a single input-output channel pair (i.e., λ_i and λ_j , or λ_k), regardless of the range of the wavelengths at their input and output.

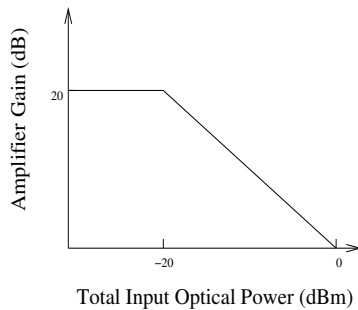


Figure 6: Amplifier Gain Model.

capability of the converter. However, more investigation is needed to study the relation between the input signal power and SNR improvement for the pre-conversion and post-conversion schemes.

Optical Amplification Impact

The final hardware-related challenge is the amplifier placement problem. In multicast communication, power loss can be due to fiber attenuation and splitter loss, if any. One of two techniques for signal amplification may be used, namely, *on-site* and *on-link* amplifications. With the on-site amplification, the amplification is performed inside the switching node itself. There are two methods to integrate the amplifier and the splitter within the node. In the first one, called the *pre-amplification* technique, the amplifier is placed before the splitter, while in the other case, called the *post-amplification* technique, it is placed after the splitter, as shown in Figure 5. Amplifiers exhibit a saturation region in the amplifier gain curve, as depicted in Figure 6 [31, 32]. For sake of comparison between these on-site amplification schemes, the following two scenarios are considered (we ignore excess loss). In the first scenario, assume that the input signal is weaker than $-20dBm$, then the output signal levels under both strategies are the same. For example, if input power is $-25dBm$, then under the pre-amplification scheme, the output power of the amplifier is $-5dBm$. Because of the $3dB$ splitting loss, the final output signal power is $-8dBm$. The same output signal power is obtained under the post-amplification scheme too.

In the second scenario, the input signal is assumed to be stronger than $-20dBm$. In the first structure, the input signal at the splitter is always $0dBm$ because of gain saturation, and the output signal is always $-3dBm$. However, with post-amplification, the output signal will be always greater than $-3dBm$, which is better than that with pre-amplification. For example, if the input power is $-15dBm$, the output power of the amplifier under the first structure is $0dBm$, and the final output signal power is $-3dBm$. However, in the second structure, the output power of the splitter is $-18dBm$ because of the $-3dB$ loss from that splitter, and the final output power is $0dBm$, which is stronger than the $-3dBm$ output signal in the first structure.

The first scheme is less expensive. However, it is to be noted that the SNR at the output signal under both schemes will be the same. This is because splitters are passive devices which theoretically do not introduce noise, and because the spontaneous emission amplification in all-optical amplifiers is proportional to the signal power.

On the other hand, the on-link amplification is performed by amplifying the signal while it is traversing the physical links. For multicasting, placing amplifiers on fiber links can increase the number of potential receivers. An optimal placement configuration can reduce the total number of amplifiers in the network. There are several studies, e.g. [31, 32], which deal with the problem of amplifier placement for unicast service. The ability of such schemes to accommodate multicast traffic needs further investigation.

It is essential to notice that one difference between multicast signal and unicast signal is the signal power level. Because of signal splitting, a multicast signal power level can be half of a unicast signal or less. Unfortunately, an ideal amplifier would amplify all input signals with the same gain, which means that a weak input signal will still be weak at the output, and the maximum gain of an amplifier is limited. Although the same situation is encountered with unicast traffic, the difference in power level of the different signals in the optical spectrum is significant and its impact is more severe in the context of multicasting. This implies that a multicast signal may not obtain enough power if it is accompanied by other stronger signals.

Moreover, to find an optimal amplifier placement for a multicast tree, at least two parameters, namely, signal power and source-destination distance, should be given. However, dynamic multicasting and the use of adjustable splitting coefficient aggravate the problem since the former results in a communication diameter that increases or decreases, while the latter increases or decreases the signal power. This problem can be solved by using the splitter with post-amplifier shown in Figure 5-(b) or introducing equalization at each amplifier stage [4].

Another problem which has been alluded to earlier is the effect of amplifiers on the SNR . All-optical amplifiers, e.g., Erbium-Doped Fiber Amplifiers (EDFA), amplify noise due to spontaneous emission [4]. Obviously, there is a lower-bound signal-noise-ratio (SNR_{sen}) for a detectable signal which is determined as

$$\frac{\textit{original signal power}}{\textit{original noise} + \textit{Amplifier noise}} \geq SNR_{sen}. \quad (1)$$

In a best scenario, an amplifier can reduce the SNR by a factor of 2, i.e., by passing through an amplifier, the SNR loses $3dB$ of its power [33]. In general, the *relative intensity noise* (RIN) per unit bandwidth is $10^{-15} dB/Hz$ [33]. Thus, for a $10GHz$ channel, $\frac{\textit{Noise}}{\textit{Power}}$ equals 1×10^{-5} , and $SNR \leq 50dB$. Moreover, the required *bit error rate* (BER) for high-speed optical communication

systems today is typically 10^{-12} [4]. Using the formula $p = Q(\frac{1}{2}\sqrt{\frac{S}{N}})$ [33] for the *BER* in terms of *SNR*, we obtain $SNR = 23dB$. Thus, total loss of *SNR* should be no more than $50 - 23 = 27dB$, and the total number of amplifier that can be installed on a path should be no more than $\frac{27}{3} = 9$. This is an approximate upper-bound on the number of amplifiers that can be placed between a source and a destination without any signal regeneration. Finding an optimal solution to minimize the number of amplifier in a multicast capable optical network is still an open issue.

IV.3 Challenges Due to Routing and Wavelength Assignment

In this subsection, we concentrate on two basic functions in wavelength-routing networks, namely, routing and wavelength assignment, from the multicasting viewpoint. The Routing (R) and the Wavelength Assignment (WA) problems have been extensively studied under unicast communication [4]. Although solving the combined Routing and Wavelength Assignment (RWA) problem results in an optimal operation, due to the difficulty of this problem, the two functions are treated separately. The multicast extension of this combined problem is referred to as MCRWA [21] or MC-RWA [22]⁷.

This problem is mainly concerned with establishing the multicast route in the network, and determining the appropriate wavelength to be assigned to it with an objective of minimizing the required network resources, especially the number of wavelengths, or maximizing the number of accepted calls. The combined problem has been proven to be NP-Complete [21, 34] since it involves finding an optimal multicast tree which is the well-known NP-Complete Steiner Minimum Tree (SMT) problem [35]. Hence, it is solved using heuristics. The combined MCRWA can also be characterized as a *wavelength-aware* approach because computing the multicast delivery structure is directly influenced by the status of the wavelengths in the network and the cost associated with their use.

Similar to the traditional RWA in unicast networks, the MCRWA can be solved as two separate problems, namely, the multicast tree is first determined, followed by appropriately selecting wavelength(s) on this route. Therefore, the uncoupled version of MCRWA is a *not-wavelength-aware* approach. The routing part of the problem is still NP-Complete since it involves the construction of the SMT, but the wavelength assignment part can be solved in polynomial time for some special cases, e.g. [34, 36]. Decoupling the routing and wavelength assignment problems simplifies the problem slightly, and also provides flexibility in enhancing routing protocol designs and integrating extra operations and features, such as QoS guarantees, fault-tolerance, and security.

⁷For sake of consistency, and in order to emphasize the All-Optical feature of multicast communication, we will introduce the AOM-RWA, AOM-R, and AOM-WA notations later in Subsection VI.1.

Furthermore, the operational alternatives of the routing and the wavelength assignment operations achieve different performance degrees with different costs. The performance-cost trade-off of these alternatives must be considered in the protocol design process. In this context, and at the routing side of the problem, the routing schemes have been traditionally characterized as either: (1) *fixed*: in which a fixed route per call is established, (2) *alternate*: a set of paths are assigned for each call which are searched sequentially in a specific order in order to pick the route, and (3) *adaptive*, in which the route is found dynamically based on the network status. The flexibility of these schemes increases from the fixed to the adaptive, but at the cost of increased overhead and complexity.

Similarly, many schemes have been proposed in the literature for wavelength assignment. However, four of these schemes were inherited from unicast communication: (1) *random*, in which wavelength selection is performed randomly from the set of available wavelengths, (2) *first-fit*, which chooses the first available wavelength, (3) *least-used*, and (4) *most-used*, in which the wavelength to be chosen is based on the degree of its usage. Other schemes, however, were proposed in the context of multicast and will be described in Section VI.1.ii. In general, the complexity of the wavelength assignment problem is a design factor that must be considered.

Finally, the routing and wavelength assignment in optical network can be performed in two operational modes: *on-line* and *off-line*. The *on-line* mode allows requests to be processed as they arrive to the system, i.e., it does not assume prior knowledge of the traffic in the network, and the admission decision of the connection is based on the instantaneous network state. If the connection request is admitted, its route and the assigned wavelengths are usually not changed during the lifetime of the connection. Otherwise, and because of the no waiting feature of circuit-switched service, the connection is blocked. The *off-line* mode, on the other hand, assumes prior knowledge of all the connection requests which makes it possible to obtain an optimal allocation for the routes and the wavelengths. The first operational mode is more practical, but optimal operation is more difficult, while the opposite is true for the second operational mode.

IV.4 Challenges Due to the All-Optical-Multicasting (AOM) Characteristics

All-Optical Multicasting (AOM) inherits all the problems of conventional multicasting. For example, the ideal operation of the multicast session entails hiding the information about the various receivers from the source node, and vice versa. These information include the nodes' addresses, locations and their total number. A dedicated address is assigned to the multicast session which is globally known. This requirement is still not yet fully achieved in most routing and wavelength assignment techniques of AOM, which are based on a key assumption of global (or partial) knowledge

about the network topology, and the nodes membership. However, [23] appears as an exception in which the source node needs not to know about the destination nodes and their locations.

In addition, different attributes of the multicast session may necessitate the use of routing strategies, which may look suboptimal, but are necessary because of the all-optical nature of the session. For example, consider the multicast session $m=(7,\{1,6,14\})$ in the NSFNET which has a short duration. Not only the number of the member nodes is small, but they are also geographically scattered and far away from each other. Under such a scenario, the designers might find it more efficient to use three separate unicast connections for this session instead of attempting to construct a single optimal tree that joins all the nodes. Such a solution may be even critical if free wavelengths exist, but wavelength continuity among all the branches cannot be maintained.

Furthermore, the connection acceptance metric may take different levels, and not just two. Nodes in the destinations set may be assigned weights, and a call is accepted if a certain total weight is achievable. For example, the multicast session is rejected (blocked) if the traffic cannot be delivered to at least one member. This can be referred to as the *full* acceptance criterion, i.e., the all or none strategy, and it provides simple and cost-effective mechanisms for connection management, but with relatively reduced network utilization. Also, one may use a partial acceptance criterion, in which a call is successful if a minimum number of destinations is reachable. The other unreachable members will be blocked. This scheme enhances network utilization, however, the cost associated with tree maintenance, network management and control-information flows is very high. The suitability of connection acceptance metric is application dependent.

On the other hand, the implementation of dynamic multicasting can be quite challenging in the all-optical environment. For example, the wavelength assignment may have to be done in a way that maximizes the probability of accepting new members, given the set of potential multicast members. This may sometimes lead to an underutilization of network resources. Moreover, in order to provide an efficient implementation of dynamic multicasting, splitters with adjustable splitting ratios may have to be deployed, which are more costly, and harder to control.

V Design Issues

This section addresses several issues involved in designing AOM networks. Such issues include network elements design and implementation, the effect of splitting on multicast group size and network diameters, network provisioning, network dimensioning, and connection provisioning.

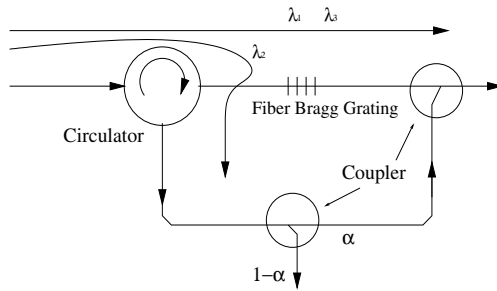


Figure 7: DaC Element Structure

V.1 Network Elements

References [41, 47, 27, 53, 21, 56] considered the various network elements that are needed for supporting multicast over optical layer. Two components were introduced, namely, Drop-and-Continue (DaC) element and Multicast-Capable (MC) switch. At least one of these two components is required in order to support multicasting in an all-optical manner in the network.

The DaC node can be set to “*drop-only*”, “*continue-only*”, or “*drop-and-continue*” [56]. This element can be implemented as shown in Figure 7, and it consists of a circulator and a fiber bragg grating to drop the expected wavelength, and a coupler to tap part of energy from that wavelength.

Several MC nodal architectures have been proposed in the literature. The splitter-and-delivery (SaD) switch structure [38] is designed such that each input signal is initially split into d sub-signals, where d is the nodal degree. The split signals are then switched to the appropriate output port using a combination of 1×2 switches. Although the SaD switch provides a non-blocking service, it does not distinguish between different traffic types. Therefore, the unicast traffic undergoes unnecessary power loss. Also, the switch design is very complicated. The Multicast-Only Splitter-and-Delivery (MOSaD) was then proposed in [39] and it is based on the idea of sharing the splitters by a group of multicast sessions. It is a power-efficient MC cross-connect architecture that supports W wavelengths. As shown in Figure V.1, after optically demultiplexing all wavelength on the fibers, an input signal is switched to a corresponding output link using a space switch. Only one output port of each space switch is connected to a split-switch bank (SSB) that is used to split the incoming multicast signal. Hence, splitting is only applied to multicast connections. Also, the advantage of this architecture is that only W splitters, one for each wavelength, are needed. which simplifies the switch design. However, due to splitter sharing concept, there is no provision to split different multicast signals coming from different input, but on the same wavelength.

The same authors propose a Tap-and-Continue (TaC) cross-connect architecture in [40]. The basic structure of this cross-connect is similar to MOSaD, except that TaC modules (TCMs) replace SSB in TaC cross-connect. When a multicast signal passes through a TCM, only a small

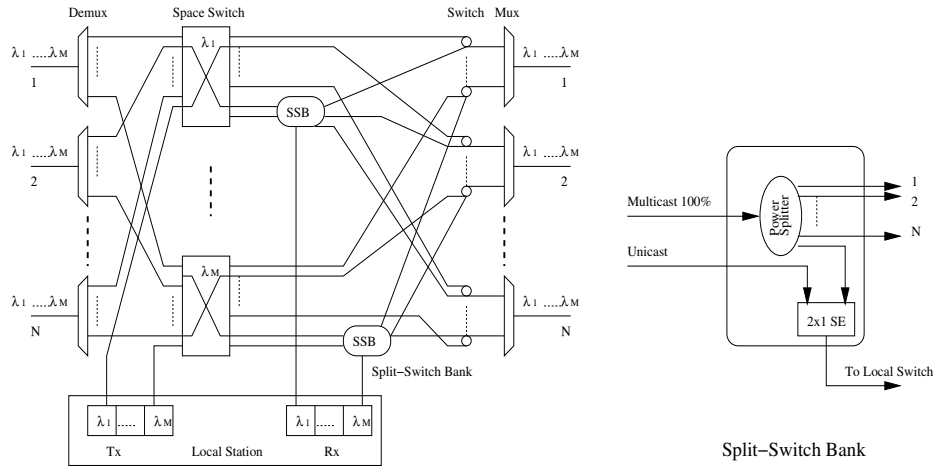


Figure 8: Multicast-only splitter-and-delivery (MOSaD) cross-connect.

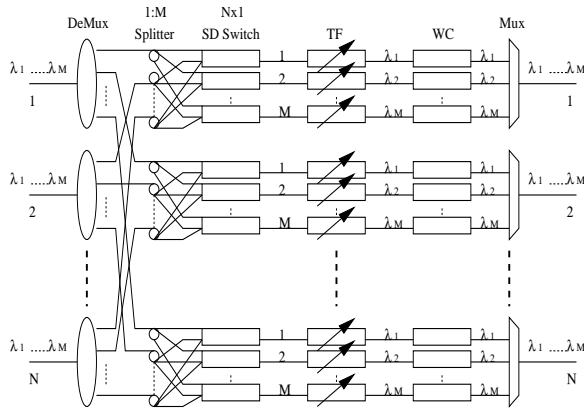


Figure 9: Multicast Capable Switch with Wavelength Conversion [56].

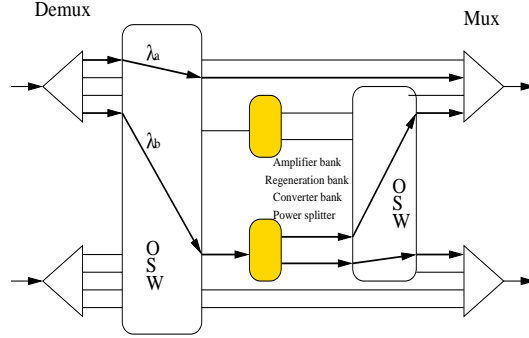


Figure 10: Multicast Capable Switch with Wavelength Conversion [53].

fraction of this signal is tapped and forwarded to the local station, while the remaining power is switched to an output port. No splitter is needed in this architecture; thus, the routing problem is more difficult.

Wavelength conversion is not employed in both MOSaD and TaC cross-connect architectures. Multicasting with wavelength conversion capable cross-connect architectures were proposed in [53, 56]. The architecture proposed in [56] employs two stages of splitters, and the output signals from the splitter pass through a space-division switch. Each output signal from the switch may be converted into different wavelengths by using a wavelength converter. The signals are then multiplexed on an output port fiber. The switch architecture is shown in Figure 9.

In [53], the proposed MC structure employs two stages of optical switches. After demultiplexing, input signals pass through the first stage of optical switch, and only multicast signals are switched to a splitter. After that, the split signals pass through an amplifier bank. Then the signal

passes through a second stage of the switch to the appropriate output port, as shown in Figure V.1. This architecture is simple and less expensive than the last one, while its multicast capability is limited.

Finally, it has been found that multistage WDM multicast switching structure has lower network cost than a crossbar multicast network [52] if the switching network cost is defined as the number of crosspoints and wavelength converters.

V.2 Network Diameter and Multicasting Group Size

In this subsection, we address several design issues regarding the multicast group size and network diameter under multicast traffic. Primarily, a multicast group size is determined by light-power limitation as well as signal-to-noise-ratio (SNR), if no signal regeneration is allowed at intermediate nodes.

V.2.i Constraints

In order to detect the received signal, the following light-power constraint must be satisfied:

$$P_{max} - \text{Power loss in Attenuation} - \text{Power loss in Splitting} - \text{Power loss in Tapping} \geq P_{sen} \quad (2)$$

where P_{max} and P_{sen} are defined as before. Using equation 2, one can calculate an upper-bound on the multicast group size when no amplifiers are employed. The use of different multicasting techniques, i.e., DaC and splitter (where the multicast route takes the form of a path and tree, respectively), and different signal detection techniques, i.e., coherent-detection and non-coherent-detection, will influence this upper-bound. The remainder of this subsection will provide an analysis of the impact of these factors on the multicasting group size, using the following assumptions:

1. Fiber Bragg Grating loss equals $0.1dB$ [4],
2. Maximum power available from a transmitter (P_{max}) equals $1mW$ [31, 32], and the transmitter power in dBm , $ToDB(P_{max})$, is $0dBm$,
3. Minimum power required for signal detection on a wavelength, P_{sen} , equals $1\mu W$ for non-coherent detection, and $ToDB(P_{sen})$ equals $-30dBm$ [4]. And, for coherent detection, P_{sen} equals $0.01\mu W$, and $ToDB(P_{sen})$ equals $-50dBm$ [4]. A node can be in a multicast group if and only if $P_{Receive} \geq P_{sen}$, where $P_{Receive}$ is the received light power at a node,
4. For common glass core/glass cladding fiber, the attenuation coefficient is $0.17dB/Km$ to $2dB/Km$ [42, 33]. Here, we assume the attenuation coefficient to be $0.17dB/Km$,

5. The distance between two adjacent nodes is assumed to be $5.9Km$, therefore, the attenuation between two nodes is $1dB$, and
6. The total number of wavelengths in WDM network is 10, and the bandwidth for each wavelength is $10Gb/s$. Thus, the maximum power available for each wavelength is $P_{max}^{wav} = 0.1mW$, and $ToDB(P_{max}^{wav}) = -10dBm$.

V.2.ii Multicast Group Size with DaC Elements and No Amplification

Two possible cases can be encountered. In the first case, the drop coefficient, α is fixed, with a value equals to, say, 0.95 for each tap. The coupling loss equals $0.22dB$. Employing non-coherent detection scheme, the input power at the tapper, P_{input} , should satisfy the following two inequalities: $P_{sen} \leq P_{input} \times 0.05$, and $P_{input} \geq 0.02mW$ for each DaC element before the last. The input power in dB should satisfy $ToDB(P_{input}) \geq -17dBm$. Thus, the power received before the fiber bragg grating element at this node, P_{recv} , should satisfy: $ToDB(P_{recv}) \geq -17dBm + 0.1dB = -16.9dBm$.

Suppose the number of receivers between the source and the receiver before the last is n , then we have the following equation:

$$P_{max}^{wav} - n \times (\text{Attenuation} + \text{GratingLoss} + \text{DropLoss}) \geq P_{sen} \quad (3)$$

where, Left-Hand-Side (LHS) = $-10 - n \times (1 + 0.1 + 0.22) \geq -16.9dBm$, which yields $n \leq 5.2$. Thus, total number of receivers is: $n + 2 \leq 7.2 < 8$.

Similarly, when coherent detection scheme is employed, the $ToDB(P_{sen})$ equals $-50dBm$, therefore, $ToDB(P_{input}) = -37dBm$, and $ToDB(P_{recv}) = -37dBm + 0.1dB = -36.9dBm$. Using equation 3 with the right-hand-side (RHS) = $-36.9dBm$, n is found to be ≤ 20.5 , and the total number of receivers can be: $n + 2 \leq 22.5 < 23$.

In the other case, α is assumed to be adjustable. Recall that the fiber attenuation equals $-1dB$, which means that only 79.4% of the signal power can be received by the next node downstream, and the grating loss equals $0.1dB$, hence, only 97.7% of the signal power can be received by next element. When the non-coherent detection scheme is used ($P_{sen} = 1\mu W$), the following equation can be used to compute the number of receivers, such that: $((0.977 \times 0.794 \times 0.1mW - 0.001mW) \times 0.977 \times 0.794 - 0.001mW) \times \dots \geq 0.001mW$. This yields $n \leq 13$ receivers. Similarly, when coherent detection method is used ($P_{sen} = 0.01\mu W$), the following equation is true: $((0.977 \times 0.794 \times 0.1mW - 0.00001mW) \times 0.977 \times 0.794 - 0.00001mW) \times \dots \geq 0.00001mW$, which yields $n \leq 31$ receivers.

V.2.iii Multicast Group Size With Light Splitter and No Amplification

Assume that a light splitter splits the power of the signal equally among two output fibers, thus, the splitter loss equals $3dB$. Considering non-coherent (coherent) detection schemes, equation 3 is solved with $RHS = -30dBm(-50dBm)$ and with the assumption that multicast tree is a balanced tree. The results show that the maximum number of receivers is $2^4 = 16$ ($2^9 = 512$) and the minimum number of receivers is 5 (10) in the worst case.

When splitters with adjustable ratios are employed, splitting the power can be proportional to the number of the members in the corresponding sub-tree and their separation.

V.2.iv Multicast Group Size with Amplification

On the positive side, installing amplifiers in the network enhances the multicast group size. In order to include the amplifier effect, equation 2 is modified as follows:

$$\begin{aligned}
 P_{max} & - \text{Power loss in Attenuation} - \text{Power loss in Splitting} \\
 & - \text{Power loss in Tapping} + \text{Power Gain of Amplifier} \geq P_{sen}
 \end{aligned}$$

However, because of the noise introduced by the amplifiers, the multicast group size can be unlimited if amplifiers are used. According to numerical results conducted from the constraint equation 1, the followings can be concluded too. For an all-optical network without splitters, and with fixed drop coefficient tapper, the maximum number of receivers for non-coherent detection system is $n \leq 8 \times 9 = 72$; and for coherent detection system, $n \leq 23 \times 9 = 207$. On the other hand, if flexible drop coefficient tapper is used, then the maximum receiver number for non-coherent detection system is $n \leq 13 \times 9 = 117$, and for coherent detection system, $n \leq 31 \times 9 = 279$. However, for an all-optical network with light-splitters, the maximum receiver number for non-coherent detection system is: $45 \leq n \leq 2^{4 \times 9} = 2^{36}$, and for coherent detection system, $90 \leq n \leq 2^{9 \times 9} = 2^{81}$. Thus, without light-splitters, the multicast group size can hardly beyond 300 members, and with light-splitters, the multicast group size can be almost infinity.

Figure 11 shows the relation between the network diameter with multicast group size. In this figure, we only compare eight methods, which are Unbalanced Tree with Equal Splitter by Coherent Detection (UTESCD) or Non-Coherent Detection (UTESNCD), Unbalanced Tree with Flexible Splitter by Coherent Detection (UTFSCD) or Non-Coherent Detection (UTFSNCD), Drop-and-Continue by Coherent Detection (DCCD) or Non-Coherent Detection (DCNCD), and Flexible Drop-and-Continue by Coherent Detection (FDCCD) or Non-Coherent Detection (FDCNCD). Here, Unbalanced tree is a tree in which each light-splitter connects directly to one group

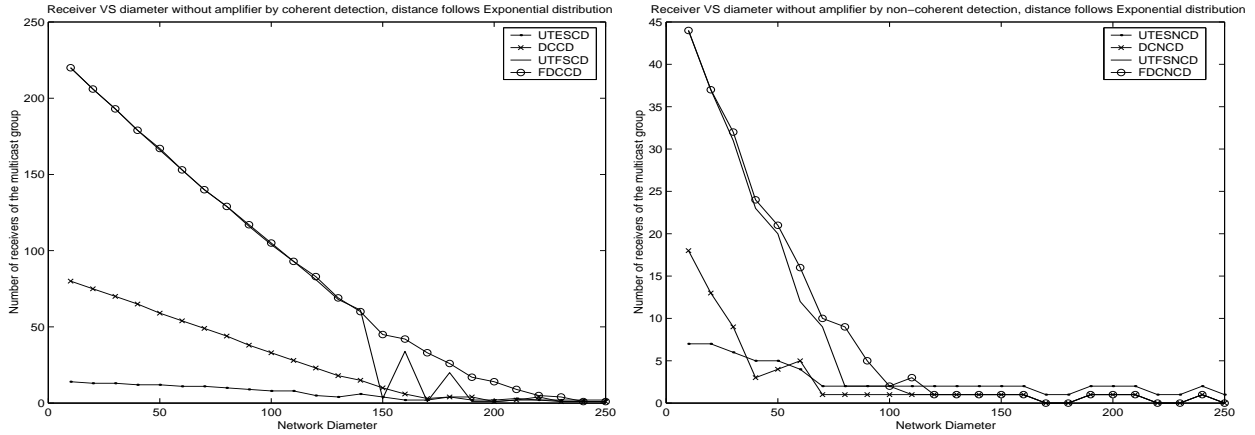


Figure 11: Network Diameter VS. Multicasting Group Size

member and one light-splitter. In this figure, FDC(N)CD and UTFS(N)CD have almost same group size, while UTES(N)CD has the smallest group size in most cases. This is because fixed splitter consumes more light power than other methods do, and flexible splitter is very similar to flexible DaC elements. We can simply expand this result if amplifiers are used in the network.

V.3 Network Provisioning

V.3.i Number of Splitter and Placement

In [27], the authors assumed that an unlimited number of wavelengths is available to avoid blocking since their only concern was to determine the total number of wavelengths needed as well as the total amount of bandwidth. They also assumed that the only traffic present in the network is of multicasting type, and the network had full wavelength conversion capability. With this assumption, both NSFNET and an arbitrary network with 30 nodes were studied for efficiency of sparse light splitters. A relationship between the number of wavelengths and required light splitters was developed too. The simulation results showed that: (1) full light splitting capability can only reduce the bandwidth consumed by an average of 8% – 18%, (2) by equipping 50% of the network nodes with splitting capability, the network could obtain around 80% – 90% of the benefit of full light splitting capability, and (3) no more than 70% of the network nodes needed to be equipped with light splitting capability to obtain almost the same benefit as full light splitting capability.

In [41], the splitter placement in wavelength-routed network problem was formulated and showed to be NP-Complete. For given constraints such as the number of the wavelengths per fiber, number of fibers per link, at least one splitter on each multicasting tree, and one splitter on at least one multicasting tree, the objective function is to maximize the number of multicasting

sessions with different proportion of splitters in the whole network. In this formulation, each multicast session could establish multiple trees. Three heuristic algorithms were developed, i.e., CPLEX heuristic, Most-Saturated Node First (MSNF), and Simulated Annealing (SA) heuristic. The SA heuristic performed much better than the CPLEX because SA can always find a feasible solution even if CPLEX cannot find one. Moreover, the blocking performance is enhanced by increasing the number of splitting nodes, however, the best design performance is obtained when no more than 50% of the total nodes are equipped with splitting capability, which agrees with the result obtained in [27].

The optimal placement of the splitters was investigated in [80] and [81]. In [80], the authors developed a Mixed Integer Linear Program (MILP) formulation for the optimal placement of a limited number of wavelength converters and power splitters. The formulation was then extended to provide a logical topology design that is based on light-trees and is able to groom several sub-wavelengths multicast streams.

In [81], although the optimal placement of power splitters was investigated, the power constraint were considered and the optimal number and placement of optical amplifiers was part of the formulation.

Reference [37] proposed a heuristic algorithm to optimize the splitting node placement in optical network. This algorithm proceeds for k iterations to place a total of k splitting nodes in the whole network. In each iteration, all nodes without multicast capability are chosen one at a time to be equipped with a splitter. The traffic blocking probability is computed and the node with minimum blocking probability is chosen as a new splitting node in the network. After k iterations, the allocation assignment is returned as the best configuration of the sparse splitting nodes placement. However, the optimization of splitter placement problem remains an open issue.

V.3.ii Amplifier Placement

The amplifier placement problem under multicast traffic is another open and challenging problem. In this subsection, we present some of the main efforts both in the unicast and multicast contexts.

An optical amplifier gain model was developed in [31] that considered fiber attenuation, amplifier saturation, sensitivity of detectors, as well as maximum power of transmitting light sources in fiber. The authors proposed a scheme that minimizes the number of amplifiers for the LAN/MAN passive-star-based optical networks without the restriction that wavelengths in the same fiber must be at the same power level. This is a reasonable assumption due to the *near-far* phenomenon in which signals on different wavelengths can originate from different locations in the network, and when they arrive at an amplifier, their power levels can be different due to the propagation at-

attenuation in fibers. The difference in power levels of the input wavelengths can significantly limit the amount of amplification available to lower powered wavelength because the higher powered wavelengths could saturate the amplifier.

The authors formulated this problem as a Mixed Integer non-Linear Program (MInLP) with the objective of minimizing the number of optical amplifiers in the network. The output of the MInLP is used to allocate the determined number of optical amplifiers off-line. Two placement policies were introduced in [31], namely, As Soon As Possible (ASAP) and As Late As Possible (ALAP) methods. For each link, the ASAP places the optical amplifiers (except the last one) on the link as soon as the total input power is low enough to allow the *maximum* gain, while the last amplifier is placed such that the input power is low enough to allow it to generate the *remaining* required gain. The ALAP method operates in a similar fashion, except that it places the amplifiers such that the total input power is at its minimum detectable level (i.e., the power of one or more of the individual input channels is at P_{sen}) which results in placing the amplifiers closer to the destination side of the link.

On the other hand, two new methods were proposed in [32]: Genetic Algorithm implementation for amplifiers placement (GA2), and Smallest Gain First (SGF) algorithm. The GA2 algorithm started with a set of placement topologies of amplifiers sorted with transmitting power on the link. The two placement topologies with lowest powers are then selected as the parents, and genetic functions are used to combine parents with a structured, but randomized information exchange to form a child, and the child is then inserted in the whole set. The SGF aims to find a better amplifier placement by assigning the minimum transmission power to every lightpath, then it would add all the lightpaths of the chosen permutation link by link according to their length (longest one first). According to the simulations results presented in [32], GA2 performs better than SGF with lower transmission powers for all lightpaths.

The influence of amplifier noise power on the transmission signal is considered in [50]. However, it only considered the effect of the noise in saturating the amplifier, without using the *SNR* concept. Such problem can be possibly solved by increasing the receiver sensitivity, which is not the case if *SNR* is very low.

With multicast service, the amplifier placement problem must consider the power loss due to light splitting or the use of DaC devices. Signal power decreases dramatically when it passes through these devices. For example, the power loss will be $3dB$ after passing through a splitter, which means another constraints needs to be added in the design. Therefore, the demand for optical amplifiers (OAs) can be more critical with multicast traffic than with its unicast counterpart. Since OAs are expensive devices, every effort must be made to minimize their number in the network.

Moreover, because of the requirement of SNR , the total number of amplifier on a link should also be considered in amplifier placement problem.

In this context, the work in [81] introduces an optimal, power-constrained design for mesh-based wavelength-routing networks in the context of AOM that takes the power loss impairment into consideration. The significance of this work is that it provides a linear formulation to the network design problem in the form of Mixed-Integer Linear Programming (MILP) with the objective of minimizing the total required optical amplifiers' gain while allowing the signals to exist at different levels at any point in the network. It also determines the optimal number and placement of the splitters, and the routing and wavelength assignment for the multicast sessions. Moreover, the output from the MILP enables us to determine the exact number of optical amplifiers per link and their exact location using any of the placement policies used in [31]. The authors are not aware of any other work that addresses the same problem in the literature.

V.4 Network Dimensioning

Basically, all-optimization strategy is based on minimizing the number of critical resources, e.g., wavelengths per fiber, fibers per link, wavelength power, amplifiers, etc. However, optimization of some resources might imply an increase in others. Thus, obtaining a global optimum solution may be hard and complex problem.

Traditionally, network dimensioning has been treated in the literature under unicast traffic. A methodology is proposed in [46] for dimensioning on the WDM optical layer and optimizing the allocation of resources after discussing a global view of the dimensioning problem of a WDM network. The dimensioning problem was defined as to determine and allocate the resources needed to support a given traffic demand on a give network topology. After establishing the optical paths for connections, the dimensioning problem changes into assignment of the appropriate resources to those paths. On the other hand, a multifiber link-load correlation model was developed in [43]. According to the analytical and simulation results, the alternate path routing requires less fibers than the shortest path routing. Moreover, it was concluded that a multifiber network has a similar blocking performance as a full-wavelength-convertible network. Moreover, several dynamic and static wavelength assignment algorithms were developed in [44] for optical broadcast-star LANs with a dynamic traffic model. Using analytical and simulation techniques, it was shown that dynamic approach significantly outperforms the static approach. However, the wavelength conversion capability in such network improves the performance marginally. In addition, for wavelength converter placement issue, a heuristic algorithm was developed in [45], and the blocking performance simulation results over NSFNET and mesh-torus were presented.

Moreover, link capacity is an important issue that restricts the maximum load of the traffic a network can support. In this context, the results of some papers ([46, 49, 51]) on Dimensioning of the WDM networks and blocking probability analysis are presented while in Section VI.1.ii, we will present some of the efforts that address the wavelength requirements in terms of wavelength assignment under multicast traffic.

Four strategies were developed in [46] in order to optimize the resource assignment, which mainly focused on number of fibers and node's transmitter- receivers. In the first strategy, wavelength conversion was employed, and the simulation results showed that a 15% – 37% reduction in the total number of fibers could be achieved. The second strategy examined several routing schemes, which included shortest path, least number of hops, and a heuristic routing scheme. The simulation results showed that the least number of hops routing scheme could reduce node's transmitter-receiver by around 3% – 7%. Finally, the third strategy was based on rerouting connections on overloaded links, which could reduce the total number of fibers around 10%.

In [49], the network dimensioning problem was defined as to determine the lightpath assignments for different time segments with the objective of achieving the minimum resource utilization with a given set of traffic pattern. The problem was formulated as a multicommodity-flow problem, and integer-programming approach and minimum variance algorithm were used. In integer-programming approach, a set of constraints were examined and the objective function was specified to achieve minimal network resources that consists of fiber-link cost, switch cost, transmitter/receiver module cost, and optical amplifier cost. Because of its complexity, a heuristic approach, i.e., minimum variance algorithm was used. In the heuristic approach, the average cost of setting up a lightpath is defined as the total cost of fiber link of one route divided by the number of lightpaths that could be established. A route with minimum average cost was chosen. Using an average wavelength utilization on a link as performance metric, this approach found a solution with minimum variance of the wavelength utilization.

The absorption probability was used in [51] instead of the blocking probability in order to analyze the traffic load on each link. Absorption probability at time t is defined as the probability that at least one lightpath request is blocked before time t . Each node is assumed to have full wavelength conversion capability; therefore, a multihop route can be assigned different wavelengths on the different links. This relaxes the constraint of wavelength continuity, and allows application of approaches used in traditional networks. Because networks with long-duration lightpaths might not reach a steady state based on changeable demands, a Markov Chain with absorbing state was used instead of the blocking state. With this method, a link capacity at time t can be obtained, compared with fixed steady state probability of Markov Chain with blocking state. In

this method, every link was assumed to be independent, and that an average of one lightpath request would arrive every month, and the average holding period would be one year. Based on these assumptions, a single link network and two links network models were analyzed, and the exact absorption probabilities were computed. Also, both approximation of upper bound and lower bound of absorption probability for these two networks were given. Compared to a formal approach, the approximation matched very good. The simulation experiments were carried out over the ARPANET, and the results obtained from these experiments showed that the link capacity allocation based on the absorption probability was much better than that based on the blocking probability within the time interval equal to or less than twice the average connection holding time.

V.5 Connection Provisioning

Network provisioning corresponds to network resource assignment for a given static traffic demand with minimum price and minimum resource. One way of studying connection provisioning problem is through blocking probability. The blocking probabilities of a link can help us increase or decrease the corresponding link capacity.

The authors in [54] used a link decomposition approach to compute call-blocking probabilities of multicast calls over arbitrary WDM network with sparse splitting and sparse conversion capabilities. Basically, the network is decomposed into a number of path subsystems, each of which is analyzed in isolation using an approximate algorithm, and the global solution is formed by combining those solutions appropriately.

The arrival of multicast calls was modeled as a state-dependent Poisson process, and the holding time was modeled using negative exponential distribution. In order to simplify the analysis, a class aggregation technique was introduced to classify all the calls into different classes according to their properties, including multicast calls. Thus, a modified model of calls was established. Using this technique, the multiclass path steady state probability was initially simplified into a single-class path steady state probability. Then the authors showed that blocking probability of a class r call is class-independent in the modified multicast model. The multicast calls are then decomposed into some simple unicast model.

For every multicast session, a multicast tree exists. By using path decomposition algorithm, the multicast tree could be broken up into several linear segments, and each segment of the tree was considered as a subsystem. In turn, each subsystem is analyzed as a unicast network with the modified model, and the blocking probability of a multicast call can then be expressed in terms of blocking probability of each segments of the multicast tree. The simulation experiments were

performed on a 3×3 torus network, and the results show that the approximation value was close to the simulation value.

The blocking probability of multicast calls in a completely connected network (with or without wavelength converter) was addressed in [47], but by using a different approach. The authors analyze the blocking probability directly by calculating the blocking probabilities for the source to the destination with different distances. If there is no wavelength available on the link between the source to a destination, then the source sends the multicast data to one of its group member which forwards it to this destination. The intermediate and destination nodes must belong to the multicasting group. For example, node b and c are both multicasting group member for node a , and link between a and b has no free channel, then the only route can be considered from a to b is through node c . By calculating the success probability of i nodes within $(n - 1)$ -hop and introducing unsuccess probability of k nodes within $(n - 1)$ -hop, the author can calculate the success probability of success probability of j nodes within n -hop. Proceeding in thus manner, success probability of m nodes within s -hop distance can be computed. The blocking probability is obtained by using $(1 - \text{success_probability})$. Because the network is completely connected, and total number of nodes in the network is n , thus no more than $n - 1$ hops is needed. The simulation results in [47] presented the values of the call blocking probability with different number of channels per fiber, and different number of destinations in each group.

VI Multicasting Techniques in WDM Wavelength-Routing Networks

Multicast traffic in wavelength routing networks can be delivered using one of three transmission domains. In the first domain, namely, Pure IP-Multicasting, or All-Electronic Multicasting (AEM), the IP layer delivers the multicast traffic over multiple IP layer hops, where each hop is a lightpath between two routers, which may consists of multiple physical layer hops. In this context, the conventional IP multicasting schemes, described earlier in Section III, are employed. If the node is a leaf destination, the message copy has been received successfully, otherwise, a transmitter is used to transform it to the equivalent optical signal, either on the same or different input wavelength, and then transmitted to the computed output port. Although this approach capitalizes on ubiquitous IP multicasting protocols by integrating them with the WDM technology, they lack signal transparency and not all-optical.

The second domain, namely, All-Optical Multicasting (AOM), delivers multicast traffic in a *transparent* manner on an end-to-end basis. With AOM, the WDM layer is involved in the task

of handling the multicast traffic. Based on the degree of coordination between the WDM and IP layers, AOM algorithms have been developed in two environments, that is, Pure WDM-Layer and IP Multicasting Over WDM schemes. In the first approach, the job is totally handled by the WDM layer alone which eliminates any role for the IP layer while in the second approach, both layers coexist and cooperate. The main focus of this paper is on the first group of AOM and they are treated extensively in Subsection VI.1. In order to provide a complete picture of the subject, the IP Multicasting Over WDM techniques are briefly presented in Subsection VI.2.

The O/E/O conversion is eliminated in AOM; thus, the transparency with respect to data type, coding format and bit rates is fulfilled. Since this transparency is already achieved in unicast traffic, the network dealing with the various types of traffic becomes consistent which simplifies its design and operation. Moreover, duplicating multicast data in optical domain is carried out by splitting the optical signal using passive devices, which makes this operation simpler, more cost effective, and encounters less delay than performing such duplication in electronic domain using buffers.

The All-Optical Internet is still evolving because the technologies of the All-Optical devices, especially All-Optical wavelength-converters, are currently immature. Therefore, Hybrid Multicasting (HM) schemes emerged as good solutions to fill the gap between the all-electronic and the all-optical multicasting approaches. As such, the multicast traffic is carried out in the optical domain as long as it is feasible. The signal is switched into the electronic domain only in limited cases. These cases, along with the various hybrid multicasting schemes are briefly explained in Subsection VI.3. The framework for our treatment is shown in Figure 12.

VI.1 Pure WDM-Layer AOM Techniques

Similar to unicast traffic, AOM employed solely at WDM layer entails both routing and wavelength assignment. Since the joint problem is NP-Hard, it is usually treated as two problems, i.e., routing (denoted as AOM-R) followed by wavelength assignment (denoted as AOM-WA). The former problem is treated in the following subsection while the later is presented in Subsection VI.1.ii.

VI.1.i All-Optical-Multicasting Routing (AOM-R) Problem

Routing multicast session reduces to the construction of a multicast tree. In order for the basic strategies discussed in Section III for multicast tree construction in IP Networks to be adopted to work in AOM networks, the properties of such networks must be taken into account. For example, the choice of the signal replication points must be congruent with the presence of signal splitters at

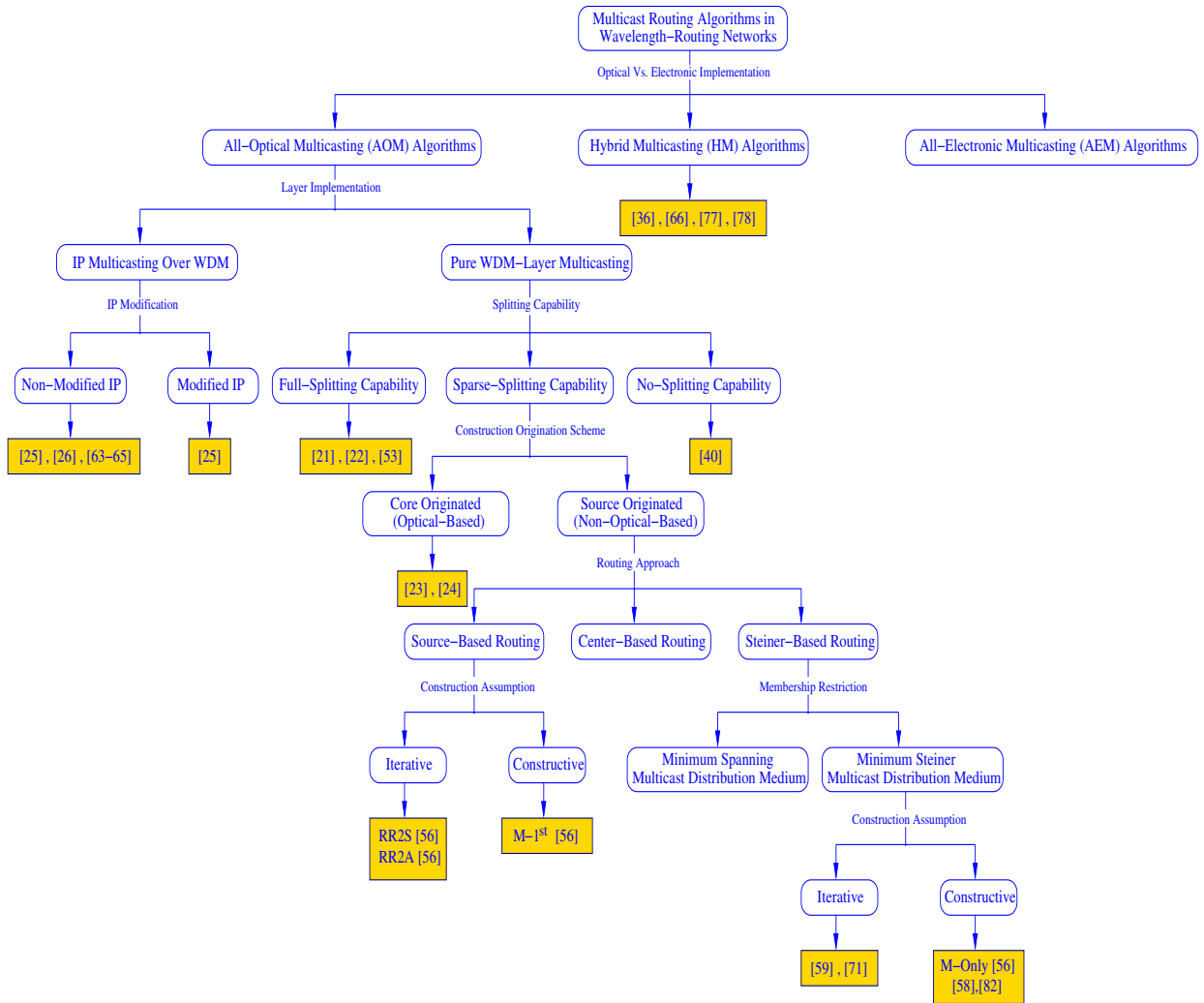


Figure 12: Classification Map for Multicast Routing Algorithms in Wavelength-Routing Networks.

such nodes. In addition, the requirement of optical amplification must be also taken into account. As indicated in Figure 12, the AOM-R techniques employed at WDM layer can be further classified to those used in networks with full, sparse and no splitting capabilities.

With full splitting capabilities, any node can be a branching node, and conventional multicast routing protocols can be used. The multicast tree in this case is called a light-tree. However, the cost of full deployment of MC nodes is too high, especially for large networks. For networks with no splitting capabilities, building a single distribution tree for multicast routing is not possible. Other routing structures may be used. For example, the use of a separate unicast session to each destination [66], or relying only on the DaC functionality to construct a trail [40]. On the other hand, in networks with sparse splitting, a single multicast tree may not be always feasible to include all destinations, e.g., [71]. Therefore, more than one tree, called the light-forest, must be used [56, 58, 58, 59].

The various techniques of AOM-R in networks with no, full and sparse splitting capability

will be discussed in the following subsections. In the rest of this paper, we use the term multicast distribution (or delivery) medium (or structure) to refer to the computed routing structure, either: tree, forest or trail. Also, in order to provide a consistent exposition while reflecting the major optical capabilities of the network model, the following notation is adopted: $[N,W,B] - S^i F^j R^k - C^x D^y - M^z$, where:

1. N, W and B: refer to the number of nodes, number of channels (wavelengths) per fiber, and number of fibers per link in the network, respectively. However, including N, W and B in the notation is optional.
2. S: represents the splitting capability of the network, such that the superscript i takes one of the symbols f , s , or n to represent *full*, *sparse*, and *no* splitting network configurations, respectively.
3. F: represents the fan-out of the splitting nodes, where the superscript j is either c for splitters with *complete* fan-out, or l for splitters with *limited* fan-out. It may also use the symbol m for a *mixed* situation in which complete and limited splitters coexist.
4. R: which indicates whether the splitting ratio of the splitting node is either *fixed* (with the superscript k equals x) or *adjustable* (with k equals a).
5. C: represents the conversion capability in the network, such that the superscript x is either f , s , or n to indicate that network has *full*, *sparse* or *no* wavelength conversion capability, respectively.
6. D: indicates the conversion degree, where that the superscript y can be either c , l or m to represent complete, limited or mixed situations, respectively, and
7. M: represents the *on-site* amplification, and has the following cases: full, sparse or no, and they are represented by letting superscript z equal f , s or n , respectively.

The above notation here can be extended to also include the *on-link* amplification. It is to be noticed that because the works we survey in this paper do not specify the availability of the on-site amplifiers explicitly, we assume that the configuration of on-site amplification is similar to that of splitting when presenting these efforts, e.g., sparse splitting case entails sparse on-site amplification. The above notation was meant to be as compact, and can even be reduced by omitting F^j and R^k parts if the network has no splitting capability, and the D^y if no conversion is employed.

A. AOM-R in Networks with No Splitting Capabilities[40]

The authors in [40] proposed a new routing algorithm for the $[N,W,B] - S^n - C^n - M^n$ network model. However, nodes are assumed to be equipped with Tap-and-Continue (TaC) capability which was described in Section V. Routing is based on constructing an optimal *trail* which originates from the multicast source node and spans all the destination nodes in the multicast group such that these nodes are visited no more than two times, that is each link is traversed at most twice.

Computing an optimal trail with the objective of minimizing the number of directed edges traversed is referred to as the problem of Multiple-Destination Minimum-Cost Trail or MDMCT. In [40], the authors followed a graph theory approach in proving that the MDMCT problem always has a solution if the network is strongly connected. They also proved that the MDMCT is an NP-Complete problem, therefore, they proposed a heuristic of polynomial time complexity in order to find a feasible trail.

The heuristic works in two steps. The first step involves the construction of a Steiner Tree for the multicast session using the Minimum Cost Path Heuristic (MCPH) proposed in [57]. This step assumes that all nodes are MC nodes. Based on the computed Steiner Tree, the multiple-destination trail is computed accordingly by re-routing around the nodes in the Steiner Tree which have an out-degree that is greater than one. The algorithm for finding the trail starts from the Steiner Tree root, i.e., multicast source, and recursively repeats at each node in the tree, referred to as the *Current* node, if this node is not a leaf node. In the downstream direction, the algorithm attempts to include all the downstream links between the *Current* node and all its children destinations. *Backtracking* is required when a leaf node is reached and there are still some destination nodes which have not been visited. In this case, backtracking is performed to the nearest branching node which has some outgoing branches that have not been visited yet. Because of backtracking, and in order to traverse as few edges as possible, the sub-trees rooted at each node are traversed according to their depths, and in ascending order.

The advantage of this method is that it eliminates the high cost of having splitting nodes, especially when the splitting ratio meant to be adjusted based on the degree of balance of the multicast tree. However, it traverses more link, and requires a longer set up time and the delay encountered by the connection becomes higher. The simulation results in [40] shows that the system degradation in terms of blocking probability when MDMCT is used is negligible.

Example

The following example illustrates the above procedure for constructing the trail for the multicast session $m = (6, \{1, 2, 3, 4, 8, 9, 10, 11, 12, 13, 14\})$ in NSFNET. The first step constructs the Steiner Tree using MCPH. According to the MCPH, the destinations are connected to the tree using the least cost path to the source node or to any already existing node in the tree. In this example, the cost function is considered to be the number of hops.

MCPH starts by adding nodes 3, 11, and 13 that are one hop away from the source node which is the only node in the tree. Then, nodes 1 and 2 are connected to the tree through node 3 while node 10 is connected through node 11. At this point, both 8 and 4 are only one hop from the tree, which is less than their distances to the source node; therefore, they are connected to nodes 1 and 2, respectively. Similarly, the remaining nodes are connected to the tree as shown in Figure 13.

The construction of the multiple-destination trail then starts. Starting at node 6 (the multicast source node), the procedure initially includes the downstream link (6,13), since it constitutes the least depth sub-tree from node 6. Because node 13 is a leaf node, and there are more unvisited members, backtracking is performed to node 6 and the next least-depth sub-tree, i.e., the one rooted at node 11, is traversed by recursively invoking the procedure at nodes 11 and 10. Backtracking is then carried out up to the source node in order to traverse the last sub-tree rooted at node 3. At this point, the sub-tree rooted at node 1 is traversed prior to the one rooted at node 2, since it is of less depth. Once the last node, i.e., node 9 is included in the trail, no backtracking is performed since all the nodes are traversed. The resulting trail is shown in Figure 13 as directed dashed arrows, where the number beside each arrow represents the sequence of constructing the trail.

B. AOM-R in Networks with Full Splitting Capabilities

The full splitting capability at all the nodes in the optical network allows easy adaptation of conventional techniques for multicast tree construction to optical networks. Therefore, a single distribution tree rooted at the source node and connects it to all the destination nodes is possible, if other required resources such as amplifiers and wavelength converters are available.

AOM-RWA in WANs [21]

Reference [21] studies the AOM-RWA problem when all nodes have complete, and fixed ratio splitting capability. However, various wavelength conversion capabilities were investigated, i.e., no/full and complete/limited. This work is important for a number of reasons. First, it introduces

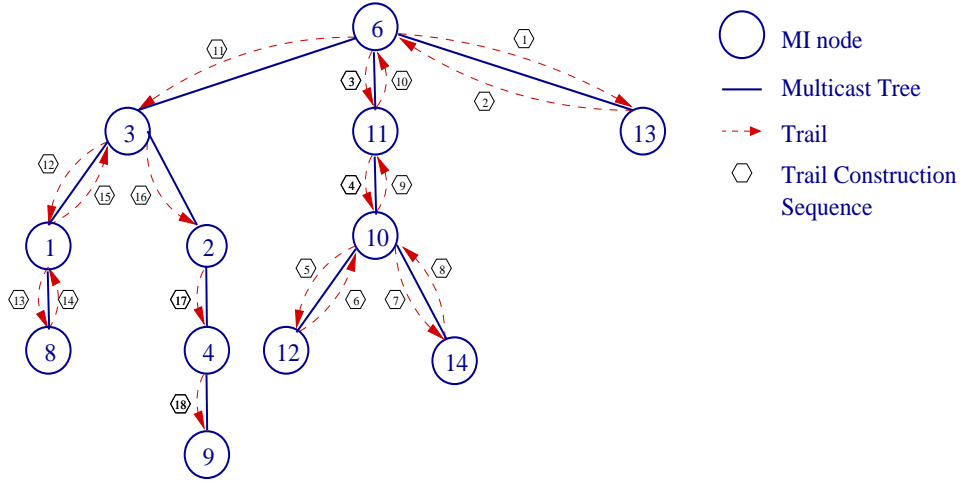


Figure 13: Computing the Multiple-Destination Trail for $m=(6,\{1,2,3,4,5,8,9,11,13\})$ in the NSFNET where all the nodes have TaC capability only.

dynamic AOM-RWA besides the static version. Second, the authors proposed two definitions for blocking policies, i.e., full and partial destination blocking policies. The paper also investigate the effect of using multiple fibers per link on the behaviour of AOM-RWA schemes. Finally, the results obtained in [21] show that the link sharing in AOM outperforms the use of multiple unicasting session.

In [21], multicast tree generation is performed using the Minimum Cost Path Heuristic (MCPH) [57]. The authors propose two schemes for selecting the multicast route. The first scheme is the *static* scheme in which the traffic demand matrix, and a set of routing trees, t_i , are pre-calculated for each multicast session, i . Assigning a routing tree to a multicast session can be performed in two fashions, namely, *fixed* or *alternate* based on the $\|t_i\|$. In the fixed option, a single tree is calculated and assigned to the multicast session, while the alternate scheme involves the computation of more than one tree per session, and the search for an available tree in t_i is performed in a certain order. Hence, fixed routing is a special case of alternate routing scheme when $\|t_i\|=1$.

Although static routing is simple, it does not make use of the instantaneous network state which results in an underutilization of the network resources and an increased session blocking probability. The second scheme, namely, the dynamic routing, solves the above problems and increases the rate of multicast calls acceptance. Unlike static routing, dynamic routing takes the link utilization into consideration when making routing decisions. Of course, this comes at the price of an increase in the computation and the communication exchange overhead.

The authors extended their approaches by defining two criteria for the multicast call acceptance, namely, full and partial destination reachability. With the first policy, a multicast session is established only when a source node is able to reach all the destinations using a single distribution

tree from t_i . This requires the availability of the required network resources (e.g., wavelengths) on this tree. This called the *Full Destination Blocking Policy* (FDBP) and its performance is characterized by the session blocking probability metric. The other session establishing policy is the *Partial Destination Blocking Policy* (PDBP) and is based on admitting the multicast connection even if some of the destinations are not reachable by the multicast tree. The destination blocking probability is used for evaluating the performance of this policy. Both policies can be used with either static (fixed/alternate) or dynamic routing schemes. However, it is worth noting that adopting any of these policies depends on the nature of the multicast application. The FDBP policy is appropriate for the applications of the collaborative nature that require the simultaneous availability of the member nodes, e.g., Teleconferencing, distributed databases and distributed computing. The PDBP seems to be a reasonable choice for less restrictive multicasting applications whose operation do not require all destinations to be connected to the multicast tree, e.g., Video-on-Demand (VoD) and newsgroups).

The simulation results in [21] reveal that the dynamic scheme outperforms both static schemes (fixed and alternate) with FDBP for single fiber, and the performance gap increases with a decreasing number of wavelengths and an increasing group size⁸. Nevertheless, both static and dynamic schemes outperform the use of multiple unicast connections. Also, the results reveal that a significant performance improvement can be achieved under the PDBP for both static and dynamic approaches. However, the static fixed scheme exhibits the best achieved performance when the PDBP is employed.

Finally, the effect of wavelength-conversion through the use of multiple fibers per link was studied. For example, full wavelength conversion can be achieved when the number of fibers per link is equal to wavelengths per fiber. The results showed a remarkable performance improvement under both static and dynamic techniques, with a leading performance in the in dynamic case. Most of this improvement is obtainable with a small number of fibers, or equivalently, a small conversion degree. A more interesting result is that even though the performance of the fixed scheme is improved with the employment of full conversion, its performance is still inferior to that of the dynamic scheme with no wavelength conversion.

Light-Trees [53]

Another cornerstone in the area of AOM is the light-tree concept introduced in [53]. A light-tree is a generalization of the lightpath in wavelength-routing WDM network. As such, the light-tree operates as a point-to-multipoint all-optical channel that provides a logical single-hop optical

⁸The group size in [21] is denoted as session fanout.

communication between the source node and the destination node(s). However, physically, the light-tree may traverse multiple links and bypass many nodes in the network. As it is the case with lightpaths, light-trees are subject to the wavelength-continuity constraint in the absence of wavelength-conversion capability. However, the treatment in [53] assume full wavelength conversion; thus, the system model is denoted as $[N,W,B] - S^f F^c R^x - C^f D^c - M^f$.

The light-tree idea makes use of a single tree to deliver one or more traffic demand to a set of destinations. Each traffic demand is assumed to be only a fraction of the total channel capacity, hence, the number of the traffic instances that are assigned to a certain light-tree depends on the capacity of the channel and on the traffic demand of each traffic instance. These traffic instances originate from the source node and can be of any type, i.e., unicast, multicast, broadcast, or mixed. Also, each receiver in the destinations set may be a destination node in one or more of these sessions, but not necessarily all.

Two main objectives of light-trees were investigated in [53], namely: (1) to minimize the average hop distance, or (2) to minimize the total number of the transceivers, for a given traffic demands. The impact of the first objective is to reduce requirement of O/E/O conversion which, according to [67], improves the network throughput.

The authors in [67] adopted a virtual topology based design approach in assuming that all nodes in the network are MC nodes. Although the primary focus in [53] was to provide a better solution to support the unicast traffic, the light-tree idea can be extended naturally for the other two main traffic types, i.e., multicast and broadcast.

The problem of light-tree virtual topology design was formulated as a Mixed Integer Linear Programming (MILP) Optimization problem. Although solutions for the unicast and broadcast traffic types only were introduced, the same concepts can be naturally extended to multipoint communication. The problem was solved with each of the previous objectives individually for the case of the unicast traffic. However, the objective of minimizing the number of the transceivers only was considered under the broadcast traffic case. The results showed that light-trees are better than lightpaths since they require fewer transceivers for both traffic types and a number of hops for unicast communication. These results regarding the number of hops were validated again in [68] which also extends these results to a number of randomly generated network under unicast traffic.

Example

Again, consider the NSFNET backbone network with three communication traffic demands for node 9. Two of them are of the unicast type: $m_1=(9,\{1\})$ and $m_2=(9,\{7\})$. However, the third

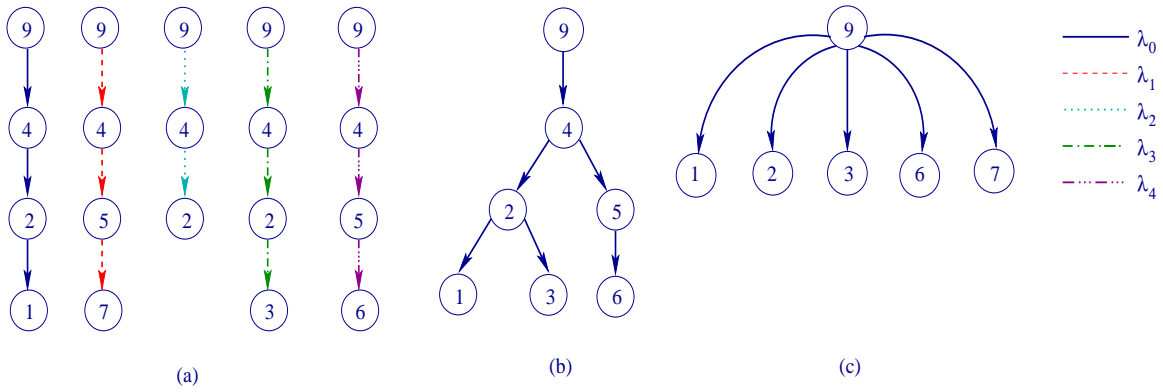


Figure 14: Significance of employing Light-Tree for routing $m_1 = (9, \{1\})$, $m_2 = (9, \{7\})$, and $m_3 = (9, \{2, 3, 6\})$ in NSFNET: (a) Lightpaths, (b) Light-Tree, and (c) Light-tree's Virtual Topology.

one is a multicast connection: $m_3 = (9, \{2, 3, 6\})$. Figure 14-a depicts the resultant lightpaths that are employed for these connections. Assume that the number of the wavelengths is sufficient over all the network links and there is no blocking due to shortage of network resources. This assumption effectively eliminates the need for wavelength conversion. On the other hand, and in order to appreciate the significance of the proposed light-tree based virtual topology, we consider the worst case scenario by assuming that electronic routing is performed at each node in the path. This requires that the lightpath to be terminated at each node of the route. With this assumption in mind, the route for unicast connection m_1 is carried over channel λ_0 and it requires three lightpaths, and hence, three electronic hops (or just hops): between 9 and 4, 4 and 2 and 2 and 1. Similarly, establishing m_2 requires 3 lightpaths: 9 to 4, 4 to 5 and 5 to 7 and it uses channel λ_1 . However, due to the absence of light splitters, the multicast connection m_3 is supported by establishing three distinct lightpaths, where each of which originates from node 9 and terminates at the individual receiver nodes, i.e., at node 2 (using λ_2 over 2 hops), at node 3 (using λ_3 over 3 hops), and at node 6 (using λ_4 using 3 hops, respectively. The total number of transceivers needed is 28 (i.e., 14 transmitters and another 14 receivers) and 5 different channels.

This number of transceivers is significantly reduced by adopting the light-tree concept. Without loss of generality, assume that the requirement for each traffic connection equals 30% of the capacity of a single channel capacity. The optical signal emitted by node 9 contains all these traffic instances and it is optically split at node 4 into two parts without being terminated at this node since it is not a destination of any of the three sessions. This eliminates the need to use the 10 transmitters and 10 receivers at node 4 that are needed previously when the lightpath based virtual topology is employed. However, a simple passive light splitter is required at node 4, instead. The two signals are independently routed to nodes 2 and 5. At node 2, which is a member in

m₃, the signal undergoes 3-way splitting. The first part of the split signal is terminated at node 2 itself while the other two parts are routed and terminated at nodes 1 and 3. Therefore, three receivers are needed at nodes 2, 1 and 3. In a similar fashion, the signal from node 4 travels all the way to node 5 where it is split into two parts: one terminated at node 5 itself and the other is routed to node 6 where it is terminated. Effectively, the total number of transmitter is one, and the total number of receivers is 5. The reduction in the required resources is extended to include the number of required channels, since a single wavelength channel is employed for these three sessions. Also, such a solution guarantees the use of a single hop and an all-optical delivery of the traffic. The actual (physical) routing light-tree is shown in Figure 14-b. However, the virtual topology induced by such a physical light-tree is depicted in Figure 14-c.

AOM-RWA for Single-Source Multicast System [22]

With $S^f F^c R^x - C^n - M^f$ system model, the AOM-RWA problem was studied in [22] with the objective of maximizing the total number of served users instead of maximizing the number of the admitted complete multicast sessions. This objective employs the Partial Destination Blocking Policy (PDBP) described earlier by accommodating partial trees. Given the network topology, and the set of multicast trees computed by any conventional algorithm, a Non-linear Integer Program (NIP) was provided for the general AOM-RWA case where the multicast sessions in the network have different sources. Due to the difficulty of solving NIP, two heuristics were proposed to support multicast with PDBP in single-source systems, which is a special case in which all the multicast sessions belong to a single source.

The first heuristic employs a linear programming (LP)-Based algorithm that consists of two Integer Linear Programs (ILP). In the first ILP, the completely served multicast sessions are accommodated, and then wavelengths are assigned to each tree. The link capacities are updated and the network topology is modified in order to exclude any link with no available wavelength. The second ILP is then iteratively executed in order to include as many users as possible from the remaining unserved multicast session. According to [22], this two-step scheme can provide a near-optimal solution.

The other heuristic is called the MAX-FIRST (MAX-1st) Algorithm. It is an iterative simple approach which accommodates the multicast sessions according to the number of their users that can be served using an available wavelength. At each iteration, the multicast session with the maximum number of *serveable* users is assigned an available wavelength, then the link capacities and the network topology are updated accordingly. Based on the new system status, the new multicast session with the maximum *serveable* users is identified, and the procedure continues

until no more users can be served.

Both heuristics are extended to allow re-routing by constructing new trees on the available wavelengths in order to accommodate more users that could not be reached according to the computed fixed trees and the wavelength availability.

The user blocking probability is used in [22] as a performance metric in the simulation experiments. Simulation results show that both LP-Based and MAX-1st heuristics have comparatively similar performance with an advantage of simplicity for the MAX-1st Algorithm. The results also show that allowing re-routing decreases the user blocking probability significantly, and 10%-20% improvement in the system performance is achieved when PDBP is employed over the FDBP counterpart.

C. AOM-R in Networks with Sparse Splitting Capabilities

Due to the complications and high cost in fabricating the wavelength-routing nodes with multicasting capabilities which result mainly from the need to use wavelength amplification, the case of full splitting capabilities may not be practical, and AOM-R schemes are needed to handle this situation.

The multicast routing algorithms that have been proposed in the literature for handling the sparse splitting situation can be classified into two basic groups based on the originating scheme adopted by the various AOM-R techniques for constructing the multicast distribution structure. The first group is referred to as the *Source-Originated* scheme because building the multicast distribution medium is initiated by the source node (i.e., multicast owner) of the multicast session regardless its nodal physical capabilities in terms of splitting, wavelength conversion and/or amplification. Therefore, this group can be also called the *Non-Optical-Capabilities-Based*, or simply *Non-Optical-Based*, technique. On the other hand, constructing the multicast delivery medium can also originate from an assisting node or a set of assisting nodes that have special optical capabilities and form a special optical structure that is referred to as the *core structure*. This core structure is employed as an original structure in the final distribution structure of the multicast session. Hence, this approach is called the *Core-Originated*, or alternatively *Optical-Based*, approach.

While Optical-Based schemes are unique for AOM in wavelength-routing networks, the *Non-Optical-Based strategy* is common with multicasting in non-optical networks. Traditionally, the *Source-Originated* routing algorithms is further divided into three main categories according to the routing approach they employ: Source-Based Routing, Steiner-Based Routing and Center-Based Routing [69]. Essentially, the Source-Based Routing approach constructs the multicast distribution medium by connecting the multicast source node to each receiver individually using the appropriate least cost (in terms of hop or delay) path in order to minimize the per source-receiver path cost. By

definition, the source-based routing scheme is applied at each sender, hence, a separate multicast distribution medium is required for each source in multicast sessions that involve multiple sources.

The objective of the Steiner-Based Routing schemes, however, is to minimize the overall cost of the multicast distribution medium, which is defined as the total cost of all its links, instead of minimizing each source-receiver pair path. These schemes are known to be NP-Complete. Considering the membership requirement of the nodes, the algorithms under this category can be furthermore subdivided into two sub-categories: Minimum Spanning- and Minimum Steiner- multicast distribution medium techniques. With the first approach, all the nodes in the routing structure are member nodes only, while the second approach eliminates such a membership restriction [70].

The objective of Center-Based Routing schemes is to provide a single shared delivery medium that supports many-to-many communication. The heart of this routing approach is the choice of a certain center node to serve as the root for the multicast medium that spans all the member nodes. The source nodes are not necessarily part of the routing structure; nevertheless, all communication must take place through the selected center node and from which it is directed to the destinations.

To the best of the authors knowledge, none of currently proposed techniques for AOM-R can be characterized as a Minimum-Spanning Steiner-Based Routing scheme or as a Center-Based Routing scheme. However, the map shown in Figure 12 includes these categories for sake of completeness and will allow future schemes to fit into it. The various techniques that are designed for Sparse-Splitting deployment are presented in the following subsections.

C.1 Source Originated (Non-Optical-Based) Multicasting Techniques: Source-Based Routing

Two schemes were employed for constructing the multicast delivery medium in a form of a light-forest, denoted by F , in the source-based routing schemes. The first scheme adopts the *iterative, i.e., improvement*, approach in which the routing medium is initially constructed as a single tree with an assumption that all the nodes in the network are equipped with splitters. The medium is then iteratively modified if any of the branching nodes, with b children, happens to be an MI node which requires it to keep one of the children only, and attempt to reconnect all the remaining nodes, denoted by R , through some other node(s) in the *Re-Join Stage*. The generalized operation of the *iterative* schemes is shown in Figure 15-a. Alternatively, the second approach is *constructive* in nature since it forms the routing medium from scratch and by checking the splitting capability of nodes during the construction stage. The Re-Route-to-Source and the Re-Route-to-Any algorithms [56] are *iterative* source-based routing algorithms while Member-First Algorithm [56] is an example for the *constructive* approach. The details of these algorithms are presented in what follows.

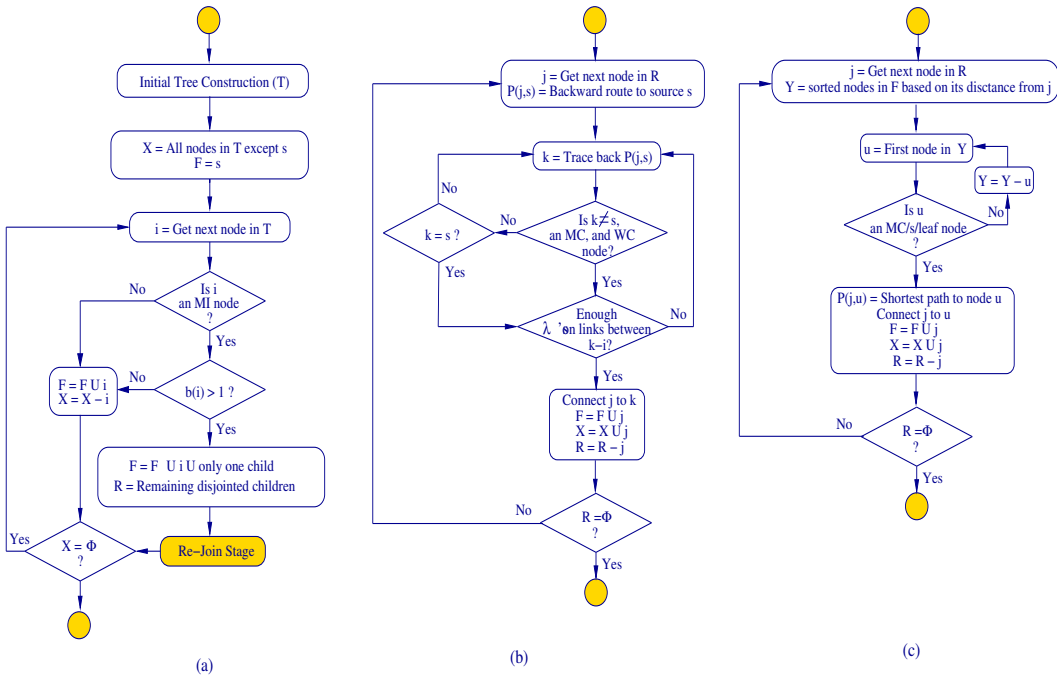


Figure 15: Flow Chart Representations for: (a) Generalized Iterative Source-Based Routing approach, (b) RR2S Algorithm, and (c) RR2A Algorithm.

Re-Route-to-Source (RR2S) Algorithm [56]

Both the Re-Route-to-Source (RR2S) and the Re-Route-to-Any (RR2A) algorithms [56] assume the following system model: $[N, W, B] - S^s - F^c - R^x - C^s - D^c - M^s$. Employing the generalized operation in Figure 15-a, both schemes, i.e., RR2S and RR2A algorithms, initially construct a single multicast tree (T) from a minimum spanning tree by pruning those links that do not lead to member nodes. Then the splitting capability of each node in T, i , is checked in breadth-first or depth-first orders and both algorithms operation terminate when all these nodes, determined by set X, are checked. However, the algorithms operation differ during the *Re-Join Stage*.

The operation of the RR2S algorithm during the *Re-Join Stage* is shown in 15-b. Basically, the RR2S re-organizes the initial light-tree, T, to form a new multicast delivery structure, F, by reconnecting every disconnected node in R, denoted by j , to an appropriate node, denoted by k , that is located along the reverse path from node j to the source node s , $P(j,s)$, on T. The candidate node k should be an MC node that is equipped with complete conversion capability. However, if such a node is not available along $P(j,s)$, then node j joins F directly at the source node s (regardless of its splitting/conversion capabilities⁹). The availability of a wavelength converter at

⁹This is possible since the source nodes are assumed to be equipped with: (1) multiple fixed transmitters, or (2) a single tunable transmitter. However, the first choice may have an advantage over the second one since it allows multiple transmissions to take place simultaneously from the source node, as shown in 16-b.

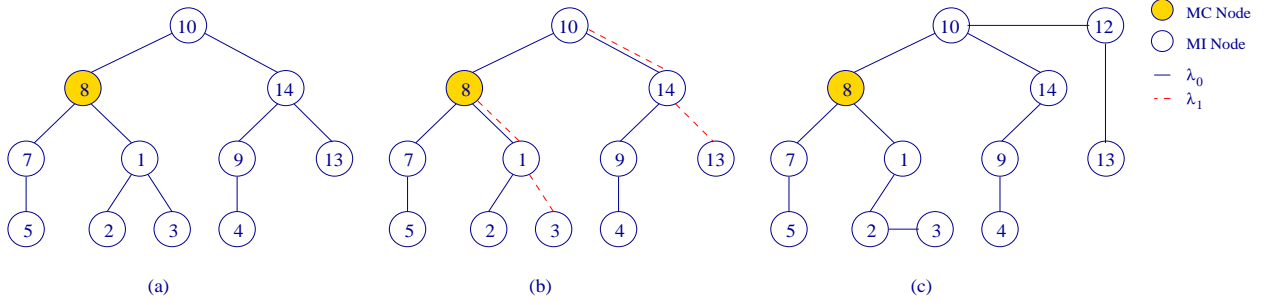


Figure 16: Example of the operation of RR2S and RR2A: (a) Initial Multicast tree (T), (b) Light-forest resulted from RR2S Algorithm, and (c) Light-forest resulted from RR2A Algorithm.

the MC node is essential because the requirement of using the same links that constitute the reverse path to the source node by all the $b(i)-1$ disconnected children from the MI node i necessitates the need of carrying out each connection over a different channel in order to avoid any signal interference and then loss. The same reasoning entails also the availability of $b(i)-1$ wavelengths on the links between nodes k and i . Although this re-joining strategy guarantees the usage of shortest paths for the cutoff members, it results in concentrating the traffic on certain links which results in exhausting the wavelengths along them rapidly.

The *Re-Join Stage* of the RR2S terminates when all the disconnected nodes in R find their alternative paths to F . The resultant multicast delivery structure, F , takes the form of a *light-forest* which is a generalization of the *light-trees*. The *light-forest* consists of a set of source-originated multicast trees that are rooted at the session source node itself and each one spans a different subset of destinations.

Example

The following example illustrates the RR2S operation for the multicast session $(10, \{1, 2, 3, 4, 5, 13\})$ in our sample NSFNET. All the nodes are assumed to be MI except node 8 which is also equipped with a wavelength converter. The initial multicast tree (T) is computed as a pruned minimum spanning tree using Dijkstra's Algorithm and it is shown in Figure 16-a. A single wavelength, say λ_0 , is assigned to T assuming its availability over all the links of T. Nodes 1, 8 and 14 are branching nodes in T. Being an MC node, node 8 is able to maintain the connections to both its children: nodes 1 and 7. On the other hand, both nodes 1 and 14 are MI nodes, which limits their capability to keep more than one of its children, say nodes 2 and 9, respectively. The cut nodes 3 and 13 are then reconnected to the new multicast delivery structure, F , using the computed reverse shortest paths on T to the source node 10 via nodes 1 and 14, respectively. In rejoining F , node 3 connects itself to node 8 which is the first MC node in its reverse path while node 13 is directly connected

to the source node 10. Because the connections from node 10 to nodes 4 and 13 share the link (10,14), the source node 10 transmits the data to node 13 over a separate wavelength, say λ_1 in order to avoid inter wavelength collision. Also, a separate channel is required over the links (8,1) and (1,3) which is achieved by converting channel λ_0 to λ_1 at node 8.

Re-Route-to-Any (RR2A) Algorithm [56]

In an attempt to achieve load balancing in wavelength usage, the Re-Route-to-Any (RR2A) algorithm eliminates the requirement proposed in RR2S of using the same backward shortest path on T by all the disconnected children of a branching MI node. Instead, the RR2A permits each of the disconnected nodes to choose its own path to reconnect itself to the multicast delivery structure, F, through the nearest node that is: (1) already in F, and (2) it is an MC, a source or a leaf node. The re-joining process in RR2A algorithm is depicted in Figure 15-c.

Although the reference [56] does not explicitly mention the requirement to have wavelength converter at the MC nodes, the importance of its existence stems from the fact that the RR2A is a generalization of the RR2S and it may happen that the shortest path to a forest node (or even the only one) is through the backward shortest path employed in RR2S.

Example

Figure 16-c depicts the multicast forest, F, after applying the RR2A algorithm to the same settings of Figure 16-a. As shown in this figure, node 3 is disconnected from node 1 and it finds that the closest candidate forest node is 2 which is kept connected to node 1. Because node 2 is a leaf node, node 3 uses the direct link with it in order to reconnect itself to F. For the other branching MI node, i.e. node 14, node 9 is again chosen to be supported by node 14 while node 13 is disconnected. Reconnecting itself to F, node 13 finds out that node 14 is the closest forest node with distance of 1 hop, but it is rejected since it is not a leaf node. The next nearest node is the source node itself and two shortest paths exist from node 13 to node 10, each of which is of 2 hops length. The first one is through node 14 and it requires the use of a different channel, say λ_1 , over the common links (10,14) and (14,13). The other possible shortest path, which is the one depicted in 16-c, is through the non-forest node 12. In the latter case, only a single wavelength (λ_0) is sufficient.

Member-First (M1st) Algorithm [56]

The objective of the Member-First (M1st) Algorithm is to take the splitting capability of the nodes into account while constructing a light-forest multicast structure that emanates from the source

node, which is unlike RR2S and RR2A which modify an initial multicast light-tree. The M1st Algorithm is designed to work with $[N, W, B] - S^s F^c R^x - C^n - M^s$ system model and it constructs one light-tree at a time such that each one of them is assembled link by link in a fashion that is similar to constructing a shortest paths spanning tree using Dijkstra's algorithm¹⁰. However, besides the shortest path feature inherited from Dijkstra's Algorithm, M1st Algorithm incorporates the nodal membership when expanding each light-tree, as will be explained in what follows.

The process of choosing the link that is to be enclosed in the constructed light-tree during the spanning stage is done through a special *Priority Queue* structure, called the *Fringe Link List(L)* which consists of the set of links traversed in the graph of the network, and are to be included in the multicast structure. Each such link is referred to as a *fringe link*. For the sake of simple referencing in this paper, the function Update_L(x) will be used in the sequel to indicate that the List L is updated by including the links that originate from node x, and it is invoked whenever a node is to be included in the light-tree.

The function Update_L(x) works as follows. It first determines the position of link (x,y) in list L according to the number of hops from the source node to node y, such that the link that lies on a shorter path is enclosed first. When several nodes are located at equal distance from the source node, then a higher priority is assigned to the link that leads to a member node (if any)¹¹.

The detailed operation of the M1st Algorithm is shown in Figure 17. The algorithm makes use of the following sets of nodes:

1. V, V' : which denotes the set of *nodes* that have been included in the current *light-tree*, T_i , where V consists of all the MC or leaf MI nodes, while the set V' includes non-leaf MI nodes.
2. UV : which denotes the set of *nodes* that have not been visited yet on T_i .
3. D^* : which determines the *members* that have not been included yet in the *light-forest*, F .

The tree expansion starts by adding the source node, s, to the light-forest and then invoking the function Update_L(s) in order to include all its outgoing links to the list L, as explained earlier. For each subsequent tree expansion, the fringe link (v,u) with the highest priority in the list L is selected. The algorithm distinguishes between the two cases where node u, which is the node to be included in the current tree, is a member node or not. In the latter case, node u is included in T_i and the list L is updated with its links. However, if node u is a member node in the multicast

¹⁰i.e., by finding the shortest path from the source node to *every* other node in the network. The M1st Algorithm is, therefore, not a prune-free method and it requires the removal of those links that do not lead to members.

¹¹Although not specified explicitly, the algorithm in [56] requires that link (x,y) is not added to L if node y has been attached to the current light-tree, and/or it is already represented in L by another link.

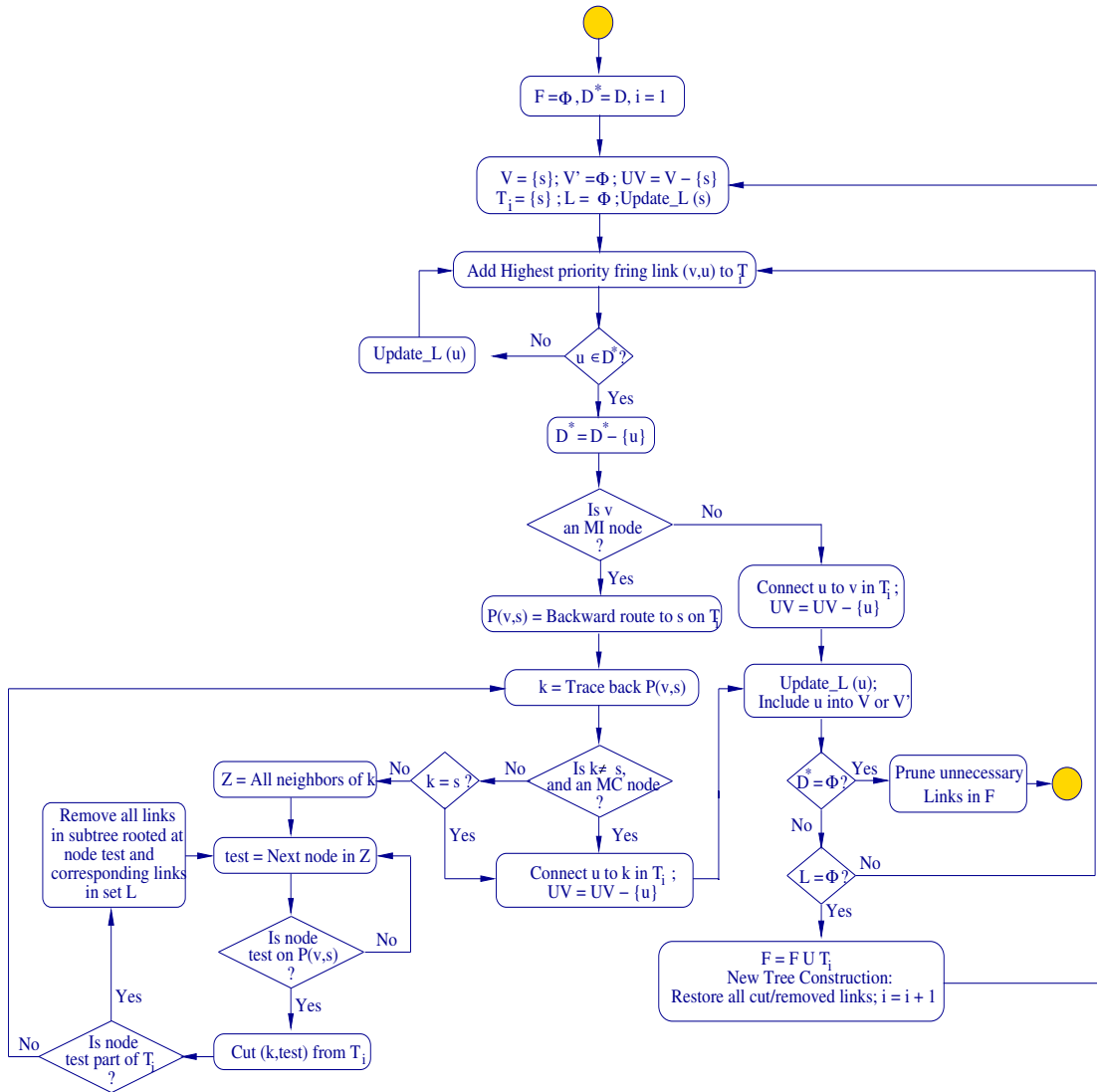


Figure 17: Flow Chart Representation for $M1^{st}$ Algorithm, where F denoted the light-forest and T_i denoted the i^{th} Light-Tree.

session, the optical-multicasting capabilities of node v , from which the tree is to be expanded, is then examined.

A direct connection is established from node u to node v if the latter is an MC node. Otherwise, the backward shortest path on T_i , $P(v,s)$, that leads to node s from node v is traced one node at a time in order to find the first MC node. If such a node is found, node u is then connected to it, otherwise a new separate connection to node s is established. A similar scheme for including the new nodes to T_i is employed in the RR2S algorithm as described earlier. However, the operation of the $M1^{st}$ Algorithm deals differently with the MI nodes along the $P(v,s)$ by employing two operations, called *cut* and *remove*, on the links as follows.

For each MI node, denoted by k , along the $P(v,s)$, the algorithm *cuts* all the links in T_i that

originate from k , except the one that leads to node u . Also, the *cut* process is performed over all those fringe links in L that also originate from k . Also, each of the cut links are examined to check whether it is part of T_i . If so, the *remove* operation is applied to all the links in the sub-tree that is rooted at node $test$ as well as the corresponding fringe links in the list L . While the *removed* links are allowed to participate in any future expansion for the current light-tree, the *cut* links are not.

More light-trees are necessary if the list L becomes empty while some of the members are not included in F yet, i.e. $D^* \neq \emptyset$. In such a case, all the links (either *removed*, *cut* or are already part of the existing light-trees) are restored and can be used for the new light-tree construction. Finally, the algorithm terminates when $D^* = \emptyset$ and the final light-forest F becomes ready after pruning all the links that do not lead to member nodes.

Example

In order to compare the light-forest produced by the $M1^{st}$ Algorithm to those generated by the RR2S and RR2A algorithms, the same multicast session $(10, \{1,2,3,4,5,13\})$ example in the NSFNET (where node 8 is the only MC node) is investigated again. The algorithm starts by including node 10 first into the first light-tree, T_1 . The list L is then initialized with all the links from node 10. Since all of them are leading to equally distant nodes that are non-members, they are assigned the same priority. Assume that the links in L are ordered as $(10,8)$, $(10,11)$, $(10,12)$, and $(10,14)$. Each one of these links is then enclosed in T_1 one link at a time and L is updated accordingly. After attaching all node 10 links to T_1 , List L will include links $(8,1)$, $(12,13)$, $(8,7)$, $(11,6)$, and $(12,9)$ in that order. Although all these links are leading to nodes which are 2 hops away from node 2, the first two links are given higher, and similar, priority since nodes 1 and 13 are members.

Link $(8,1)$ is considered next. Since node 1 is a member, and node 8 is an MC node, then node 1 is connected directly to it and fringe links $(1,2)$ and $(1,3)$ are attached to the end of list L since nodes 2 and 3 are farther at higher distance away from node 10 than all the other nodes represented currently in L . However, when node 13 (a member) is attached to T_1 via node 12 (MC node), it is connected directly to the source node over the backward shortest path $(13,12,10)$. Moreover, link $(12,9)$ is *cut* from L ; thus it cannot be used to expand T_1 . List L cannot be updated by links originated from node 13 because node 6 is already represented in L , and node 14 has been included in T_1 in a previous iteration.

Node 7 is directly connected to the MC node 8, and then node 6 is connected to node 11. However, when link $(1,2)$ is considered next, a scenario similar to that applied to node 13 is

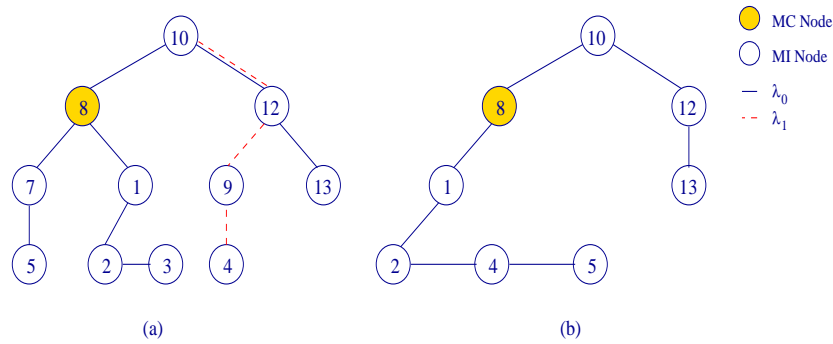


Figure 18: The multicast forest generated for routing $m=(10, \{1,2,3,4,5,13\})$ in NSFNET using: (a) $M1^{st}$ Algorithm, and (b) M-Only Algorithm.

performed. In that sense, node 2 is directly connected to node 8, and the fringe link (1,3) is *cut* from L which is updated by links (2,3), and (2,4) at the end of the list.

The algorithm proceed to include links (7,5) and (2,3) and then cutting link (2,4). At this point of time, the list L becomes empty but one member node, namely node 4, is still not visited. Therefore, a new light-tree is needed and the algorithm resumes again after restoring all the cut and used links. Applying the same steps, light-tree T_2 will consist of one single path (10,12,9,4). Because link (10,12) is shared between the two light-trees, different wavelengths are assigned for each one. Including all the members, the procedure terminates by pruning the links (10,14), (10,11) and (11,6) from the light-forest F since they do not lead to members and the resultant light-forest is shown in Figure 18-(a).

C.2 Source Originated (Non-Optical-Based) Multicasting Techniques: Steiner-Based Routing

Similar to the Source-Based Routing schemes of Subsection C.1, the Minimum Steiner-Based schemes can be described as being either *constructive*, e.g., Member-Only [56], and Virtual-Source Capacity-Priority Algorithm [58], or *iterative*, e.g., Centralized Splitting Algorithm [59], and Tabu-Search Based AOM [71]. These schemes are presented below.

Member-Only (M-Only) Algorithm [56]

The first Steiner-based source-originated routing scheme to be presented is the Member-Only (M-Only) algorithm. Like $M1^{st}$ algorithm, the M-Only algorithm aims to construct a light-forest, F, with the source node as the root for each light-tree in $[N,W,B] - S^s F^c R^x - C^n - M^s$, system model. Unlike the $M1^{st}$ algorithm, the M-Only algorithm uses the member nodes only to expand each light-tree, T_i , such that a single member is added at each iteration. Hence, no pruning is

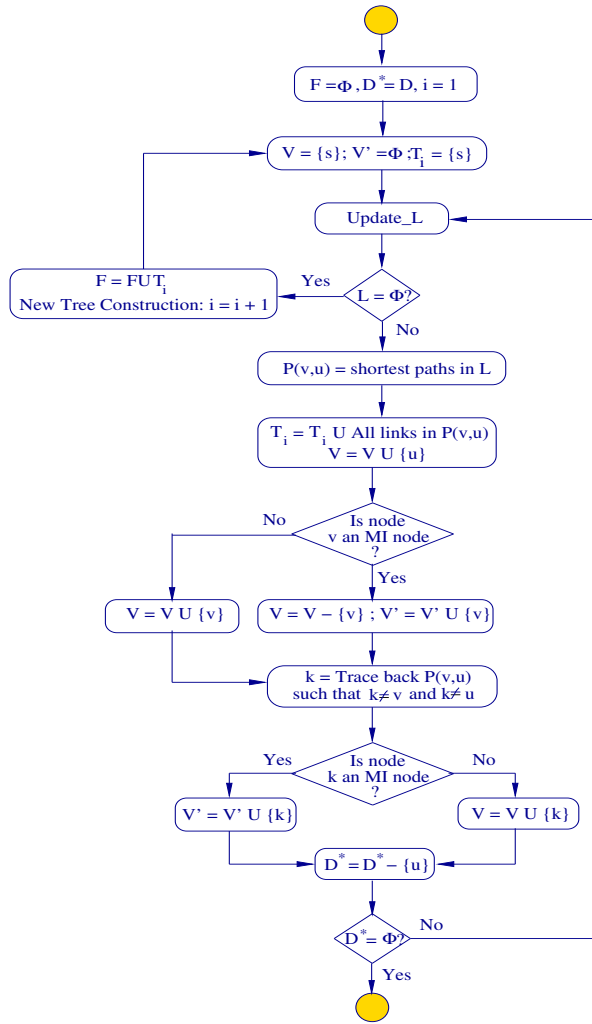


Figure 19: Flow Chart Representation for M-Only Algorithm, where F denoted the light-forest and T_i denoted the i^{th} Light-Tree.

required.

The operation of the algorithm is depicted in Figure 19 and it makes use of the same sets V, V', UV and D^* defined earlier in the $(M1^{st})$ Algorithm. Additionally, we use an ordered list, L , for the sake of easy demonstration. List L consists of a set of shortest paths sorted according to their length, and it is updated using the function `Update_L` whenever a new member is included in the current light-tree¹². The function `Update_L` finds the shortest paths from every node $u \in D^*$ to a single node $v \in V$, such that these paths do not involve any node in V' . If more than one shortest path exists from node u to node v , one of the paths is chosen arbitrarily.

The shortest qualified path among all those in list L , denoted by $P(v,u)$, is then chosen for inclusion in T_i along with all its links. Since node u is a leaf node, it is qualified to join set V . On

¹²Other schemes can be employed to select the nodes $u \in D^*$ in order to find their shortest path, e.g., the authors in [58] choose to compute the shortest path for the nodes in UV according to their distance from node s .

the other hand, node v can no longer belong to V if it is an MI node since it is not a leaf node any more; otherwise, it will remain in V . The algorithm also examines the splitting capability of every node along the path $P(v,u)$ and it updates the sets V and V' accordingly.

In this manner, the algorithm attempts to include as many member nodes as possible, and it terminates only when all member nodes are visited. If this is not the case and it happened that list L becomes empty, i.e., no more qualified paths exist for expanding T_i , then a new light-tree is needed and the procedure repeats.

Simulation results in [56] reveal that the M-Only Algorithm requires the least number of branches per forest (bandwidth) among all the remaining algorithms proposed in [56], i.e., RR2S, RR2A and $M1^{st}$. However, both $M1^{st}$ and M-Only result in almost the smallest number of channels per link, while the RR2S requires the largest number of channels per link, as well as the highest bandwidth requirement. On the other hand, RR2S achieves the shortest delay while M-Only exhibits the longest delay.

Example

The same example of the multicast session $(10, \{1,2,3,4,5,13\})$ in the NSFNET is considered again with the M-Only algorithm. The resultant F is shown in Figure 18-b.. The algorithm starts by including node 10 first into T_1 and V . Having no nodes in V' , Update_L includes the shortest paths to node 10 from all other members. The shortest paths from nodes 1 and 13 are of length 2 hops each, which are the shortest among all others. Chosen arbitrarily, the links of the shortest path to member 1, i.e., $P(10,1)$ is added to the light-tree. Currently, V includes nodes 10, 8 and 1 since node 8 is an MC node and node 1 is a leaf-MI node. List L is updated to reflect the addition of node 1 in T_1 and it turns out that each of nodes 2 and 3 is one hop away from node 1. Therefore, node 2 is arbitrarily chosen. The algorithm continues in the same manner until all the members are included in T_1 .

Virtual-Source Capacity-Priority (VS-CP) Algorithm [58]

The Virtual-Source Capacity-Priority (VS-PC) Algorithm proposed in [58] is another Steiner-based source-originated technique that constructs a multicast structure in the form of a light-forest, F , in $[N,W,B] - S^S F^C R^x - C^S D^C - M^S$ system model. In this work, the authors introduced the concept of a *Virtual Source*, which is MC node that is equipped with wavelength-converter. VS-CP Algorithm is an enhancement to the M-Only algorithm [56] described earlier. The enhancement is based onto two key observations that lead to more savings in the network resources in terms of

the number of wavelength channels per forest (i.e., the number of the links in the forest) and the required wavelengths (i.e., maximum number of wavelengths) per link.

The first observation is about the selection of the node, v , from which the light-tree is to be expanded when several qualified shortest paths of the same length exist in list L . Because such selection is performed without considering the optical capabilities of the various nodes to which the connection is to be made, it may happen that an MI node is chosen while MC or VS alternatives are available. This exhausts the expansion capability of the MI node and may force the remaining unvisited members to search for longer qualified paths (which means more links). This may also necessitate the construction of a new light-tree if no qualified shortest path is available (which means more wavelengths per link are required). Therefore, the authors in [58] propose assigning different priorities to the nodes based on their optical capabilities. The highest-priority is assigned to the VS nodes, followed by the MC nodes, the MI nodes with conversion capability (referred to as wavelength converter -or WC- nodes), and finally the MI nodes with no wavelength capability in that order. This scheme reflects the flexibility of each node in expanding the tree. The higher the priority of the node, the more is its flexibility. Therefore, the priorities assigned to the VS and MC nodes are higher than the MI-nodes. However, the significance of VS over the regular MC node stems from its ability to support more than one connection (the split signals) over the same link using different wavelengths; hence, it is assigned higher priority. Moreover, the WC-MI nodes provide more flexibility than the non-WC MI nodes in terms of their ability to carry the connection over any wavelength although they both have the same expansion capability.

The priorities of the nodes are exploited such that the node with the highest priority is selected in order to expand the light-tree. This postpones the use of the various MI nodes, and may give more chance to utilize their expansion capability more efficiently. This scheme is called the *Capability-Based-Priority Heuristic*.

The other observation, which makes the most use of the VS nodes, is related to the mechanism of constructing the new light-tree. The M-Only Algorithm starts the construction of the new tree from scratch and tries to connect the new node(s) to the source node itself. A saving in the wavelength channels (links) results if connecting these new node(s) is made to a VS node that is closer than the source node. This approach is called *spawn-from-VS Heuristic*, and it gets its name from the fact that the VS behaves like a source from which a sub light-tree is spawned.

Incorporating both heuristics into the original M-Only Algorithm, the operation of the VS-CP Algorithm is described as follows. Beside the sets V, V', UV and D^* defined earlier, the VS-CP Algorithm makes use of set Z which represents the set of *nodes* that are VS in the current light-tree, T_i . In addition to the function `Update_L` employed in the M-Only Algorithm, the function

Update L' is used to find the shortest paths between the node to be added to T_i and those nodes in Z , if any. However, the restriction about the optical capabilities of the nodes along these paths is relaxed as it is the case with Update L .

Like the M-Only Algorithm, the VS-CP Algorithm attempts to include as many destinations (each denoted by u) as possible and it deals with the member nodes only; thus, it is a prune-free procedure. Also, each iteration of the algorithm starts by updating the List L with the links of the node that has been added in the previous iteration. Yet, unlike the M-Only Algorithm, the Capacity-Based Heuristic is invoked to break the tie between those paths of the same smallest length in L such that the link that leads to the highest priority node is chosen. Once this node, v , is determined, the algorithm investigates the possibility of connecting the destination node u to it if and only if the length of this path, namely d_{uv} is less than or equal to its distance from the source node, namely d_{us} . If this is the case, the algorithm proceeds similar to the M-Only counterpart by connecting node u to node v , and updating sets V and V' with the various nodes on this path, as shown earlier. Additionally, set Z includes all the VS nodes in the path between nodes u and v .

On the other hand, if the path¹³ between u and s is shorter than that between nodes u and v , or even when there is no node v through which a connection may take part, then the Spawn-From-VS Heuristic is invoked. It simply determines the closet node, $w \in Z$, to u using the function Update L' . Node w is chosen to connect node u to it if and only if the distance from u to w , i.e., d_{uw} , is less than that from node u to s , i.e. d_{us} . Otherwise, the connection is established all the way back to node s . In both cases, the optical capability of all the nodes along the chosen path is examined and sets V, V' and Z are updated accordingly.

When any of the destinations cannot be included in T_i , which is the case determined by the emptiness of list L' when D^* is not empty, T_i is attached to the light-forest, F , and the algorithm is repeated for the new tree construction. Eventually, the algorithm terminates when no more destinations are to be added to F . The operation of the VS-CP algorithm is depicted in Figure 20.

Comparing the performance of the VS-CP Algorithm with that of the M-Only Algorithm, simulation results in [58] show a considerable saving in the network resources, in terms of number of wavelengths per link and the number of links per light forest, especially with large group sizes.

Example

In this example, we compare the light-forest construction of the multicast session $(10, \{1, 2, 3, 4, 6, 7, 9, 13\})$ in the NSFNET using the VS-CP Algorithm with that generated by the M-Only Algorithm. Node

¹³Such path may involve MI nodes

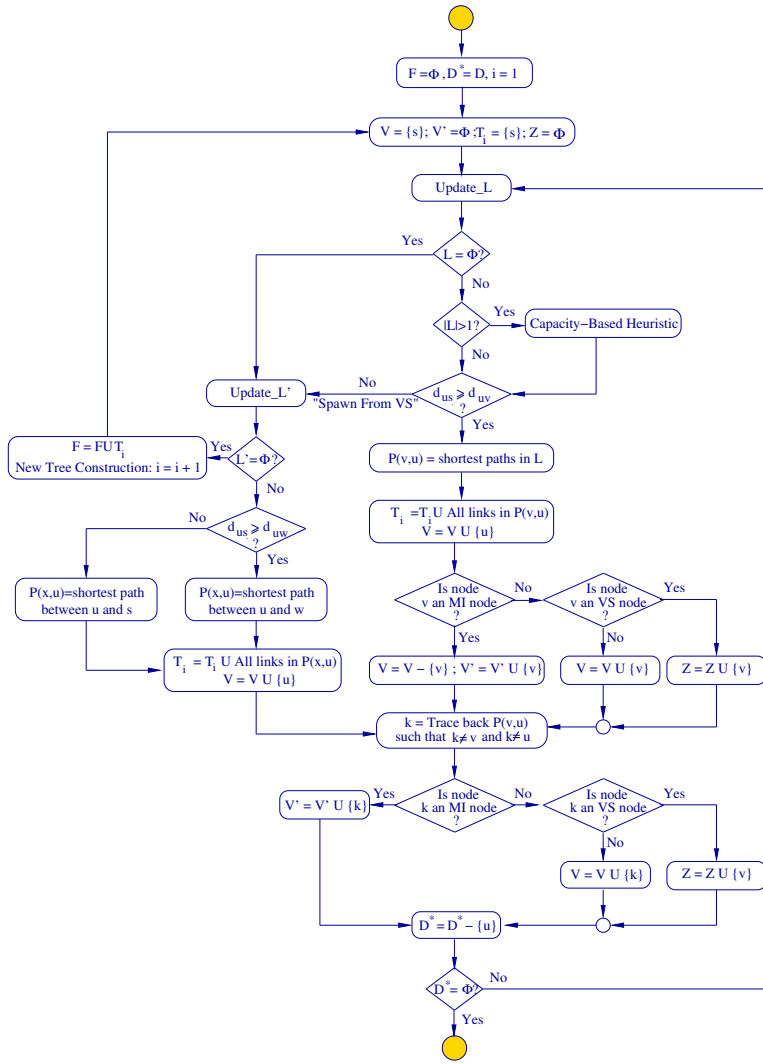


Figure 20: Flow Chart Representation for VS-CP Algorithm, where F denoted the light-forest and T_i denoted the i^{th} Light-Tree.

6 is an VS node, node 14 is an MC node while the others are regular MI nodes. Figure 21-a shows the light-forest, F , generated using the M-Only Algorithm. In this context, nodes 13 and 9 connect to source node 10 through nodes 12 and 14, respectively. These arbitrary choices yield to an unnecessary use of two extra wavelength channels in the time that both nodes could be connected through the MC node 14. The M-Only Algorithm continues in the manner described earlier till all the members are included in T_1 except node 1, which cannot be connected to any of its neighbor nodes 2, 3, and 8 because these MI-nodes were already used in expanding the tree. Hence, a new tree, T_2 , is needed in order to connect node 1 to source node via node 8 using different wavelength.

When the VS-CP Algorithm is applied, initially both nodes 13 and 9 choose node 14 to connect to the tree T_1 , since they both are 2 hops away from node 10 and node 14 has priority higher than node 12. Then, node 6 is considered next and it is connected to node 13 which is 1 hop away.

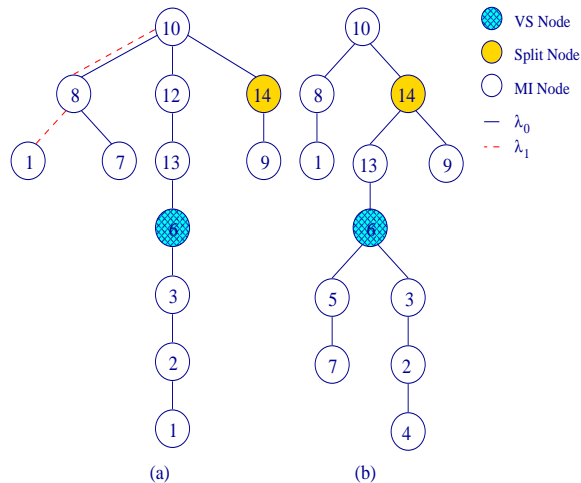


Figure 21: Constructing light-forest, F , for $m=(10,\{1,2,3,4,6,7,9,13\})$ in NSFNET, using (a) M-Only Algorithm, and (b) VS-CP Algorithm.

Similarly, nodes 3, 2 and 4 are added to T_1 in that order. At this stage, both nodes 1 and 7 have not been included yet. Node 7 is of 2 hop away from node 10 and 6. It chooses node 6 to connect itself to¹⁴ which allows node 1 to connect itself directly to node 10. The resultant F is shown in Figure 21-b. Not only one wavelength per fiber is saved with the VS-CP Algorithm, but also the number of links has been reduced.

Centralized Splitting Algorithm (CSA) [59]

The Centralized Splitting Algorithm (CSA) is a Steiner-based source-originated routing scheme that it designed to work in the system model $[N,W,B] - S^S F^C R^x - C^S D^C - M^S$. CSA takes care of the power-budget requirements when constructing the multicasting delivery structure. As such, CSA aims to construct a minimum Steiner tree that achieves an efficient utilization for the network resources (in terms of the maximum number of required wavelengths, and the number of used wavelength channels), and a reasonable delay while achieving small power loss in order to maintain the delivered optical signal above the sensitivity threshold.

The strategy presented in [59] is *adaptive* in nature. An initial multicasting delivery structure is constructed first using the M-Only Algorithm [56] and without considering the power-level requirements. since this initial structure, denoted by F_{M_Only} , may not satisfy the receiver sensitivity requirements, it is therefore, adjusted by following certain set of rules in order to minimize the maximum power loss.

The rules for reconstructing F_{M_Only} are based on the following two main observations. The

¹⁴Although the algorithm operation suggests that node 6 should connect to source node 10 as its distance to source node 10 equals its distance to node 6, we choose this selection in order to present the example easily.

first observation is related to the effect of signal splitting dimension (i.e., performing splitting *in-depth*, or *in-breadth*) on the signal power loss. While it is true that power loss is multiplicative when the splitting nodes are concatenated in a particular (sub) light-tree, such an effect is diminished when the splitting capability of a single node is increased, and fewer splitters are concatenated. For example, assume nodes x and y are two splitting nodes with splitting fan-outs A and B, respectively, and that node x is the parent for node y in the light-tree, either directly, or through some other non-splitting node(s). Neglecting the signal attenuation due to light propagation in fiber, the power at the output of node y is $\frac{1}{A \cdot B}$ of the total power inserted at node x. On the other hand, and without loss of generality, if the fan-out for node x is increased from A to A+1, then the growth in the power loss due to splitting, i.e. $(\frac{1}{A+1} - \frac{1}{A})$ of the input Power, becomes small as A increases. These two observations suggest that increasing the number of branches spawned by an MC-node in the light-tree is more preferable from the point of view of power loss than concatenating various MC nodes to each other. Therefore, the authors were motivated to propose the concept of the *Centralized Splitting Node*, in which a splitting node is chosen (based on a certain criterion) to replace a set of concatenated splitting nodes.

The second observation concerns the splitter location with respect to the root of the light-tree. Although signal distribution can be more balanced if the light-splitting occurs near the root, this has a side-effect of increasing the probability of exposing more member nodes to this power reduction, which will make them unable to detect the signal. Hence, what can be called *As Late As Possible* (ALAP) splitting scheme seems to be more attractive for the power-level reduction issue¹⁵.

The details of the CSA are shown in Figure 22. It starts by applying the M-Only Algorithm to construct F_{M_Only} . Then, the modification stage starts and it is performed on each light-tree, T_i , in F_{M_Only} that has concatenated splitting nodes. For the sake of simplicity of the algorithm presentation, let the variable n represent the number of light-trees in F_{M_Only} , and the set κ consist of all the branching (splitting) nodes in T_i .

The CSA begins by attempting to find a centralized splitting node, h, for T_i . Node h is chosen such that it is any MC-node in the network (thus, it can be chosen from outside the set of nodes in T_i) that has the smallest average distance to the *members* in the sub-tree. In order to find h, a path structure, denoted as the *Main Path* (P), is determined as the shortest path between the

¹⁵A fourth observation was made in [59] regarding the optical capability of the nodes. It was stated that the average power loss is smaller if attaching a new node to the light-tree is done through an MC node or a leaf MI node than if it is done through non-leaf MI-node. However, such a connection to a non-leaf MI-node does not achieve the all-optical requirement, therefore this observation is excluded in our treatment.

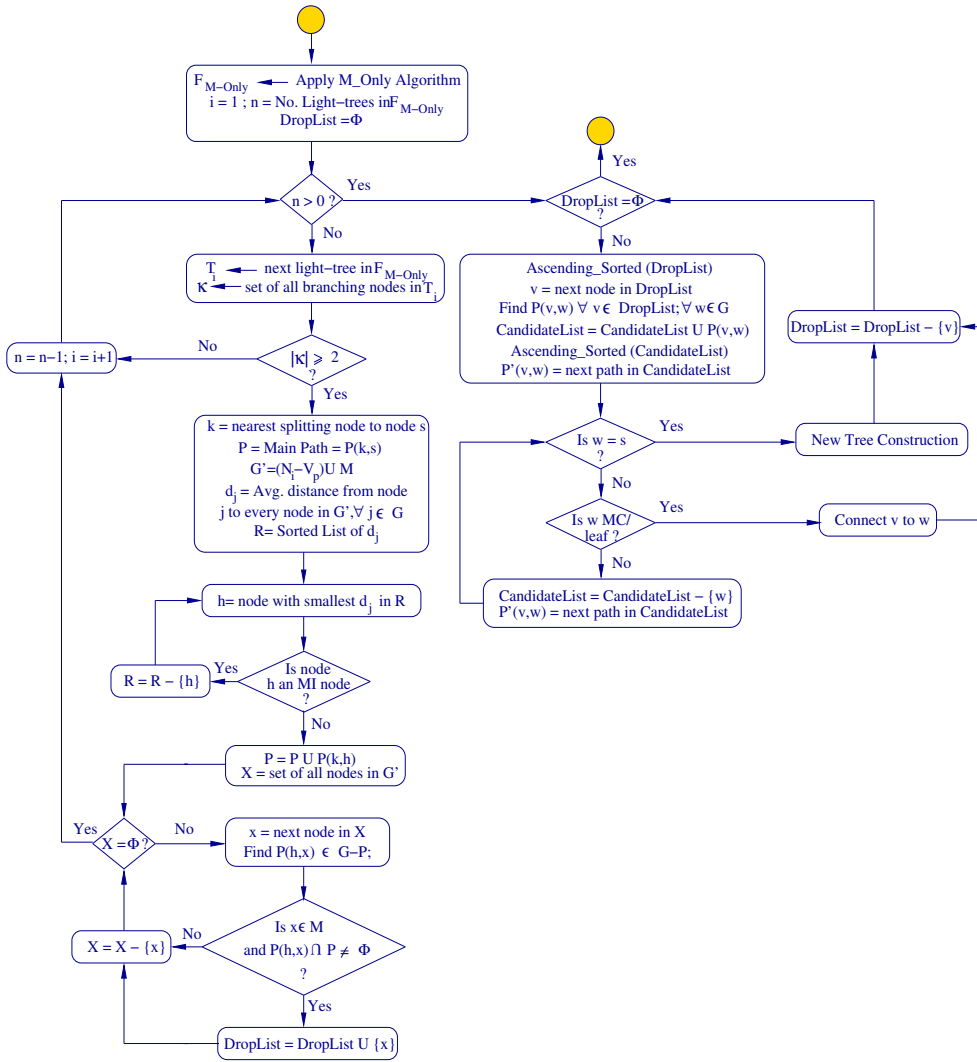


Figure 22: Flow Chart Representation for the operation of the Centralized Splitting Algorithm (CSA), where M : members set, $T_i = (V_i, E_i)$: i^{th} Light-Tree, $P = (V_p, E_p)$: Main Path

source node, s , and its nearest splitting node, k . Also, a special set of nodes, G' , is computed such that it consists of all the nodes in T_i after excluding those non-member nodes that lie on P . The average distance between all the nodes in the network and each node in G' are computed and node h is chosen as the MC node with the smallest computed average distance.

By determining node h , the next step involves modifying T_i such that two requirements are achieved: (1) light-splitting at node k is removed and is *pushed* away from the root to node h , and (2) node h represents a new branching node in the modified tree where all the remaining *members* are reconnected to. The first requirement is achieved by extending the main path P to include node h through the shortest path from node k . The second requirement, on the other hand, is fulfilled through establishing shortest paths connections from node h to each *member* such that

these connections do not involve nodes on the extended P ¹⁶. If this restriction is not satisfied, the member node is added to a special list, called DropList.

After checking each light-tree, CSA attempts to reconnect those members in DropList (if any). Each member, v , is considered at a time according to its distance from the source node, s , in an increasing order, and then the shortest path from node v to all other nodes in the network is found. Node v is connected to the nearest node, w , if and only if w is an MC-node or an MI-leaf node, or in the worst case, to s by constructing a dedicated light-tree.

The simulation experiments in [59] compared the CSA with the M-Only Algorithm [56] using the NSFNET. The results show that the CSA can achieve bandwidth utilization similar to that achieved by the M-Only, while it can reduce the power loss by 17% when the multicast traffic generation rate is high and group size is large. This comes with an insignificant increase in the algorithm complexity, the average delay between the source node to destination nodes, and wavelength usage.

Example

The following example illustrates the CSA operation in constructing the light-forest for $m = (10, \{1,2,3,4,6,7,8,9,12,13\})$ in the NSFNET, where the MC-nodes set includes nodes 1, 2, 8 and 12. The Initial light-forest, F_{M_Only} , constructed by the M-Only Algorithm is depicted in Figure 23-a, and it consists of two light-trees. One of the light-trees (denoted by T_1), which is the one that is sub-rooted at node 8, is modified since it has two concatenated splitting sites, i.e., at nodes 8 and 1, while only one light splitting is performed at the other light-tree which is spawned from node 12.

In order to modify T_1 , the main path (P) is initially determined by the link (10,8) since node 8 is the nearest splitting node to the source node 10. The average distance, d_j , from every node, j , in the network to the nodes in $G' = \{1,2,3,4,6,7,9,12,13\}$ is then computed, and is shown in Table 1. Node 4 is then selected as the centralized splitting node, h , since it has the least d_j and it is an MC-node. The main path is then extended to include node 4 along the path: (4,5,7,8). The shortest paths from node 4 to all the remaining unvisited members, i.e., nodes 1, 2, 3, and 6, are then computed and node 2 is directly connected to node 4 while both nodes 1 and 3 are connected to node 4 via node 2. Since the shortest path of node 6 passes through P , node 6 is added to DropList and it will be the only node in that list. When DropList is eventually considered, node 6

¹⁶This restriction is trivial since P is part of the modified tree and the only node on P , beside h , that can support new connection(s) through it (by light-splitting) is node k which contradicts the main purpose of pushing the splitting process away from the root.

d_j	d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8	d_9	d_{10}	d_{11}	d_{12}	d_{13}	d_{14}
Value	1.88	1.77	1.77	1.66	1.88	1.77	2.33	2.22	2.0	2.22	2.4	2.0	2.0	2.11

Table 1: Average distance, d_j , computed for every node, j , in the network from the nodes in $G'=\{1,2,3,4,6,7,9,12,13\}$

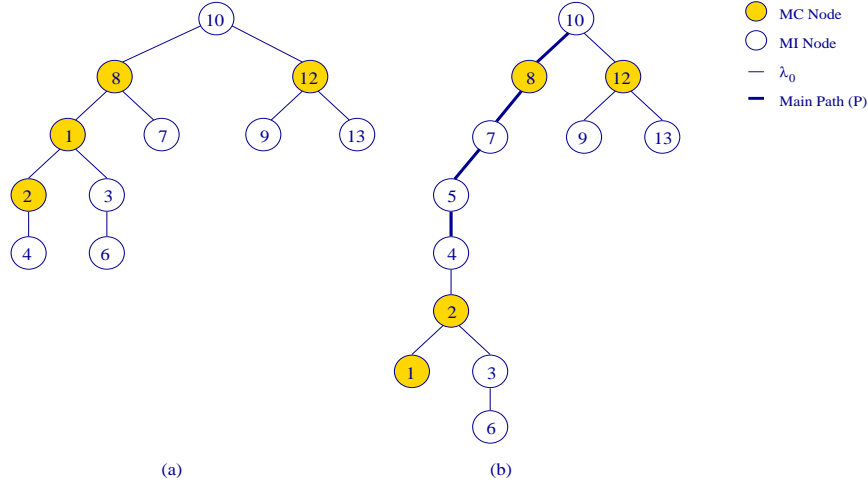


Figure 23: Computing the light-forest, F , for $m=(10,\{1,2,3,4,6,7,8,9,12,13\})$ in NSFNET using CSA: (a) F_{M_Only} , and (b) Modified light-forest generated by the CSA.

finds its way either to T_1 via node 3, or to T_2 via node 13. Arbitrarily, node 6 chooses to connect to T_1 and the resultant light-forest is depicted in Figure 23-b.

Tabu-Search Based AOM (TS-AOM) [71]

The algorithm proposed in [71] makes use of the Tabu-Search scheme in proposing a heuristic for computing a single multicast light-tree that reaches all the destinations while minimizing the number of links in the $S^S F^C R^x - C^n - M^S$ system model. Generally, Tabu-Search (TS) scheme is an *improvement* local search algorithm. As such, it starts with an initial solution, either chosen arbitrarily or computed by certain procedure, and then at each iteration it searches for a better solution in the neighborhood of the current one. Two solutions are considered to be neighbors if one solution is obtained from the other through some well-defined operations. A set of forbidden actions or mutations, called a Tabu List, is also considered while selecting the solution. Hence, the goal of the TS Algorithm is to allow *good* moves in each iteration without revisiting solutions that have already been considered. Because the TS Algorithm does not guarantee an optimal solution, it terminates when a certain stopping criterion is satisfied.

Within this context, the TS-AOM heuristic [71] considers the light-tree computed at each

iteration as a solution. Two solutions are computed using the MCPH and a modified version of the M-Only Algorithm¹⁷, and the better one is chosen as the initial solution. During each iteration, the neighborhood of the current solution is defined with the assistance of two operations, i.e., *Insertion*, and *Deletion*, that are performed over all the nodes, except the source and the destinations. The *Insertion* operation is carried out if the node under consideration is not part of the current solution, and is done as follows. A path is established between the new node and only one node in the current solution. This computed path is called the *disjoint path*. Connecting to an MC or a leaf node in the current solution is given a higher priority, otherwise, the connection is made to an MI node. Once the disjoint path is established, the connections from the new node to the remaining nodes in the current solution is done (if possible) one at a time and starting with the nearest. With each added node, its path in the current solution is pruned in order to maintain the tree structure. On the other hand, the *Delete* operation is performed on each node that is part of the current solution. When a node is to be deleted, the pruning operation is performed to delete its path that leads to the source node and all its outgoing links in the current solution. The cut children are then reconnected to the source node. The reader is encouraged to refer to [71] for more details of the TS-AOM heuristic.

An ILP was developed in [71] for the sake of comparison with the TS-AOM heuristic. Simulation results show that the TS-AOM heuristic is able to determine a solution that is within 10% of the optimal solution almost all the time, and within 5% in about half the time.

Power-Aware Multicasting (PAM) Algorithm [82]

In [82], the AOM Routing and Wavelength Assignment on already dimensioned wavelength routed networks is investigated while taking the optical power impairments into consideration. This problem is called RWA with Power Aware Multicasting (RWA-PAM). Unlike [81], the problem in [82] is a connection provisioning problem during the network operation phase.

Only subsets of the nodes are MC nodes, while no wavelength conversion capability is assumed in the network. Therefore, the system model is characterized as: $[N, W, B] - S^S F^C R^x - M^S$. Furthermore, only three sources of power loss are considered in [82], namely, loss due to fiber attenuation, splitting loss at tree branch nodes, and tapping loss at each node. Other sources of power loss (e.g., due to multiplexing, demultiplexing, and crosstalk) can be easily incorporated in the proposed solutions.

With the objective of minimizing the session blocking probability, the problem is first formu-

¹⁷In the modified M-Only Algorithm, the algorithm resorts to MI nodes whenever a new light-tree is to be generated; therefore, a single light-tree is constructed.

lated as a Mixed-Integer Linear Program (MILP). Due to the fact that some power-constraints are non-linear, the formulation was linearized using a unique set of developed mapping schemes. Due to space constraints, the reader is referred to [82] for more details about the MILP formulation.

Because the MILP formulation is not scalable to large network sizes and traffic volumes, an efficient greedy heuristic, called **PAM** Algorithm, is introduced in order to solve the problem in a fast and efficient manner. The heuristic decomposes the RWA-PAM into three subproblems, namely, Routing (**R**), Wavelength Assignment (**WA**) and Power Assignment (**PA**) subproblems. Although solved separately for sake of simplifying the solution, the heuristic considers the impact of these subproblems on each other by using a special cost function for the network links. In addition, the PAM heuristic has two characteristics. First, the algorithm deals with the sessions in a parallel fashion, rather than sequentially, such that all (or a subset of the sessions) are treated together. This accounts for the interaction between the sessions which can reduce blocking due to *service disruption*¹⁸. Second, since assigning the minimum power to each light-tree does not necessarily produce the best solution, the power assignment module of the heuristic adopts a semi-random scheme that is governed by a set of probabilities for determining the best combination of power levels at the source nodes. Based on the results reported in [82], these two characteristics prove to provide better solutions over the minimum power assignment and sequential treatment counterparts.

The link cost at any time is defined as the ratio of the maximum number of wavelengths that can be supported and the number of free wavelengths over the link. Due to power constraints, not all available channels are usable. Hence, the maximum and free number of wavelengths are determined not only by the wavelengths availability, but also by the signals ability to reach the first Optical Amplifier, if any, and still be detectable. In this sense, the number of (maximum or available) usable channels is the number of available channels or the result of dividing the available power over the link by the minimum power that is needed to detect the optical signal by the first optical amplifier, whichever is minimum.

The basic operation of PAM algorithm is depicted in Figure 24 and its output is the number of admitted sessions, their corresponding RWA, and power values. The core operation of PAM algorithm is the RWA-PAM module. The input to the RWA-PAM module is the set of sessions to be constructed, the sharing degree and the construction mode. The sharing degree represents the number of sessions to be considered at each algorithmic step. If the sharing degree is 1, the solution is obtained by adding one light-forest at a time, i.e., in a sequential manner. Otherwise, RWA-

¹⁸A service disruption is a case that results when adding new connection to the network causes the power constraints of at least one already provisioned connection to be violated.

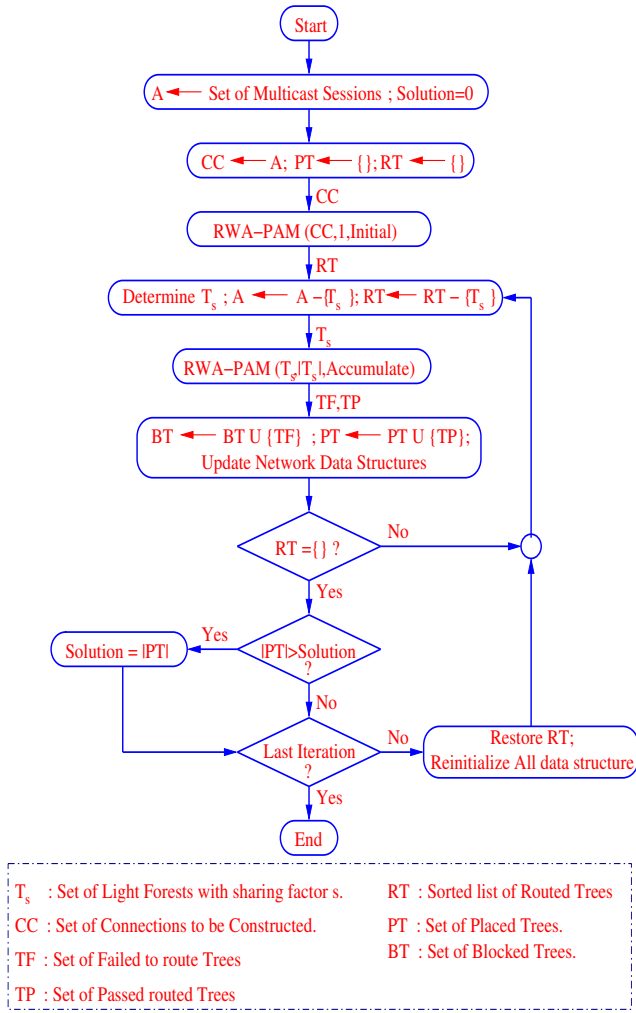


Figure 24: PAM Algorithm Operation.

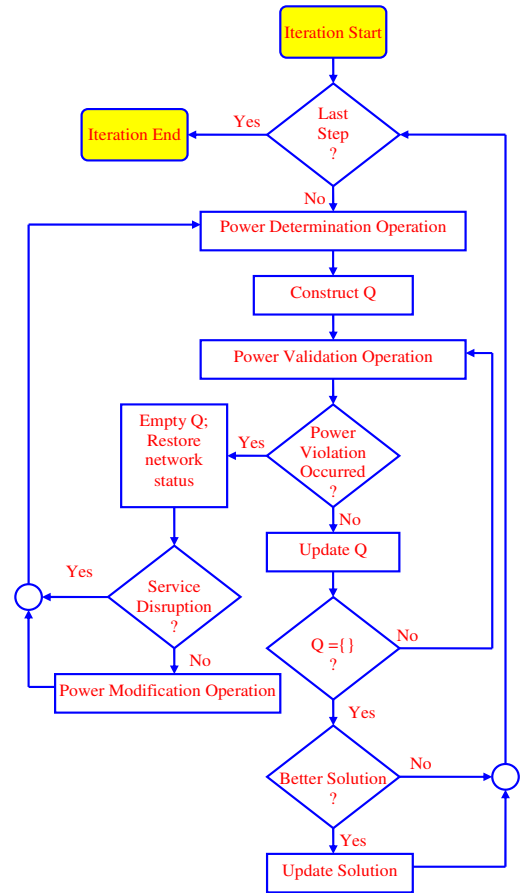


Figure 25: Power Assignment Module

PAM considers more than one light-forest. Moreover, two construction modes are defined, namely, *Initial* and *Accumulate*. In the first mode, each session's light-forest is constructed using the initial setup of the network, i.e., as if no other sessions exist. In the Accumulate mode, solutions are conducted on the current network status¹⁹. Output of RWA-PAM module is the set of provisioned and blocked sessions.

PAM algorithm consists of two phases. In the first phase, called the Initial Phase, RWA-PAM stage operates at the Initial mode in order to identify those sessions that are blocked purely due to the physical setup of the network, and eliminate them from further investigation. The second phase of PAM algorithm is the Iterative Phase. During each iteration of this phase, the RWA-

¹⁹Current network status refers to the current power values at each point in the network and the wavelength availability.

PAM stage operates in the Accumulate mode and the number of sessions considered depends on the sharing degree. The solution found from each iteration is computed with respect to the current network status which is re-initialized before the next iteration starts. The best solution out of all the iterations is selected to be the final solution.

RWA-PAM module performs three basic operations, namely, (1) Routing, (2) Wavelength Assignment, and (3) Power Assignment. Routing is performed using an extended version of the Member-Only Heuristic (M-Only) [56] by employing k-shortest paths for constructing the tree instead of a single path. In this context, the member node is connected to the tree through the path that causes the least increase in the network cost and which has sufficient power to reach the destination. Wavelength Assignment is performed using the First-Fit scheme in which the first available wavelength over the tree links is chosen. The details of the Power Assignment (PA) module is depicted in Figure 25.

The PA module makes use of a queue structure, called Q . The queue consists of unique entities of the links identity. A link becomes part of the queue if it is part of the current light-forests under investigation and/or it is part of the affected links. Multiple traversals of the link is permitted by the algorithm. However, Q contains at most one instance of the link at each algorithm step.

The PA module runs for a certain number of iterations and it consists of three main operations, namely, Power Determination Operation, Power Validation Operation, and Power Modification Operation. Given the power values of the already provisioned sessions, these three operations work together during each iteration in order to determine the best power value to be launched at the source nodes of the sessions under investigation. The first operation determines the power value set (called Power Vector) at the source nodes. The second operation determines if this Power Vector is valid over the link, i.e., it does not result in any power constraint violation, while the last operation is needed only to determine if a gain drop can be tolerated in the network.

In the first iteration of the PA module, the Power Vector is initialized with the minimum power values at the source nodes. Q is then populated with all the outgoing links from the source nodes of the current sessions under investigation. The order of adding these links is immaterial. The module proceeds on a link by link basis, starting from the link at the queue head and it checks if power levels are valid on every points on this link. If so, the link is removed from Q while its outgoing links are added to Q , if (1) they have not been already included in Q , and (2) they are part of the current sessions light-trees. The PA algorithm continues with the next link in Q and it stops when Q becomes empty.

If the Power Validation Operation indicates that a gain drop occurs, but no service disruption is encountered, then the Power Modification Operation is invoked. In this operation, the algorithm

identifies the set of all sessions on the link where a gain reduction took place, and we refer to them as the affected sessions set. These sessions include the old placed sessions as well as the new ones. It then continues its normal operation. However, queue Q is now populated with links from the affected sessions set.

If power violation is encountered at any point in the network, Q is emptied, the network status is restored, and the Power Determination Operation is called in order to determine the new Power Vector. In this operation, the power violation type is identified, and the set of sessions that are involved in this power violation is determined to be the new sessions on the link where the violation occurs. From these sessions, the operation then determines, randomly, the sessions to be blocked from participating in the next iteration of the PA module. For each remaining session, the module determines, randomly, whether to increase, decrease or not change the current power based on the power violation type. This randomness is governed by certain probabilities that are carefully chosen to reflect the method used to resolve this power violation type. For instance, if power violation is due to the power level dropping below P_{Sen} , then it is more probable that increasing the power value might solve the problem. Therefore, increasing power value is given high probability in this case. On the other hand, power violation due to service disruption is not necessarily always due to high-power input. Therefore, we assign moderate values to the probabilities of increasing the power in the case of Service disruption. Finally, this probability is low in the case of the total power exceeding P_{Max} .

The numerical results obtained by PAM Algorithm were compared against the optimal solutions of the MILP formulation. The results obtained were near optimal and in some cases they were optimal which indicates the robustness of the PAM Heuristic. In addition, PAM algorithm was used to investigate the impact of the power constraints on upgrading the networks. In this context, the results showed that increasing the number of channels may reduce the system blocking probability if the main source of call blocking due to shortage in the number of available wavelengths. However, such performance improvement reduces and then stops when violation of power constraints becomes the main source of call blocking and the OA placement [81] becomes the bottleneck for the PAM algorithm. Hence, increasing the number of channels does not improve the system blocking and these channels are wasted.

C.3 Core-Originated (Optical-Based) Multicasting Techniques

The set of routing algorithms that belong to this category is based on employing a special *core structure* in constructing the multicast delivery tree. The core structure, denoted by CS, connects a subset of nodes, called core nodes, who have special optical capabilities in terms of light-splitting

and wavelength conversion²⁰. The multicast session is then established with the assistance of this core structure. Although the connection is still rooted at the source node, the process of establishing the multicast delivery structure does not originate from it, and instead it originates from the CS²¹. Currently, two different approaches belong to this category of algorithms: [23] and [24]. They differ in their implementation of the CS concept in terms of the nature of the intra-CS connections (i.e., the connections established between the various core nodes which can take the form of a tree, or a collection of interconnected unicast connections, etc), the membership requirements of the core nodes and the time of establishing the core tree.

Virtual Source (VS-) Based Algorithm [23]

The combined optical capabilities of light-splitting and wavelength conversion are exploited again in the Virtual Source (VS-) Based Algorithm proposed in [23]. The system model proposed in [23] is attributed as $[N,W,B] - S^s F^c R^x - C^s D^c - M^s$. A single CS that is shared among all the *active* multicast sessions is pre-established before starting any multicast session, and connects all the virtual sources, VS, regardless of whether they are member nodes, or source nodes in any active multicast session, or not. Such CS may include some non-VS nodes that lie on the established connections between each pair of virtual sources. These connections in the CS are carried out over unicast connections using a dedicated wavelength for each VS-to-VS light-path; thus, two distinct wavelengths are required to establish the bidirectional connection between each pair of VS nodes. The basic operation of the VS-Based Algorithm is then to extend these light-paths to form a single light-tree for each connection, as will be explained below.

The VS-Based Algorithm runs into two stages. The first stage is of long term significance and it concerned about allocating the VS nodes in the network, connecting them, and then *clustering* the network according to VS nodes locality such that each cluster forms a sub-tree with the VS as the root, and the nearest non-VS nodes as its children. The VS nodes are selected such that its nodal degree is high, so that it can be connected to a large number of nodes, and the average distance from the VS node to its cluster-nodes is nearly the same for all the clusters²².

The second stage is the tree generation stage and it exploits the pre-established CS, and the virtual node-clustering organization resulting from the first stage for setting up the multicast tree. This stage is repeated for each new multicast session. For a successful establishment for the

²⁰Thus, the name optical-based is used.

²¹Thus, the name non-source/optical-based originated is used.

²²Although the CS is physically established, the clustering is performed virtually and will be employed when the connection for multicast session is actually taking place.

multicast session, two conditions must be satisfied. The first one is related to the VS-availability while the second one is concerned about the wavelength-availability.

Each multicast session must be established through a *free* VS node, i.e., through a VS node that is not associated with any other multicast session. The VS nodes are checked in sequence according to the source node distance from them and the first free VS, called the Primary VS (PVS), is chosen. Due to the absence of any wavelength conversion away from the CS, the same wavelength should be available over all the links from the source node to the selected VS; otherwise, the connection cannot be established and the next nearest VS is examined. The connection is blocked if no VS node is free and/or no single wavelength is available over the links of the shortest path to a free VS. This node joining mechanism relieves the source node from determining the routes to destinations, therefore makes algorithm scalable.

On the receiver(s) side, each member node establishes a connection to its cluster VS, called the Secondary VS (SVS), through a unicast connection to the SVS using a single wavelength. The destination is blocked only if no single wavelength is available over all the links between the destination node to its SVS, since wavelength conversion is assumed to exist only at VS nodes. This scheme allows the VS-Based Algorithm to support dynamic multicasting since it provides an easy means for the receivers to join/leave the multicast session at anytime during the session lifetime.

The Member-SVS connections are done in parallel which reduces the required setup time for the multicast session. This reduction in the setup time is a direct result of the pre-establishment of the CS. However, reserving two wavelengths per VS-VS connection pair to maintain the CS forms a bottleneck on the system performance in terms of the number of multicast sessions that can be supported by the system, which is equal to the number of the VS nodes in the network. Therefore, the cost of increasing the number of multicast sessions in the system is very high and is determined by the cost of increasing the number of VS nodes in the network. Alternatively, more wavelengths can be reserved for maintaining the CS (which increases the possibility of the inefficient usage of the network resources), or an efficient time division multiplexing scheme over the wavelength channel may be employed. However, it is worth noting that this bottleneck affects only the number of *outgoing* multicast sessions that a single VS can support, i.e., as a PVS, but not the number of *incoming* multicast sessions, i.e., as a SVS.

Moreover, the conditions of choosing the VS nodes do not take the relative locations of the VS nodes among themselves into consideration. Hence the CS may result into long routes between the VS nodes. Also, because of the absence of any splitting node away from the CS, the CS construction is done inefficiently by unicast connections which means that more wavelengths are

wasted to maintain the CS. Also, the construction process of the multicast tree may not be loop-free, e.g. when the multicast source is part of the CS. Although [23] does not specify a mechanism to detect/eliminate such loops, special treatments can resolve them.

Example

Assume that the NSFNET has 3 VS nodes only while the remaining nodes are MI nodes with no wavelength conversion. Four multicast sessions are to be supported and they are denoted as: $m_1 = (12, \{3, 7, 11, 13\})$, $m_2 = (4, \{1, 2, 6, 13, 14\})$, $m_3 = (11, \{2, 5, 9, 13\})$, $m_4 = (2, \{6, 9\})$. Since the number of multicast sessions is greater than the number of available VS nodes, one of the multicast sessions, say m_4 , will be blocked. Assuming no shortage in wavelengths, the other three multicast sessions are successfully established.

In order to generate the multicast trees of these sessions, the VS nodes need to be first identified. The degrees of nodes 6 and 10 is 4 which is the maximum degree among all the nodes in NSFNET. Therefore, those nodes are the best candidates for being VS nodes. Also, node 4 is chosen as the third VS node. Then, the CS is constructed using 6 unicast connections, i.e., 2 connections for each VS-VS nodes pair, and four distinct wavelengths, as shown in 26-a. Notice that the notation λ_i/λ_j in the figure refers to wavelengths λ_i and λ_j being used in opposite directions on the link. Also, the figure shows the virtual cluster for each VS node.

Each multicast session is considered one by one. The source node for m_1 , i.e. node 12, selects VS node 10, which is the nearest *free* VS node, as its PVS and it connects to it. Member nodes 3 and 13 are connected to their cluster's SVS, i.e. node 6. Destination nodes 7 and 11 lie on the CS, and are thus able to acquire the multicast data using their DaC capability²³. The resultant tree is shown in Figure 26-b. Similarly, the multicast trees for m_2 and m_3 are constructed using VS nodes 4 (which is the source node itself) and 6 as their PVS and are shown in Figure 26-c and 26-d, respectively.

Partial Packet Replication All-Optical Multicast Heuristic (PPR-AOMH) [24]

Similar to the VS-Based Algorithm, the Partial Packet Replication All-Optical Multicast Heuristic (PPR-AOMH) proposed in [24] uses the CS concept to construct the multicast distribution trees. However, the PPR-AOMH differs from the VS-Based Algorithm in many ways. First, the CS implementation in the PPR-AOMH takes the form of a tree that connects the MC nodes instead of the interconnected unicast paths. Second, the core nodes are basically MC nodes that are

²³Such a situation where a member node is also a core node was not discussed explicitly in [23]. Therefore, other scenarios are possible too, e.g., a special channel is established from the destination core node to its cluster VS node.

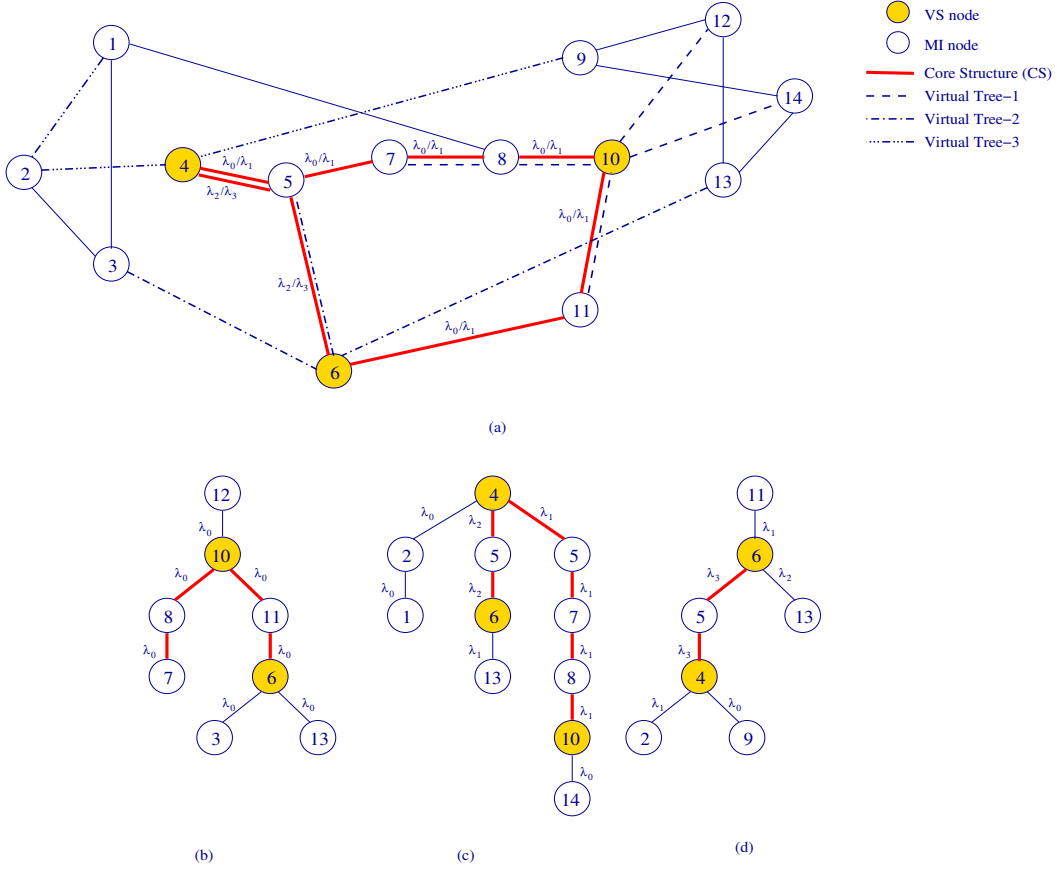


Figure 26: Example of the VS-Based Algorithm in the NSFNET where nodes 4, 6 and 10 are the only VS nodes: (a) The CS and the Virtual Trees organizations, (b) The multicast tree for m_1 , (c) The multicast tree for m_2 , and (d) The multicast tree for m_3

members in the multicast session, but they may also include some *assisting* non-member MC nodes besides some MI nodes. Third, a single CS is constructed per multicast session at the time of connection establishment. In addition, the AOMH takes into consideration the fan-out constraint while constructing the multicast trees as will be shown below.

The PPR-AOMH is designed to support multicast traffic in $S^s F^c R^x - C^f D^c - M^s$, or $S^s F^c R^x - C^s D^c - M^s$, where wavelength conversion can be deployed fully or sparsely, respectively. However, the MC nodes are the only nodes which have the DaC capability while the MI nodes can support the Drop-Only (DO) or Continue-Only (CO) capabilities using Add-Drop Multiplexer (ADM)²⁴. Due to this hardware limitation, each of the destination nodes that are MI must be connected to an MC destination through a dedicated unicast connection. Hence, the MI nodes will be leaf nodes in the constructed tree.

²⁴Although the ADMs are one possible hardware implementation at the MI nodes, other implementations that achieve the DaC capability in an all-optical fashion are possible too. Therefore, the authors of this paper disagree with the argument made in [24] regarding that the algorithms in [56] are not for All-Optical Multicasting.

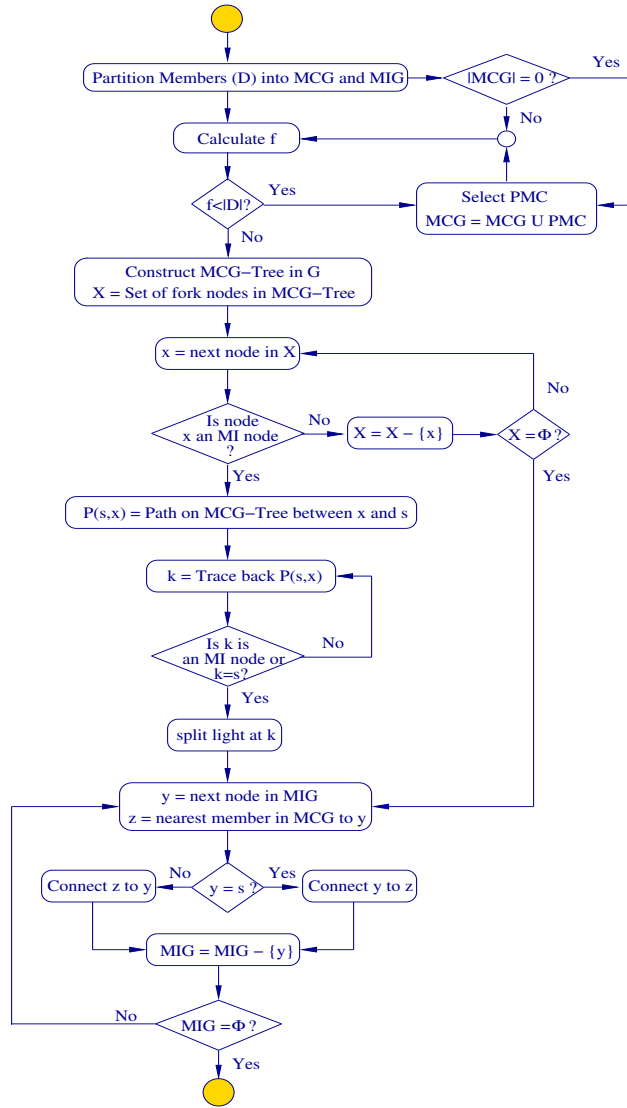


Figure 27: Flow Chart Representation for the PPR-AOMH.

The construction of the multicast delivery tree is triggered by a new connection request, and employs three phases. The first phase is the *Preparation Stage*. In this stage, the algorithm divides the *member nodes*, D , into two disjoint sets, i.e. Multicast Capable Group (MCG) and Multicast Incapable Group (MIG), based on their splitting capability. Since each connection to an MI node is established using a dedicated channel per MI destination, the number of these connections will be limited by: (1) the potential splitting capability over the path to the destination nodes, which is referred to as the fan-out constraint $f = W \sum \delta_i > D$ where δ_i determines the nodal degree of the i 's MC destination and W is the number of channels per link, and (2) the number of available wavelengths on the links, i.e. link capacity constraint. If the MCG is empty or the fan-out constraint is not satisfied, then the MCG is expanded to include some assistant MC nodes, called Proxy MC or PMC nodes, until such constraint is satisfied.

Then, the second stage starts to construct the CS. The CS, or the MCG-tree, is constructed to include all the nodes in the MCG, and assumes that all the nodes in the network have the splitting capability. Any source-based, shared-tree or Steiner-based tree construction algorithm can be used to construct the MCG-Tree. Then, the branching nodes in the MCG-Tree are examined one by one to check if they are MC node, and are hence capable of supporting splitting in the tree. If a fork node happens to be an MI node, the algorithm pushes light splitting to the nearest ancestor MC node along the shortest path from this node to the source node. In the worst case, if no MC nodes are encountered, then data duplication takes place at the source node using multiple wavelengths. Finally, the third stage connects each remaining MI node to the nearest MC node in the CS using a dedicated unicast connection, even though if more than one MI destination lie on the same shortest path to the nearest MC core node because of the limitation in the optical-capabilities explained earlier. This results in different MIG sub-trees than the ones produced by the VS-Based Algorithms. Moreover, the direction of each connection is determined by whether the MI node is the source node or not. If the MI node is the source node itself, then the connection will be directed towards the core MC node; otherwise, the direction is reversed. The resulting sub-trees are called MIG-Trees. The PPR-AOMH operation is shown in Figure 27.

Two deployment strategies for the MC nodes were proposed in [24]. The first one deploys the MC nodes randomly in the network, hence the name RAND. The second one, called PRIOR, makes use of the nodal degree and deploys the nodes semi-randomly such that a node with a higher degree is chosen to be an MC node with higher priority. Simulation results show that the number and the deployment strategy of the MC nodes have direct impact on the system performance. The PPR-AOMH exhibits good performance when 20% of the nodes in the network are MC nodes and they are deployed using the PRIOR scheme. However, the PPR-AOMH may result in high delay, which makes it inappropriate for delay-sensitive applications.

Example

The following example illustrates the operation of the PPR-AOMH for constructing the multicast tree for the multicast session $(2, \{1, 3, 4, 6, 9, 10, 11, 13, 14\})$ in the sample NSFNET, with two wavelengths only. For the sake of simplicity, we assume that the conversion and splitting capabilities are integrated in the same set of MC nodes: 4, 6 and 10. This selection of the VS node set is based on the *PRIOR* scheme since they have the highest nodal degree in the network. When the multicast request arrives, the algorithm starts to build the CS, i.e., the MCG-Tree. The MCG-Tree rooted at node 4 is constructed assuming that all the nodes have the splitting capability. The MIG-trees are also constructed such that each MI node, including the source node itself, is connected to the

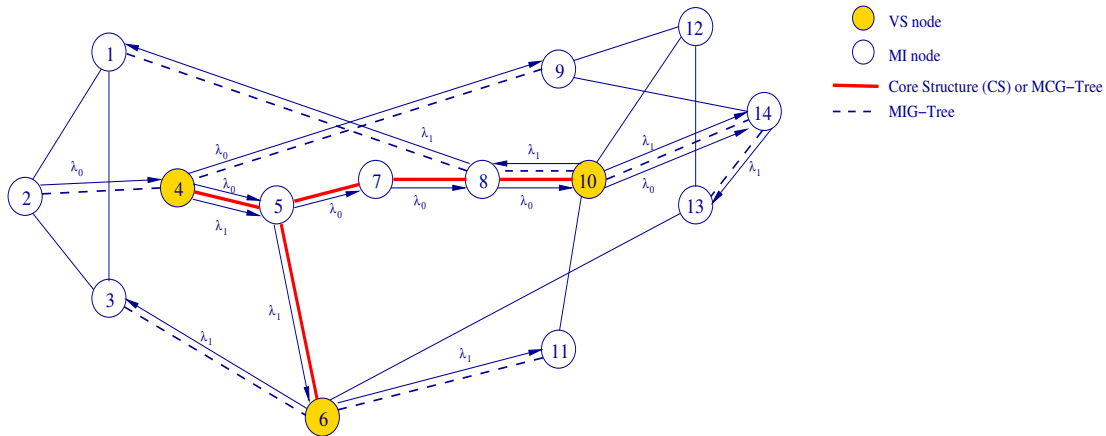


Figure 28: Constructing the light-tree using PPR-AOMH for $(2, \{1, 3, 4, 6, 9, 10, 11, 13, 14\})$ in NSFNET, where nodes 4, 6, 8 and 10 are VS nodes.

nearest MC member node. The resulting MCG-/MIG- Trees are shown in Figure 28. However, the MIG-Tree includes one branching node, i.e. node 5, that is an MI node. Hence, light splitting is performed at node 4, which is the nearest ancestor MC node to node 5.

When the multicast session starts, the source node 2 transmits over λ_0 to node 4 which splits the light into three signals: one copy on link (4,9) over λ_0 , which terminates at node 9, and two copies over link (4,5). The first copy uses λ_0 , and continues over the subsequent links to node 10, while the second copy uses λ_1 , and continues to node 6 where it is split again to feed the destination nodes 3 and 11. The first signal over λ_0 is split at node 10 into three copies to support the three unicast connections to its children, i.e., nodes 1, 13 and 14. Note that unlike the splitting at node 4 which was necessary because the branching node 5 is not an MC node, the splitting at node 10 is performed because node 14 cannot support the drop and continue operations simultaneously; hence, two separate unicast connections are needed. Moreover, although link (8,10) involves two bidirectional connection for the multicast tree, the tree is still loop-free since both connections are carried over different wavelengths.

VI.1.ii All-Optical-Multicasting Wavelength Assignment (AOM-WA) Techniques

The AOM-WA has been investigated in the literature in two contexts. In the first, researchers determine the wavelength requirements for supporting multicast in the network. For example, in [72] some properties from the expander graphs were exploited to derive an asymptotic upper bound on the number of wavelengths needed to support AOM. According to [72], such bound is impractical, however, it still can be used as a bound on the rate of growth of the number of wavelengths.

In [73], the authors derived bounds on the minimum number of required wavelengths in some

regular network topologies, which include linear arrays, rings, hypercubes and meshes. Also, in [74] the author considers power limitation constraint and assumes that a multicast message can be dropped or split only at a limited number of nodes along a lightpath or a light-tree. Based on this, the upper and lower bounds of wavelength requirements for establishing a multicast connection in different topologies, such as mesh torus, mesh, hypercube, and general topologies, are proposed and proved. Similar to [74], the work in [75] computes the wavelength requirements of a multicast communication when a limited number of light-dropping or tapping is allowed in a multi-hop fashion. The author proved that determining the minimum wavelength requirements in an arbitrary network is NP-Complete; however, a solution for such a problem can be found in polynomial time in some regular topologies, such as rings, tori and hypercubes.

In the other context, several techniques were developed to assign wavelengths to multicast sessions such that the number of wavelengths is minimized. These techniques were either formulated as an optimization problem or solved using heuristics.

The wavelength assignment optimization problem in [76] was studied in the context of several Quality of Service (QoS) requests. Initial multicast trees along with their wavelength assignments are computed. Then, minimizing the number of required wavelengths in the network can be solved by rerouting the multicast trees that either contain the links of the maximum load, or are initially assigned the least used wavelengths. Simulation results in [76] revealed that the second approach performs better than the first one.

Also, the AOM-WA was formulated in [40] as an ILP. On the other hand, the AOM-WA is solved in [58] using a simple heuristic in which the light-forest is divided into segments, each of which is a sequence of links such that a wavelength converter resides at one end of each segment, or both. Then all the links that constitute the segment are assigned one wavelength.

Moreover, [34, 48] discussed wavelength assignment with multicast service in WDM networks, and also proposed two different heuristic for wavelength assignment. Both methods separate the wavelength assignment for multicast over WDM into two parts: constructing a routing tree and assigning wavelength. The difference between them is that: [34] constructs multicast routing tree before assigning wavelength, while [48] assigns wavelength for a multicast session before finding a multicast routing tree.

The objective function of wavelength assignment scheme proposed in [34] is to minimize the number of wavelength conversion times in the multicast tree. For each node, if its incoming wavelength is different from its outgoing wavelength, then the wavelength conversion cost is a non-zero value. Before assigning wavelength for this multicast session, the multicast routing tree is built, then the nodes on that tree are processed in the bottom-up fashion, i.e., computing a

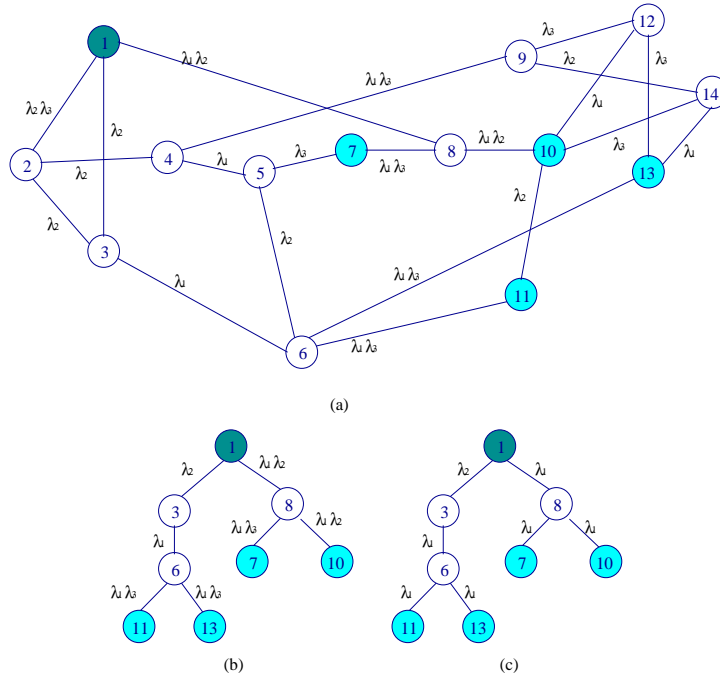


Figure 29: Example for Wavelength Assignment Object to Minimum Wavelength Conversion

specific wavelength conversion cost from leaf nodes first. Each leaf node selects a wavelength with minimum wavelength conversion cost as its incoming wavelength from its parent node. The parent node adds its child nodes' conversion cost, and includes its own wavelength conversion cost into that sum if wavelength conversion is necessary between an incoming wavelength and an outgoing wavelength. Thus, when a node is processed, all the required values in its sub-tree are already available. This paper proved that such selected wavelength set is optimal. An example using NSFNET is given in Figure 29. Node 1 is the source node, while nodes 7, 10, 11, and 13 are multicast group members. Figure 29-(b) shows a multicast tree for that session. The final result is given in Figure 29-(c).

The authors in [83] extended the work in [34] by allowing multiple available wavelengths in a link to carry the multicast signal from the source node instead of one wavelength. In this scenario, each wavelength is launched from the source node as a separate light-tree and it can share some links with other light-trees. This scheme proves to significantly reduce the wavelength conversion cost since instead of converting the signal at a node, a separate signal from the source is used. This approach is employed with the assumption that the wavelength conversion cost is much higher than the wavelength and splitting costs.

The objective function in [48] is to minimize the number of wavelengths used. The scheme consists of three steps:

1. Generate an auxiliary bipartite graph from the given network, whose vertex-sets correspond

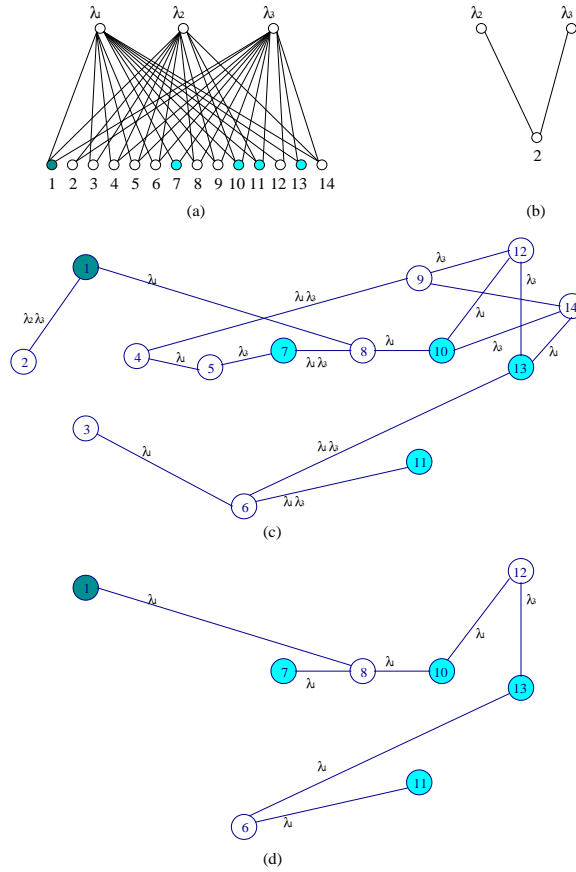


Figure 30: Example for Wavelength Assignment Object to Minimum Wavelength Number

to the set of wavelengths, and the set of nodes in the original network. There is an edge between a wavelength and a node if and only if this wavelength can cover that node.

2. Choose one wavelength that can cover the maximum number of nodes in the auxiliary bipartite graph, and then remove those covered nodes as well as this wavelength from the auxiliary graph. This step is repeated until all nodes are removed. Thus, a wavelength set that can cover all nodes is obtained, and that paper prove this set is in minimum number of wavelength.
3. Connect all nodes in the given network with those edges with chosen wavelength set, and based on this new graph, produce a shortest path tree or a Steiner minimum tree.

Using the above NSFNET example, Figure 30 shows the procedure of this algorithm. The auxiliary bipartite graph is given by Figure 30-(a). Figure 30-(b) shows the first iteration of step 2: wavelength λ_1 is removed because it covers the largest number of nodes, and all those nodes covered by λ_1 are also removed. Figure 30-(c) gives the subgraph with selected wavelength set as $\{\lambda_1, \lambda_3\}$. The final multicast tree is given in Figure 30-(d).

VI.2 All-Optical IP (AO-IP) Multicasting Algorithms

The various schemes discussed so far support multicast communication in WDM networks all-optically and at the optical layer. These techniques are of circuit-switched nature, and they are best suited for applications with huge bandwidth, and long duration requirements. However, integrating IP operation with WDM networks, in what is called the Optical Internet (OI) [26], has received an increasing interest recently. Surely, AOM is one of the main services to be supported in the OI. The efforts in this field are strongly motivated by the advantages of AOM that were discussed previously and by the need to inter-work AOM with IP-Multicasting which is already deployed in the Internet. These efforts, however, are targeting another class of applications that require significantly reduced amount of bandwidth, have relatively shorter lifetime, and their traffic is bursty.

In this section, we provide an overview of the schemes that have been proposed in the literature for supporting All-Optical IP (AO-IP) Multicasting. These schemes make use of the newly proposed Optical Burst/Label Switching (OBS/OLS) mechanism [60] as a means for transporting the data using virtual circuits under the framework of the Multiprotocol Label Switching (MPLS) [61] in the OI. In this context, no reservation for wavelength takes place in the network, i.e., it is done on-demand, and the transportation of the multicast traffic over the delivery tree is performed in the form of bursts and using Label Switched Paths/Trees²⁵ (LSP/LST) instead of Light-paths/Light-trees. Moreover, like the wavelength-routing networks, some of the nodes in the OI are equipped with the optical splitting capability and they are controlled using a special controller that is separate from the upper IP router.

IP multicasting cannot be employed directly in the OI networks to achieve the AOM because of the operational incompatibilities between the two schemes. For example, the concept of the Multicast capable/incapable (MC/MI) nodes in IP multicasting does not include the splitting capability of the nodes. As such the MC (MI) node in the IP multicasting is determined as a node that runs (does not run) the appropriate multicasting protocol at the IP layer only²⁶. Also, because the operations of the currently deployed schemes do not take the splitting capability of the nodes into consideration, the multicast trees computed by any IP multicasting scheme may not achieve good utilization for the network resources since it may include (ignore) electronic MC (MI) *nodes* that are equipped with optical MI (MC) *switches*. Additionally, these MI nodes are either ignored (as in the MOSPF) or bypassed (as in the DVMRP by using the concept of IP

²⁵i.e., one-to-many LSP.

²⁶To distinguish between the two cases, we refer to the Capable/Incapable IP Multicast node as an MC/MI *node* (*router*) while these capabilities are referred to as MC/MI *switches* when optical splitting capabilities are determined.

encapsulation/tunneling) [62] while computing the multicast tree. The strategy of ignoring or bypassing the MI nodes cannot be extended to work with the MI switches in order to achieve AOM in OI. Moreover, supporting AOM necessitates the global deployment of the same multicast protocol at the optical layer of each switch regardless of its optical splitting capabilities, while this condition is not necessary for the operation of IP multicasting. Last but not least, the WDM switch and the IP layer at a node suffer from inconsistency since they may have different views of the network topology or it may happen that more than one IP router is running over WDM layer which results in no one-to-one mapping between the two layers.

Two strategies are followed in order to overcome these incompatibilities and to achieve AOM in the OI [25]. In the first approach, the option of not modifying the existing IP multicasting protocols is adopted. Each node in the OI has two independent entities: the IP controller that runs the IP routing protocol, and the WDM multicast controller which runs at, and controls each optical switch. In this configuration, the IP controller computes the multicast tree without considering the splitting capability of the underlying switches. In order to enable the IP controllers to maintain this *conventional* view of the network, the WDM multicast controller is responsible for: (1) handling any modification in the IP-based tree that is necessitated by the existence of branching nodes at MI switches²⁷, and (2) hiding these modifications from the IP router. In achieving this aim, the operation of the WDM multicast controller does not interfere with that of the IP, however, the resulting routing tables computed by the IP layer, along with the membership information in the network and the splitting capabilities of the nodes are available to the WDM multicast controller.

Reference [25] investigated the applicability of this approach and the strategies to reconnect the affected nodes when two IP multicast routing protocols, i.e. DVMRP and MOSPF, are used. When the IP router employs DVMRP, three scenarios were proposed in order to initiate the reconnection process of the cut node(s). In these scenarios, the reconnection process is initiated by the parent node of the cut node, the affected cut child itself or any of its neighbor nodes that is not the source node (called a relative node), respectively. However, in the case of MOSPF, the LSP is established from an MC node in the shortest path to the source node.

The work in [26] proposed a practical AOM protocol that keeps the employed IP multicasting protocol (DVMRP in this case) intact and modifies its constructed tree into a forest. The protocol works as follows. The initial DVMRP-based tree is constructed first. The modification stage is then initiated by the source node which sends a *repair* message along the initial tree to all the downstream nodes. At a branching node, the *repair* message is forwarded to all its children if and

²⁷Such modifications are performed by reconnecting each disconnected children using a dedicated LSP.

only if this branching node is equipped with MC switch, otherwise, the *repair* message is relayed to one child only while a *purge* message is sent to the remaining children.

All the nodes that receive the *repair* message are considered at the light-forest, while each purged node has to reconnect itself to this forest through an LSP and using the *grow request* message which is sent to all its neighbors. Two modes of operation were proposed in [26] for growing back the cut children. In the first technique, which is called the direct grow, the *grow request* message is processed and relayed at a node only if it is part of the original multicast tree while in the other scheme, i.e., the indirect grow, the *grow request* message is always processed.

Like [26], reference [63] proposed another source-rooted scheme that relies on the modification of an initial IP-based tree. However, unlike [26], the modification process is receiver-initiated, instead of being source-initiated, which suits heterogeneous networks with insufficient information at the WDM layer. The core operation of the proposed protocol depends on providing the receivers with sufficient information about the various capabilities of the nodes along a path using a certain mechanism. The collected information include not only the splitting and conversion capability, but also the tapping and add/drop capabilities. However, such information does not include the network topology, global capabilities and/or membership information.

On the other hand, in spite of the advantages achieved by the OBS, its operation encounters control overhead because a control packet has to be sent for each burst in order to setup the switches along the LSP. In addition, Guard Bands (GBs) of bursts on different channels are needed in order to accommodate timing jitter in the burst in the intermediate nodes. This overhead was considered in designing the three multicasting schemes proposed in [64].

The first algorithm, namely Separate Multicast (S-MCAST), delivers the assembled multicast bursts separately from the unicast ones using a dedicated source-rooted tree for each multicast session. The second technique is called Multiple Unicasting (M-UCAST), and it basically assembles the multicast traffic and the unicast counterpart into the same bursts that are destined to the same destination. Although this scheme results in an inefficient bandwidth usage since it treats the multicast as unicast, it results in reducing both the control packets and the GBs overheads and it is a good solution for certain conditions.

The third scheme is the Tree-Shared Multicasting (TS-MCAST). Its operation benefits from the overlapping degree in the membership of the nodes. Basically, the multicast session originating from an edge router are divided into classes, called Multicast Sharing Classes (MSCs), according to certain criterion²⁸. Each MSC constructs a single tree to deliver the multicast burst. This technique

²⁸Three options were proposed in [64]: (1) *Equal Coverage* scheme groups the sessions with the same members and constructs one tree for all, (2) *Super Coverage* scheme constructs a single tree for a superset session with all the

results in longer bursts, thus the impact of the control and GBs overheads is less compared to their effects in the other algorithms.

Moreover, in the OBS-Based OI, a burst can be blocked and lost, at the intermediate nodes if no wavelength is available at the outgoing link and the nodes are not equipped with any buffering mechanisms (i.e., using the Fiber-Delay Line, FDL). The authors in [65] proposed a reliable multicast scheme to recover from any burst lost.

The other approach that can be adopted is to modify the existing IP multicasting protocols such that they take care of the optical capabilities of the nodes (i.e., splitting, conversion, DaC, etc) when constructing the multicast tree. This option eliminates the need for the WDM multicast controller and, hence, reduces the complexity of the control overhead and can produce good tree structure since the multicast tree computed by the IP layer will be able to avoid branching at MI switches. However, the practical deployment of such modifications in the network is questionable.

VI.3 Hybrid Multicasting (HM)

The Hybrid Multicasting (HM) scheme arises as an intermediate solution that attempts to benefit from the advantages of All-Optical Multicasting (AOM) and the flexibility provided by the matured technology of All-Electronic Multicasting (AEM). Basically, this strategy attempts to maintain the multicast signal in the optical domain as long as possible. Conversion to the electronic domain is performed in extreme cases, e.g., when wavelength conversion is done using electronic-based converters, or when AOM-R protocol results in constructing a multicast delivery structure with an unacceptable delay or bandwidth wastage, etc.

With the $S^f F^c R^x - C^n - M^s$ system model, the MSCH2 Algorithm [66] initially computes the multicast light-tree for the multicast session using any conventional multicast algorithm, and then it tries to assign a single wavelength to it. If no single wavelength is available along the computed tree, the *tunnels*, which are the branching points where the wavelength continuity constraint is violated, are found [66]. At these nodes, the optical connection is terminated and a new optical connection is established over another wavelength.

The scheme proposed in [77] is designed for the $S^f F^c R^x - C^n - M^s$ system model. Each routing node is equipped with an ATM switch/SONET framer that operates on the top of an MC optical switch. The MC optical switch is used to switch the multicast traffic all-optically. However, electronic switching is used for the following cases:

other sessions are subset of it, and (3) *Overlapping Coverage* scheme which repeatedly combines the sessions into the MCS in order to increase value of a certain overlapping degree criterion.

1. *Wavelength Translation*: Whenever data must be forwarded from a wavelength to another, and this is performed by retransmitting the signal,
2. *Signals Merging Operation*: When two or more distinct signals of the same wavelength and are input from different ports (links) are directed to the same output port (link),
3. *Signals Splitting Operation*: When two or more combined signals of the same wavelength that are input from the same port (link) are directed to the different output ports (links).

The structure of the routing node in [77] is capable of resolving any contention in the optical and electronic switches by the use of Fiber Delay Line (FDL) and output queues, respectively. Also, a dynamic multicast routing algorithm was proposed in [77] that allows nodes to be added to the tree in such a way that the number of required electronic copies is minimized.

The system model in [78] is similar to that employed in [40], i.e., $[N, W, B] - S^n - C^n - M^n$, where all the nodes have the TaC capability. The authors in [78] investigate the problem of constructing optimal multicast virtual topology with minimum average *optical* hop distance, in the presence of splitting loss constraint. The problem is proven to be NP-Complete in arbitrary topologies. However, the work in [78] proves that polynomial time solutions of this problem exist in two particular topologies, i.e., linear and ring topologies. All nodes in these topologies are assumed to be part of the multicast communication, and the number of light taps is bounded by a certain constant, that is technology-dependent. When the limit on the number of light-taps is reached, the lightpath is terminated and a new lightpath is established if more destinations need to be included. The optimal solution for the virtual topology is carried out using dynamic programming.

Finally, the work in [36] considers a special variant of the $S^f F^c R^x - C^n - M^s$ system model. This study is mainly concerned about the difficulty that stems from the limitation in the number of available transmitters and receivers at the intermediate nodes, especially when the availability of wavelengths on the links is limited. Therefore, wavelength translation is required by terminating the connection at a receiver and re-establishing the connection by transmission over another wavelength. To overcome this transceiver constraint, such wavelength translation is performed at an ancestor node (or at worst, at the source node) and multiple copies are then transmitted on the same link(s) to the bottleneck node. Nevertheless, each multicast session is allowed to use a maximum number of wavelengths per single link. With this configuration, the combined RWA problem was proven to be NP-Complete, while a solution for the WA problem can be found in linear time in the number of nodes if the number of wavelengths per link, number of transceivers per node and the switch degree are constant.

Tech.	Ref.	System Model	Structure	Membership	Policy	Pow.
MDMCT	[40]	$S^n-C^n-M^n$	Trail	Static	FDBP	No
RWA/WANs	[21]	$S^f F^c R^x-C^f/nD^c/l-M^f$	Tree	Static/ Dynamic	FDBP/ PDBP	No
Light-Tree	[53]	$S^f F^c R^x-C^f D^c-M^f$	Tree	Static	FDBP	No
RWA/SS	[22]	$S^f F^c R^x-C^n-M^f$	Tree/ Forest	Static/ Dynamic	PDBP	No
RR2S	[56]	$S^s F^c R^x-C^s-D^c-M^s$	Forest	Static	FSBP	No
RR2A	[56]	$S^s F^c R^x-C^s-D^c-M^s$	Forest	Static	FSBP	No
M1 st	[56]	$S^s F^c R^x-C^n-M^s$	Forest	Static	FDBP	No
M-Only	[56]	$S^s F^c R^x-C^n-M^s$	Forest	Static	FDBP	No
VS-CP	[58]	$S^s F^c R^x-C^s D^c-M^s$	Forest	Static	FDBP	No
CSA	[59]	$S^s F^c R^x-C^s D^c-M^s$	Forest	Static	FDBP	Yes
TS-AOM	[71]	$S^s F^c R^x-C^n-M^s$	Tree	Static	FDBP	No
VS-Based	[23]	$S^s F^c R^x-C^s D^c-M^s$	Tree	Dynamic	FDBP	No
PPR-AOMH	[24]	$S^s F^c R^x-C^f/sD^c-M^s$	Tree	Static	FDBP	No
MSCH2	[66]	$S^f F^c R^x-C^n-M^s$	Tree	Static	FDBP	No
ATM-WDM	[77]	$S^f F^c R^x-C^n-M^s$	Tree	Dynamic	PDBP	No
Vir. Top.	[78]	$S^n-C^n-M^n$	Path	Static	FDBP	Yes
Multi-Hop	[36]	$S^f F^c R^x-C^n-M^s$	Tree	Static	FDBP	No
logical-Network	[80]	$S^s F^c R^x-C^s D^c-M^s$	Forest	Static	FDBP	No
Amplifier Placement	[81]	$S^s F^c R^x-C^n-M^s$	Forest	Static	FDBP	Yes
PAM	[82]	$S^s F^c R^x-C^n-M^s$	Forest	Static	FDBP	Yes

Table 2: Comparison between multicasting techniques in wavelength-routing networks in terms of: system model, multicast delivery structure (Structure), membership policy (Membership), destinations blocking policy (Policy) and power-budget awareness (pow.) .

VII Conclusions and Open Research Problems

This paper has considered multicasting in wavelength-routing optical networks. The paper has discussed the design and operation of wavelength-routing networks in general, and the special provisions required for multicast support in such networks. In this context, the challenges were outlined, and some of the proposed solutions were introduced. The paper also introduced a classification

and a comparison of multicasting techniques, and outlined their advantages and disadvantages. Table 2 provides a summary and a comparison of those schemes.

The area of multicasting in wavelength routing networks is a relatively new area of research, and there are still several open research issues. For example, the dimensioning of wavelength-routing networks for multicasting still requires more study. The problem is in general NP-Hard, and efficient algorithms are required to determine the number of fibers, the number of wavelengths, and the numbers and locations of splitters, amplifiers and wavelength converters. Although network dimensioning is an off-line problem, connection provisioning may have to be done on-line, especially in highly dynamic environments. This puts more emphasis on the need for time, and storage efficient algorithms to determine how to route multicast calls, and how to assign wavelengths to such calls. The design of optimal, or near optimal multicast trees which takes into account device limitations, as well as the power budget constraints still needs more study. This may necessitate that use of, not just a single multicast tree, but rather a forest of trees. Optimal algorithms for the determination of each of those trees, and the partitioning of the multicast group are still needed. In addition, non-traditional multicast tree construction methods may be needed. The evaluation of blocking probabilities under multicasting is not an easy task, and very little work has been done in this area. Accurate, and efficient models need to be developed. Although most of the models which have been developed assume full call acceptance, partial call acceptance must also be modeled, since network operators will be offering this service in order to maximize their revenues.

References

- [1] C.A. Brackett, “*Dense Wavelength Division Multiplexing Networks: Principles and Applications*”, IEEE JSAC, Volume 8, Issue 6, pp. 948-964, Aug. 1990.
- [2] R. Ramaswami, “*Multiwavelength Lightwave Networks for Computer Communication*”, IEEE Communications Magazine, Volume 31, Issue 2, pp. 78–88, Feb. 1993.
- [3] “*Special issue on lightwave networks*”, IEEE Journal of Lightwave Technology, Volume 14, Issue 6, Jun. 1996.
- [4] R. Ramaswami and K N. Sivarajan, “*Optical Networks: A Practical Perspective*”, 2nd Edition, Morgan Kaufmann Publishers, 2002.
- [5] T.E. Stern and K.Bala, “*Multiwavelength Optical Networks: A Layered Approach*”, Addison-Wesley, 1999.

- [6] Lucent Technologies, “*Bell labs uses ultra-dense WDM to transmit 1,022 channels over fiber*”, www.lucent.com/news/1999/november/10/2.html.
- [7] B. Mukherjee, “*WDM-Based Local Lightwave Networks-Part I: Single-hop systems*”, IEEE Network, Volume 6, Issue 3, pp. 12-27, May 1992.
- [8] B. Mukherjee, “*WDM-Based Local Lightwave Networks-Part II: Multihop systems*”, IEEE Network Magazine, Volume 6, Issue 4, pp. 20-32, Jul. 1992.
- [9] J. Turner, “*The Challenge Of Multipoint Communication*”, 5th ITC seminar on ISDN Design and Planning, pp. 263-279, May 1987.
- [10] A. Hamad and A. Kamal, “*A Survey of Multicasting Protocols For Broadcast-and-Select Single-Hop Networks*”, IEEE Network Magazine, Vol. 16, Issue 4, pp. 36-48, Jul/Aug. 2002.
- [11] J. Leuthold and C.W. Joyner, “*Multimode Interference Couplers With Tunable Power Splitting Ratios*”, Journal of Lightwave Technology, Volume 19, Issue 5, pp. 700-707, May 2001.
- [12] A Girard, “*Routing and Dimensioning in Circuit-Switched Networks*”, Addison-Wesley, 1990.
- [13] D. R. Cheriton and S. Deering, “*Host Groups: A multicast extension for datagram internet-works*”, in Proc. Data Communications Symposium, 1985, pp. 172-179.
- [14] S. Deering, “*Host Extensions for IP Multicasting*”, RFC 1112, Aug. 1989.
- [15] Y. K. Dalal and R. M. Metcalfe, “*Reverse Path Forwarding of Broadcast Packets*”, Communications of the ACM, Volume 21, pp. 1040-1048, Dec. 1978.
- [16] D. Waitzman and C. Partridge, “*Distance Vector Multicast Routing Protocol*”, RFC 1075, Nov. 1988.
- [17] H. Eriksson, “*MBone: The Multicast Backbone*”, Communications of the ACM, Volume 37, pp. 54-60, Aug. 1994.
- [18] B. Fenner, M. Handley, H. Holbrook and I. Kouvelas, “*Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)*”, IETF Internet Draft, draft-ietf-pim-sm-v2-new-06.txt, December 2002.
- [19] A. Adams, W. Siadak, J. Nicholas, “*Protocol Independent Multicast-Dense Mode (PIM-DM) Protocol Specification (Revised)*”, IETF Internet Draft, draft-ietf-pim-dm-new-v2-02.txt, Oct. 2002.

- [20] E. W. Dijkstra, “*Note on two problem in connection with graphs*”, Numerische Mathematik, Volume 1, pp. 269-271, 1959.
- [21] G. Sahin and M Azizoglu, “*Routing and Wavelength Assignment in All-Optical Networks with Multicast Traffic*”, European Transactions on Telecommunications, Volume 11, Issue 1, pp. 55-62, Jan/Feb. 2000.
- [22] J. He, S. -H. G. Chan and D.H.K. Tsang, “*Routing and Wavelength Assignment For WDM Multicast Networks*”, IEEE GLOBECOM’01, Volume 3, pp. 1536-1540, 2001.
- [23] N. Sreenath, N. K. M. Reddy, G. Mohan and C. S. R. Murthy, “*Virtual Source Based Multicast Routing in WDM Networks With Sparse Light Splitting*”, IEEE Workshop on High Performance Switching and Routing, pp. 141-145, 2001.
- [24] W. Tseng and S Kuo, “*All-Optical Multicasting on Wavelength-Routed WDM Networks with Partial Replication*”, 15th International Conference on Information Networking, pp. 813-818, 2001.
- [25] X. Zhang J. Wei and C. Qiao, “*On Fundamental Issues in IP over WDM Multicast*”, IEEE IC3N’99, Oct. 1999.
- [26] C. Qiao, M. Jeong, A. Guha, X. Zhang and J., “*WDM Multicasting in IP Over WDM Networks*”, IEEE ICNP’99, pp. 89-96, Nov. 1999.
- [27] R. Malli, X. Zhang and C. Qiao, “*Benefit of Multicasting in All-Optical Networks*”, SPIE proceedings of Conf. All-optical Networking, Volume 3531, pp. 209-220, Nov. 1998.
- [28] J.M.H. Elmirghani and H.T. Mouftah, “*All-Optical Wavelength Conversion: Technologies and Applications in DWDM Networks*”, IEEE Communications Magazine , Volume 38, Issue3, pp. 86-92, Mar. 2000.
- [29] D. Wolfson, T. Fjelde and A. Kloch, “*Technologies for All-Optical Wavelength Conversion in DWDM Networks*”, 4th Pacific Rim Conference on Lasers and Electro-Optics, pp. 574-575, 2001.
- [30] D. Wolfson, B. Mikkelsen, S. L. Danielsen, H. N. Poulsen, P. B. Hansen, K. E. Stubkjaer, “*Experimental and Theoretical Investigation of Electro-Optic and All-Optical Implementations of Wavelength Converting 2R-Regenerators*”, OFC’98 Technical Digest, pp. 103-105, 1998.

- [31] B. Ramamurthy and B. Mukherjee, “*Optimizing Amplifier Placements in a Multiwavelength Optical Lan/Man: The Unequally Powered Wavelengths case*”, IEEE/ACM transactions on networking, Volume 6, Issue 6, pp. 755-767, Dec. 1998.
- [32] L. Zhong and B. Ramamurthy, “*Optimization of Amplifier Placements in Switch-Based Optical Networks*”, IEEE ICC’01, Volume 1, pp. 224-228, 2001.
- [33] J. H. Franz and V. K. Jain, “*Optical Communications: Components and Systems*”, CRC press, 2000.
- [34] B. Chen and J. Wang, “*Efficient Routing and Wavelength Assignment for Multicast in WDM Networks*”, IEEE JSAC, Volume 20, Issue 1, pp. 97-109, Jan. 2002.
- [35] F. K. Hwang and D. S. Richards, “*Steiner Tree Problems*”, Networks, Volume 22, Issue 1, pp. 55-89, 1992.
- [36] R. Libeskind-Hadas and R. Melhem, “*Multicast Routing and Wavelength Assignment in Multihop Optical Networks*”, IEEE/ACM Transactions on Networking, Volume 10, Issue 5, pp. 621-629, Oct. 2002.
- [37] M. Ali, “*Optimization of Splitting Node Placement in Wavelength-Routed Optical Networks*”, IEEE JSAC, Volume 20, Issue 8, pp. 1571-1579, Oct. 2002.
- [38] W. S. Hu and Q. J. Zeng, “*Multicasting Optical Cross Connects Employing Splitter-and-delivery Switch*”, IEEE Photon. Technol. Lett., Volume 10, pp. 970-972, Jul. 1998.
- [39] M. Ali and J. Deogun, “*Power-Effective Design of Multicast Wavelength-Routed Networks*”, IEEE JSAC, Volume 18, Issue 10, pp. 1852-1862, Oct. 2000.
- [40] M. Ali and J. Deogun, “*Cost-Effective Implementation of Multicasting in Wavelength-Routed Networks*”, Journal of Lightwave Technology, Volume 18, Issue 12, pp. 1628-1638, Dec. 2000.
- [41] M. Ali and J. Deogun, “*Allocation of Multicast Nodes in Wavelength-Routed Networks*”, ICC’01, Volume 2, Jun. 2001.
- [42] Bob Chomycz, “*Fiber Optic Installer’s Field Manual*”, McGraw-Hill, 2000.
- [43] L. Li and A. K. Somani, “*Fiber Requirement in Multifiber WDM Networks with Alternate-Path Routing*”, IEEE IC3N’99, pp. 338-343, Oct. 1999.
- [44] A. K. Somani and M. Azizoglu, “*All-Optical LAN Interconnection with a Wavelength Selective Router*”, IEEE INFOCOM’97, Volume 3, pp. 1278-1285.

- [45] S. Thiagarajan and A. K. Somani, “*An Efficient Algorithm for Optimal Wavelength Converter Placement on Wavelength-Routed Networks with Arbitrary Topologies*”, Proceedings of IEEE INFOCOM’99, Volume 2, pp. 916-923, 1999.
- [46] M. Garnot, F. Masetti and L. Nederlof, G. Eilenberger, S. Bunse and A. Aguilar, “*Dimensioning and Optimization of the Wavelength-Division-Multiplexed Optical Layer of Future Transport Networks*”, IEEE ICC’98, Volume 1, pp. 202-206, 1998.
- [47] A. E. Kamal and A. K. Al-Yatama, “*Blocking Probabilities in Circuit-Switched Wavelength Division Multiplexing Networks Under Multicast Service*”, Performance Evaluation, Volume 47, Issue 1, pp. 43-71, Jan. 2002.
- [48] D. Li, X. Du, X. Hu, L. Ruan and X. Jia, “*Minimizing Number of Wavelengths in Multicast Routing Trees in WDM Networks*”, Networks, Volume 35, No. 4, pp. 260-265, Jul. 2000.
- [49] C. Lee, C. R. Hui, F. F.-K. Tong and P. T.-S. Yun, “*Network Dimensioning in WDM-Based All-Optical Networks*”, IEEE GLOBECOM’98, Volume 1, pp. 328-333, 1998.
- [50] C. Li, F. F. Tong, C. J. Georgiou and M. Chen, “*Gain Equalization in Metropolitan and Wide Area Optical Networks Using Optical Amplifiers*”, IEEE INFOCOM’94, Volume 1, pp. 130-137, 1994.
- [51] T. K. Nayak and K. N. Sivaraman, “*A New Approach to Dimensioning Optical Networks*”, IEEE JSAC, Volume 20, Issue 1, pp. 134-148, Jan. 2002.
- [52] X. Qin and Y. Yang, “*A Cost-Effective Construction for WDM Multicast Switching Networks*”, IEEE ICC’02, Volume 5, pp. 2902-2906, 2002.
- [53] L. H. Sahasrabudde and B. Mukherjee, “*Light-Trees: Optical Multicasting for Improved Performance in Wavelength-Routed Networks*”, IEEE Communications Magazine, Volume 37, Issue 2, pp. 67-73, Feb. 1999.
- [54] S. Ramesh, G. N. Rouskas and H. G. Perros, “*Computing Blocking Probabilities in Multiclass Wavelength-Routing Networks With Multicast Calls*”, IEEE JSAC, Volume 20, Issue 1, pp. 89-96, Jan. 2002.
- [55] R. Srinivasan and A. K. Somani, “*On the Blocking Performance of Tree Establishment in Time-Space Switched Optical Networks*”, OPTICOMM’01, 2001.

- [56] X. Zhang, J. Y. Wei and C. Qiao, “*Constrained Multicast Routing in WDM Networks with Sparse Light Splitting*”, Journal of lightwave technology, Volume 18, Issue. 12, pp. 1917-1927, Dec. 2000.
- [57] H. Takahashi and A. Matsuyama, “*An Approximate Solution For the Steiner problem in Graphs*”, Math. Japonica, Volume 24, Issue 6, pp. 573-577, 1980.
- [58] N. Sreenath, C. Siva Ram Murthy and G. Mohan, “*Virtual Source Based Multicast Routing in WDM Optical Networks*”, Photonic Network Communications, Volume 3, Issue 3, pp. 213-226, Jul. 2001.
- [59] K. Wu, J. Wu and C. Yang, “*Multicast Routing with Power Consideration in Sparse Splitting WDM Networks*”, IEEE ICC’01, Volume 2, pp. 513-517, 2001.
- [60] C. Qiao and M. Yoo, “*Optical Burst Switching (OBS)- A New Paradigm for an Optical Internet*”, Journal of High-Speed Networks, Volume 8, Issue 1, pp. 69-84, 1999.
- [61] R. Callon, N. Feldman, A. Fredette, G. Swallow and A. Viswanathan, “*MPLS Multiprotocol Label Switching Architecture*”, RFC 3031, Jan. 2001.
- [62] L. H. Sahasrabudhe and B. Mukherjee, “*Multicast routing algorithms and protocols: a tutorial*”, IEEE Network Magazine, Volume 14, Issue 1, pp. 90-102, Jan/Feb. 2000.
- [63] M. R. Salvador, S. H. de Groot and D. Dey, “*Supporting IP Dense Mode Multicast Routing Protocols in All-Optical Lambda-Switched Networks*”, SPIE Optical Networks Magazine, Routing Architectures and Technologies for Next-Generation WDM-Based Internet Networks issue, Nov/Dec. 2001.
- [64] M. Jeong, Y. Xiong, H. Cankaya, M. Vandenhoute and C. Qiao, “*Efficient Multicast Schemes for Optical Burst-Switched WDM Networks*”, IEEE ICC’00, pp. 1289-1294, 2000.
- [65] M. Jeong, C. Qiao and Y. Xiong, “*Reliable WDM Multicast in Optical Burst-Switched Networks*”, Opticomm’00, pp. 153-166, 2000.
- [66] Y. Sun, J. Gu and D. H. K. Tsang, “*Multicast Routing in All-Optical Wavelength-Routed Networks*”, Optical Networks Magazines, Volume 2, Issue 4, pp. 101-109, Jul/Aug. 2001.
- [67] B. Mukherjee, S. Ramamurthy, D. Banerjee and A. Mukherjee, “*Some Principles for Designing a Wide-Area Optical Network*”, IEEE/ACM Transactions on Networking, Volume 4, Issue 5, pp. 684-696, Oct. 1996.

- [68] Y. Zhang and D. Sidhu, “*An Analysis Comparing Light-Tree and Lightpath in wavelength Routed Optical Networks*”, IEEE LCN’01, pp. 486-487, 2001.
- [69] C. Siva Ram Murthy and G. Manimaran, “*Resource Management in Real-Time Systems and Networks*”, MIT Press, Apr. 2001.
- [70] S. Paul, “*Multicasting on the Internet and Its Applications*”, Kluwer Academic Publishers, 1998.
- [71] S. Yan, M. Ali and J. Deogun, “*Route Optimization of Multicast Sessions in Sparse Light-Splitting Optical Networks*”, IEEE GLOBECOM ’01, Volume 4, pp. 2134-2138, 2001.
- [72] R. K. Pankaj; “*Wavelength Requirements for Multicasting in All-Optical Networks*”, IEEE/ACM Transactions on Networking, Volume 7, Issue 3, pp. 414-424, Jun. 1999.
- [73] Y. Wang and Y. Yang, “*Multicasting in a Class of Multicast-Capable WDM Networks*”, Journal of Lightwave Technology, Volume 20, Issue 3, pp. 350-359, Mar. 2002.
- [74] S. Gao, X. Jia, X. Hu and D. Li, “*Wavelength Requirements and Routing for Multicast Connections in Lightpath and Light-Tree Models of WDM networks With Limited Drops*”, IEE Communications Proceeding, Volume 148, Issue 6, pp. 363-367, Dec. 2001.
- [75] R. Libeskind-Hadas, “*Efficient Collective Communication in WDM Networks with a Power Budget*”, IEEE IC3N’02, pp. 612-616, 2000.
- [76] X. Jia, D. Du, X. Hu, M. Lee and J. Gu, “*Optimization of Wavelength Assignment for QoS Multicast in WDM Networks*”, IEEE Transactions on Communications, Volume 49, Issue 2, pp. 341-350, Feb. 2001.
- [77] N. Huang, T. Liu, Y. Wang and B. Li, “*On Multicast Routing for Wavelength-Routed WDM Networks with Dynamic Membership*”, Photonic Network Communications, Volume 4, Issue 2, pp. 179-190, 2002.
- [78] R. Libeskind-Hadas, J. R. Hartline, K. Dresner, E. Drucker and K. Ray, “*Multicast Virtual Topologies in WDM Paths and Rings with Splitting Loss*”, IEEE IC3N’02, pp. 318-321, 2002.
- [79] G. Rouskas, “*Optical Layer Multicast: Rationale, Building Blocks, and Challenges*”, IEEE Network, Volume 17, Issue 1, pp. 60-65, Jan/Feb., 2003.
- [80] D. Yang and W. Liao, “*Design of Light-Tree Based Logical Topologies for Multicast Streams in Wavelength Routed Optical Networks*”, IEEE INFOCOM’03, Volume 1, pp. 32-41, 2003.

- [81] A. Hamad and A. Kamal, “*Optimal Power-Aware Design of All-Optical Multicasting in Wavelength Routed Networks*”, IEEE ICC’04, Volume 3, pp. 1796-1800, 2004.
- [82] A. Hamad and A. Kamal, “*Routing and Wavelength Assignment with Power Aware Multicasting in WDM Networks*”, IEEE/CreateNet BROADNETS’05, pp. 33-42, 2005.
- [83] Y. Zhou and G. Poo, “*A New Multi-Wavelength Multicast Wavelength Assignment (MMWA) Algorithm in Wavelength-Routed WDM Networks*”, IEEE ICC’04, Volume 27, Number 1, pp. 1786-1790, 2004.

Appendix - Acronyms

Acronym	Meaning
WDM	Wavelength Division Multiplexing
AOM/AEM	All-Optical/All-Electronic Multicasting

HM	Hybrid Multicasting
MC/MI	Multicast Capable/Multicast Incapable
NSFNET	National Science Foundation Network Backbone
QoS	Quality of Service
DEM	Design Effectiveness Metric
IGMP	Internet Group Management Protocol
RPF	Reverse Path Forwarding
DVMRP	Distance Vector Multicast Routing Protocol
MBone	Multicast Backbone
PIM-dm/PIM-sm	Protocol Independent Multicast-Dense Mode/Sparse Mode
OSPF	Multicast Open Shortest Path First
CBT/OCBT	Core Based Tree/Ordered Core Based Tree
HIP	Hierarchical IP
MALLOC	Multicast Address Allocation
BGMP	Border Gateway Multicast Protocol
M-IGP	Multicast-Interior Gateway Protocol
IETF	Internet Engineering Task Force
DaC/TaC	Drop-and-Continues/Tap-and-Continue
dB	Decibel
NPP/PPR/FPR	No/Partial/Full Packet Replication
SNR	Signal Noise Ratio
ASAP/ALAP	As Soon As Possible/As Late As Possible
EDFA	Erbium-Doped Fiber Amplifier
RWA	Routing and Wavelength Assignment Problem
SMT	Steiner Minimum Tree
SSB	Split Switch Bank
SaD/MOSaD	Splitter-and-Delivery/Multicast-Only Splitter-and-Delivery
AON	All-Optical Network
UTESCD/UTESNCD	Unbalanced Tree with Equal Splitter by Coherent/Non-Coherent Detection
UTFSCD/UTFSNCD	Unbalanced Tree with Flexible Splitter by Coherent/Non-Coherent Detection

DCCD/DCNCD	Drop-and-Continue by Coherent/Non-Coherent Detection
FDCCD/FDCNCD	Flexible Drop-and-Continue by Coherent/Non-Coherent Detection
ARPANET	Advanced Research Project Agency Network
MSNF	Most-Saturated Node First
SA	Simulated Annealing
O/E/O	Optical/Electronic/Optical
MDMCT	Multiple-Destination Minimum-Cost Trail
MCPH	Minimum Cost Path Heuristic
FDBP/PDBP	Full/Partial Destination Blocking Policy
LP/ILP/MILP	Linear/Integer Linear/Mixed Integer Linear Programming
MAX-1 st	MAX-First Algorithm
RR2S/RR2A	Re-Route-to-Source/Re-Route-to-Any
M1 st /M-Only	Member-First/Member-Only Algorithm
VS/PVS/SVS	Virtual Source/Primary and Secondary Virtual Source
VS-CP	Virtual-Source Capacity-Priority Algorithm
CSA	Centralized Splitting Algorithm
TS-AOM	Tabu-Search Based All-Optical Multicasting Algorithm
VS-Based	Virtual Source-Based Algorithm
PPR-AOMH	Partial Packet Replication All-Optical Multicasting Heuristic
ADM	Add-Drop Multiplexer
MCG/MIG	Multicast Capable/Incapable Group
AO-IP	All-Optical IP
OBS/OLS	Optical Burst/Label Switching
MPLS	Multi-Protocol Label Switching
LSP/LST	Label Switched Path/Tree
GB	Guard Band
S-MCAST/TS-MCAST	Separate/Shared-Tree Multicast
M-UCAST	Multiple Unicast
MSC	Multicast Sharing Class
FDL	Fiber Delay line

Table 3: Table of Acronyms