Department of Telecommunications and Media Informatics
Budapest University of Technology and Economics

Traffic Driven Optimization of Resilient QoS-aware Metro Ethernet Networks

András Kern

Theses

Supervisors:

Dr Tibor Cinkler, Dr. József Bíró and Dr. Gyula Sallai
Highspeed Networks Laboratory
Department of Telecommunications and Media Informatics
Budapest University of Technology and Economics

Budapest
2007

1 Introduction

The continuously increasing internet access speeds of the home customers raise the demand for new value added services like Voice Over IP (VoIP) and Interactive Video on Demand (iVoD). These new services, however, not only consume higher bandwidth but they also require more strict Quality of Service (QoS) than the traditional internet services. The increased QoS demands call for QoS-aware provider network architectures. At the same time Ethernet becomes the sole Local Area Network (LAN) technology both for home and enterprise solutions.

The increasing line speeds and decreasing equipment prices make the Ethernet a cost efficient technology not only for LANs but also for Metropolitan Networks. However, to make Ethernet a viable solution, several issues have to be dealt with, e.g., scalability, high availability and QoS.

The standardization bodies and the main vendor providers are continuously extending the limits of the Ethernet. For customer and traffic separation the concept of Virtual Local Area Networks (VLANs) (IEEE 802.1Q [IEEE802.1Q2003]) were introduced that allows defining up to 4096 logical networks. Besides, by introducing 3 priority bits 8 traffic classes can be differentiated. Further solutions for scalability are the hierarchically organized VLANs and addresses [IEE05b, IEE05a].

The simple frame forwarding mechanism of the Ethernet, that is based on “reverse learning” and broadcasting, requires a loop-free topology. By building loop-free physical topologies the raised problem is eliminated, but these topologies has very limited resilience and QoS capabilities. Therefore the Spanning Tree Protocol (STP) [LSJ98] and Rapid Spanning Tree Protocol (RSTP) [IEEE802.1w, GP] proposed by IEEE present methods to define such a loop-free logical topology over an arbitrary physical one. Furthermore, both protocols are able to react to the changes of the topology, so they recover the connection between the nodes in case of a failure. However, only one tree is defined in the network, thus, the usage of redundant links is inhibited. The Multiple Spanning Tree Protocol (MSTP) [IEEE802.1s] deploys more, independent RSTP instances, so the trees can be configured independently to each other. By properly configuring the trees the redundant links become available for the traffic and Traffic Engineering (TE) can be realized.

During my work, I assumed the QoS Architecture (depicted on Figure 1) that contains three main network segments. Among them I focus on the metro network that are assumed to be switched Ethernet one. Based on their role the switches can be put into three distinct groups. The Access Nodes (ANs) aggregate the traffic of couple of dozens or hundreds of customers, and the Edge Nodes (ENs) provide the connections to the internet and the application providers. These types of nodes perform traffic control tasks, while the simple internal nodes only forward the frames. Thus, the network intelligence is focused to these nodes and to a further central entity called Network Manager Entity (NME). The NME overviews the overall network and performs the network management.

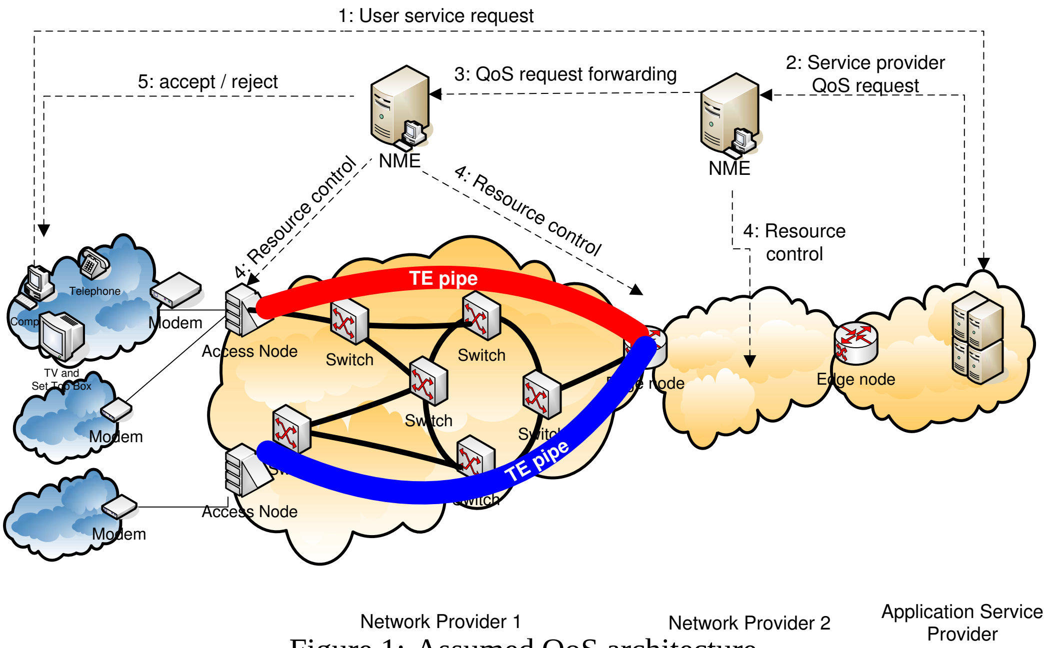


Figure 1: Assumed QoS architecture

The resource reservation is done by pre-provisioned resources and so called logical TE pipes run between the ANs and ENs. Although the standard allows up to 8 classes in most cases only four classes are differentiated (see ITU-T G.1010 [G.1010] and 3GPP TS 22.105 [3GP06] recommodations). Since the resource reservation on these pipes is logical, it is vital to control the amount of traffic assigned to them to avoid exceeding the size of the pipe. This control can be realized either in centralized [BvdSBP03] or in distributed way [**C3**].

Most of the Ethernet switches support the absolute priority scheduling. Then, the traffic classes are ordered based on their priority, and a frame belonging to a lower priority class will not be served until there is higher priority traffic in the queue. To provide the aimed QoS level and to avoid large delays or starvation, the traffic of the different classes has to be limited. In the assumed QoS model for each traffic class static and global ratios are given by the operator. Therefore, the traffic belonging to a certain traffic class must not exceed the predefined limit that is a proportion of the speed of the considered link.

The recovery capability of the RTSP is the prime tool to provide resilience; however, it has two main drawbacks. First, the recovery time can be even some seconds that may not be acceptable for some real time applications. Second, there is no any guarantee that the required amount of capacity is available along the new, recovered tree. The MSTP does not give satisfing solution for addressed problems. To provide the acceptable resilience level, VLAN switching based methods are assumed [G.8031][FTAW05]: for each pipe two disjoint VLANs are defined. In normal operation one of them transmits the traffic and in case of a failure that affects the transmitting VLAN the endnodes switch to the secondary VLAN.

To efficiently implement such an architecture accurate design of paths for TE pipes and allocation of resources are needed. My dissertation is devoted to this issue.

2 Objectives of Research

The solutions provided so far cover only parts of the problem discussed in the introduction. They mainly focus on the main drawback of the application of spanning trees: on the high recovery times after failures [SGNcC04, FTAW05]. Several proposals are presented to implement the support of QoS and TE [LYD$^+$03b, LYD$^+$03a, LLN02]; however, these solutions do not deal with resilience at all. In my dissertation, assuming the architecture introduced in Section 1, I will propose methods and efficient algorithms that enable efficient traffic engineering while the required level of QoS and resiliency are guaranteed. I defined the following research goals:

1. Identification of the goals and issues of traffic engineering and the formalization of these aims as an Integer Linear Program that provides optimal solution.
2. Investigation of the applicability of traffic engineering in typical metro Ethernet topologies with the help of this ILP model. Further aim is the discussion of several protection schemes, such as dedicated and shared protection, and the investigation of their applicability in the model.
3. Proposal of heuristic algorithms in order to efficiently implement TE, in other words, the determination of near-to-optimal solutions within acceptable time constraints. The improvement of the proposed heuristic algorithms in order to exploit statistical multiplexing gain.

3 Research Methodology

The infocommunication networks are modeled with graphs, where the switches are represented by graph nodes, while the connections between them are described with two edges with opposite direction. A positive value is assigned to each edge that defines the available capacity on the link. Therefore, the defined tasks can be derived to graph problems and the graph methods and approaches can be applied.

The optimization problems are formulated as Integer Linear Programs and they are solved by ILOG CPLEX solver tool [CPL]. The obtained results are globally optimal. The ILP, however, is proved to be NP-complete [GJ79], thus, it is hardly scalable: the running time in the case of real life topologies may be high and this method can be applied only for static TE, where the reconfiguration of the network is performed occasionally.Therefore, I sought for heuristic algorithms, such algorithms that wander (but not exhaust) the state space of a problem assuming guiding principles or heuristics, that although, provide only near optimal solutions, they have polynomial running times making possible the dynamic Traffic Engineering.

The evaluation of the proposed solutions and algorithms are performed through simulations. Since my research focused on Metro Ethernet networks the same simulation environment is applied in all theses. During the simulations typical metro network topologies are considered. The aggregation parts concentrate the traffic of the Access Nodes (ANs) to core part. The core part, which is formed by rings or a mesh, transmits the aggregated traffic to the proper Edge Nodes (ENs). Based on [CS06] two topology classes are defined and they differ in the structure of the aggregation parts. The two classes are *Tree-Ring* and *Dual-Homing*. In both classes more topologies are defined with the sizes from 12 up to 48 nodes. The average nodal degree is about 3.0.

The traffic flows only between the border nodes (AN-EN), then, for each service a logical pipe is defined between every AN-EN pairs. The sizes of these pipes depend on the *class of traffic* and the *global traffic level* parameters. Among the four classes three are prioritized (47% of overall traffic) the other 53% belongs to the Best Effort class.

As reference the STP and $MSTP$ protocols are considered that are configured using the default port costs defined in the standard. Due to the configuration schemes they are referred to as "topology driven" STP and $MSTP$. Other cost sets would also be considered, however, the application of such "rules of thumb" would introduce optimization. During the simulations the following parameters are investigated:

- *Achievable (available) throughput* describes the performance of the proposed methods. This property is calculated via scaling the traffic by maximizing the *global traffic level* parameter.

- *The allocated network capacity (or load)* is used to present the capacity efficiency of the methods that is important especially when protection is assumed.

- *Running time* allows us not only to investigate the scalability of the algorithms, but to present their applicability. For instance, the presented algorithms are able to support dynamic traffic engineering or not.

4 New Results

I have summarized my results in three theses each consists of 2–3 subtheses.

Thesis 1 *I have given an optimal method for traffic driven tree configuration of QoS-aware and reliable Metro Ethernet networks.*

Introduction The traffic driven spanning tree optimization task consists of three tasks. First, the logical channels are to be mapped to VLANs. Second, these VLANs are to be assigned to tree instances, and finally, these tree instances have to be spanned in the network in order that the logical channels fulfill both the QoS and reliability requirements. The architecture discussed in the introduction allows several simplifications. Since the traffic flows only between the border nodes (Access Nodes (ANs) and Edge Nodes (ENs)) we can make an assumptions without loss of generality: the let the roots of the trees be placed at the ENs. Furthermore, due to administration and accounting issues the traffic cannot be flown between the access nodes, but through one of the Edge Nodes.

Subthesis 1.1 *I have formulated the problem of optimal tree instance spanning and demand to tree instance mapping problem in QoS-aware Metro-Ethernet Networks as an Integer Linear Program (ILP) ($MSTP_{ILP}$). This model allows the use of the inhibited redundant links, thus, in the case of typical metro network topologies [CS06] the achieved throughput can be even doubled compared to the "topology-driven" MSTP ([**C4**, **C8**]).*

The inputs of the proposed optimization task are the network topology, the capacities of the links, the traffic class descriptors and the traffic matrices, one for each traffic class. To describe the task, I have declared two binary variable sets. $x_{link}^{demand,\ tree}$s describe the route of the traffic demands and the mapping to the tree instances. While, variables y_{link}^{tree} formulate how the trees are spanned. Typical optimization goals are to minimize the number of links used by trees and the amount of allocated capacity. The proposed optimization aim combines these two goals. The weight is characterized by parameter α. The objective is as follows:

$$\min \sum_{\forall links} \left[\alpha \sum_{\forall trees} y_{link}^{tree} + (1-\alpha) \sum_{\forall demands, \forall trees} x_{link}^{demand,\ tree} \cdot \{\text{bandwidth demand}\} \right]. \qquad (1)$$

The first constraint states that each traffic flow must obey the *flow conservation law*, i.e., when a demand goes from its source to its destination and when it enters a node, that is not its destination, it has to leave it as well. And vice versa, if a demand leaves a node it also has to enter it. For the demands belonging to different traffic classes separate, *own* *capacity constraints* are given. They ensure that the traffic of a certain class will not exceed the predefined capacity ratio. The “best effort” traffic, however, can use all the remaining capacities, thus, I consider the traffic of all four classes.

Finding a proper mapping of demands and trees is among the optimization tasks. One option is to define a set of additional constraints. The other option is the extension of the used binary variables [**C4**]. The last group of constraints is responsible for ensuring that the tree instances would be surely trees while their roots are placed at ENs and the ANs would be the leaves. This latter type of constraints does not limit the generality because of the problem definition.

Using simulations the I have shown that in the case of the supposed topologies the achievable throughput can be considerably increased compared to the “topology driven” reference solutions. The calculated throughput of the proposed method is detailed in Figure 2, where 100% is the reference $RSTP$. Compared to the “standard based” MSTP the throughput gain is between 50%–100% depending on the structure of the topology. Outstanding increase is observed in the case of the considered dual-homing structures, since the redundant links can be more efficiently used.

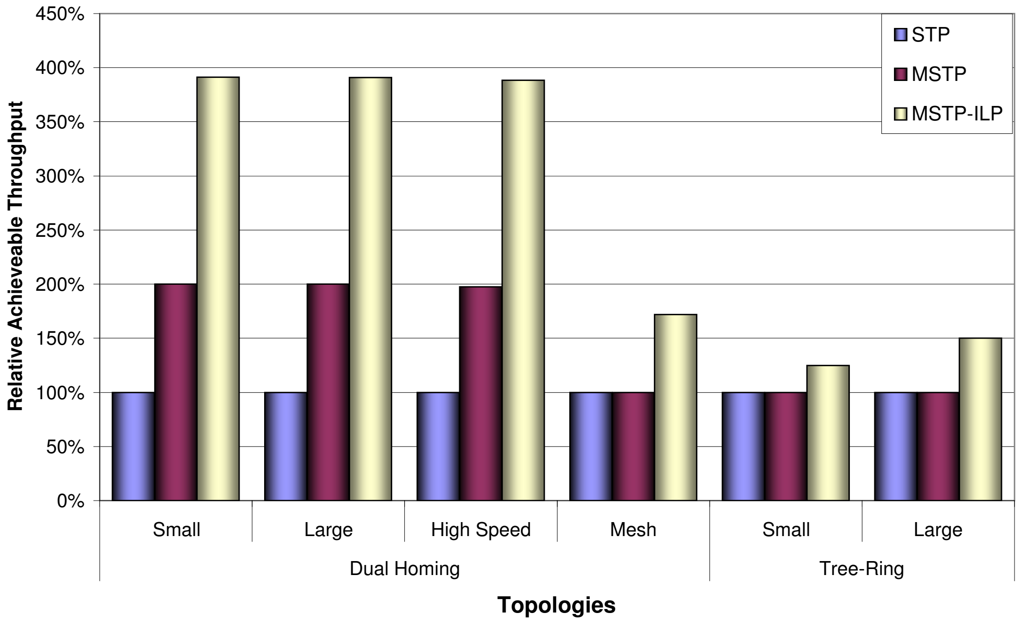
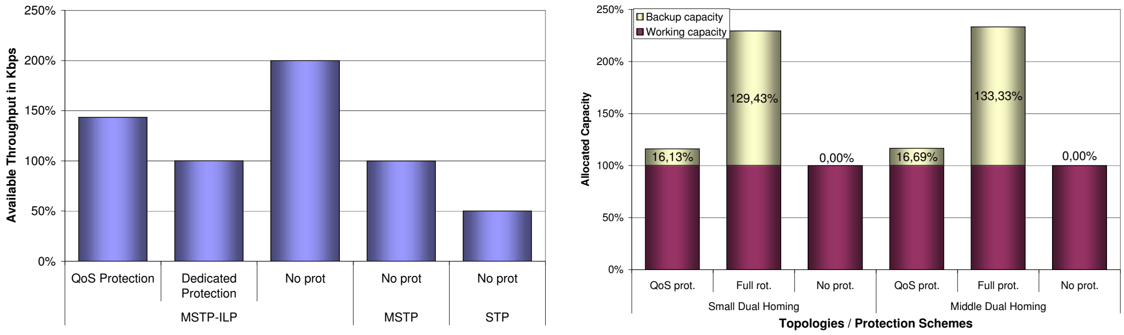


Figure 2: Achieved throughput in different topology cases



(a) Achieved throughput (b) Working and backup capacities

Figure 3: Achieved throughput and resource allocation in case of 1:1 protection.

Subthesis 1.2 *For resilient and QoS-aware Metro Ethernet networks I have given an optimal method based on ILP in case of dedicated protection scheme ([**C6**, **C8**]). I have shown that in case of Dual-Homing topology class the proposed method achieves the same throughput as the "topology driven" ones, but it is able to provide QoS guarantees after a single link failure.*

The VLAN switching based method belongs to the family of path protection schemes. Here, the backup paths also have to obey the flow conservation rule and they also have to be considered in capacity constraints. Besides, a further constraint is added prohibiting that the working and the backup paths use the same link. This ensures that the two paths will be disjoint. I make no difference between the two path classes (working and backup ones) when they are assigned to the trees.

In the case of the investigated topologies the throughput gain resulted by the optimization is consumed by the dedicated protection, that is why the same throughput level was achieved as the standard based MSTP (Figure 3(a)). However, in case of link failure the MSTP reconstructs the trees without taking the traffic conditions into account; therefore, neither the desired QoS nor the allocated capacity is not surely provided. On the contrary, the proposed method gives QoS and protection guarantees.

Subthesis 1.3 *I have proposed a novel protection scheme for reliable Metro Ethernet networks that provides protection paths only for the QoS traffic and shares the capacity allocated for protecting QoS traffic with BE traffic ([**C6**]). This way, in the case of Dual-Homing topology class and if the amount of the Best Effort traffic is 53% of the total one, the throughput can be increased by about 40% compared to the dedicated protection.*

As is can be seen on Figure 3(b), dedicating resources to backup path wastes the capacity. Thus, a goal is to apply such protection schemes that consume less resources. From this point of view one of the most efficient solutions is the shared protection [GDC+02], however, the ILP formulation of shared protection does not give solutions even for small networks within acceptable time constraints.

At the same time, for the different traffic classes different level of reliability is usually defined: e.g., for BE traffic no protection paths are defined. Further capacity saving can be realized by sharing the capacity allocated for the BE traffic with the backup paths of the priority traffic. Taking it into account, I propose a protection scheme (*QoS protection*) that defines dedicated protection for the priority traffic, however, the capacity allocated for backup can be also used by BE traffic.

Assuming that at least 50% of the traffic is best effort, the QoS can be guaranteed for the priority traffic in case of link failure. The drawback of this approach is that it does not provide protection for BE traffic. Whenever, some BE demands have to be protected, it is possible to define a backup path for them.

The simulations showed that assuming the Dual-Homing topology class, in case of QoS protection 40% higher throughput can be achieved than for dedicated protection. Although 47% of the traffic is protected (3(a)). For protection purposes this method consumes only 16% capacity compared to the dedicated protection that needs 130% 3(b). This shows the efficiency of the method.

Thesis 2 *I have developed and evaluated scalable methods for the optimization of trees in QoS-aware and reliable Metro Ethernet networks.*

Introduction Although using ILP formulation of the traffic driven tree optimization provides the optimal solution, it is hardly scalable. This unscalability is resulted by the fast increase of problem state space especially in the case of increasing the number of available tree instances and the number of demands. Moreover, the state space cannot be systematically exhausted. I propose, therefore, such heuristic algorithms that decompose the problem to several smaller tasks and these subtasks are solved using special heuristics. During the design of these algorithms I pursued the their efficiency. Efficiency means here that the algorithms have to be able to find near optimal solutions within acceptable time even in the case of larger networks.

Inspecting the structure of the ILP formulation I have divided the subtasks defined in Thesis 1 into two groups as follows:

- *Demand routing task* is to define path for the traffic demands, and to route the demands considering the network capacity and QoS constraints. This task is the modified version of the well known multicommodity flow problem since QoS constraints appear. Several general heuristic algorithms are known (e.g.,[RR01, Pio97]); however, I have improved them in order to make them able to handle the constraints of multiple QoS classes.

- *Tree covering problem* is to cover the calculated paths with as few tree instances as possible. Since two demands belonging to two different ENs cannot be assigned to the same tree instance, the problem can be further decomposed to the problem of minimal tree covering of the paths terminating at an EN.

Both identified subtasks are performed one after the other. Since the both subtasks are much smaller than the original problem, this decomposition can provide a well scalable method if efficient algorithms are developed for the subtasks. However, this method does not guarantee the optimal solution and a quality of the given solution highly depends on the partial results given for the subtasks.

Subthesis 2.1 *I have developed a scalable algorithm for the minimal tree covering problem. For the considered topology classes, the algorithm covers a given path set with minimal number of trees in 60%–100% of the cases depending on the topology [**C9**].*

The aim of the proposed Tree Assigner & Placer (TAP) algorithm is to cover a set of paths given in advance with minimal set of trees. The demands terminating at different ENs are not allowed to form common tree instances. Therefore, the path set can be decomposed into several disjoint subsets based on their EN, and if these subsets are covered by trees one by one, the quality of the solution will not be deteriorated. The proposed algorithm describes this latter problem as the sequence of the atomic operations of *tree extension* and *removal*. In order to ensure the convergence of the algorithm, the probability of selecting the extension operation must be larger than probability of removal: here, the probability of selection is set to 0.8. When all the paths are assigned to a tree instance, the solution is stored. After a predefined number of iterations that solution among the stored ones will be selected which requires the fewest tree instances. The two atomic operations are as follows:

Tree extension selects among the unassigned paths randomly with uniform distribution. Then it looks for a tree instance to which the selected path can be assigned. In other words, the selected path and the tree will not form a loop. Let us assume that after step i there are n tree instances. Then the algorithm tries to fit the selected path to the first tree instance. If it fits, i.e. it does not form a cycle, the operation finishes. Otherwise, it tries to fit the path to the next tree. If the path fits none of the trees a new tree will be created containing the considered path.

Tree removal deletes one tree among the existing tree instances that contains the less paths. The paths assigned to this tree will be released and they are added to the set of non-routed paths.

Checking path fitness: Checking whether a path forms a cycle with a tree is not trivial. In the assumed model, however, we can exploit the fact that the paths are directed to the ENs (to the root of the trees). Thus I have proposed a method with $o(n)$ complexity that takes the edges of the considered path from the source to the target one by one and it checks that the recently selected edge does not form a cycle with the tree. If not, the edge will be temporarily added to the tree, otherwise the path is supposed not to fit.

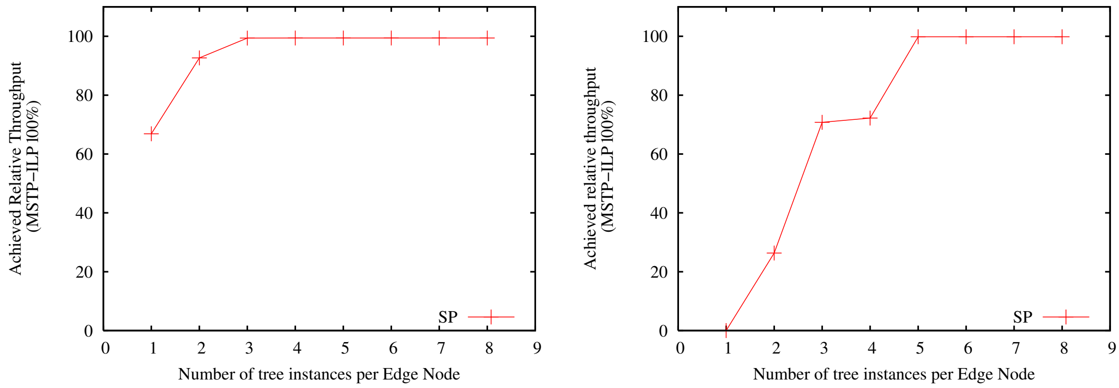
Since the performance of the algorithm highly depends on the path set obtained, to evaluate it the path given by the method proposed in Section 1.1 set is used. Through simulations I have shown that over the investigated topologies (Tree-Ring and Dual-Homing classes), in 60–100% of the considered problem instances the path set can be covered by minimal number of trees.

Subthesis 2.2 *I have developed a scalable algorithm to calculate the path sets. Combining it with the algorithm proposed in Thesis 2.1, results in a scalable method. Over the considered topology classes, this method provides the same throughput as $MSTP_{ILP}$, however, it uses more (+2-4) tree instances [**C9**].*

The algorithm proposed for the demand routing problem (Demand Router (DR)) applies the Simulated Allocation metaheuristic: it decomposes the optimization problem to the sequential search of paths for demands considering the given capacity and QoS constraints. For this purpose Dijkstra algorithm [Dij59] is a well known option that finds paths with minimal weight in a directed graph. When the demands are routed, besides the capacity constraints a further capacity-like constraint is given for each edge that is resulted by the QoS constraints. During the whole process (DRTAP), the TAP algorithm (proposed in thesis 2.1) uses the result of the DR algorithm. Depending on the topology the TAP was able to cover the path set with optimal (minimal) number of trees calculated by DR.

The performance of DRTAP is compared to the $MSTP_{ILP}$ (proposed in Thesis 1.1). The measured throughputs in case of allowing the use of different numbers of tree instances are shown on Figure 4. As can be seen, the throughput achieved by the proposed heuristic approaches the optimal solution; however, it uses more tree instances than the $MSTP_{ILP}$ does. At the same time, as the number of used tree instances increases the throughput increases as well, however, to a certain limit only. For example for Dual Homing topologies (Figure 4(a)) more than six trees (3 for each ENs) does not introduce further throughput increase. While in case of the other topology class (Figure 4(b)) no more than five trees per edge node is enough to reach the throughput limit.

Obviously, to provide higher throughput more network resources are required. Nevertheless, the amount of allocated resources describes well the efficiency of the algorithm. The lengths of the paths also characterize the efficiency. I have summarized the amount of allocated resources and the average lengths of the paths in Table 1. The standard based STP, $MSTP$ and the ILP based optimal solution are depicted as a reference. The proposed DRTAP algorithm allocates 6–10% more resources than the optimal solutions found



(a) Dual Homing

(b) Tree and ring

Figure 4: The achieved throughput with different number of tree instaces

by $MSTP_{ILP}$. At the same time, the lengths of the paths are slightly longer, that explains the increased resource reservation.

Table 1: Efficiency of DRTAP algorithm: amount of allocated resources and the average path lengths.

Topologies	**Methods**			
	References			**DrTAP**
	STP	$MSTP$	$MSTP_{ILP}$	> 4 TpR
Relative allocated capacities (compared to STP)				
Dual homing	1.00	1.33	2.66	2.80
Tree like	1.00	0.85	1.07	0.82
Average path length (hop count)				
Dual homing	4.50	3.00	3.00	3.10
Tree like	2.95	2.95	2.99	3.03

The time required for finding the solution is a criterion of the applicability of the algorithm. The measured running times in different topology cases are shown in Table 2. As it can be seen, the DRTAP has lower running time than the $MSTP_{ILP}$ in all test cases and the difference between them increases as larger topologies are investigated. In the case of a topology having 42 nodes, the $MSTP_{ILP}$ needed 5 hours. On the contrary, the DRTAP required only 5 minutes. These tendencies show that the algorithm is expected to support dynamic traffic engineering even in the case of larger networks.

Table 2: The running time of DRTAP algorithm in case of different topologies [sec]

Number of nodes Internal/Edge/Access	Algorithms ILP	DrTAP
12 / 2 / 4	5	1,2
18 / 2 / 8	7	5
24 / 2 / 12	170	15
42 / 2 / 24	18100	240

Subthesis 2.3 *I have proposed a scalable algorithm that solves the demand routing and tree cover problems* together *and it is able to deal with both, dedicated and shared protection schemes.*

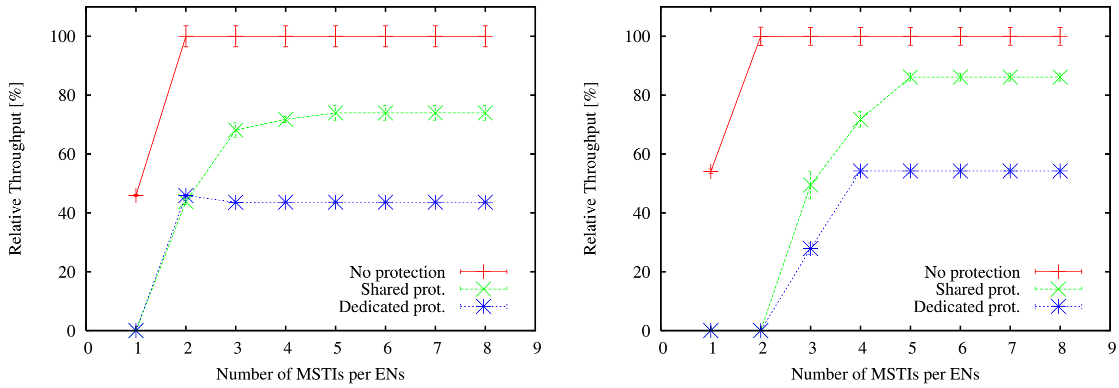
In Thesis 2 I have presented the decomposition of the optimization problem and I have proposed a scalable algorithms for both steps (Theses 2.1 and 2.2). However, in the demand routing step the algorithm does not take into account that the paths will be covered by trees that may deteriorate the solution. To avoid this problem a novel algorithm is proposed by this thesis.

The proposed algorithm – Joint Router and Tree Placer (JRTP) – decomposes the original problem to the sequence of demand routing and tree release operations. The demand ordering problem is solved considering the Simulated Allocation (SAL) metaheuristic. The opposite of the "TAP tree append operation" the "demand routing" operation of the proposed algorithm assigns the demands to a tree instance *before* it looks for a path for the demand. This step enables us to avoid the case when the path to be assigned to the demand forms a loop with the tree instance. That is why the path search is performed over a reduced graph that is calculated as follows:

1. All edges, where the capacity or the QoS constraints are violated, are temporarily deleted.

2. All the edges that "point out" of the considered tree instance (i.e., their source nodes are in the tree but their targets are not) are pruned.

To realize path protection schemes two or more edge disjoint paths (or VLANs) are calculated for each traffic demand. For this purpose more algorithms are known, for instance, Edmonds-Karp [Kar75], Suurballe [ST84], Double Dijkstra [**C1**]. In the proposed algorithm I have selected this latter one. Furthermore, I had to extend the demand routing operation: two VLANs are routed one by one. First, the working path is defined and assigned to a tree instance, then, the same steps are performed for the backup path as well. However, in this second step, all the edges, that are used by the working path, are also



(a) Small (12 bridges)

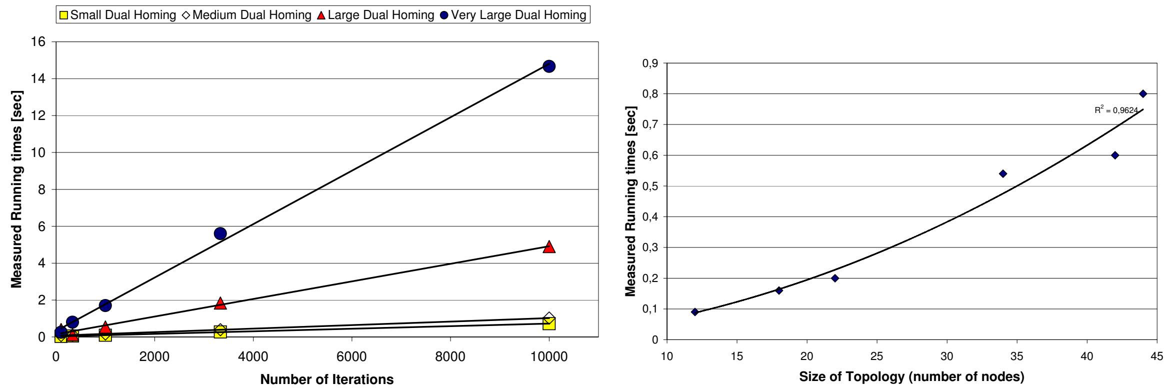
(b) Medium (18 bridges)

Figure 5: Achievable throughput in the cases of dedicated and shared protection schemes for different number of tree instances.

removed. In the case of Shared Protection, the method of calculating the reserved backup capacities are changed, a Backup Reservation Matrix [**C1**] is applied.

The JRTP is evaluated the same way as the Demand Router and Tree Assigner & Placer (DRTAP): the achievable throughput is determined allowing different numbers of tree instances per edge node. Figure 5 summarizes the achieved throughput of the proposed JRTP while the 100% is the results provided by $MSTP_{ILP}$ without protection. During the simulations the Dual-Homing topology case is assumed. Using dedicated protection the throughput is roughly 50%, while sharing the backup capacities among the backup paths the throughput is 75-80% compared to the case when no protection is considered. Like DRTAP algorithm, the continuous increase of the throughput to a limit is observed. This limit depends not only on the topology, but on the type of the applied protection: to provide protection paths more than two times more trees are required. Moreover, to fully exploit the resource gain of shared protection further tree instances are used.

The scalability of the proposed JRTP is investigated through measuring the dependence of the running time on the number of iterations and on the size of the topology (this latter criterion is described by the number of nodes). The measured running times are depicted on Figure 6. We can make two major observations. First, the running time is roughly linear function of the number of iterations as we expected. At the same time the running time has polynomial dependence on the size of the topology. To better illustrate it, I have fitted a second order function to the measured running times. Note that, these running times cannot be compared to the measured one for DRTAP because the two simulation environments differ (different computers are used).



(a) Running times vs. number of iterations

(b) Running times vs. number of switches

Figure 6: The running times of JRTP.

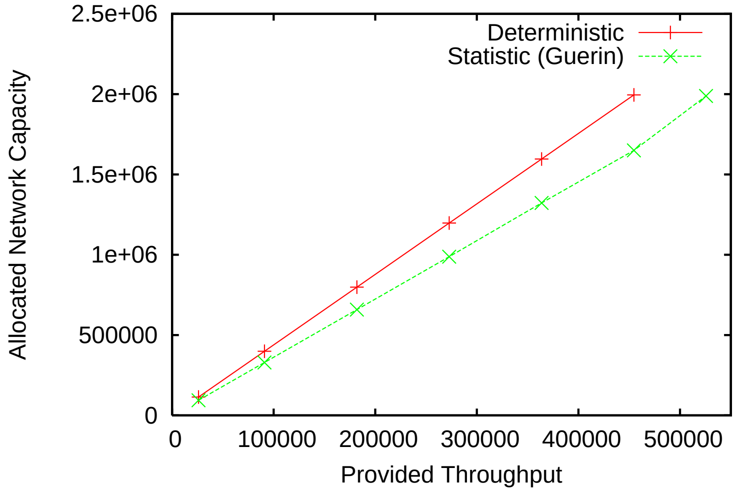
Thesis 3 *I have extended the proposed model and algorithms with statistical multiplexing capabilities and the support of two-layer optical Ethernet architecture.*

Subthesis 3.1 *Based on the algorithm in Thesis 2.3 I have proposed a method that is able to exploit the gain of statistical multiplexing during the network optimization [**C7**]. In Dual-Homing topology class 20% increase of throughput is achieved compared to the deterministic multiplexing.*

The efficiency of resources usage can be increased if the statistical behavior of the traffic flowing in the pipes is considered during network optimization. Among the investigated models [GAN91, Flo96, Lin94] I have selected the Guerin based one because of its performance [**M1**]. This model assumes that bandwidth consumption of the traffic has Gaussian distribution. For n logical pipes the total amount traffic to be allocated is: $BW = \sum_{i=1}^{n} m_i + \alpha \cdot \sigma$, where m_i is the average bit rate for pipe i, σ is the standard deviation of the aggregated traffic and α depends on the packet loss ratio and set by the operator. The deviation (σ) of the aggregation can be estimated based on the deviation (and on the peak-to-mean bit ratio) of the basic flows.

In order to implement this aggregation model, I have improved the heuristic algorithm proposed in thesis 2.3.

During the simulations Triple-Play service model is assumed. Beside the basic parameters of the pipes (like global traffic level and traffic class dependent ratios) a further parameter has to be defined that defines a ratio between the average and peak bit rates. The exact value depends on the class of service and it is set to 1.0 for VoIP, 1.2 for video streaming and 2.0 for high speed internet services, respectively. The question is the impact of considering statistical multiplexing gain at demand routing on the amount allocated capacities as well as on the achievable throughput. On Figure 7 the resources allocated



(a) Middle Dual Homing Topology

Figure 7: Allocated network resources in different throughput levels.

by both, the Guèrin based statistical and the deterministic multiplexing models in different throughput levels can be seen. Those throughput levels, where no solutions are found, are not depicted. These results present that by using Guerin's model the allocated capacity decreases 15%. Moreover, this decreased resource usage makes possible to increase the achievable throughput with 15–20%. Emphasize that, although the achieved throughput increase seems to be trivial, in the proposed extension of that algorithm the effects of statistical multiplexing are considered *during* the network configuration task, i.e., during the optimization.

Subthesis 3.2 *I presented the applicability of Coarse Wavelength Division Multiplexing (CWDM) in Metro Ethernet environment and I have modified the ILP model in order to solve the critical problems resulted by the introduction of protection schemes. I have shown a further advantage of CWDM: with the help of optimizing of wavelength paths the number of optical interfaces can be significantly (50%) decreased at the cost of a minor, still acceptable throughput loss [**C13**].*

In the past few years, optical transmission technologies have appeared not only in core networks, but also in metropolitan and access ones to satisfy the steadily increasing bandwidth hunger. A viable solution to further increase the capacities of the connections is Wavelength Division Multiplexing (WDM), especially the Coarse Wavelength Division Multiplexing (CWDM). In CWDM systems, up to 8 parallel wavelength channels are used and they are aggregated using IEEE 802.3ad. Due to the constraints of the topology and the high costs of transponders, a more efficient (cheaper) network would be established, if a logical topology formed by wavelength paths was defined over the physical one. This way, the distant nodes are interconnected with "direct links" using these wavelength paths. This approach, however, involved critical protection problems, since the fail of a physical link affects more wavelength paths and, therefore, more logical links. To solve this problem I have extended the ILP model proposed in thesis 2.1 with the support of Shared Risk Link Groups (SRLGs).

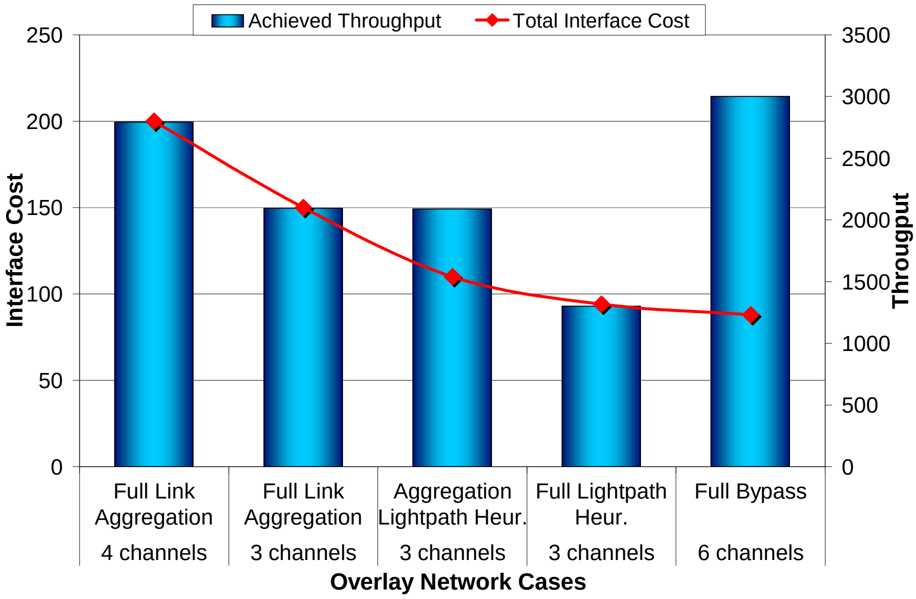


Figure 8: Achievable throughput and the cost of the installed interfaces.

The capabilities of the further developed model are illustrated over a test topology formed by four interconnected rings. Over this physical one different logical topologies are defined. I have considered 1:1 protection switching for sake of simplicity. The building costs (described with the number of optical interfaces and shown on Figure 8 with the solid line) of the logical topologies are obviously different. The full link aggregation requires the most interfaces. Figure 8 also shows that by establishing direct connections between distant nodes and bypassing them under the intermediate switches significantly decreases the number of used optical-electric converters (transponders). At the same time, the decrease of the served traffic (depicted by the columns) is only because of the decreased link capacities. Based on this results the Full Bypass seems to be the best option, however, it required many dedicated lightpaths making it inflexible against the change of traffic.

A further problem of spanning trees is that the RSTP performs better over topologies with smaller diameter when "swallower" trees are to be spanned. Table 3 presents the

Table 3: The smallest / average / highest path lengths.

Test case	**Without deep constraint on trees**		**Depth of trees <= 7**	
	Working	**Backup**	**Working**	**Backup**
Full Link Aggregation	1.00 / 3.17 / 5.00	4.00 / 6.17 / 8.00	-	-
Full Bypass	1.00 / 1.67 / 2.00	2.00 / 2.34 / 3.00	1.00 / 1.67 / 2.00	2.00 / 2.34 / 3.00
Aggregation Lightpath Heuristic	1.00 / 2.16 / 3.00	2.00 / 3.76 / 5.00	1.00 / 2.16 / 3.00	2.00 / 3.76 / 5.00
Full Lightpath Heuristic	1.00 / 2.00 / 3.00	2.00 / 5.03 / 11.00	1.00 / 2.73 / 5.00	2 / 4.28 / 7.00

shortest, the average and the longest working and backup paths in different test cases. Since the TE pipes runs from the leaves to the root their length estimates the depth of the tree. Furthermore, the diameter of the logical network can be decreased, thus, the lengths of the AN-EN paths also decrease.

5 Application of the Results

The main goal of my dissertation was to develop traffic engineering methods with high availability and QoS support in a well-defined network architecture. This architectural environment implies the applicability of the proposed algorithms as well.

All proposed methods, the Integer Linear Program based and the heuristic ones was able to determine efficient configurations of Ethernet based metropolitan networks while both the QoS and availability requirements were satisfied. Therefore, value added services like Triple-play, where telephone, television and internet services could be provided over a common network infrastructure. Besides, through finding the efficient configurations the methods support even dynamic traffic engineering to adapt the network to the changing traffic requirements.

6 Acknowledgements

This thesis would not be born without my advisors Dr Tibor Cinkler, Dr József Bíró and Dr Gyula Sallai. The every day discussions, and their deliberate revisions made this thesis more adequate and useful. I want also to thank István Moldován for guiding me in the forrest of technological aspects of Metro Ethernet. Furthermore, I thank High Speed Networks Laboratory for the technical support, especially Dr. Tamás Henk for his support.

Last, but not least, I would most like to thank to my family for their patience.

The work in connection with this thesis has been done within the research co-operation

framework of three European IST-FP6 projects: NoE e-Photon/ONe (http://www.e-photon-one.org/), IP Multiservice Access Everywhere (http://www.ist-muse.org/) and IP Next generation Optical network for Broadband European Leadership (http://www.ist-nobel.org/).

References

[3GP06] 3GPP. *TS-22.105: Technical Specification Group Services and System Aspects Service Aspects; Services and Service Capabilities (Release 8)*, June 2006.

[BvdSBP03] Christele Bouchat, van den Sven Bosch, and Thierry Pollet. QoS in DSL Access. *IEEE Communications Magazine*, 41(9):108–114, Sept. 2003.

[CPL] ILOG CPLEX v9.1, http://www.ilog.com/products/cplex/index.cfm.

[CS06] Amit Cohen and Ed Shrum. Migration to Ethernet-Based DSL Aggregation. Technical Report TR-101, DSL Forum, April 2006.

[Dij59] Edsger. W. Dijkstra. A Note on Two Problems in Connexion with Graphs. In *Numerische Mathematik*, volume 1, pages 269–271, 1959.

[Flo96] Sally Floyd. Comments on Measurement-based Admissions Control for Controlled-Load Services. submitted to CCR, July 1996., July 1996.

[FTAW05] János Farkas, Gábor Tóth, Csaba Antal, and Lars Westberg. Distributed Resilient Architecture for Ethernet Networks. In *The 5th International Workshop on Design of Reliable Communication Networks (DRCN'2005)*, October 2005.

[GAN91] Roch Guérin, Hamid Ahmadi, and Mahmud Naghshineh. Equivalent Capacity and its Applications to Bandwidth Allocation in High-Speed Networks. *IEEE JSAC*, 9(7):968–981, September 1991.

[GDC$^+$02] Wayne Grover, John Douchette, Matthieu Clouqueur, Dion Leung, and Demetrios Stamatelakis. New Option and Insights for Survivable Transport Networks. *IEEE Communications Magazine*, 40(1):34–41, January 2002.

[GP] Michael Galea and Marzio Pozzuoli. *Redundancy in Substation LANs with the Rapid Spanning Tree Protocol (IEEE 802.1w)*. RuggedCom Inc. - Industrial Strength Networks, Concord Ontario Canada.

[IEEE802.1w] "Media Access Control (MACc) Bridges: Rapid reconfiguration of Spanning Tree", IEEE 802.1w, IEEE 2001, incorporated into IEEE 802.1D-2004.

[IEEE802.1s] "Virtual Bridged Local Area Networks: Multiple Spanning Trees", IEEE 802.1s, IEEE, 2002.

[IEEE802.1Q2003] "Local and Metropolitan Area Networks, Virtual Bridged Local Area Networks", IEEE 802.1Q-2003, IEEE, 2003.

[IEE05a] Draft Standard for Local and Metropolitan Area Networks - Virtual Bridged Local Area Networks - Amendment 5: Connectivity Fault Management, 2005.

[IEE05b] Virtual Bridged Local Area Networks - Amendment 4: Provider Bridges, 2005.

[G.8031] "Ethernet Protection Switching (DRAFT)", ITU-T.G.8031, ITU, February 2006.

[G.1010] "End-User Multimedia QoS Categories", ITU-T G.1010, ITU, November 2006.

[Kar75] Richard Manning Karp. On the Computational Complexity of Combinatorial Problems. In *Networks*, volume 5, 1975.

[Lin94] Karl Lindberger. Dimensioning and Design Methods for Integrated ATM Networks. In *The 14th International Teletraffic Congress*, pages 897–906, 1994.

[LLN02] King-Shan Lui, Whay Chiou Lee, and Klara Nahrstedt. STAR: a Transport Spanning Tree Bridge Protocol with Alternate Routing. In *ACM SIGCOMM Computer Communications Review*, volume 32, page 22//46, July 2002.

[LSJ98] William P. Lidinsky, Mick Seaman, and Tony Jeffree. Media Access Control (MAC) Bridges. standard 802.1D, ANSI/IEEE, 1998.

[LYD$^+$03a] Yujin Lim, Heeyeol Yu, Shirshanka Das, Scott Seongwook Lee, and Mario Gerla. Efficient Building Method of Multiple Spanning Tree for QoS and Load Balancing. In *GLOBECOM 2003*, volume 7, pages 3620–3625, December 2003.

[LYD$^+$03b] Yujin Lim, Heeyeol Yu, Shirshanka Das, Scott Seongwook Lee, and Mario Gerla. QoS-aware Multiple Spanning Tree Mechanism over a Bridged LAN Environment. In *GLOBECOM 2003*, volume 6, pages 3068–3072, December 2003.

[Pio97] Michael Pioro. Simulation Approach to the Optimization of Multicommodity Integral Flow Networks. In *Selected Proceedings of the Third INFORMS Telecommunication Conference, Telecommunication Systems*. Balzer Science Publishers, 1997.

[RR01] Mauricio G. C. Resende and Celso C. Ribeiro. Greedy Randomized Adaptive Search Procedures. Technical report, AT&T Labs Research Technical Report, September 2001. Revision 2, August 29, 2002. To appear in "State of Art Handbook in Metaheuristics", F. Glover, G. Kochenberger, eds. Kluwer, 2002.

[SGNcC04] Srikant Sharma, Kartik Gopalan, Susanta Nanda, and Tzi cker Chiueh. Viking: a Multi-spanning-tree Ethernet Architecture for Metropolitan Area and Cluster Networks. In *INFOCOM 2004. Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies*, volume 4, pages 2283–2294vol.4, 7-11 March 2004.

[ST84] John W. Suurballe and Robert E. Tarjan. A Quick Method for Finding Shortest Pairs of Disjoint Paths. *Networks*, 14:325–336, 1984.

List of Publications

[J] **Journal Papers**

[**J1**] Mátyás Martinecz, **András Kern**, Zalán Heszberger and József Bíró "Architecture and Configuration of Broadband Access Networks Supporting Multimedia Applications" *International Journal of Computer and Their Application* vol. 14, No. 1, pp. 2–12, March, 2007.

[**J2**] **András Kern**, István Moldován, Tibor Cinkler, "Efficient TE and Protection in Metro Ethernet", SUBMITTED TO *Mediterrian Journal of Computers and Networks*.

[I] **Journal Papers in Hungarian Journals**

[**I1**] **András Kern**, György Somogyi, Tibor Cinkler, "Applying statistical multiplexing and traffic grooming in optical networks jointly"*Híradástechnika*, LXI, July 2006.

[M] **Hungarian Journal Papers**

[**M1**] **Kern András**, Somogyi György and Cinkler Tibor, "Statisztikus nyalábolás és forgalom kötegelés együttes hatása optikai hálózatokban", *Híradástechika* LXI, February 2006.

[C] **Conference Papers**

[**C1**] Balázs Gábor Józsa, Dániel Orincsay and **András Kern** "Surviving Multiple Network Failures Using Shared Backup Path Protection", In Proceedings *The 8th IEEE Symposium on Computers and Communications (ISCC 2003)*, vol. 2, pp. 1333–1340, Antalya, Turkey, June 30 - July 3 2003.

[**C2**] Dániel Orincsay, Balázs Gábor Józsa and **András Kern** "On the Use of Routing Optimization for Virtual Private Network Design" In Proceedings *The 7th IFIP Working Conference on Optical Network Design & Modelling (ONDM 2003)*, February 3 - 5, Budapest, Hungary.

[**C3**] **András Kern**, Mátyás Martinecz, Zalán Heszberger and Gyula Sallai "Architecture and Configuration of Broadband Access Networks Supporting Multimedia Applications" In Proceedings *The 10th IEEE Symposioum on Computers and Communications (ISCC'2005)* pp. 173–178, June 27–30, Cartagena, Spain.

[**C4**] Tibor Cinkler, István Moldován, **András Kern**, Csaba Lukovszki and Gyula Sallai, "Optimizing QoS Aware Ethernet Spanning Trees", In Proceedings *1st International Conference on Multimedia Services Access Networks (MSAN'2005)* pp. 30–34, Orlando, Fl, USA, June, 2005.

[**C5**] Tibor Cinkler, Géza Geleji, Márk Asztalos, Péter Hegyi, András Kern and János Szigeti, "Lambda-path Fragmentation and De-Fragmentation through Dynamic Grooming" *The 7th International Conference on Transparent Optical Networks (ICTON 2005)*, vol. 2, pp 1–4, July 3 - 7, 2005, Barcelona Spain.

[**C6**] Tibor Cinkler, **András Kern** and István Moldován "Optimized QoS Protection of Ethernet Trees" In Proceedings *The 5th International Workshop on Design of Reliable Communication Networks (DRCN'2005)* Ischia, Naples, Italy, 16–19 October, 2005.

[**C7**] **András Kern**, István Moldován and Tibor Cinkler, "On the Optimal Configuration of Metro Ethernet for Triple Play" *2nd Conference on Next Generation Internet Design and Engineering (NGI 2006)* pp. 334–341, València, Spain 2006.

[**C8**] **András Kern**, István Moldován and Tibor Cinkler "Traffic-driven Optimization of Routing for Metropolitan Ethernet Networks" In Proceedings *World Telecommunications Congress 2006* May 2006, Budapest, Hungary.

[**C9**] **András Kern**, István Moldován and Tibor Cinkler "Scalable Tree Optimization for QoS Ethernet" *The 11th IEEE Symposioum on Computers and Communications (ISCC'2006)* pp. 578–584, Cagliari, Italy, 26–29 June 2006.

[**C10**] **András Kern**, György Somogyi, Tibor Cinkler "On the Gain of Statistical Multiplexing Over Traffic Grooming" *8th International Conference on Transparent Optical Networks*, vol 3., pp. 112–115, 18–22 June 2006, Nottingham, UK.

[**C11**] Cinkler Tibor, Hegyi Péter, Asztalos Márk, Geleji Géza, **Kern András**, Szigeti János, "Multi-Layer Traffic Engineering through Adaptive Lambda-path Fragmentation and De-Fragmentation: The Grooming-Graph and Shadow-Capacities", *Networking 2006*, pp. 715–726, 15–19 May 2006, Coimbra, Portugal.

[**C12**] Tibor Cinkler, **András Kern** and István Moldován, "Dimensioning Transport Networks for VPNs over Capacities with Stepwise Costs", *Networks 2006*, pp. 1–4, 6–9 November, 2006, New Delhi, India.

[**C13**] **András Kern**, István Moldován, Péter Hegyi and Tibor Cinkler, "Ethernet over WDM: Optimization for Resilience and Scalability", *The 6th International Workshop on Design of Reliable Communication Networks (DRCN'2007),* 7–10 October 2007, La Rochelle, France.

[**C14**] **András Kern**, István Moldován and Tibor Cinkler, "Bandwidth Guarantees for Resilient Ethernet Networks through RSTP Port Cost Optimization", *Accessnets 2007*, Ottawa, Canada, 22–24, August, 2007.