Traffic Engineering in IP over Optical Transport Networks for Metropolitan and Wide Area Environments

SALVATORE SPADARO

ADVISOR: DR. JOSEP SOLÉ PARETA

COMPUTER ARCHITECTURE DEPARTMENT
UNIVERSITAT POLITÈCINICA DE CATALUNYA

A THESIS PRESENTED TO THE UNIVERSITAT POLITÈCNICA DE CATALUNYA IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF

Doctor en Ingeniería de Telecomunicación

December 2004

Dr.	
President	

Dr.	
Secretari	

Dr.	
Vocal	

Dr.	
Vocal	

Dr.	
Vocal	

Data de la defensa pública	
Qualificació	

Acknowledgements

Es realmente muy difícil dar las gracias a todos aquellos que han contribuido de alguna forma u otra a que esta Tesis fuese posible, realmente debería ser una lista interminable.

Ante todo quiero agradecer a mi familia, que aunque lejana siempre ha sido un apoyo en los momentos difíciles y una referencia continua.

Quiero dar las gracias a Josep Solé Pareta. Quiero agradecerle la confianza que siempre ha demostrado en mí a lo largo de esta Tesis y también por haberme introducido, a parte de en su grupo de investigación, en el mundo de los proyectos europeos, contribuyendo de forma determinante a mi desarrollo como investigador y también como persona.

Pero quiero darle las gracias también por los continuos consejos y ánimos que me han hecho seguir adelante también en los momentos de dificultad. Nunca me ha faltado un consejo o un *avanti* cuando lo he necesitado. Le agradezco también haberme ayudado a descubrir los muchos aspectos de Catalunya; sin él, mis conocimientos de catalán no habrían llegado al nivel en que están.

Asimismo quiero dar las gracias a todos los componentes del CCABA sin los cuales esta tesis no se habría podido hacer. En particular quiero agradecer a Pere Barlet por haberme ayudado en cada momento que lo he necesitado.

Esta tesis se ha desarrollado básicamente en el entorno del proyecto europeo LION. Participar en este proyecto, no sólo me ha ayudado con mi tesis sino me ha permitido también conocer a personas brillantes, tanto desde un punto de vista técnico como humano. En particular, quiero agradecer a Antonio Manzalini y Marco Quagliotti de Telecom Italia Lab y Chris Wajda de AGH por su incomparable ayuda en el desarrollo de las ideas contenidas en esta tesis.

Quiero dar las gracias a Davide. Han sido varios años compartiendo en lo profesional casi todo, como las muchas noches de trabajo muy próximas a los *deadlines*. Siempre ha sido una referencia por su forma de trabajar. Además es un amigo y hemos pasado juntos (y pasamos) muchas cosas desde nuestra llegada a Barcelona. En especial quiero agradecerle el apoyo que siempre me ha brindado en cualquier circunstancia tanto en el trabajo como en la vida privada y por haber aguantado mis infinitas quejas durante las comidas o cafés. Realmente le deseo el éxito en la vida que se merece.

Quiero agradecer a todos aquellos que, haciendo su proyecto final de carrera, me han ayudado. Doy las gracias a David Alarcón, Lorenzo Marqués, Marc Nogueras, Óscar Guell y Raimundo Alonso. Un especial agradecimiento es para Joaquim Recio y Óscar Pisa que han contribuido de forma muy importante a esta Tesis.

En particular quiero dar las gracias a dos verdaderos amigos; a Javier Feria por su ilimitada disponibilidad, por la calidad de su trabajo y por su calidad humana y a Fabrizio, cuya ayuda ha sido fundamental para finalizar la Tesis.

Sin embargo, la persona a la que más tengo que agradecer es a Sandra, que ha vivido esta Tesis como si fuera la suya. Gracias por la paciencia, por el apoyo incondicionado que me ha dado siempre y por soportarme sobre todo en el último período.

Resumen

La arquitectura de las redes de transporte actuales está basada en la tecnología de transporte *Synchronous Digital Hierarchy* (SDH). Las redes SDH se han diseñado y están optimizadas básicamente para el transporte del tráfico de voz. Actualmente, se está experimentando un crecimiento exponencial del volumen de tráfico de datos. Este crecimiento se debe a que el protocolo IP se está consolidando como capa de integración para servicios múltiples, algunos de ellos con requerimientos de Calidad de Servicio (*QoS*) y también a la introducción de tecnología de acceso de alta velocidad. Las características estadísticas del tráfico de datos son diferentes respecto a las del tráfico telefónico. De hecho, el tráfico IP se caracteriza no solo por su asimetría sino por su naturaleza dinámica, ya que presenta fluctuaciones o picos difíciles de predecir a priori.

Como consecuencia, ha surgido la necesidad de emigrar desde las actuales redes hacia una estructura más flexible y dinámica, optimizada para el transporte de tráfico de datos.

La evolución de las actuales redes de transporte incluye trasladar todas las funcionalidades de SDH (conmutación, monitorización de la calidad de la señal, protección frente a fallos) a nivel óptico. El resultado consistirá en una red de transporte óptica (*Optical Transport Network*, OTN) basada en tecnología DWDM, con *Optical Cross Connects* (OXC) para encaminar canales ópticos de forma permanente o conmutada (*Automatic Switched Optical Network*, ASON).

Uno de los principales problemas a solucionar por las operadoras de red es la eficiente gestión de la capacidad disponible, y así evitar por un lado la necesidad de sobredimensionar la red de transporte y por el otro optimizar la utilización de los recursos mediante la definición de estrategias de ingeniería de tráfico.

La introducción de las redes de transporte a conmutación automática (ASON), capaces de proporcionar conexiones ópticas bajo demanda, es considerada como la solución de red que puede proporcionar el rápido y flexible aprovisionamiento de ancho de banda. Tal funcionalidad, posible gracias a la definición de un plano de control basado en el paradigma GMPLS, puede ser usada para gestionar de manera dinámica los recursos disponibles, tanto a nivel SDH como a nivel óptico, respondiendo de forma eficiente a las fluctuaciones del tráfico generado por la red cliente.

Sin embargo, el problema que surge es el diseño de un mecanismo para disparar automáticamente las peticiones de establecimiento de circuitos SDH/canales ópticos conmutados.

En este sentido, la primera contribución de esta Tesis es el diseño de un mecanismo de disparo de peticiones de circuitos SDH/canales ópticos basado en la monitorización y predicción del tráfico de la red cliente (IP). Además, el mecanismo diseñado incluye la definición de políticas de

ingeniería de tráfico para la optimización de la utilización del elevado ancho de banda proporcionado por las conexiones ópticas. Concretamente, el mecanismo diseñado se caracteriza por la interoperabilidad entre la capa cliente y la capa de transporte.

La Tesis incluye también una contribución sobre el diseño de una metodología para el dimensionado de la redes ASON, basada en la caracterización del tráfico de llegadas de peticiones de establecimiento de conexiones, mediante su valor medio y el factor de *peakedness*.

Por otro lado, la optimización de los recursos disponibles es muy crítica cuando se produce un fallo en la infraestructura de red debido a la necesidad de encontrar rutas alternativas para el tráfico afectado. Debido al gran volumen de tráfico a transportar, un fallo en la infraestructura de red puede tener graves consecuencias económicas. Por ejemplo, un corte de una única fibra óptica produce el fallo de todas las longitudes de onda que transporta; de esta manera la pérdida de cada longitud de onda operante a 2.5 Gbps o 10 Gbps puede resultar en el corte de un enorme número de conexiones en curso. Por lo tanto, a mayor capacidad, mayor es la importancia de la rapidez y rendimiento de los mecanismos de protección y recuperación.

Las estrategias de protección frente a fallos deben ser simples, minimizar las pérdidas de tráfico y deben utilizar eficientemente los recursos disponibles.

La recién estandardizada tecnología para redes de entornos metropolitanos, *Resilient Packet Ring* (RPR) se caracteriza por mecanismos de protección optimizados para minimizar el tiempo de recuperación en caso de fallos. Además, tales mecanismos no requieren la asignación a priori de recursos de red a utilizar solamente en caso de fallos.

Por lo que respecta a los mecanismos de recuperación, se puede optar por una estrategia de recuperación en una sola capa (*single layer recovery*) o alternativamente por una estrategia de recuperación en múltiples capas (*multi-layer recovery*), donde en la recuperación intervienen diferentes capas de la estructura de red. El esquema de recuperación multi-capas más fácil de implementar es el consistente en ejecutar los mecanismos de protección/recuperación de los distintos niveles de manera paralela e independiente. Esta estrategia no es, sin embargo, la más eficiente. La interoperabilidad entre los mecanismos de protección de las diferentes capas permite reaccionar más rápidamente a los fallos que se pueden producir.

La segunda contribución de esta Tesis es el diseño de una política de coordinación entre los mecanismos de protección proporcionados por RPR y los mecanismos de protección definidos por la capa óptica. Concretamente, la estrategia diseñada se basa en la interoperabilidad entre la capa

RPR y la capa de transporte (OTN) para redes de entornos metropolitanos. La estrategia diseñada permite, además, la optimización de los recursos de red.

Table of Contents

Acronyms

APS	Automatic Protection Switching
ASON	Automatically Switched Optical Network
ASTN	Automatically Switched Transport Network
ATM	Asynchronous Transfer Mode
AU	Administrative Unit
BoD	Bandwidth on Demand
CAPEX	Capital Expenditure
CBR	Constraint Based Routing
CC	Connection Controller
CIR	Commited Information Rate
CT	Craft Terminal
DCN	Data Communication Network
DPT	Dynamic Packet Transport
DWDM	Dense Wavelenght Division Multiplexing
EGP	Exterior Gateway Protocol
ESCON	Enterprise Stems Connection
FDDI	Fibre Distribution Data Interconnection
FICON	Fibre Connection
FSC	Fibre Switch Capable
GFP	Generic Framing Procedure
GMPLS	Generalized Multi-Protocol Label Switching
GoS	Grade of Service
HDLC	High Level Data Link Control
HOVC	High Order Virtual Container
IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet Engineering Task Force
IP	Internet Protocol
ISDN	Integrated Services Digital Network
IS-IS	Intermediate System – Intermediate System
ITU-T	International Telecommunications Union – Telecommunications Sector
LCAS	Link Capacity Adjustment Scheme
LDP	Label Distribution Protocol
LMP	Link Management Protocol
LoL	Loss of Light
LoS	Loss of Signal
LOVC	Low Order Virtual Container
LRM	Link Resource Manager
LSP	Label Switched Path
LSR	Label Switched Router
MAC	Medium Access Control
MEMS	Micro Electro Mechanical Systems
MIB	Management Information Base

MPLS	Multi-Protocol Label Switching
MS	Multiplex Section
NNI	Network Network Interface
OADM	Optical Add and Drop Multiplexer
ODU	Optical Digital Unit
OIF	Optical Internetworking Forum
OPEX	Operational Expenditure
OSPF	Open Shortest Path First
OSS	Operational Support System
OTN	Optical Transport Network
OUT	Optical Transport Unit
OVPN	Optical Virtual Private Network
OXC	Optical Cross-Connect
PC	Permanent Connection
PDH	Plesiochronous Digital Hierarchy
PDU	Payload Data Unit
PoS	Packet over SONET/SDH
POTS	Plain Old Telephone Service
QoS	Quality of Service
RC	Routing Controller
RPR	Resilient Packet Ring
RS	Regeneration Section
RSVP	Reservation Resource Protocol
RWA	Routinf and Wavelenght Assignment
SC	Switched Connection
SDH	Synchronous Digital Hierarchy
SF	Signal Fail
SNMP	Simple Network Management Protocol
SONET	Synchronous Optical Network
SPC	Soft Permanent Connection
STM	Synchronous Transport Module
TC	Traffic Controller
TD	Topology Discovery
TDM	Time Division Multiplexing
TE	Traffic Engineering
UNI	User Network Interface
VC	Virtual Container
VCAT	Virtual Concatenation
VCG	Virtual Containers Group
VCS	Virtual Container Selector
VPN	Virtual Private Network
WDM	Wavelength Division Multiplexing

List of Figures

List of Tables

Abstract

The main objective of the traffic engineering (TE) strategies is the efficient mapping of the actual traffic onto the available network resources.

Legacy Time Division Multiplexing-based networking architecture was basically designed to transport symmetric voice traffic. However, the volume of data traffic is increasing at explosive rate and already dominates the voice traffic. This is due to a progressive migration of many applications and services over the Internet Protocol (IP) and also to a deeper and deeper introduction of high-speed access technologies. Also there is the convergence towards the IP of real-time applications (i.e. multimedia applications) which have very strict QoS requirements.

The statistical characteristics of the data traffic are rather different from those of telephone traffic. Specifically, IP traffic is highly dynamic showing predictable and unpredictable traffic surges/peaks. Such surges are caused by unexpected events such as user' behaviours, weather conditions, accidents, faults, etc. This can cause significant fluctuations of the aggregated data traffic to be carried by the transport networks.

The current SONET/SDH transport networks (but also the incoming Optical Transport Networks) tend to be static, which means that connections (SONET/SDH circuits and light paths) are provided manually through the Network Management System (NMS). The manual configuration is time consuming, which means that weeks or even months are needed to provide high bandwidth connections.

The highly dynamic IP traffic pattern does not match with the static provisioning of capacity of the optical transport networks, leading to non-optimal utilization of the resources (i.e. network congestion or under-utilization of resources).

Thus, the problem that arises for Network Operators is how to efficiently manage the network resources in the transport network to efficiently respond to the changes in the traffic demands reaching, in such a way, traffic engineering objectives.

The introduction of the Automatic Switched Optical Networks (ASON), which is able to provide dynamically switched connections on demand, is recognized as the enabling solution to meet the requirement of fast and flexible end-to-end bandwidth provisioning. The automatic set up and tear down of optical connections can be used for the dynamic management of the transport network resources to track significant variations in the volume of the network client traffic. In such a context, a mechanism that triggers demands to set up/tear down light paths as a function of the variation of the client traffic to be transported is required.

The design of a multi-layer traffic engineering (MTE) strategy for IP/MPLS over ASON/GMPLS networks to face with the dynamic traffic demands is the first contribution of this Ph.D. Thesis. It has to be underlined that the policies for the set up of the light paths are out of the scope of this work. In fact, it is assumed that the set up/tear down of the switched connections is in charge of the ASON control plane, namely the GMPLS-based routing and signalling protocols.

As a second contribution, it is presented a practical approach for ASON networks dimensioning purposes based on the approximate characterization of the traffic arrival process, through its mean and the peakedness factor.

On the other hand, the optimization of the utilization of network resources is very critical when failures occur in the network as a consequence of the need of rerouting the affected traffic. The increase of the capacity and number of wavelengths that can be multiplexed onto the same fibre, each one carrying 2.5 or 10 Gbit/s client signals, implies that outages of the network infrastructure can have serious economical and social consequences.

Network recovery/resilience, i.e., the capability of the networks to efficiently recover from failures, has become of vital importance. Thus, optical transport networks need to be very robust to face failures. The protection mechanisms should be designed basically with the aim to be simple, to minimize the traffic losses and to optimize the utilization of the network resources.

Survivability strategies in current transport networks are based on the pre-allocation of network resources (spare resources) to be used only to switch (route) the affected traffic in case of failures.

In legacy multi-layer networks, each layer (e.g. IP, SDH) has its own protection mechanism built in, independent from the other layers. Network recovery basically relies on the SONET/SDH network layer. Indeed, different mechanisms, based on the protection approach, have been proposed that allow fast recovery within the target of 50 ms. Nevertheless, SONET/SDH protection is mainly limited to ring topologies and it is not able to distinguish between different priorities of traffic and it has not vision of higher layer failures.

The emerging packet-based Resilient Packet Ring (RPR) technology for metropolitan networks provides powerful protection mechanisms that minimize the time needed to restore the traffic without the pre-allocation of resources.

To face to failures, the resilience single-layer strategy (a single layer has the responsibility for the recovery) is very simple from the implementation point of view. However it may not be able to efficiently recover the network from all kind of failures that can occur. Therefore, multi-layer resilience (various network layers can participate to the recovery actions) provides better performance not only in terms of protection but also in terms of resources optimization.

Multi-layer resilience strategies require coordinating the recovery mechanisms provided by each layer. In such a context, another contribution of this Ph.D. Thesis is the design and evaluation of a multi-layer resilience mechanism to be used in the IP over RPR over intelligent optical transport network for metropolitan environment to efficiently face with a wide range of network outages, while optimizing the utilization of the network resources. Its novelty relies on the interworking required between the RPR and the optical transport layer.

Finally, the fourth contribution of the Thesis deals with the optimization of the bandwidth utilization of the RPR rings taking benefits from the automatic switching of optical connections capabilities of the underlying ASON/GMPLS networks.

1 Introduction

Legacy networking architecture is based on a client layer (Layer 3) on top of the transport network. The classic telecom mapping was based on a multi-layer architecture composed by the Internet Protocol (IP), which is actually the clear dominator among layer 3 protocols, while the transport network was composed by Asynchronous Transfer Mode (ATM) layer over Synchronous Optical Networks/Synchronous Digital Hierarchy (SONET/SDH) layer over Wavelength Division Multiplexing (WDM) technology [1]. This networking architecture (Figure 1) is an effective solution in a voice-centric scenario where IP is just one of the clients of the transport network. In such a network context, IP packets, whose length is variable (from 40 bytes to 1500 bytes) are segmented into ATM cells of fixed length [1]. Then, the ATM cells are assigned to different Virtual Connections (switched through ATM core switches), packed into SONET/SDH frames according to the SONET/SDH multiplexing hierarchy, and transported through the SONET/SDH transport network (onto bi-directional circuits). The transport capacity of the fibres interconnecting the SDH Digital Cross Connects (DXCs)/Add Drop Multiplexers (ADMs) is highly increased through the utilization of the WDM technology that allows different wavelengths to be multiplexed on a single fibre [2], [3]. The resulting network is thus a multi-layer architecture requiring different equipments for each layer and where each layer has to be maintained and managed independently from the others. Moreover, it presents many drawbacks such as heavy overall overhead, partial overlapping of functions (e.g., protection, management functions) and very high costs for Network Operators.

1.1 Transport Network evolution path

The first step in the networking architecture evolution has been the elimination of the ATM layer. The fact that the SONET/SDH network was kept at layer 1 still ensured all the advantages of

SONET/SDH transport networks, among which protection, digital performance monitoring, and complete network management capabilities [1].

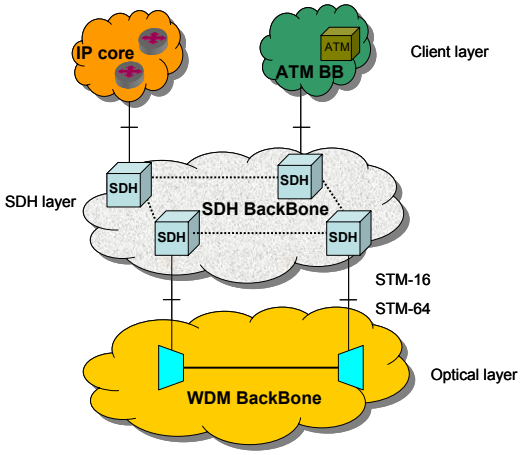


Figure 1: Legacy networking architecture

Time Division Multiplexing (TDM) legacy transport networks (e.g., SONET/SDH) have been basically designed for voice and leased line services (SONET/SDH technology was initially designed to optimize transport of 64 kb/s-based TDM services). In the past ten years many network operators have largely deployed SONET/SDH transport infrastructures in both long haul and metropolitan/regional networks.

However, today it is widely recognized that the traffic expected to be carried by the public transport networks will be progressively dominated by data, which is growing at explosive rate. This is due to the migration of many applications and services over IP and also to a progressive introduction of high-speed access technologies [4]. Indeed, emerging applications are fuelling the growing of data traffic. Some examples of such emerging applications are: high-bandwidth multi-media applications (real-time and no-real-time), shared access to remote resources, network-wide computation and data services (grid-computing, data-grid), storage networking, disaster recovery, etc. In this context, another important point of attention is that the convergence towards the IP of real-time applications (i.e. multimedia applications) imposes very strict QoS requirements such as latency, jitter, packet loss, etc.

On the other hand, traditional IP routing algorithms (i.e., Open Shortest Path First (OSPF), Interior Gateway Protocol (IGP) and Intermediate System-Intermediate System (IS-IS)) do not provide an efficient distribution of traffic load onto available network resources [5]. Indeed, when IP traffic is carried by bi-directional SONET/SDH circuits, large portions of the network bandwidth may remain under-utilized, while the bandwidth at the opposite direction can be congested [6]. IP

traffic is characterized by its typical asynchronous, burst and asymmetric nature in contrast to the traditional symmetric telephone traffic [7] and [8].

With the rapid growth and dynamicity in demand for network resources, both fast bandwidth provisioning and Traffic Engineering (TE) are key requirements for the next generation transport networks. Some of the TE objectives are efficiently mapping the actual traffic to available network resources and manage traffic as well as network capacity rapidly and effectively in response to changes in the traffic pattern. These dynamic changes in the client traffic pattern are the consequence, for example, of sudden fluctuations of the traffic due to unexpected events (e.g. disasters) or equipments failures. According to [9], TE is the process to control traffic flows in a network in order to optimize resources utilization and network performance. Practically, this means choosing routes taking into account traffic load, network state and user requirements such as QoS or bandwidth, and moving traffic from more congested paths to less congested ones.

In legacy networks, TE issues relied on the ATM technology, basically thanks to its connection-oriented nature and due to the decoupling of routing (control plane) and forwarding (data plane) functionalities [1]. Nevertheless, the IP over ATM mapping increases the network complexity since IP is ATM unaware and hence two separate control planes were needed. Thus, Internet Engineering Task Force (IETF) developed the Multi-Protocol Label Switching (MPLS) technology as a tool to provide TE features to the IP networks [10]. Nevertheless, it is recognised that TE in the IP-MPLS layer may not be sufficient in backbone networks [4].

The evolution path of current transport networks is characterized by a data-centric networking scenario (Figure 2). Such evolution will encompass different steps depending on the time scale it is considered for such evolution [11], [12] and [13].

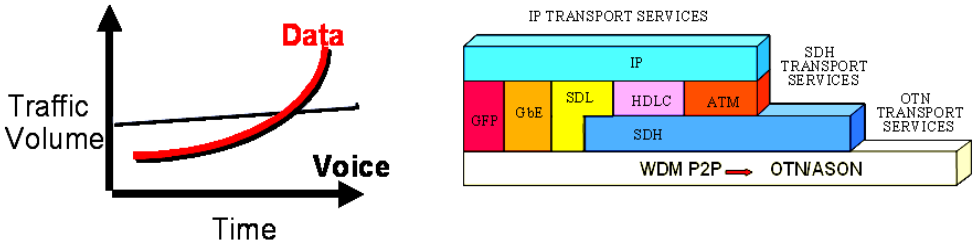


Figure 2: Evolution path towards data-optimized networks

The short-term step in the network evolution is based on the addition of new network functionalities (i.e., dynamic bandwidth allocation) to legacy SONET/SDH networks through the implementation of the Virtual Concatenation (VCAT) and Link Adjustment Capacity Scheme (LCAS) paradigms [14], [15]. The SONET/SDH payload bit rates are rigid since they were

originally designed for traditionally voice networks. VCAT tries to relax this characteristic. It consists basically on breaking the payload into individual channels (called members of the Virtual Concatenation Group, VCG), transporting each channel separately, and then recombining them into a contiguous bandwidth at the endpoint of the transmission. This functionality is needed only at the path termination equipments.

With the LCAS definition, the number of concatenated payloads (members) may be increased or decreased at any time without affecting traffic currently being sent. Moreover, LCAS will automatically decrease the capacity if a member of VCG experiences a failure in the network. Specifically, when one of the constituent channels experiences a failure, it will be automatically removed while the remaining channels are still working. Thus, the available bandwidth will be lowered but the connection will be maintained.

The second step in the network evolution path encompasses, in a short/middle-term scenario, for both metropolitan and wide area environments, the elimination of SONET/SDH network layer.

Indeed, the transport capabilities of SONET/SDH (e.g. protection and accommodation of various bit rates through tributaries) are being absorbed by the optical layer thanks to the advances of optical technology [16], and the standardization of the Optical Transport Network (OTN) [17]. An OTN is composed by a set of optical network elements (Optical Cross-Connects (OXC) and Optical Add Drop Multiplexer (OADM)) connected by optical fibre links. OTNs are able to provide the network functionality of transporting, multiplexing, routing, management, supervision and survivability of optical channels carrying client signals, avoiding the electronic process of the data at intermediate nodes. A distinguishing characteristic of the OTN is its provision of transport for any digital signal independently of client-specific aspects, i.e., it provides client layer independence.

The transport networks (both legacy SONET/SDH and incoming OTN) tend to be static, which means that connections (both SONET/SDH circuits and light paths in OTNs) are provided "manually" through the Network Management System (NMS) (i.e., permanent connections). This way of provisioning is rather time consuming, which means that weeks or even months are needed to provide high bandwidth connections.

The further step of the network evolution path, in a long-term scenario, comprises the introduction of the Automatically Switched Optical Networks (ASON) [18]. While current optical networks only provide transport capacity, the main features of an ASON are the ability to automatically increase/decrease the transmission capacity on demand (client request), i.e., setting up/tearing down optical channels automatically, and the automatic neighbouring discovery. To

provide such network functionalities, a Control Plane (CP) has to be defined. The GMPLS (Generalized Multi-Protocol Label Switching) paradigm is proposed to be the control plane for the ASON networks [19]. It is an extension of the set of protocols designed for the MPLS technology and it encompasses time-division (e.g., SONET/SDH, Plesiochronous Digital Hierarchy (PDH), G.709 [111]), wavelength, and spatial switching (e.g. incoming port or fibre to outgoing port or fibre). The implementation of a GMPLS-based control plane allows an integrated vision of the network, namely the integration between the client layer and the optical transport layer. It is recognized that such integration leads to lower networks cost and complexity [20] and [21].

The introduction of intelligence by means the CP is recognized as the enabling solution to meet the requirements, among others, of fast and flexible end-to-end bandwidth provisioning, automatic topology discovery and fast restoration.

In this context, one of the main topics to be solved is how to dynamically manage the capacity available at optical level, in order to transport the client traffic in a cost-effective way, while optimizing the utilization of the network resources, reducing both the complexity and the inefficiency of legacy transport networks.

Focusing particularly in the metropolitan network context, many networks use a physical ring structure. This is a natural environment for the SONET/SDH networks that constitute the bulk of current metropolitan network infrastructure. As mentioned above, SONET/SDH, however, was designed for point-to-point circuit-switched services. Alternatively, Ethernet technology offers a simpler and inexpensive solution for data traffic. However, because Ethernet is optimized for point-to-point or meshed topologies, its use of the available bandwidth is inefficient and it does not take advantage of the ring topology in order to implement a fast protection mechanism [22].

The emerging Resilient Packet Ring (RPR) Layer 2 technology, recently standardized by the Institute of Electrical and Electronics Engineering (IEEE) as IEEE 802.17 RPR, fills this gap by being a multi-service transport protocol based on packets rather than circuits [23]. RPR systems are seen by many carriers as the inevitable successors to SONET/SDH ADM-based rings in metropolitan networks, allowing moreover higher revenues expectations [24]. This is due to the fact that RPR networks may provide performance-monitoring features similar to those of SONET/SDH and, at the same time, maintain Ethernet's advantages, i.e., low equipment cost, high bandwidth granularity and statistical multiplexing capability. Furthermore, the RPR may run over the fibre (dark fibre or WDM).

For carriers, RPR promises the delivering of all the necessary end-user services, such as TDM voice, Virtual Private Networks (VPN) data and Internet access, at dramatically lower equipment, facility and operating costs.

The optimization of the utilization of network resources is very critical when failures occur in the network (as a consequence of the need of rerouting the affected traffic). Indeed, network survivability/resilience/recovery, namely the capability of the network to recover traffic affected by failures, has become of vital importance in current networks and next generation networks will need to be very robust to face failures [25], [26]. Network Operators, therefore, have to take special precautions in order to prevent this, which means doing their networks survivable. As it is difficult to prevent failures in the network infrastructure (equipment failures, cable breaks, etc.) the objective is to maintain service availability even under failure conditions. In order to make the networks more reliable the networks has to be reconfigurable. This reconfiguration has to be fast and optimizing the utilization of the network resources. Finally, it has not to increase too much the cost of the network.

The protection mechanisms implemented in RPR are fast. In fact, they aim to achieve recovery times of about 50 ms and to protect against any single failure in the ring. No bandwidth is dedicated for recovery purposes and, therefore, in a failureless state the resource utilization is very high.

On the other side, the achievements in the optical layer thanks to the standardization of the OTN and the ASON/GMPLS paradigm allow recovery capabilities directly in the optical layer.

To face to failures, the resilience single-layer strategy (a single layer has the responsibility for the recovery) is very simple from the implementation point of view. However it may not be able to efficiently recover the network from all kind of failures that can occur. Therefore, multi-layer resilience (various network layers can participate to the recovery actions) provides better performance not only in terms of protection but also in terms of resources optimization. However, it requires coordinating the recovery mechanisms provided by each layer.

In metropolitan IP over RPR over intelligent optical layer, to improve the overall network resources utilization in case of failures, efficient strategies to coordinate the recovery mechanism implemented at RPR layer and at the OTN layer have to be defined. The aim is to provide not only survivability but also overall network resources optimisation, which means meeting TE objectives.

1.2 Problems addressed in this Ph.D. Thesis

This Ph.D. Thesis deals with the design of multi-layer traffic engineering (MTE) strategies both for metro and wide area networks. Specifically, two MTE procedures have been designed and evaluated, namely one for IP/MPLS over ASON/GMPLS network scenario to face with the dynamic traffic fluctuations at the client network, and another for metropolitan IP over RPR over intelligent optical transport networks scenario to face with the efficient recovery from failures.

The optimization of network resources means the efficient transport of client network traffic on the optical transport network. The aim is to avoid network congestion due to the unexpected events, such as traffic variations or network failure, as well as optimizing the optical resources bandwidth utilization.

In the case of the Wide Area Networks environment, we concentrate on transporting IP/MPLS over intelligent optical transport networks. The automatic set up/tear down of optical connections introduced by the definition of ASON/GMPLS paradigm allows the design of efficient MTE to cope with dynamic bandwidth demands. Specifically, we designed the TRIDENT procedure.

Moreover, we designed a practical approach for ASON networks dimensioning purposes based on the approximate characterization of the traffic arrival process, through its mean and the peakedness factor.

In the case of the metropolitan networks, we concentrate on the design and evaluation of interworking mechanisms for a short-term network solution in the evolution path, which implies to suppress the SONET/SDH layer and considering the very promising solution represented by Resilient Packet Ring technology. Given the necessity to make the networks reliable, a multi-layer resilience strategy in the IP over RPR over optical transport networks to recover efficiently from a vast range of possible networks outages is designed and evaluated. Its novelty relies on the interworking required between the RPR and the optical transport layer.

The optimization of the bandwidth utilization of the RPR rings taking benefits from the automatic switching of optical connections capabilities of the underlying ASON/GMPLS networks has also been addressed.

It has to be underlined that the work here presented has been part of the research activities performed by the Advanced Broadband Communication Centre (CCABA) of the Universitat

Politècnica de Catalunya. In particular, the work was carried out within the framework of the international research project IST-1999-11387 “Layers Interworking in Optical Networks (LION)” funded by the European Commission and within two national projects, namely “Evaluación de arquitecturas de conmutación de paquetes para redes ópticas (ECOPAQ)” funded by the Spanish Ministry of Education under contracts TIC99-0572-C02-02 and “Transporte de tráfico IP sobre redes ópticas: diseño y evaluación (TRIPODE)” funded by MCyT (Spanish Ministry of Science and Technology) under contracts FEDER-TIC2002-04344-C02-02.

This Ph.D. Thesis is organized in three parts. The first one is devoted to the capacity management problem in next generation networks. It includes the Chapters 2, 3 and 4. Specifically, Chapter 2 focuses on the multi-layer TE in IP/MPLS over ASON/GMPLS presenting the considered scenario and the related works. Chapter 3 describes the procedure suggested on this topic. Specifically, it presents the TRIDENT procedure highlighting its characteristics and merits. Chapter 4 deals with the traffic modelling for the dimensioning of the ASON networks. In particular, it presents the investigations about the suitability of classical teletraffic models for ASON dimensioning purposes.

The second part deals with the design of a multi-layer resilience strategy able not only to efficiently react to failures but also to optimize the network resources utilisation in case of failures. It includes Chapter 5 and Chapter 6. Specifically, Chapter 5 discusses the related work on the multi-layer recovery approach while Chapter 6 discusses on one hand the strengths and weakness of the RPR technology and on the other hand, it presents the MTE strategies designed to be used in IP over RPR over intelligent optical networks for metropolitan environments.

Chapter 7 draws some conclusions while Chapter 8 discusses some possible future works which arise from the different contributions.

Finally, the Ph.D. Thesis includes a third part, including three Chapters. Chapter 9 presents in detail the TRIDENT procedure and Chapter 10 presents the experimental implementation of the TRIDENT procedure in a real environment, namely in the ASON-GMPLS testbed developed at the Telecom Italia Lab premises. In particular, the feasibility of the procedure is evaluated and some experimental results are depicted and discussed. Finally, Chapter 11 describes the generalization of the TRIDENT procedure to dynamic SDH-based networks, namely legacy SDH networks improved to better meet the current client’s requirements by the application of the VCAT and LCAS functionalities.

PART I: Multi-layer Traffic Engineering (TE) in ASON Networks

This part of the Ph.D. Thesis deals with the capacity management issues in next generation transport networks. Firstly the problem is formulated, some related works are discussed and finally we present TRIDENT, a multi-layer traffic engineering procedure designed to efficiently track the fluctuations of the IP/MPLS traffic, while optimizing the available network resources at the optical layer. Specifically, its characteristics and merits are discussed and evaluated by simulation.

This part is concluded by a practical contribution of this Ph.D. Thesis which deals with the suitability of classical teletraffic models for ASON dimensioning purposes.

2 Multi-layer TE in IP/MPLS over ASON/GMPLS networks

Next generation transport network infrastructures have to cope with the growing demand for bandwidth generated by the client layer as well as have to be multi-service, namely able to support several traffic classes with different requirements in terms of Quality of Service (QoS) [27].

It is widely recognized that the traffic in evolutionary transport networks will be progressively dominated by data. As a matter of fact, the statistical characteristics of data traffic are rather different from those of traditional voice traffic, for which legacy TDM-based transport networks have been designed and optimized.

In fact, firstly IP traffic is characterized by its asymmetry and secondly it is highly dynamic; it is not as easily predictable and stable as the voice traffic, and it shows predictable and unpredictable surges/peaks as well. Such surges are caused by unexpected events such as users' behaviour, weather conditions, accidents, faults, etc., which cause significant and unexpected fluctuations over time (e.g. on a daily basis) of the aggregated data traffic to be transported by the telecommunication networks.

With such data traffic nature, a very simple dimensioning approach relies on the bandwidth over-provisioning, namely over dimensioning the network resources in order to take into account the peaks of the traffic that has to be carried. Nevertheless, such approach does not represent a cost-effective solution since Network Operators want to reduce both the infrastructure Capital Expenditure (CAPEX) and Operational Expenditure (OPEX).

Alternatively, the data traffic peaks/burst can be handled by the provisioning of bandwidth through the Network Management System (NMS), using a bandwidth scheduled approach [28]. In such a case, the bandwidth is allocated for example on the basis of the time of the day. The

easiest way but, at the same time the most inefficient, to implement this approach is over dimension the entire network and modify the traffic policy done at the client router depending on the time. Another possibility to make the system more efficient, from a bandwidth point of view, is to trigger, as we point out in the next Section, User Network Interface (UNI) set up signalling from the client or server router depending on the time to allocate more bandwidth.

However, when IP traffic has to be carried, although the periodic nature of the traffic pattern (similar periodic pattern can be observed on network links during the same periods, See Figure 3 and Figure 4), the data traffic volume to be transported is not predictable since it is difficult to know how huge the surges are. As an example, if we compare the traffic monitored from the Catalan Research &Academic Network [29] for two different working weeks, the traffic volume is not easily predictable.

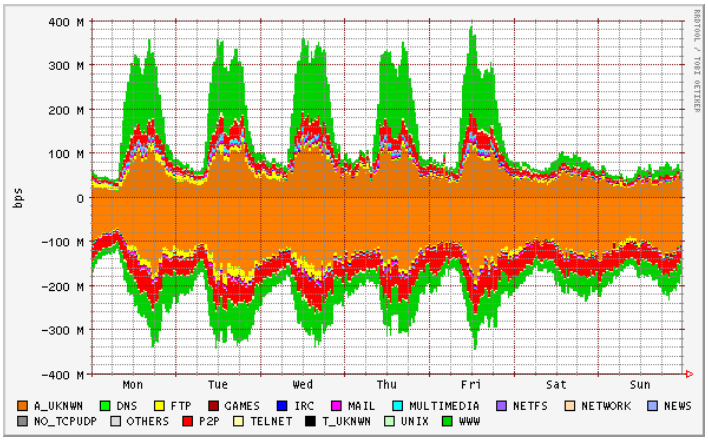


Figure 3: Weekly incoming/outgoing traffic from the Catalan R&A Network, November 2004, 8-15

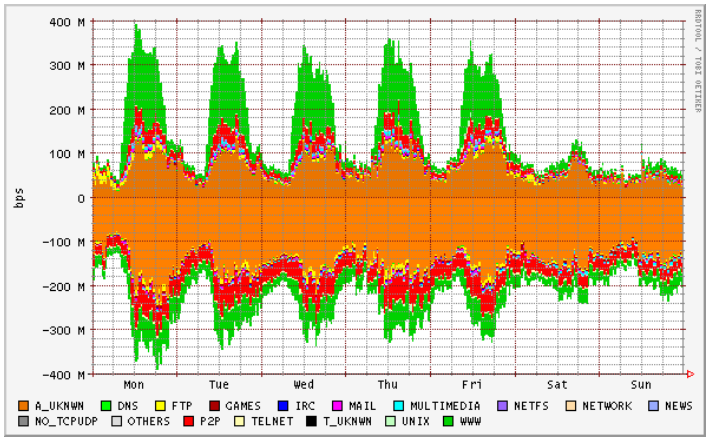


Figure 4: Weekly incoming/outgoing traffic from the Catalan R&A Network, October 2004, 25-31

Moreover, such approach requires detailed investigations about the behaviour of the different clients of the transport network.

As a consequence, the basic requirements for next-generation optical transport networks are the flexibility and the ability to dynamically react to traffic demand changes, allowing at the same time the optimization of the network resources [4], [30].

2.1 Capacity Management for network resources optimization

In order to provide TE capabilities in the Internet network context, the Internet Engineering Task Force (IETF) introduced MPLS technology [10], and the Constraint-Based Routing (CBR) feature [31].

MPLS decouples routing and signalling (control plane) and forwarding (data plane). Such technology is based on the encapsulation of IP packets into labelled packets that are forwarded in a MPLS domain along virtual connections called Label Switched Paths (LSPs). MPLS-enabled routers are called label switched routers (LSRs). Each LSP can be set up at the ingress LSR by means of ordered control before packet forwarding.

Constraint-based routing means that when a route is calculated (for example for a MPLS LSP) not only network topology but also user's requirements have to be taken into account [27]. An LSP can be forced to follow a route that is calculated a priori thanks to the explicit routing function. Then, the network resources are reserved on the specific path by means of suitable signalling protocols (e.g., Resource Reservation Protocol (RSVP) [32] and Label Distribution Protocol (LDP) [33]). Moreover, each LSP can be set up, torn down, rerouted if needed, and modified by means of the variation of some of its attributes. Furthermore, pre-emption mechanisms on LSPs can also be used in order to favour higher priority data flows at the expense of lower priority ones, while avoiding congestion in the network.

MPLS is IP aware and introduces a single IP control plane and thus the network scalability is easier with respect to the IP over ATM architecture. However, it is recognised that MPLS is not sufficient to provide efficient mechanisms for TE in the network backbone segments [4].

The efficient management and control of the growth of bandwidth demands from data traffic and the need of the Network Operators to reduce investments and operative costs (CAPEX and OPEX) represent two of the major drivers moving the evolution of current transport networks, both in the long-haul and metro segments [34], [35].

The introduction of intelligence in the transport networks (i.e. the ability to set up/tear down high bandwidth connections dynamically), is considered a promising solution to meet the above mentioned emerging requirements. This new functionality is obtained by the implementation of a distributed control plane (CP), which consists of a set of functionalities such as signalling and routing distributed throughout the network.

2.1.1 Automatically Switched Optical Network (ASON)

The standardization of the Automatically Switched Optical Network (ASON) [18], which is able to provide dynamically optical bandwidth (i.e. light paths) on the basis of the client layer requests, is recognized as the enabling solution to meet the requirement for fast and flexible end-to-end bandwidth provisioning. An ASON is an optical transport network supporting a Transport Plane, a Control Plane and a Management Plane (Figure 5). The Transport Plane provides bi-directional or unidirectional information flows transfer between users while the Management Plane is responsible for fault, performance, configuration, accounting, and security management functions both for the transport and control planes. ASON networks provide leased optical lines (Permanent Connections established through the NMS) and other two innovative transport services: Soft Permanent Connections (SPC) and Switched Optical Connections (SC).

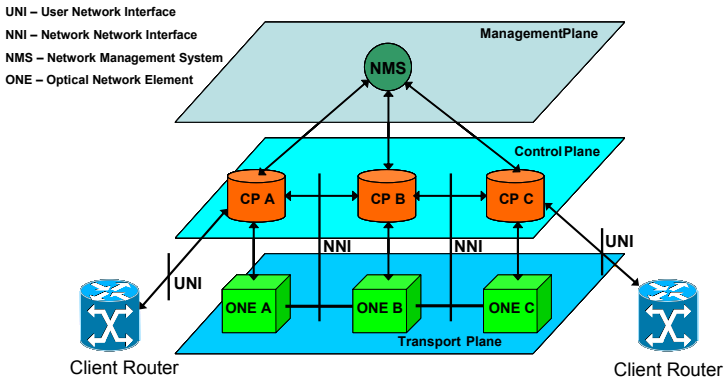


Figure 5: ASON Reference network model

The SPC (Figure 6 (a)) is requested by the management plane, which uses network generated signalling and routing protocols via the Network to Network Interface (NNI) to establish the connections, while the switched service (Figure 6 (b)) is requested directly by the customers via UNI signalling; then the connections are set-up using NNI signalling and routing protocols. A distributed GMPLS-based Control Plane (CP) has to be implemented to achieve the above automatic switching network features. Basically, the CP consists of the set of routing and signalling

protocols needed to set up or release switched connections according to users' requests or restore connections in case of failures.

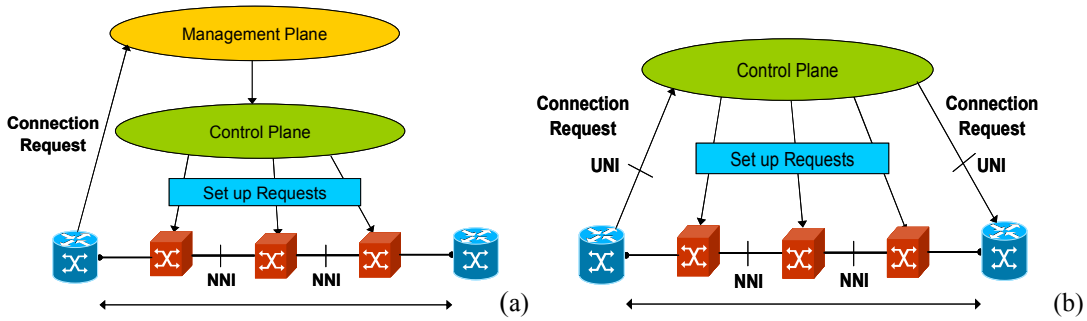


Figure 6: ASON transport services, (a) soft-permanent connection, (b) switched connection

Basically the Internet Engineering Task Force (IETF) [36] and Optical Internetworking Forum (OIF) [37] have been defining the generalization of MPLS control plane (Generalized MPLS) to control the optical components (OXCs and OADMs) and thus being viable the implementation of the CP functions for ASON networks.

Basically, the CP is in charge of: 1) Fast and efficient configuration of connections within the transport layer network to support both switched and soft permanent connections, 2) Reconfiguration or modification of connections that support calls that have previously been set up and 3) Performing restoration functions directly in the optical layer in case of failures.

The control plane defined for ASON consists of components with diverse functionalities. Interactions of these components and the information needed for a communication between components are done via interfaces (e.g., UNI and NNI). The main components of the ASON CP are (Figure 7) [18]:

- **Connection Controller (CC)**: it is the responsible for the coordination among all the control plane components for the purpose of the management and supervision of connection set ups, releases and the modification of connection parameters for already established connections.
- **Routing Controller (RC)**: it responds to requests from connection controllers for path information needed to set up connections and respond to requests for topology information for network management purposes.
- **Link Resource Manager (LRM):** it is responsible for the management of connections among ports, lambdas, of the transport plane elements. Moreover it provides information about the physical characteristics of the nodes to the CC.

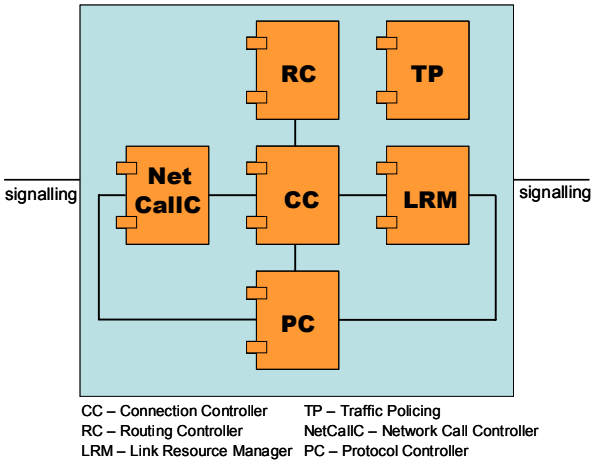


Figure 7: ASON Control Plane

For the communication among the control plane components of the network nodes a "signalling network" is required. Such signalling network is called Data Communication Network (DCN) and among its requirements, it is important to highlight the reliability and the transmission capacity. The former relies on the ability to recover from failures and the latter is essential to avoid high propagation delays with consequent high set up times. The set up time is strongly related for example to the recovery time in case of failures recovered at the optical layer (e.g., fast restoration).

The implementation of a control channel (i.e., unidirectional channel between two adjacent nodes of the signalling network) can be obtained as follows [38]:

- **Out-of-fibre/Out-of-band**: the DCN is implemented with an independent network such as Ethernet or IP networks.
- **In-fibre/Out-of-Band**: the DCN is implemented by using one among the WDM channels, which is dedicated to the signalling.
- **In-Fibre/In-Band**: the signalling information is transmitted using for example the overhead of the SONET/SDH frames.

Providing intelligence to the optical layer (automatic provisioning of light paths on client demands) opens the possibility to define multi-layer traffic engineering (MTE) strategies, and offers opportunities to network operators for cost savings (both CAPEX and OPEX) [34], [35]. The implementation of ASON/GMPLS networks also opens the possibility for providing Bandwidth on Demand (BoD) services.

The following Figure 8 illustrates an exemplary network scenario in which the functionalities provided by the ASON definition can be used [39].

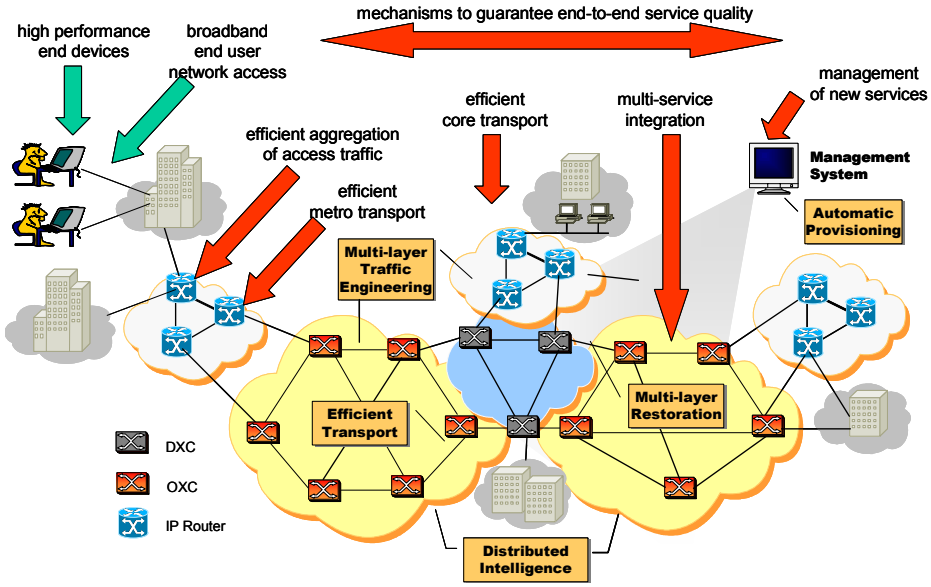


Figure 8: Reference network scenario

What ASON switched connections are needed for?

A well-designed network architecture should give to Network Operators and to Services Providers (SP) the control of their network, while providing fast and reliable connection set up. It is expected to be sufficiently generic to support different technologies, differing business needs and different distribution of functions by vendors.

The main feature of an ASON is the ability to increase/decrease the transmission capacity on demand, i.e. set up/tear down optical channels automatically (switched connections) or at least very dynamically via the OSS (soft permanent connections). This feature may reply at least to two needs:

- To cope with dynamic fluctuations of the client traffic.
- To recovery very dynamically from failures (e.g., optical fast-reroute).

The purpose of this Subsection is to show some general scenarios with clear indication who and in what cases is likely to request an automatic switched optical channel. Figure 9 and Figure 10 show two different situations where the new ASON switched connections are required [40].

Specifically, in Figure 9, two sub-networks of an individual ISP are interconnected with a permanent optical connection. Both the expected and the unexpected increases of the traffic carried between the sub-networks results in demands for switched optical channels.

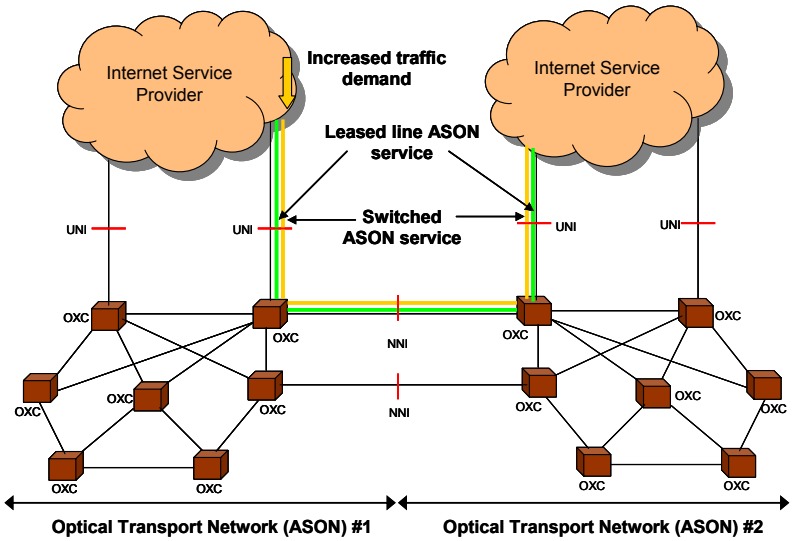


Figure 9: Switched service requested by ISP to support extra traffic

Figure 10 illustrates the other considered example. Two sub-networks of an individual ISP are interconnected with permanent optical connections. In this case, it is supposed that there are not significant changes in the traffic pattern between ISP sub-networks.

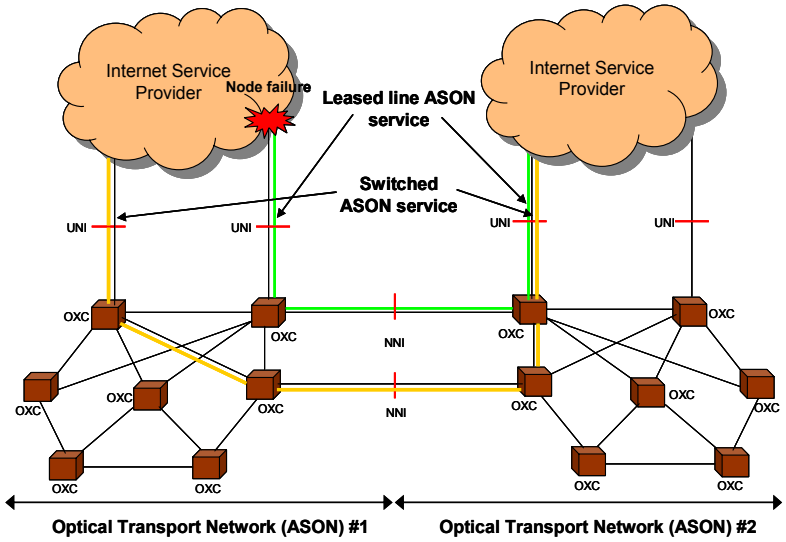


Figure 10: Switched service triggered by node failure

The node failure (link failure) triggers the optical restoration, and thus the Control Plane is requested to set up a switched optical connection to be used to recover the traffic from the failure.

When intelligent optical transport networks are taken into account, both the traffic engineering (TE) and the capacity management (CM) approaches can be considered to face with the dynamic IP traffic bandwidth demands [41].

On one hand, the TE approach ensures the maximization of the network performances both in normal and in failure conditions, trying to move "*the traffic where the bandwidth is*". As an example, it is based on centralized and/or distributed routing algorithms such as dynamic routing (e.g., time-dependent or state-dependent routing path selection) or QoS routing algorithms.

On the other hand, the CM approach provides connections to meet traffic demands minimizing the size of the network, and therefore reducing both the operational and the capital costs. It is specifically based on moving "*the bandwidth where the traffic is*".

This part of the Ph.D. Thesis deals with dynamic capacity (light paths) management procedure aiming at the optimization of the utilization of the network resources.

2.1.2 Capacity Management: Related Work

The provisioning of the connections/light paths needed to cope with a given traffic demand, is defined as the virtual topology design, which represents the set of all the light paths established in the optical network. The design of the logical topology is a very important aspect since it may lead to optimize the use of the network resources for a given traffic demand.

As an example, the virtual topology is employed by Internet Service Providers (ISPs) to efficiently connect their end equipments such as IP routers by leasing optical bandwidth from network operators who own the optical devices (i.e., fibres and optical cross-connects) [42].

Two approaches have been considered for the logical topology design, namely the off-line and the on-line approaches. The off-line approach aims at the optimization of the logical network topology according to a given and known traffic demand, while the on-line approach is applied when the future traffic demand is not known a priori (i.e., when having an unexpected traffic demand).

Concerning the off-line approach, many studies present different procedures for the optimal design of the virtual topology in a static wavelength routed scenario (no automatic bandwidth provisioning capability) and a given traffic demand [43], [44] and [45]. Moreover, a survey of such logical topology design algorithms can be found in [46].

However, the data traffic fluctuates over time, and a virtual topology optimized for a given traffic profile is not able to efficiently manage with the dynamic traffic demands. In such cases, a

reconfiguration of the logical topology is needed. Different reconfiguration procedures in static wavelength routed networks, able to react to traffic changes have been proposed in [47], [48]. In these studies, the reconfiguration of the logical topology is carried out in two steps. The first step encompasses the design of the virtual topology for the new traffic demands while the second step includes the transition from the old topology to the new one. However, the dynamics of the traffic demand are assumed to be known.

In the case of the on-line network reconfiguration, it requires two complementary steps, namely a first one consisting on deciding when a reconfiguration has to be triggered and a second one consisting on to manage the reconfiguration itself (i.e., setting up/tearing down of LSPs and/or light paths). Next, some related works referring basically to the latter one are discussed, while this part of the Thesis refers to the former issue.

In [4], an on-line algorithm to be used by ISPs to reconfigure MPLS networks is suggested. It proposes a traffic engineering system able to dynamically react to traffic changes while at the same time fulfilling QoS requirements for different classes of service. The solution consists of a hybrid routing approach, based on both off-line and on-line methods, and a bandwidth management system that handles priority, pre-emption mechanisms, and traffic rerouting in order to concurrently accommodate the largest amount of traffic and fulfil QoS requirements. More specifically, the TE system invokes an off-line procedure to achieve global optimization of path calculation, according to the expected traffic matrix, while invoking an online routing procedure to dynamically accommodate the actual traffic requests (LSPs establishment requests), which allows to react to traffic changes. This solution allows ISPs to efficiently manage different class of traffic in their MPLS networks. Therefore, in this work, the reconfiguration of MPLS networks was taken into account.

In [42], the authors propose an on-line centralized approach to be used by ISPs for the virtual-topology reconfiguration of a static WDM wide-area mesh network under dynamic traffic demand. The key idea of the approach is to adapt the underlying optical connectivity by periodically measuring the actual traffic load in the light paths, and react to the load imbalances caused by the traffic fluctuations, by either adding or deleting one or more light paths at a time. When a load imbalance occurs, it is fixed either by tearing down a light path that is lightly loaded or by setting up a new light path when congestion occurs. Once the reconfiguration is done, the rerouting of the client traffic over the new logical network topology is required, which means running the routing and signalling protocols to solve the routing and wavelength assignment (RWA) problem.

In such a way, ISPs can optimize the operational cost of their virtual topology by leasing only the appropriate amount of light paths to transport the actual traffic.

The above mentioned studies ([4] and [42]) propose solutions, which only concern with the reconfiguration of MPLS networks and static WDM networks and both of them provide to ISPs with procedure to efficiently respond to dynamic traffic demands. In other words, these studies do not take into account interworking features between the client and the transport layer. The automatic bandwidth provisioning capability provided by the new ASON/GMPLS paradigm is neither taken into account.

Providing intelligence to the optical layer opens the possibility to define multi-layer traffic engineering (MTE) strategies offering in such a way opportunities to network operators for cost savings, as well as the provisioning of new emerging services such as the Bandwidth on Demand (BoD).

The ASON/GMPLS paradigm is based on an integrated vision of the network, in which the different layers can interwork among them. This is recognized to provide better overall network performance as well as reduction of OPEX and CAPEX [34]. In such context, the application of the concepts of multi-layer traffic engineering (MTE) [49] or Integrated Traffic Engineering (ITE) concept [50] leads to significant improvements.

Different studies define multi-layer traffic engineering policies for IP over ASON scenarios. In such studies, TE actions are carried out involving both the IP (MPLS) layer and the optical layer. The aim of these on-line mechanisms is the reconfiguration of the network topology in order to optimize the network resources responding efficiently to dynamic traffic demands. These studies can be found in [30], [49] and [51].

In [30], the author defines a centralized integrated traffic engineering method based on the traffic routing stability. Incoming data traffic (i.e., LSPs) is classified as high priority, which can tolerate limited rerouting, and low priority, which can be rerouted after a timer has expired. Specifically, it consists of the design of a mechanism that reacts to traffic variations in MPLS over ASON/GMPLS networks. Such traffic variations are due to dynamic requests for LSPs establishments. The LSP are characterised by a fixed bandwidth requirement (10 Mbps), which means that it assumed that the traffic load supported by the light paths carrying the LSPs varies only due to the set up/release of LSPs. The suggested mechanism reacts to new high priority (HP) and low priority (LP) MPLS LSP requests accommodating them on light paths according to a routing-stability constraint. When a new LSP request cannot be accommodate on existing light paths (the

required bandwidth does not match with the available light path bandwidth), the set up of a new light path is requested. A traffic monitoring is considered, which means that the number of LSPs allocated in each light path has to be advertised to the network nodes.

In [51], the authors define an on-line virtual topology adaptation procedure based on the actual traffic load, in MPLS over GMPLS networks. A network architecture where different MPLS networks (for different traffic classes) are built over the optical network is considered. Therefore each light path will be assigned to LSPs carrying an aggregation of traffic flows of the same traffic classes. An optimal routing policy is designed to set up and tear down LSPs and light paths (λSP) in response to new traffic demands. The aim is to optimise the accommodation of the bandwidth requests minimizing the costs involving bandwidth, switching and signalling.

In [49] a multi-layer traffic engineering strategy is defined to be used in IP/MPLS over ASON networks. The strategy is based on the dynamic reconfiguration of the logical topology of the network. Congestion experimented at IP layer can be solved by the reconfiguration of the logical topology of the network automatically setting up and tearing down switched connections in the optical layer.

All the works discussed take into account the network resource optimisation from the Internet Service Providers perspective, namely how the ISPs can reconfigure the logical topology connecting their IP routers by leasing the appropriate number of light paths from the network operators. Moreover, these MTE mechanisms react to the bandwidth requests generated from the client layer.

2.1.3 Problem addressed

The approach presented here is rather different. In this part of the Ph.D. Thesis, it is defined a procedure for the dynamic management and control of the available light paths in order to keep limited the size (in terms of number of resources) of the optical transport network. The aim is to provide, at any moment, the bandwidth required to transport through the ASON, the MPLS-LSPs already established at the client layer, avoiding congestions and under-utilisation of the light paths.

We consider that the aggregated traffic load carried by the light path fluctuates over time due to many reasons, such as the actual bandwidth (not the nominal) required by already established LSPs, the request for the establishment of new LSPs and the actual increase/decrease of the traffic carried by LSPs without a predefined maximum bandwidth.

Our objective consists on the design of a proper capacity management/traffic engineering procedure to trigger the requests for the automatic set up and tear down of switched connections to adapt the bandwidth available at the transport layer to carry the aggregated client traffic. We assume that the light paths set up/tear down is in charge of the GMPLS routing and signalling protocols and therefore, routing and signalling issues are out of the scope of this Thesis.

This TE strategy is to be applied at the edge routers of the optical transport networks which collect the data traffic from the different client networks (i.e. Internet Services Providers) and provides to the Network Operators, who own the transport network and the access router, with a cost-effective management of the transport network resources. At the same time, it provides the client networks with a new transport service such as Bandwidth on Demand (BoD).

Specifically, Next Chapter is devoted to describe this procedure which can be classified as a multi-layer traffic engineering approach to track the fluctuations of the client traffic.

3 Capacity management in IP/MPLS over ASON/GMPLS networks

Currently there is an increasing interest of vendors, operators, *fora* and different international research projects around ASON, namely around the introduction of automatic switching capabilities in the optical transport networks. However, up to now there is not a market model which states, for example, potential customers of ASONs, potential statistics of the requests of switched connections, nor the real advantages of ASON implementation for Network Operators. Definitely, a traffic demand model for ASON is not yet defined which leaves some open issues such as:

1. The design of a proper procedure to automatically trigger demands/requests for setting up/tearing down switched connections to adapt the bandwidth available at transport layer according to the incoming client traffic fluctuations.

2. The characterization of the statistical distribution of the connection demands for the most significant clients of ASON. In particular, we refer to the Holding Time (HT) and the InterArrival Time (IAT), where the IAT is the time elapsing between two consecutive requests for the set up of switched connections while the HT is the time elapsing between the seizure and the release of a connection.

In this part of the Ph.D. Thesis, we concentrated on IP networks as clients of the ASON transport networks. In particular, in this Chapter we define TRIDENT, a procedure for the efficient capacity management for IP/MPLS over ASON/GMPLS networks, describing its characteristics and merits. TRIDENT is a procedure to be used by the Network Operators to dynamically manage the bandwidth available at the optical transport layer in order to track the fluctuations of the client

networks, in an IP/MPLS over ASON/GMPLS environment. It consists on avoiding both the over- and the under-utilization of the transport network resources (and aiming at keeping limited its size) while coping with the dynamic fluctuations of the incoming client traffic. Moreover, it is a way of implementation of a new *telco* service such as Bandwidth on Demand (BoD). However, it is worth noting that there are significant ASONs clients different from the IP clients. Therefore, in the next Section we discuss, as related work, the research activity done within the framework of the IST-1999-11387 Layers Interworking in Optical Networks (LION) project funded by the European Union, by the University of Science and Technology (AGH) [40], [52], in which the authors identified some potential customers of ASON transport services (apart from the IP clients). Moreover, they investigated the switched connection demands imposed on ASON networks by such kind of clients, focusing specifically on the HT and the IAT statistics.

3.1 Potential customers of ASON services: Related Work

This Section describes some of the potential applications, which users could benefit from the introduction of automatically switched transport services. Specifically, we discuss the banking sector, the video service delivering and the health care service.

3.1.1 Banking Sector

The most promising candidate for being ASON customer is the banking sector. Currently, banks are very important clients of telecommunications operators. They are using permanent connections or virtual private networks (VPN) for their purposes e.g., online data exchange, management and backup. They will probably maintain at least low capacity protected permanent connections but the use of high capacity switched connections would be more convenient and economical for them (especially in case of backup). Huge banks could be the first to replace traditional leased lines solutions with switched connections.

A typical structure of a bank is shown in Figure 11. It usually consists of bank headquarters (BH), dozens of branch offices (BO) and hundreds of local offices and automated teller machines. The whole system is guard against severe disasters by at least one backup centre (BC).

To determine possible demands the banking sector may impose on ASON networks, the authors made some assumptions basing on non-confidential data about Polish banks from [53], [54] and Polish official statistics [55]. They assumed that there are 70 banks in a hypothetical European country. All of them have headquarters in the capital and 35 banks have backup centres in the

capital. Each bank has on average 4 branch offices in the capital and 34 in the country. It holds 420,000 customers' accounts of average size of about 250 KB. On the above assumptions, a data volume between BH and BC as well as BH and BO (105 GB and 3.1 GB respectively) was made. This, in turn, led to estimate the Holding Time and Inter Arrival Time values, which can be found in the upper part of Table 1.

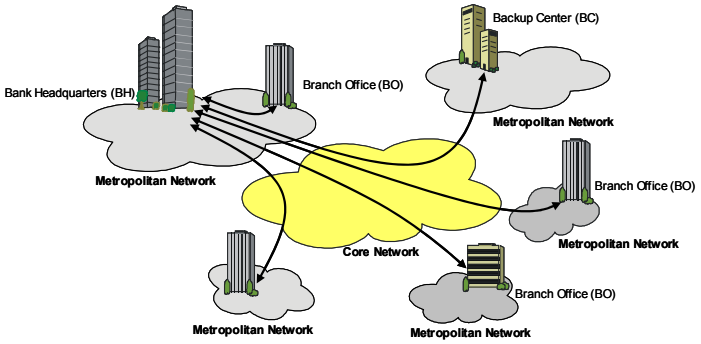


Figure 11: Structure of a typical bank

Data are sent once per month in Bank Headquarter-Backup Centre (BH-BC) case. For BH-BO connections data are sent once per day (only data that were modified during the day). Values in the Table 1 refer to 3 hours after Branch closing.

In terms of requirements, the banking sector will probably not have strong requirements on blocking probability and may accept a delay in connection set up.

3.1.2 Video delivering service

Switched connection services may be used for delivering not only data but also other kind of content (e.g. voice, video) to the customers. The chain of digital cinemas is a good example of potential customer of these services.

For years celluloid has been used as the medium for recording, storing and projecting images. Nowadays, computer workstations and high-resolution electronic video projectors are replacing conventional 35 mm film and projectors. Such system is called Digital Cinema. Digital Cinema is a system capable of delivering full motion pictures, and other audio/visual cinema-quality programs to theatres throughout the world using digital technology. Digital Cinema systems is used to distribute motion pictures which have been digitised and delivered to theatres using either physical media distribution (e.g. DVD) or digital transmission methods like fibre cable. At each showing, the digitised information is retrieved from the local storage and displayed using cinema-quality projectors. Digital Cinema ensures high quality motion picture at every showing without film

degradation. Digital projector may be used for pay per view events, interactive entertainment, corporate presentations, e-commerce, distance learning and video games. Moreover, distribution of digital releases may permit studios to rapidly change distributing pattern accordingly to the market demand. Additionally, Digital Cinema systems may hamper piracy problem by worldwide movies exhibition before illegal copies become available on the market.

According to [56], 10 Digital Cinemas are placed in Europe (about 50 in the world [57]). Although, a few Digital Cinemas exist today in Europe, it may be assumed that in the near future the number of such theatres will rise. However, it is hard to estimate how many digital cinemas will operate. The number of such cinemas may vary with population, GNP and preferences of particular population.

To determine possible demands imposed on ASON by the deployment of Digital Cinemas, the authors assumed that 40 multiplex cinemas would operate in a hypothetical country. Multiplexes seem to be main client of companies distributing digital movies in the beginning. In the capital city, there will be a specialized server storing digital movies as well as 5 multiplexes will operate. The number of new movies at each multi-screen cinema depends on number of screens, popularity of already projected films, etc. But it may be assumed that, on average, every two days a new film is introduced. Moreover, it may be presumed that there is a high probability of downloading movies during eight-hour workday. It may be assumed that average size of movie file is about 80 GB for two-hour movies (on the basis of specification of equipment offered by [58]). The exact size varies depending on content, source quality, frame rate, etc. The above assumptions lead to the values of mean Inter-Arrival Time and mean Holding Time included in the central row of Table 1.

3.1.3 Health care service

Another possible client of the ASON is a health care sector. After the analysis of the sector structure and various possible tele-medicine services the need for ASON services emerges in case of tele-consulting and tele-surgery.

Some surgeries are very difficult and complicated and must be leaded by doctors of high specialization. The best specialists are usually concentrated in only a few hospitals or health centres in the country, sometimes remote from the hospital where the patient is. Transportation is expensive and takes a lot of time while in some case even single minutes decide about human life. In many cases such a surgery could be remote-controlled if appropriate tools were available. The possible solution is establishing specialist centres serving for all hospitals in the country according to the

structure such as depicted in Figure 12. Since the strong centres where the best specialists are employed currently exist (e.g. in Poland) and remote surgeries occur, this idea seems to be very probable.

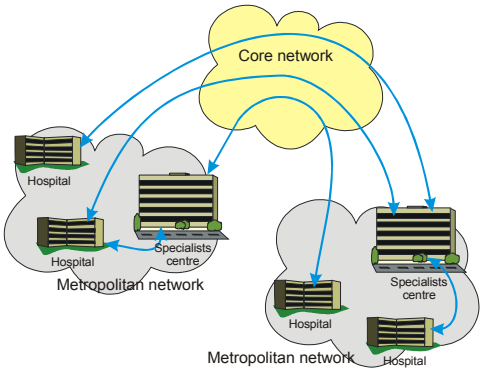


Figure 12: Structure of health care sector consisted of hospitals and specialists centres

During a tele-surgery a lot of different information must be sent between the remote centres. It encompasses one or more high-resolution video streams and on-line data from medical equipment. Hence, the need for high capacity connection emerges. Maintaining permanent connections between all hospitals and specialists centres would be too expensive. Instead, switched connections could be used.

It should be noted that cutting off the channel during the surgery might have dangerous consequences. This causes a strong demand for connection reliability (i.e. "zero" probability of connection blocking and interrupting). On the basis of Polish official statistics [55] the authors made the following assumptions, that is, there are 700 hospitals in the hypothetical European country; about 25 hospitals are located in each main city; 15 specialist centres in the country; up to 4 in some main cities. Moreover, it is assumed that: duration of a typical surgery varies from 1 to 9 hours; the required connection capacity is 155Mbit/s (one or more high-resolution video streams and on-line data from medical equipment); a hospital requires a service once per 3 days. On the above assumptions, the minimum and the maximum value for the Holding Time and the Inter-Arrival Time was obtained. These values can be found in lower row of Table 1.

		HT				**IAT**	
		2 Mbps	155 Mbps	622 Mbps	2,5 Gbps	Metro	Core
Banking sector	BH-BC	166 h	90 min	22.5 min	5.6 min	2.4 h	4.8 h
	BH-BO	3.4 h	2.6 min	40 s	10 s	4.5 s	5.1 s
Video delivering		88h	1h 8 min	17 min	4.2 min	24 min	27 min
Health care		52min	1-9 h		5.6 min	9.2 min	7.2 min

Table 1: Mean HT and mean IAT for Banking sector, Video delivering and Health care sector

In the next Subsections, the problem of the efficient transporting of the IP traffic through ASON networks is formulated and then we present the procedure we designed and evaluated within this Thesis, highlighting its characteristics and merits.

3.2 Efficient and cost-effective transport of IP traffic over ASON/GMPLS networks

Internet Protocol is the clear dominator among the layer 3 protocols. In fact, IP is expected to be the layer integrating most of the emerging video, voice and data applications. IP-based networks are likely to become the main client of ASON networks. Thus, special attention has to be paid to the case of transporting IP traffic over ASONs.

First of all, it can be considered that IP traffic varies at two different time scales, namely long-term variations (i.e. monthly, weekly) and short-term variations (i.e. seconds, hours). Moreover, IP traffic presents very fast fluctuations/peaks as well as slow fluctuations/peaks (See Figure 13).

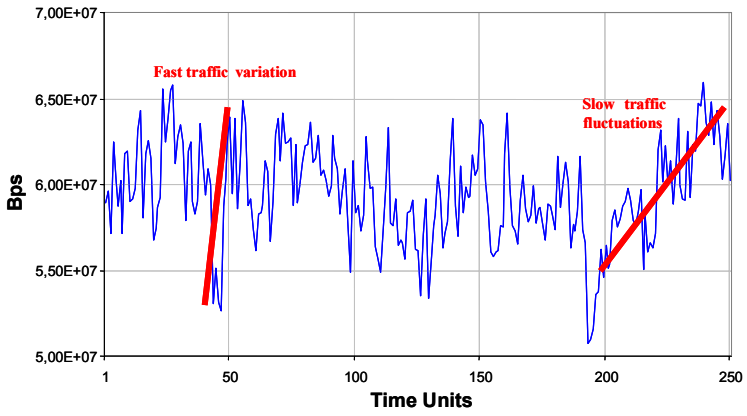


Figure 13: IP traffic fluctuations

For carrying the IP traffic, two ASON transport services can be used, namely the permanent (or soft-permanent) service to support the expected average bit rate and to cope with the IP traffic long-term variations, and the automatic switched service to cope with the short-time IP traffic fluctuations. While the permanent channels are directly set up by the NMS, the switched channels are dynamically set up/released by the CP.

It has to be highlighted that IP layer is characterized to be connectionless, while ASON/GMPLS networks provide a connection-oriented transport service. In such a context, the following questions arise: 1) when a request for the automatic set up/tear down of optical connections (light paths) has to be triggered? And 2) which are the statistical characteristics of the ASON switched connections?

These questions lead, on one hand, to the necessity to design a mechanism that triggers demands to set up/tear down connections/light paths as a function of the fluctuations of the aggregated IP traffic. The aim is to efficiently manage the capacity available in the optical transport layer to cope with the near real-time bandwidth requirements of the client traffic. On the other hand, they lead to define a suitable traffic model for the ASON dimensioning.

These two issues have been addressed in this work, the former in current Chapter and the latter in Chapter 4.

3.2.1 Triggering demand model: Definition of specifications

Concerning the above mentioned former issue, the first step in the design of this procedure consisted on evaluate the specifications it has to cope with. Basically, the clear requirements are: 1) Maximization of the bandwidth utilisation of the light paths (in order to reach TE objectives), 2) Limitation of the number of set up requests in order to avoid, on one hand, both higher layers instabilities and to excessively increase the control plane routing and signalling functions and, on the other, to obtain HT and IAT statistics compatible with the time required by the CP to establish an end-to-end switched light path and, 3) Minimization of the IP packet losses.

Thus, we carried out different studies in order to evaluate the best strategy to manage the capacity available at the optical transport layer to track the fluctuations of the incoming traffic to cope with the above mentioned specifications. In particular, this Subsection deals with the steps we carried out which allowed us to design a proper procedure.

In IP/MPLS environment, a simple LSP set up policy based on the traffic-driven approach has been proposed in [59], in which an LSP is established whenever the number of bytes forwarded within one minute exceeds a threshold. In IP over ASON environment, we use a similar approach. Specifically, we propose a procedure including a monitoring function of the client network traffic offered to the ASON, and triggering the requests for the set up/tear down of switched connections according to what results from this monitoring [52]. As an example, Figure 14 shows a possible architecture based on Traffic Engineering (TE) Servers for an IP over ASON scenario that includes the following functions:

- Traffic monitoring at IP layer.
- Signalling in the control plane, management coordination between IP and ASON.
- Establishing/releasing light paths according to the routing scheme.

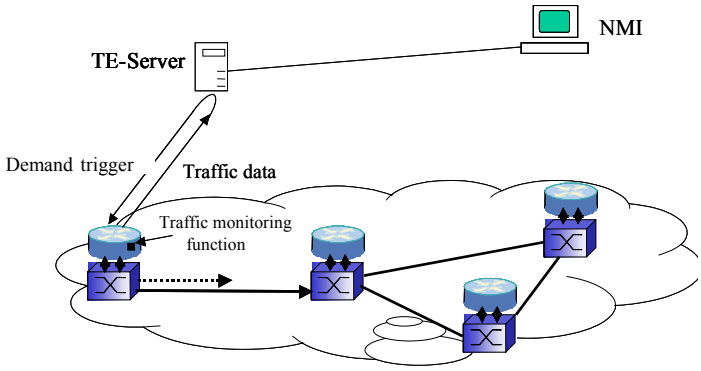


Figure 14: Possible architecture based on TE Servers for IP over ASON

The traffic monitoring function is installed at the ASON client (e.g., the "access" router of the transport network, which owns to the Network Operators), and has to be designed to monitor a parameter of the IP traffic injected in the transport network. Some examples of the monitored parameters can be the amount of the data, the link utilization or the occupancy of the electrical buffer of the router interfaces.

The monitoring function can be done by monitoring, either the instantaneous value to track even the fast variations, or an average value (and thus computed periodically) of the considered parameter. As a result of the monitoring function, node congestion due to a traffic burst/surge or under-utilization of the optical connections condition is detected. In order to do this, a threshold-based policy is used. Specifically, the reaction (triggering of a request for the set up/tear down of switched connections to cope with the burst) is produced once that monitored value goes above/below a predefined threshold. Either a single threshold or two can be used, one for setting up connections (congestion threshold) and the other for releasing them (under-utilisation threshold).

As first step, we considered the monitoring of the instantaneous fluctuations of the aggregated IP traffic from a source node (collecting the incoming traffic from the client network) towards a given destination node, and we based the request for the set up/tear down of the switched connections (SC) using a rule based on a ***static scheme*** [60], which means that the request for switched optical connections is triggered if the traffic in the link exceeds the predefined threshold. The switched optical connection is going to be torn if the traffic comes back below the threshold.

Figure 15 depicts the procedure to request the set up/tear down of switched connections (SC) using the static scheme.

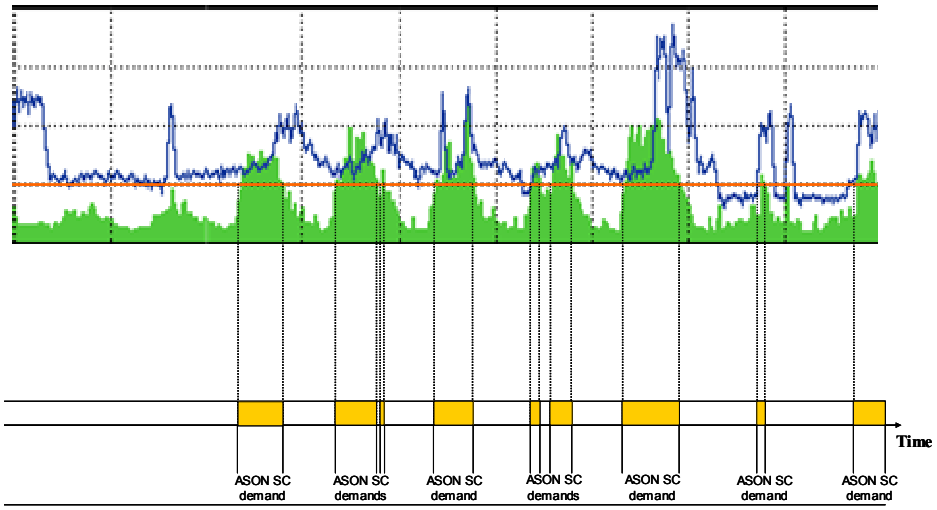
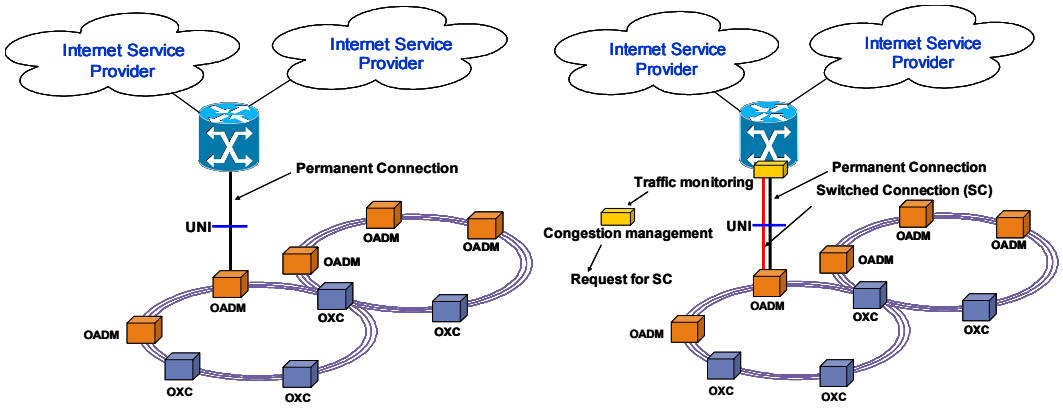


Figure 15: Triggering connection requests procedure based on the Static scheme

Figure 16 (a) and Figure 16 (b) depict how the procedure based on the monitoring of the client traffic works. For example, initially, a permanent connection/light path is established through the NMS to carry the IP traffic towards a given destination edge node (Figure 16 (a)). When a congestion situation, due to a traffic burst, is detected (the monitored parameter is higher than the congestion threshold), then the access router triggers the request to the CP for the establishment of a switched connection towards the same destination node (Figure 16 (b)). Then, some TE rules (basically Load Balancing) are applied to optimize the bandwidth utilization of the light paths.



(a) (b)

Figure 16: IP over ASON, (a) Initially conditions, (b) using SC to track the traffic burst

To evaluate the triggering request procedure using the static scheme we built a simulation model consisted of an access IP Router on top of an OXC. Since the data traffic is characterised by

a self-similar nature [61], the aggregated incoming from the IP clients was modelled using a self-similar traffic pattern and according to the IP packet size distribution suggested in [62].

In the literature, several methods have been elaborated in order to generate artificial traffic traces of self-similar processes to be used as traffic model for simulation case study. A simple method consists in multiplexing *n* flows, each of them being an ON/OFF process with ON and OFF periods having a heavy-tailed distribution (e.g., Pareto distribution) [63].

Let us consider the process resulting from the multiplexing of *n* independent instances of such processes. Let $S_n(t)$ be the process obtained by counting the number of active sources at time *t*. As *n* increases, the process is asymptotically self-similar. Each ON and OFF period is Pareto-distributed. The Pareto distribution has the following probability distribution function (pdf) and mean value:

$$f(x) = \alpha \frac{k^{\alpha}}{x^{\alpha+1}} \quad \alpha > 0 \quad x \geq k > 0$$

$$\mathrm{E}[x] = k \frac{\alpha}{\alpha - 1} \quad \alpha > 1$$

Traffic from a Pareto source is generated using the inverse transform:

$$X_h = \frac{k}{(r_h)^{1/\alpha}}$$

being r_h uniformly distributed between 0 and 1, and relation between the Hurst parameter (H) [61] and the alpha parameter of the Pareto distribution α is given by H=(3–α)/2. From the mean length of the periods, the average load each source offers and thus the overall load may be derived. The advantage of this method relies on its simplicity. Its drawback is the need to choose a large value for *n*, as the property holds asymptotically.

As expected, the obtained simulation results lead to conclude that the procedure for triggering automatically set up demands based on monitoring the instantaneous variations of the IP traffic parameters may cause Control Plane instabilities [64] because it requires to set up/release connections too often (IAT and HT have very low values, in the order of ms or even lower). The same conclusions were drawn by the simulation case study in which we improved the static scheme by using a **hysteresis scheme**, on the basis of which, the request for switched optical channel is triggered if the traffic in the link exceeds the predefined threshold and this link state lasts for at least τ_{up}. The switched optical connection is going to be torn if the traffic comes back below the threshold and this link state lasts at least τ_{down}.

So then, we discarded the monitoring of the IP traffic instantaneous fluctuations to track also the fast variations moving to monitor the traffic periodically. Specifically, we used a scheme consisting of applying an Observation Window (OW). However, using a scheme based on the average traffic calculated over an OW has the strong drawback to not track the instantaneous fast variations which can highly increase the IP packets losses. Therefore, we considered a scheme which allows on one hand to monitor the incoming traffic periodically (tracking the slow variations) and on the other that allows to absorb the peaks due to the fast variations.

Such a scheme/procedure was characterised by the introduction of an electrical buffer which collects the aggregated incoming traffic to be transported towards a specific destination node (See Figure 17). In particular, the monitored parameter was the average occupancy of this electrical buffer (Average Buffer Occupancy, ABO). The procedure includes computing the average value of the BO during each OW and triggering a dynamic connection set up demand accordingly (Figure 17) [52].

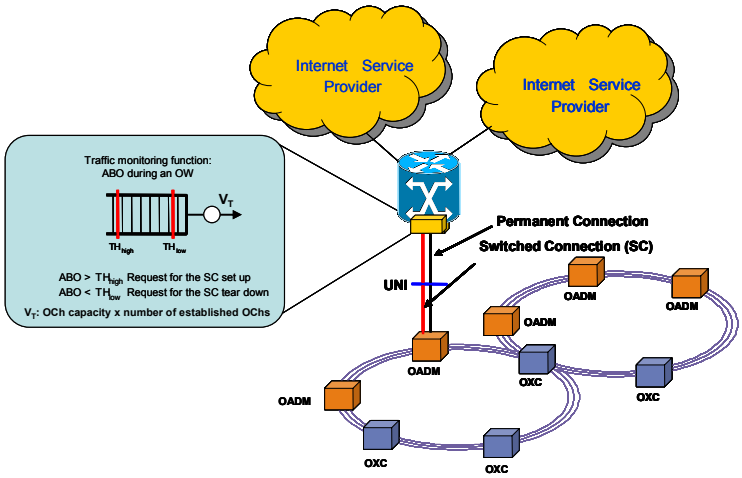


Figure 17: Capacity management based on the ABO monitoring, IP over ASON/GMPLS scenario

As above mentioned, the congestion and/or under-utilisation conditions are detected using a threshold-based policy. Specifically, two thresholds are considered, namely the congestion threshold (TH_{high}) and the under-utilisation threshold (TH_{low}). At the end of each OW, if the ABO is higher than the TH_{high}, the procedure requests to the CP the automatic establishment of a switched connection towards the same destination edge node (e.g. node Y). An example of how this procedure allows to track the traffic fluctuations is depicted in Figure 18. It presents two cases, namely the case indicated as A, in which the incoming traffic is stable and then no requests for setting up switched connections are triggered and, in the case indicated as B, the increase of the

incoming traffic results on the fact that the ABO is greater than the congestion threshold and then the request for the establishment of a switched connection is triggered.

Once the switched connection is established by the CP, the aggregated traffic towards node Y is distributed among the established optical connections. The requests to the CP for the set up/tear down of the switched connections are done via UNI interface or via internal signalling, depending on the architecture of the ASON edge node. To optimise the bandwidth utilisation of the established light paths, a TE policy has to be applied, such as the Load Balancing (LB) feature, supported by current commercial IP routers [65].

On the other hand, if an under-utilization condition is detected (end of the burst), which means that ABO < TH_{low}, the procedure reacts requesting to the CP the tear down of a switched light path once the aggregated traffic to be carried towards the destination node has been reallocated to the remaining established light paths.

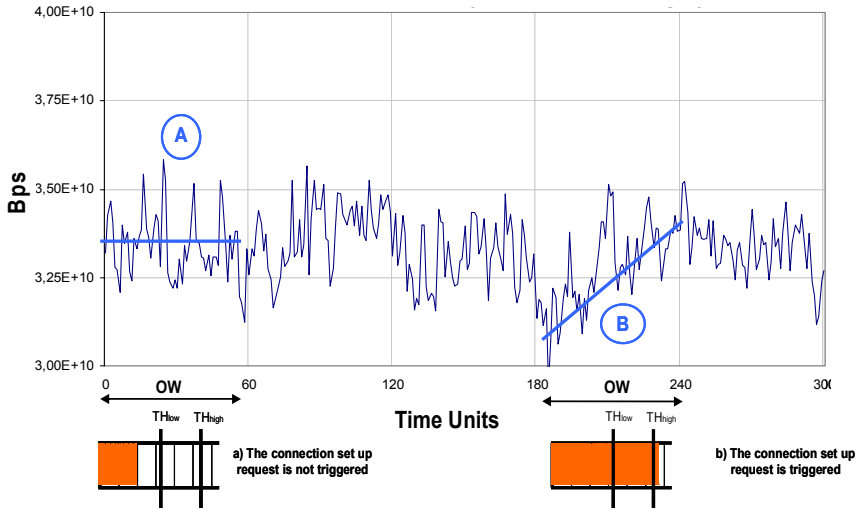


Figure 18: Capacity management based on the ABO monitoring, tracking the traffic variations

By adequately dimensioning the OW size, it is possible to fix the minimum values of IAT and HT, and make them feasible for the ASON Control Plane. Nevertheless, as shown in Figure 19, the proper dimension of the OW is a trade-off since, make it too small (case A in the Figure 19) leads to fall into the same drawbacks of the instantaneous monitoring schemes (very high number of requests for setting up/tearing down switched connections), and making it too large may lead to react behind the actual bandwidth requirements, since the triggering request decision is taken after traffic bursts appear (case B in Figure 19). This latter case may provoke very high IP packets losses due to the buffer overflow, depending on the size of the peaks.

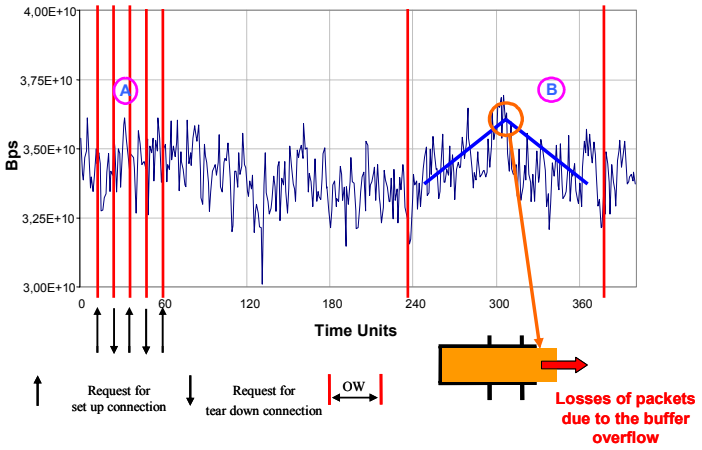


Figure 19: OW size dimensioning issue

To evaluate the feasibility and the effectiveness of such procedure, mostly to characterize the HT and IAT statistics for the switched connections, we built a simulator, in which we considered connections of 2 Mbps, the IP traffic between the couple of source and destination node taken into consideration modelled by the self-similar model above described (with an Hurst parameter, H = 0.9) with a peak rate of 140 Mbps and an average rate of 60 Mbps. It worth stressing that this is a simulation scenario to test the procedure and it is not intended as a potential real network scenario where the applicability of optical connections (DWDM) is competitive for high-capacity bandwidth requests (far above the 2 Mbps). We supposed that the size of the buffers of the router interfaces is big enough in order to avoid IP packet losses. Initially, it is assumed that at least one permanent (or soft-permanent) ASON connection is established towards the destination node. Then switched connections are set up/torn down according to the short-term evolution of that traffic by applying the above discussed procedure. A sample of the results obtained with this case study is reported in Table 2. Specifically, some results for the mean IAT and mean HT of the switched connections for different configurations for the Observation Window (OW) and the thresholds were considered.

OW (s)	TH_{high} (KB)	TH_{low} (KB)	Mean HT (min)	Mean IAT (min)
10	4	4	10.3	22.3
30	4	4	31.4	67
60	4	4	68.1	143.4
10	50	50	10.1	23
30	50	50	30.7	69
60	50	50	61	138
10	450	50	14.5	33
30	450	50	45.6	103
60	450	50	119	214

Table 2: Procedure based on monitoring the ABO, mean HT and IAT

As expected, the bigger is the OW, the higher are the mean IAT and HT. The use of two different thresholds instead of a single one produces higher values for the mean IAT and HT statistics. These results show the feasibility of the procedure since such HT and IAT times (tens or hundreds of minutes) are compatible with the time requirements for the establishment of switched connections of the control plane (i.e., of the signalling and routing protocols).

Nevertheless, as above mentioned, it is worth noting that this triggering procedure, in which the decision for requesting the setting up/tearing down of the switched connections is based on the ABO calculated periodically may produce buffer overloads which lead to high figures for IP packets losses (higher when increasing the OW). In particular, investigating the probability distribution function (pdf) of the buffer occupancy, we obtained that for the different configurations simulated the obtained packet losses figures were unacceptable for the most of the IP applications, in particular for those with stringent QoS requirements. As an example, Figure 20 shows the pdf of the buffer occupancy when TH_{high} = TH_{low} = 4 KB and OW = 60 s. It also depicts the probability that the buffer occupancy is greater than 512 KB and 1 MB respectively, considering these latter as reference values.

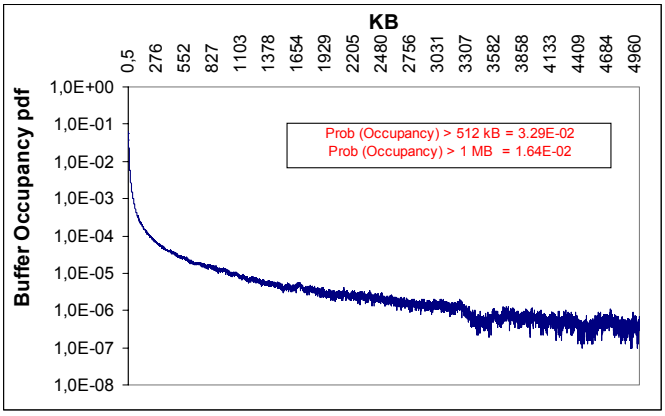


Figure 20: Probability Distribution Function of the Buffer Occupancy

As a consequence, based on the above discussed results, we introduced, apart from the monitoring step, a traffic prediction step. This arises from the consideration that the implementation of a prediction step advantageously allows detecting in advance the occurrence of a traffic burst, so that the control function has enough time to take the suitable decision in order to cope with the burst, reducing in this way the experimented IP packets losses. Nevertheless, using the buffer occupancy as the monitored parameter on which the decision for the set up of additional connection is based makes very hard to take benefits from the prediction step. In fact, as an example, Figure 21

shows the instantaneous buffer occupancy over time. Specifically, in Figure 21 a representative short time interval is considered.

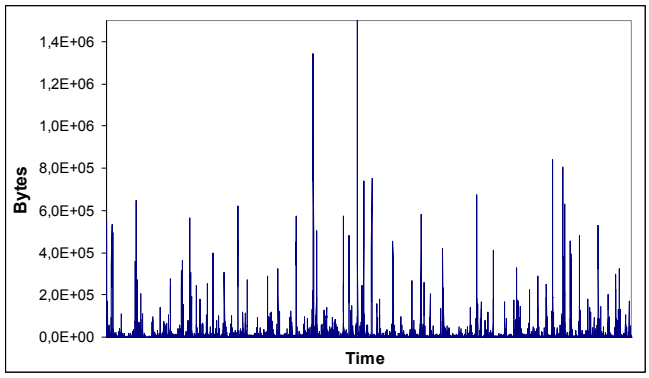


Figure 21: Instantaneous buffer occupancy over time

The strong and continue variations of the buffer occupancy over time as a consequence of the set up and/or torn down of switched connections, make to discard to implement a prediction step in the procedure described so far.

On the basis of these studies about the best approach to define a procedure to efficiently manage the bandwidth available at the optical transport layer to track the fluctuations of the client traffic, next, we propose the procedure for automatically triggering demands to set up/tear down connections in IP/MPLS over ASON/GMPLS networks.

We call this procedure TRIDENT (TRIggering DEmands mechanism for the connection set up and tear down in optical transport NeTworks based on traffic monitoring and prediction). This procedure is the matter of a Patent Application submission to the European Patent Office [66] which has been jointly authored by the Universitat Politècnica de Catalunya (UPC) and Telecom Italia Labs (TILAB), the R&D department of the Italian Network Operator Telecom Italia [67].

3.3 TRIDENT: A procedure for the automatic demand for setting up/tearing down connections in IP/MPLS over ASON/GMPLS networks

Data traffic fluctuations and network resources utilization are two opposite aspects in a circuit switching network. Thus, coping with the fluctuations of the aggregated data traffic to be transported by ASON/GMPLS networks and keeping limited the size of the transport networks, requires a proper dynamic management of the ASON switched connections.

In such a context, we define and evaluate TRIDENT, a distributed capacity management procedure for establishing switched connections to automatically track the dynamic changes of the incoming traffic. The procedure also defines traffic engineering (TE) rules required to optimize the transport network resource utilization, reducing in this way the CAPEX of the Network Operators.

TRIDENT interworks between the client network (IP/MPLS) and the circuit switched server layer of an ASON/GMPLS network supporting permanent, soft-permanent and switched connections service for the dynamic use of transmission resources. It can be classified as a multi-layer traffic engineering approach to track the fluctuations of the client networks traffic. Figure 22 shows the scenario where to apply TRIDENT.

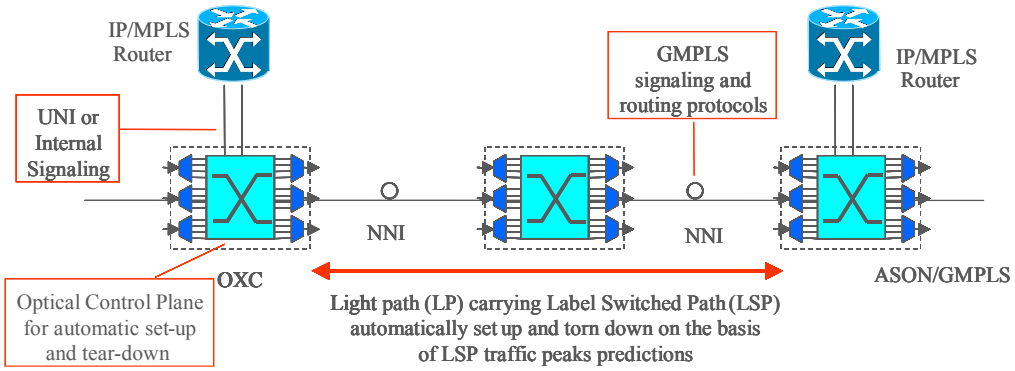


Figure 22: IP/MPLS over ASON/GMPLS network scenario

Specifically, Figure 23 depicts, as an example, an edge node integrating an IP/MPLS router and an optical switch to which TRIDENT is applied.

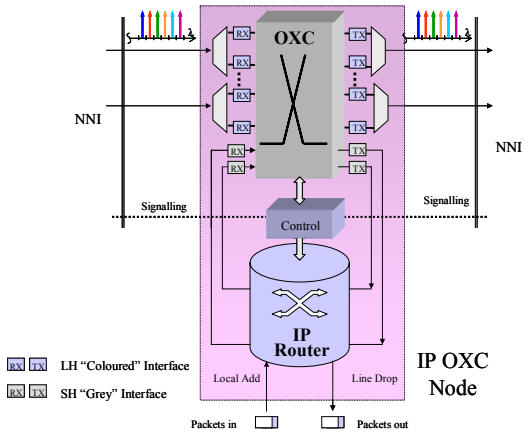


Figure 23: Peer-to-peer edge node architecture

On the basis of the growing penetration of IP-based applications which have QoS requirements (e.g., latency, jitter, packet losses), in the definition of the TRIDENT procedure we consider multi-

service networks which means to take into account traffic priorities (specifically two priorities, namely high and low priority traffic). Specifically, client traffic is classified as high priority (HP) and low priority (LP) Label Switched Paths (LSPs) according to the carried applications and/or according to the established Service Level Agreements (SLA). In particular, TRIDENT implies that HP LSPs are bundled onto HP light paths and LP LPSs are bundled onto LP light paths (Figure 24). In such a way, the two types of LPs may have different routing (e.g. path length, number of hops) and survivability policies (e.g. protection, not protection, restoration, etc.) in order to meet their requirements within the transport network.

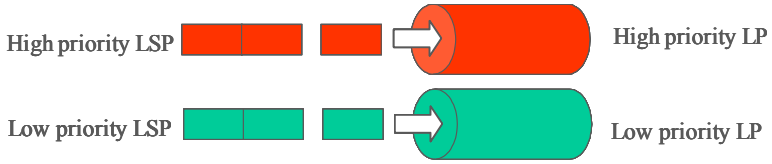


Figure 24: Label Switched Paths (LSPs) carried by Light Paths (LP) of the same priority

To implement the TRIDENT procedure, some interfaces of the network nodes (Layer 2 and Layer 3) are supposed to be directly allocated to high priority traffic (supporting for example a number of permanent and soft permanent/switched light paths) and some interfaces are allocated to low priority traffic (supporting for example a number of switched light paths) (Figure 25).

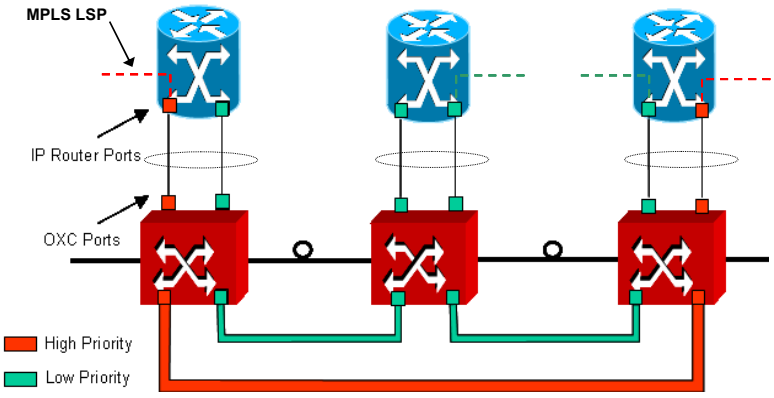


Figure 25: TRIDENT procedure, System configuration

This procedure is suggested to be implemented in the edge nodes of the IP/MPLS over ASON networks and it consists of the following steps: 1) Monitoring and short-term prediction of the data traffic entering a node at client layer, 2) Congestion management, taking benefits form the different traffic priorities, and resource utilization optimization, and 3) Automatic set up/tear down of switched optical connections.

Figure 26 depicts the interactions among the three steps which are described in detail in the next Subsections.

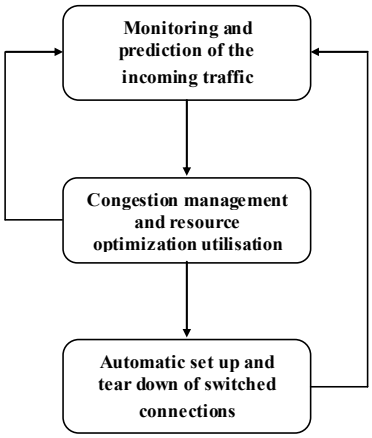


Figure 26: Steps of the TRIDENT procedure

3.3.1 Monitoring and prediction of incoming data traffic

IP traffic typically fluctuates irregularly during a day and therefore peaks of the traffic can only be identified by traffic measurements [68]. Thus, we used periodic network traffic monitoring enforced by short-trend traffic estimation in order to automatically adapt the network resources to the current network conditions in near real time, reacting to prevent possible localized congestion due to high priority traffic bursts/peaks.

According to [69], [70], [71], traffic monitoring can be performed on: Flow-based, Interface-based, Link-based, Node-based, Node-pair-based and MPLS LSP-based. Specifically, TRIDENT requires the interface-based and MPLS LSP-based traffic monitoring.

As a first option to measure the actual traffic crossing the router interfaces, the Network Management System (NMS) can be used, which means to carry out the traffic monitoring via the Simple Network Management Protocol (SNMP). It is an Internet-standard protocol for managing and control devices (e.g., routers) on IP networks. SNMP is characterised by a simple set of operations that allows the devices to be remotely managed and controlled [72]. For example, SNMP protocol can be used to shut down an interface of the managed device or check the speed at which an Ethernet interface is operating. Any sort of devices status or statistical information that can be accessed by the NMS has to be defined in a Management Information Base (MIB). On the network devices to be remotely managed, a piece of software (called *agent*) has to run. The NMS obtains the required information on the basis of the requests sent to the *agent* of the network devices. As an example, the *agent* built into the IP router will respond to the NMS requests for the variables (*objects*) defined by the MIB standard [73]. In addition, the router may have some significant new features that are worth monitoring but are not covered by any standard MIB. So, basically, the

vendor defines its own MIB (proprietary MIB) that implements managed objects for the status and statistical information of their new router.

One of the most useful objects for network monitoring defined in the standard MIB is the so-called *interface* table [72]. It contains, among others, the following variables:

- *ifInOctets*: It represents the total number of octets (bytes) received by the polled interface.
- *ifOutOctets*: It represents the total number of octets (bytes) sent by the polled interface.

As already mentioned, TRIDENT relies on both the interface-based and MPLS LSP-based monitoring. We have investigated if current commercial IP/MPLS routers support MIBs able to store information about the traffic carried by the LSPs. Specifically, we found that MPLS routers from Cisco Systems implement the proprietary MIB called *mplsLsrMibCapabilityV12R0* [74], on which the *mplsInSegmentOctets* variable is defined, while we found that IP/MPLS routers from Juniper Networks support the *jnxMibs.mpls* MIB [75]. Specifically, the latter defines the variable *mplsLspOctets* which allows monitoring "*the number of octets (bytes) that have been forwarded over current LSP active path*". On the other hand, there are currently a number of third-party measurement/monitoring products available [76]. Hence, another option is to deploy such equipments, which might have performance advantages but also introduces additional cost. As an example, very recently a novel concept for passive network monitoring of DWDM optical networks has been developed [77]. Specifically, the collection and real-time analysis of IP packet data from any one of the active 10 Gbit/s wavelength carriers on a DWDM optical network link is enabled.

In the TRIDENT procedure, the collection of the data traffic samples is carried out periodically (according to the monitoring tool) and averaged over an Observation Window (OW). Starting from these samples, traffic prediction is carried out to forecast the short-time evolution of the data traffic entering the node interfaces for the next OW. To carry out the short-term prediction of the incoming traffic to each IP/MPLS high priority router interface, the adaptive Normalized Least Mean Square (NLMS) error linear predictor ([78], [79]) is used. Since it is based on an adaptive approach, it is used as an on-line algorithm for forecasting the aggregated traffic bandwidth requirements. Generally speaking, a *k*-step linear predictor (Figure 27) is concerned with the estimation (prediction) of the next *k-th* sample of the signal *x(n)*, which means that by using a linear combination of the previous values of *x(n)*, *x(n+k)* is predicted.

A *pth*-order linear predictor has the form:

$$\hat{\chi}(n+k) = \sum_{l=0}^{p-1} w(l)x(n-l)$$

where *w(l),* for *l = 0, 1, ..., p-1* are the prediction filter coefficients.

The aim of the linear predictor is to minimize the mean square error defined as:

$$e(n) = x(n+k) - \hat{\chi}(n+k)$$

The update equation for the filter coefficient is:

$$w(n+1) = w(n) + \frac{\mu\, e(n)\, x(n)}{\|x(n)\|^2}$$

where μ is a constant called step size.

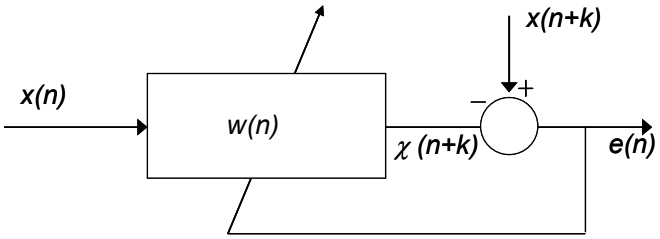


Figure 27: Adaptive linear predictor

An important advantage of using the NLMS predictor relies that is not highly sensitive to the step size. In fact, it has to be underlined that predictor design parameters should be chosen as a trade-off between low prediction error and fast convergence for prediction algorithm. Using large values for μ results in a faster convergence and quicker response to signal changes while using small μ results in slower convergence and less fluctuations after the convergence. After different simulation tests, in our implementation of the algorithm we set μ = 0.01.

In our procedure, *x(n)* represents the traffic crossing the router interfaces and the following variables are defined:

- Prediction sample period = τ
- Number of prediction sample periods used to predict the *k-th* future value of the traffic: *p*

Thus, at time t_0, the past *p* samples are used to predict the traffic value for at the time (t_0+k·τ). Moreover, it has to be considered that we use a linear algorithm in order to avoid increasing excessively the complexity of the procedure.

3.3.2 Congestion management and resource utilization optimization

The objective of this step is, as a result of the previous one, checking, at the end of each OW, the potential occurrence of node congestions (due to the high priority traffic) and deciding whether or not automatically trigger the request to set up/tear down switched light paths to cope with such congestions. The procedure relies on the threshold-based policy depicted in Figure 28. Comparing the predicted traffic crossing the HP interface *n* (i.e., B_n) with the thresholds of congestion (TH_{high}) or under-utilization (TH_{low}) (hereafter high threshold and low threshold respectively), a decision making function automatically manages the way to admit the high priority data traffic to proper connections.

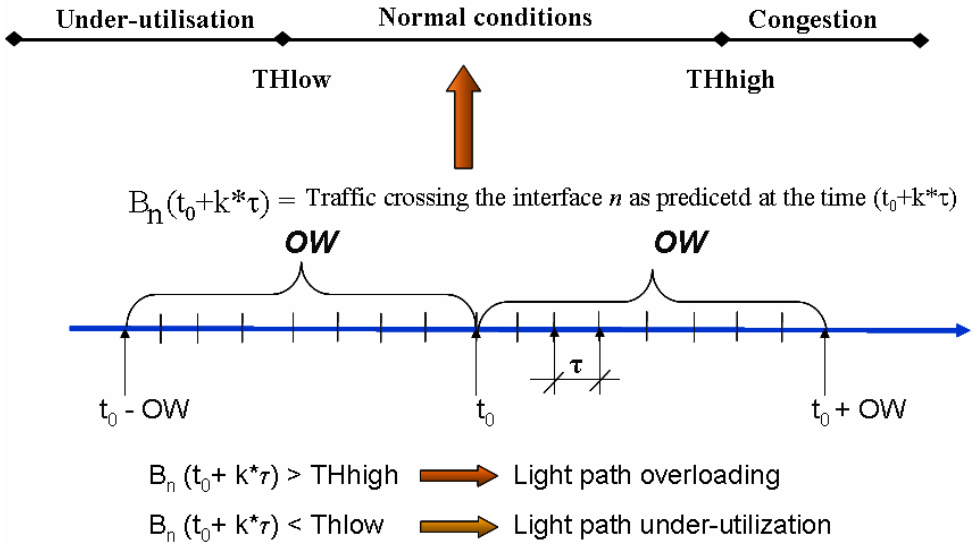


Figure 28: Threshold-based policy for congestion management and resource utilization optimization

Figure 29 and Figure 30 depict an example of how the procedure works. Initially one router interface is allocated to HP traffic (i.e. it supports a permanent HP light path) and the other router interface is allocated to LP traffic (supporting a LP light path).

When high priority traffic surges are predicted on the permanent connection (i.e., B_n higher than the high threshold), the client requests some network resources to the server layer, even if previously dedicated to low priority traffic, to be torn down in order to make them available for the high priority traffic.

The node interface made available at the server layer is used for setting up the HP switched light path to accommodate the traffic surges, thus avoiding the network congestion (Figure 30).

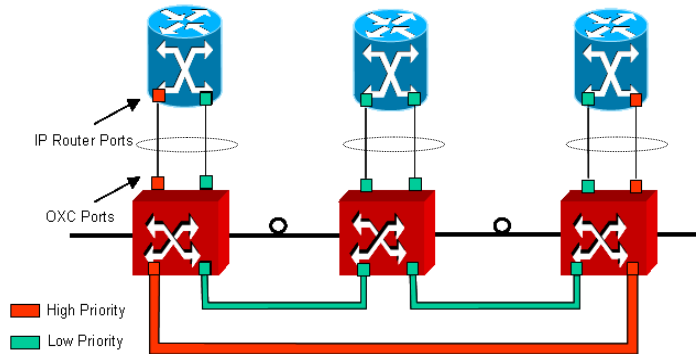


Figure 29:IP/MPLS over ASON/GMPLS network scenario with the logical LP and HP optical connections, initial conditions

As above mentioned, the detection of both the over- and under-utilisation of the light paths relies on the thresholds. It has to be underlined that the election of the low and the high thresholds value has to be the result of the compromise among the higher layers stability as well as the routing and signalling cost functions (i.e., it has to be avoided to set up/tear down connections too often), the packets lost at the high priority router interfaces, which are equipped with limited buffers, due to the sudden increases of the HP traffic (hereafter Packet Loss Rate, PLR), and the bandwidth utilization of the optical connections.

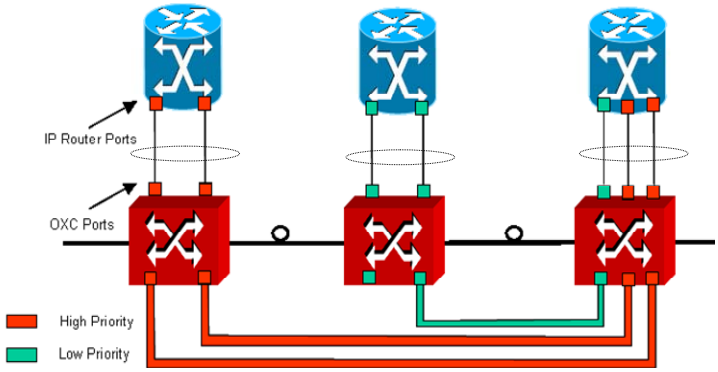


Figure 30: IP/MPLS over ASON/GMPLS network scenario with the logical LP and HP optical connections, conditions after tracking the HP traffic surges

Once the requested switched light path is established by the control plane, the predicted high priority traffic is then accommodated into the new light path and traffic engineering rules are applied to optimize the LSPs allocation in order to increase the bandwidth utilization of the light paths. When the traffic surge expires (B_n lower than the low threshold), the HP switched light path previously allocated is torn down and the network resources are restored to the initial conditions.

Figure 31 shows the pseudo-code for the procedure to handle an HP traffic burst on an HP interface. When the traffic surges expire (the predicted traffic on the HP permanent light path crosses the low threshold) the switched HP light path previously allocated is torn down and the

network resources are restored to the initial conditions. Figure 32 shows the pseudo-code for the procedure which handles the tear down of an HP light path according to the end of the HP traffic burst. The third part of this Ph.D. Thesis includes the description of TRIDENT procedure in detail, including the TE rules designed to optimise the bandwidth utilization of the light paths.

Input: monitoring of the traffic crossing the interface n, B_n

Output: Automatic triggering of the request for the set up/tear down of switched optical connections

Algorithm:

At the end of each OW, on the interface n:

```
While (Bn > High Threshold)
    if (other HP light path towards the same destination exists (t))
        while ((Bt < High Threshold)&(Bn > High Threshold)
            1. Tag of the LSPs to be rerouted for load balancing
            2. Update of Bt and Bn
    endif

    if (Bn <= High Threshold)
        reroute of the tagged LSPs to the light path supported by t
    else

        1. Request for the tear down of LP light path on interface t
        2. Triggering of the request for the set up of a HP light path
           on interface t
           if (GMPLS signalling OK)
               While (Bn > High Threshold)
                   Rerouted of the LSPs for load balancing
               endwhile
           Else
               Reroute of the tagged LSPs
```

Figure 31: TRIDENT procedure, handling HP traffic burst at HP interface *n*

Input: B_n, B_t

Output: Triggering the request for the set-up/tear down of switched optical connections

Algorithm:

At the end of each OW, on the interface n:

```
if (Bn < Low Threshold)
    while ((Bt > 0) & (Bn < High Threshold))
        1. Tag of the IP/MPLS traffic to be rerouted on the HP light path
           supported by the router interface n
        2. update Bn and Bt
    endwhile
    if (Bt = 0)
        1. Reroute of the tagged IP/MPLS traffic
        2. Tear down of HP light path supported by router interface t
        3. Triggering request for the set up of LP light path on router
           interface t

endif
```

Figure 32: TRIDENT procedure, handling the end of the HP traffic burst on HP interface *n*

3.3.3 Automatic Set up/Tear down of switched connections

The automatic set up/tear down of the switched connections is done via the Control Plane (CP) and Network-Node Interface (NNI) signalling [80], [18]. Figure 33 depicts an example of switched light path establishment through the UNI and NNI interfaces.

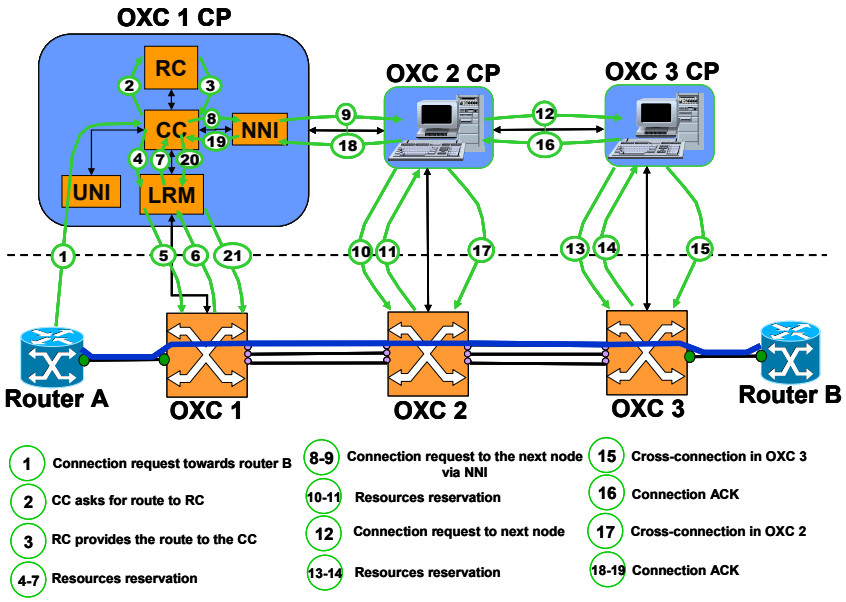


Figure 33: Example of connection establishment via UNI and NNI interfaces

Figure 34 (a) represents the block diagram of the Optical Control Plane for setting up/tearing down switched light paths in ASON/GMPLS-based nodes. To achieve that, a light path is identified by appropriate information such as source and destination node Identifier (ID), port IDs, light path ID and payload type.

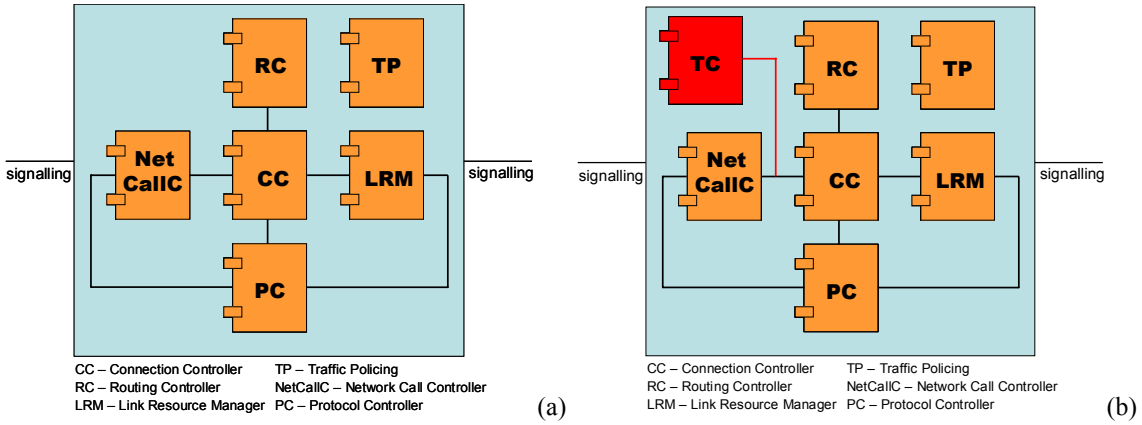


Figure 34: (a) Optical Control Plane for setting up/tearing down switched light paths in an ASON/GMPLS, (b) TRIDENT procedure: adding the Traffic Control (TC) component

The TRIDENT procedure requires the introduction of a Traffic Control (TC) component in the Control Plane of the switched transport network (Figure 34 (b)).

The TC component is responsible for the collection of the raw data on the router interfaces, elaborating these data to predict the traffic trend and then making a decision whether or not requesting (e.g., via UNI or internal signalling) to the Connection Controller (CC) of the Control Plane, the set up/tear down of a switched connection.

3.3.4 Techno-economic advantages of the TRIDENT procedure implementation

This Section is devoted to highlight potential benefits obtained through the implementation of the TRIDENT procedure. It is designed as a cost-effective solution for Network Operators to transport different classes of traffic (i.e., high and low priority traffic). On the other hand, it allows the optimized utilization of the network resources according to the typical daily data traffic fluctuations as well as unexpected traffic variations due for example to equipment failures or unexpected traffic burst. Specifically, it avoids dimensioning the network resources destined to a certain class of traffic to cope with the peaks of that traffic class. Figure 35 shows that if C is the capacity of a light path, then the number of permanent light paths potentially required in a static network to account for the peaks (B_{max}) of high priority traffic is $N = B_{max}/C$ (e.g., in At2). The result of such kind of transport network dimensioning is the no-optimal bandwidth utilization of the N light paths according to the daily traffic evolution.

Contrarily, implementing the TRIDENT procedure, only $P = B_{nor}/C$ permanent light paths (for example set up by the NMS) are constantly used. In such a way, there are other additional $S = N-P = (B_{max}-B_{nor})/C$ light paths (switched connections) to cope with HP traffic peaks on the basis of periodical traffic monitoring and short-time prediction. Such light paths can be used to carry low priority traffic in absence of high priority traffic bursts or used for fast restoration in case of equipments failure.

Therefore, according to Figure 35:

- During At1: P light paths are carrying HP traffic whilst S light paths are carrying LP traffic
- During At2: N light paths are carrying HP traffic (after dropping the interface for LP traffic)
- During At3: P light paths are carrying HP traffic whilst S light paths are carrying LP traffic

For a given number of interfaces (N), the procedure allows to handle HP traffic peaks and at the same time to carry (B_{max}-B_{nor})*(At1+At2) low priority data traffic (i.e. when there are no HP traffic peaks).

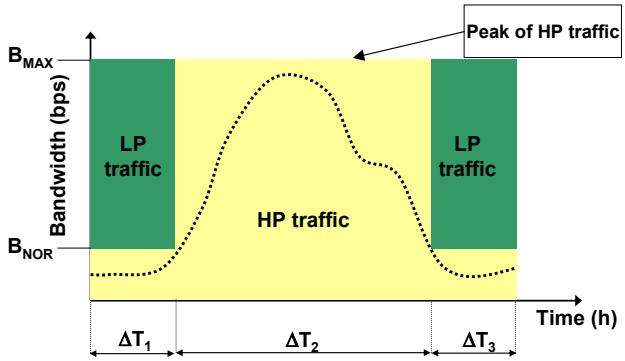


Figure 35: Basic techno-economic considerations

Another advantage provided by the procedure is that low priority traffic is transported only by low priority optical switched connections thus reserving high priority resources for high priority traffic requests. This way helps on guarantee the QoS parameters such as delay, jitter and PLR required by the real-time applications.

3.3.5 TRIDENT procedure: Performance Evaluation

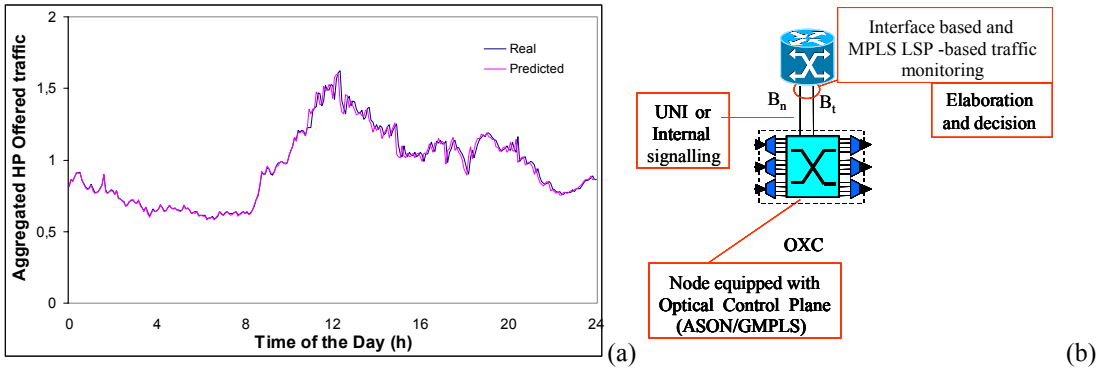
Simulation case studies were carried out to evaluate the performance as well as the effectiveness of the procedure to handle high priority traffic surges incoming to the router interfaces tagged as HP. The simulated network was composed by IP/MPLS edge routers connected through a meshed ASON/GMPLS transport network.

Firstly, two case studies were carried out. The first one aimed at obtaining the number of the high priority light paths required to carry a certain high priority traffic pattern. In this first case study, the influence of the different parameters of the procedure on its performance was also evaluated. Specifically, we evaluated the influence of the high and the low threshold values as well as of the size of the Observation Window (OW). The second case study aimed at illustrating the effectiveness of the procedure in case of unexpected traffic surges, showing how it allows to promptly react to the unexpected fluctuations of the client traffic diverting LP network resources to cope with the surges.

The simulation results hereby reported refer to the HP light paths established between a couple of nodes (source-observed sink), according to a given daily HP traffic pattern between the two nodes. Focusing specifically on the first case study, Figure 36 (a) shows the daily HP traffic profile

(normalized to the router interfaces capacity and averaged over 5 minutes) between the two nodes. In order to test the procedure with real daily traffic profiles, it has to be underlined that the traffic pattern we used in our simulation studies has been extracted from the traffic monitoring of the Catalan R&A Network (*Anella Cientifica*). Specifically, the one-day long traffic profile of Figure 36 (a) was extracted from the real trace obtained on September 22, 2003.

By applying the TRIDENT procedure, the monitoring and prediction of the traffic crossing the router interfaces is carried out periodically during the OWs. At the end of every OW, the set up or the tear down requests are triggered on the basis of the predicted traffic for the next OW, as the result of applying the NLMS prediction algorithm described in Section 3.3.1. In particular, for the prediction algorithm the following parameters were used: p = 12, k = 3 and τ = OW/3. Figure 36 (a) also depict the predicted traffic resulting form the application of the NLMS algorithm. It was assumed that each router interface is equipped with buffer which size was set to 1 Mbytes.



(a) (b)

Figure 36: (a) Daily HP IP/MPLS traffic profile between the source and the destination nodes, (b) Source node at which the procedure is applied

Figure 36 (b) shows the source node at which the procedure is applied. It consists of an IP/MPLS router which represents the "access" router collecting all the incoming traffic from different clients (i.e. ISPs) of the transport networks and an OXC, equipped with a control plane.

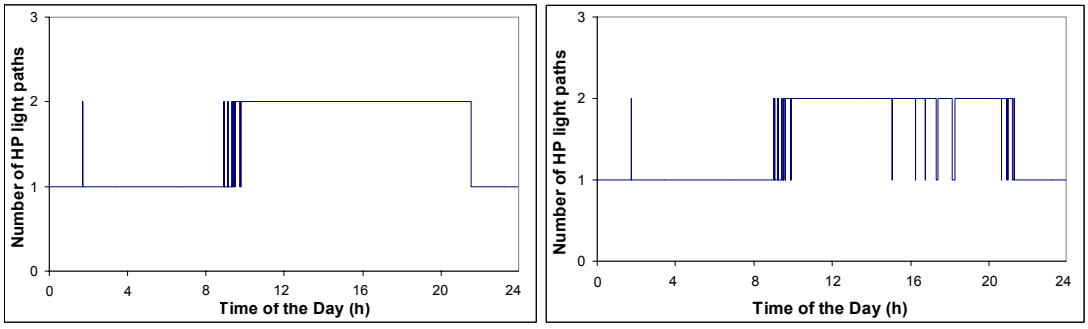
Initially, one high priority router interface (supporting a high priority permanent connection) is allocated to HP traffic towards the observed sink node, whilst the remaining ones are allocated to HP and LP traffic towards other sink nodes.

Figure 37 depicts the number of established HP light paths versus time, which are used to transport the HP traffic of Figure 36 (a) between the source and observed sink node; we set the high threshold (TH_{high}) to 95% of the interface capacity while the low threshold (TH_{low}) was set to 40% (Figure 37 (a)) and 60% (Figure 37 (b) respectively. Observation Windows of 1 minute long were used. It can

be observed that the number of HP light paths established between the source-sink nodes under simulation dynamically rises and falls following the HP traffic dynamics. If we had used the static dimensioning (i.e., over-provisioning approach), two HP light paths would have been used during all the simulated time, according to the client traffic peak. By applying the suggested procedure, the second interface is used to set up the second HP light path only when the permanent connection is not able to carry the aggregated HP traffic; otherwise it can be used, for example, to provide new potential services (e.g. dynamic connections for Storage) connecting the edge router towards other destination nodes. This means that, by applying the TRIDENT procedure, there is some additional capacity available in the network with the same number of network resources.

As stated above, extensive simulations have been carried out to evaluate the impact on the performance of the procedure of the different parameters characterizing the procedure itself. Firstly, it has to be highlighted that the applied thresholds policy constitutes a compromise taking care of PLR and the bandwidth utilization of the optical connections/light paths. Secondly, it has to be specially avoided to request to set up/tear down connections too often. This is due, basically, on one hand, to avoid potential instability at the higher layers and, on the other hand, to avoid to excessively increase the cost of the routing and the signalling functions of the control plane.

Figure 37 shows the effect of the increase of the low threshold value. The mean number of the light paths required to carry the high priority traffic when the low threshold is set to 60% of the interface capacity is lower than the case of fixing it to the 40%. This is due to the fact that the higher the low threshold is, the earlier the procedure starts to restore the initial conditions and therefore the lower is the mean holding time of the switched connection.



(a) (b)

Figure 37: Number of light paths needed to carry the high priority IP/MPLS traffic, a) TH_{low} = 40% of the interface capacity, b) TH_{low} = 60% of the interface capacity

The low threshold also influences the bandwidth utilization of the permanent connection as well as the experimented PLR. In fact, to restore the initial conditions (i.e., to request to the Control

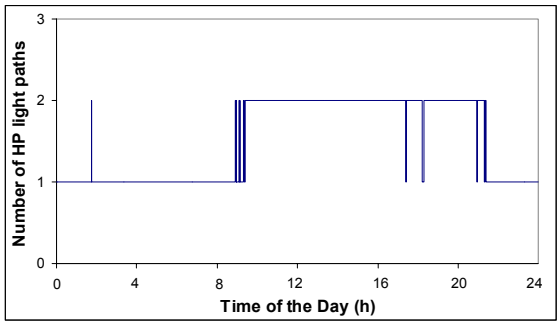
Plane the tear down of the switched connection), the IP/MPLS traffic that is being carried by the high priority switched connections has to be switched back to the permanent connection. Therefore, the higher the low threshold is, the higher is the bandwidth utilization of the permanent channel (it is shown in Table 3). Nevertheless, as pointed out above, the number of requests for the set up/tear down of the switched connection is higher, increasing thus, the routing and signalling cost functions. Besides, the higher the low threshold is, the higher is the PLR (Table 3).

TH_{high} = 95% **TH_{low} (%)**	**Mean Bandwidth utilization**	**PLR**	**TH_{low} = 60%** TH_{high} (%)	**Mean Bandwidth utilization**	**PLR**
40	68%	4.79E-05	90	74%	3.5E-05
60	77%	0.97E-04	95	77%	0.97E-04
80	78%	1.0E-04	97	78%	1.7E-04

Table 3: Impact of the high and low thresholds

As a solution to avoid requesting the setting up/tearing down of the switched connection too often, we define a conservative approach. It consists on holding the switched connection even if not strictly required, which means that the request for the tear down of the HP switched connection is triggered only when the experimented under-utilization condition is repeated again for a certain number of consecutive (n) OWs. As an example, Figure 38 shows the light paths established to carry the client traffic by applying the conservative approach with n = 3. The result is referred to the case in which the high threshold is set to 95% and the low threshold is set to 60% of the router interfaces capacity.

The mean number of requested light paths is lower than the case without the conservative approach. Also, in this case, the PLR is lower. Nevertheless, holding the switched connection implies that the bandwidth utilization of the permanent connection is lower (Figure 38 (b)).



(a)

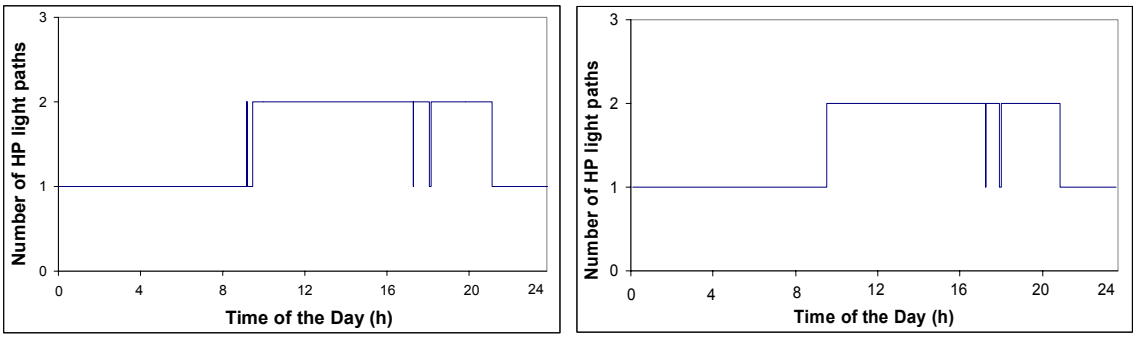
TH_{low} = 60% Tear-down counter = 3 TH_{high} (%)	**Bandwidth utilization**	**PLR**
95	73%	4.4E-05
97	77%	1.1E-04

(b)

Figure 38: (a) Number of light paths using the conservative approach, (b) Permanent connection mean bandwidth utilization and experimented PLR

On the other hand, to reduce the number of requests for the set up/tear down of the switched connection, another approach is based on increasing the size of the OWs. As an example, Figure 39

shows the number of the light paths established to carry the HP traffic when OWs of 3 and 5 minutes long respectively were considered. In fact, in this case, the number of requests for the set up/tear down of switched connections is lower than the previous cases (i.e., OW of 1 minute).



(a) (b)

Figure 39: Number of light paths increasing the OW, a) OW = 3 min, b) OW = 5 min

However, with the increasing of the OW size, the experimented PLR is higher (See Table 4). Basically, this is due to the fact that the higher the OW is, the worse is the prediction of the traffic crossing the router interface.

TH_{low} = 60% TH_{high} = 95%	**PLR**
OW (min)	
1	1.0 E-04
3	2.38E-04
5	3.4E-04

Table 4: PLR when increasing the OW

However, by implementing the prediction step the PLR is lower with respect to the case without prediction. In fact, in all the simulated configurations, using the traffic prediction provides better PLR figures than the cases without prediction. As an example, the Table 5 shows the comparison when using OW of 3 and 5 minutes respectively, and setting the TH_{high} and TH_{low} to 95% and 60% of the interface capacity respectively.

TH_{low} = 60% TH_{high} = 95%	**PLR**	
OW (min)	With prediction	Without prediction
3	2.38E-04	2.65E-04
5	3.4E-04	1.46E-03

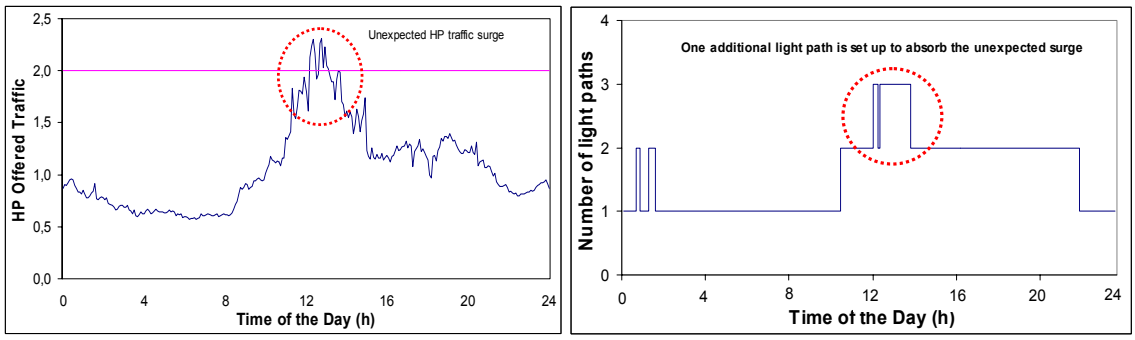
Table 5: Improving PLR by using the prediction step

We also evaluated the influence of the high threshold when maintaining fixed the low threshold. It influences both the bandwidth utilization of the permanent connection and the PLR. Specifically, as depicted in Table 3, the higher the high threshold is, the higher is the bandwidth utilization. This

is because, when the high threshold is increased, the congestion condition, and consequent request for the HP switched connection, is detected when the actual bandwidth utilization of the permanent connection is higher.

Nevertheless, increasing the high threshold implies that there are more packets lost at the router interfaces are as a consequence of the sudden increases of the HP traffic. In fact, as shown in Table 3, the experimented PLR is higher. On the other hand, the thresholds value can be can be dynamically changed on the basis of the actual bandwidth utilization of the light paths and/or on the basis of the actual PLR. For example, we can introduce an Updating Window (UW), much larger than the OW (e.g., UW = m·OW). Sizing properly the UW, the thresholds values can be dynamically modified on the basis of the bandwidth utilization and PLR during the previous UW.

The second case study aimed at evaluating the effectiveness of the procedure in case of unexpected traffic bursts. The HP traffic profile showed in Figure 40 (a) is the daily traffic profile between the two nodes. If compared with the previous traffic profile, it is characterized by an unexpected traffic burst and, thus, the traffic volume to be transported is higher than the expected one. In this case the procedure, contrarily to both the over-provisioning and the scheduled approach, is able to promptly react to the traffic burst adapting the bandwidth to the traffic that has to be carried. The plot of Figure 40 (b) was obtained with the high threshold set to 95% and the low threshold set to 60% of the router interface capacity. OW of 5 minutes long and the conservative approach with n = 3 were used.



(a) (b)

Figure 40: (a) HP client traffic with unexpected burst/surge, (b) Number of HP light paths needed

As stated above, extensive simulations were carried out to test the procedure considering different values for the parameters of the procedure itself. The following Table 6 summarizes the results obtained when OW 1 minute long was considered. Firstly, the TH_{high} is fixed to 95% of the interface capacity while varying the TH_{low}, the number of set up requests, the percentage of time the

additional HP light path is required and the mean HT and IAT for the switched connection were obtained. Table 6 shows that the percentage of time (calculated over the simulation time, i.e., 24 hours) the HP switched connection is required, is about the 50% (about 12 hours). This means, that during about half a day, the second interface can be used to carry different traffic (e.g., LP traffic) or to provide new emerging services such as storage service. On the other hand, the obtained mean HT and IAT are compatible with the time required to provide a switched connections by the control plane.

OW = 1min $\mathbf{TH_{high}}$**= 95%** TH_{low} (%)	Number of set up requests	Time percentage of using the SC[1]	Mean Holding Time (min)	Mean InterArrival Time (min)
40	11	51%	61	121
60	21	48%	34.45	69.57
80	21	47.2%	32.38	72.35

OW = 1 min, $\mathbf{TH_{low}}$ **= 60%** TH_{high} (%)	Number of set up requests	Time percentage of using the SC[1]	Mean Holding Time (min)	Mean InterArrival Time (min)
90	19	52.5%	39.8	76.79
95	21	48%	34.45	69.57
97	25	46.2%	27	58.64

[1] calculated over the simulated time

Table 6: OW = 1 min, summary of simulation results

In the case of maintaining fixed the TH_{low} and increasing the TH_{high} threshold, the effect is to increase the number of set up requests. This due to the fact that the higher is the TH_{high}, the higher is the amount of traffic switched back from the switched light path when the tear down procedure is activated. Then, the probability to cross again the TH_{high} on the permanent connection is higher.

The TRIDENT procedure does not require the knowledge of the future traffic pattern. It is an adaptive procedure to efficiently manage the bandwidth available at the optical level and able to provide in near-real time the bandwidth required by the LSPs established at the client layer. Therefore, as additional simulation case study, we applied the designed procedure also to different traffic profile extracted from the Catalan Academic Network (CAN). Specifically, an one-day long traffic profile was extracted from the real trace obtained on May 13, 2004. The simulation results hereby reported refer to the HP light paths established between the same couple of nodes, according to the daily HP traffic pattern between the two nodes. Initially, one high priority router interface (supporting a HP permanent connection) is allocated to HP traffic towards the observed sink node, whilst the remaining ones are allocated to HP and LP traffic towards other sink nodes. As a sample of the obtained results, the following Figures and Tables refer to the number of HP light paths used

to carry the considered HP traffic profile. First of all, it can be observed that in this case, the HP traffic volume between the two observed nodes is higher than the previous case. Figure 41 report the number of HP light paths established in the case of using OW 5 minutes long, TH_{high} = 95% and TH_{low} = 60% of the interface capacity, and by applying the above described conservative approach with n = 1 and n = 7 respectively. First of all, the number of HP light paths established between the source-sink nodes under simulation dynamically rises and falls following the HP traffic dynamics. The switched connections are used only when strictly required. By increasing the parameter n, the number of set up requests decreases making the procedure feasible with the CP requirements.

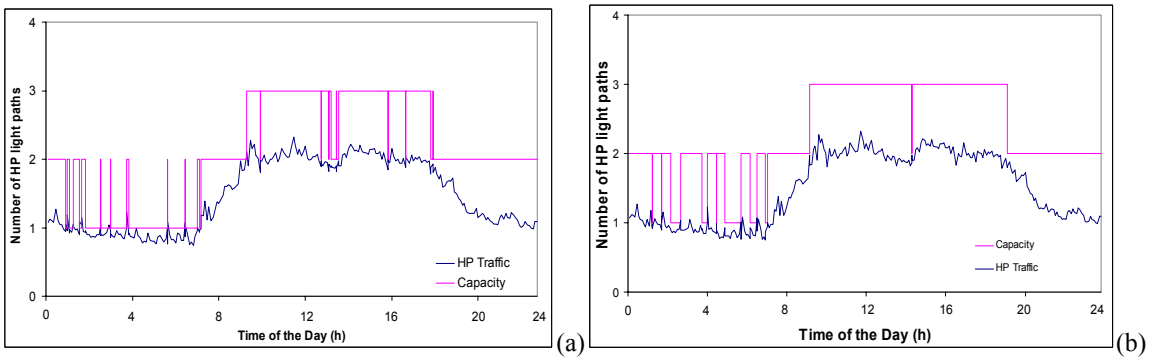


Figure 41: Number of HP light paths, OW = 5 min, conservative approach; (a) n = 1, (b) n =7

The following Table 7 and Table 8 summarize the different simulations results we obtained. OWs of 3 and 5 minutes long were considered respectively. Specifically, it is reported the number of the set up requests, the percentage of time over the simulated one in which the switched connections are required to transport the HP traffic, the mean HT and IAT for the switched light paths and the PLR.

OW = 3 min, TH_{high} = 95 % TH_{low} = 60 %	Number of set up requests	Time percentage of using the first SC[1]	Time percentage of using the second SC[1]	PLR
n = 0	41	76.87%	33.3%	2.4E-03
n = 3	35	82.1%	35.62%	2.16E-03
n = 6	19	90.62%	38.12%	1.59E-03

OW = 3 min, TH_{high} = 95 % TH_{low} = 60 %	Mean Holding Time for the first SC[1] (min)	Mean Holding Time for the second SC[1] (min)	Mean IAT for the first SC[1] (min)	Mean IAT for the second SC[1] (min)
n = 0	41	32.2	22.44	41
n = 3	56.28	39.46	23.4	50
n = 6	93.21	91.5	39.7	117

[1] calculated over the simulated time

Table 7: Summary of results using the conservative approach, OW = 3 min

In particular, we consider the impact of the conservative approach, since we compare the different figures considering different values for the parameter *n*.

OW = 5 min, TH_{high} = 95 % TH_{low} = 60 %	Number of set up requests	Time percentage of using the first SC[1]	Time percentage of using the second SC[1]	PLR
n = 0	27	78.47%	35.07%	2.46E-03
n = 3	14	84.02%	40.62%	1.92E-03
n = 5	12	87.5%	39.58%	1.91E-03
n = 7	9	89.23%	39.58%	1.6E-03

OW = 5 min, TH_{high} = 95 % TH_{low} = 60 %	Holding Time for the first SC[1] (min)	Holding Time for the second SC[1] (min)	IAT for the first SC[1] (min)	IAT for the second SC[1] (min)
n = 0	66.47	45.90	37.81	60
n = 3	108	600	47.72	*
n = 5	140	142	63.75	195
n = 7	142.77	285	63.75	310

[1] calculated over the simulated time
* no samples to calculate the IAT

Table 8: Summary of results using the conservative approach, OW = 5 min

Summarizing, it can be drawn the following conclusions:

1. Increasing the parameter *n* implies reducing the number of requests for the dynamic connections.
2. The higher is the parameter *n* the lower is the experimented PLR.
3. The mean HT and IAT for the SC are compatible with the time required by the CP to establish/tear down a connection.
4. The higher is the parameter *n* the higher is the time percentage in which the SC is used to carry the HP traffic.

On the basis of such simulation results, we suggest to use the TRIDENT procedure implementing the conservative approach since it allows to cope with the PLR requirements imposed by the real-time client applications.

To further reduce the PLR, the TH_{high} can be decreased. As an example, for example, of using Observation Windows 5 minutes long and TH_{high} = 90% of the interface capacity and *n* = 3, the experimented PLR is 1.10 E-03 which is lower than the case reported in Table 8 (1.92 E-03).

4 Traffic Modelling for ASON/GMPLS networks dimensioning

Automatic Switched Optical Networks represent a very promising technology in offering flexibility for bandwidth allocation in future telecommunication networks. However, they open also new perspectives in providing high bandwidth connection services to a wide range of clients. For example, in this context, it can be highlighted that a variety of transport services have to be carried by the same transport network. In particular, as we stated in Chapter 2, the way to provide connections ranges from the permanent (by means the NMS) to the switched (by means the CP).

Concerning the dynamic traffic (that related with the switched transport services), it could probably show various kinds of statistics for the Holding Time (HT) and the Inter-Arrival Time (IAT), depending on the clients.

Recent research activities have been related basically to ASON architectural issues. However, it has to be considered that all the new functionalities introduced by the ASON/GMPLS paradigm need a solid methodological background to be properly adopted.

The evolution towards intelligent optical networks is driven by the growing importance of traffic that requires on-demand switched connections for its efficient transport.

Planning such a novel type of transport network hinges on investigating traffic models suited to represent (at the connection level) the dynamic component of the traffic. In the past, some teletraffic models were developed that describe the user request patterns offered to a circuit-switched network, namely Plain Old Telephone Service (POTS)/Integrated Services Digital Network (ISDN). Indeed, the models available in literature are mainly oriented to the classic telephone networks.

There are some analogies but also some important differences between ASONs and traditional circuit-switched networks. Specifically, an ASON allows the user to dynamically establish connections from one UNI to another in a way similar to that of a telephone conversation. The process to set up, hold and release connections is very similar. But, contrarily to the traditional switched networks, in ASON networks both permanent and switched connections have to coexist simultaneously, and while the permanent connections are provided as leased lines and established by the NMS, the switched connections require signalling between the customer premises and the network, and they require the support of an intelligent control plane.

This part of the Ph.D. Thesis is devoted to present the simulation case study we carried out about the applicability of the classical teletraffic models to dimension the switched part of the ASON networks. It requires the characterization of the switched connections demands statistical distribution (i.e., traffic arrivals/intensity process statistics). The traffic arrival process is generated by applying the triggering demands procedure described in Chapter 3. Specifically, for this case study, we considered the procedure based on the monitoring the average buffer occupancy (ABO).

Our case study consisted in characterizing the traffic arrivals process through its two first statistical moments, namely the mean and the variance. Then, we evaluated by simulation the suitability of the classical teletraffic models for ASON switched part dimensioning purposes.

The organization of this Chapter is as follows: firstly we introduce the most important classical teletraffic models, namely the *Poisson*, the *Engset* and the *Fredericks* models, discussing their characteristics and their applicability to ASON networks environment and then we present in detail the simulation case study we carried out to evaluate their suitability to be used as analytical models to calculate the blocking probability for ASON network dimensioning.

4.1 Introductory notations

Arrival traffic process is characterized by the busy circuits distribution induced by it in an infinite pool of resources [81], [82]. The average traffic intensity, offered by a set of users generating call accommodation requests, is the average number of the observed busy resources, evaluated during a given period of time; its measure unit is *Erlang*.

When a pool of resources is finite, there is the possibility of resource exhaustion; this occurrence is called congestion. There are some ways to measure congestion in a system. Two

important congestion indicators are the blocking congestion (the fraction of call attempts which observes the system busy) and the traffic congestion (the fraction of offered traffic that is not carried and is consequently lost).

The difference between the two indices lies in the fact that the former is based on users call attempts and the latter on the actual offered traffic.

To obtain the traffic arrival process distribution the so-called *moment-matching technique*, commonly used in teletraffic theory to study overflow streams in telephone networks, can be used [83] and [84]. This technique consists of choosing an equivalent process that yields the same moments and whose distribution is easily obtained, and using that equivalent process to obtain the busy circuit distribution. In general, the first two statistical moments are considered, although it represents an approximation of the traffic arrival process [85].

The two moments considered are the average number (A) and the variance (V) of busy circuits. With the first two statistical moments, it can be calculated the peakedness factor (Z) of the traffic arrivals process, since it is defined as the ratio of the variance to the average number of busy circuits (Z = V/A) in an infinite-circuits system.

The peakedness factor is a measure of the deviation of the traffic from Poissonian nature. If Z is greater than 1 the traffic is said to be *peaked* while if it is lower than 1 it is said to be *smoothed*. While peaked traffic is characterized by irregular arrivals (i.e., arrivals in "groups"), *smoothed* traffic is characterised by regular arrivals [86].

Figure 42 presents the simulation results obtained to depict the traffic arrival process characteristics according to the peakedness factor Z. It shows that increasing Z the connection establishment requests arrives in groups.

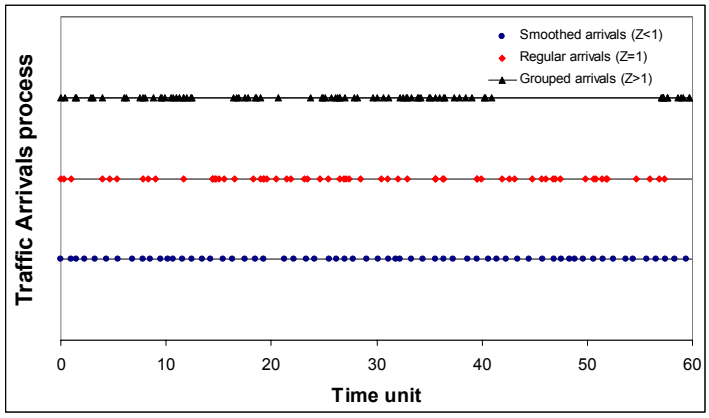


Figure 42: Traffic arrivals process according to the peakedness factor Z

4.2 Classical Teletraffic Models

In the past, some teletraffic models that describe user request patterns offered to a circuit-switched network (typically POTS/ISDN) were developed. The three most significant teletraffic models we considered are the *Poisson*, *Fredericks* and *Engset* models.

In this Subsection, we summarise these basic traffic models that could be used in the ASON context:

- ***Poisson*** model is characterized by exponentially distributed inter-arrival times of connection requests and exponentially distributed holding times. The variance of traffic intensity (V) is equal to the offered traffic (A = HT/IAT) and therefore the peakedness factor is $Z = 1$. Blocking probability experimented by a Poisson process on a set of n resources (circuits) is the well-known Erlang-B formula [86].

$$B^{Erlang}(n, A) = \frac{A^n / n!}{\sum_{k=0}^{n} A^k / k!}$$

  It is possible to demonstrate that the Erlang-B formula does not require exponentially distributed holding time but it is valid for arbitrary holding time distributions.

- ***Fredericks*** model gives an approximate formula for evaluating the congestion probability in a system of n circuits when the offered traffic is not poissonian and it is characterized by mean A and peakedness factor Z.

  The principle is to create equivalence between a system characterized by the tern of parameters (n, A, Z) and a poissonian equivalent system (n/Z, A/Z, 1).

  This assumption can be interpreted as an ideal model composed of a set of Z equivalent processes, each one fed by one of the Z simultaneous requests arriving in a group. Request groups follow a Poisson process. This model transformation is illustrated in Figure 43. Under this hypothesis the expression for the congestion can thus derived from the Erlang-B formula in the following way [87]:

$$B^{Fredericks}(n, A, Z) = B^{Erlang}(n/Z, A/Z)$$

  where n is the number of servers, A is the offered traffic in Erlang and Z is the peakedness factor. The Fredericks model was developed to capture the characteristics of the overflow telephone traffic, when an alternative routing strategy is used in the network. The offered

traffic (that is supposed to be Poisson) tries to get the first routing choice, and if rejected for congestion reasons, is offered to the next choice. The Fredericks approach allows modelling the carried as well as the overflow traffic by its mean and variance. The overflow traffic is shown to be characterized by a Z factor greater than 1, while the opposite (Z <1) for the carried traffic.

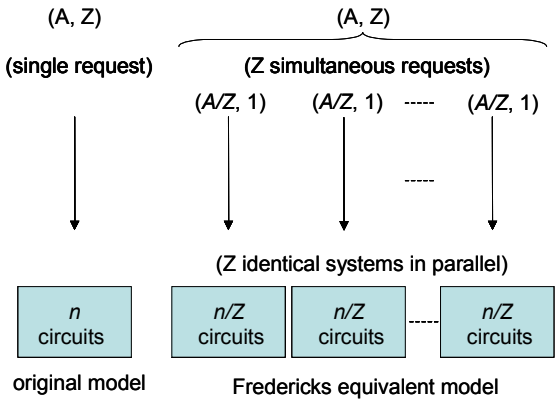


Figure 43: Fredericks model

In general Fredericks model require the evaluation of the Erlang-B formula with a non-integer number of channels and thus it needs for numeric computations an extended version of the Erlang-B formula for a continuous number of channels.

- ***Engset*** model assumes a finite number of sources (S), each requiring one connection at a time. A source may be idle or active (when it has a connection in place): the sojourn time in both states is exponentially distributed. When a source requires a connection, it switches from idle to active state if it finds a free resource otherwise it falls in the idle state again. The blocking congestion for the Engset model is given by the following formula [87]:

$$B^{Engset}(n,S) = \frac{\binom{S-1}{n}\beta^n}{\sum_{k=0}^{n}\binom{S-1}{k}\beta^k}$$

where $\beta = \alpha/(1-\alpha)$ is the so-called offered traffic per idle source and α is the offered traffic per source (the carried traffic if the sources were never blocked); $A = S\cdot\alpha$ is the total offered traffic. The peakedness factor is given by $Z = 1-\alpha$.

4.3 Applicability of the teletraffic models to the ASON networks

The models above summarized can be applied in an ASON context with the following assumptions:

- Each user can request just one optical channel at a time.
- α = A/S is the offered traffic per user and it indicates the ratio of the time in which the user is using an optical channel.
- *n* is the number of circuits required on a specific trunk.
- Z is the degree of simultaneity of user requests.

As a matter of fact, in an ASON network, switched connections are likely to present dynamics substantially different from those observed on a traditional telephone network. Due to their huge bandwidth (typical bandwidth for optical channels is 2.5 or 10 Gbit/s), these connections are likely to be required by very demanding services like network-to-network interconnections, big business customers, etc.

The variability associated to these kinds of services in terms of IATs and HTs is far from being known and we cannot therefore test the conformity of the models with the real behaviour. The behaviour of people calling on a phone network usually fits very well with a model of IAT and HT exponentially distributed, while potential customers that require connections to an ASON could show a very heterogeneous behaviour in terms of IAT and HT. This is because traffic sources in ASON networks could be residential users that require huge bandwidth connections for broadband applications but more realistically, in a medium term perspective, they could be ISPs, enterprises, branch offices of banks or some other kind of business customers that need bandwidth on demand. In addition, the behaviour could be very different depending on the type of customer. As a consequence, the offered traffic could not be poissonian in many cases. Thus, studies on statistical behaviour of customers of ASON will be necessary when such networks and related customer applications are available.

Our intention is to make a survey of the available, analytically tractable models, and verify what kind of behaviour they are able to catch and taken into account for sizing the network resources. Then, by applying these models for a single link dimensioning, we investigate, by simulation, their applicability for ASON dimensioning purposes.

4.4 Suitability of classical teletraffic theory for ASON network dimensioning: Simulation case study

As a comparison among the above-discussed models, in Figure 44, a graph depicting the number of circuits required using these models is reported, assuming a blocking target probability of 0.1% and a population of 100 users. All the curves are drawn as a function of the traffic specific for each user. The first two curves (starting from below) refer respectively to the Engset model and to the Erlang (Poisson) model while the rest refers to the Fredericks model calculated for various values of peakedness factor.

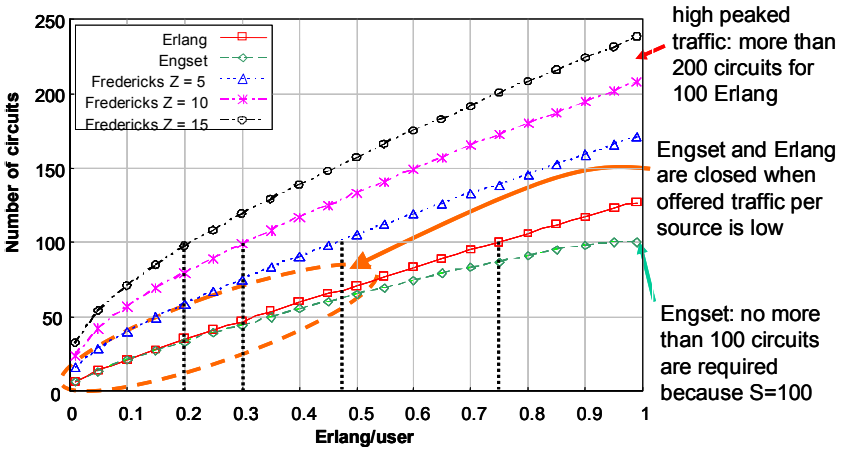


Figure 44: Number of circuits required as a function of the traffic intensity per user

It can be seen that the differences between Erlang and the Engset models are quite small. In particular, Engset model gives values always less than 100 because it is accounting for the finite number of users while the Erlang model starts to exhibit values greater than 100 for traffic of 76 Erlang or more, which corresponds to a 0.76 Erlang/user. For values lower than 0.2 Erlang/user the two models give the same results, while for higher values the Erlang model gives greater values than the Engset model in such a way that the dimensioning approach that uses the Erlang model is always conservative.

Regarding the Fredericks model, the case Z = 1 is equivalent to the Erlang case while for higher values of Z, when increasing the simultaneity of the users requests, the number of required circuits increases in order to maintain the assigned level of performance.

The intersection of these curves with the horizontal line (corresponding to a value of 100 circuits) is emphasized in the graph. Under these types of traffic, the points of intersection represent the break-even points between a full resource allocation and a statistical allocation planning strategy. For traffic values lower than the identified thresholds, some statistical gain can be

achieved, while for higher values the reduction of just one circuit with respect to the allocation of one circuit per user would degrade the performance to unacceptable values.

The presented models have the advantage to be simple general purposes models. Especially in the case of systems loaded by high number of sources the Fredericks model could be flexibly used to represent the traffic by only two parameters, the offered traffic and the peakedness factor. This is an appreciable advantage because no additional information on statistical distributions is required. The drawback concerns the accuracy of the model that cannot be very high as it is an approximated model that does not take into account distributions but only the first two moments of the traffic intensity.

Assuming the above summarized as reference models, we investigated whether such models can be effectively used as "*black box*" models with the aim to estimate the congestion probability in a simulated system reproducing connection requests to an ASON network. Such models could be taken into consideration as a tool for dimensioning ASON network resources and employed in network planning processes.

In order to obtain numerical results on the applicability of the three above-mentioned teletraffic models to ASONs, we have evaluated a simulation case study in a simple IP over ASON/GMPLS scenario. It consists of an IP router collecting the traffic from different IP networks (e.g., ISPs) on top of an optical switch (OXC) provided with a control plane (Figure 45).

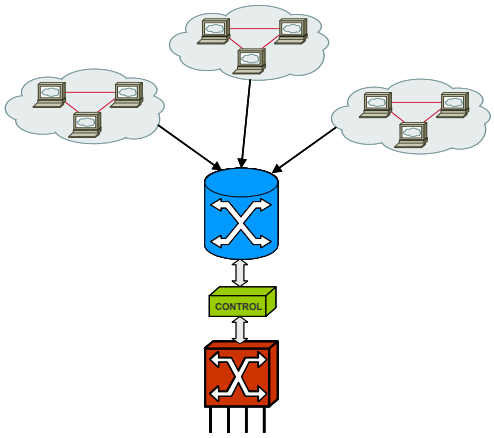


Figure 45: Simulated scenario, IP router on top of an OXC

The control function in Figure 45 implements the procedure (hereafter triggering mechanism) based on the monitoring the average buffer occupancy (ABO) of the IP router interfaces (described

in Chapter 3), which triggers the requests for switched connections on the basis of the ABO monitoring.

The aim of this simulation case study was to study the suitability of the classical teletraffic models presented in the previous Section to be used as analytical models for dimensioning the OXC outgoing trunk (single link dimensioning). The simulation conditions are the same of those described in Section 3.2.

Specifically, three steps composed our study, namely:

Step 1: Running simulations for different values of the OW and triggering thresholds in order to characterize the switched channels demand statistics. Both the case of a single threshold and of different thresholds for set up and tear down triggering were considered. The number of available channels on the OXC trunks was assumed big enough to avoid any connection request blocking. As a result, we obtained the statistical distribution for IAT and HT, which allowed us to compute the total offered traffic intensity A as $\sum_i (HT_i/IAT_i)$ for any channel *i* that was switched on and off during the simulation time. Moreover, we obtained the peakedness factor Z as V/A, where V is the variance of the number of channels that were switched on and off during the simulation time. The results in Table 9 show that the peakedness factor obtained in all the cases is lower than one.

OW (s)	**TH_{high}** (KB)	**TH_{low}** (KB)	**Traffic Intensity** (A)	**Peakedness Factor** (Z)
60	4	4	2.483	0.645
60	50	50	2.653	0.721
60	50	4	1.894	0.792
60	450	50	2.234	0.680
30	4	4	2.507	0.763
30	50	4	2.038	0.731
10	4	4	3.49	0.735
10	50	4	2.31	0.837

Table 9: A and Z obtained for different configurations of the triggering mechanism

This suggests that the connection requests traffic generated by the particular above-mentioned triggering mechanism is smoothed, which discards a Poissonian assumption (we recall that the Erlang model is characterized by a peakedness factor of 1). Therefore, the Fredericks and the Engset models are potentially suitable to model the traffic arrivals process.

Step 2: Calculating the analytical blocking congestion probability applying the teletraffic models (Erlang, Fredericks and Engset) as a function of the offered traffic A and peakedness factor

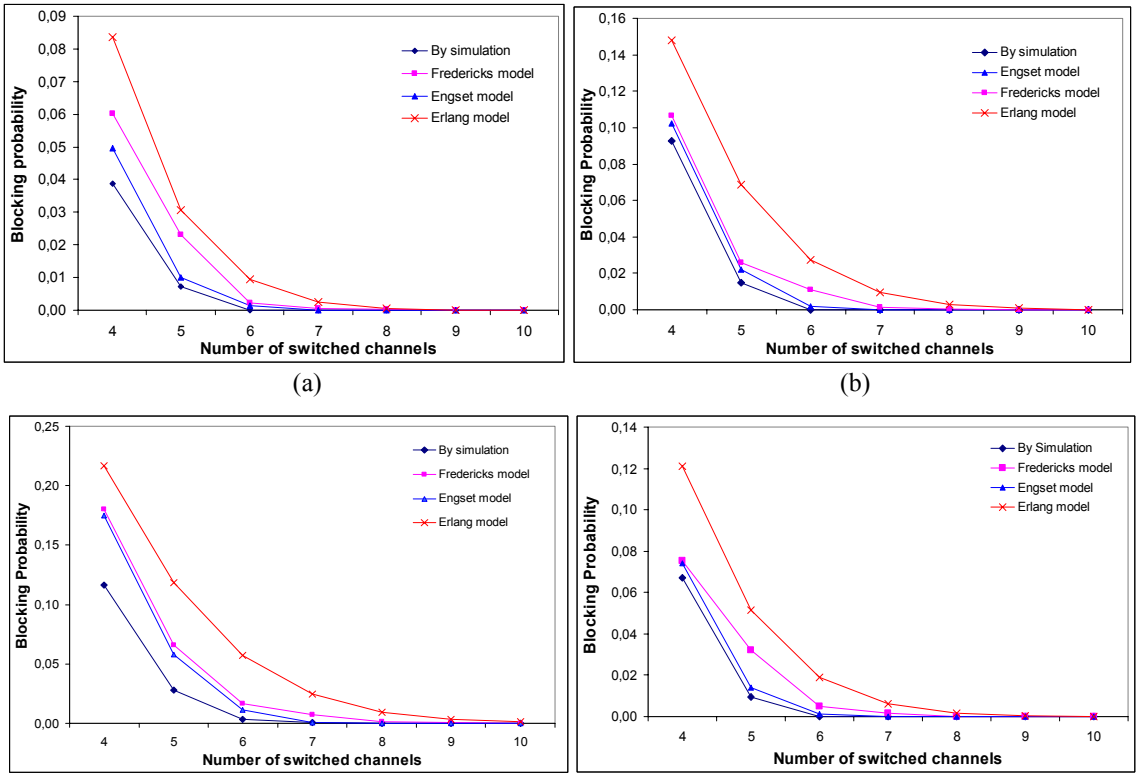
Z obtained by simulation in the first step and assuming a limited number of available switched channels. In particular, for the Engset model, an "equivalent number of sources" S was obtained by applying the values of A and Z to the formula $S = \frac{A}{1-Z}$, and the offered traffic per source was obtained by applying Z to the formula $\alpha = 1 - Z$. The blocking probability was then computed using the Engset formula given in Section 4.2.

The physical meaning of the traffic sources (S) is not clear because the model is applied in a "black box" way. There is no set of users generating directly the traffic process that loads the system, but the traffic process at connection level is the result of the packet arrival process, filtered by the triggering mechanism. Nevertheless the notion of "equivalent" number of sources could be retained as the index of the potential number of channels the aggregated traffic could require if they offer traffic with their potential maximum intensity.

Step 3: Comparison of the figures obtained in step 2 with the actual blocking probability obtained by simulation, in which the number of available channels is assigned in such a way that the blocking congestion assumed significant values. Specifically, the applicability of the Poisson, Fredericks and Engset models was then evaluated by comparing their blocking probability (calculated on the basis of the IAT and HT results obtained in the first set of simulations) to the blocking probability obtained trough the second set of simulations.

Figure 46 shows, for the various configurations simulated, that the Poisson model is not suited to characterize the optical switched connection demand generated by the IP traffic filtered by applying the designed procedure to trigger requests for switched connections. In fact, it experiences a significantly higher blocking probability compared to the results obtained through the simulation. In contrast, the Engset and Fredericks model provide only a slightly higher blocking probability than the simulated case, i.e., slightly overestimate the number of optical channels that have to be provisioned in order to guarantee the required Grade of Service (GoS).

The Engset model is the one that best approximates the simulated traffic pattern, and could thus be chosen as candidate to dimension an ASON that has to cope with the traffic offered by an IP client network.



(a) OW = 1 minute, TH_{high}/TH_{low} threshold = 4 KB

(b) OW = 1 minute, TH_{high} = 50 KB, TH_{low} = 4 KB

(c) OW = 1 minute, TH_{high}/TH_{low} threshold = 50 KB

(c) OW = 1 minute, TH_{high} = 50 KB, TH_{low} = 450 KB

Figure 46: Blocking probability as a function of the number of switched channels available

However, it has to be underlined that this is a simple case study and that the simulation results are strongly influenced, on one hand, by the offered IP traffic and, on the other, by the parameters of the triggering mechanism (e.g., OW, TH_{high} and TH_{low} thresholds).

Moreover, it is worth mentioning that the suitability of classical teletraffic models in the analysed scenario is considered under particular conditions. In fact, the model identification has been done for a small range of model parameter values (offered traffic about 2 Erlang and peakedness between 0.65 and 0.84), while in telephone networks, where such models have been successfully applied since the beginning of the past century, it is of the order of tens or hundreds of Erlangs. Anyway, it has to be taken into account that probably just few channels will be requested automatically in an ASON environment.

Summarizing, as a minor contribution of this Ph.D. Thesis, we suggest an approximated method for modelling and dimensioning the switched part of ASON networks, which consists of:

1. Characterizing the connection demand statistics generated by applying the triggering request mechanism, by means of the average traffic intensity (A) and the peakedness factor (Z).

2. Applying the resulting values of A and *Z* to compute the well know Engset model to obtain the number of switched channels required to cope with a given GoS.

PART II: Multi-layer TE in Metropolitan Area Networks

Traffic Engineering is concerned with the performance and resources optimization in response to dynamic traffic demands and/or node and link failures [41].

In this second part of the Ph.D. Thesis, we concentrate on the design and evaluation of efficient resilience strategies to face with different failure scenarios which can occur in the networks. Specifically, we concentrate on an IP/MPLS over Resilient Packet Ring (RPR) over optical transport network scenario for metropolitan area networks. We present the DHOT coordination approach between the protection mechanisms provided by the RPR technology and the ones defined in the optical layer. The novelty of the suggested approach relies on the required interworking between the RPR and the optical layer.

Firstly we introduce the multi-layer resilience concept, highlighting the related work on this topic. Then, we describe the strength and weakness of the recent standardized packet-based RPR technology and finally it is presented the DHOT approach, discussing its characteristics and its performance evaluation. As a further contribution, an interworking strategy between RPR and the optical layer, based on taking benefits from the automatic switching of optical connections capability to optimize the RPR bandwidth utilization, is also presented and discussed.

5 Multi-layer Resilience

The failures occurrence produces a strong impact on network performance. In fact, when failures occur in the network, there is the need of rerouting/switching the affected traffic along alternative routes. For example, this possibly implies longer routes and therefore higher propagation delays and thus higher end-to-end delays. For time-sensitive applications (e.g., multimedia applications, voice over IP), this is an important factor in assessing the performance of the recovery strategy.

In case of failures, the optimization of the utilization of the network resources is very critical. Indeed, network survivability/recovery/resilience, namely the capability of the network to recover traffic affected by failures, has become of vital importance in current and next generation networks [25], [26], [88] and [89]. With the growth of data traffic that has to be transported, the networks need to be very robust to face the different kinds of failures that can occur. Therefore, network operators have to take special precautions in order to do networks survivable. Since it is difficult to prevent failures in the network infrastructure (equipment failures, cable breaks, etc.) the network design objective is to maintain services availability even under failure conditions. In order to make the networks more reliable, they have to be reconfigurable and such reconfiguration has to be fast in order to minimize the traffic lost and services interruptions. At the same time, the recovery strategies have to optimize the utilisation of the network resources while, from the Network Operator's point of view, they should not increase too much the network cost (both CAPEX and OPEX).

In the case of metropolitan area networks, the emerging Resilient Packet Ring (RPR) technology provides powerful protection mechanisms which minimize the time needed to restore

the traffic and, thus, the traffic lost due to the failures. RPR recovery mechanisms do not need to pre-allocate spare network resources to be used in case of failures.

This second part of the Ph.D. Thesis deals with the design and performance evaluation of multi-layer resilience strategies to be used in an IP/MPLS over RPR over intelligent optical networks.

5.1 Network Survivability

With the exponential growth of the traffic that has to be transported, network resilience has gained a critical role in the design of telecommunication networks. The failures can occur at different layers composing the network architecture and it is very important to decide at which layer the recovery is implemented. This is due to the fact that, for example, lower layers could not aware of failure occurred at the higher layers. The recovery mechanisms have to be designed basically with the aim to be simple from the implementation point of view, to minimize the traffic losses due to the failures and finally to optimize the utilization of the network resources, reaching TE objectives.

Among others, some of the most common failures that have to be taken into account in order to design a survivable network are [25]:

- **Node Failure:** A node can fail because of different reasons such as power down or heat. Thus, the connections that use that node are interrupted and the communications to the neighbouring nodes are lost.
- **Link Failure:** Link failure is a common failure, which interrupts the communication between two neighbouring nodes. The failure of a link can be detected at several layers.
- **Router Failure:** This failure occurs at the IP layer and therefore the detection and recovery is done at this layer. This failure can be detected by timeouts and then, a backup router replaces the failed router.
- **Optical Path failure:** Optical path failure interrupts the communication between the sending and receiving nodes of a light path. This failure can be caused by a bad functioning of lasers, by a bad-established switch connection, etc. This failure affects one optical path and therefore, only the data belonging to this interrupted path have to be restored.

To make a network survivable, two approaches can be used, namely the protection and restoration approaches [90], [3]. The first one is much quicker than the second one and it implies the use of fixed, pre-calculated routes and pre-allocated spare capacity, eventually used to transport

low priority traffic. Specifically, in point-to-point links, basically two types of protection mechanisms are used namely 1+1 and 1:1 (more generally 1:N). In 1+1 protection, traffic is sent simultaneously on two different physically disjoint paths between source and destination nodes; one of the path is called the working path while the second is called the protection (recovery) path. The destination node, in absence of failures, selects one of the two paths for reception. In the case the failure of that path is produced, the receptor node switches to the other path. In such a mechanism, no signalling between nodes to recovery from the failure is required. In 1:1 protection, two paths are available between source and destination nodes. In this case, the data are sent only through the working path. In case a failure occurs, both the source node and the destination node switch to the other pre-defined path. In this case, a signalling mechanism is required between the source and the destination node. However, the advantage is that in absence of failure, the second path can be used to transmit, for example, low priority traffic.

On the other hand, the restoration approach implies the rerouting of the affected traffic calculating an alternative route once the failure has occurred. It is based on using any available capacity in the network. The restoration mechanisms are much more efficient that protection in terms of network resource utilization since no spare resources are needed to be pre-allocated for recovery purposes; however, since the affected traffic has to be rerouted, this leads to higher recovery times, which is the time needed to recover from the failure [3].

In the current multi-layers networks, each layer (e.g., IP and SONET/SDH) has its own protection mechanism built in, independent from the other layers [25], [91]. Reliability basically relies on protection at the SONET/SDH network layer. Indeed, different protection mechanisms have been designed for survivable SONET/SDH networks that allow fast recovery within the target of 50 ms [92], [93]. Nevertheless, SONET/SDH protection is mainly limited to ring topologies and it is not able to distinguish between different priorities of traffic and it has not vision of higher layer failures. On the other hand, the IP layer has limited recovery functionalities (i.e., rerouting). The routing protocols can reroute the traffic in case of failures, but the time needed for the routing algorithms to re-converge is in the order of seconds. Thus, this rerouting time compared to the 50 ms of the SONET/SDH is extremely poor, especially in the case of real time applications. MultiProtocol Label Switching (MPLS) technology can be used to enhance the survivability of IP networks. Basically, mechanisms defined for protection in MPLS rely on pre-established protection LSPs, used as backups for the working LSPs, achieving better protection switching times than IP networks [94]. The backup (or alternative) LSPs are set-up (signalled) at the moment the failure is

detected by the IP/MPLS router. These mechanisms rely on the control plane functionalities of MPLS [31].

5.2 Multi-layer Resilience: Related Work

As before mentioned, in the current multi-layers networks each layer has its own protection mechanism built in. Various technologies at different layers may provide protection and/or restoration capabilities at different temporal granularities (i.e., in terms of time scales) and at different bandwidth granularity (from packet-level to circuit level) [90]. The recovery actions rely on a single-layer strategy, which means that a single layer takes the responsibility to recovery from the failure. These are taken either in the lowest (bottom) layer or in the highest (top) layer. The resilience single-layer strategy is very simple from the implementation point of view. Its major drawback concerns that it may not be able to recover the network from all kind of failures that can occur within the network [26]. Moreover, deciding at which layer the recovery actions have to be carried out is very challenging.

A more efficient approach, consisting on to combine recovery mechanisms in more than one layer, has been proposed in [88]. Recovery at multiple layers (multi-layer resilience) increases the reliability of the multi-layer networks, since the network is very robust to a wider range of failures scenario. It is beneficial in order to avoid contention between the different single-layer recovery schemes and it takes benefits from the advantages of each layer recovery mechanism.

Indeed, the definition of multi-layer resilience approaches/strategies leads to decrease the investments costs required to ensure a certain survivability target and leads to overall better utilization of the network resources after the network reconfiguration [88].

Generally speaking, to evaluate the effectiveness of multi-layer resilience strategies, an important issue deals with the definition of their actual performance parameters. In [25], the authors defined the cost, the complexity and the feasibility of the recovery strategy.

The cost is strictly dependent on the properties of the used resilience techniques such as the extra resources (spare resources) required to enable the recovery actions. This is closely related to the planning of the network, which must enable the provisioning of network connections throughout occurring network failures.

Aside from the actual complexity of the resilience mechanisms, routes have to be calculated and set up, and this may increase the complexity even though the resilience protocol itself can be fairly simple (like for example the protection mechanisms).

Finally, feasibility is related to complexity, but is to be considered on a more general level and considers whether a resilience strategy is feasible/achievable to be applied.

When considering multi-layer resilience, the simplest way to implement it is to run the different mechanisms in parallel and independently from each other; it is called the *uncoordinated approach* [88], [94]. As each layer detects the failure, it starts to run its own recovery mechanism.

Such solution is very simple from the implementation point of view since no standardization of coordination signals between layers is required. It is also simple from the operational point of view. The most important drawback is that multiple layers can start the recovery action contemporarily leading in such a way to potential networks instability (above all at the higher network layers) and unnecessary reduction of the overall available bandwidth. Thus, coordination among the different recovery mechanisms is required.

A way for the coordination of the recovery mechanisms relies on using the so-called *sequential approach*, namely one layer tries to restore the traffic and the following layer only takes over if the current layer does not succeed to recover the affected traffic.

Specifically, two sequential approaches have been proposed in [89], namely:

- ***The bottom-up approach***: The lower layer starts the recovery actions. In the case that it cannot restore all the traffic, then higher layer actions are triggered.

- ***The top-down approach***: Recovery actions are initiated at the top/highest possible layer and only if the higher layer cannot restore all traffic, lower layer actions are triggered. An advantage of this approach is that a higher layer can more easily differentiate traffic with respect to the service classes (service-based recovery) and thus it may try to recover high priority traffic first and then try to restore low priority traffic. A major drawback is that a lower layer may not detect whether a higher layer is able to restore traffic or not and thus an explicit signal between the layers is required for this purpose.

The implementation of these multi-layer resilience strategies implies the need to define some rules to coordinate the recovery actions between the different network layers. Authors in [94], [95] and [96] have proposed the following three different rules.

The first one is based on the hold-off timer concept. It can be applied both for the bottom-up and top-down approaches. Specifically, a hold-off timer is set at the moment the server (client) layer starts attempting to restore the traffic. If this hold-off timer expires and the traffic is not restored, then the client (server) layer will take over the recovery actions while the server (client) layer ceases its attempts. It does not require any interworking between layers. Therefore, it is probably less appropriate for the top-down approach, since the lower layer should be notified with an explicit signal whether the higher layer managed to restore the traffic or not. The main drawback of a hold-off timer is that higher (lower) layer recovery actions are always delayed, independent of the failure scenario.

To overcome this delay, a second strategy, based on using a recovery token signal between layers, has been proposed [26], [94]. It is based on the fact that the server layer sends the recovery token, by means of an explicit signal, to the client layer from the moment that it knows that it cannot restore traffic anymore. Contrarily to the hold-off timer approach, it requires the interworking between the layers involved in the recovery. Moreover, when comparing the recovery token approach with the hold-off timer one, its major drawback consists on that a recovery token signal needs to be incorporated in the standardization of the interface between network layers. Therefore, even though its complexity is rather low, its feasibility is rather high.

Finally, in [94], a third possible strategy, namely the integrated approach, has been proposed. It is based on a single integrated multi-layer recovery scheme. This implies that this has a full overview of all the network layers and that it can decide when and in which layer (or layers) to take the appropriate recovery actions. Although it is the most flexible coordination mechanism, the major issue is its implementation. It is unlikely to develop a single recovery scheme, controlling and having an overview of all network layers, in current overlay-based networks. The integrated approach might represent an interesting solution if a peer-to-peer network model is used.

5.3 Problem addressed

The second part of this Ph.D. Thesis is related with the design and evaluation of a multi-layer resilience strategy for metropolitan area networks. Specifically it takes benefits from the characteristics of the recovery mechanisms of the emerging Resilient Packet Rings technology and

the ones designed for the optical layer. The proposed strategy/mechanism is based on the definition of interworking rules between the RPR layer and the optical layer with the aim to recovery faster from failures while achieving the optimization of the utilization of the network resources, reaching in such a way traffic engineering objectives. Specifically, firstly, the double hold-off timer approach, which improves the hold-off timer one, is presented and its performance evaluation discussed.

Secondly, if the failure is recovered at the RPR layer, this recovery action implies the substantial reduction of the bandwidth available at the logical level. To avoid such bandwidth reduction, the automatic provisioning of connections capability provided by the intelligent optical layer (i.e., ASON/GMPLS networks) is used. Specifically, we present and discuss a procedure based on monitoring the traffic load carried by the light path connecting two IP/RPR routers and, on the basis of such monitoring, the request to the GMPLS-based control plane for a switched connection to be used temporarily as additional light path connecting the routers is triggered.

6 Multi-layer Recovery Strategy in RPR over Optical Transport Networks

This Chapter is devoted to describe the suggested resilience interworking strategies to be applied in an IP over Resilient Packet Ring over intelligent optical transport networks. Firstly, we discuss the RPR technology which is an emerging packet-based transport technology recently standardized by the IEEE. Specifically, we concentrate on the resilience mechanisms provided by RPR highlighting their advantages and drawbacks. Secondly, the recovery at the optical layer is discussed. Then, the DHOT approach, which is based on the interworking between the RPR and the optical layers is presented and evaluated. Finally, as a further contribution of this Ph.D. Thesis, a procedure based on the use of the automatic switching of optical connections capability of the ASON networks to improve the bandwidth utilization of the RPR rings in case of failures or to respond to the client traffic fluctuations is presented and discussed.

6.1 Resilient Packet Ring technology

Resilient Packet Ring (RPR) is a new packet-based transport technology for ring-based metropolitan area networks. It includes a new MAC layer technology as well as powerful automatic recovery mechanisms. Indeed, RPR systems are seen as the successors to SONET/SDH ADM-based rings for the efficient delivery of IP-based data traffic achieving both the optimization of the bandwidth utilization. RPR technology has been recently standardized by the Institute of Electrical and Electronics Engineers (IEEE) as IEEE 802.17 RPR [97].

Many legacy metropolitan networks use a physical ring structure. It is a natural environment for the SONET/SDH networks that constitute the bulk of current metropolitan network infrastructure. SONET/SDH networks, however, as we discussed in Chapter 2, was designed and optimized for

point-to-point circuit-switched services such as voice services. On the other hand, for metropolitan environments, Ethernet technology may offer a simpler and cost-effective solution for the transport of the data traffic. However, because Ethernet is optimized for point-to-point or meshed topologies, its use of the available bandwidth is inefficient and it does not take advantage of the ring topology in order to implement fast protection mechanisms.

RPR technology fills this gap by acting as a multi-service transport protocol based on packets rather than circuits. While Resilient Packet Ring paradigm may provide performance-monitoring features similar to those of SONET/SDH, it maintains the advantages of the Ethernet technology such as low equipment cost, high bandwidth granularity and statistical multiplexing capability. The IEEE 802.17 RPR standard defines a set of protocols for detecting and initializing the shared ring configuration, recovery from failures and regulating the fair access to the shared medium.

For Carriers, RPR promises to deliver all the necessary end-user services, such as TDM voice, Virtual Private Networks, data and Internet access, at dramatically lower equipment, facility and operating costs. It is a very promising transport technology, since most of the major carriers and vendors (among others Cisco Systems, Luminous and Nortel Networks) have actively participated in the IEEE 802.17 standardization process and have shown much interest in the evolution of the standard [98]. In fact, the unique features of RPR were sufficiently interesting to trigger many pre-standard installations by important players in the telecommunications market (e.g., Cisco Systems, Nortel Networks, Sprint, Luminous, Bell Canada, Worldcom and SUNET). The first major pre-IEEE 802.17 RPR standard deployments RPR technology-based networks introduced by Sprint in 1999 and Macedonia Telecom as well as China Telecom in 2001.

Ring topology based on RPR is also studied by ITU-T and a preliminary version of Recommendation on Multiple Services Ring (X.msr) is available [99]. Table 10 summarizes the most significant differences between X.msr and IEEE 802.17 RPR.

Feature	ITU-T (X.msr)	IEEE 802.17 (RPR)
Topology	Two counter-rotating rings, max. 32 stations	N×dual counter-rotating rings, max. 256 stations
MAC address	Local with fixed addresses (4 octets) – possibly IP address	Globally unique MAC address (6 octets)
MAC transit	Unspecified buffer, 8 priorities	Single or dual buffers, 2 priorities
Protection	Wrapping	Wrapping and steering
Spatial reuse	Supported	Supported
Fairness	Not necessary (pre-planned bandwidth)	Fairness algorithm for unprovisioned traffic (Class C)
Multicast	Supported	Supported

Table 10: Significant differences between X.msr and IEEE 802.17 RPR

The introduction of RPR-based metropolitan networks is gaining importance and this emerging technology represents a very promising networking solution to transport data traffic in a short/medium term. In fact, the IEEE 802.17 RPR standard has been approved in June 2004 and thus, network equipments standard-compliant will not be available before the beginning of 2005. As a consequence, the earliest deployment of IEEE 802.17 RPR networks will be in the timeframe of two or three years. As concerns different geographical areas' readiness to implement RPR technology, we believe that Asia (mainly China) seems to be in first position, followed by the United States of America (USA). In Europe, the prospects for RPR seem not to be not particularly promising. This conclusion is based on the current deployment of IEEE 802.17 RPR pre-standard technology such as DPT-based products from Cisco Systems and OPTera Packet Edge Systems series 3000 from Nortel Networks. China is currently a good market for RPR products because there is not a great deal of SONET/SDH infrastructure installed, and which thus opens the market to new and more efficient technologies (China Netcom already deployed Luminous' RPR-Based Metro Platform in several cities in 2002). Pre-RPR systems, such as OPTera-3000, are better positioned in the USA than in Europe, simply because OPTera 3000 is ready for SONET and not for SDH. Indeed, it is considered able to provide increased revenues for carries to transport data traffic in metropolitan environments.

Summarizing, in comparison with the currently enabled technology in the metropolitan environment, which RPR pursues to substitute (i.e., SONET/SDH rings), the RPR technology exhibits the following advantages: 1) Overcome of the limitations that TDM/circuit-based architectures impose on data communications allowing direct connectivity without circuit provisioning; 2) Increase of the bandwidth efficiency by implementing the spatial reuse of bandwidth and the statistical multiplexing of packets; 3) Enabling service integration by supporting various traffic priorities and 4) Reduction costs and complexity by eliminating intermediate layers between the IP layer and the optical layer.

6.1.1 Fundamentals of RPR technology

RPR-based networks enable efficient transfer of data traffic as well as fast protection mechanisms. It is a standard which consists of a superset of features derived from various proprietary solutions such as Cisco Systems' DPT [100] and Nortel Networks' Optera Packet Edge System [108].

Network Operators claim that the functionalities of RPR and its implementation in real commercial environments present many advantages, namely:

- Advanced protection mechanisms,
- Distributed control,
- Interoperability with major transmission standards,
- Scalability in speed and number of nodes,
- Plug-and-play operation,
- Performance monitoring capabilities,
- Support for a limited number of priorities (two or three),
- OAM and advanced traffic and bandwidth management,
- Support for unicast, multicast and broadcast data traffic.

It has been designed to operate over a variety of physical layers, including SONET/SDH, Gigabit Ethernet (IEEE 802.3ab), DWDM and dark fibre, and is expected to work over higher-speed physical layers. The minimum supported data rate is 155 Mb/s.

RPR networks are based on two symmetric counter-rotating rings (external and internal ring) that carry data and control information (Figure 47). The nodes/stations may send data on either of the two ringlets. Basically, the shortest path to the destination is used. Therefore, the nodes use a topology discovery protocol to obtain a topology map of the ring, which is then used for the shortest path computation.

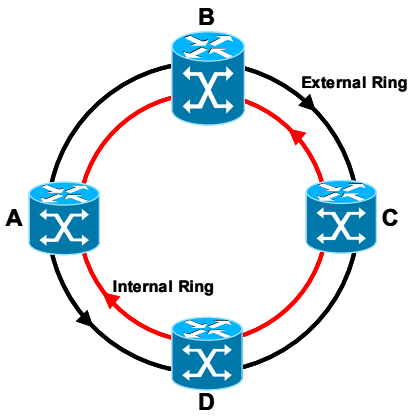


Figure 47:4-nodes Resilient Packet Ring network

An important feature of RPR technology is the spatial reuse, which increases the overall aggregate bandwidth of the ring. Unicast frames are removed from the ring at their destination,

which means that they only occupy bandwidth on the links from source to destination. This is in contrast to earlier techniques, such as Token Ring, in which each frame had to traverse the whole ring so that spatial reuse could not be exploited.

The IEEE 802.17 RPR standard supports three types of services, namely Class A (high priority), Class B (medium priority) and Class C (low priority) [97]. The Class A service is designed to support real-time applications that require a guaranteed bandwidth and low jitter. This service has absolute priority over the other types of services, and must be shaped at the ingress. A token bucket shaper (Shape A on Figure 48) is provided to ensure that the client traffic does not exceed the allocated rate. Each node/station advertises the amount of bandwidth it needs for its Class A service. This allows calculating how much bandwidth is reserved for Class A in the ring and how much is left unreserved for Class B and C services. Traffic above the allocated rate is rejected.

The Class B service is dedicated to near real-time applications that are less delay-sensitive but that still require some bandwidth guarantees. It provides guaranteed information transfer at the Committed Information Rate (CIR) and best-effort transfer for excess traffic (beyond the committed rate). In contrast to Class A, the bandwidth for Class B CIR traffic is not statically allocated. In the presence of congestion, the node sends messages that throttle Class C transmissions from other stations to leave bandwidth for its Class B traffic.

The Class C service implements the best-effort traffic class. This service is subject to weighted fairness mechanisms, which ensure that each station gets its fair share of the bandwidth available. The traffic is shaped by the IEEE 802.17 RPR Medium Access Control (MAC), which uses a token bucket shaper. A fairness mechanism decides on the amount of bandwidth each station may currently use for its Class C transmission. The calculation involves determining the amount of Class A and B traffic present in the ring and divides the remaining bandwidth in proportion to administratively configured node weights.

The allocated rates for Class A and Class B services and node weights for Class C are configured in each station by a provisioning mechanism.

Each node has two MAC datapath entities, one for each ringlet (in the IEEE 802.17 RPR standard, the external ring is called Ringlet 0 while the internal ring is called Ringlet 1). Figure 48 presents an example of a three-node IEEE 802.17 RPR ring and a more detailed view of the MAC datapath entity. Specifically, it shows that a frame received from the IEEE 802.17 RPR ring is checked against bit errors and time-to-live expiration. Once this is performed, a filter module decides whether the frame should be copied to the client (in case of multicast traffic), passed to the

control sublayer or neither. The adjust function is responsible for stripping frames from the ring, adjusting frame fields (e.g. the time-to-live field) and placing the frame in the correct transit queue. The node described has two transit queues, one for Class A service and the other for classes B and C. An alternative implementation is characterized by a single transit queue.

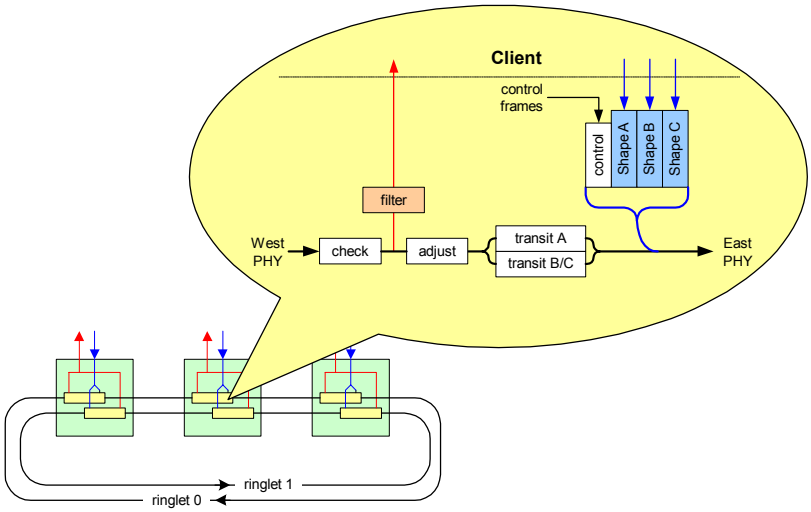


Figure 48: A three-node IEEE 802.17 RPR ring with a simplified structure of the MAC datapath entity

More than one frame may be ready for transmission at a given moment. Two transit buffers, the control queue and three queues corresponding to Class A, B and C local services may simultaneously hold a frame ready to be transmitted. A set of precedence rules is defined to maintain traffic priorities and to avoid loss of frames in transit [97].

We carried out a simulation case study to evaluate the performance of the RPR MAC protocol. Specifically, a RPR ring composed by 5 nodes/stations was simulated. The traffic inserted by each node is uniformly distributed among the rest of the nodes. Specifically, 20% of the generated traffic by each node corresponds to Class A (i.e. high quality video traffic), other 20% to Class B (high priority IP traffic) and the rest to Class C (low priority IP traffic). Figure 49 depicts the throughput of the ring when increasing the offered load (ρ), being the offered load the ratio between the offered traffic and the maximum network throughput ($Throughput_{Max}$). Since the network is assumed error-free and no packets are lost, network throughput and offered traffic are equal until saturation. This occurs when the utilization factor of each link between nodes is equal to 1. It can be calculated that the $Throughput_{Max}$ is equal to $8 \cdot R_b/(n+1)$, where R_b is the RPR bit rate interface and n is the number of nodes composing the ring. In Figure 49, being $n = 5$ and $R_b = 2.5$ Gbps, the $Throughput_{Max}$ is equal to 16.67 Gbit/s.

In extreme load traffic conditions (i.e., the ring is saturated), the network is not able to support all the offered traffic. Thus, RPR MAC protocol reacts and thus the throughput of Class C decreases, while the throughput of the Class A and B still continue to increase.

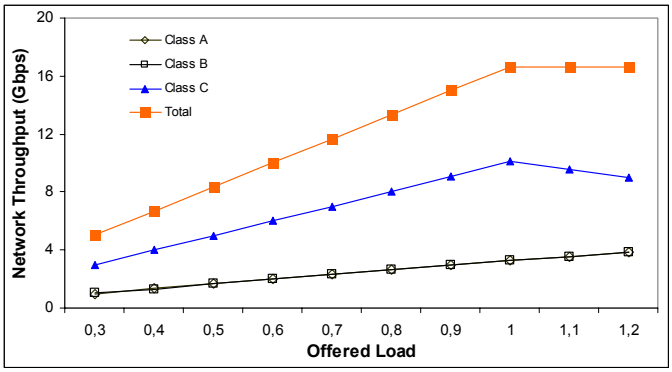


Figure 49: IEEE 802.17 RPR MAC: performance evaluation

If high priority traffic is used in RPR rings, the traffic must be shaped at ingress, and the service that uses this type of traffic must be carefully engineered. In fact, no mechanisms are provided to solve contention among high priority traffic streams. If the high priority traffic admitted exceeds the capacity of a given span, low priority traffic is blocked. Thus, the amount of high priority traffic injected into the ring must be controlled and limited by the higher layers, especially in the case of failure. Specifically, each failure scenario has to be investigated in turn to determine whether a given load is handled properly.

An IEEE 802.17 RPR node may use virtual output queuing to avoid head-of-line blocking for frames destined to nodes that are physically closer than the congestion point. This is called multi-choke implementation, which requires a detailed awareness of congestion points in the whole ring but increases ring utilization and spatial reuse.

6.1.2 RPR resilience mechanisms

At the same time, RPR standard offers powerful protection methods, namely the *ring wrapping* and the *packet steering*. They are based on the ring wrap at the nodes surrounding the failure or on the packet steering by causing the source node to redirect packets [97]. Both of them have been designed to minimize the traffic losses in case of failures and aim to achieve recovery times of about 50 ms and no spare resources are required to be pre-allocated [97]. RPR allows the full ring bandwidth to be utilized under normal conditions and protects traffic in case of failure obviating the need for SDH/SONET-based protection.

There are few steps to recover from failure by RPR layer, which include the indication of a failure (or significant signal degradation) and the final wrapping or steering of the ring.

The RPR complete recovery time is the time required by the network to return to a steady state after a failure has occurred in the ring. Focusing specifically on the wrapping, the complete recovery time comprises the *response time*, the *topology discovery reconfiguration time*, and the *MAC protocol convergence time*. The *response time* includes the detection of the failure, the generation of protection messages and the node state transitions, and finally the ring wrap. Wrapping mechanism works as follows: if failure is detected (either equipment or link failure), packets directed towards failure direction are wrapped back in opposite direction. It is made possible because of internal node structure with dual homing (connection to external and internal ring). Figure 50 shows an example of how ring wrapping works. The example shows RPR ring composed of four nodes in which link failure (e.g., fibre cut) on the fibre from node A to B is considered. The interchanged control messages needed to run the ring wrapping are figured as *{Request type, Source Address, Wrap Status, Path Indicator}*. In absence of failure (Ring in idle), each node periodically sends control message (IDLE). When the failure occurs in the external ring between A and B, node B detects a signal fail (SF) on the external ring, (e. g., not receiving the periodic keep-alive message sent periodically from node A). Thus, it changes to wrapped state performing a wrap and transmits towards A on the internal ring (short path) the message {SF, B, W, S} and on the external ring (long path) the message {SF, B, W, L}.

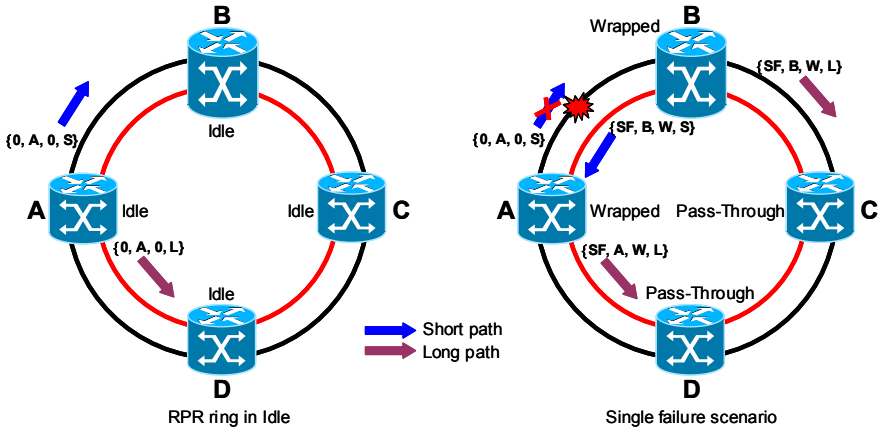


Figure 50: RPR wrapping protection

After receiving a protection request from node B on the short path node A changes to wrapped state (it has performed the ring wrapping). Then, it transmits towards B on the short path the message {0, A, W, S} and on the long path the message {SF, B, W, L}. When the nodes C and D

receive long path control packets (switch requests), they change from idle to pass-through mode (in each direction) [97].

After the wrapping of the ring, the update of the topology nodes map is needed because RPR uses the shortest path for the communications between node pairs. To do this, each node runs a topology discovery algorithm [97].

On the other hand, packet steering is based on the ability to choose the ringlet on which the data is sent. If the preferred path is unavailable due to failure, the other path will be used.

The implementation of the wrapping protection in the nodes is optional. Both protection modes may be mixed in a wrap-then-steer mode where the wrapping protection is activated first to avoid the loss of frames in transit; then nodes switch to steering to improve ring utilization.

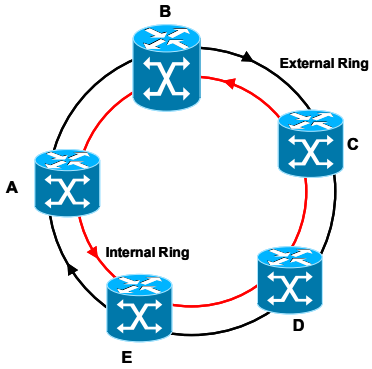
All the stations in the ring must use the same protection method; the default method in IEEE 802.17 RPR is steering. If, however, all the nodes support wrapping, the ring may be configured to use the wrapping protection.

At RPR layer, the failure detection may be carried out in two ways. The first is based on messages received from the physical layer, for example, Loss of Signal (LoS) from the SONET/SDH or the optical layer, and the second on periodical continuity checks within the IEEE 802.17 RPR layer.

6.1.3 Topology Discovery algorithm

The Topology Discover (TD) protocol is used for the network reconfiguration after the ring wrapping/packet steering. In normal RPR ring operation, both rings are utilized to carry traffic. After the wrapping of the ring, the available bandwidth is reduced, and low priority data traffic is reconverged fairly to the lowered bandwidth, which is accomplished dynamically by the fairness algorithm [97]. In order to optimize the bandwidth utilization it is necessary to run the TD protocol because RPR technology supports the basic version of traffic forwarding based on number of spans, which means that the shortest path is chosen towards destination RPR node (Figure 51). Each RPR node performs this action by sending out special discovery packets on one or both rings. The originator of a topology discovery packet sets the egress ring identifier (internal or external ring) adding its own MAC address and length field. Such packet is sent hop-by-hop around the ring (however in nature it is a point-to-point packet). Each traversed node appends its MAC address, updates length field and forwards packet towards destination. After reaching again the originator of

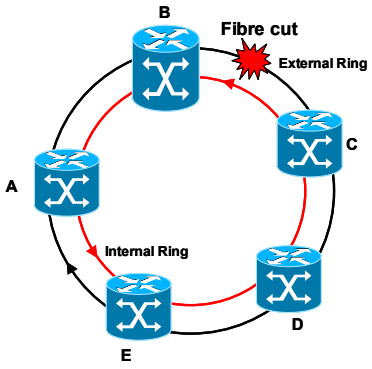
this packet, topology discovery packet has all nodes MAC addresses in proper order and with relevant length field.



Node A Topology map		
Destination node	**Outgoing ring**	**Number of hops**
B	External	1
C	External	2
D	Internal	2
E	Internal	1

Figure 51: Topology map for node A before failure

When sending topology discovery packet over wrapped ring, the wrapped node indicates this situation and wraps the packet (i.e., sends it further along the reconfigured ring). On the way towards originator after passing wrapped node, MAC addresses are not added since it is travelling the same route in opposite order of nodes. Then, the topology map at each node is updated (Figure 52).



Node A Topology map		
Destination node	**Outgoing ring**	**Number of hops**
B	External	1
C	Internal	3
D	Internal	2
E	Internal	1

Figure 52: Topology map for Node A after the running of the TD algorithm

The topology map of the ring is changed after receiving two identical TD packets. This is done to avoid changes of topology in transient conditions. The delay introduced by the TD protocol is the time of traversing the whole ring (already being wrapped) plus necessary time to service packet inside the subsequent nodes.

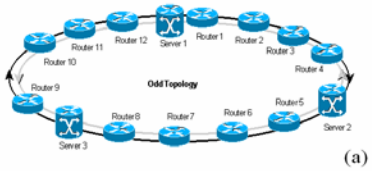
6.1.4 RPR resilience mechanism: Performance evaluation

We have carried out simulation case studies to assess the resilience features of RPR rings [102]. Specifically, a metropolitan IP/RPR network for the city of Milan was simulated using the OPNET

simulation tool [103]. A similar scenario can be assumed for any major European city. We consider service classes for both elastic traffic (web browsing, http-based services and e-mail services) and streaming traffic with stringent delay requirements (telephone services and video streaming), which are the most common IP-based applications. The simulated network was an RPR ring connecting twelve IP routers, and three Internet servers, and we assumed that all of them use the wrapping protection mechanism.

The distance among nodes was set to 3 km, which results in the propagation time between nodes of 15 μsec. We considered OC-48 (2.4 Gb/s) RPR node interfaces, with video and voice traffic sent as high priority (Class A) traffic, and data traffic (web browsing, http and e-mail services) sent as low priority (Class C) traffic.

The network consisted of a logical topology composed of three different segments, each one including four routers logically attached to one server. Each segment represents a geographical zone of the metropolitan environment. Moreover, we assumed traffic homogeneity in the three different segments and the same traffic matrices for all of these. As an example, Figure 53 includes the traffic matrices for one of these segments, namely the one composed of Server 1 (S1) and Routers 2, 3, 6 and 11.



(a)

	Server 1	Router 2	Router 3	Router 6	Router 11
Server 1	-	4.78	71.8	148.4	253.8
Router 2	104.15	-	8.48	-	8.48
Router 3	78.8	-	-	16.66	15.55
Router 6	45.42	-	7.21	-	19.46
Router 11	46.74	-	1.14	9.12	-

(b)

	Server 1	Router 2	Router 3	Router 6	Router 11
Server 1	-	-	34.5	-	34.5
Router 2	-	-	-	-	-
Router 3	73.9	-	-	-	38.1
Router 6	-	-	-	-	-
Router 11	17.2	-	8.9	-	-

(c)

	Server 1	Router 2	Router 3	Router 6	Router 11
Server 1	-	(*) VTx0.25	(*) VTx0.29	(*) VTx0.25	(*) VTx0.21
Router 2	-	-	-	-	-
Router 3	15.75	-	-	-	-
Router 6	13.5	-	-	-	-
Router 11	-	-	-	-	-

(d)

(*) Concerning the video traffic (VT) generated by the servers, we considered four different cases:
VT = 0,
VT = 0.33 Gb/s,
VT = 0.43 Gb/s and
VT = 0.83 Gb/s

Figure 53: (a) RPR network topology; Traffic matrix in Mb/s: (b) data traffic, (c) voice traffic, and (d) video traffic

These traffic matrices were obtained from the estimation, carried out within the IST LION Project, of traffic flows in a realistic environment (the city of Milan) [104], and we also used in

[105] and [106]. The estimation took into account not only the characteristics of each service but also the potential penetration (percentage of customers) for these kinds of services.

As concerns the traffic model, we used the ON-OFF model, with a burstiness (peak rate/average rate) of b = 10 and a mean burst length of BL = 10 packets for data traffic sources, and the Poisson model with a mean packet arrival intensity of λ packets per second for voice and video traffic sources. For data traffic, we considered the statistical distribution for the IP packet size given in [62], while for voice and video packets we used fixed packet sizes of 44 and 512 bytes, respectively. On this scenario, we simulated two different cases study, namely a fibre cut between two routers and a router (not server) failure. The simulated operation time was 200 ms for the case of the fibre cut and 300 ms for the case of the router failure. In both cases, it was assumed that the failure occurred at the instant t = 70 ms.

The aim of these simulations was to evaluate the impact of a failure, both on the mean end-to-end delay and on the network throughput.

Figure 54 depicts the mean end-to-end delay experienced by the Class A and Class C traffic before the failure (in this case the fibre cut) occurs and after the reconfiguration of the ring once the failure has occurred. The results show that the average end-to-end delay suffered a quite significant increase (about 50%). This is due to the fact that after the wrapping reconfiguration of the ring the end-to-end path is longer for some traffic streams. It has to be noted that in the particular case of Class A traffic, the end-to-end delay is below the typical requirements for voice and video services even after the ring failure recovery.

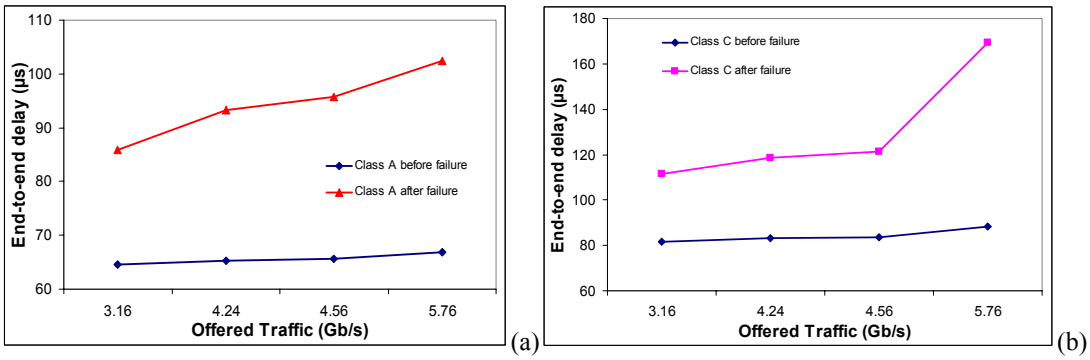
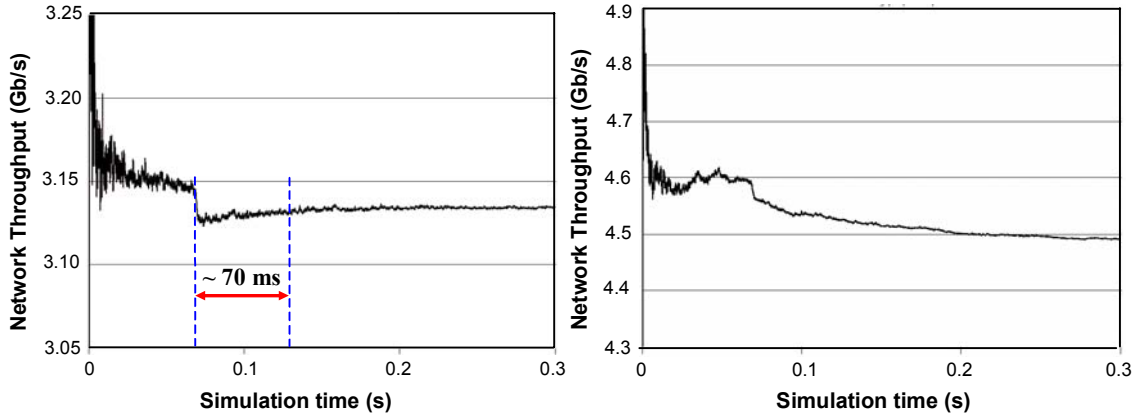


Figure 54: Impact of link failure on end-to-end delay for (a) high priority and (b) low priority traffic

Figure 55 depicts the throughput as a function of the simulated operation time in the case of the router failure (300 ms). Both plots of this Figure show that after node failure the throughput suddenly decreases and subsequently, after the complete recovery, converges towards a final throughput, which is lower than the throughput value before the failure. This is because a router,

after it fails, no longer injects traffic into the ring and the remaining nodes stop sending traffic towards that router once they know that it has been excluded from the network. Figure 55 (a), which was obtained for the case of no video service, shows that the network throughput evolves towards a steady state after the stabilization of the MAC protocol, and the RPR ring continues to work efficiently. We estimated that, in this case, the time needed to return to stability (to the same network throughput value) is 70 ms. Figure 55 (b), which is obtained in case of higher load (including video traffic: VT per server = 0.43 Gbit/s), shows that, after node failure, the throughput decreases, and it takes longer time to reach a stable situation.



(a) (b)

Figure 55: Network throughput evolution after a node failure: (a) no video traffic, (b) average video traffic generated by the servers is 0.43 Gbit/s

In the case of failure the most important objective is maintaining network connectivity and minimization of the packet losses. The results of the simulation experiments discussed above show that the RPR protection mechanisms have been optimized to do so. On one hand, traffic losses can only occur during the response time (few ms), which is comparable to that of SONET/SDH networks. On the other hand, the complete recovery time (including the time required for the reconvergence of the RPR MAC protocol) depends on the ring size and on the actual traffic load, but if the traffic is well engineered the network, after the ring reconfiguration, reaches the stability and will not saturate.

Although fast and efficient, the RPR recovery mechanisms imply that the available bandwidth is reduced. The reduction factor strongly depends on the actual load and distribution of traffic. The next Section discusses this problem in detail.

6.1.5 Potential hazardous situation[1]

The aim of this Section is to describe a situation in which a given traffic assignment leads to a significant degradation of network performance when failure occur. The presented example is valid both for steering and wrapping protection.

The consequence of the failure is that the routes traversed by frames switch from short to long ones. Additionally, the use of the fairness algorithm causes bandwidth to be shared between all active streams. This inevitably leads to potentially hazardous situations, e.g., a significant decrease in the bandwidth allocated to each Class C stream.

As an example, consider the situation in Figure 56, in which a RPR ring composed by N nodes is taken into consideration. Nodes in one part of the ring (here, on the right) send Class C traffic to their neighbours while the remaining nodes send class C traffic to a given hub node (depicted here as Node 3).

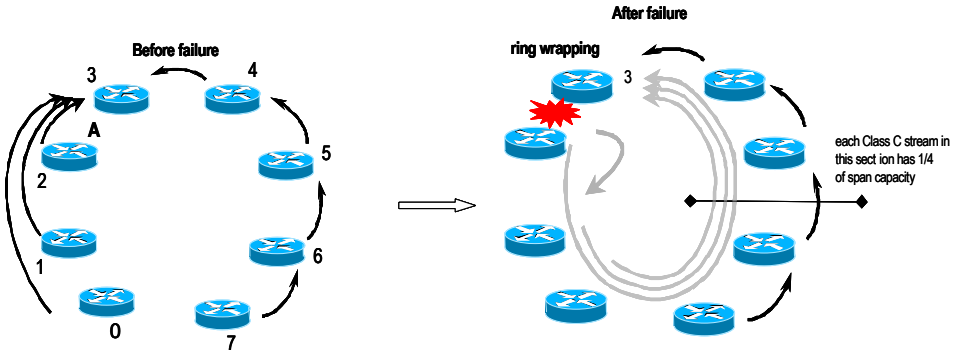


Figure 56: RPR ring with a worst-case traffic stream assignment

Let N_n be the number of nodes sending traffic to their neighbours and N_{to_3} the number of nodes sending traffic to node 3 (Figure 56, on the left).

$$N_n + N_{to\_3} + 1 = N \qquad (1)$$

The total bandwidth (i.e., the maximum bandwidth available for all traffic streams in the case described) before failure is equal to:

$$B = N_n + 1 \qquad (2)$$

The bandwidth unit used here is the full bandwidth of the RPR span, (e.g. 2.4 Gbit/s). Component 1 in the equation above is the result of all N_{to_3} streams in the left part of the ring sharing the bandwidth and therefore each stream reaches a steady state of $1/N_{to\_3}$ units when

[1] This Section includes the work carried out by the University of Science and Technology (AGH), Krakow, Poland, as a contribution to reference [102]

sending traffic to node 3. Component N_n is the aggregated bandwidth of the streams associated with nodes sending traffic to their neighbours (on the right side of the ring). Obviously, the situation shown in Figure 56 is a simplification, since both optical rings are, in fact, used.

After the failure, the rings are wrapped and the traffic streams previously directed to node 3 must now travel in the opposite direction, thus interfering with traffic to neighbouring nodes. Due to the fairness feature for Class C traffic, each span of the ring shares its full capacity evenly among (N_{to_3}+1) streams (as shown in Figure 56, where N_{to_3} equals 3), so the capacity of single streams equals 1/(N_{to_3}+1). The total bandwidth for RPR after failure is:

$$B = \frac{1}{N_{to\_3}+1} \cdot (N_n + N_{to\_3}) \qquad (3)$$

The total loss of traffic after failure is equal to (after maximizing it against N_n):

$$Loss = 1 - \frac{4(N-1)}{(N+1)^2} \qquad (4)$$

For N = 63 (in principle, N may be as high as 255), the loss of traffic is equal to 94% of the traffic sent before the RPR ring reconfiguration. This result is somewhat discouraging and leads to the following conclusions, namely: 1) RPR rings should be engineered to avoid such a traffic pattern (or similar) and 2) When network performance is an important issue it may be necessary to verify network operation in each of the assumed failure scenarios.

Apart from the theoretical case study carried out by AGH University, we carried out a simulation case study to evaluate the bandwidth reduction due to the failure. Specifically, we simulated a 5-nodes RPR network. The offered traffic (ρ) was set to 0.6. The simulated operation time was 120 ms and the failure (e.g. fibre cut connecting two underlying OXCs) occurs at the instant t = 50 ms. The wrapping method was used.

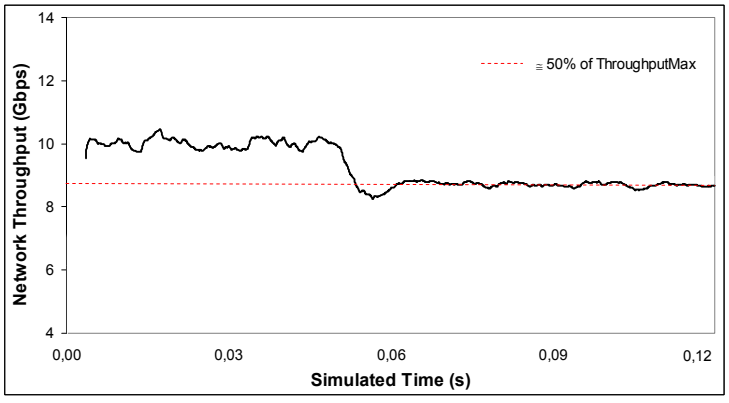


Figure 57: Effect of the RPR network reconfiguration: ring saturation

As it can be seen in Figure 57, once the RPR layer detects the failure and after the consequent reconfiguration of the ring, the available bandwidth is substantially reduced (in this case halved), so then the network is not able to carry the entire offered load (the RPR ring is saturated).

6.1.6 Summary of strengths and weakness of RPR technology

RPR technology has attracted great interest during the last 3 years. It can be considered a *niche* technology. Important issues related to the use of RPR technology are discussed below, to point out its advantages and discuss its disadvantages.

The protection mechanisms implemented in RPR are fast: they aim to achieve recovery times of approximately 50 ms and to protect against any single failure in the ring. No bandwidth is dedicated for recovery purposes and, therefore, in a failureless state the resource utilization is high. However, in the case of failure, the bandwidth available is substantially reduced. The reduction factor depends on the actual load and distribution of traffic.

If high priority traffic is used in an RPR ring, the traffic must be shaped at ingress, and the service that uses this type of traffic must be carefully engineered. No mechanisms are provided to solve contention among high priority traffic streams. If the high priority traffic admitted exceeds the capacity of a given span, low priority traffic is blocked. Thus, if congestion problems have to be avoided, the amount of high priority traffic injected into the ring must be controlled and limited by the higher layers, especially in the case of failures. We suggest that each failure scenario be investigated in turn to determine whether a given load is handled properly.

The RPR would seem to be a wise choice for efficient and reliable transport of best-effort traffic. It may be used to transport traffic with strict bandwidth and delay requirements, although in this case one would need to verify whether RPR would satisfy the necessary parameters for all the conceivable traffic-flow patterns. With regard to the use of different classes of traffic, RPR requires external measures to prevent congestion. These measures are not standardized or otherwise defined at present, so it is up to the user to provide them. However, it is possible that such measures will be defined as RPR technology matures and its use becomes widespread.

Finally, an important issue in modern telecommunications networks is interoperability among different layers. A new protocol should interwork smoothly with existing protocols. Interoperability with several physical layer techniques was explicitly considered during the standardization process

of the IEEE 802.17 RPR. From the upper layer point of view RPR may be seen as a shared medium technology, and as such the problem was not widely studied during the definition of the standard.

6.2 Resilience interworking strategy in RPR over intelligent optical networks

The increase of the number of wavelengths that can be multiplexed onto the same fibre (up to 160) each one carrying 2.5 or 10 Gbit/s client signals implies that outages of the network infrastructure (e.g. fibre cut) can have serious consequences (economical as well as social) [26].

As before mentioned, in current network infrastructure recovery is carried out at SONET/SDH layer. Protection at the SDH layer, based for example on the Automatic Protection Switching (APS) protocol [107], although very robust (allowing network recovery within 50 ms) is not efficient from the network resources optimization point of view. In fact, it needs to pre-allocate spare network resources to be used for protection purposes and the bandwidth reserved for the backup paths is not used to carry traffic, increasing in such a way the required CAPEX.

On the other hand, the achieved advances in optical components as well as the introduction of intelligence (i.e. Control Plane) to the optical layer lead to the definition of recovery mechanisms directly in the server optical layer. The introduction of resilience mechanisms in the optical layer is very useful because the optical layer provides better management for certain kind of failures. As an example, let us suppose that a single optical fibre carries multiple wavelengths which correspond to some SONET/SDH streams. If recovery at SONET/SDH layer is used, a fibre cut therefore results in that all the streams are restored independently by this layer. As a consequence, the network management system is flooded with a large number of alarms generated by each of these independent entities. Contrarily, if the fibre cut is recovered at optical layer this operational inefficiency can be avoided.

Hence, failures such as fibres cut, or optical equipment damages can be more efficiently handled.

Both protection (1+1 and/or 1:1/1:N) and restoration can be used in the optical layer [3]. The fault management is based on the detection of the failure, the notification of the detection of the failure, the failure localization, and finally on the recovery procedure. The latter can be based for example on dedicated path protection (Figure 58) and/or restoration.

To implement such fault management the GMPLS-based control plane can be used. Specifically, the optical nodes are equipped by OXC switch controllers connected through

signalling networks and generating the optical signalling. Different ways can be used to implement signalling between the controllers of the optical equipments.

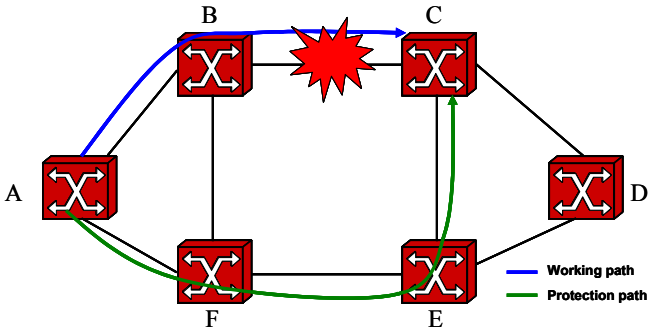


Figure 58: Recovery at the optical layer, 1:1 dedicate path protection between node A and C

As described in Chapter 2, the way to implement signalling, the in-fibre in-band signalling, the in-fibre out-of-and signalling and the out-of-fibre out-of-band signalling methods can be used.

Specifically, Figure 59 shows an example in which the in-fibre out-of-band signalling is considered.

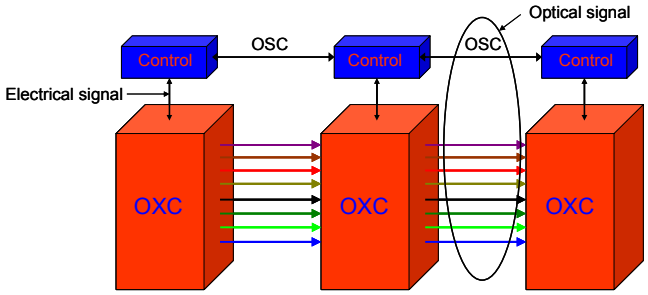


Figure 59: Control structure in the optical layer, in-fibre out-of-band signalling

To monitor the connectivity of the control channels, the Link Management Protocol (LMP) defined by the Internet Engineering Task Force (IETF) is used [108]. In such a context, the supervision of the control channels is based on a fast keep-alive procedure based on the interchange of *HELLO* messages between the optical Connection Controller (CC). Each node sends the *HELLO* messages periodically (each *HelloInterval*) to its neighbours. Nevertheless, the interchange of these messages could not provide data link failure detection since the failure of the control channels can be due to the failure of the transmission equipment of the signalling messages (i.e., control channel lasers). Thus, the data link failure detection is done combining the signalling based on the *HELLO* messages with the monitoring of the optical signals at the OXC interfaces/ports, in order to detect the Loss of Light (LOL). In such a scheme, when the optical layer detects the fault, first it signals to

the client layer that the failure has been detected and then it launches the predefined procedure for the management of the fault.

It has to be underlined that the signalling network can be used not just for the connectivity of the control channels but also for the management of the routing and signalling instances.

However, it worth noting that carrying recovery in the optical layer presents some drawbacks such as: 1) The optical layer is not aware of failures that occur at higher layers and 2) Link budget constraints limits the recovery capability because the length of the protection route or the number of nodes the protected traffic passes through may be physically constrained [3].

If we consider that RPR runs over intelligent optical transport networks, using a single-layer strategy, the only layer responsible for taking the recovery actions in case of a failure is the optical layer. This is due to the fact that it is better to recover from the failure at the optical layer since the RPR protection mechanisms imply the reduction of the available bandwidth.

The obvious advantage of this single-layer approach is that it does not require any interworking feature between RPR and optical transport layer. However, not all kinds of failures will be handled efficiently with this strategy, since the optical layer is not able to detect the failures occurred at the higher layers, such as the RPR line card failures or RPR site failures.

Hence, this Ph.D. Thesis proposes a multi-layer recovery strategy in order to efficiently coordinate the different mechanisms of each layer. It is to be used in an IP over RPR over OTN/ASON metropolitan network scenario.

When combining RPR over OTN/ASON, we are using technologies with similar reaction times but different features. While RPR recovers from failure by ring wrapping around failed span or by packet steering, the optical layer relies, for example, on dedicated resources to recover from failures (i.e., 1+1 and/or 1:1 dedicated protection).

If the *uncoordinated approach* is used, both RPR and optical layer resilience mechanisms act independently of each other. Nevertheless, due to the fact that resilience mechanisms detect the failure in similar time, it is very likely, that both layers will try to restore connectivity at the same time (Figure 60). This can lead to significant performance degradation for the layer above RPR (i.e., IP). If both RPR and OTN/ASON start recovery actions, independently of the procedure in the optical layer, once detected the failure, RPR will wrap its ring, thus reducing the available

bandwidth depending upon the traffic pattern usage before and after the wrap. As a result, a failure at the optical level that could be efficiently managed by the optical layer moreover implies the reduction of the bandwidth at the client layer (RPR layer).

The problem of harmful interaction between RPR and underlying layer is only mentioned in [97]. There, RPR over SONET/SDH interworking scenario is considered and not specific conclusions are given. The document suggests avoiding such a case where RPR and SDH protection are used simultaneously and recommends either using a single layer protection (e.g. SDH APS) or implementing a hold-off timer (hereafter single hold-off timer approach).

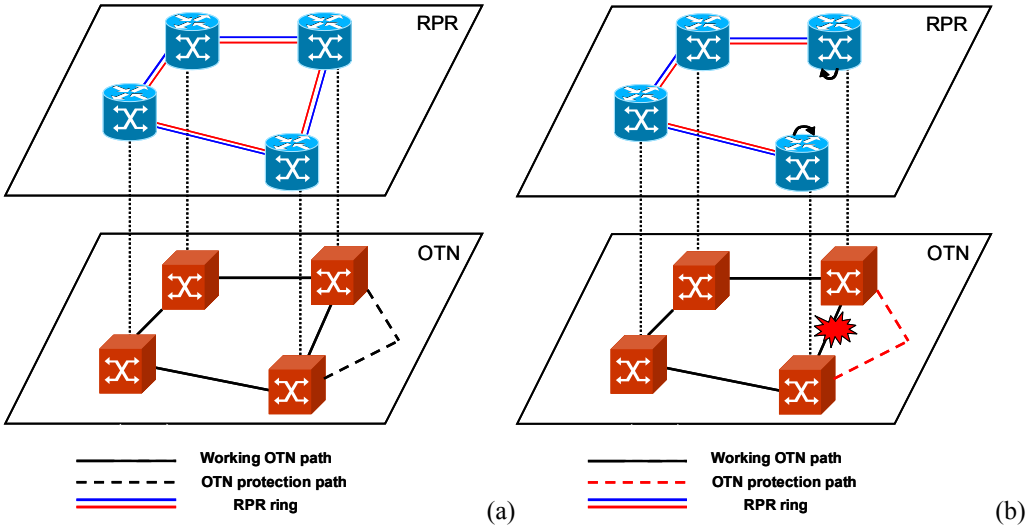


Figure 60: RPR/OTN scenario: a) basic arrangement, b) uncoordinated approach

Regarding the single hold-off timer, using this approach, the recovery action in OTN/ASON is launched immediately after a failure is detected, while the recovery in RPR is delayed for some time, needed to complete recovery tasks. If the optical layer manages to re-establish connectivity, there is nothing to do for the RPR protection and no action is thus taken at this level. If, on the other hand, the optical layer is unable to recover from a failure, after the hold-off timer has expired RPR will trigger its own protection, and will recover from the failure. The main advantage of this approach is its simplicity. Moreover, in case the failure is resolved in the optical layer, there is no need for RPR re-convergence (which slows down recovery process) and no bandwidth reduction occurs as it was in case of the uncoordinated recovery.

Nevertheless, the single hold-off timer (hereafter SHOT) approach presents a very important drawback, namely its length. If it is too short, the protection in RPR will be needlessly triggered before OTN/ASON has finished the recovery process. If it is too long and OTN/ASON cannot cope

with the failure, RPR will wait with its protection without any reason. Summarizing, if the failure occurs at the optical layer, it works properly but its efficiency decreases very much when the failure occurs at the higher layers.

6.2.1 Double Hold-Off timer approach

This Ph.D. Thesis proposes a novel multi-layer resilience strategy, based on the interworking between RPR and the optical layer. It consists on implementing the double hold-off timer (hereafter DHOT) approach. RPR can detect a failure in different ways, depending on the used sources of information about failures: some of them are independent from other layers and one is the information signalled from the underlying layer (i.e., optical layer). RPR, detecting a failure (through signal fail (SF) signalling [97]), is able to distinguish between two cases: 1) When the optical layer has also detected the failure (it has occurred at optical layer) and 2) When the optical layer has not detected the failure (and RPR is the only layer that is able to do a successful recovery). In the latter case, the failure has occurred in the upper layers.

The suggested DHOT approach is based on dividing the entire single hold-off timer into two parts, namely the H1 (short) part and the H2 (long) one. The first part (H1) is activated after RPR detects the failure. It serves to give to the optical layer some time to detect the failure, to notify and signal it to the RPR layer. It has to be underlined that the time required for the detection of the failure at the optical level is strongly influenced by the optical components themselves and their management. Anyway, according to [109], it takes very few ms.

After the expiration of H1, if the optical layer has not detected the failure, RPR triggers its protection immediately. On the other hand, if the failure is signalled to RPR layer by the optical layer, the RPR layer waits during the H2 timer to give time for recovery in the optical layer. The DHOT approach is described in detail in the flow-chart of Figure 61.

The recovery procedure in the optical layer encompasses both fault localization and the recovery mechanism (i.e. dedicated path protection or restoration). Even in the case the optical layer detects the failure, it could be unable to solve it, for example due to the unavailability of resources in case of using restoration. Therefore, if after the expiration of H2 the failure is not recovered, then RPR protection mechanism is launched.

The required functionalities for the DHOT implementation have already been incorporated both in optical transport layer and in RPR. As stated above and according to [17], OTN is required to

signal to its client layer both signal degradation and signal failure while the RPR standard is able to accept such signals from the underlying layer [97].

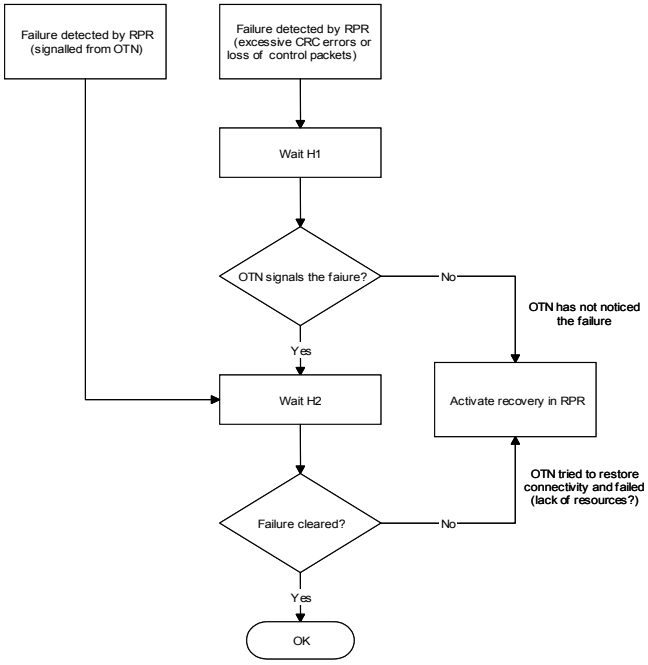


Figure 61: Coordinated approach: double hold–off timer (DHOT)

Figure 62 depicts the failure management both in the case the failure occurs at the RPR layer and in the case it occurs at the optical layer.

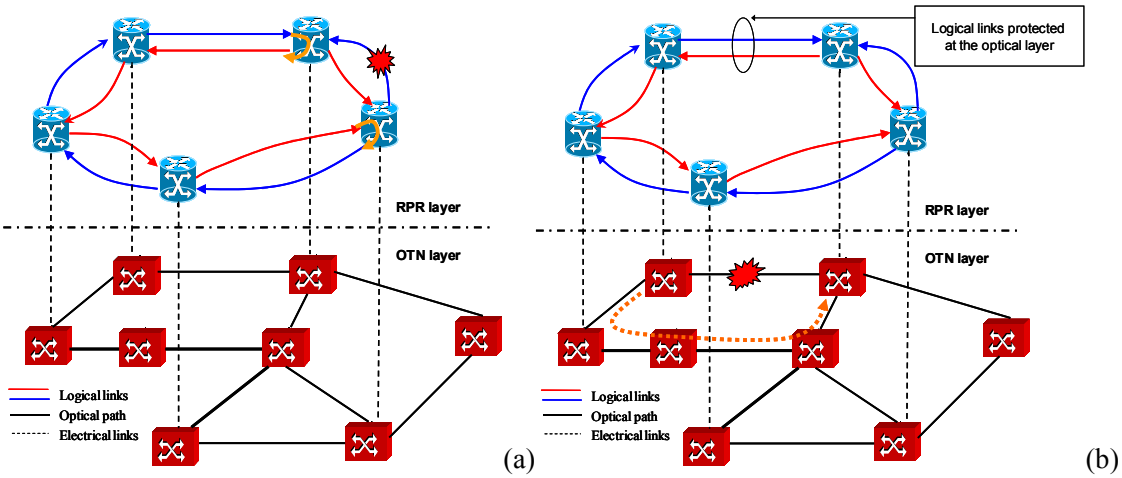


Figure 62: Failure management, a) at RPR layer and b) at the optical layer

The main advantage of the DHOT approach, when compared to the SHOT one is that the recovery time is much shorter when failure root is above the optical layer, allowing in this case minimizing the traffic lost due to the failure. Moreover, being based on the interworking between

the two layers, it is able to react to the failure more promptly and independently from the failure scenario optimizing, at the same time, the utilization of the network resources.

6.2.2 DHOT approach: Performance Evaluation

We carried out various simulation case studies in order to compare the SHOT and the DHOT approaches for different failure scenarios. The simulated scenario consists of 5 IP routers equipped with RPR cards, logically connected through a meshed optical transport networks composed by optical cross-connects (OXCs). The optical nodes are connected through bi-directional optical paths (i.e. two physically disjoint optical fibres) and the 1:1 optical path protection was implemented. For the fault management in the optical layer, we implemented the GMPLS-based Link Management Protocol. In our simulation model, the fault detection is carried out through the implementation of the *HELLO* messages signalling between the optical nodes controllers combined with the Loss of Light (LoL) alarms from the OXCs. Specifically, an in-fibre out-of-band signalling approach has been implemented. To avoid to get the ring saturated after the RPR recovery process, the offered load (ρ) was set to 0.45. Class A represented the 20% of the offered traffic, the same for the Class B traffic and the rest represented Class C traffic. We also assumed that the traffic inserted in the ring by each node was uniformly distributed among the rest of stations/nodes. The simulated operation time was set to 120 ms and the bottom-up coordination approach was used. The failures occur at the instant t = 50 ms.

Two case studies were carried out. The first one concerned the case in which the failure at the optical level (e.g., cut of the fibre connecting two OXCs breaking the logical connection between two IP/RPR routers) and the second one concerned the case in which the failure occurs at RPR layer (e.g., failure of RPR card of one of the routers composing the ring). In both case studies, the H1 timer was set to 10 ms while the H2 timer was set to 30 ms.

Focusing in the first case study, Figure 63 shows the network throughput versus the simulated operation time. According to the defined DHOT approach, the RPR layer once detected the failure, instead of launching immediately the recovery action waits for H1 in order to leave to the optical layer the objective to recovery from the failure. In this case, since the failure root is in the optical layer, this signals the failure detection to the RPR layer and then it launches the optical path protection. Once the optical level has recovered the network from the failure (we implemented path protection recovery), the network throughput comes back again to the same value before the failure occurred. When the optical layer is able to handle the failure, the behaviour of the network

throughput is the same using both the DHOT and the SHOT approaches (In Figure 63 only the case of DHOT is plotted). We obtained that the time required returning to the steady condition, which includes the fast failure detection and the switch to the optical protection path, is about 12 ms.

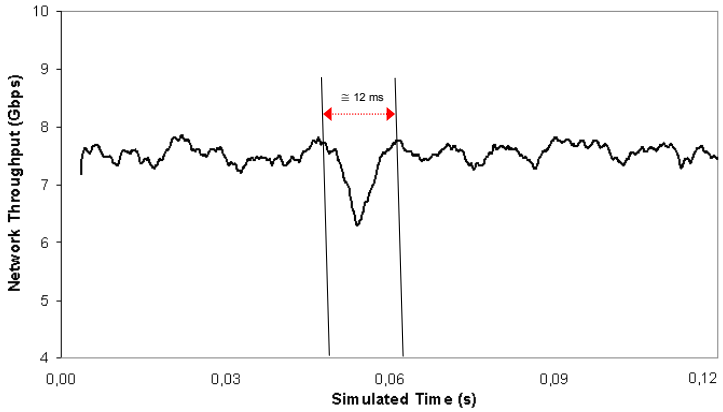


Figure 63: Recovery from failure at optical level

We insist that the detection failure at the optical level is strongly influenced by the optical components themselves and by the way they are managed. Anyway, as stated before, it takes very few ms.

The second case study deals with the comparison between the SHOT and the DHOT when the failure occurs at the RPR layer or upper layers. The comparison of the network throughput in case of using the SHOT and in case of using the DHOT is depicted in Figure 64. The SHOT foresees that the RPR layer waits for the entire hold-off timer (H1+H2). Once the timer has expired and the failure has not been recovered by the optical layer, then RPR starts to recovery from failure. By using the DHOT approach, the RPR layer just waits for the first short timer (H1). If H1 expires and the optical layer has not signalled the failure detection, then the RPR starts immediately the recovery action (in this case, we implemented the wrapping mechanism).

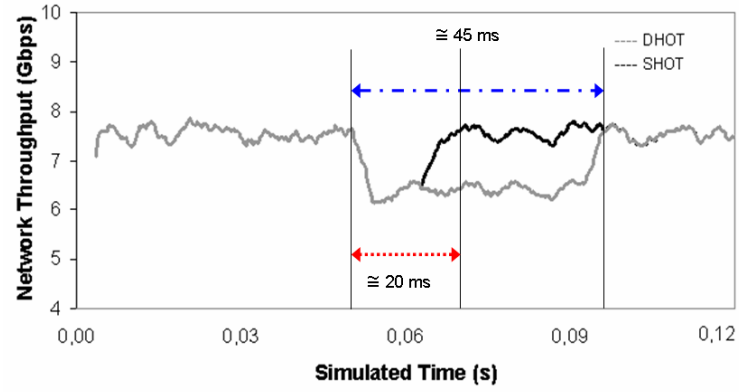


Figure 64: SHOT vs. DHOT, failure at RPR level

In such a case, after H1 expires and the RPR layer runs its recovery procedure, the network throughput comes back again to the value before the failure. We can estimate that the time required to the network throughput to come back to the same value before the failure in case of SHOT is about 45 ms while in case of using DHOT in the same conditions is about 20 ms.

We also carried a third simulation case study to compare the SHOT and the DHOT approaches, using the relative traffic losses. Also, in this case, we assumed that the optical layer is unable to detect the failure occurrence and we considered various values for the H1 and H2 timers. We calculated the ratio (R) between the packets lost obtained with the SHOT and the DHOT approaches, that is $R = (\text{Packets Lost})_{DHOT}/(\text{Packets Lost})_{SHOT}$.

Table 11 and Table 12 depict the gain in terms of percentage of reduction of the traffic losses (i.e., 100*(1-R)) arising from the implementation of the DHOT approach with respect to the SHOT. For such comparison, both RPR protection mechanisms have been considered.

Specifically, the results showed in Table 11 indicate that the traffic losses reduction ranges from the 62% to 77%, according to the values of the H2 timer and the RPR protection mechanism.

H1 = 10 ms	***DHOT vs. SHOT: Traffic Lost reduction***	
H2 (ms)	*RPR wrapping*	*RPR steering*
20	62.5%	62.0%
30	71.5%	71.1%
35	74.5%	74.2%
40	77.0%	76.6%

Table 11: DHOT vs. SHOT: Packets lost

Table 12 depicted this gain when fixing the value for the H2 timer (30 ms) and considering different values for the H1 timer. The aim is to consider that the implemented failure detection at the optical switched strongly depend from the optical components. Specifically, in this case, the results indicate that the traffic losses reduction ranges from the 52% to 71%, according to the values of the H2 timer and the RPR protection mechanism.

H2 = 30 ms	***DHOT vs. SHOT: Traffic Lost reduction***	
H1 (ms)	*RPR wrapping*	*RPR steering*
10	71.5%	71.1%
15	63.9%	63.6%
20	57.5%	57.2%
25	52.7%	52.5%

Table 12: DHOT vs. SHOT: Packets lost

We carried out a fourth simulation case study in order to complete the comparison between the SHOT and the DHOT approaches when the failure occurs at the RPR layer. In particular, we

evaluated the recovery time that is the time required to recover from the failure. Specifically, this is the time required for the network reconfiguration after the failure and it encompasses the time required for the ring wrapping plus the time required by the TD algorithm. Table 13 reports the recovery times when the H1 timer is set to 10 ms and various values for H2 timer are considered. If the DHOT approach is used, the recovery time is given by the H1 timer plus the time required by RPR layer to recover from the failure (few ms). Contrarily, if the SHOT would be used, the recovery time, since the optical layer is not able to recover from the failure, is given by the total hold-off timer (H1+H2) plus the time required by the RPR mechanisms.

H1 = 10 ms	***DHOT***	***SHOT***
H2 (ms)	*RPR wrapping (ms)*	*RPR wrapping (ms)*
20	12.55	32.57
30	12.55	42.58
35	12.55	47.56
40	12.55	52.57

Table 13: DHOT vs. SHOT: Recovery Time

Specifically, it can be observed that DHOT performs much better than the SHOT approach. In fact, the time required by DHOT for recovery is about from the 23% to the 38% of the time which would be required by using the SHOT.

Table 14 reports the recovery time for different values of H1 while H2 is fixed to 30 ms. Also in this case, it can be observed how the recovery time is much lower using the DHOT than using the SHOT. Specifically, the recovery time required by using the DHOT is from the 29% to 48% of the recovery time required by using the SHOT.

H2 = 30 ms	***DHOT***	***SHOT***
H1 (ms)	*RPR wrapping (ms)*	*RPR wrapping (ms)*
10	12.55	42.58
15	17.57	47.56
20	22.57	52.60
25	27.56	57.55

Table 14: DHOT vs. SHOT: Recovery Time

Finally, it has to be noted that this percentage depends on the actual traffic load, the failure scenario and the set of the double hold-off timers.

6.3 Resilience Interworking in RPR over ASON/GMPLS networks

As stated in the previous Sections, the protection at RPR layer implies "some" reduction of the available bandwidth at logical level.

Let us suppose that the traffic transported by the RPR ring and generated from the higher layers (i.e., IP/MPLS) is uniformly distributed among the ring nodes, hence the maximum offered traffic to avoid the ring saturation (after the ring reconfiguration) is $\rho = 0.5$. In fact, in this situation, the available bandwidth is halved after the wrapping/steering. If the offered traffic is higher than 0.5, the recovery action implies the saturation of the ring, basically blocking the low priority (Class C traffic). On the other hand, both the Class A and B have to be well-engineered in order to avoid packet losses due to the ring saturation.

On the other hand, as vastly discussed in Chapter 2 and 3, in an IP/MPLS over RPR environment, the client traffic offered to the ring is characterized by its fluctuations over time (e.g. on a daily time basis).

We propose here a procedure which aim consists in using the automatic switching of optical connections capability provided by the ASON networks to face with both the potential ring saturation in case of failures and the fluctuations over time of the traffic inserted in the RPR ring. By implementing this procedure, the available bandwidth of the ring is automatically increased/decreased when strictly required by the ring status.

Basically, this procedure is based on introducing at the IP/RPR routers a monitoring function in order to compute periodically (e.g., each Observation Window) the traffic being carried by the light paths connecting a couple of RPR nodes.

As illustrative example, let us suppose that IP/RPR routers are connected through light paths (e.g., permanent optical connections set up by the NMS). This is depicted in Figure 65 (a). If failure occurs and it has to be handled at the RPR layer by using, for example, the ring wrapping mechanism, the monitoring function is used to detect an overloading condition (i.e., ring saturation) as a consequence of the ring reconfiguration. When the over-loading condition is detected on the light path connecting two RPR nodes, then the IP/RPR router requests via UNI interface to the optical connection controller (CC) of the optical components (e.g., OXC) the dynamic establishment of a switched connection. If the GMPLS signalling is able to provide the switched connection, the two RPR nodes are connected through two light paths (Figure 65 (b)).

Then, by applying some TE rules (i.e., Load Balancing), the traffic to be transported is distributed between the two light paths.

On the other hand, when the monitoring function detect that the traffic between the two routers can be carried by only the permanent connection (under-utilization condition), then the router requests to GMPLS-based control plane the tear down of the switched connection.

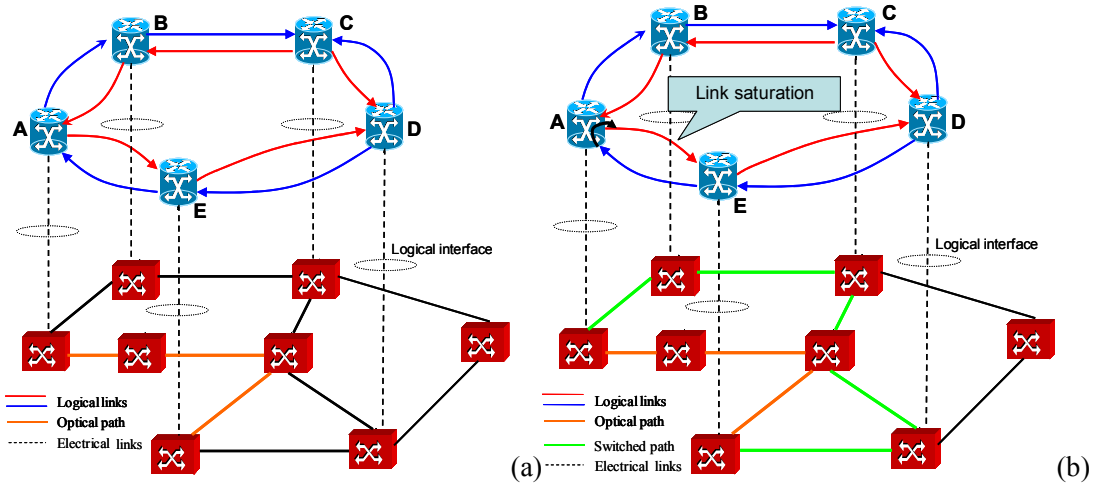


Figure 65: RPR over ASON interworking in case of failure, (a) Initially condition, (b) Requesting a switched connection

For the over-loading/under-utilization of the light paths, a threshold-based policy (the one defined for the TRIDENT procedure) is used, which means using a high threshold for the congestion detection and a low threshold to detect the under-utilization condition [110].

Of course, the tear down of the switched connection is also requested also when the failure has been physically solved and the ring logical bandwidth comes again to initial conditions.

Nevertheless, the implementation of such a mechanism implies the use of spare RPR cards to be used when saturation occurs

The aim of this interworking procedure is to keep limited the size of the transport networks. Indeed, when the failure has to be recovered at the RPR layer, the automatic switching capability offered by the ASON/GMPLS networks can be used to avoid to dedicate spare resources to protect at the optical level the logical links between routers. In general, it has to be considered that a transport network support different client networks and thus, avoiding to over dimension the transport network allows reducing the CAPEX for Network Operators.

The following Figure 66 illustrates the flow-chart of the suggested procedure.

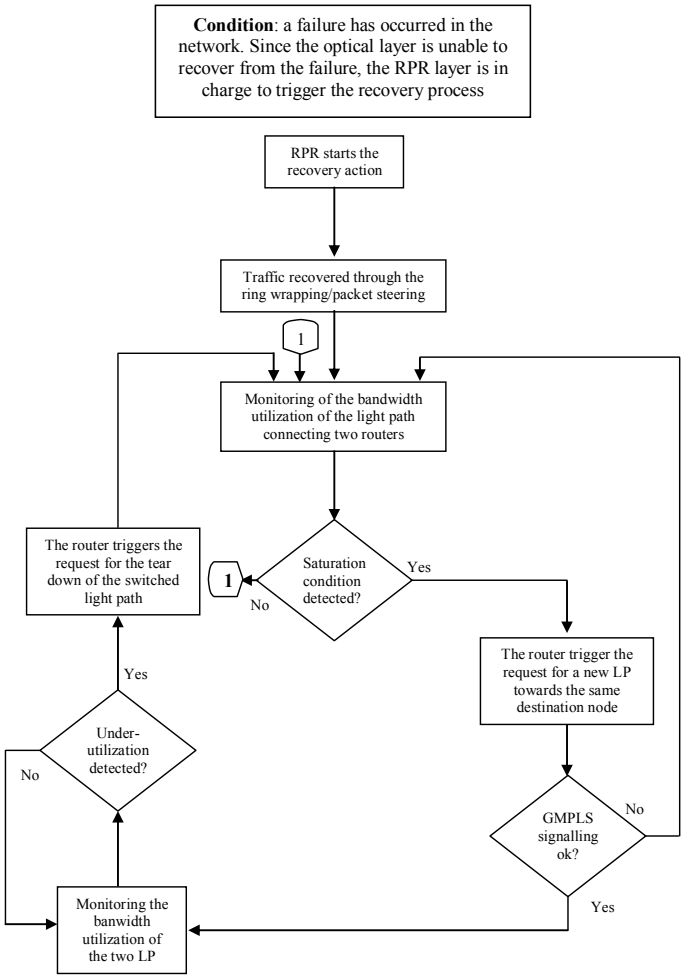


Figure 66: RPR over ASON interworking: flow-chart

7 Summary and Final Conclusions

The Information Age is consecrating IP as the integrating layer of most applications. This is fuelling the increase of data traffic, characterized by its typical asynchronous, burst and asymmetric nature.

On one hand, the introduction of ASON/GMPLS network solutions is, in a medium/large term the most interesting approach to meet network emerging requirements, not only to overcome the four fundamental network problems (bandwidth, latency, packet loss, jitter) for transporting traffic in real time, but also to enable flexible, automatic and fast provisioning of bandwidth, automatic discovery, multi-layer traffic engineering and multi-layer resilience.

On the other hand, for metropolitan environments, the emerging RPR packet-based technology seems to be, in a short term, the most promising technology to substitute the legacy SONET/SDH networks, which have been designed and optimized for voice-based applications. In contrast, RPR was specifically designed to transport data (IP) traffic.

This Ph.D. Thesis includes four contributions. The first one proposes the capacity management/traffic engineering TRIDENT procedure for the automatic set up and the tear down of switched optical connections based on the monitoring and prediction of the aggregated IP/MPLS packet traffic to be transported by an ASON/GMPLS network. Specifically the procedure deals with the dynamic establishment of optical connections in order to track the client traffic fluctuations.

Without requiring any knowledge of the future aggregated traffic pattern to be carried, the suggested procedure allows to provide, time-dynamically, the bandwidth required to transport through the ASON/GMPLS transport network the MPLS-LSPs already established at the IP client layer, while avoiding network congestions and bandwidth connections under-utilization.

Both the simulation and the experimental results show the feasibility of the procedure, which allows to promptly reacting to dynamic changes in the client traffic. The main requirements of the procedure are to maximize the utilization of the bandwidth of the optical connections, to limit the number of connections set up requests to have HT and IAT statistics compatible with the CP requirements and finally to minimise the IP packet losses.

The extensive simulation results presented in this Ph. D. report show that by adequately configuring the procedure parameters, it provides high figures for the bandwidth utilization of the light paths reaching in such a way traffic engineering objectives. At the same time, by applying the conservative approach, TRIDENT allows to keep limited the IP packet losses as well as the number of request to the GMPLS-based control plane for connections establishments. The former implies that TRIDENT allows meeting the QoS requirements imposed on the transport network by real-time applications while the latter implies that it is able to track the client traffic fluctuations without excessively increase the cost of the routing and signalling functionalities.

Also we present the experimental results that show the feasibility of the procedure in a real network environment.

Summarising, the TRIDENT procedure offers two opportunities to the Network Operators, namely: 1) providing, automatically, Bandwidth on Demand services according to the dynamic bandwidth needs and 2) avoiding, in real-time, network congestions and allowing resource utilization optimization by applying traffic engineering rules. Furthermore it allows offering different classes of transport services since low priority traffic is transported only by low priority optical connections reserving thus high priority resources for the high priority traffic.

Thus, TRIDENT procedure really represents a cost-effective solution for Network Operators since it avoids the over-dimensioning of the network (enabling network CAPEX and OPEX reductions) while cope with both the requirements of the client network and of the GMPLS-based control plane.

The second contribution proposes a preliminary practical approach for ASON networks dimensioning purposes based on the approximate characterization of the traffic arrivals process. The simulation results indicate that the classical teletraffic theory (in particular the classical Engset analytical model) seems to be suitable for ASON dimensioning while an exact model is not available. The proposed analysis framework, in spite of its limits, aims to be useful as a reference scheme in real business cases both for Network Operators and Vendors.

The optimisation of the network resources is very critical in case of failures. Since failures can occur at each layer composing the network, as the third contribution we propose a novel multi-layer resilience strategy based on the interworking between the RPR layer and the optical layer to be applied in an IP/RPR over OTN metropolitan network scenario. Specifically, we propose the Double Hold-Off Timer (DHOT) as a coordination strategy between the recovery mechanisms of RPR and those implemented at the optical layer. It is a novel approach which improves the performance of the already proposed hold-off timer coordination approach (SHOT), allowing to the network to better and faster cope with different failure scenarios.

A recovery strategy to be implemented in real network environments, need to be simple from the implementation point of view, feasible, and need to minimize the time of the interruption of the provisioned services, which means to minimize the losses arising from the failure occurrence.

The suggested DHOT approach is simple since it is based on timers, it is feasible since it is even compatible with the current status of RPR and optical technologies and it allows to react to network outages in a fast and effective way. Moreover, simulation results demonstrate that the DHOT, compared with the hold-off timer approach, works better in terms of recovery since lower recovery times and lower traffic losses are obtained in the case of failures in the higher network layers (for above the optical transport layer). In this case, the DHOT approach allows a traffic losses reduction, in certain conditions, of about the 70% with respect to the SHOT.

Nevertheless, the RPR protection mechanisms, although very efficient and fast, provoke the substantial reduction of the available bandwidth due to the required ring reconfiguration. In order to avoid the ring saturation, the traffic inserted in the ring by the higher layers has to be kept limited, which implies very low bandwidth utilization in a failureless state.

The fourth contribution of this Ph.D. Thesis proposes the use of the automatic switching capability provided by the ASON/GMPLS networks to allow to keep the very high ring bandwidth utilisation provided by the RPR technology, even when the network reconfiguration is required (in case of failures). Specifically, we propose to implement in each IP/RPR router a monitoring function to monitor periodically the bandwidth utilization of the light paths connecting the IP/RPR routers and by using a threshold-based policy, to use the switched services to cope with the ring saturation and under-utilization. This procedure allows to react to the traffic fluctuations and on the other hand it avoids to limit the traffic inserted in the RPR ring, improving the bandwidth utilization.

8 Future Work

From the contribution provided by this Ph.D. Thesis many different lines of future work may arise.

Regarding the TRIDENT procedure, future investigations can deal with its signalling and routing aspects. As an example it will worth to investigate if current routing algorithms could be used in a network environment implying the frequent change of the network topology, above all in inter-domain environments.

Another line of future work regards the dimensioning of the ASON networks. The methodology suggested in this Ph.D. Thesis strongly depends from the procedure used to automatically adapt the available bandwidth at the optical transport layer on the basis of the fluctuations of the traffic of the client network. A further step implies to investigate the suggested methodology when the TRIDENT procedure is applied.

Regarding the RPR technology, when a new strategy is defined, it should interwork smoothly with existing protocols. For this reason, it will be future work the study of the interworking with the higher layers (for example the MPLS protocol) and with the underlying optical transport layer, not only related to the protection mechanism but also for the optimization of the network resources. As an example, the investigation of the applicability of the TRIDENT procedure to IP/RPR over ASON/GMPLS networks will be another point of future investigations.

PART III: System and Method for the automatic set up of switched circuits based on traffic prediction in a telecommunications networks (confidential)

This third part includes three Chapters. Chapter 9 presents in detail the TRIDENT procedure and Chapter 10 presents the experimental implementation of the TRIDENT procedure in a real environment, namely in the ASON-GMPLS testbed developed at the Telecom Italia Lab premises. In particular, the feasibility of the procedure is evaluated and some experimental results are depicted and discussed. Finally, Chapter 11 describes the generalization of the TRIDENT procedure to dynamic SDH-based networks, namely legacy SDH networks improved to better meet the current client's requirements by the application of the VCAT and LCAS functionalities.

9 TRIDENT procedure: Detailed description

The TRIDENT procedure deals with the light paths management in order to track, in an efficient and cost-effective way, the fluctuations of the client traffic allowing both CAPEX and OPEX reductions. The detailed description of the procedure is reported in this Chapter since there are still in progress patent protection activities and, thus, this Part III of the Ph. D. report will be kept confidential till December 2005.

The flow charts hereby reported refer to the handling of HP traffic bursts/surge between a couple of edge nodes (X and Y), initially connected through HP permanent light paths (LightP in the flow charts) supported by HP router interfaces.

Let us concentrate on the HP traffic carried by the HP light path ($LightP_n$) supported by the HP router interface *n*. The NLMS prediction algorithm allows estimating, at the end of an OW, the actual traffic carried by each of the LSPs being carried by $LightP_n$ for the next OW.

Let us assume that at the time t_0 (end of an OW), the predicted incoming data traffic at the HP edge router interface *n* ($B_n(t_0+OW)$) crosses the threshold of congestion (TH_{high}), which means that following congestion condition is detected:

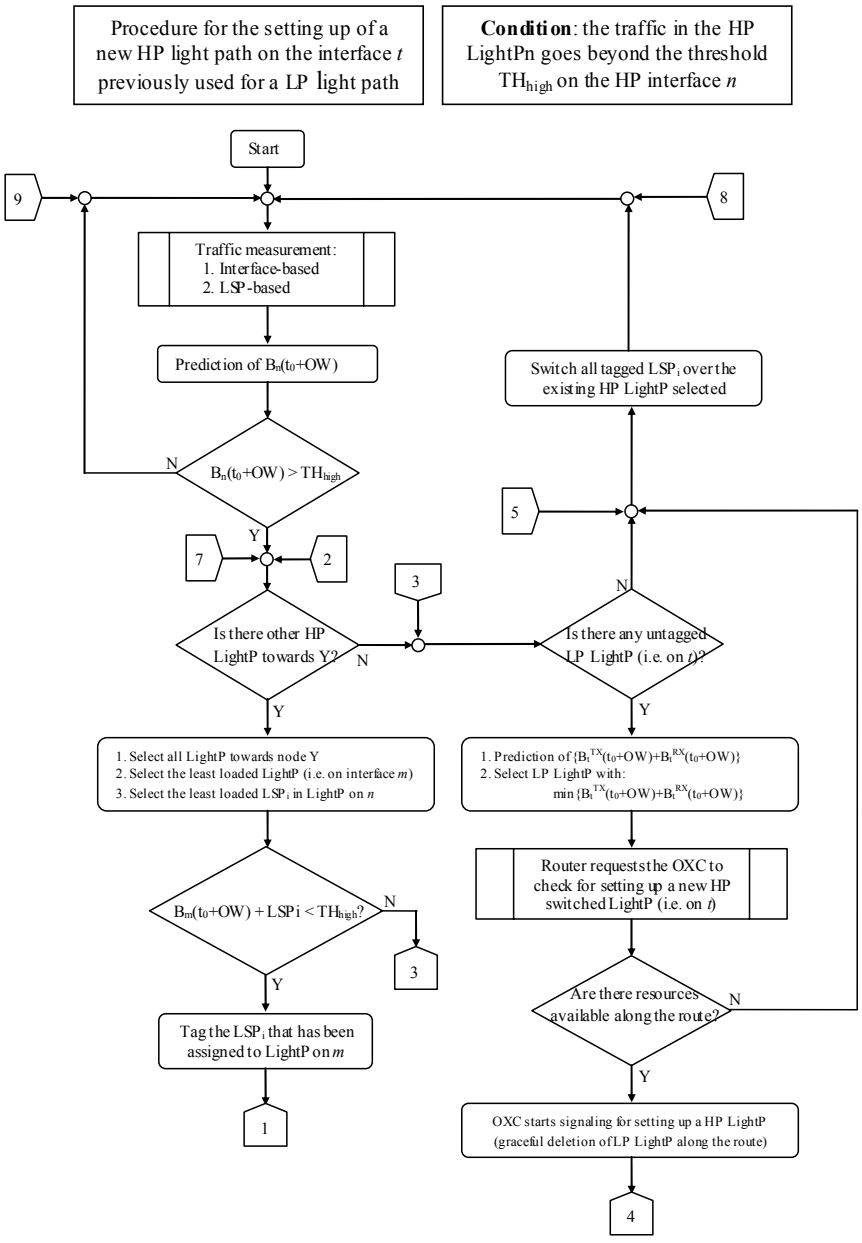
$$\sum_i LSP_i^{x,y}(t_0 + OW) = B_n^{x,y}(t_0 + OW) > TH_{high}$$

Figure 67 reports the complete flow chart describing how the TRIDENT procedure handles the HP traffic surge detected at the source node X (hereafter head-end node). Firstly, to manage the node congestion, a tentative to avoid, if possible, diverting low priority (LP) interfaces with consequent LP traffic losses is done. In order to do this, TRIDENT implies the rerouting of some LSPs from the interface *n* to other HP router interface supporting another established HP light path towards the same destination edge node Y (e.g., the HP light path supported by the HP router

interface *m*). In this case, the procedure starts to select the LSPs to be rerouted towards the router interface *m*. The procedure selects and tags the least loaded LSP, namely:

$$LSP^{x,y}{}_{selected} = \min_i \{LSP_1^{x,y}, LSP_2^{x,y}, ..., LSP_i^{x,y}\}$$

Then, both $B_n(t_0+OW)$ and $B_m(t_0+OW)$ are updated. This process is repeated until the congestion condition is not solved ($B_n^{x,y}(t_0 + OW) < TH_{high}$) or the potential reroute of the LSP_i generates a congestion condition on the HP interface *m* ($(B_m^{x,y}(t_0 + OW) + LSP_i) > TH_{high}$).



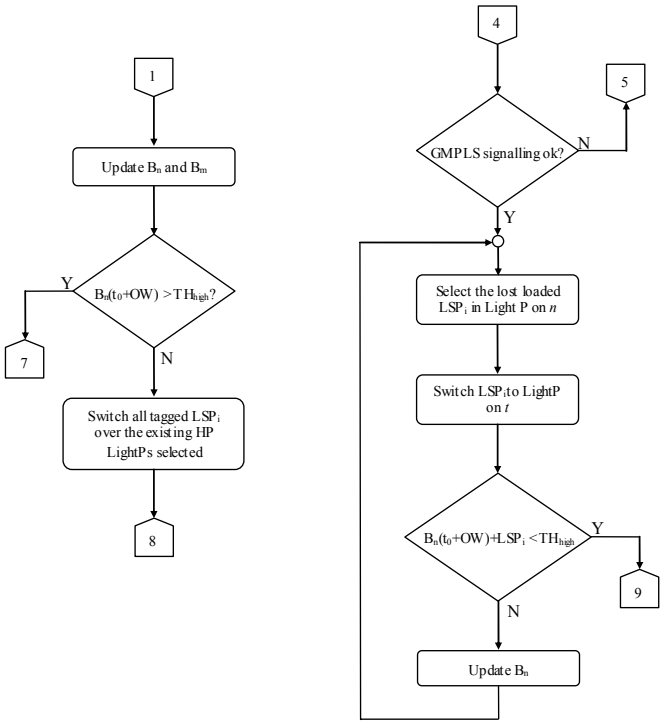


Figure 67: Handling HP traffic burst at Head-end node (node X)

The reroute of the selected LSPs is carried out at the end of the congestion management. This way implies that if the congestion cannot be solved by using the already established HP light paths towards node Y, the LSPs rerouting is not carried out. This is to avoid to stress the CP with continues rerouting. In case the available bandwidth on the other HP light paths is enough to accommodate the traffic burst, the tagged LSPs are rerouted once the congestion condition is solved.

On the other hand, if another already established HP light path towards node Y does not exist or the congestion condition cannot be solved by the rerouting of some LSPs, TRIDENT foresees the tear down (with a graceful deletion) of a LP light path (if available) to be used temporarily as HP light path (e.g., on the interface *t*) to cope with the HP traffic burst on the HP interface *n*. To minimize the traffic losses, the least loaded low priority (LP) bi-directional light path is selected to be torn down. The LP traffic can be rerouted towards other LP interfaces (if available) or it is dropped. Then, the router requests the OXC connection controller (CC), via UNI or via internal signalling, to check resource along the route calculated by the control plane for setting up the new HP light path. If the resources are available along the route, the OXC starts signalling for the setting up the new HP light path.

In the case the GMPLS-based signalling is able to set up the required switched light path between nodes X and Y, the procedure includes the reroute of some LSPs (starting from the most loaded one) from the permanent light path experimenting the congestion ($LightP_n$) to the new set up HP light path (e.g., $LightP_t$) until the predicted traffic at the router interface *n* goes down the threshold of congestion (TH_{high}). Taking into account also signalling aspects for the TRIDENT procedure, Figure 68 describes, as a consequence of the head-end OXC (node X) request for a new HP light path, the way the tail-end OXC (node Y) searches an interface *t* at the tail-end client router to serve the request of setting up a new light path. It searches for an available LP interface (e.g., *t*) and it is chosen on the basis of the LP traffic carried by the $LightP_t$. If such interface is available, the tail-end node once rearranged the traffic being carried by light path on *t*, tags the interface as HP and it signals to the head-end node the acceptance of its request. If LP interfaces are not available, the tail-end node refuses the head-end node request and the temporarily required HP switched light path cannot be established.

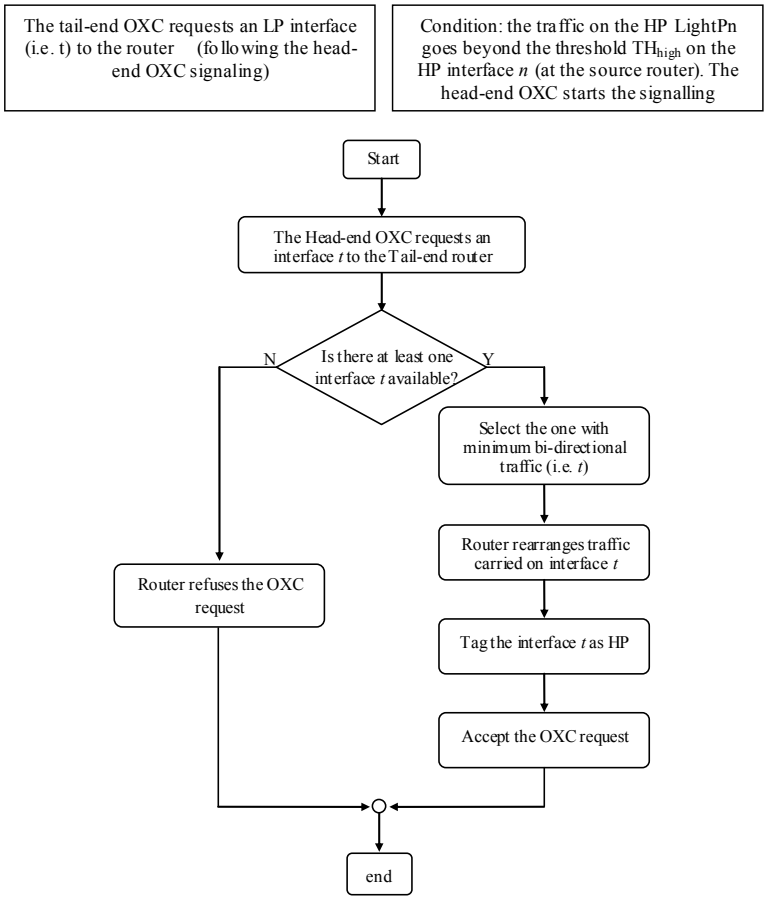


Figure 68: Handling HP traffic burst at Tail-end node (node Y)

In such a way, an additional HP switched light path is used temporarily to face with the HP traffic bursts. When such bursts end, the TRIDENT procedure implies the tear down of the additional HP switched light path. Specifically, Figure 69 describes the procedure of the tearing down (at the head-end node X) of the HP light path on the interface *t*, previously set up as HP light path to handle the HP traffic burst. It is due to the detection of the end of the burst. The action is started when the predicted traffic at the router interface *n* ($B_n(t_0+OW)$) crosses the threshold of under-utilization TH_{low}.

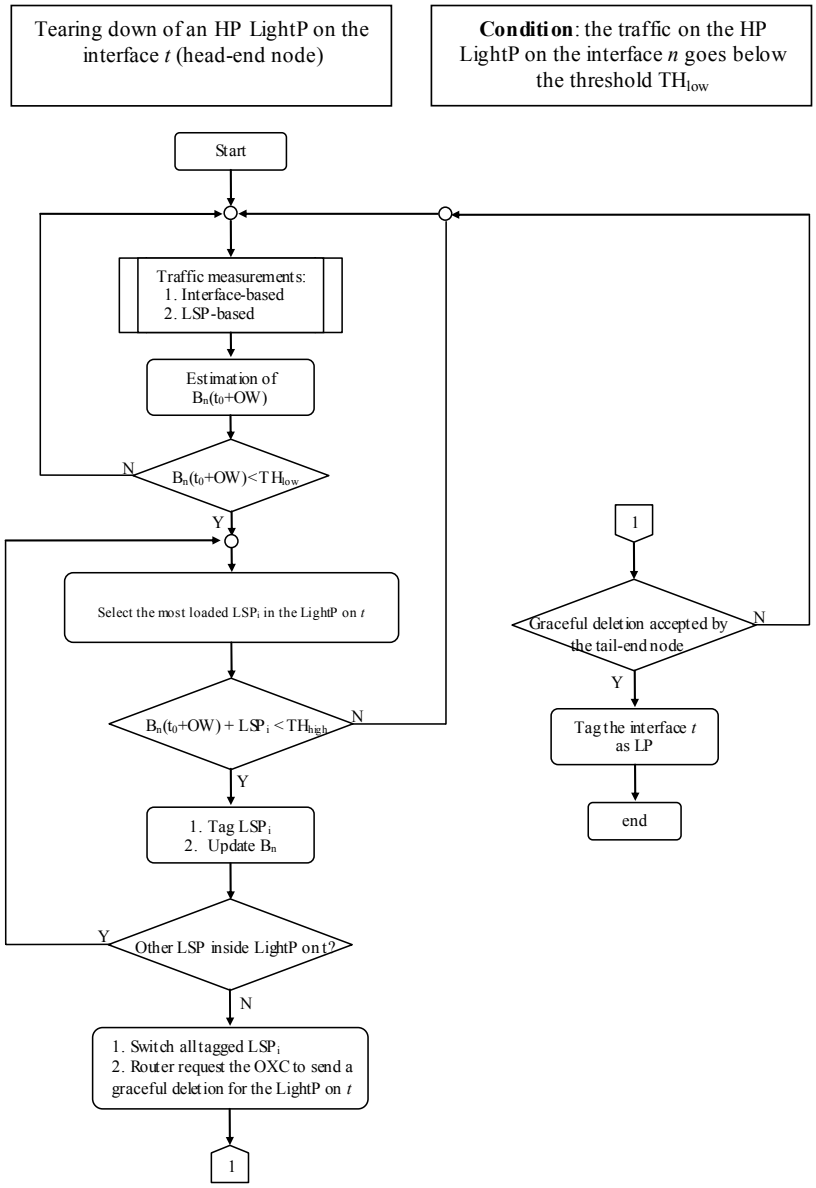


Figure 69: Tearing down the HP switched light path at the Head-end node (node X)

The TRIDENT procedure foresees that when on the permanent HP light path, the predicted traffic goes down the threshold of under-utilization (TH_{low}), all the HP LSPs that are being carried by the switched connection (supported by interface *t*) temporarily used as HP light path, are rerouted to the permanent light paths. If it is possible, then the head-end node requests the tear down of the HP light path in order to be set up as LP light path to carry LP traffic (Figure 69).

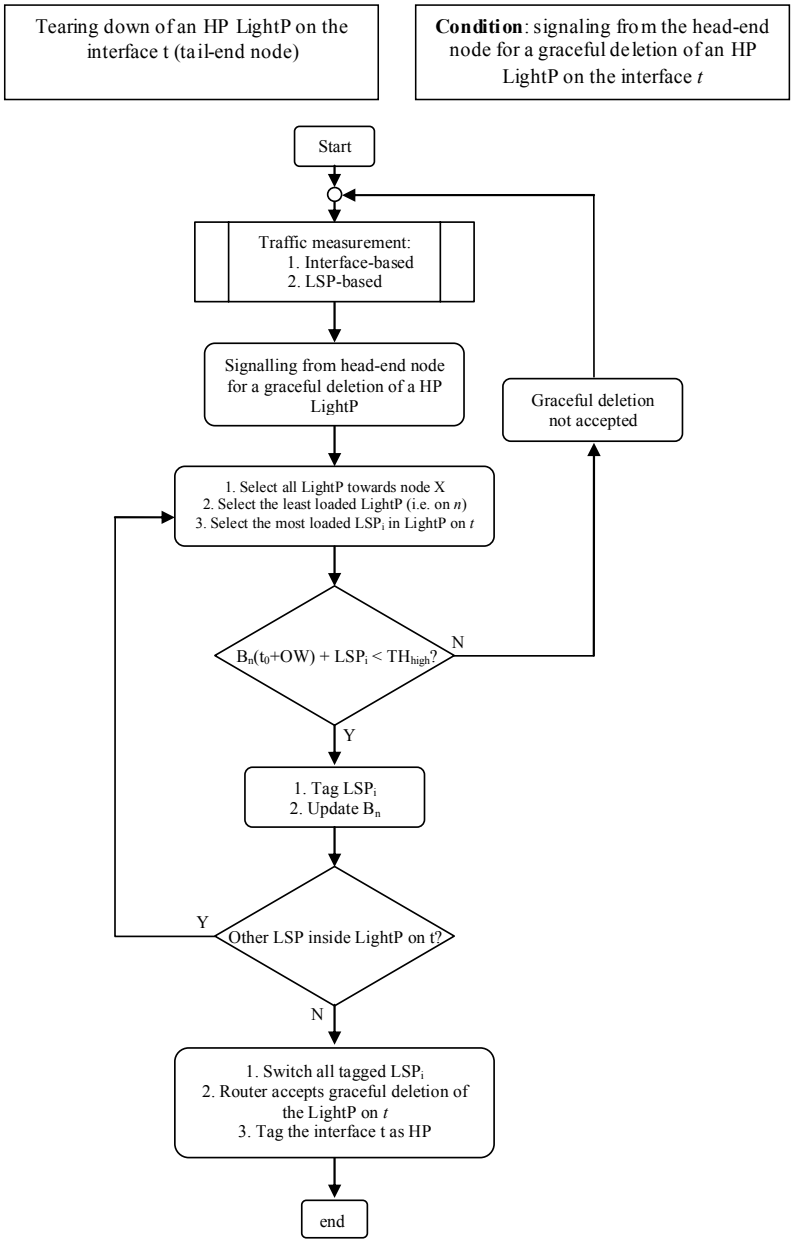


Figure 70: Tearing down the HP switched light path at the Tail-end node (node Y)

Figure 70 describes the graceful deletion of the HP light path on the interface *t* on the tail-end node as a consequence of the signalling message from the head-end node for a graceful deletion of the HP switched light path. Since it was assumed to consider bi-directional light paths, the graceful deletion is accepted by the tail-end node only if all the LSPs being carried by the HP $LightP_t$ can be rerouted towards other HP light paths already established towards the head-end node (node X).

10 TRIDENT procedure: Experimental implementation

In order to verify the feasibility of the designed TRIDENT procedure in an experimental environment and to complement the simulation results discussed in Chapter 3, in this Ph.D. Thesis we also carried out its physical implementation in an ASON/GMPLS testbed. Specifically, the experiments were done in the testbed developed within the framework of the IST-1999-11387 LION project and currently located at Telecom Italia Lab (TILAB) premises in Turin, Italy.

10.1 Network environment

The TILAB ASON/GMPLS testbed is composed by 6 Fibre Switch Capable (FSC) Optical Cross Connects (OXCs) and IP routers 12000 series from Cisco Systems (Figure 71) [109]. The transport network and the Data Communications Networks (DCN), which supports the control plane, are independent. Specifically, the DCN is based on Fast Ethernet technology while for the transport a FSC-based ASON network is implemented.

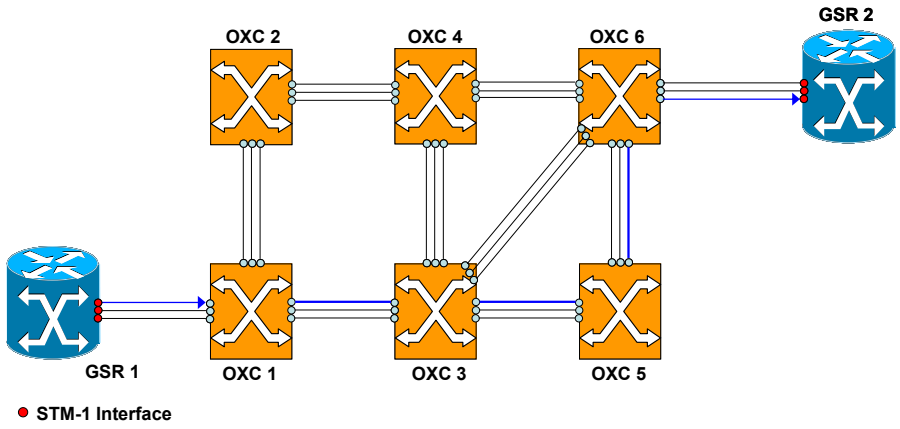


Figure 71: TILAB ASON/GMPLS testbed

Considering the GMPLS hierarchy [19], the FSC-capability means that the OXCs switch from one incoming fibre to outgoing fibres. This switching capability is due to the utilization of optical switching matrixes based on Micro Electro Mechanical Systems (MEMS) technology [2]. These optical switching matrixes are moreover able to measure the incoming optical power. In this way, for example a fibre cut can be detected and then the pre-defined resilience mechanism can be triggered (e.g. protection or fast-restoration).

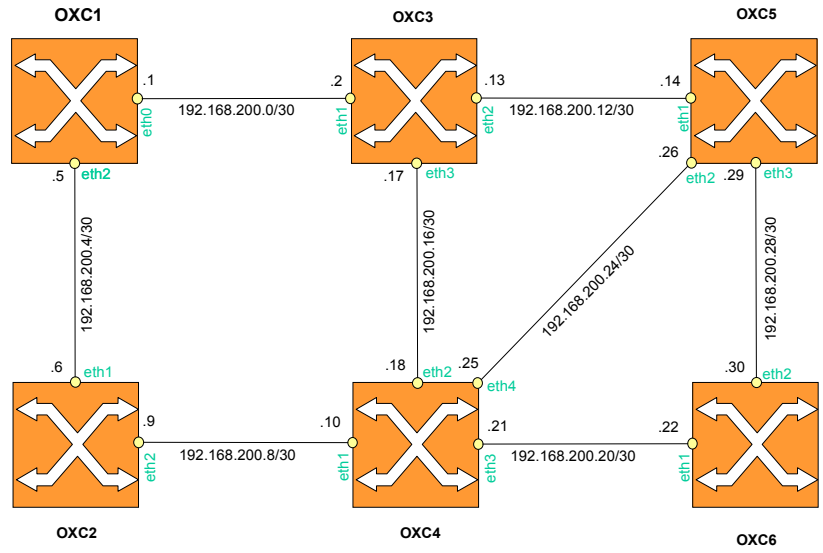


Figure 72: CP implementation using Fast Ethernet technology

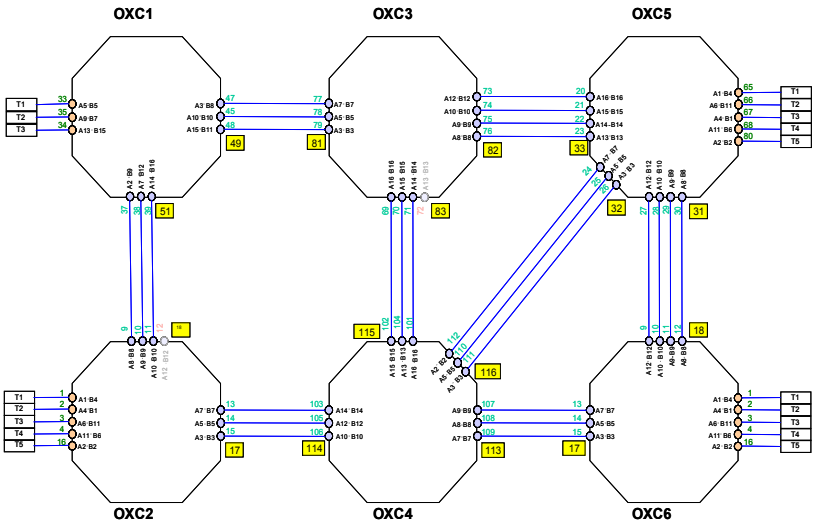


Figure 73: Transport Plane topology

On the other hand, the control plane is composed by 6 PCs which execute the process corresponding to LRM, NNI, CC and the RC (see Chapter 2).

It has to be highlighted that the current development of the testbed has limited GMPLS functionalities; therefore the implementation relies on a simplified version of the TRIDENT procedure. Specifically, the MPLS level is not currently available at the testbed and therefore the testbed CP is not able to carry out the TE rules (discussed in the previous Chapter and based on the rerouting of the LSPs) designed to optimize the utilization of the bandwidth of the light paths. Moreover, to make the implementation of the procedure easier, only the high priority (HP) traffic was considered and spare Synchronous Transport Module (STM)-1 interfaces instead of low priority interfaces are used when the resources dedicated to the HP traffic are not sufficient (i.e., in case of high priority traffic surge).

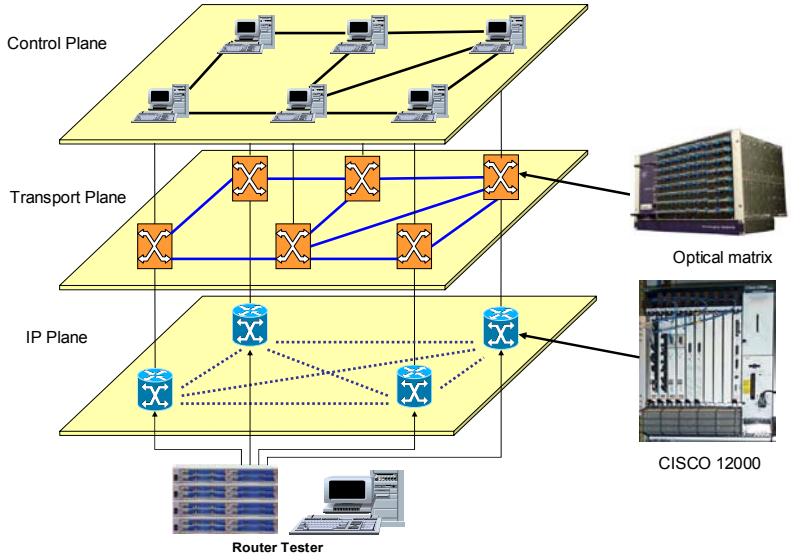


Figure 74: ASON test-bed: Transport and Control Plane

The first component required to implement the procedure is the HP traffic generator. To this purpose, a Router Tester from Agilent Technologies was used (Figure 75) [115]. Using the C programming language, a program was implemented in the router tester in order to generate traffic between the two routers (from GSR1 to GSR2). The program is able "to read" from a file of samples of traffic measurements extracted from the Catalan R&A network monitoring. The generated HP traffic is in the order of hundreds of Mbit/s.

Initially some STM-1 interfaces are set up and additional STM-1 interface are used as spare interfaces when the procedure requires to set up/tear down an additional interface to cope with the traffic bursts/surges.

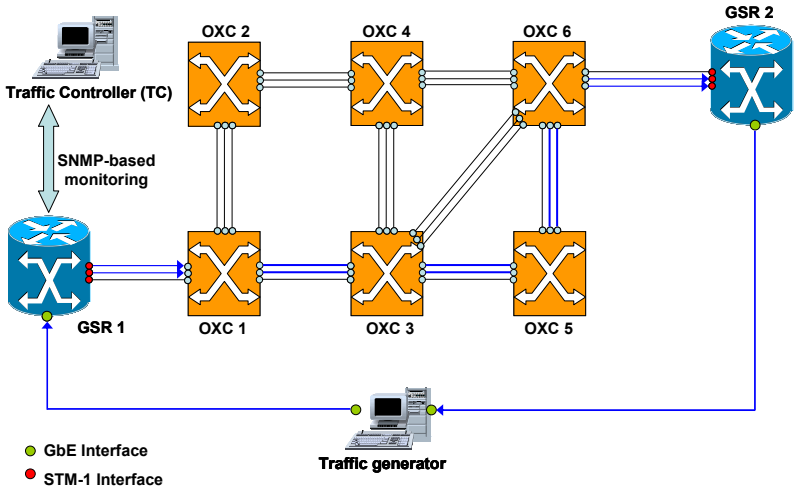


Figure 75: Complete system: Initially conditions

The traffic generator basically is in charge of the following functions: 1) Generation of traffic through a Gigabit Ethernet (GbE) interface, 2) Reception of this traffic after crossing the optical network through another GbE interface and, finally 3) Calculation of the packets lost at the router interface (PLR) as the ratio between the generated and the received traffic. The PLR parameter is used as one of the metrics for the performance evaluation of the procedure.

As showed in Figure 75and Figure 76 the whole system includes some blocks required to test the feasibility of the procedure. Next, we will describe the different blocks.

The physical implementation of the TRIDENT procedure requires the introduction of a new element called Traffic Controller (TC) that implements the procedure components and hence manages the available bandwidth. These components are the traffic monitoring and short-term prediction and the automatic set up/tear down of switched connections. The implementation of the TC component arose two main problems, namely: 1) How to carry out the monitoring of the traffic and thus of the required bandwidth and, 2) How to make the set up and/or the tear down of an additional connection. To resolve the first problem we used SNMP software (See [72]) to make a polling of the number of bytes that cross a router interface. Every 20 seconds, a request to the IP router is triggered for how many bytes have passed through the interface. After two consecutive requests at time t_0 and t_1 respectively, then, the required bandwidth by the traffic crossing the interface is calculated using the following formula:

$$Required\ Bandwidth\ (Mbit/s) = \frac{\left(OutOctets_{t_1} - OutOctets_{t_0}\right)\cdot 8}{\left(t_1 - t_0\right)\cdot 10^6}$$

On the other hand, to make feasible the request for the automatic set up/tear down of connections, a Craft Terminal (CT) has been developed. Basically the CT opens a socket to talk to the Control Plane (CP) of a node interchanging strings, and this way, the CT allows the request to the CP the connection set up/tear down (Figure 76).

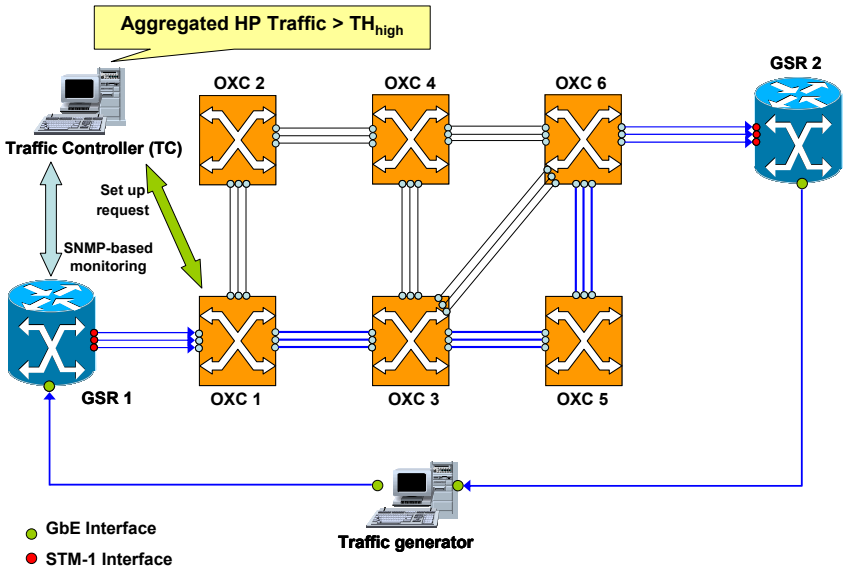


Figure 76: Complete system: After tracking the HP traffic burst

For traffic routing purposes, the link bundling concept was used, which means that the data traffic towards the destination node is routed using different interfaces/connections. This is called inverse multiplexing (Figure 77).

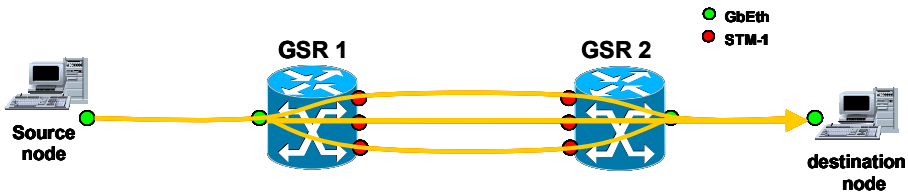


Figure 77: Inverse multiplexing

Specifically, a per-packet load balancing scheme was used. Specifically, if three interfaces are activated to reach the router destination the first packet is sent through the first interface, the second packet through the second interface and the third packet through the third interface (Figure 78).

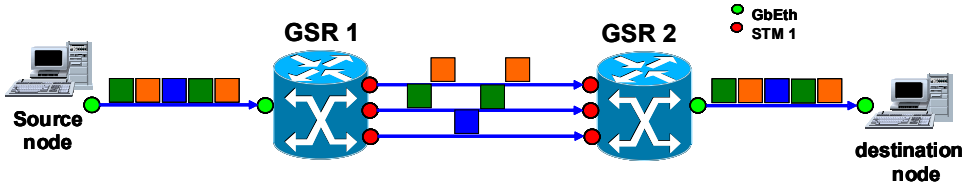


Figure 78: Per-packet load balancing

10.2 Set up delays

The set up time is the delay between a set up request and the time the first packet is carried trough the new connection. This delay is due to a group of delays including, among others, the switching time of the optical matrixes. However, the most important delays are the anti flap time and the routing time. The anti flap time is a protection measure of the Cisco Systems routers aimed to avoid routing recalculations in presence of spurious carrier detections. This delay is useful in a traditional scenario but in a dynamic scenario could increase the number of loss packets. However it can be modified by the users.

The second big delay is produced by the routing time, this is the time that a router needs to recalculate the paths and refresh the routing tables. This delay depends on the routing algorithm used and its configuration. As an example, using OSPF routing algorithm the "*Hello interval*" and the "*time between SPF*" has to be set to the minimum value.

In our tests we have measured a routing time of 2 seconds using static routing and 7 seconds using OSPF when "*Hello intervals*" of 1 s long were considered.

10.3 TRIDENT procedure: Experimental results

As a sample of the results obtained in the experiments, the following Figures depict the number of STM-1 connections required to transport the exemplary of HP traffic generated by the router tester and based on the monitoring of the Catalan R&A network. The aim of the experiments has been twofold: 1) To evaluate the feasibility of the implementation of the TRIDENT procedure, although simplified, in a real network environment and 2) To complement and confirm the conclusions obtained by the simulation results.

The generated traffic represents the HP traffic from one source node (GSR 1) towards a destination node (GSR 2) from 7 a.m. to 21 p.m. (about 14 hours). Initially, two STM-1 connections were established between the nodes and a third STM-1 interface has been used as spare interface to be automatically set up/torn down when required by applying the procedure.

Observation Window 5 and 1 minutes long were used and the TH_{high} and the TH_{low} were set to 99% and 96% of the capacity equivalent to two STM-1 interface capacity (296 Mbps).

Specifically, as Figure 79 shows when OW of 5 minutes was used, the number of STM-1 connections needed to carry the HP traffic rises and falls according to the traffic dynamics. The

number of the set up requests is equal to 8 and the mean Holding Time is 135 min. The measured PLR is 3.8 E-04.

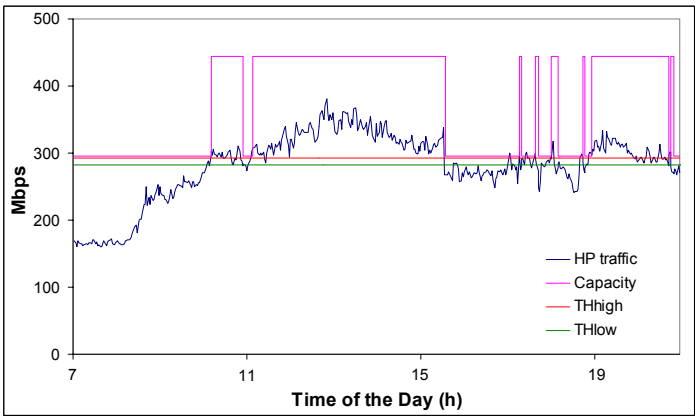


Figure 79: Experimental result, number of STM-1 connection used

As discussed in Chapter 3, to avoid requesting the set up/tear down of connections too often (increasing in this way the routing and signalling functions cost and the potential instability at router level), we implemented a conservative approach. Such conservative approach resides on to request the set up of an additional STM-1 connection only when the experimented congestion condition (i.e., HP Traffic > TH_{high}) is repeated again for *m* consecutives OWs. On the other hand, the requests for the tear down of a switched connection is requested only when the experimented under-utilization condition (HP traffic < TH_{low}) is repeated again for *n* consecutives OWs. Figure 80 reports the number of STM-1 connections when the conservative approach is applied and OW is set to 5 minutes.

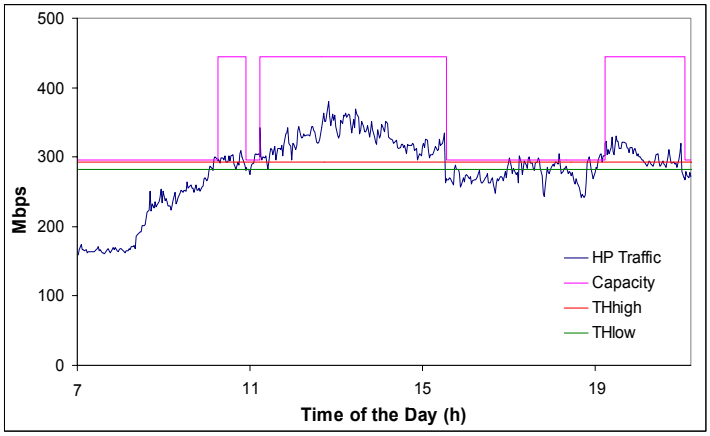


Figure 80: Experimental result, number of STM-1 connection used using the conservative approach

Specifically, the figure reports the case with $m = 1$ and $n = 1$. It can be observed that the number of set up requests in this case is 3 (lower than the previous case) while the PLR is equal to 5.53E-04 (higher than the previous case). The mean HT is equal to 135 min.

The following figures depict the number of STM-1 connections required to carry the HP traffic considering different values both for m and n; specifically, increasing the number of m or n, the number of the requests to the CP for the set up of the spare connection is lower and the percentage of time that the spare resource is used to transport the traffic is higher, reducing in this way the potential maximum capacity available to transport LP traffic.

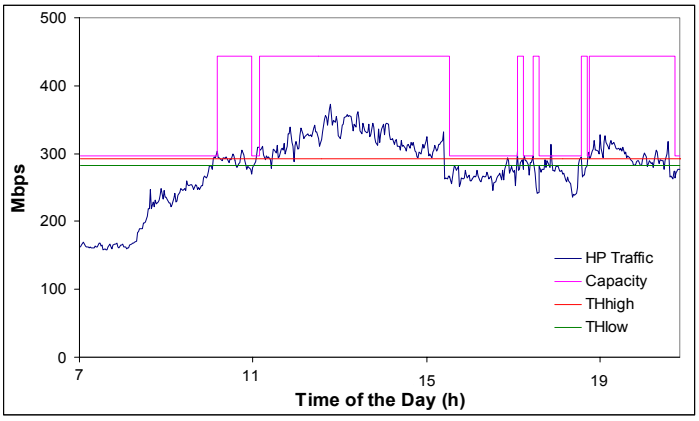


Figure 81: Experimental results, conservative approach (m=0, n=1)

However, increasing the parameter m means to delay the set up of the spare resource required once the congestion condition is detected. Thus, the higher m is, the higher is the measured PLR. On the other hand, increasing parameter n means holding on the spare connection even if not strictly required. The effect is to decrease the bandwidth utilization of the permanent connections but, at the same time, to decrease the PLR.

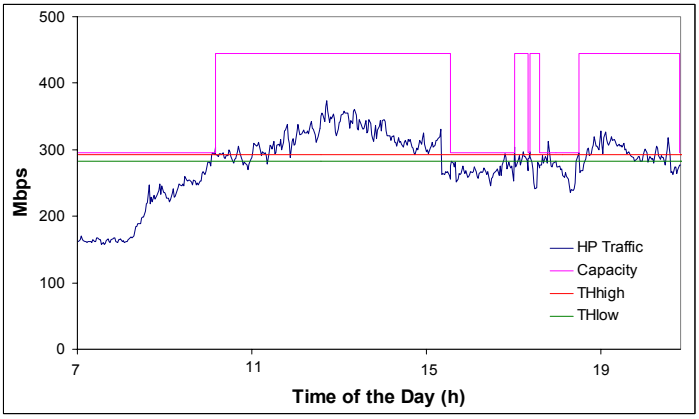


Figure 82: Experimental result, conservative approach, (m=0, n=2)

From the analysis of the results, we conclude that the best compromise is to set the conservative approach, but it is better to use it for the tear down procedure. For completeness reason, the following Table 15 summarize the results we obtained considering different configurations of the conservative approach and OW = 5 minutes.

OW = 5 min TH_{high} = 99%, TH_{low} = 96% TRIDENT using the conservative approach			
		SC usage time[1] (%)	PLR
m = 0	n = 1 n = 2 n = 3	54.3 59.3 63.2	3.76E-04 2.42E-04 8.98E-05
m = 1	n = 1 n = 2 n = 3	42.7 44.5 45.6	8.24E-04 6.44E-04 6.43E-04
m = 2	n = 1 n = 2 n = 3	39.7 41.5 42.7	1.32E-03 1.14E-03 1.13E-03
m=3	n = 1 n = 2 n = 3	38.5 41.5 41.6	1.54E-03 1.35E-03 1.34E-03

[1]calculated over the total time

Table 15: Conservative approach; OW = 5 min, Summary of results

We carried out experiments also considering OW 3 minutes long and applying the conservative approach. As an example, the Figure 83 reports the number of STM-1 connections required to carry the traffic between the considered nodes. Specifically, the Figure refers to the case of *m*=2 and *n*=2.

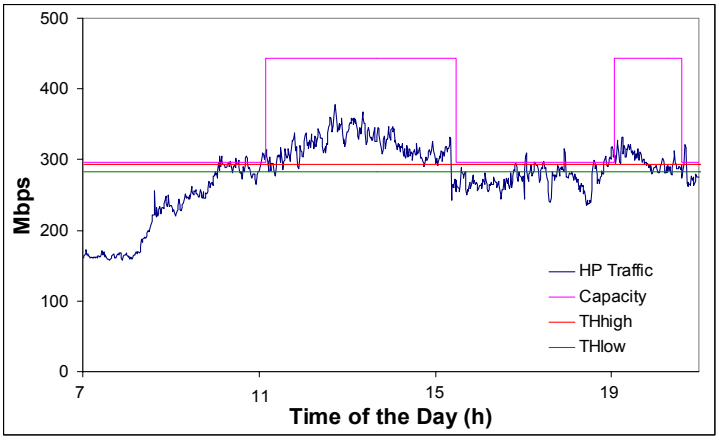


Figure 83: Experimental result using the conservative approach, OW = 3 min

Table 15 summarizes the results obtained with OW of 3 minutes and with the different configurations of the conservative approach. The behaviour is the same observed in case of using OW of 5 minutes.

OW = 3 min TH_{high} = 99%, TH_{low} = 96% Using the conservative approach			
		SC usage time[1] (%)	PLR
m = 0	n = 1	55.10	3.52E-04
	n = 2	57.65	3.49E-04
	n = 3	59.78	3.48E-04
m = 1	n = 1	47.33	8.59E-04
	n = 2	49.11	7.57E-04
	n = 3	50.18	7.56E-04
m = 2	n = 1	39.85	1.40E-03
	n = 2	42	1.30E-03
	n = 3	43.42	1.04E-03
m=3	n = 1	39.15	1.59E-03
	n = 2	41.28	1.49E-03
	n = 3	42.70	1.23E-03

[1]calculated over the total time

Table 16: Conservative approach, OW = 3 min, Summary of results

These experimental results show that the TRIDENT procedure can be applied to a real environment since it is feasible both the traffic monitoring and short-term prediction even with the current technology. Once the testbed will be further developed, the complete version of the procedure will be implemented and tested. On the other hand, the experimental results are in line with the simulation results since they show that the parameters of the procedure have the same influence on the performance of the procedure itself.

11 TRIDENT procedure: Generalization to dynamic SONET/SDH networks

The TRIDENT procedure has been explained till now by making specific reference to reconfigurable DWDM networks, specifically to an exemplary IP/MPLS over ASON/GMPLS optical network, in which a single, circuit switched server layer (ASON/optical WDM layer) is associated to a packet switched client layer (IP/MPLS). In this Chapter, we describe the generalization of the procedure in order to take into account different transport layers. Specifically, we considered such generalization taking into account the current enhancements for the SONET/SDH networks (Next Generation SDH networks). The corresponding scenario corresponds to a medium-term solution for the optical networking.

Indeed, circuit switched networks in which other server layers are used in combination with an optical WDM layer can take benefits from the implementation of the procedure. As an example, the transport network may be configured as a TDM network (e.g., SONET/SDH network) using TDM circuits in combination with WDM circuits (i.e., light paths). This basically means introducing more switching granularity and therefore to increase the bandwidth utilization of the light paths.

In such a case, TDM circuits (e.g., STM circuits, and/or virtual container circuits [112]), can be also tagged as high priority and low priority circuits, and be subjected to the TRIDENT procedure.

SONET/SDH networks were initially designed to optimize transport of 64-kbit/s-based TDM services and a rigid capacity of payload as well as a coarse fixed-rate multiplexing hierarchy was defined. Today, SDH/SONET systems are built with bit rates as high as 10 Gbit/s (STM-64/OC-192), with 40 Gbit/s (STM-256/OC-768) on the horizon. Current SDH/SONET core networks are characterized by a switching granularity of VC-4/STS-3.

By the use of Virtual Concatenation (VCAT) and Generic Frame Procedure (GFP), SDH/SONET may be improved to better meet today's bandwidth requirements, e.g., various switching granularities. Virtual Concatenation allows flexible concatenation of several SDH/SONET payloads [14] and [112]. In such a way, it assures an effective use of the SDH/SONET capacity. Virtually concatenated payloads (members) constitute a Virtual Concatenation Group (VCG). Members of a VCG, as opposed to the so-called contiguous concatenation, may not reside in the same STM-N/OC-N contiguously. They may even reside at different STM-N/OC-N interfaces and are treated within the network separately and independently (See Figure 84). As a consequence, members of a VCG may reach the destination node through various routes. Intermediate nodes do not need to handle virtual concatenation and then the VC functionalities must be implemented only at path termination nodes.

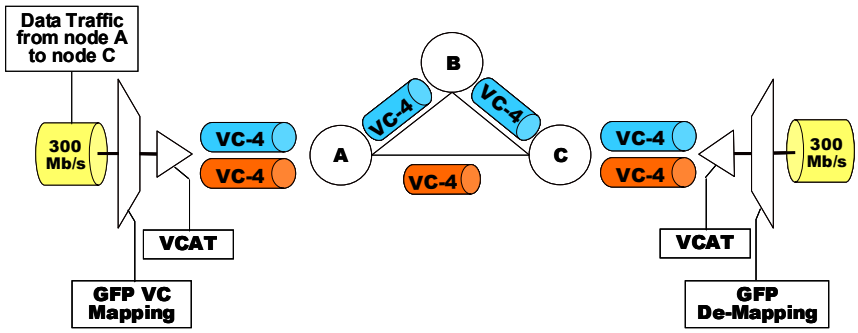


Figure 84: SONET/SDH Networks, Virtual Concatenation concept

There are different typologies of virtual concatenation. The way of specify them is VC-n-Xv where n defines the capacity of the containers and X, the number of the virtual containers concatenated (VC-4-Xv in case of concatenation of VC-4). A VC-4-Xv provides a contiguous payload area of X Container-4 (VC-4-Xv) with a payload capacity of X*149,60 Mbit/s. The container is mapped in X individual VC-4s which form the VC-4-Xv. Each VC-4 of the VC-4-Xv is transported individually through the network. Due to different propagation delay of the VC-4s, a differential delay will occur between the individual VC-4s. This differential delay has to be compensated and the individual VC-4s have to be realigned for access to the contiguous payload area. The sequence indicator (SQ) identifies the sequence or order in which the individual VC-4s of the VC 4-Xv are combined to form the contiguous container VC-4-Xv. Each VC-4 of a VC-4-Xv has a fixed unique sequence number in the range of 0 to (X–1). The VC-4 transporting the first time slot of the VC-4-Xv has the sequence number 0, the VC-4 transporting the second time slot the sequence number 1 and so on up to the VC-4 transporting time slot X of the VC-4-Xv with the sequence number (X–1).

This feature makes it possible to deploy virtual concatenation on legacy SDH/SONET equipment of existing networks, thus it constitutes a short-term solution and a smooth transition to enhanced data-driven transport networks. Another advantage of virtual concatenation is its ability to divide STM-N/OC-N bandwidth into several sub-rates. Each of the sub-rates may be used for accommodation of a different service. The bandwidth of STM-N/OC-N may be shared, for example, by both telephone service and data signals. The authors in [14] make an example of a practical use of virtual concatenation when Gigabit Ethernet frames have to be transported. Under conventional SDH networks, an STM-16 circuit is required to accommodate Gigabit Ethernet signals. However, the capacity of 1.4 Gbit/s is then wasted. On the other hand, contiguous concatenation of four VC-4 containers provides too small capacity to fully accommodate Gigabit Ethernet signals. In such a case, the best solution would be the concatenation of seven VC-4 payloads. It is possible with virtual concatenation. Bandwidth of 1.05 Gb/s provided by a VC-4-7v VCG is suitable for Gigabit Ethernet.

For the mapping of the client traffic to VCs, the GFP procedure can be used. The Generic Framing Procedure defines a very effective way of mapping a wide variety of data signals into the transport network [113] and [114]. It is a flexible and simple mechanism to adapt client signals. Specifically, it adapts the traffic from higher-layer client signals over SONET/SDH, OTN or dark fibre into a common format. The ITU-T recommendation defines two transport modes. The first one, referred as *Frame-Mapped GFP* (GFP-F) is optimized for the adaptation of Payload Data Unit (PDU)-oriented streams such as IP, MPLS or Ethernet traffic. The second mode, optimized for block-code-oriented streams, is called *Transparent GFP* (GFP-T). This mode is used for GbE, Fibre Channel, FICON and ESCON traffic.

From the functional point of view, the procedure is divided in two levels. The first one takes charge of common aspects of client signals, particularly defines a frame structure, independently of the kind of data signal. The second one specifies operations that it has to be done for adapting the client signal to the GFP frame structure.

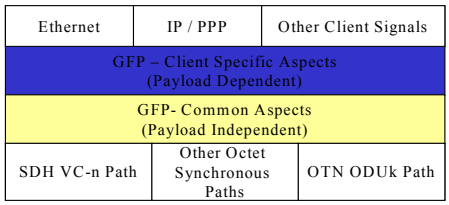


Figure 85: GFP's relationship to payloads and SONET paths

The GFP anticipates two topologies of frame structure, Client Frame and Control Frame. The first one has variable length and is divided in the field Core Header and the field Payload Area. It is used to transport client signals and for the operation OA&M (Client Management Frame). The second one has fixed length and it is only used as "stuffing" in the Rate Adaptation Procedure.

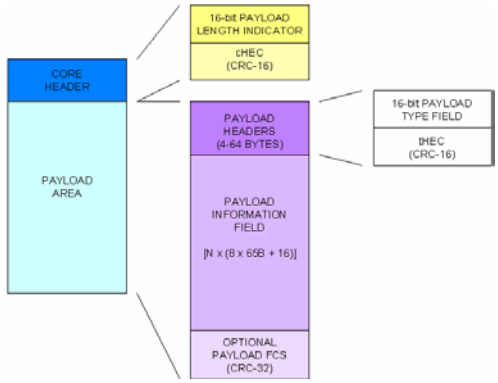


Figure 86: GFP Frame Format

The transport of the GFP frames can be done by VC-4 virtual or contiguous concatenated. The mapping process is divided in two stages. In the first stage, the GFP frame is inserted in a C-4 multiple payload. In the second stage the typology of the VC-4 concatenated is selected, that means the VC-4-X typology selection, where X can be 4, 16, 64 or 256 in the contiguous concatenation and 2 to 256 in the virtual concatenation. This selection is done by the typology of the client signal.

The mapping procedure of GFP in the SDH structure is specified in [112] and therefore is out of the scope of the Thesis.

Figure 87 shows, in a schematized view, different possible server layer segmentations used by the client IP/MPLS packets.

The packets can be mapped directly on switched circuits at the optical server layer (i.e., light paths, indicated as OCh in Figure 87), as in the so far description of the TRIDENT procedure. In another possible scheme, packets are first mapped in Optical Digital Unit (ODU) circuits, and then the ODU circuits are mapped in OCh circuits; in another possible scheme packets are first mapped in Higher Order Virtual Container (HOVC) circuits, and then the HOVC circuits are mapped in OCh circuits; in another possible scheme, packets are first mapped in Lower Order Virtual Container (LOVC) circuits, then LOVC circuits are mapped in HOVC circuits, then HOVC circuits are mapped in ODU circuits, then ODU circuits are mapped in OCh circuits, thus exploiting all possible segmentation server layers.

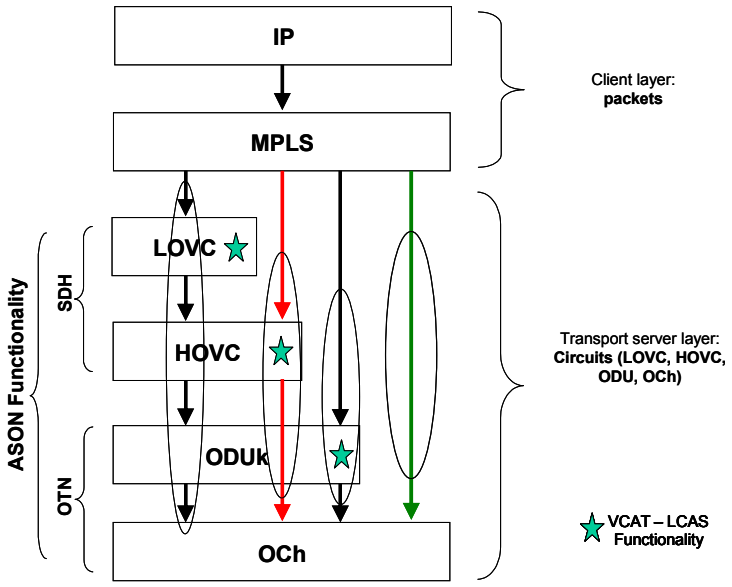


Figure 87: Server layer segmentation, General Model

The classification in high priority and low priority can be applied to switched circuits belonging to any "server" layer. Preferably, if the client traffic is mapped onto different nested switched circuits, the classification in high priority and low priority is performed at all server layers used, so that the lowest hierarchy high priority server layer is adapted to transport high priority traffic packets, and higher hierarchy high priority server circuits are adapted to transport lower hierarchy high priority server circuits. The same it can be applied for low priority traffic, as well as for lower hierarchy and higher hierarchy low priority switched circuits. However, it has not to be excluded that lower hierarchy low priority switched circuits could be mapped onto higher hierarchy switched circuits tagged as high priority switched circuits, during periods of under-utilization by the high priority traffic. This allows to increase the bandwidth utilization of the circuits. As a consequence, the basic idea is that the TRIDENT procedure can also be applied to any and/or all the server layers of Figure 87. Specifically, after the detection of a high priority traffic burst, the tearing down of a low priority switched circuit, and a consequent set up of a new, temporarily, high priority switched circuit can be adopted at any suitable server layer, according to the needing.

The main advantage in using different nested server layers is that data traffic can be managed more efficiently, since a number of possible routing solutions can be adopted, among which the best one can be eventually chosen. As explained above, VCAT in a SONET/SDH network allows a breaking of the traffic bandwidth into individual virtual containers belonging to the same VCG. The individual virtual containers can be routed onto different light paths having the same destination

node, and then they will be recombined together at the destination node. This may avoid the set up of a new higher hierarchy switched circuit in order to manage a possible congestion of a node interface. Furthermore, a higher granularity of intervention, even in case of detection of a burst, can be exploited in a network using a plurality of server layers. For example, in case of the detection of a possible congestion in a network node due to the high priority traffic, a first intervention may include an increase of the capacity assigned to a VCG, by addition of a suitable number of VCs, until the maximum capacity of the virtual container group is reached. If such procedure does not sufficiently cope with the traffic burst, a tearing down of a low priority virtual container, and/or of a higher hierarchy low priority switched circuit, may be performed in order to make available resources within the network for the burst of the high priority traffic. On the other hand, when an imminent end of the high priority traffic burst is detected, a first intervention in the opposite direction may be of progressively decreasing the capacity of a new temporary high priority virtual container group previously set-up after the detection of the burst, before performing a complete tearing down of the temporary virtual container group.

A further granularity of intervention may be provided by the Link Capacity Adjustment Scheme (LCAS) functionality ([14], [15]), which may act to vary the bandwidth assigned to at least a portion of the virtual containers when a traffic fluctuation is detected. LCAS in the virtual concatenation source and sink adaptation functions provides a control mechanism to hitless increase or decrease the capacity of a VCG link in order to better meet the bandwidth required by the application. Generally speaking, it also provides the capability of temporarily removing member links that have experienced a failure.

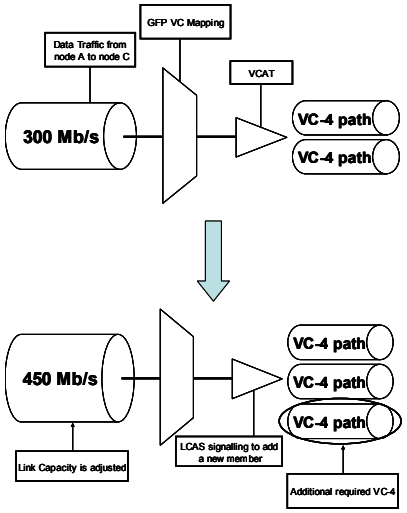


Figure 88: Example of the Link Capacity Adjustment Scheme (LCAS) functionality

As explained before, some of the IP/MPLS routers interfaces are tagged as high priority (HP) and the rest as low priority (LP). The client LSPs are classified as HP or LP according to the carried applications and their requirements in terms of QoS requirements. Then, LSPs are mapped to the virtual containers (VCs) according to their priority. Specifically, HP LSPs are mapped over HP VCs while LP LSPs are mapped on LP VCs (Figure 89). In order to do this mapping, the Generic Framing Procedure is used.

In such generalization of the network context in which the procedure can be applied, it has to be underlined that, in order to improve the bandwidth utilization, the high priority light paths can transport both high and low priority VCs, while low priority ones can only transport low priority VCs.

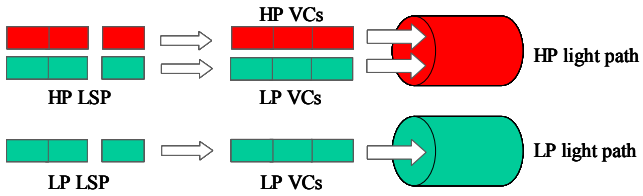


Figure 89: Transport hierarchy: Light paths/VCs/LSPs

Figure 90 schematically shows an exemplary network node of a network exploiting multiple server layers, i.e. a lower hierarchy switched circuit (SDH High Order Virtual Container) and a higher hierarchy switched circuit (OCh, or light path).

Data traffic coming from an edge node with tributary interfaces is inserted through the interfaces in the mapping/demapping subsystems depicted in the Figure 90. The incoming packets are then mapped into lower hierarchy circuits of suitable payload (e.g., HOVC at 150 Mbit/s). A first portion of the interfaces is allocated to high priority traffic, whereas a second portion thereof is allocated to low priority traffic. A Virtual Concatenation Selector (VCS) connects the mapping/demapping subsystems to the available HOVC termination points. Different HOVC may be virtually concatenated in a Virtual Container Group if the carried traffic should reach the same destination, even via differently routed light paths towards the same destination. The HOVC Fabric allows cross-connection of the HOVC towards an OCh Fabric (e.g. in an Optical Cross Connect), through Adaptation/Termination Points. In such Adaptation/Termination Points, the HOVC circuits are properly adapted/terminated (according to the SDH multiplexing scheme) for mapping into optical WDM higher hierarchy circuits (i.e. light paths) of suitable payload (e.g. 2.5 Gbit/s). The OCh Fabric separates in ordered manner the light paths carrying high priority Higher Order Virtual Containers and low priority Higher Order Virtual Containers, i.e. low priority light paths and high

priority light paths, according to the destination and priority policies. When the data traffic has been mapped as VCs, the system adds a pointer to form an Administrative Unit (AU). This process allows the necessary adaptation between the High Order Path layer and the Multiplex Section layer (MS). After the multiplexing process, the signals enter to the Regeneration Section (RS) for being regenerated. Finally, the system will compose a Synchronous Transport Module (STM-N) which cope with the light paths capacity. Once the STM-N signal is composed, this will enter in the OXC, which allows the cross-connection of optical channels (OCh).

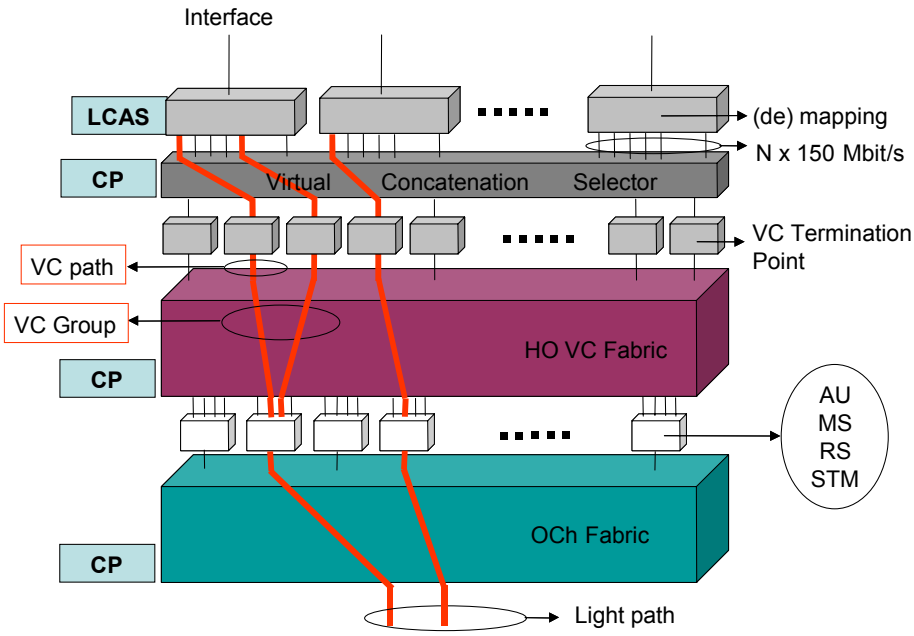


Figure 90: Proposed Architecture

In the exemplary network node shown in Figure 90, a monitoring is performed at the interfaces, in order to detect high priority traffic bursts, as explained above with reference to the IP/MPLS over WDM networks. The Control Plane component to be introduced to support the procedure (i.e. Traffic Controller component), performs the monitoring and the short-term prediction of the traffic crossing the interfaces. Then, it performs the calculation of the number of Virtual Containers and/or of the number of WDM circuits (light paths) required in order to cope with the burst of the high priority traffic, according to the following equation:

$$VC_{HP}^{x,y} = \frac{\hat{\chi}(n)}{\text{VC Bit Rate}}$$

where X and Y are the origin and destination of the HP VC and VC Bit Rate is the payload of the virtual container (e.g., 150 Mbit/s in case of VC-4). $\hat{\chi}(n)$ is the predicted traffic between nodes X and Y, as the result of applying the NMLS prediction algorithm by the TC component.

Based on the result of such calculation, the Control Plane can act at different levels, for example suitably driving a LCAS controller in order to modify, at the Virtual Concatenation Selector, the bandwidth of at least a portion of the Virtual Containers (See Figure 91). However, if the bandwidth adjustment is not sufficient to cope with the burst, the Control Plane CP may act in order to tear down low priority circuits, at the HOVC layer and/or at the OCh layer. Specifically, two different cases can be considered:

1. If the traffic prediction implies the need to increase the number of HP VCs and such increase cannot be supported by the current established light path towards the destination edge node, the procedure in this case triggers the request to the node CC for the set up of a new HP light path towards the same destination node. This may imply the tear down of a LP light path.

2. If the traffic prediction implies the request for the reduction of the member of the HP VCs, and such reduction implies that a light paths towards the destination node does not carry HP traffic, the procedure implies the request for the tear down of such light path, freeing network resource.

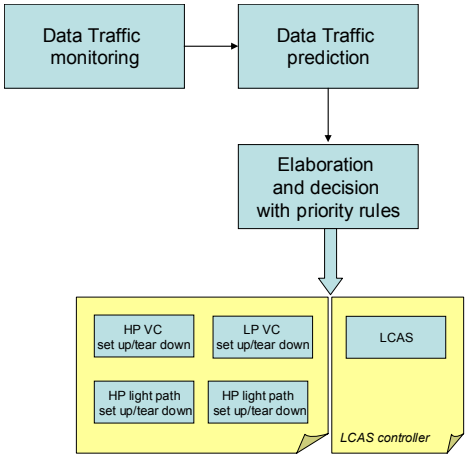


Figure 91: Functional Diagram of the generalized procedure

Interfaces corresponding to the torn down low priority circuits are also depleted from low priority traffic, as previously described. In such way, resources made available within the network by the tearing down of low priority circuits can then be used in order to set up new, temporary high priority circuits to carry the high priority excess traffic.

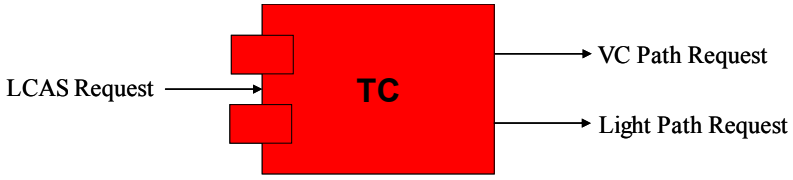


Figure 92: Traffic Controller Component Interfaces

In the case the traffic burst can be accommodated by increasing the bandwidth (number of members of the VCG by using the LCAS functionality), the TC component contacts with the LCAS controller for adding and deleting members to/from a virtual concatenation group. For doing that, TC must provide the virtual concatenation group identifier and the number of virtual containers to add.

Next, we describe the most important blocks required by the TRIDENT procedure to be applied to dynamic SONET/SDH networks. The first block consists of the monitoring and prediction of the incoming data traffic to the HP router interface *n* (B_n) and to be carried towards a certain destination edge node (e.g., node Y). As a result of this block, the number of the high priority virtual containers (VC^{HP}) is calculated. In case the number of the required VC^{HP} has to be decreased, the procedure implies, on one hand, the elimination of the VC as members of the VCG and, on the other hand, the signalling for the release of the path of the removed VC. To do this, the TC component interworks with the LCAS controller in order to signal a *remove* state to the member that is wanted to be deleted. When the sink node ACK with a *status fail*, the VC is removed and the TC interworks with CC which starts the signalling process for releasing the VC path. If, additionally, a HP light path has to be torn down, the CC starts the signalling for the HP light path tear down.

On the contrary, if the result of the monitoring and prediction indicates that the number of VCs has to be increased (in presence of a HP traffic bursts), the procedure tries to allocate the required VC^{HP} over the already established HP light paths towards the destination node. This step implies, if necessary, the deletion of low priority virtual containers (VC^{LP}) and consequent loss of the low priority traffic.

If the HP traffic cannot be allocated over the already established light paths, the procedure starts for the signalling for the set up of a new HP light path to be used temporarily to carry the HP traffic. Then, by using the LCAS signalling, the number of members of the VCG is increased and the different VC are individually routed by the Control Plane through the two different HP light paths.

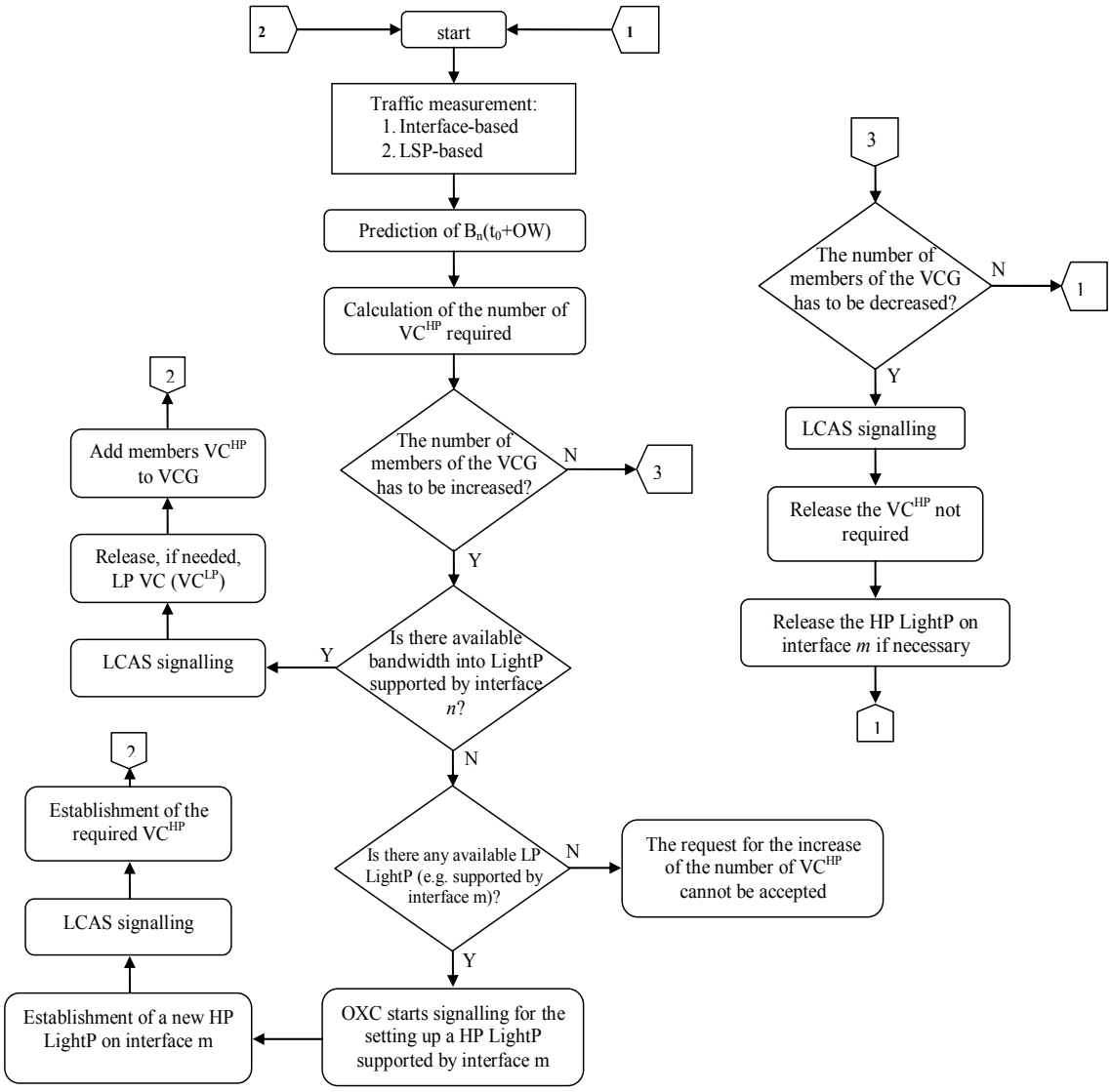


Figure 93: IP/MPLS over SDH/ASON networks, handling HP traffic bursts

It has to be underlined that in such generalization of the procedure, HP light paths can carry both low and high priority VCs, in order to increase the bandwidth utilization of the lightpath, reaching traffic engineering objectives.

12 Bibliography

[1] IST-1999-11387 LION Project, "Network Scenarios and Requirements", Deliverable 6, October 2000.

[2] S. V. Kartalopulos, "Introduction to DWDM Technology", Wiley-IEEE Press Publication, 2001

[3] R. Ramaswami, K. N. Sivarajan, "Optical Networks, A practical perpective", Morgan Kaufmann Publishers, 2nd Edition, 2002.

[4] P. Iovanna, R. Sabella, M. Settembre, "A Traffic Engineering System for Multilayer Networks Based on the GMPLS Paradigm", IEEE Network, Vol. 3, pp. 28-27, March 2003.

[5] F. Baker, "Requirements for IP version 4 of Routers", RFC 1812, June 1995

[6] T. W. Chung et al., "Architectural and Engineering Issues for building an Optical Internet", http://www.6pop.canet2.net/, September 1998.

[7] M. Listanti et al., "Architectural and technological issues for future optical internet networks" IEEE Communications Magazine, Vol. 38, No. 9, Sept. 2000.

[8] S. De Maesschalck et al., "Asymmetric IP Traffic and its consequences for the Optical Layer", in Proceeding of European Conference on Optical Communications (ECOC), Amsterdam, The Netherland, September 2001.

[9] D. Awduche, J. Malcom, J. Agogbua, M. O'Dell and J. McManus, "Requirements for Traffic Engineering Over MPLS", IETF Request for Comments 2702, September 1999.

[10] E. Rosen et al., "Multiprotocol Label Switching Architecture," IETF Request for Comments 3031, January 2001.

[11] IST-1999-11387 LION Project, "Recommendations for Network Operators", Deliverable 27, December 2002.

[12] IST- FP6-506760 NOBEL Project, "Preliminary definition of drivers and requirements for core and metro networks supporting end-to-end broadband services for all", Deliverable 6, September 2004.

[13] IST- FP6-506760 NOBEL Project, "Preliminary definition of network scenarios and solutions supporting broadband services for all", Deliverable 11, December 2004.

[14] D. Cavendish, K. Murakami, S. Yun, O. Matsuda, M. Nishihara, "New Transport Services for Next-Generation SONET/SDH Systems", IEEE Communications Magazine, Vol. 40, Issue 5, pp. 80-87, May 2002.

[15] ITU-T Rec. G.7042, "Link Capacity Adjustment Scheme (LCAS) for Virtual Concatenated Signals", August 2001.

[16] P. Green, "Progress in Optical Networking", IEEE Communications Magazine, Vol. 39, Issue 1, pp. 54-61, January 2001.

[17] ITU-T Rec. G.872, "Architecture of Optical Transport Networks", February 1999.

[18] ITU-T Rec. G.8080 "Architecture for the Automatic Switched Optical Network (ASON)", November 2001

[19] A. Banerjee, J. Drake, J. P. Lang, B. Turner, K. Kompella, Y. Rekhter, "Generalized Multiprotocol Label Switching: An Overview of Routing and Management Enhancements", IEEE Communications Magazine, Vol. 39 , nº 1, pp. 144-150, January 2001

[20] S. Dixit, "IP over WDM: Building the Next-Generation Optical Internet", John Wiley & Sons Inc., 2003.

[21] N. Ghani, S. Dixit, T. Wang, "On IP over WDM Integration", IEEE Communications Magazine, Vol. 38, nº 3, March 2000.

[22] Resilient Packet Ring Alliance, "An Introduction to Resilient Packet Ring Technology", http://www.rpralliance.org/articles/Whitepaper10.01.pdf.

[23] IEEE 802.17 RPR Work group, http://grouper.ieee.org/groups/802/17.

[24] RPR Alliance, IEEE 802.17 RPR Standard Approved, http://www.rpralliance.org

[25] IST-1999-11387 LION Project, "Failure Scenarios of Resilience in multi-layer networks", Deliverable 7, October 2000.

[26] IST-1999-11387 LION Project, "Resilience interworking strategies in multi-layer networks", Deliverable 16, October 2001.

[27] X. Xiao, A. Hannan, B. Bailey, "Traffic Engineering with MPLS in the Internet", IEEE Network, Vol. 14, nº 2, pp. 28-33, March 2000.

[28] J. Kuri, N. Puech, M. Gagnaire, E. Dotaro, R. Douville, "Routing and wavelength assignment of scheduled lightpath demands", IEEE Journal on Selected Areas of Communications, Vol. 21, nº 8, pp. 1231-1240, October 2003.

[29] http://www.fcr.es/cas/recerca/033.asp.

[30] G. Chiruvolu, "System and Method for Routing Stability-based Integrated Traffic Engineering for GMPLS Optical Networks", Alcatel USA, US Patent Application nº US 2003/0067880 A1, April 2003.

[31] D. Awduche et al., "Overview and Principles of internet Traffic Engineering", IETF Request for Comments 3272, May 2002.

[32] R. Braden, L. Zhang, S. Berson, S. Herzog, S. Jamin, "Resource Reservation Protocol, Functional Specification", IETF Request for Comments 2205, September 1995.

[33] L. Andersson, P. Doolan, N. Feldman, A. Fredette, B. Thomas, "Label Distribution Protocol specification", IETF Request for Comments 3036, January 2001.

[34] A. Manzalini, K. Shimano, C. Cavazzoni, A. D'Alessandro, "Architecture and functional requirements of control planes for automatic switched optical networks: experience of the IST project LION", IEEE Communications Magazine, Vol. 40, nº 11, pp. 60-65, November 2002.

[35] A. Manzalini, C. Cavazzoni, A. D'Alessandro, R. Morro, "Opportunities and Challenges of ASON/GMPLS Transport Networks", in Proceedings of 29th European Conference on Optical Communications (ECOC), Rimini, Italy, September 2003.

[36] The Internet Engineering task Force (IETF), www.ietf.org.

[37] Optical Internetworking Forum (OIF), www.oiforum.com.

[38] Architecture, OAM&P, PLL, & Signaling Working Groups, "User Network Interface (UNI) 1.0 Signaling Specification", Optical Internetworking Forum (OIF), December 2000.

[39] IST- FP6-506760 NOBEL Project Web page, www.ist-nobel.org.

[40] IST-1999-11387 LION Project, "Optimisation of Network Architectures", Deliverable 18, December 2001.

[41] J. Ash, "Traffic Engineering & QoS Methods for IP-, ATM-, & TDM-Based Multiservice Networks", <draft-ietf-tewg-qos-routing-01.txt>, April 2001.

[42] A. Gençata, B. Mukherjee, "Virtual-Topology Adaptation for WDM Mesh Networks under Dynamic Traffic", IEEE/ACM Transactions on Networking, Vol. 11, nº 2, April 2003.

[43] R. Ramaswami, K. N. Sivarajan, "Design of logical topologies for wavelength-routed optical networks", IEEE Journal on Selected Areas of Communications, Vol. 14, pp. 840-851, June 1996.

[44] B. Mukherjee, "WDM optical communication networks: Progress and challenges", IEEE Journal on Selected Areas of Communications, Vol. 18, pp. 1810-1824, October 2000.

[45] R. M. Krishnaswamy, K. N. Sivarajan, "Design of logical topologies: A linear formulation for wavelength-routed optical networks with no wavelength changers", IEEE /ACM Transactions on Networking, Vol. 9, pp.186-198, April 2001.

[46] R. Dutta, G. N. Rouskas, "A survey of virtual topology design algorithms for wavelength-routed optical networks", Optical Communications Magazine, Vol. 1, nº 1, pp. 73-89, January 2000.

[47] J. F. P. Labourdette, G. Hart, A, S, Acampora, "Branch-exchange sequences for reconfiguration of lightwave networks", IEEE Transactions on Communications, Vol. 42, pp. 2822.2832, October 1994.

[48] I. Baldine, G. N. Rouskas, “Traffic Adaptive WDM networks: A study of reconfiguration issues”, IEEE Journal on Lightwave Technology, Vol. 19, pp. 433-455, April 2001.

[49] B. Puype at al., “Multi-layer Traffic Engineering in Data-centric Optical Networks”, in Proceedings of the 7th IFIP Conference on Optical Network Design & Modeling”, pp. 211-226, February 3-5, 2003, Budapest, Hungary.

[50] A. Banerjee et al.,”Generalised MultiProtocol Label Switching: An overview of Signalling Enhancement and Recovery Techniques”, IEEE Communications Magazine, Vol. 39, no 7, pp. 144-151, July 2001.

[51] T. Anjali, C. Scoglio, I. K. Akyildiz, “LSP and λSP Setup in GMPLS Networks”, in Proceedings of IEEE Infocom 2004.

[52] S. Spadaro et al., “Network applications and Traffic modelling for ASONs”, In Proceedings of European Conference on Optical Communications (ECOC 2002), Copenhagen, Denmark, September 2002.

[53] National Bank of Poland, http://www.nbp.pl/statystyka/index.html.

[54] Bank BPH, http://www.bph.pl.

[55] Polish Official Statistics, http://www.stat.gov.pl/english/index.htm.

[56] The Motion Picture Association, http://www.mpaa.org.

[57] http://ca.movies.yahoo.com/ap/20020402/101779340400.html.

[58] http://www.quvis.com/products/qubitST.htm.

[59] S Uhlig, O. Bonaventure, “On the Cost of Using MPLS for Interdomain Traffic”, in Proceedings of Quality of Future Internet Service’00, Berlin, Germany, September 2002.

[60] J. Filipiak, ”Real Time Network Management”, Elsevier Science Publishers, 1991.

[61] W. E. Leland, et al, "On the self-similar nature of Ethernet traffic”, IEEE/ACM Transactions on Networking, vol.2, n.1, February 1994.

[62] K. Thompson, G. J. Miller, R. Wilder, “Wide-Area Internet Traffic Patterns and Characteristics”, IEEE Network, vol. 11, n. 6, pp. 10-23, November-December 1997.

[63] M. S. Taqqu, W. Willinger, and R. Sherman, “Proof of a fundamental result in self-similar traffic modelling”, ACM Computer Communications Review, vol. 27, no. 2, pp. 5-23, April 1997.

[64] X. Xiao et al, "Internet QoS: A Big Picture", IEEE Network Magazine, March/April 1999, pp. 8-18.

[65] Cisco Systems Products, htttp://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122newft/122t/122t4/fteibmpl.pdf.

[66] A. Manzalini, A. D'Alessandro, S. Spadaro, J. Solé-Pareta, O. Pisa, "System and Method for the Automatic setup of switched circuits based on traffic prediction in a Telecommunications Network", Patent Application nº PCT/EP03/14800, submitted to European Patent Office, December 2003.

[67] Telecom Italia Lab, www.tilab.com.

[68] T. Anjali, C. Scoglio, J. de Oliveira, L. C. Chen, I. F. Akyildiz, J. A. Smith, G. Uhl, A. Sciuto, "A New Path Selection Algorithm for MPLS Networks Based on Available Bandwidth Measurement", in Proceedings of QofIS 2002, Zurich, Switzerland.

[69] T. D. Nadeau, "Multiprotocol Label Switching (MPLS) Label Switch Router (LSR) Management Information Base", IETF draft-ietf-mpls-lsr-mib-11.txt, January 2002.

[70] D. Kim at al., "A Requirement of the Network State Information Database for Traffic Engineering Over GMPLS", IETF draft-kim-ccamp-gmpls-nsid-01.txt, November 2002.

[71] W. S. Lai, R. W. Tibbs, S. Van den Berghe, "A Framework for Internet Traffic Engineering Measurement", IETF draft-ietf-tewg-measure-05.txt, February 2003.

[72] D. R. Mauro, K. J. Schmidt, "Essential SNMP", O'Reilly & Associates, July 2001.

[73] K. McCloghrie, M. Rose, "Management Information Base for Network Management of TCP/IP-based internets: MIB II", IETF Request for Comments 1213, March 1991.

[74] T. D. Nadeau, "Agent capabilities for MPLS-LSR-MIB", Cisco System proprietary MIB, September 2001.

[75] F. Mikus, "MIB.jnxMibs.mpls", Juniper Network proprietary MIB, January 2004.

[76] W. S. Lai, R. W. Tibbs, S. Van den Berghe, "A Framework for Internet Traffic Engineering Measurement", draft-ietf-tewg-measure-06.txt, July 2003.

[77] NLANR Measurement and Network Analysis, "Prototyping lambdaMONs", http://pma.nlanr.net/lambdamon.html.

[78] A. Adas, "Using Adaptive Linear Prediction to support real-time VBR video", IEEE/ACM Transactions on Networking, Vol. 6, nº 5, October 1998.

[79] S. Haykin, "Adaptive Filter theory", Prentice Hall, 1991.

[80] ITU-T Rec. G.807 "Architecture for the Automatic Switched Network", May 2001.

[81] A. Girard, "Routing and Dimensioning in Circuit-Switched Networks", Addison-Wensley, 1990.

[82] E. A. Van Doorn, "Some aspects of the peakedness concept in teletraffic theory", Elektronische Informationsverarbeitung und Kybernetik, vol. 22, no. 2-3, pp. 93-104, 1986.

[83] L. E. N. Delbrouck, "A unified approximate evaluation of congestion functions for smooth and peaky traffics", IEEE Trans. Commun., vol. 29, no. 2, pp. 85-91, 1981.

[84] R. I. Wilkinson, "Theories for toll traffic engineering in the USA", Bell Systems Technical.

[85] S. Subramaniam, A. K. Somani, M. Azizoglu, and R. A. Barry. A performance model for wavelength conversion with non-Poisson traffic. In Proceedings of IEEE Infocom 1997, Kobe, Japan, April 1997.

[86] J. R. Boucher. Traffic System Design Handbook. Telecommunications Handbook Series, IEEE Press, 1993.

[87] V. B. Iversen, "Teletraffic Engineering Handbook", ITU-D SG 2/16&ITC, Draft 2001-06-20, www.tele.dtu.dk/teletraffic.

[88] P. Demeester et al. "Resilience in Multilayer Networks", IEEE Communications Magazine, Vol. 37, nº 8, pp. 70-76, August 1999.

[89] ACTS-PANEL Project, "Overall Network Protection", Deliverable D2, 1997.

[90] M. Gryseels, R. Clemente and P. Demeester, "Protection strategies for SDH-over-WDM networks", in Proceedings of Networks and Optical Communications (NOC), 1998.

[91] T. H. Wu, "Emerging Technologies for Fiber Network Survivability", IEEE Communications Magazine, Vol. 33, nº 2, pp. 58-74, 1995.

[92] ITU-T Rec. G.841, "Types and characteristics of SDH network protection architectures", October 1998.

[93] ITU-T Rec. G.774.03, "Synchronous Digital Hierarchy (SDH) management of multiplex-section protection for the network element view", November 1996.

[94] D. Colle et al., "Data-Centric Optical Networks and Their Survivability", IEEE Journal on Selected Areas in Communications, Vol. 20, nº 1, pp. 6-20, January 2002.

[95] J. Manchester, "Fault Detection and Propagation in Transport Networks", in Proceedings of 1st International Design and Reliable Communications Networks (DRCN), Brugge, Belgium, 1998.

[96] A. Autenrieth et al., "Simulation and Evaluation of Multi-layer Broadband Networks", in Proceedings of 1st International Design and Reliable Communications Networks (DRCN), Brugge, Belgium, 1998.

[97] IEEE 802.17 Working Group, "Resilient Packet Ring access method & physical layer specifications", IEEE 802.17 RPR Aprroved Standard, June 2004.

[98] Resilient Packet Ring Alliance, www.rpralliance.org.

[99] ITU-T, Draft Rec. X.msr, "Multiple Services Ring", March 2002 (Temporary Document 2053).

[100] D. Tsiang and G. Suwala, “The Cisco SRP MAC Layer Protocol”, IETF Request for Comments 2892, August 2000.

[101] Nortel Network Product, Optical Packet Edge System, formerly known as: OPTera Packet Edge System / Resilient Packet Ring (RPR) Technology.

[102] S. Spadaro, J. Solé-Pareta, D. Careglio, C. Wajda, A. Symansky, “Positioning of RPR standard in contemporary operators’ environment”, IEEE Network magazine, vol. 18, no. 2, Mar/Apr. 2004.

[103] OPNET simulation tool, www.opnet.com.

[104] IST-1999-11387 LION Project, “Multilayer resilient network planning and evaluation: preliminary results”, Deliverable 10, January 2001

[105] J. Moyano, S. Spadaro, B. Bostica, J. Solé-Pareta, "Performance Evaluation of the Spatial Reuse Protocol fairness algorithm (SRP-fa) used in DPT Networks", Proc. of IEEE International Conference on Telecommunications (ICT), Bucharest, Romania, June 2001.

[106] S. Spadaro, J. Solé-Pareta, D. Careglio, K. Wajda, A. Symansky, "Assessment of Resilience Features for the DPT Rings", Proc. Eurescom Summit, Heidelberg, Germany, October 2002.

[107] ITU-T Recommendation G.841, “Types and characteristics of SDH network protection architectures”, 1998.

[108] J. Lang, “Link Management Protocol (LMP)”, IETF draft-ietf-ccamp-lmp-10.txt, October 2003.

[109] C. Cavazzoni et al., “The IP/MPLS over ASON/GMPLS test bed of the IST project LION”, IEEE Journal on Lightwave Technology, Vol. 21, nº 11, pp. 2791-2803, November 2003.

[110] S. Spadaro, A. D’Alessandro, A. Manzalini, J. Solé-Pareta, “A Procedure for the Automatic Set-up and Tear-down of Switched Connections Tracking Traffic Fluctuations in IP/MPLS over ASON/GMPLS Networks”, in Proceedings of 30th European Conference on Optical Communications (ECOC 2004), Stockholm, Sweden, September 2004.

[111] ITU-T Rec. G.709, “Interfaces for the Optical Transport Network (OTN)”, March 2003.

[112] ITU-T Rec. G.707, “Network node interface for the synchronous digital hierarchy (SDH)”, October 2000.

[113] ITU-T Rec. G.7041, “Generic Framing Procedure”, December 2001.

[114] E. Hernandez-Valencia, M. Scholten, Z. Zhu, “The Generic Framing Procedure (GFP): An Overview”, IEEE Communications Magazine, Vol. 40 Issue 5, May 2002.

[115] Agilent Technologies, www.agilent.com

Appendix: List of Publications

1. **S. Spadaro**, A. D'Alessandro, A. Manzalini, J. Solé-Pareta, "A procedure for the automatic set up and tear down of switched connections tracking traffic fluctuations in IP/MPLS over ASON/GMPLS networks", submitted to the Elsevier Computer Networks Journal.

2. **S. Spadaro**, A. D'Alessandro, A. Manzalini, J. Solé-Pareta, "A procedure for automatically triggering the set up/tear down of switched connections in IP(MPLS) over ASON network", Workshop on Recent Advances in Computer Networking, April 10, 2004, Atlanta, USA.

3. A. Manzalini, A. D'Alessandro, **S. Spadaro**, J. Solé-Pareta, O. Pisa, "System and Method for the Automatic setup of switched circuits based on traffic prediction in a Telecommunications Network", Patent Application nº PCT/EP03/14800, submitted to European Patent Office, December 23rd 2003, in collaboration with the network operator Telecom Italia Labs (TILAB).

4. **S. Spadaro**, A. D'Alessandro, A. Manzalini, J. Solé-Pareta, "A Procedure for the Automatic Set up and Tear down of Switched Connections Tracking Traffic Fluctuations in IP/MPLS over ASON/GMPLS Networks", in Proceedings of 30th European Conference on Optical Communications (ECOC 2004), Stockholm, Sweden.

5. **S. Spadaro**, J. Solé-Pareta, D. Careglio, C. Wajda, A. Symansky, "Positioning of RPR standard in contemporary operators' environment", IEEE Network magazine, vol. 18, no. 2, Mar/Apr. 2004.

6. **S. Spadaro**, M. Quagliotti, J. Solé-Pareta, D. Careglio, A. Manzalini, F. Saluta, R. Stanckiewicsz, J. Rzasa, A. Lason, "Teletraffic engineering methods for Intelligent Optical Networks", in Proceedings of 8th IFIP Optical Networks Dimensioning and Modelling, February 2-4, 2004, Ghent, Belgium.

7. J. Solé-Pareta, **S. Spadaro**, "IP over WDM: The IST LION Project approach", chapter of the book "Deploying and Managing IP over WDM", Artech House Publishers, June 2003.

8. S. De Maesschalck, D. Colle, P. Demeester, I. Lievens A. Manzalini, M. Quagliotti, F. Saluta, **S. Spadaro,** J. Prat, J. Comellas, , J. Sole Pareta, M. Jaeger, I. Shake, G. Kylafas, L. Raptis, J. Soldatos, R. Leone, J. Derkacz, A. Lason, J. Rzasa, A. Matzke, "Advantages of Intelligent Optical Networks", submitted to IEEE Optical Communications, Quartely Supplement to IEEE Communications Magazine.

9. J. Solé-Pareta, X. Masip, S. Sanchez, **S. Spadaro**, D. Careglio, "Some Open Issues in the Optical Networks Control Plane", in Proceedings of 5th IEEE International Conference on Transparent Optical Networks (ICTON), Varsow, Poland, July 2003.

10. **S. Spadaro**, J. Solé-Pareta, A. Lason, J. Rzasa, R. Stankiewicz, A. Manzalini, A. D'Alessandro, D. Colle, S. De Maesschalck, I. Lievens, I. Shake, K. Shimano, "Network Applications and Traffic Modelling for ASONs", in Proceedings of 28th European Conference on Optical Communications (ECOC), Copenhagen, Denmark, September 2002.

11. A. Manzalini, M. Quagliotti, A. Lason, J. Rzasa, R. Stankiewicz, J. Solé-Pareta, **S. Spadaro**, "Teletraffic Engineering for modelling and Dimensioning ASONs", in Proceedings of Workshop on High-Capacity Optical Networks, Turin, Italy, October 2002.

12. **S. Spadaro**, J. Solé-Pareta, D. Careglio, C. Wajda, A. Symanszky, "Assessment for Resilience in the DPT Rings", in Proceedings of Eurescom Summit Powerful Networks for Profitable Services, Heidelberg, Germany, October 2002.

13. **S. Spadaro**, J. Moyano, B. Bostica, J. Solé-Pareta, "Perfomance Evaluation of SRP-fa", IP over DWDM Conference, November 25-30, Paris, France.

14. J. Moyano, **S. Spadaro**, B. Bostica, J. Solé-Pareta, "Performance Evaluation of SRP-fa algorithm used in DPT Networks", in Proceedings of IEEE International Conference on Telecommunications (ICT), Bucharest, Romania, June 2001.

Projects Deliverables

1. IST-1999-11387 LION project "Multilayer resilient network planning and evaluation: preliminary results", Deliverable D10, January 2001.

2. IST-1999-11387 LION, "Resilience Interworking strategies for Multi-layer Networks", Deliverable D16, October 2001.

3. IST-1999-11387 LION, "Multilayer resilient network planning and evaluation: intermediate results", Deliverable D19, December 2001.

4. IST-1999-11387 LION, "Optimisation of Network Architectures", Deliverable D18, December 2001.

5. IST-1999-11387 LION, "Multilayer resilient network planning and evaluation: final results", Deliverable D24, November 2002.

6. IST-1999-11387 LION, "Preliminary Recommendations for Network Evolution", Deliverable D27, December 2002.

Other Publications

1. D. Careglio, J. Solé-Pareta, **S. Spadaro**, "Novel contention resolution technique for QoS support in connection-oriented optical packet switching", accepted for Publications in IEEE International Conference on Communications (ICC), Seoul, Korea, May 2005.

2. Bianco, D. Careglio, J. Finochietto, G. Galante, E. Leonardi, F. Neri, J. Solé-Pareta, **S. Spadaro**, "Multi-class resource allocation for interconnected WDM rings in the DAVID metro network", IEEE Journal on Selected Areas of Communications, vol. 22, nº 8, October 2004.

3. M. Klinkowski, D. Careglio, X. Masip-Bruin, **S. Spadaro**, S. Sanchez-López, J. Solé-Pareta, "A simulation study of combined routing and contention resolution algorithms in connection-oriented OPS network scenario", in Proceedings of 6th IEEE International Conference on Transparent Optical Networks (ICTON2004), Wroclaw, Poland, Jul. 2004.

4. **S. Spadaro**, J. Comellas, E. Torrecilla, G. Junyent, Josep Solé-Pareta, "Multicast-like approach for optical networks resources optimisation", in Proceedings of 9th European Conference on Networks and Optical Communications (NOC 2004), Eindhoven, The Netherlands, June 2004.

5. E. Escalona, **S. Spadaro**, X. Masip, G. Junyent, J. Solé-Pareta, J. Domingo, S. Sanchez, "Evolution of current transport networks to ASON/GMPLS", III Workshop on MPLS networks, Girona, Spain, March 25-26, 2004.

6. D. Careglio, J. Solé-Pareta, **S. Spadaro**, "Heuristics for QoS in DAVID network", in Proceedings of 29th European Conference on Optical Communications (ECOC), Rimini, Italy, September 2003.

7. D. Careglio, A. Rafel, J. Solé-Pareta, **S. Spadaro**, G. Junyent, "Quality of Service Strategy in an Optical Packet Network with a Multi-class Frame-based Scheduling", in Proceedings of Workshop on High Performance Switching and Routing (HPSR 2003), June 2003, Turin, Italy.

8. D. Careglio, J. Solé-Pareta, **S. Spadaro**, "Optical Slot Size in IP over OPS Networks", in Proceedings of 7th International Conference on Telecommunications (CONTEL 2003), June 11–13 2003, Zagreb, Croatia

9. D. Careglio, J. Solé-Pareta, **S. Spadaro**, G. Junyent, "Performance Evaluation of Interconnected WDM PONs Metro Networks with QoS Provisioning", in Proceedings of 7th IFIP Optical Networks Dimensioning and Modelling (ONDM), February 2003, Budapest, Hungry.

10. D. Careglio, G. Giner, J. Solé-Pareta, **S. Spadaro**, G. Junyent, "Evaluación de la red óptica metropolitana multi-anillo del proyecto DAVID", in Proceedings of 12th Telecom I+D, November 2002, Madrid, Spain.

11. D. Careglio, J. Solé Pareta, **S. Spadaro**, "Performance Evaluation of Metro Optical Networks based on Multiple WDM PONs Interconnected through a PWRN", in Proceedings of 2nd International Workshop on All-Optical Networks (WAON'2001), Zagreb, Croatia.

12. J. Solé Pareta, D. Careglio, **S. Spadaro**, J. Masip, J. Noguera, G. Junyent, "Modelling and Performance Evaluation of a National Scale Switchless Based Network", Fifht International Symposium on Interworking '2000, October 3-7, Bergen, Norway.