Network resilience requirements and algorithms for multicasting and broadcasting digital TV

JUAN PEDRO FERNANDEZ-PALACIOS GIMENEZ, RAUL DUQUE
Telefónica Investigación y Desarrollo, Telefónica, Spain
{jpfpg, e.rt1}@tid.es

TIBOR CINKLER, PÉTER FODOR, PÉTER SOPRONI, MARCELL PERÉNYI, JÁNOS TAPOLCAI, ANDRÁS GULYÁS, GYULA SALLAI
Department of Telecommunications and Media Informatics,
Budapest University of Technology and Economics, Hungary
{cinkler, fodorp, soproni, perenyim, tapolcai, gulyas, sallai}@tmit.bme.hu

JAVIER ARACIL RICO
Departamento de Ingeniera Informática, Universidad Autonoma de Madrid, Spain
javier.aracil@uam.es

RUPERT GRUENZINGER
Alcatel Lucent, Nürnberg, Germany
rgruenzinger@alcatel-lucent.com

Keywords: digital TV, metro and core networks, network resilience

In this paper we give an overview of the network architecture and of the resilience requirements for both metro and core networks. We propose and evaluate various protection mechanisms for the metro network and restoration mechanisms for the core network, and evaluate them by extensive simulations, showing that the quality of the tree obtained after the failure is much less important than the restoration time, that strongly depends on the algorithm that determines the new tree. We evaluate how the interrupts of protection switching and restoration affect the experienced quality of service for different video formats and resolutions.

1. Introduction and motivation

Broadcast TV is one of the key services offered by most telecom operators nowadays. This service presents strict requirements in terms of survivability since a very high number of users would be affected by a failure in its distribution. Therefore, broadcast TV transport connections should be able to face multiple simultaneous link failures in order to achieve 99,999% availability.

Currently, there exist multiple transport alternatives for broadcast video distribution in metro and core networks. Operators can choose among layer 1 (NG-SDH, OTH), layer 2 (PBB, PBB-TE, T-MPLS, RPR) and layer 3 (IP/MPLS) transport solutions. While most of these technologies already include protection mechanisms for multicast connections, the development of restoration mechanisms for multicast traffic is still an open issue. In this context, this paper provides a performance ana-

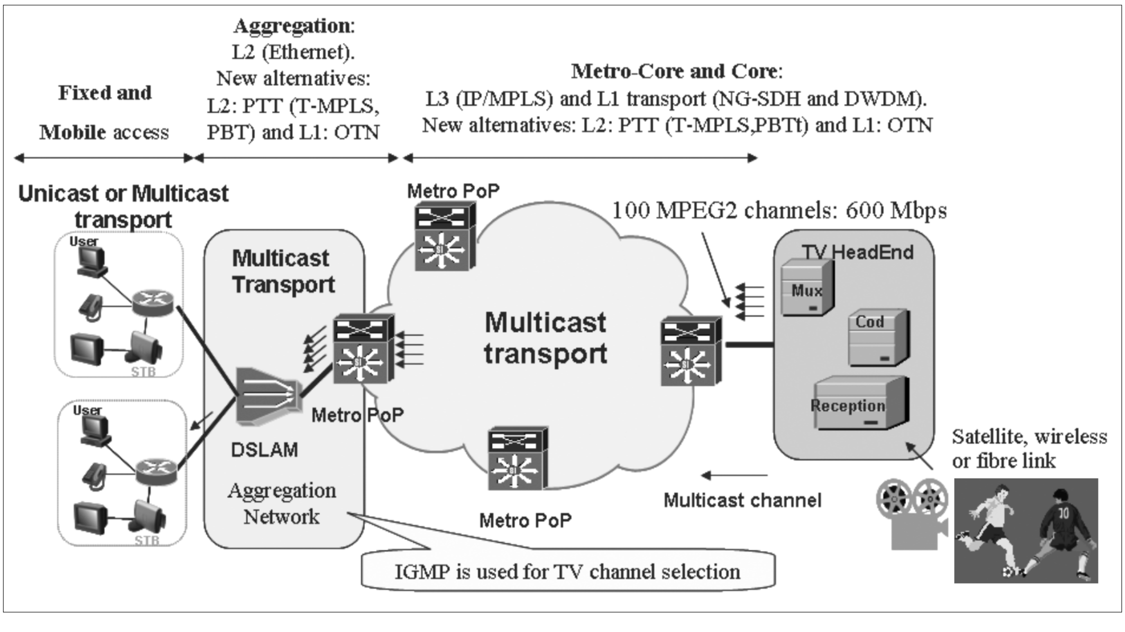


Figure 1. Multicast distribution over metro and core networks [1]

	Bandwidth [Mbps]			
Type of service	**Peak down**	**Peak up**	**Mean down**	**Mean up**
Video Broadcast 0 (Mobility TV)	0.384	0	0.256	0
Video Broadcast 1 (SDTV mpeg2)	6	0	6	0
Video Broadcast 2 (SDTV mpeg4)	3	0	3	0
Video Broadcast 3 (HDTV mpeg2)	20	0	20	0
Video Broadcast 4 (HDTV mpeg4)	10	0	10	0

Table 1. The bandwidth requirements of the most common Broadcast TC services

lysis of resilience solutions for broadcast TV services in metro and core networks based on a combination of both multicast transport and restoration.

2. Distribution architecture and requirements for broadcast TV services

Broadcast TV service is based on the digitalization, compression and transmission and decoding of video signals over IP networks.

The distribution architecture of this type of services is explained as follows *(Figure 1)*:

- A regional TV head-end receives all the national and international TV channels by means of satellite, fibre or wireless connections.
- Then TV channels are broadcasted within the regional network following two steps: a multicast transport over IP in the backbone and a multicast transport over Layer 2 in the metro area.
- Therefore, each DSLAM can receive all TV channels. However, due to the bandwidth limitations of the access segment, the end user only receives the selected channels.

Table 1 shows the bandwidth requirements of most common Broadcast TV services:

Broadcast TV traffic volume does not depend on the number of customers but on the number, definition and coding 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 coding, will need 1 Gbps from the TV head-end to the rest of Service PoPs in the metro core, and from the Service PoP to the access nodes in the metro access.

Considering resilience issues of IP-TV services, we can mention the following topics:

- This service requires resilience mechanisms for high capacity multicast traffic.
  Total capacity depends on the number of TV channels, and the capacity per channel (definition and coding technique).
- A very high number of users would be affected by a total service cut. Therefore, broadcast TV transport connections should be able to face multiple simultaneous link failures.
- Recovery speed should be lower than 50 ms in case of a single failure and lower than 1 s in case of multiple failures [1].
- Neither retransmission nor FEC techniques are used due to strict jitter requirements.
  Therefore very low packet loss rates are required.

Table 2 shows a summary of the main resilience requirements for Broadcast TV services.

3. Transport and resilience alternatives for Broadcast TV distribution

A possible alternative for Broadcast TV distribution is Multi-Protocol Label Switching over IP (IP/MPLS) whose Fast Reroute (FRR) mechanism supports both unicast and multicast traffic. FRR is based on pre-planned protection schemes which are specially adapted to single failures situations.

Virtual Private LAN Service (VPLS) is a layer 2 multipoint VPN that allows multiple sites to be connected in a single bridged domain over a provider managed IP/MPLS network. So, MPLS-based Fast Reroute mechanism can be used to ensure sub-50 ms protection [2].

Provider Backbone Bridges (PBB) support multicast over Carrier Ethernet networks. Packets are forwarded and replicated according to their Backbone MAC. Service is still connectionless, flooding is used when destination MAC addresses are not recognized, and spanning tree protocol (STP) is used to prevent loops.

Provider Backbone Bridges – Traffic Engineering (PBB-TE) is solving the survivability problems of PBB by disabling STP and implementing 50 ms recovery with fast 802.1ag CFM OAM. However 802.1ag only implements protection mechanisms for unicast traffic. Restoration mechanisms are not available yet due to the lack of a distributed control plane. Current PBB-TE standards work does not address P2MP architectures.

Table 2. The main resilience requirements for Broadcast TV services

	Multicast	Multiple failure survivability	Max. Jitter	Recovery Time	
				Single failure	Multiple failures
IP-TV	YES	YES	20 ms	50 ms	1 s

Provider Link State Bridging (PLSB) is an extension of PBB-TE which supports multicast transport. It could include either protection or restoration mechanisms. However, there are no standardization initiatives for PLSB yet.

Resilient Packet Ring (RPR) that is based on a dual counter-rotating ring can be used for video multicasting.

Transport-Multiprotocol Label Switching (T-MPLS) is a connection-oriented packet switched transport layer technology. T-MPLS supports almost all the protection mechanisms of typical transport networks with sub-50 ms protection switching time (relying on hardware-based OAM implementation). Some open issues of this technology are the standardization of OAM and resilience mechanisms for P2MP connections and the T-MPLS control plane definition.

Next Generation-Synchronous Digital Hierarchy (NG-SDH) is one of the most extended technologies in metro-core and core networks, and is mainly aimed at providing a bridging point between the legacy TDM architectures and new IP and Ethernet transport networks. Currently, NG-SDH networks can also include a GMPLS control plane. According to it, both legacy protection and new restoration mechanisms can be implemented.

There are two more technologies to provide Broadcast TV service. *Optical Circuit Switching* (OCS) implements both dedicated and shared protection mechanisms in order to allow sub-50 ms lambda recovery. If optical nodes include add&drop functionalities, then such mechanisms could be easily used for multicast protection. It is important to highlight that the implementation of restoration mechanisms in all optical networks is much more complex. *Optical Burst Switching* (OBS) is another possibility. It can be seen as a long-term alternative for video transport and it could be a good mechanism for fast video downloads. From a resilience-based perspective, OBS is more fault-tolerant than OCS.

Table 3 summarizes some characteristics of the previous technical transport alternatives for delivering IP-TV traffic.

Although T-MPLS and PBB-TE do not support multicast distribution solution by now, there are a lot of technical alternatives providing it, as it can be observed in the table. Then, we will focus in multicast mechanisms for delivering Broadcast TV services. In addition, it is checked that it achieves bandwidth savings in comparison with unicast distribution solution.

In this paper, we propose and evaluate different resilience mechanisms for multicast connections.

4. Resilience mechanisms for metro networks

Once we have chosen multicast transport as the most efficient one in terms of resource consumption, we aim to quantify the possible differences between two resilience strategies, such as 1+1 protection and restoration, measuring the provided service availability.

4.1 Case studies

In the metropolitan area, we have developed the simulations over the Madrid's Metro-Core reference scenario, presented in *Figure 2*.

As we mentioned previously, we will focus on multicast transport solution. So, a point-to-multipoint connection is established from the TV head node to the metro access nodes to transport all the TV channels.

In this part of the study, we have analyzed the following case studies:

- **Multicast distribution combined with protection:** the use of global 1+1 protection for the multicast tree is analyzed in a multiple failure scenario.

	Multicast	Recovery speed		Resilience Standards
		1+1 Protection	Restoration	
IP/MPLS	YES	< 50 ms	IP rerouting process (sec or min)	Unicast: RFC 4090; Multicast: work in progress
PLSB	YES	TBD	TBD	None
RPR	YES	< 50 ms	NO	802.17
VPLS	YES	< 50 ms	IP rerouting process (sec or min)	Just like MPLS
PBB	YES	NO	STP convergence (sec or min)	802.1D (STP)
PBB-TE	Under definition	< 50 ms	NO	802.1ag
T-MPLS	Under definition	< 50 ms	NO	Y.1720/G.8131
OBS	YES	TBD	TBD	None
NG-SDH (GMPLS)	YES	< 50 ms	< 100 ms	Y.1720/G.8131
OCS (GMPLS)	YES	< 50 ms	seconds	None

Table 3. Some characteristics of the previous technical transport alternatives for delivering IP-TV traffic

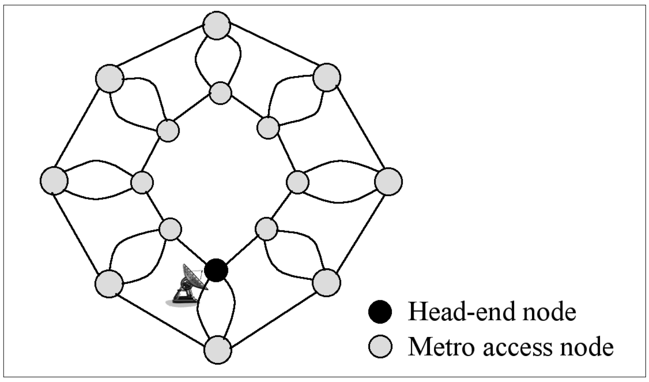


Figure 2. Madrid's Metro-Core reference network

- **Multicast distribution combined with restoration:** in this case, we analyze a global restoration mechanism to recover the branches of the multicast tree in a situation with multiple failures.

4.2 Recovery procedures

In this section we are going to explain the operation of the two resilience mechanisms that we expect to compare in terms of service availability:

- **1+1 Protection**: it consists of pre-calculating a working multicast tree and a backup multicast tree, both according to shortest path algorithm. There is another condition to compute the backup tree: it must be link-disjoint respect to the working one *(Figure 3)*.

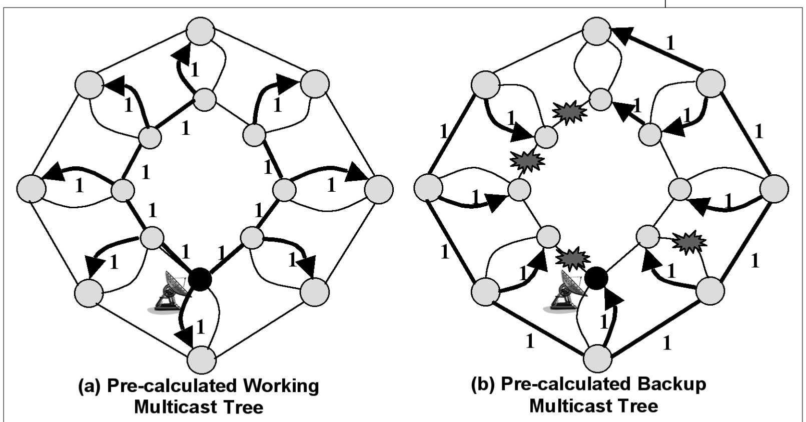


Figure 3. Multicast with 1+1 protection.

At the initial situation, we have established the pre-calculated working multicast tree (Fig. 3a). After a multiple link failure event, if there is an error affecting any link of the primary multicast tree, the resilience mechanism (1+1 protection) switches to the pre-calculated backup multicast tree, if possible (Fig. 3b).

In this situation, if a link failure affects any link of the backup multicast tree, the resilience mechanism tries to re-establish the pre-computed working multicast tree (Figure 3a).

- **Restoration**: as soon as the network starts to work, a multicast tree is computed and established according to shortest path algorithm (Fig. 4a). After a multiple link failure event, if there is an error affecting any link of the established multicast tree, the resilience mechanism (restoration) searches another possible multicast tree according to the available network resources (Fig. 4b), avoiding the broken links. In this manner, it is tried to maintain the IP-TV service in all the access nodes.

4.3 Study parameters

The simulations have been carried out using OMNET++ [3]. OMNET++ is an object-oriented modular discrete event network simulator. The most common application area of OMNET++ is the simulation of telecommunications networks.

The simulation model includes, as well as the network nodes and bidirectional links, a central module that computes the corresponding multicast trees depending on both the considered resilience mechanism and the available network resources (i.e., non-cut links) at every moment. Also, it calculates the number of metro access nodes that can not receive the IP-TV traffic and the period of time being out of service.

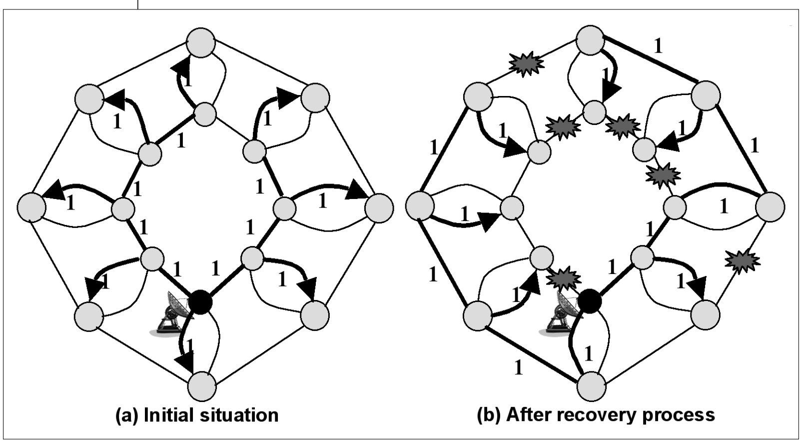
For all the simulations, we have chosen two input parameters: *mean time between link failures* (MTBF) and *mean time to repair the link failures* (MTTR).

The chosen values have been the following:

- MTBF: 30 days, 45 days, 2 months and 6 months.
- MTTR: 6 hours, 12 hours, 1 day, 2 days, 5 days, 1 week and 2 weeks.
- Simulation time: 5 years.

We have carried out simulations combining all the considered MTTR values for every MTBF one separately. The random number generator has been fed with different seeds in order to obtain statistically reliable results for each pair of MTTR and MTBF values. Specifically, we have used

Figure 4. Multicast with restoration. Recovery procedure example.



the L'Ecuyer random number generator [4] with a generation period of 2^{191} that guarantees a great amount of independent streams.

In this study, a subset of the obtained results is presented in the next section.

The output parameter has been the service unavailability percentage obtained with each considered resilience mechanism. To evaluate this parameter, every time when a subset of metro access nodes can not be reached, we measure the number of access nodes that are out of service and the period of time when they are not receiving the IP-TV traffic, and finally, we carry out the next operation:

$$SU(\%) = \frac{\sum_{j \in P_{out}} AN_{out}^{j} \cdot T_{out}^{j}}{AN_{total} \cdot T_{sim}}$$

where SU is the total service unavailability percentage, AN_{total} is the total number of metro access nodes in the network, T_{sim} is the total simulation time, AN_{out} is the number of access nodes being out of service during a certain period of time T_{out}. Each of these pairs of values (AN_{out}, T_{out}) represents an element j of the group P_{out}.

4.4 Results

For these input parameters, MTBF = 6 months, MTTR varies from 6 hours to 14 days, we have measured the service unavailability percentage for both IP-TV distribution mechanisms. This MTBF value is evaluated as a worst case and also it can be considered as a typical value for an air network deployment. We have obtained the following graphics in *Figure 5* and *Figure 6*.

Figure 5. Service unavailability percentage with Multicast with 1+1 Protection

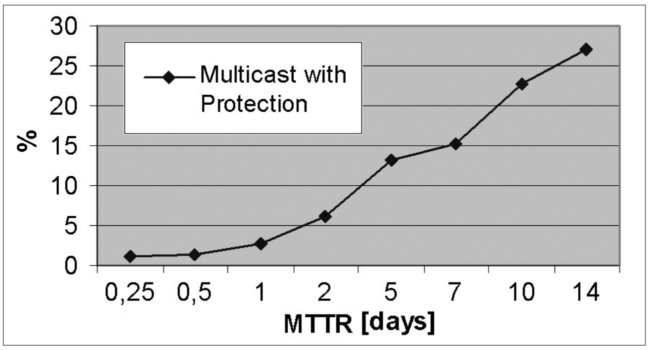
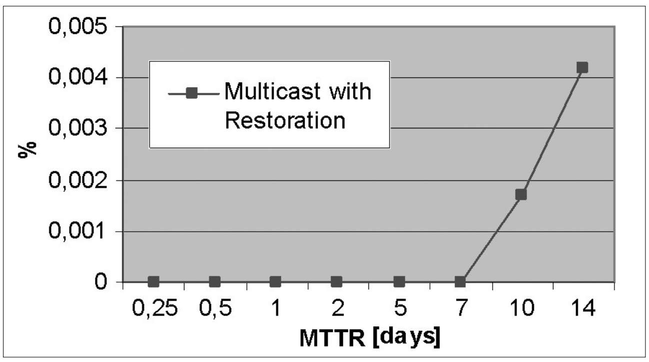


Figure 6. Service unavailability percentage with Multicast with Restoration



Thanks to the adaptability capacity of the restoration mechanism to the available network resources in case of multiple failure, the service unavailability time is zero or negligible for multicast with restoration (Fig. 6), for all the considered MTTR range. In the case of the 1+1 protection resilience mechanism, as we have only two possible multicast trees, the probability that a failure affects a branch of any of the two possible trees is many greater, increasing the access nodes that can not be reached (Fig. 5).

Summing up, according to our simulation results, only restoration-based solutions can achieve 99.999% figures for providing Broadcast TV services, regardless of the MTTR (mean time to repair) parameter.

4.5 Description of assumptions

Our analysis is based on the following assumptions:

1. There is enough bandwidth to carry the TV broadcast channel.

2. However, failures may happen at any link, with Poisson rate λ, which is equal to 1/MTBF and recovery rate μ, which is equal to 1/MTTR.

3. The possible transmission strategies are either multicast with restoration or protection.

4. The maximum number of failures is equal to four. Namely, recovery will happen immediately if four failures occur. This implies that the operator will make every effort not to have four failures whatsoever.

Based on these assumptions, we will evaluate:

1. Number of failures that lead to loss of connectivity in the multicast tree.

2. Distribution of the time to disconnection, namely, time it takes to loose connectivity in the multicast tree.

By multicast tree, we mean the (Hamiltonian) path in the network, such that every node is visited once. The goal is to assess the network performance in terms of fault tolerance of the multicast tree, in terms of MTBF and MTTR.

4.6 Number of failures to loose connectivity

For both the protection and restoration case we evaluate the number of failures that lead to loss of connectivity. Clearly, the chances of disconnection increase with increasing number of failures. Our aim is to evaluate the minimum number of failures are necessary for an eventual disconnection.

4.6.1 Multicast tree with restoration

Our preliminary results show that there are four possible link-disjoint paths originating at the head-end node to any other node *Figure 7*.

Being the node outdegree equal to four, a Hamiltonian path exists (Theorem 6.10 [5]). On the other hand, since the maximum number of link-disjoint paths is four then the minimum number of "cut-sets" that disconnects the subgraph is four (Menger's Theorem, [5]). As a result, the minimum number of failures is equal to four. However, the failure location is essential. Namely, four failures may no lead to loss of connectivity in the tree.

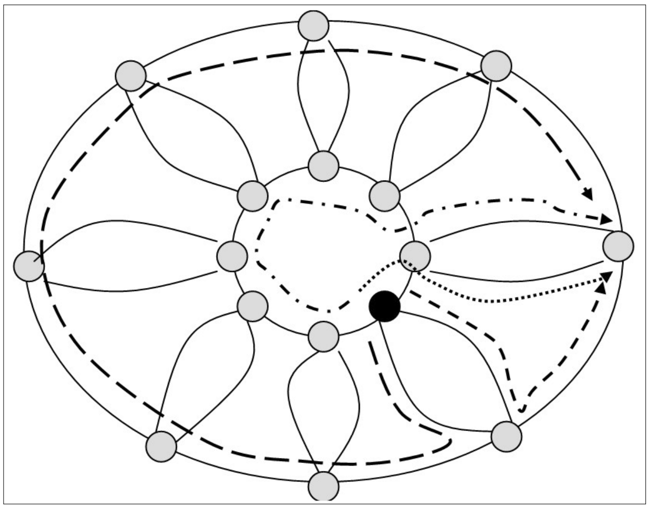


Figure 7. Link disjoint paths in a restoration scheme

4.6.2 Multicast tree with protection

In this case, there are two possible link-disjoint multicast trees, a primary and a backup one *(Figure 8* and *9)*.

Figure 8. Primary multicast tree

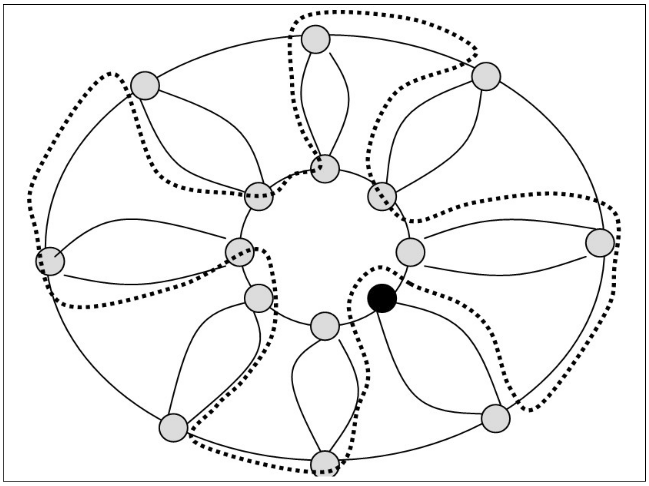
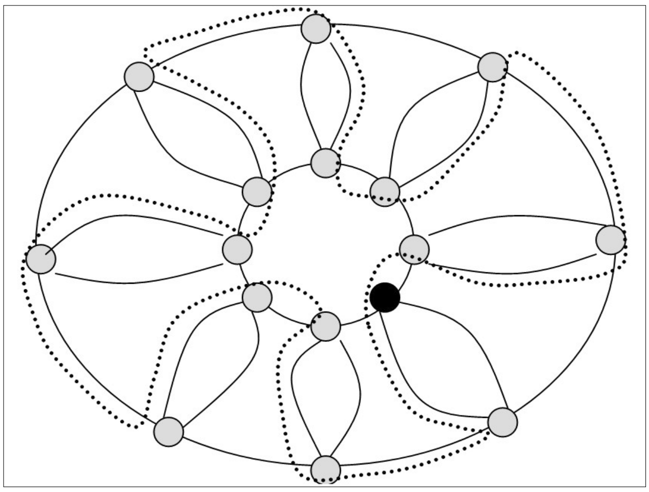


Figure 9. Backup multicast tree



Since the number of trees is equal to two the minimum number of failures is also equal to two, one per tree, in contrast to restoration where typically more than two failures can be "survived".

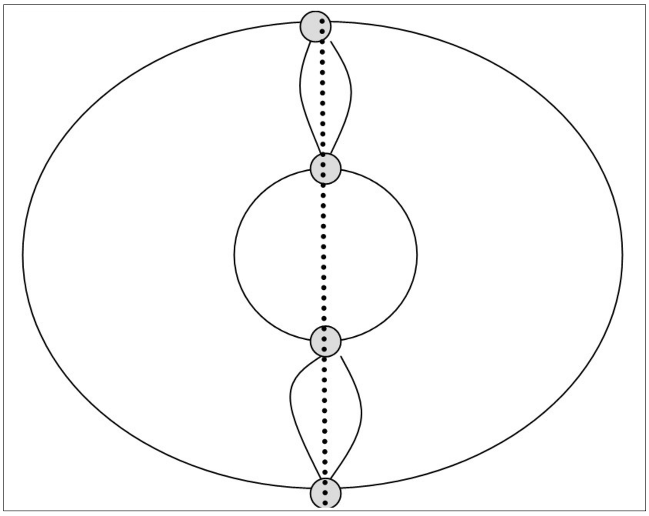
4.7 Distribution of the time to disconnection

In this section, we will characterize the time to disconnection, as a Continuous Time Markov Chain, with the hypothesis set forth in previous sections. First we will derive the set of states. Then, we will evaluate the hitting time to the state "disconnection of the multicast tree".

4.7.1 The case with four nodes

Let us examine the case with four nodes, in order to provide a lower bound for the time to disconnection *(Figure 10)*.

Figure 10. The case with 4 nodes



First note, that a minimum of four failures are necessary to produce a disconnection. Secondly, we note that the graph is symmetric. In what follows, we exploit the symmetry to derive the time to disconnection.

The following graph *(Figure 11)* is obtained by folding the previous graph along the axis of symmetry.

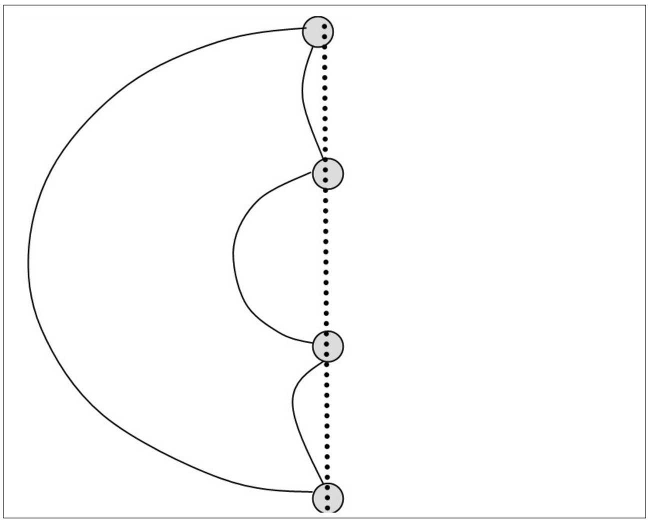


Figure 11. The folded graph with 4 nodes

Interestingly, let us consider the *reflected failures across the symmetry axis from right to left.* For example, a failure in the right outer edge can be reflected to the left outer edge as shown in the following figure *(Figure 12)*.

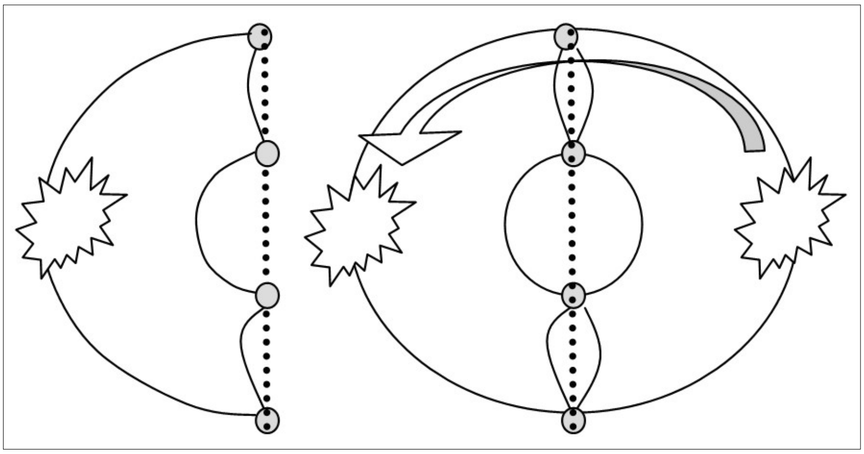


Figure 12. Reflected failures from left to right

Now, if we consider the folded graph only, it turns out that *disconnection happens if and only if there are four reflected failures, in two adjacent edges*.

4.7.2 Time to disconnection with four nodes

In this section we evaluate the time to disconnection, i. e. the distribution of the time elapsed until the multicast tree is disconnected. To do so, we consider the Continuous Time Markov Chain of the number of failures and *ratio of sample paths that lead to disconnection*.

The former can be easily derived using the M/M/4/4 model. The latter requires an explicit calculation of the number of scenarios that give raise to disconnection, with a grand total of four failures.

We distinguish the following cases:

- One reflected failure per link in four links: we only have one possibility.
- Two reflected failures per link, non-adjacent: we have two possibilities.
- Two reflected failures per link, adjacent: we have 4 possibilities.

Thus, the ratio of the number of scenarios leading to disconnection is 4/(4+2+1)=4/7=0.57.

On the other hand, the distribution function of the time to disconnection can be obtained from the M/M/4/4 as follows. First, we note that the infinitesimal generator is given by

$$P=\begin{pmatrix} -\lambda & \lambda & 0 & 0 & 0 \\ \mu & -\lambda-\mu & \lambda & 0 & 0 \\ 0 & 2\mu & -\lambda-2\mu & \lambda & 0 \\ 0 & 0 & 3\mu & -\lambda-3\mu & \lambda \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Let X be the time until four failures, let $M=e^{-Pt}$. Then $P(X\leq t)=M(1.5)$. Finally, note that the latter is the distribution function of the time elapsed until four failures occur, regardless of whether they bring disconnection or not. Consequently, this is a conservative analysis. *Figure 13* shows the results for MTTR=15 days and MTBF=60 days (worst case).

Our preliminary result shows that four failures will happen in a time interval of 160 days with probability 0.9975. Namely, we may expect four failures in a time interval of approximately five months. However, only 4/7 of them will lead to disconnection.

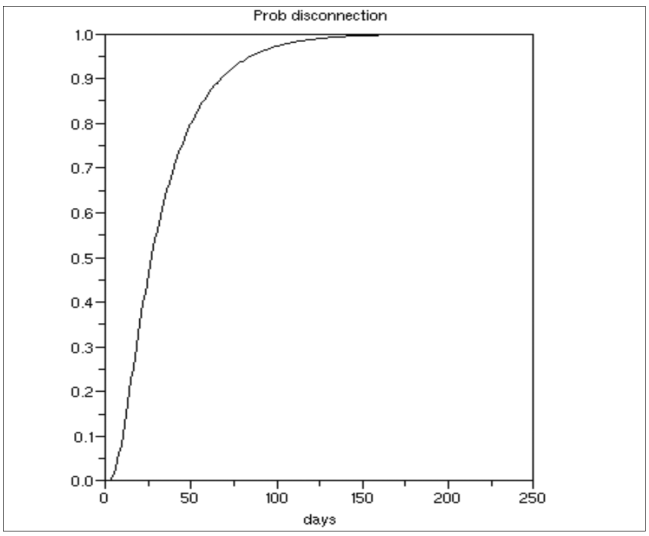
5. Multicast/Broadcast solutions for core networks

In core networks the video content is distributed (multi-casted) 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 multicasts having smaller bandwidth requirement can share a single wavelength path, while others that exhaust the capacity of a wavelength channel may even require multiple bundled wavelength channels.

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 multicasting capability at both layers. At the upper, IP/MPLS/Ethernet layer multicast 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 multicast 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.

Figure 13. The folded graph with 4 nodes



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 utilization.

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 [6] and [7].

5.1 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 multicast 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 [6] and [7]: 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-leaves demand of the same tree is set to zero, that means it can be used for free for all future root-to-leaves demands of the same tree. Of course the chosen sequence significantly influences the result.

• **MPH**: Minimal Path Heuristic

We have adapted [8] 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 [9] 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 [6].

5.2 Methods for restoring multicast sessions

If a link or a node fails in the network it will affect all the multicast connections that use that element. However, if this element is just a leaf (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 trees upon failures. The four methods (ASP, ASP partial, ILP and ILP partial) for restoration that we propose here 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 leaf or multiple (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.

5.3 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 5.1 and shown as the leftmost triplet of bars in *Figure 14(a) – 14(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 5.2.

The evaluation criteria were as follows.

First we have evaluated the cost of the obtained tree as shown in *Figure 14(a)*. The failure-less 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 14(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.

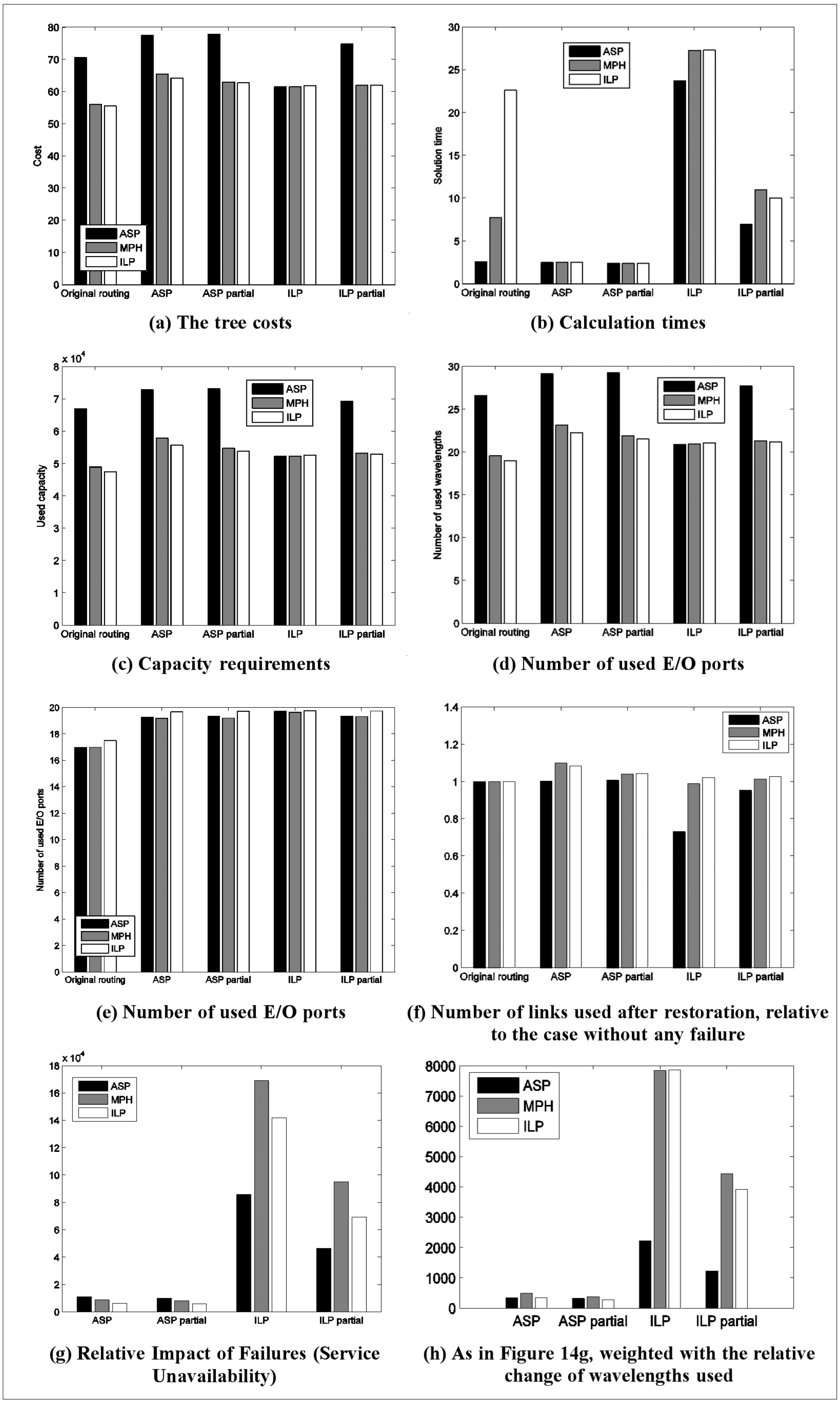
The amount of used capacity shown in *Figure 14(c)* has similar character to that of the cost (Figure 14(a)).

Figure 14(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 14(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 (Fig.14(d)).

If more wavelengths are used, slightly less E/O and O/E conversions are requested, since in some cases ‘E’ (electronic) multicasting can be substituted by the ‘O’ (optical) multicasting. Any failure will cause significant growth in using O/E and E/O ports.

In *Figure 14(f)* it is interesting to note that the size of the network relative to the failure-less case can be somewhat smaller, particularly for the ASP tree set-up with ILP tree-restoration! The explanation of this beha-



(a) The tree costs

(b) Calculation times

(c) Capacity requirements

(d) Number of used E/O ports

(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 Figure 14g, weighted with the relative change of wavelengths used

Figure 14.
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, ILP for setting up trees initially. The leftmost triplet of columns is the failureless reference case in Figure 14(a)–14(f).

Format	Resolution	Frames/sec	Declared bit-rate	Real average bit-rate
SDTV slow (Spiderman)	720x576	25	4.16 Mbps	~ 4,5 Mbps
SDTV fast (Polar express)	720x576	25	4.16 Mbps	~ 5 Mbps
HDTV (Magic of flight)	1920x1080	30	38.8104 Mbps	~ 9,6 Mbps

Table 4. The parameters of the 3 videos evaluated

viour is, that in the failure-less 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, *Figure 14(g)* and *Figure 14(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 (Figure 14(g)), and its variant (Figure 14(h)) weighted by the relative change of the number of wavelengths used, i.e., by the ratio of the number of wavelengths in the failure-less 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 which failure is being considered (the longer the link is the more prone to failures is, i.e., has lower availability and will fail more often, therefore, it is taken with higher weight into the average).

In Figure 14(g) and Figure 14(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 not acceptable!

6. Quality of experience for video streaming in case of short interrupts caused by network failures

In this section we present our experimental study to evaluate how the different protection and restoration times upon failures affect the experienced quality of different video formats and contents.

In our experiments we have used the *ACR* (Absolute Category Rating) method [10]. The IETF has a similar framework, the MDI, Media Delivery Index [11], where the packet delay and packet loss are mostly considered.

However, since the failures happen very rarely, e.g., a few times per year that hardly affects the quality of the service in its classical sense we have condensed the failure events in the following way. We assume that there are exactly three failures of equal duration at random time instants with at least 10 seconds of difference between two failures within each 40-second video clip.

In all cases MPEG2 encoding has been used, with maximum packet size of 1310 bytes. The video frames have been carried by UDP over IP, around 400 packets per second, 16 packets per video frame on average. Three videos have been evaluated as shown in *Table 4*. The bit-rates of videos were analyzed by the Elecord Stream Application software. In all cases there were 15 evaluators. First, the video clip with no failures has been shown, then failures of duration of 30, 50, 200,

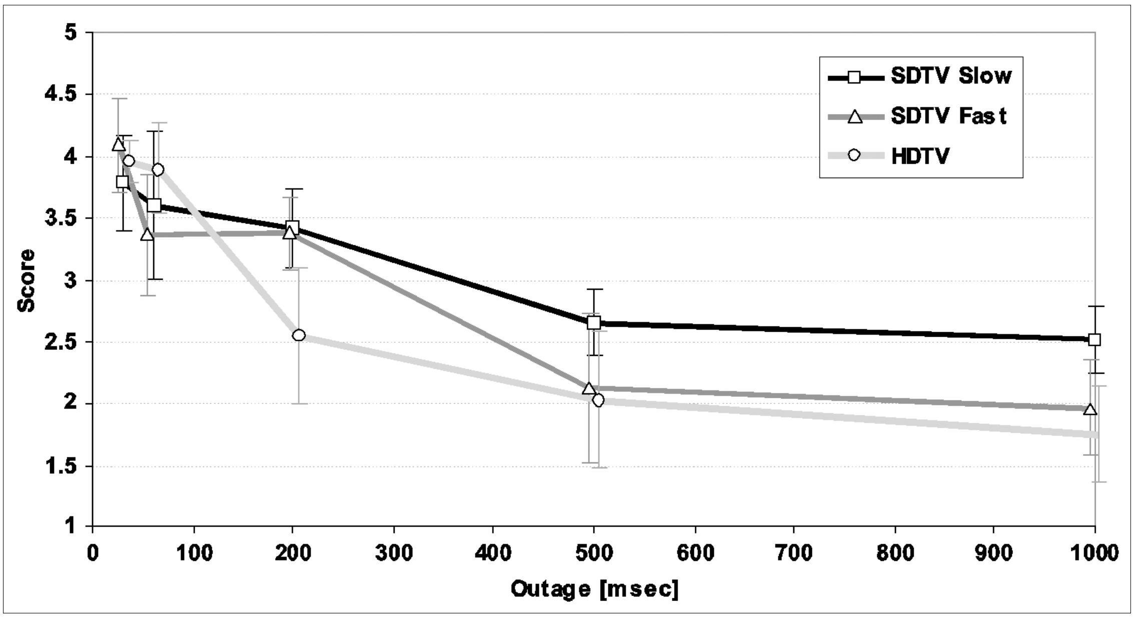


Figure 15. Mean opinion score with variance for the three videos for different durations of protection or restoration. Higher score means better quality.

500 and 1000 ms in random order and again these five failure durations in another random order with different failure instants.

Figure 15 shows the average and the variance of scores provided by the 15 evaluators. The effect of failures typically was that the motion stopped, the sound disappeared, and the picture was in part covered by squares off different sizes of colors mostly similar to the picture, however sometimes very different colors appeared as well.

It can be seen that slow motion SDTV is less affected by failures than the HDTV, while fast motion SDTV is between them for longer outages. For shorter outages the experienced quality is in general better. Two interesting properties can be noticed. First, that the quality of HDTV that was most seriously affected by longer outages is least impacted by shorter outages. Some short interrupts could be noticed, however, the squares were much smaller than for SDTV that explains the better experienced quality. Second, for shorter interrupts it does matter where the outages occur, i.e., what kinds of frames are lost.

In MPEG encoding each frame (GOP: Group of Pictures) consists of a series of frames, where the first one is the so-called I-frame which corresponds to a whole fixed image. It is followed by other frames that do not carry the whole fixed image, only differences relative to the picture carried in the I-frame. Therefore, if an I-frame is lost it is more critical than loosing only differences to this frame. Also, if there is sound or particularly speech when the outage happens it is more critical from the perspective of the user. This explains that for very short interrupts the subjective scoring of experienced quality can vary depending on the exact timing of failures.

Regarding the effects of failures onto the experienced quality of video streaming we can conclude, that interrupts of length from 30 to 1000 ms can be all noticed, they cause a minor disruption; however, considering that they happen a few times per year only, they are not critical at all. Although for SDTV interrupts over 100 ms, while for HDTV interrupts of over 50 ms can be annoying, if the service is restored within 50 ms the user will not loose any content that could hinder him understanding a sport event, a movie or news.

7. Conclusions

At present, IP-TV distribution in the metro and core networks is based on packet transport technologies such as IP/MPLS at level 3, including sub-50 ms 1+1 protection and restoration mechanisms (i.e., fast rerouting) and NG-SDH and DWDM technologies for transport at level 1.

At this point, it is important to mention the new existing alternatives for delivering IP-TV services, like PBT (T-MPLS or PBB-TE) at level 2, and OTN for level 1. However, these technologies only implement protection mechanisms for unicast traffic. In addition, restoration mechanisms are not available yet due to the lack of a distributed control plane in both technologies. So, in the future, these technologies can be planned to be used for the distribution of IP-TV service, as long as the standardization of OAM and resilience mechanisms for P2MP connections and the control plane definition for both technologies will be achieved.

In this paper we have analyzed the resilience requirements of IPTV based video streaming (multicast, broadcast) services, and also compared a wide range of resilience mechanisms and evaluated their capabilities and performance for both metro and core networks.

Firstly, in Section 4.4, it has been demonstrated by means of simulation that the utilization of restoration mechanism is the most appropriate one in order to support an acceptable quality of service when it is highly dependent on the service availability time, in a multiple failure scenario. This is the case of IP-TV broadcast service since a very high number of users would be affected by a total service cut. So, we recommend the application of restoration as the resilience mechanism in combination with multicast transport, since it provides total service availability in all the metro access nodes, regardless of the mean time to repair.

Finally, in Section 5.3 and 6, our results show that while there are few failures at a time the protection is fast enough not to affect the understandability and enjoyability of the video content. However, if there are multiple failures at a time, and instead of protection restoration has to be used that can last 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 milliseconds the content should be cached and streamed again as soon as the network, or the cut branches of the tree have recovered.

Acknowledgement

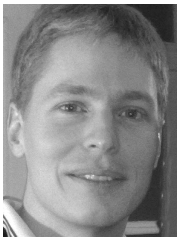
The authors are grateful to all colleagues in the FP6 IST IP NOBEL project, particularly within WP2.

Authors

TIBOR CINKLER has received M.Sc.(1994) and Ph.D. (1999) degrees from the Budapest University of Technology and Economics (BME), Hungary, where he is currently associate professor at the Department of Telecommunications and Media Informatics. His research interests focus on optimisation of routing, traffic engineering, design, configuration, dimensioning and resilience of IP, Ethernet, MPLS, ngSDH, OTN and particularly of heterogeneous GMPLS-controlled WDM-based multilayer networks. He is author of over 130 referreed scientific publications and of three patents. He has been involved in numerous related European and Hungarian projects including ACTS METON and DEMON; COST 266, 291, 293; IP NOBEL I and II and MUSE;NoE e-Photon/ONe andNoE e-Photon/ONe+; CELTICPROMISE; NKFP, GVOP, ETIK; and he is member of ONDM, DRCN, Broad-Nets, AccessNets, IEEE ICC and Globecom, EUNICE, CHINACOM, Networks, WynSys, ICTON, etc. Scientific and Programm Committees. He has been guest editor of a Feature Topic of the IEEE ComMag and reviewer for many journals.

PÉTER FODOR received his M.Sc. degree in electrical engineering from the Budapest University of Technology and Economics (BME), Budapest, Hungary, in 2001. He was consultant manager for a couple of years at the Department of Telecommunications and Media Informatics and a member of the High Speed Networks Laboratory. His research interests focus on network controlling, modeling and generalization. Currently he is working for Ericsson Hungary as Tool Expert.

PÉTER SOPRONI received his M.Sc. degree in Computer Science from the Budapest University of Technology and Economics (BUTE), Hungary, in 2008. He is currently a Ph.D. student at the Department of Telecommunication and Media Informatics in the same institute. He has participated in several research projects supported by the EU and the Hungarian government. His research interests include simulation, algorithmic optimization and planning of optical networks, as well as soft-computing especially bacterial algorithms. He has experience in .NET and Linux based software development.



MARCELL PERÉNYI received his M.Sc. degree in Computer Science from the Budapest University of Technology and Economics (BUTE), Hungary, in 2005. He awarded the Ph.D. title from the Doctoral School of Informatics of BUTE in 2009. He participated in several research projects, supported by the EU and the Hungarian government, at the Department of Telecommunication and Media Informatics. He is a member of IEEE and HTE. His research interests include simulation, algorithmic optimization and planning of optical networks, as well as identification and analysis of traffic of IP networks, especially P2P, VoIP and other multimedia applications. He has experience in planning and maintaining of database systems, web services and Microsoft infrastructures. Currently he is working for Ericsson Hungary as System Engineer.

JÁNOS TAPOLCAI received his M.Sc. (2000 in Technical Informatics), and Ph.D. (2005 in Computer Science) degrees in Technical Informatics from Budapest University of Technology and Economics (BME), Budapest, Hungary. Currently he is an associate professor at the High-Speed Networks Laboratory at the Department of Telecommunications and Media Informatics at BME. His research interests include applied mathematics, combinatorial optimization, linear programming, communication networks, routing and addressing, availability analysis and distributed computing. He has been involved in several related European and Canadian projects. He is an author of over 40 scientific publications, and is the recipient of the Best Paper Award in ICC'06.

ANDRÁS GULYÁS received M.Sc. and Ph.D. degree in Informatics at Budapest University of Technology and Economics, Budapest, Hungary in 2002 and 2008 respectively. Currently he is a research fellow at the Department of Telecommunications and Media Informatics. His research interests are complex and self-organizing networks, network calculus and traffic management.

GYULA SALLAI: see his cv after the Guest Editorial.

References

[1] NOBEL 2, "Deliverable 2.3: Report on Resilience Mechanism for NOBEL Solutions in Medium and Long Term Scenarios", IST NOBEL Phase 2, 2008. http://www.ist-nobel.org/Nobel2/imatges/D2[1].3_v1.1am.pdf

[2] NOBEL, "Deliverable D27: Final report on Traffic Engineering and resilience strategies for NOBEL solutions", IST NOBEL, 2005.

[3] OMNET++ object-oriented discrete event simulation system. URL reference: http://www.omnetpp.org, 1996.

[4] P. L'Ecuyer, R. Simard, E.J. Chen, W.D. Kelton, An Objected-Oriented Random-Number Package with Many Long Streams and Substreams. Operations Research, 50(6):1073–1075, 2002. Source code can be downloaded from: http://www.iro.umontreal.ca/~lecuyer/papers.html

[5] Alan Gibbons, "Algorithmic graph theory", Cambridge University Press, 1985.

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

[7] M. Perényi, P. Soproni, T. Cinkler, D. Larrabeiti, "Regular Reconfiguration of Light-Trees in Multilayer Optical Networks", ONDM 2008, 12th Int. Conference on Optical Networking Design and Modeling, Vilanova i la Geltrú, Spain, March 12-14, 2008.

[8] M. Ali, J.S. Deogun, "Cost-effective implementation of multicasting in wavelength-routed networks", Journal of Lightwave Tech., Vol. 18, Issue 12., 2000.

[9] Thomas H. Cormen et al., "Introduction to Algorithms", 2nd Edition, MIT Press and McGraw-Hill, 2001. Section 23.2: "The algorithms of Kruskal and Prim", pp.567–574.

[10] [P910] "Subjective video quality assessment methods for multimedia applications", ITU-T Recommendation P.910, September 1999.

[11] J. Welch, J. Clark, " A Proposed Media Delivery Index (MDI)", RFC4445, April 2006.