Restoration of Multi-Cast Trees in Optical-Beared Grooming-Capable Two-Layer Networks

Tibor Cinkler, Marcell Perényi, Péter Soproni
Department of Telecommunications and Media Informatics (www.tmit.bme.hu)
Budapest University of Technology and Economics (www.bme.hu)
Magyar tudósok körútja 2, H-1117 Budapest, Hungary
E-mail: {cinkler, perenyim, soproni}@tmit.bme.hu

***Abstract*—We consider dynamically changing multicast trees (light-trees) in two-layer optical-beared grooming-capable networks. The continuous changing of the tree "leaves" causes the degradation of the tree in time. Therefore, a huge amount of network resources can be spared by reconfigurations performed periodically or upon failures and reparations.**

In this paper we focus onto restoration of trees if a link (or any other network element) fails. The failures are more critical if they affect the tree closer to its root, while less critical if closer to a leave. We propose and evaluate four simple restoration strategies and investigate their performance for different multicast routing algorithms.

I. On Multicast and Broadcast

Multicasting and broadcasting TV programs over the networks is one of the key services offered by most telecom operators nowadays. This service requires high network availability since a failure can affect a very large number of users. These connections should be able to survive even multiple simultaneous link failures.

There are similar services that require setting up multicast trees. These include caching for VoD, and particularly for streaming video services, peer-casting the decoded and transcoded content, VPLS (Virtual Private LAN Services), anycast for GRID services, etc.

Currently, there exist multiple transport alternatives for multicasting video distribution in metro and core networks. Operators can choose among layer 1 (ngSDH, OTH (OTN)), layer 2 (PBB, PBB-TE, T-MPLS, RPR) and layer 3 (IP/MPLS) transport solutions. While most of these technologies already include protection mechanisms, the development of restoration mechanisms for multi-cast services is still an open issue. This paper proposes four simple restoration schemes and provides a performance analysis for multicast TV services in metro and core networks.

Broadcast TV traffic volume does not depend on the number of customers but on the number, definition and encoding of TV channels. So, TV traffic volume would be similar in the metro access and metro core segments. For example, 100 HDTV channels, with MPEG 4 encoding, will need 1 Gbps from the TV Head-End to the rest of Service PoPs (points of presence) in the metro core, and from the Service PoP to the access nodes in the metro access.

The work described in this paper was carried out with the support of the project BONE ("Building the Future Optical Network in Europe"), a Network of Excellence funded by the European Commission through the 7th ICT-Framework Programme.

II. Multicast/Broadcast Solutions for Core Networks

In core networks the video content is distributed (multicasted) in bundles of tens to hundreds of programs to the metro networks. Depending on the resolution and encoding of certain program channels this requires a capacity from 100 Mbps to a few Gbps. Therefore, some multi-casts having smaller bandwidth requirement can share a single wavelength path, while others that exhaust the capacity of a wavelength channel may even require multiple wavelength channels bundled together.

We assume a two-layer network architecture, where the upper layer is an asynchronous time switched one, e.g., IP transport over Ethernet and/or MPLS while the lower layer is a circuit switched one, based on wavelength division multiplexing (typically DWDM eventually with OTN framing).

In such a two-layer architecture we assume multi-casting capability at both layers. At the upper, IP/MPLS/Ethernet layer multi-cast is supported, by sending the same packets to two or more outgoing ports. This increases the load of the backplane of the switch. At the lower, optical layer the multi-cast is done physically, i.e., the signal, as well as its power is divided among two or multiple outgoing ports. This approach requires splitters in the optical switches that although not yet supported by many manufacturers can be done by a simple and cheap splitter.

We also assume grooming in our approach as follows. If there are two or more sub-lambda traffic streams that use the same path in a part of the network, they can be groomed together into a single wavelength channel by any grooming capable node as well as they can be separated (de-multiplexed) again by any other grooming capable node. This leads to much better resource utilisation.

This two-layer network is represented as a single graph, with as many parallel edges between certain nodes as many wavelengths are supported over that link and using sub-graphs connecting these parallel edges in nodes to model different functionality including cross-connecting, grooming, etc. The model of this network is discussed in our earlier papers, including [1] and [2].

III. Optical (O) or Electronic (E) Multicasting?

Assuming two network layers,

- an upper, electronic, packet switching capable one and
- a lower, optical, circuit switching capable one,

multicast can be performed by any of them. However, while in the electronic domain an incoming packet is forwarded to two or more output ports, in the optical domain the whole optical signal (e.g., a wavelength) is physically split (e.g., by a directional coupler) to two or more parts without access to sub-lambda granularity multiplexing or switching capability.

Clearly, a compound of multiple channels that fills the capacity of a wavelength path, the 'O' multicast seems more efficient, while for smaller granularity 'E' multicasting is preferred.

IV. Methods for Multicast Routing

We have assumed that a single source (the root of the tree) supplies a few sinks, (destinations, leaves of the tree). This is a special Steiner tree, where the idea is to carry the information in a single exemplar (copy) as long as possible and to multiply at the farthest node to use as few capacity as possible for the whole multi-cast connection (tree). However, there are two constraints. Both upper (electronic) and lower (optical) layer multi-cast capabilities have breadth limitations, i.e., each node has limitation to how many output ports can it copy the same content. Furthermore, the depth of the tree, i.e., the largest source-destination distance has to be limited as well.

Here we have evaluated the following three multi-cast routing methods that we proposed earlier in [1] and [2]: ASP, MPH and ILP.

- **ASP**: Accumulative Shortest Path (Dijkstra)
  This method is the fastest and simplest one however the results it provides are suboptimal. The root-to-leave demands are not routed at once simultaneously, but in a sequence one after the other using Dijkstra's algorithm. The idea is that the cost of elements (links in the wavelength graph) already used by a root-to-leave demand of the same tree is set to zero, that means it can be used for free for all future root-to-leave demands of the same tree. Of course the chosen sequence significantly influences the result.
- **MPH**: Minimal Path Heuristic
  We have adapted [3] to our wavelength graph model. The idea is that we calculate the shortest path in our wavelength graph model between all leaves and between the leaves and the route. This results in a complete graph where the number of vertices equals to the number of leaves plus one for the root. In this simpler graph Prim's algorithm [4] is used to find the least cost spanning tree. This minimum spanning tree is then traced back to the wavelength graph. Analogously to the ASP, where a new demand joins the tree, here while reconnecting the cut leaves the costs of all already used edges are set to zero.
- **ILP**: Integer Linear Programming
  Since this method provides always the global optimum in terms of the objective function this was the reference method to compare other methods to. The time requirements for ILP were the largest among the three methods ranging from a few to a few hundred seconds in our case. The ILP formulation was proposed and explained in our earlier paper [1].

V. Methods for Multicast Restoration

If a link or a node fails in the network it will affect all the multi-cast connections that use that element. However, if this element is just a leave (a single user) its failure will affect only that user, however if an element close to the source (to the root of the tree) fails, than typically many leaves (end users) will be cut from the source. We propose methods for all the cases that reconnect the cut leaves (users) or whole branches (groups of users) to the healthy part of the tree or directly to the source.

Here we propose and discuss the different methods for restoring the multicast trees upon failures. The four methods we propose for restoration(ASP, ASP partial, ILP and ILP partial) are based on methods for routing as follows.

- **ASP**
  ASP restoration can be applied to any tree that was set up by any algorithm. Its idea is that if a link fails it can cut a single, or multiple, or even all the leaves from the root. We use here Dijkstra's algorithm to find a new path from each cut leave to the root, where the costs of already used links are set to zero as explained for the ASP routing.
- **ASP Partial**
  ASP partial restoration is a kind of link restoration, i.e., if a branch of the tree is cut, then the whole branch as it is will be reconnected to the closest point of the tree.
- **ILP**
  The whole tree is configured from scratch in optimal way. Instead of the original graph we use the graph without the elements that failed. This is the optimal new tree. However, it can be very different from the original one. This is a drawback, since many connections will have to be interrupted for reconfiguration purposes.
- **ILP Partial**
  This is very similar to the ILP restoration approach with the difference, that the part of the tree that is not affected by the failure is kept, i.e., all unaffected links will have zero cost.

Our earlier results [5] show that while there are few failures at time restoration is fast enough not to affect the understandability and enjoyability of the video content. However, if there are multiple failures at time, and instead of protection restoration has to be used that can last even for seconds the users will not be satisfied with the quality. The probabilities of having such a failure pattern that will interrupt the streaming for more than half a second is very rare. In case of interrupts longer than a few tens of mili-seconds the content should be cached and streamed again as soon as the network, or the cut branches of the tree have recovered.

VI. Simulation Results

The simulations have been carried out on the COST 266 BT European reference network that consists of 28 nodes and 41 links. Each tree consisted of one 'root' and 5-27 'leaves' all randomly chosen with uniform distribution.

First, we have optimally configured the multi-cast trees using 'ASP', 'MPH' and 'ILP' as explained in Section IV and shown as the leftmost triplet of bars in Figures 1(a)-1(f).

Then, we have simulated link failures one-by-one for all links used by the considered tree, and for each such failure scenario we have restored the tree using the four methods 'ASP', 'ASP partial', 'ILP', 'ILP partial' as explained in Section V.

The evaluation criteria were as follows.

First we have evaluated the cost of the obtained tree as shown in Figure 1(a). The failureless tree was always the 'cheapest' particularly that obtained by 'ILP'. After the failure, the 'ILP' has best restored the tree, regardless what was the initial tree set up method. For other restoration methods 'MPH' had roughly the same performance as 'ILP', while 'ASP' was the worst.

Second, the time required to calculate the multi-cast tree as well as to recalculate the restoration of the tree was evaluated as shown in Figure 1(b). Here we see the drawback of the 'ILP' method for both routing and restoring the tree. However, it gives the global optimum in terms of its cost-based objective function. 'ILP' has the most significant time requirement, while 'ASP' and 'ASP partial' are the fastest.

The amount of used capacity shown in Figure 1(c) has similar character to that of the cost (Figure 1(a)).

Figure 1(d) shows how many wavelengths are used by the different methods to set up and restore the trees. For both, ILP is followed by ASP. For restoration the partial methods have better performance than the simple full ASP.

Figure 1(e) shows how many E/O ports are required to perform multi-cast in the electronic (upper) layer. This is slightly related to the number of wavelengths used (Figure 1(d)). If more wavelengths are used, slightly less E/O and O/E conversions are requested, since in some cases 'E' (electronic) multi-casting can be substituted by the 'O' (optical) multi-casting. Any failure will cause significant growth in using O/E and E/O ports.

In Figure 1(f) it is interesting to note that the size of the network relative to the failureless case can be somewhat smaller, particularly for the ASP tree set-up with ILP tree-restoration! The explanation of this behaviour is, that in the failureless case ASP did not find a good tree, so relative to it ILP resets the whole tree from scratch, resulting a much better tree even if a link is unavailable due to its failure!

Finally, Figures 1(g) and 1(h) show how the tree set-up method and the restoration strategy upon a failure impact the users. For this purpose we have defined two metrics, the Relative Impact (Figures 1(g)), and its variant (Figure 1(h)) weighted by the relative change of the number of wavelengths used, i.e., by the ratio of the number of wavelengths in the failureless case to that in the case of failures.

We have defined the relative impact of failures as the average of the following products for all failure scenarios:

- The ratio of leaves cut from the root of the tree by the considered failure to all the leaves of the tree.
- The time of restoring the tree, i.e., calculating and setting up the new tree.
- The length of the link, the failure of which is being considered (the longer the link is the more prone to failures it is, i.e., has lower availability and will fail more often, therefore, it is taken with higher weight into the average).

In Figures 1(g) and 1(h) it is to be noted that regardless of the tree set-up methods, the faster 'ASP' and 'ASP partial' methods should be used for restoration upon the failure, since although they provide slightly cheaper trees, their calculation times are unacceptably long!

VII. Conclusions

In this paper we have analysed what are the resilience requirements of IPTV based video streaming (multi-cast, broadcast) services, and also compared a wide range of resilience mechanisms and evaluated their capabilities and performance for metro and core networks.

The results clearly show, that the speed of calculating the restoration is crucial, while the method how the tree was set up and how good the resulting tree will be is less significant. I.e., the length of the service interrupt is more important than the quality of certain trees before and after the failure.

References

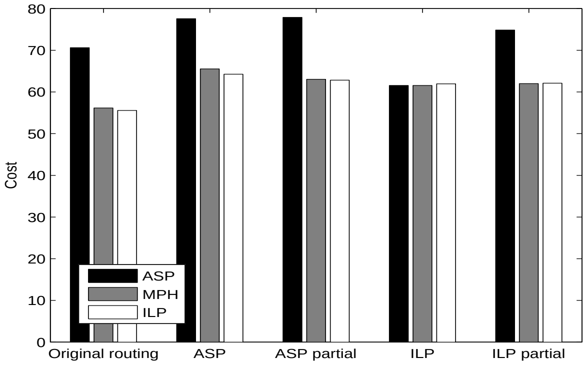
[1] P. Soproni, M. Perényi, T. Cinkler: “Grooming-Enhanced Multicast in Multilayer Networks”, ONDM2007, 11th Conference on Optical Network Design and Modelling, May 29-31, 2007, Athens, Greece

[2] M. Perényi, P. Soproni, T. Cinkler, D. Larrabeiti: “Regular Reconfiguration of Light-Trees in Multilayer Optical Networks”, ONDM 2008, The 12th International Conference on Optical Networking Design and Modelling - ONDM 2008 - Vilanova i la Geltru, Catalonia, Spain, March 12-14, 2008

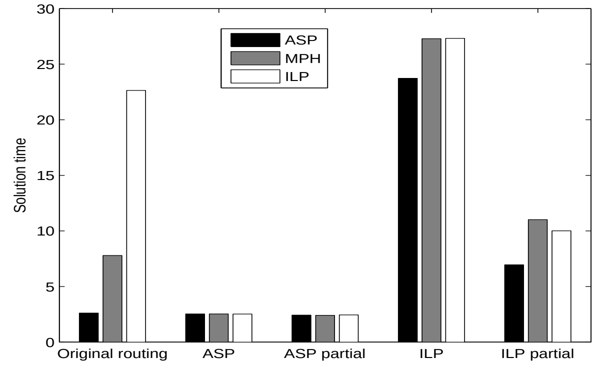
[3] M. Ali, J.S. Deogun: “Cost-effective implementation of multicasting in wavelength-routed networks”, Journal of Lightwave Technology, Vol. 18. Issue 12., 2000

[4] Thomas H. Cormen et al.: “Introduction to Algorithms”, Second Edition, MIT Press and McGraw-Hill, 2001, Section 23.2: ”The algorithms of Kruskal and Prim”, pp.567-574.

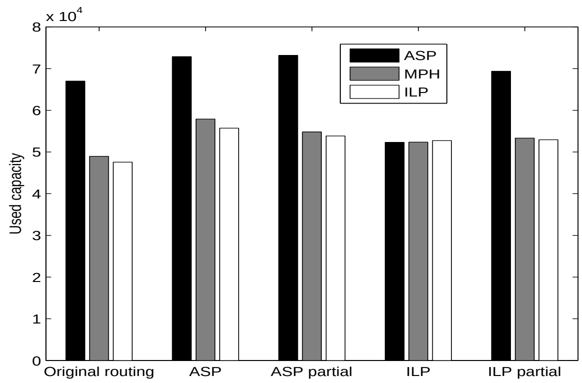
[5] J. Fernandez-Palacios, R. Duque, T. Cinkler, P. Soproni, M. Perényi, J. Tapolcai, P. Fodor, A. Gulyás, J. Hernandez, J. Aracil, R. Gruenzinger: “Network Resilience Requirements and Algorithms for Multicasting and Broadcasting Digital TV”, Networks 2008, The 13th International Telecommunications Network Strategy and Planning Symposium, Budapest, September 28-October 2, 2008



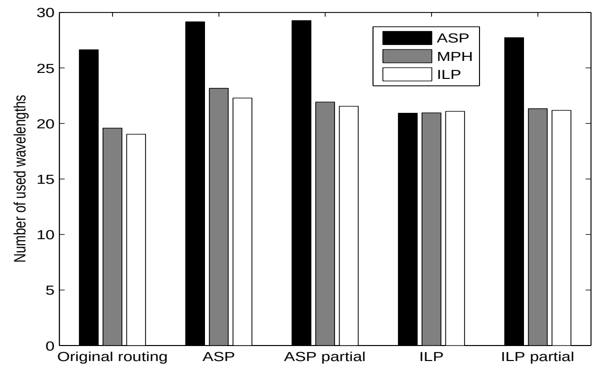
(a) The tree costs.



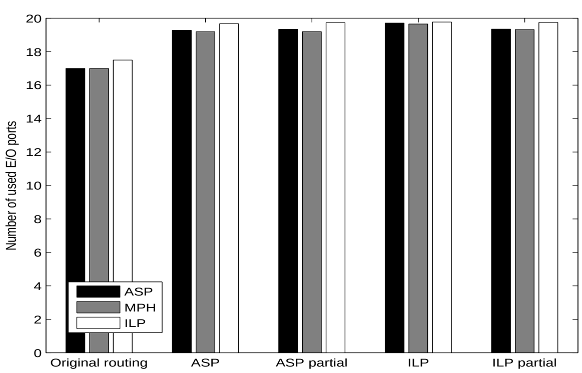
(b) Calculation times.



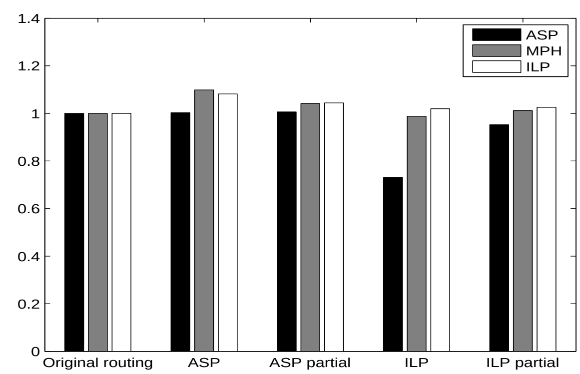
(c) Capacity requirements.



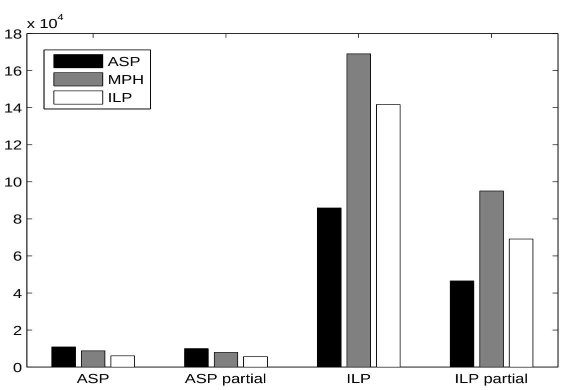
(d) Number of used wavelengths.



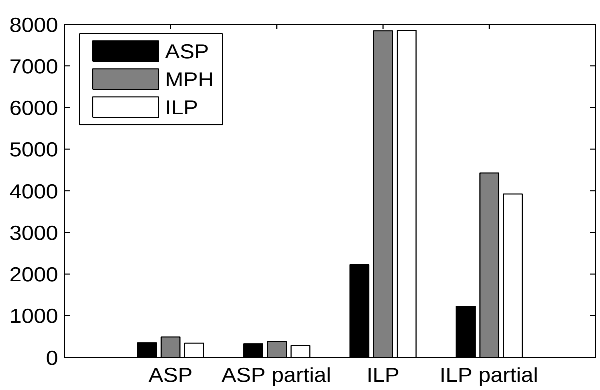
(e) Number of used E/O ports.



(f) Number of links used after restoration, relative to the case without any failure.



(g) Relative Impact of Failures (Service Unavailability)



(h) As in Fig. 1(g), weighted with the relative change of wavelengths used.

Fig. 1. The results of simulating failures and recovering after them using four methods: ASP, ASP partial, ILP, ILP partial. The triple columns show the three methods ASP, MPH and ILP for setting up trees initially. The left-most triplet of columns is the failureless reference case in Figures 1(a)-1(f).