Optical Multicast in Wavelength Switched Networks
- Architectural Framework -
draft-poj-optical-multicast-02.txt

1. Introduction

During the past years, the concept of multicast has been widely discussed for packet-oriented networks. This proposal extends the multicast concept to optical networks, which provide enhanced performance for multicast as well as broadcast-based applications. The applications that can benefit from the optical multicast concept are numerous and cover a wide range of current client needs. These applications mainly include Optical Broadband (Video, HDTV, Multimedia, etc.) applications, Content distribution for Server Mirror Sites and Optical Storage Area Networks (O-SAN).

The architecture of optical multicast-capable networks needs to balance potentially conflicting design and performance criteria, including:
- Minimize the number of nodes traversed (in particular nodes including optical splitters)
- Minimize some combination of the number of transceivers, optical amplifiers and optical splitters
- Maintain a reasonable optical power budget provisioning
- Maximize the connectivity between nodes of interest (i.e. improve the optical network performance)
- Maximize the use of shared resources such as optical splitters
- Minimize the blocking probability (the wavelength continuity problem in multicast networks also referred to as WCRWA Problem; MultiCast Routing and Wavelength Assignment Problem)
- Solving the Routing and Wavelength Assignment (RWA) problem for
both unicast and multicast connections. This contribution reviews the relationships between these tradeoffs.

In the optical domain, whilst optical unicast connections carried on lightpaths are point-to-point connections, optical multicast refers to point-to-multipoint connections which are realized by using light-trees. Nodes traversed by a light-tree must include optical signal power splitters. An optical power splitter divides the power of the input signal into more than one output signal; thus reducing the input signal power by a factor proportional to the number of outputs. The input power can be either equally divided or non-equally divided between several outputs. In the latter case, when the splitting ratio between the output and the input power is tunable per output one refers to a dynamic optical splitter. When this ratio is fixed one simply refers to a static optical power splitter.

Therefore, optical splitter is the key component to realize multicast in optical networks. The concept of optical multicast is becoming possible since 1:2 and 1:4 splitters are widely available and more recently 1:8 splitters have been developed.

The distinction between an optical splitter and an optical tap is related to the output signal ratio. With a splitter the output signal power is distributed in contrast, with a tap the output while signal power ratio is non-equally distributed only a fraction of the signal power is tapped (e.g. sufficient for feeding a host, for instance).

2. Rationale for Optical Multicast

By extending the multicast technology into the optical domain, one offers the following features and advantages which are related to critical aspects of current optical networks:

- Protection: efficient optical channel and line dedicated protection
- Performance: improved performance (no store and forward) compared to packet-oriented multicast technology
- Cost Reduction: reduction of the total number of transceivers deployed in optical networks
- Resource utilization: overall network throughput improvement by reducing the number of wavelengths used per fiber link (i.e. minimizing the overall bandwidth usage per fiber link)

However, the major drawback to overcome in optical multicast-capable networks is to compensate the power penalty introduced during the optical signal splitting leading generally to additional optical amplifiers within the network. This can in turn lead to additional penalties related to the Amplified Spontaneous Emission (ASE) which induces a degradation of the Optical SNR.

Moreover, the multicast problem in communication networks, described by the Steiner Tree Problem, is NP-Complete [6]. The Steiner Tree Problem is defined as follows. Given a Graph G(V,E) where V is a Vertex and E an Edge; a cost function applied to each of the edge of the graph G; and a set of node N included within the graph G; find a sub-tree T = (V[T], E[T]) spanning N and such that its cost C (defined as the sum of individual edge cost C(E)) is minimized. For more detailed explanations concerning this problem and related computational aspects, we refer to [14].

Since the Steiner Tree problem is NP-Complete, some heuristics are needed to setup multicast trees in an efficient way. Notice that this problem is not restricted to the optical domain. Consequently, even if a method to optimize the multicast tree topology of an optical network exists in theory, in practice it is not possible to analytically solve this optimization problem and a heuristic approach is required. For instance, in packet-based multicast networks, shortest-path trees are computed in order to implement a fully distributed on-line computation of the Multicast-tree:\nThe

3. Lightpath and Light-tree Concepts

This section details the concepts of lightpath (Section 3.1) and light-tree (Section 3.2).

In the context of this memo, we assume that an Optical Cross-Connect (OXC) is defined as an O/E/O capable device on each of its interfaces while a Photonic Cross-Connect (PXC) is defined as an O-O device. We also assume that an OXC-based network defines a non-transparent network while a PXC-based network defines an all-optical (or transparent) network. Opaque devices (TDM-based matrixes) are obviously not within the scope of this document.

3.1 Lightpath

A lightpath is defined in the context of wavelength (or optical channel/path) switched networks including either Photonic Cross-Connect (PXC) when referring to all-optical networks or Optical Cross-Connect (OXC) when referring to a non-transparent optical networks.

A unicast connection i.e. a point-to-point connection, in a non-transparent (or all-) optical network is carried on a lightpath. Such connection is defined as a non-transparent (or all-) optical channel and is used to transparently transport packets or circuits...
throughout the optical network. In non-transparent optical networks, the number of transmitters (Tx) needed to setup unicast connections equals the number of receivers (Rx). In all-optical networks, this equality applies as well because we include within the total number of transmitters and receivers the ones belonging to the source and the destination node interfaces of the point-to-point connections (i.e. not only the intermediate nodes).

Example:

Let's consider an optical network including N nodes and M source or destination nodes as indicated here:

```
Papadimitriou et al. Expires May 2002                           4
draft-poj-optical-multicast-02.txt
November 2001
---| D | = D[1]
      /     ---
      /      ...
      /      ---       ---          ---       ---         ---
---       ---          ---       ---         ---
| S |-----| I |-. . . -| I |-----| I |-------| D | = D[m]
---       ---          ---       ---         ---
       ...        ...        ...
      ---       ---          ---       ---         ---
    ----| D | = D[M-1]
    ---      ---

Thus one has, N-M intermediate nodes i.e. nodes which do not terminate optical point-to-point connections. If each of these nodes is defined as OXC then the number of transceivers needed to setup M-1 unicast unidirectional connections from a specific source equals: ((N-M) x (M-1)) + (M-1) Transmitters and ((N-M) x (M-1)) + (M-1) Receivers, in total 2 x ((N-M) x (M-1)) + (M-1) Transceivers. However, if each of the intermediate nodes is defined as PXC then the number of transceivers needed to setup M-1 unidirectional connections from a specific source equals: M-1 Transmitter and M-1 Receivers, in total, 2 x (M-1) transceivers.

More generally, we consider an optical network including N nodes and M source or destination nodes and an average network diameter d coverage for each lightpath. Then, the number of transceivers needed to setup M-1 unicast unidirectional connections from a specific source equals: d x (M-1) Transmitters and d x (M-1) Receivers, in total 2 x (d x (M-1)) Transceivers.

Under the wavelength continuity constraint (i.e. in the absence of wavelength converters) within transparent optical networks, the routing and wavelength assignment (RWA) is the most challenging problem. The RWA problem can be formulated as follows: select the best combination of route and wavelength for each unicast connection such that the number of lightpaths established is maximized and such that two lightpaths sharing a common fiber link do not use the same wavelength. However, the routing and wavelength assignment problem for unicast connections in optical networks has been demonstrated in [4] to be NP-Complete. Therefore, heuristics have been developed in order to decrease the complexity level of the RWA problem.

3.2 Light-tree

The concept of a light-tree has been introduced in [3] and explicitly refers to point-to-multipoint non-looping optical channels established in optical networks. A light-tree is formally defined as a directed Steiner tree [7] that is rooted at the source node and spans all the destination nodes. Each of the intermediate nodes of a light-tree provides optical multicast function i.e. requires optical signal power splitters.

Papadimitriou et al. Expires May 2002                           5
draft-poj-optical-multicast-02.txt
November 2001

A light-tree enables an upstream node to have more than one logical downstream neighbor. This means that one source node can reach through a unique point-to-multipoint connection more than one downstream node considered as destination nodes. Consequently, the use of optical point-to-multipoint connections in all-optical networks reduces the number of transmitters compared to the number of receivers. For non-transparent networks, this difference does not apply because the optical channel is terminated at each O/E/O interface (e.g. a link Amplifier - one transmitter and one receiver). Therefore, for all-optical networks, one additional objective of optical multicast is to minimize the total number of transceivers (e.g. reduce the number transmitters compared to the number of receivers).

Example:

Let's consider an optical network including N nodes, M source or destination nodes and we assume that the N-M intermediate nodes include optical splitters. Thus one has, N-M intermediate nodes i.e. nodes which do not terminate optical point-to-multipoint connections. If each of these nodes is defined as OXC then the number of transceivers needed to setup one multicast (point-to-multipoint) connections from a specific source to M-1 destinations equals: ((N-M) x (M-1)) Transmitters and ((N-M) x (M-1)) Receivers, in total 2(M-1) Transceivers. However, if each of the intermediate nodes is defined as PXC then the number of transceivers needed to setup this multicast connection equals: 1 Transmitter and M-1 Receivers, in total, M-1 Transceivers.

For instance, if N = 32 and M = 8, we need 62 transceivers for an
ONC\textsuperscript{a} based non-transparent network and 8 Transceivers for a PXC-based all-optical network in order to setup a multicast tree from a specific source to M-1 destinations.

Therefore, in all-optical multicast-capable networks, and per light-tree, the number of transmitters always equals 1 and the number of receivers always equals M-1 (i.e. the optimization of the total number of transceivers applies only to source and destination nodes). Moreover, and compared to the situation occurring in non-transparent networks, the minimization of the number of intermediate transceivers is directly transposable to the reduction of the number of PXC sub-interfaces.

More generically, we consider an optical network including N nodes and M source or destination nodes and an average network diameter \( d \) coverage for each lightpath. If each of these nodes is defined as OXC, then the number of transceivers needed to setup one multicast (point-to-multipoint) connections from a specific source to M-1 destinations equals: \( \log(d) \) Transmitters and \( \log(d) \) Receivers, in total \( 2 \times \log(d) \) Transceivers. This by assuming that the relation between the number of destinations (M-1) and the number of intermediate nodes covered by the corresponding tree is logarithmic.

### 3.3 Light-tree Examples

The routing problem in optical multicast-capable networks referred to as multicast-routing and wavelength assignment (MCRWA) problem is NP-Complete since it includes the RWA problem which is NP-Complete as demonstrated in [147]. In other words, the RWA problem is defined as a special case of the MCRWA problem. Therefore, as mentioned in Section 2, one has to define and develop heuristics in order to decrease the MCRWA problem complexity. A complete formulation of the MCRWA problem can be found among other references in [260].

For the purpose of this memo, we propose to extend the multicast tree definition proposed in [3], by allowing light-trees to also cover shortest-path trees (SPT) meaning in this context non-Steiner trees.

Let's take for instance the following optical multicast-capable network where each node has optical signal power splitting capability:

\[
\begin{array}{c}
F \rightarrow \cdots \rightarrow S \\
| 5 \quad 1 \quad 5 \\
A \quad B \quad C \quad D \quad E \\
| 1 \quad 5 \\
I \quad J \\
| 1 \\
\end{array}
\]

**Fig.1 - Optical Multicast-capable Network**

Let's assume that the node A is defined as the root of the following light-tree (see Fig 2.) which covers every node. The members of the light-tree are equal cost nodes while an integer value is assigned to the link cost. The resulting light-tree has a minimized cost starting at node A reaching each of the leaf (or destination nodes) included in the topology.

In this optical network topology, node D, for instance, executes a form of drop-and-continue function since it terminates (i.e. drop) one of the output of the split signal and continues the other output of the split signal toward Node E. The same drop-and-continue operation applies at node B. In this latter case, the incoming signal is split at the output of node B and continued toward node C, F and H.

\[
\begin{array}{c}
F \rightarrow G \\
| 1 \quad 1 \\
5 \\
\end{array}
\]

**Fig.2 - Drop and Continue Function**

If we consider only a subset of destination nodes, for instance D, E, G and J then we obtain the following light-tree which does not correspond to the shortest-path tree but to an optimized tree since signal splitting occurs at the more distant node (i.e. node D) from the root node A. Noticed that this optimized tree corresponds to the Steiner tree.
Fig. 3 - Optimized Multicast Tree

From A to D: Cost = 11
From A to E: Cost = 12
From A to G: Cost = 12
From A to J: Cost = 12
Steiner Tree Cost C(E) = 14

If we had to compute the shortest-path tree for the same group of destination nodes {D, E, G, J} we would obtain the following tree:

Fig. 4 - Shortest-Path Multicast-Tree

From A to D: Cost = 11
From A to E: Cost = 12
From A to G: Cost = 11
From A to J: Cost = 12
Tree Cost C(E) = 19

Therefore, the cost of the Shortest-Path Tree is higher than the one of the corresponding Steiner Tree. The former is thus referred to as a sub-optimal tree.

Notice that the node and link cost minimization is only one among several of the possible constraints applicable to an optical multicast-capable network. If the jitter accumulation has to be minimized in a non-transparent optical network, then a shortest-path tree computation can be used for that purpose.

Consequently, it results from this discussion that point-to-multipoint connections that the logical connectivity of the optical network is increased with respect to the number of wavelengths used within the optical domain. Therefore, when using point-to-multipoint optical channels (or light-trees) we enable communication between a source (root) and a set of destination nodes. This in turn increases the overall network throughput since spare resources (in particular, wavelengths) can be used for other purposes.

4. Virtual Topologies

As described in [5], virtual topology is an abstract concept used to describe a set of light-trees established on top of a specific (physical) optical network topology, which by definition is a fixed topology. Since a light-tree generalizes the optical point-to-point connection carrier i.e. the lightpath, the hierarchical relationship between light-tree and virtual topology can be depicted as follows:
In [55], an efficient methodology has been described to achieve a new constraint-based virtual topology (from the current virtual topology) by simultaneously minimizing the changes required to obtain the new virtual topology. Though the proposed method and computational model are beyond the scope of this document, they demonstrate that topological changes of the virtual topology can also be minimized.

Moreover, since from the network resource utilization perspective a light-tree is an efficient generalization of a lightpath, the set of light-tree based virtual topologies is a superset of the set of lightpath-based virtual topologies. Consequently, an optimum light-tree-based virtual topology is guaranteed to have better performance than an optimum lightpath-based virtual topology.

When considering the following example, the virtual topology can be defined as the set of point-to-multipoint connections (S, L, D) where S is the source node (here Node A and Node K), L is the multicasting wavelength and D is the set of destination nodes. Node A is multicasting on wavelength L1 and Node K is multicasting on wavelength L2. In this example, the sets of destination nodes D are Nodes {C, F, H, J} and Nodes {B, E, G, I} respectively since wavelength L1 reaches Nodes {C, F, H, J} and wavelength L2 reaches Nodes {B, E, G, I}.

Moreover, as mentioned above, the network resource optimization problem in optical multicast-capable networks includes both the minimization of the number of links (distance) on top of the physical network topology and the minimization of the total number of transceivers (and/or sub-interfaces) used in the non-transparent optical network.

5. Optical Splitters and Virtual Topologies

5.1 Optical Splitters

In an optical multicast network, the splitter is the key optical component since it provides the splitting of the optical input signal. An N-way optical splitter is defined as an optical component, which splits the power of the input signal among N outputs; thus reducing the power at the output AoA to 1/k[n] times the original input signal power (∑k/n = 1). In this case, the k[n] value (n = 1, ..., N) is defined as the optical splitting ratio of the output AoA.

A static optical splitter is defined as a particular case when the optical splitting ratio is equally distributed and constant. More precisely, the input power is equally split among N outputs; thus reducing the optical power of each output to (1/N) times the original signal power (N value is called the splitting ratio). Therefore, in this case, the optical splitting ratios k[n] (n = 1, ..., N) are equal and constant.

A special case of an optical splitter is a tap. A tap does not split equally the power of the input signal over the outgoing ports, it only taps a small amount of the signal power from the incoming signal and allows the signal to be continued with negligible power degradation.

A dynamic optical splitter offers dynamic control over the splitting ratio. Such optical splitters provide a tunable splitting ratio per output. Consequently, a static optical splitter appears as a dynamic optical splitter whose optical splitting ratios are fixed. The amount of control depends on the technology being used to realize the splitter. A high granularity dynamic 1:N splitter may offer a continuous splitting capability from 0 to 100% of the input power between any of the output ports to within 1 dB granularity, while a discrete dynamic splitter, based on switching between different static 1:N splitters, may offer only several discrete splitting ratios, for instance 0%, 25%, 50%, 75% and 100%.

This memo is focused on optical multicast implemented through the use of optical static and dynamic splitters only.

Consequently, an optical multicast-capable network will require a more detailed power budget computation since any optical power splitter introduces optical power loss. Therefore, a large number of splitters require a large number of optical amplifiers to balance the optical power loss. Consequently, one of the major constraints when deploying an optical multicast network is to determine a design that reduces the total number of optical splitters allocated with minimal effects on network performance (i.e., keeping the network...
blocking probability low) while minimizing the number of optical amplifiers.

5.2 Relation with Virtual Topologies

Simultaneously, one has to reduce the impact on the power budget when multicast-tree virtual topology (see Section 4) changes by minimizing the total number of optical amplifiers used on each link. This means that either (1) the power budget is over-provisioned i.e. any topology fits into the optical splitter allocation within the network or (2) the power budget (in addition to the optical splitter availability constraint) decrease the potential number of virtual topologies. In the first case, the optical splitters allocation within the network give the capability to split the optical signal without any knowledge about the optical characteristics of that signal (BER, Q-Factor, etc.). As proposed in [8], a power-efficient design methodology can be based on a two-dimensional design space:

- The splitter-sharing dimension focused on minimizing the number of splitters inside the PXC itself.
- The splitter-cover dimension focused on minimizing the number of multicast-capable PXCâs in the overall optical network taking into account the blocking probability.

The optical splitter placement problem (or allocation problem) is also NP-complete (since it includes the RWA problem which is NP-complete). Notice that the most important result obtained in [9] is that for a wide variety of optical network topologies and different traffic patterns no more than 50% of the PXCâs need to include optical power splitters. Having more than 50% of PXCâs with multicast-capable hardware only slightly enhances the blocking performance of the optical network.

5.3 Dynamic Optical Splitters

By considering that optical splitter are static components if one wants to avoid blocking during multicast tree setup due to power budget constraints the optical network power budget has to be over-provisioned. This implies that the power loss due to the optical signal splitting does not restrict the multicast tree virtual topologies and this independently of the launched power. In order to avoid such over-provisioning requirement, one has to consider dynamic optical splitters as defined in Section 5.1.

The consequence of using such dynamic splitters is that one assumes implicitly that multicast tree setup might be blocking for certain particular requests due to power budget constraints. However, in most cases, the dynamic optical power splitting ratio must allow allocating the exact output power in order to reach a given destination (a multicast tree leaf). This implies that nodes dynamically splitting the optical signal must be aware of the underlying optical routing impairments such as Optical SNR.

Several alternatives must be considered when dynamic optical splitters are deployed:

1. Localized Approach: the optical splitting ratio of each output is determined independently so that optical splitting ratio fine-tuning is only performed locally; thus only impacting the downstream nodes; for instance, the dynamic process is restricted to the node performing the last splitting before reaching the destination nodes (i.e. leaves).

2. Recursive Approach: the optical splitting ratio of each output is recursively determined by a cascading of optical splitting ratio fine-tuning from the root node until the destination nodes or from the destination nodes to the root node; in this case the optical splitting ratio fine-tuning is not (necessarily) performed locally.

3. Global Approach: the optical splitting ratio is determined by optimally fine-tuning the optical splitting ratio along the downstream direction from the root towards given destinations (leaves).

5.3.1 Localized and Recursive Approach

Let us consider the following optical network topology in order to illustrate both localized and recursive alternatives:

![Fig.5 - Dynamic Tuning of Optical Splitting Ratio](image)

The multicast tree having node A as root and (E, G, I, J) as set of destination nodes is considered as the initial tree configuration. In this situation node D performs equal ratio optical signal splitting since each of the downstream links has the same cost. Now, consider that Node L wants to join the set of leaf of this tree. Since the link connecting node L to node J has a cost of 5, node J must now request more power in order to be able to locally drop the signal and continue it to fulfill the node L request. Notice that
one assumes here that there exists an explicit relationship between the cost and the optical power available on the corresponding link. Therefore, since the input signal power reaching node J is determined by the optical splitting ratio of the node J output (at node D), node D has to change the power ratio it allocates towards node J. In such conditions, (1) if the input power P arriving at node D was higher than needed to reach simultaneously the initial set of destination nodes and (2) if the difference between the additional power needed to reach additionally node L and the unused power is greater than zero then adding node L to the multicast tree can be performed at node D by allocating additional output power toward node J.

If such conditions are not met, then depending on selection criteria such as node cost and precedence either the request is rejected or node D initiates another input signal power change request toward node C. Another alternative would be to remove some of the leaves belonging to the initial set of destination nodes in order to recuperate the corresponding output signal power and allocate it to new request.

Therefore, the generic procedure to dynamically determine the optical splitting ratio of the output $\bar{A}_n$ at level $M-1$ is defined by $k[n,M-1]$; at level $M$ (leaf), the input power signal equals the output power signal (dropped signal)

When adding a leaf at level $M$ (leading to a tree of depth $M+1$) from output $n$:
- Apply the following recursive procedure until reaching a node enabling optical splitting ratio fine-tuning (e.g. for $i=M; i>0; i=i-1$)
  - If level $\bar{A}_i$ is not equal to 0 (root level) and if input power signal $k[n,i]$ does not enable to continue the signal until the requestor node (leaf) then initiate an optical splitting ratio change request for $k[n,i-1]$
  - Otherwise fine-tune the optical splitting ratio $k[n,i-1]$ and exit this recursive procedure

As expected this procedure is rather straightforward and some constraints must be defined in order to reduce as much as possible the impact of an optical splitting ratio for a given output at a specific node. The most obvious is that when such change occurs, the N-1 other output should not be disturbed by such modification. This means that to some extent if one would like to achieve a dynamic optical multicast-capable network, launched power over-capacity will be needed which may in turn, request for some additional attenuation device within the optical network.

Notice, as well that this procedure must be reversible in order to enable the re-distribution of the optical power when a delete-branch request is initiated by receiver node. This allows the AretunA of the optical power in the upstream direction therefore, avoiding the starvation of optical power in the middle of the multicast tree.

## 5.3.2 Global Approach

In order to illustrate the global alternative, let us again refer to the network topology in Fig.-6. The multicast tree having node A as root and {E, G, I, J} as set of destination nodes is again considered as the initial tree configuration.

However since now a global power distribution scheme is considered, when Node L wants to join the set of leaves of this tree the request is allowed to progress to the root (node A) along with the necessary link information. The root node may then decide globally how to allocate the optical power along the optical path (in the multicast tree) that connects it to the new leaf (node L). Notice that the power re-distribution must be done without disturbing the N-1 other outputs.

The distribution of power throughout the path can be handled using two possible alternatives:
- A minimum power approach that allocates just enough optical power to each of the leaves.
- A best performance approach that allocates excess optical power throughout the network.

The minimum power approach may use a distribution algorithm similar to the recursive approach detailed earlier in this section assuming that all the optical power reserve is kept at the root node A. Therefore, the generic procedure to dynamically determine the optical splitting ratio throughout the optical path would follow these steps when a specific leaf is attempting to join a multicast tree of deep $M$:

Determine the new optical splitting ratios along the optical path:
- Initial conditions:
  - The optical splitting ratio for a specific output $\bar{A}_n$ at
Level M-1 is defined by \( k[n, M-1] \).
- At level M (leaf), the input power signal equals the output power signal (dropped signal).

When adding a leaf at level M (leading to a tree of depth \( M+1 \)) from output n:
- Apply the following recursive procedure to determine the needed optical power level/splitting ratio until reaching the root (e.g. for \( i=M; i>0; i=i-1 \))
  - If level \( i \) is not equal to 0 (root level) then calculate the new needed optical splitting ratio (and total optical power needed at level \( i \)) using the needed power to reach each tree node at level \( i+1 \).
  - Then, send an optical splitting ratio change request for \( k[n,i-1] \) towards the upstream direction.
  - Otherwise fine-tune the optical splitting ratio \( k[n,i-1] \) and exit this recursive procedure.

As in the recursive procedure this procedure is rather straightforward but some constraints must be defined in order to reduce as much as possible the impact of an optical splitting ratio for a given output at a specific node. The most obvious is that when such change occurs, the \( N-1 \) other outputs should not be disturbed by such modification. This means that to some extent if one would like to achieve a dynamic optical multicast-capable network, launched power over-capacity is needed which may in turn require some additional attenuation devices within the optical network and that the optical splitting ratio fine tuning must be performed in a fast closed loop. This, in turn, ensures that once the optical power at the input to the node at level \( N \) increases this increase is only distributed to the desired output \( n \) and that the transients are minimal.

The best performance approach requires a different power distribution scheme that maps the reserve power throughout the network. In this case once the root node has determined the new minimum required power distribution it must specify how to redistribute the reserve power and initiate optical split ratio fine-tuning throughout the network. This approach has the advantage of better use of the extra optical power available, however, it complicates the power redistribution scheme (especially when one does not wish to disturb the \( N-1 \) other connections).

Notice, as well, that (like in the recursive case), both these procedures must be reversible in order to enable the re-distribution of the optical power when a delete-branch request is initiated by receiver node.

6. Applications of Optical Multicast

Applications that can be realized using the optical multicast capabilities include Multimedia Applications (covering HDTV and Video), Content Distribution in Server Mirror Sites and in Optical Storage Area Networks.

Moreover, optical multicast provides a very efficient implementation of dual-homing protection i.e. dedicated 1+1 protection for optical channels and optical lines. Such protection scheme can be applied to deploy efficient (all-)optical rings survivability mechanisms. Optical multicast would in turn allow defining so-called sub-network connection protection between optical signal splitter capable devices.

To illustrate this mechanism, which is not restricted to optical ring topologies, let\(\text{\`s consider the example given in Section 3.}\)

\[ \begin{align*}
&5 \\
&\downarrow \\
&6 \\
&\downarrow \\
&1 \\
&\downarrow \\
&A \\
&\downarrow \\
&B \\
&\downarrow \\
&C \\
&\downarrow \\
&D \\
&\downarrow \\
&1 \\
&\downarrow \\
&E \\
&\downarrow \\
&F \\
&\downarrow \\
&G
\end{align*} \]

Fig. 6 - Optical Multicast-Capable Optical Domain

If node B and node D have optical signal splitting capabilities, then the following scenario can be considered. One needs to establish a protected connection between source node A and destination node E knowing that:
- C is defined with a high failure probability (or all resources on this node have been already allocated) so that one must exclude node C for such connection service request.
- Alternate paths through nodes F and G as well as through nodes H, I and J have enough free resources to accommodate such request.

Consequently, one could have the following primary (working) and secondary (protection) segments established between nodes B and D:
- path \( [A, B, F, G, D, E] \) is defined as the primary while \( [A, B, H, I, J, D, E] \) is defined as the secondary optical channel. Notice that due to the signal splitting at node B, node D receives both primary
and protection signals and selects the primary one. Therefore, when a failure occurs at one of the network resources allocated to the primary segment \([B, F, G, D]\) switching to the protection segment \([B, H, I, J, D]\) is straightforward.

Moreover, the application of the optical multicast concept will also facilitate the implementation of the Drop-and-Continue functionality between optical rings as described in \[10\].

However, as described in \[11\], the scheduling of multicast traffic in an optical multicast-capable network needs to take into account two conflicting objectives: low bandwidth utilization at the packet-layer and high optical channel capacity utilization. Moreover, as stated in \[12\] more practical traffic patterns and real network scenarios must be considered, since in optical network topologies both unicast and multicast traffic will be simultaneously present within the same optical network.

Papadimitriou et al. Expires May 2002 17

draft-pojo-optical-multicast-02.txt November 2001

Taking into account these considerations is key to a successful development and deployment of optical multicast-capable networks.

7. Signalling in Optical Multicast-capable Networks

It is widely accepted that point-to-multipoint applications such as HDTV and Video Streams will comprise a large portion of the multicast application space. The proposed signalling model is designed to efficiently handle such cases where sources are well known. This is specifically the case with optical multicast were the sources \(S\) are in most cases well defined (location, address, etc.)

Signalling in optical multicast-capable networks is realized through GMPLS signalling as described in \[GMPLS-SIG\]. Since an optical multicast tree is unidirectional, the suggested label mechanism defined for bi-directional LSP setup is not applicable here.

In this section, a node is independently referred (at the signalling plane level) to as a PXC or an OXC or using the GMPLS terminology \[18\] as an LSR. A lightpath is referred here to as a Lambda-LSP (L-LSP) or simply an LSP.

7.1 Optical Network Architecture

The Optical Cross-Connect (OXC) system architecture for lambda switching has been introduced in the \[16\] and \[17\]. The OXC interfaces considered includes, as described in \[18\], Lambda Switch Capable (LSC) interfaces that switches the lightpath segment based on the incoming wavelength and Fiber-Switch Capable (FSC) interfaces that switches the lightpath based on the spatial position of the incoming data stream in the physical space.

An OXC has several incoming and outgoing LSC interfaces or ports, connected to adjacent OXCs, and several incoming and outgoing LSC interfaces or ports attached to an edge device that can be a router (or any other kind of device supporting termination capable interfaces). An OXC includes mainly two functional parts: an OXC Switch Controller (OSCtrl) and OXC optical matrix.

The OSCtrl communicates through Optical Supervisory Channels (OSC) i.e. out-of-band signalling transport mechanism or through dedicated and physically diverse control network i.e. out-of-network signalling transport mechanism. OSCtrl and OSC define the signalling plane of the optical network.

When receiving a signalling message, the OSCtrl translate to an internal control command, and sends this command to the OXC optical matrix. The commands to control the OXC optical matrix are as follows: connect (and disconnect) between an incoming LSC interface and an outgoing LSC interface. The OSCtrl is also capable to process status requests, lightpath modification requests as well as notification messages. The same commands apply when the OXC is connected to an LSC capable edge device.

Based on these commands, a chain of connections through OXCs can form a point-to-point optical channel i.e. a lightpath (as described in section 3.1). The ingress OXC in the node where the multicast tree starts (sender), the egress OXC(s) are the destination nodes (receivers) and the OXCs in the middle of the tree are referred to as intermediate OXCs.

An OXC optical matrix receives commands from the OSCtrl, and replies whether the command was successful or not. The OSCtrl then converts the result into a message that it sends via the OSC channel throughout the signalling plane of the optical network.

To cover any kind of optical network, we consider as specified in \[16\] the following distinction between Optical Cross-Connect (OXC) in non-transparent optical networks and Photonic Cross-Connect (PXC) in all-optical networks. Basically, OXC devices included within non-transparent optical networks performs O/E/O conversion at each of their interfaces while PXC devices included within optical networks do not perform O/E/O conversion at all so they are defined as pure O/O devices. We consider here that all the interfaces of an OXC as well as all the interfaces of a PXC are Lambda Switching Capable (LSC) interfaces. Moreover, in order to enable optical multicast functionality, OXCs and PXCs must include optical splitters contained into splitter banks.
In the remaining parts of this document, nodes are considered as optical PXCs but these considerations can also be applied to OXCs.

### 7.2 Shortest-Path Trees

Typically in packet-networks, multicast routing protocols create shortest-path trees. These trees are non-optimal (as described in section 5), but the algorithm is distributed and allows the dynamic adding and removal of the multicast tree branches without affecting the existing tree. The mapping of these trees onto LSPs is discussed for Protocol Independent Multicast (PIM) Sparse Mode (PIM-SM) in [19] and more specifically in [20]. The PIM-SM protocol is used in two multicast models: Any-Source Multicast (ASM) and Source-Specific Multicast (SSM). Therefore, one refers to the models as PIM-SM for ASM and PIM-SM for SSM, respectively.

PIM Sparse Mode (PIM-SM) is defined as receiver initiated approaches. One of the key issues with this approach is the merging of user signalling-plane with the optical signalling-plane when considering a unified service model between the edge router and the optical domain.

#### 7.2.1 Source/Share Multicast Trees

In classical packet multicast [19], IP multicast routing protocols create either source trees (S, G), i.e. a tree per source (S) and per multicast group (G), or shared trees (*, G), i.e. one tree per multicast group (Figure 1).

![Fig.7 - Shared and Source Trees](image)

In packet-based multicast-capable networks, the advantage of using shared trees, when label switching is applied, is that shared trees consume less labels than source trees (1 label per group versus 1 label per source and per group).

#### 7.2.2 PIM-SM for ASM

PIM Sparse Mode (PIM-SM) enables to construct unidirectional shared or source trees (see Fig.10) to forward data from senders to receivers of a multicast group. Moreover in [20], PIM-SM (PIMv1 and PIMv2) has been extended to combine MPLS label distribution with the distribution of (*,G) join state, (S,G) join state, or (S,G) RPT-bit prune state. Note that (*,G) join state implications need to be evaluated in optical multicast because the elimination of the shared tree concept is the key to implement SSM. This because SSM eliminates the need for starting with a shared tree and then switching to a source-specific tree (see section 7.2.3).

In PIM-SM MPLS extension, labels and multicast routes are sent together in one message. Therefore, we propose here to extend this method for optical multicast-capable networks. We will refer to this method as Optical PIM-SM.

The design described in [20], specifies that Multicast label distribution procedures should not depend on the media type. Note also that when a multicast routing table change requires a label distribution change, the latency between the two should be minimized, both to improve performance and to minimize the possibility of race conditions.

On receiving PIM Join/Prune messages from an edge router, a PXC that supports optical multicast (i.e. that includes optical splitters) sends PIM Join/Prune messages to upstream neighboring PXCs toward the Rendezvous Point (RP) for the shared-tree (*,G) or toward a source for a source-tree (S,G). Wavelength Labels, described in [18], are distributed by being associated with IPv4 multicast addresses (address issue is considered in section 7.2.3) in the join list or the prune list. As described in [19], the multicast tree setup is performed using downstream label distribution with independent LDP control so that requesting nodes can join independently the multicast tree.
Router R and PXCs are included in the same signalling plane i.e. the mechanism described here applies to a unified service model [18]. So that the generalization of the mechanism defined in packet multicast-capable networks to the optical multicast-capable networks need to be considered as described in [22].

7.2.3 PIM-SM for SSM

Source Specific Multicast (SSM) defines a method of multicast forwarding restricted to shortest path trees to specific sources. The key difference between ASM and SSM is that the latter eliminates the need for starting with a (*,G) shared tree and then switching to a source-specific tree. More precisely, when a LSR (for instance, a boundary PXC) identifies request to join a specific source with a SSM group address (232/8), it always initiates a (S,G) join and never a (*,G) join.

SSM enables that a single source S can transmit to a channel (S,G) where G is an SSM address. Each receiver is capable of specifying the specific sources from which it would like to receive content. This would in turn enable a direct mapping of SSM to optical multicast.

7.2.4 PIM-SM: ASM versus SSM

As discussed in the previous sections, shared trees (*,G) join state are only used to discover new sources of the group and after a switchover to a source tree is performed. However, the end-to-end mapping of a (*,G) shared tree implies the setup of multipoint-to-multipoint trees (i.e. light-trees). The problem is that in the optical domain, labels represent wavelengths, which are by definition non-mergeable. Consequently, optical multipoint-to-multipoint light-trees are not applicable in optical multicast since label merging is not feasible.

Since the difference between ASM and SSM is that the latter eliminates the need for starting with a (*,G) shared tree and then switching to a source-specific tree, this restricts the choice to SSM for shortest-path tree receiver initiated approach.

7.3 Non-Shortest-Path Trees

The shortest-path tree approach could be taken for optical multicast but there are several reasons to consider other means:

- The diversity of the nodes in optical networks: PXCs without and with support for multicast, the latter can be Splitter and Delivery (SaDs) or Multicast optical SaDs (MoSads) [8]
- The multicast light-trees tend to be more stable (branches added and removed sporadically)
- Shortest-path trees are not optimal (as described previously)
- Internal PXCs might not be aware of global routing information (when using an overlay signalling plane), so a PXE would be able to determine where to send a PIM Join if the source S is outside of the optical domain
- MEWS constraints (see section 4)
- MCRWA constraints (see section 4)

Because of the above reasons, it might be useful to consider a non-shortest-path tree approach combined with an offline traffic-engineering tool. The offline tool performs the multicast tree route computation by taking into account the variety of constraints set by optical networks. The offline tool computation can provide a more optimal multicast tree than the shortest-path tree one. Many heuristics exist to approximate e.g. the optimal Steiner Tree [18].

7.3.1 Root Initiated

The explicit routing signalling is based on the computation of the...
explicit multicast route from the root to each of the (intermediate) destination nodes.

In the root-initiated case, the signalling flows from root to destinations. This approach requires that a multicast tree route object [19] is signaled. For instance, if we refer to the previous example:

```
F                       G
|                          |
5          1           5
A ------- B --------- C --------- D          E
|                          |
1
H I J
```

Fig.9 Â Root Initiated Approach

The explicit multicast tree route that the node A (or the offline tool) has computed can be represented as: A [B [C [D [G, J]]]]

More precisely, as defined in [20], each of the link bundles of a PXC (i.e. PXC-LSR) can be numbered (by an IPv4 address) or unnumbered (by a Node ID and a Link bundle index). So that the identification of the hops included in the explicit multicast tree route can be expressed by using this addressing scheme.

Therefore, multicast explicit route can be included within the Explicit Route TLV (ER-TLV) described in [21]. The content of an ER-TLV is defined as a series of variable length ER-Hop TLVs. We consider here ER-Hop Type being IPv4 prefix. The IPv4 prefix represents the IP address of the link bundle of each node through which the multicast tree need to be established. The L bit defines whether the value of the attribute is loose. In this case, the value of the attribute is strict so that the L bit is not set.

```
+---------+
---------------------| Offline |---------------------
|             --------|  Tool   |--------             |
|            |        +---------+        |            |
|            |             |             |            |
|            |             |             |            |
|            |             |             |            |
|            |             |             |            |
---          ---           ---           ---          ---
| A |--------| B |---------| C |---------| D |--------| E |
--- =======> --- ========> --- ========> ---          ---
| ^    (1)    |     (2)     |     (3)     |x           |
|+           |             |             |x (4)       |
|+           |             |             |x           |
|+          ---           ---           --- Leaf      |
```  

Then the root of the multicast tree sends a Label Request message along with the explicit tree TLV. The subsequent PXC-LSR looks up its downstream PXC-LSR in the explicit tree object of the label request message. Then it sends the Label Request message to these downstream routers. After the PXC-LSR receives the Label Mapping messages from the downstream PXC-LSR, it allocates a label itself, installs this point-to-multipoint forwarding entry into the forwarding table and sends a Label Mapping message to its upstream PXC-LSR.

This tree can be torn down by the Label Release messages sent from the root to all the leaves. When a node receives a Label Release message, it takes the MPLS forwarding entry out of the forwarding table, and sends a Label Release message to every downstream PXC-LSR.

For an edge router to become a member of a particular optical multicast group, the router has to register to the multicast group membership with a specific query to the PXC that handles the multicast group membership. Protocol like IGMP could be used for that purpose. To construct a distribution tree connecting all the edge routers, multicast-capable PXC-LSRs then exchange messages with each other by using Multicast Traffic-Engineering extension to IGP routing protocols. Details concerning these multicast traffic-engineering extensions are left for further study.
7.3.2 Receiver Initiated

Alternatively the signalling can start at the receiver side [20]. The signalling starting at a receiver is sent in the direction of the root along an explicit path. In this case, the (wavelength) label distribution node is a downstream-on-demand distribution and the LSP control is ordered. However, the signalling goes in the upstream direction from the destination node (receiver) to the source node (sender).

The Downstream on Demand procedures apply to multicast distribution trees. Independent LSP control is needed so that different downstream branches of a multicast distribution tree can join the tree independently.

When using the receiver initiated approach, the multicast-tree is built from the leaves to the root. In this case, the tree can be centrally calculated by an offline tool and the reverse path from the root to each leaf or the reverse path from the leaf to root can also be calculated by the leaf of the tree, based on e.g. Multicast TE-Routing information. Note that in former case, the pruning of one leaf and the subsequent new tree calculation by the offline tool (not necessarily centralized) can affect other branches than the one of the removed leaf.

Each leaf node sends a Join message (note that this is not a multicast routing Join message, but an extension to an MPLS signalling protocol as defined in [26]) with the explicit reverse path and a label towards the root. At the subsequent upstream PXC-LSR, the Join messages of the same tree are merged, a label is allocated, the point-to-multipoint forwarding entry is installed into the forwarding table, and a Join with the newly allocated label and the explicit reverse path object is sent to the upstream PXC-LSR. Notice, that this approach encompass both (S,G) and (*,G) Join messages, however as stated in Section 7.2, only (S,G) Join messages will be used to setup optical multicast trees.

When a Join message reaches an on-tree PXC-LSR, it processes the message, modifies the forwarding entry with the label assigned by the newly joined downstream PXC-LSR and finishes the join procedure.

When a leaf node wants to leave the group, it sends a Label Withdraw to its upstream router. When all the downstream neighbors of a PXC-LSR leave the group, it should send a Label Withdraw to its upstream neighbor.

7.3.4 Additional Considerations

After calculating the multicast tree, the offline tool can decide to already assign the wavelength for every link on every branch. Alternatively, it can leave it up to the individual PXCs and send the explicit multicast route to the source node and setup the corresponding light-tree through GMPLS signalling.

In packet networks unicast and multicast have their own separate label spaces, this is no longer true for optical networks since wavelengths are physical, non-shareable and non-mergeable entities. If the PXCs perform the label allocation it would be advisable to define an allocation mechanism (downstream, downstream-on-demand, etc.) that is similar to the one used for unicast traffic. This in order to avoid simultaneous allocation of the same wavelength by the upstream and the downstream PXC.

7.4 Other GMPLS Issues

Other GMPLS signalling considerations such as hierarchical label distribution, whether a multicast connection can hold point to point higher layer connections over it and related traffic-engineering aspects are left for further study.

7.5 User-plane Signalling

Up to now, this memo does not address one of the key issues with optical multicast: the merging of user signalling-plane with the optical signalling-plane when considering a unified service model between the edge router and the optical domain.

By applying one of the approaches defined in the previous sections, we obtain an optical light-tree (which by definition is strictly unidirectional) defined at the transport-plane level from the source S to the destination. So that a single source router [root] can reach simultaneously several destinations belonging to the same multicast group. However, this document does not describe a mechanism allowing specific destinations to reach a source router at the packet LSP level (i.e. at the user-plane level) since these considerations are out of the scope of this document.

If the same administrative authority manages both the edge routers and the optical network, then a peer relationship between the PIM instance of the optical network and the client network can be defined as depicted in the following figure:
Otherwise, if the edge routers and the optical network belong to distinct administrative authorities, then a mapping between PIM instances need to be defined (see Figure 14).

Fig. 11 Â Peer Relationship

Fig. 12 Â Layered PIM Instances

Needless to say that such considerations need further developments which are clearly outside of the scope of this document.

8. Security Considerations

Security-related issues are left for further considerations.

9. References

Papadimitriou et al. Expires May 2002


[12] H. Pernos et al., ÂScheduled Combined Unicast and Multicast Traffic in Broadcast WDM Networks,Â Photonic Network
Communications, Volume 2, Number 2, May 2000


Acknowledgments
The authors would like to thank Bernard Sales, Emmanuel Desmet, Hans De Neve, Fabrice Poppe and Gert Grammel for their constructive comments.

Author’s Addresses

Papadimitriou Dimitri
Alcatel
Francis Welleplein 1,
B-2018 Antwerpen, Belgium
Phone: +32 3 240-8491
Email: dimitri.papadimitriou@alcatel.be

Ooms Dick
Alcatel
Francis Welleplein 1,
B-2018 Antwerpen, Belgium
Phone: +32 3 240-7893
Email: dirk.ooms@alcatel.be

Jim Jones
Alcatel
3400 W. Plano Parkway,
Plano, TX 75075, USA
Phone: +1 972 519-2744
Email: jim.d.jones1@usa.alcatel.com

Eric Mannie
Ebone
Terhulpsteenweg 6A
1560 Hoeilaart, Belgium
Phone: +32 2 658-5652
Email: eric.mannie@ebone.com

Roy Appelman
Civcom (www.civcom.com)
Phone: +1 972 3 922-9229

Papadimitriou et al. Expires May 2002 28
draft-poj-optical-multicast-02.txt November 2001
Full Copyright Statement

"Copyright (C) The Internet Society (date). All Rights Reserved.
This document and translations of it may be copied and furnished to
others, and derivative works that comment on or otherwise explain it
or assist in its implementation may be prepared, copied, published
and distributed, in whole or in part, without restriction of any
kind, provided that the above copyright notice and this paragraph
are included on all such copies and derivative works. However, this
document itself may not be modified in any way, such as by removing
the copyright notice or references to the Internet Society or other
Internet organizations, except as needed for the purpose of
developing Internet standards in which case the procedures for
copyrights defined in the Internet Standards process must be
followed, or as required to translate it into languages other than
English.

The limited permissions granted above are perpetual and will not be
revoked by the Internet Society or its successors or assigns.

This document and the information contained herein is provided on an
"AS IS" basis and THE INTERNET SOCIETY AND THE INTERNET ENGINEERING
TASK FORCE DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING
BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION
HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF
MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE."