



The INIM system: in-service non-intrusive monitoring for QoS enabled transparent WDM

C. Pinart, G. Junyent

Publication:	IEEE Journal of Selected Topics in Quantum Electronics
Vol.:	12
No.:	4
pp.:	635-644
Date:	July/August 2006

This publication has been included here just to facilitate downloads to those people asking for personal use copies. This material may be published at copyrighted journals or conference proceedings, so personal use of the download is required. In particular, publications from IEEE have to be downloaded according to the following IEEE note:

©2007 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

The INIM system: in-service non-intrusive monitoring for QoS enabled transparent WDM

Carolina Pinart and Gabriel Junyent

Abstract—This paper presents the design and experimental implementation of an in-service performance monitoring system that uses a combination of non-intrusive spectral OPM and IP metrics and is based on a low-complexity, distributed architecture. Apart from guaranteeing QoS in real-time (less than 1 sec) in an IP/WDM dynamic network, the system provides link-state information for impairment-aware RWA. Another novel aspect is the nature of the SLAs proposed, which are ‘all-optical’ (transparent). Performance delays of the system are evaluated in a real testbed featuring DWDM and transparent OADMs. Strategies to estimate link-state parameters from real-time monitoring information are also proposed.

Index Terms—Performance monitoring, optical management, IP/WDM, physical impairments.

I. INTRODUCTION

OPTICAL networks evolved from telephony systems designed to transmit digitized voice circuits across a fiber link. With the growth of bursty, packet-based Internet Protocol (IP) data traffic, these network architectures are being used to deliver Internet content. However, the core technologies of these networks were originally designed for voice and high-priority data traffic, which makes them difficult to adapt to the nature of IP-based traffic, the most dominant factor in data transport. Equipment for Wavelength Division Multiplexing (WDM), reconfigurable Optical Cross-Connects (OXC) and Optical Add-Drop Multiplexers (OADM), along with emerging approaches of optical intelligence (especially the Generalized Multi-Protocol Label Switching, GMPLS [1]), have matured sufficiently to build the very high-capacity networks that will be needed to transport the ever-increasing amount of information, known as IP/WDM. The combination of these elements and the removal of opto-electronic conversions in the core nodes will result in the efficient transportation of any type of data traffic, regardless of its payload or format. This gives future IP/WDM the chance to provide new on-demand network services in a transparent way and with different quality levels (QoS), but results in a major challenge for performance monitoring, principally due to the lack of electrical regeneration in the core of transparent IP/WDM, that is, in the accumulation of impairments along the lightpaths, but also because Routing and Wavelength Assignment (RWA) mechanisms are being enhanced to consider information about the status of physical resources with a view to integrate QoS in

the provisioning process. This is known as impairment-aware RWA (IRWA).

This paper presents an experimental, on-line monitoring system, named In-service Non-Intrusive Monitoring (INIM), which combines distributed elements and non-intrusive monitoring techniques to guarantee service level agreements (SLA) that can be verified with optical and IP parameters. The system can also provide information on a per link basis to perform IRWA. The INIM system is implemented in a laboratory testbed featuring dense WDM and GMPLS enabled transparent optical nodes, which provides on-demand lambda services for IP traffic. The remainder of the paper is organized as follows. Section II outlines the principal monitoring techniques for transparent WDM. Section III proposes service-intrinsic performance parameters for an ‘all-optical’ SLA. In Section IV we describe the modelling and architectural elements of the INIM system. Section V includes a performance evaluation of INIM focusing on delays and normal operation to provide link-state information. Finally, in Section VI we draw conclusions.

II. MONITORING TECHNIQUES FOR TRANSPARENT WDM

Optical performance monitoring (OPM) is an indispensable element for the quality assurance of an optical network. OPM aims at measuring the undesirable effects that an optical signal may suffer, such as frequency misalignment in the optical components, gain irregularities, in-band and out-of-band crosstalk, dispersion or fiber non-linearities. OPM may include monitoring of different analog or digital parameters of the transmission system, including the wavelength drift, optical power, signal integrity, optical signal to noise ratio (OSNR), bit error rate (BER), Q factor and dispersion.

A transparent network is an optical network in which optical signals traverse the network elements and links without opto-electronic conversions but at the ingress and egress of the network. Contrarily, in opaque networks optical signals are regenerated electrically, and hence monitoring functions that use the overhead information carried in optical signals can be located at the regenerators. This cannot be done in transparent networks without losing transparency. Moreover, in such networks the monitoring system may not have prior information about the protocol, format, or data rate of the optical signals. Therefore, non-intrusive OPM schemes that do not decode the overhead information are to be employed. If we add to this the fact that in future optical networks signals will be routed, added/dropped, (de)multiplexed or regenerated independently of the higher electrical layers and in a reconfigurable way, the monitoring schemes needed must be insensitive to the signal origins and its transverse path history.

Manuscript received August 10, 2005 and reviewed February 28, 2006.

C. Pinart is with the Centre Tecnològic de Telecomunicacions de Catalunya, 08860 Castelldefels, Spain (phone: +34-93-645-29-22 fax: +34-93-645-29-01; e-mail: carolina.pinart@cttc.es). G. Junyent is with the Universitat Politècnica de Catalunya, 08034 Barcelona, Spain (e-mail: junyent@tsc.upc.edu).

Performance monitoring techniques can be intrusive or non-intrusive. Examples of intrusive techniques are BER calculations and optical time-domain reflectometer measurements. Examples of non-intrusive techniques are the spectral analysis (power, frequency drift, OSNR), Q factor estimation and pilot-tone methods. Ideally, desirable monitoring techniques in a transparent network should be non-intrusive, fast, accurate, with large dynamic range, simple, compact, low-cost, scalable and comprehensive. The rationale behind this is fourfold: avoidance of opto-electronic conversions, low capital and operational expenses, rapid location of faults and use in systems with high number, dense-spaced WDM channels.

At the following we review the most common non-intrusive OPM techniques. In spectral analysis techniques, OSNR can be measured in- or out-of-band. In-band OSNR monitoring techniques can be classified as noise spectrum analysis [2], polarization-assisted analysis [3], subcarrier multiplexing [4], and based on Mach-Zehnder interferometer [5]. The traditional out-of-band OSNR monitoring technique involves measuring and interpolating the noise power from adjacent channels. Some of these concepts can be applied to signal level (power) and wavelength drift measurements. The Q factor (Q_{BER}), which is calculated by measuring the minimum BER at the optimum decision level, is utilized for evaluating the performance of optically amplified systems. Since it takes an extremely long time to obtain the BER measurements (e.g. for a channel at 10 Gbps and BER of 10^{-12} , recording 10 errors will take 10^4 seconds), Q_{BER} is frequently estimated in much less time (a few seconds) [6]. Finally, the pilot tone technique consists in superimposing a tone on the optical signal with a transmission capacity determined by the carrier to noise ratio [7].

If we want to avoid opto-electronic conversions to monitor the status of active optical signals, non-intrusive monitoring is preferable in transparent networks. However, non-intrusive techniques provide a limited subset of performance parameters, which leads to the need for estimating ‘electrical’ parameters such as BER, and to use other complementary measurements, for example at the IP layer (packet metrics).

III. ‘ALL-OPTICAL’ SLA FOR LAMBDA SERVICES

A lambda service is an optical channel dedicated to carry a given traffic in a lightpath from a source to a destination. This work considers a transparent wavelength-routed network that offers WDM lambda services on demand (connections established by a GMPLS control plane). If we wish to guarantee lambda-service SLAs real-time monitoring in the optical network, we should define service classes with limited complexity. In this context, this work proposes three classes, named *VoIP-like*, *IPTV-like* and *Internet+*, which are inspired in the requirements of Triple Play services: voice over IP (VoIP), IP television (IPTV) and Internet data.

Table I lists the service-intrinsic parameters proposed. VoIP-like is built around VoIP characteristics, but it takes into account as well some intrinsic features of emergency-related (very low delay for connection setup, real-time transmission, almost error-free, stringent availability) and conversational (high requirements on latency and data loss, low transmission

TABLE I
SERVICE-INTRINSIC PARAMETERS OF AN ‘ALL-OPTICAL’ SLA

SLA parameter (service-intrinsic)	Lambda service class		
	VoIP-like	IPTV-like	Internet+
Setup delay	< 1sec	< 10sec	< 1.5min
Availability (BP)	10^{-3}	10^{-2}	10^{-1}
Throughput	Up to maximum laser bit rate		
Packet delay/ α	< 50msec	< 500msec	< 5sec
Packet loss	1%	0.1%	10%
OSNR	$OSNR_{targetBER} + \Delta OSNR_{BERest}$		

error rate, small delay for connection setup, highly available) services. IPTV-like is built around IP television requirements, and considers as well streaming features, which have lower demands on latency and errors because the end systems can compensate irregularities, for example by buffering data, and less stringent requirements about availability and setup delay (uncritical). Finally, Internet+ combines prioritized non-realtime traffic transmission with typical best effort. These classes can be verified in real-time by the INIM system through non-intrusive monitoring in the optical layer, and probe packets in the IP layer, as described in Section IV. The parameters considered are briefly discussed here:

Setup delay and blocking probability (as a means to measure availability in the establishment process) depend basically on the behavior of the optical control plane, because this work considers dynamic circuit-switched connections. The setup delay values listed in Table I assume that the functions of a control plane are real-time. IPTV-like and Internet+ classes have setup delays an order of magnitude higher than the previous class, respectively. VoIP-like adopts an order of magnitude more than the typical blocking probability for telephony networks, which is 1%, to cope with emergency-like traffic. The remaining classes have blocking probabilities of one order of magnitude less consecutively. Although best effort has usually no guaranteed availability, Internet+ considers 10% blocking to deal with prioritized non-real-time traffic. As for *throughput*, the service offered by the optical network framework considered in this work is connection-oriented analogue optical transmission, that is, a WDM channel between two endpoints, and therefore the bandwidth allocated can range from the minimum to the maximum bit rates of the laser source. Therefore, the maximum throughput is this maximum bit rate, which may be adjusted or not on a per class basis, and also depending on the user. It can also be adjusted to meet certain packet delay requirements.

Since the classes listed in Table I are built around the requirements of VoIP, IPTV (including high-definition) and Internet data, the metrics of *packet loss and delay* are basically those derived from these traffic classes. Before discussing the values given, two comments must be done. The first one is related to packet delay; the latency added by the optical network, being circuit-switched and transparent, is only transmission delay, plus processing time at the edges in order to perform opto-electronic conversion. Roughly, we can model this latency (in sec) as $D \cdot n/c$, where D (in km) is the length of the lightpath, n is the index of the fiber and c is the speed of light in vacuum. Moreover, since packet delay

metrics are end-to-end, we assume that the values in the optical network are a factor α of the end-to-end metrics, where $\alpha \leq 1$, which will be determined in the SLA negotiation phase. The second comment is about packet loss; since the optical network does not process the packets sent over it, we can model lost packets from bit errors, which may be caused by transmission impairments and/or faults in the network. Assuming that a packet has N bits, of which Z are errored, the expected number of errors per packet is $E[Z] = N \cdot BER$. Depending on the location of these Z errors, a packet is considered as errored (errors in the payload) or lost (errors in the header). Therefore, the packet loss rate (PLR) can be expressed as:

$$PLR = 1 - P(IP \text{ header correct}) = 1 - p_z(0)|_{N=IP_h} \quad (1)$$

where p_z is the probability density function for Z and IP_h is the length of the IP header in bits, which in IPv4 is 192. This expression is useful because a packet loss ratio can be derived from a target BER, and inversely. For example, by using Markov or a Neuman-A distribution to model the time evolution of correct and erroneous bits [8], p_z can be computed and hence BER can be estimated from the PLR.

The values of packet delay and loss in Table I are derived from the following requirements: in VoIP, round trip latencies above 300 msec result in users experiencing annoying talk-over effects, and emergency-like services have much more stringent requirements, for what we allocate a maximum (one-way) packet delay of 50 msec for the VoIP-like class, and approximately two orders of magnitude more for each of the remaining classes, which complies with the requirements of streaming, IPTV and Internet data [9]. Note that IPTV includes buffering, and therefore a 500-msec setup delay will be perceived as less by final users. Without FEC, VoIP targets a 1% PLR, which can be raised to 5% with correction methods. If IPTV is not very sensitive to packet delay, because its buffer can be up to 10 sec, it is to packet loss, because in multicast, large-extent retransmission is not possible. Therefore, IPTV targets a PLR of less than 0.5%, and our IPTV-like class targets 0.1%. Finally, the Internet+ class is very elastic with respect to packet losses, mainly due to retransmission, which must be bounded so that the maximum tolerated latency is not exceeded. Therefore, we target 10%.

Last but not least, the proposed ‘all-optical’ SLA includes OSNR as an estimation of BER, since there is no opto-electronic conversion in the optical network but at the edges, which complicates BER calculations, and we envisage real-time monitoring. OSNR measurement as a means to estimate BER is based on the assumption that the Q factor can be used as intermediate parameter. Marcuse [10] and Humblet and Azizoğlu [11] derived widely-used approximate expressions for the Q factor as a function of the SNR of the electric current and as a function of the OSNR, respectively. In both studies, the authors assume that the receiver consists of a rectangular optical filter, a square-law photodetector, and an integrate-and-dump electrical filter. They also assume that the optical signals have a perfect extinction ratio and that the optical noise is Gaussian and white prior to the optical filter. Finally, they assume that the signal is polarized and that the optical

noise is either unpolarized or is completely polarized and is co-polarized with the signal. While the the Q factor can be directly converted to an electrical SNR value, the relationship to the OSNR is unfortunately not so simple. It is defined only for systems with optical amplifiers, and as a ratio of the optical signal power to the ASE of the amplifiers. Combining the results of Humblet and Azizoğlu [11] and Becker et al. [12], the relation between the Q factor and the OSNR can be approximated as follows:

$$Q = \sqrt{\frac{B_o}{B_e} \frac{2OSNR}{\sqrt{4OSNR + 1} + 1}} \quad (2)$$

where B_o is the optical bandwidth and B_e is the electrical bandwidth. For example, DWDM systems may consider B_o values of 70 GHz, and typical values of electrical bandwidth range from 9 to 25 GHz. The Q factor can be estimated using the decision-circuit method introduced by Bergano et al. in [6] (Q_{EYE} , estimated from the eye diagram). By considering the case $Q_{BER} \simeq Q_{EYE}$ (e.g. intensity modulation direct detection systems with low inter-symbol interference), and combining the well-known BER expression with equation 2 we obtain:

$$BER \approx \frac{1}{2} \text{erfc}\left(\sqrt{\frac{B_o}{2B_e} \frac{2OSNR}{\sqrt{4OSNR + 1} + 1}}\right) \quad (3)$$

For example, if we target a BER of 10^{-8} , by using equation 3 we obtain an OSNR value of about 6.6 dB for a system with $B_o/B_e = 7$, and of 8.8 dB for a system with $B_o/B_e = 5$. To this OSNR value ($OSNR_{targetBER}$ in Table I), we add a $\Delta OSNR_{BERest}$ which is obtained empirically from the difference between real and estimated BER, and is meant as a kind of ‘offset’. Real BER is measured in the deployment phase, and can be measured while in service to keep $\Delta OSNR_{BERest}$ updated. Assuming $B_o/B_e = 7$, the VoIP-like class has $OSNR_{targetBER}$ of 8.2 dB (BER of 10^{-10} for conversational services [9]), IPTV-like between 5 and 6.6 dB (BER in the range $10^{-6} - 10^{-8}$ for streaming and prioritized elastic traffic) and Internet+ 5 dB, for a target BER of 10^{-6} .

A question concerning SLAs that arises in IP/WDM is how to provide QoS using a single transport technology. In this work we consider QoS in two ways: setup process and traffic constraints. In the former, the network offers bounded setup delays and blocking probabilities depending on the traffic type (class). In the latter, the lightpaths are chosen by IRWA depending on the traffic constraints of the class, basically target BER and maximum latency. However, other types of QoS can be envisioned: protection schemes, recovery times, etc.

IV. THE INIM SYSTEM

This section addresses the major aspects of design and implementation of the INIM system. This on-line monitoring system presents the following novelties: use of performance information obtained exclusively from non-intrusive techniques combination of physical- and IP-layer performance parameters and use of performance information from active channels to infer link-state information for real-time IRWA. The INIM

system is based on distributed elements, and uses the IP control channel of the Data Communication network (DCN) to transfer performance monitoring information.

A. Requirements and system modelling

The first step for designing the INIM system is to analyze its requirements. The input parameters of the system are the *network topology*, *nodes*, *links* and *users*, the *SLA parameters* of the offered lambda services and the *link-state parameters* necessary for IRWA. So, the system needs to be fed with information about nodes, links, channels, and sources of performance monitoring events, as well as the service-intrinsic parameters of the lambda service classes defined in Section III, and link-state thresholds for IRWA. The main processes of the INIM system are the filtering, correlation and aggregation of events, the verification of the SLA of each connection established, and the monitoring of the status of optical resources on a link-state basis. The last two processes correlate data from network topology, events and SLA/link-state parameters, and decide if any SLA fails, as well as the status of the optical links. The system provides two outputs; in normal operation, it logs performance information, including SLA and link-state validation. If an SLA is violated, INIM raises an alarm to the service management system, which is responsible for handling service-level issues. Similarly, if one or more parameters of an optical link exceed the thresholds defined as input parameters, the system raises an alarm to the GMPLS node that is ingress of this link, so that its local link resource management (LRM) module [13] can be updated. After this, the GMPLS mechanisms will flood the performance changes reported by INIM so that this information can be updated globally [14].

The information modelling of the INIM system is based on the above requirements and encompasses data and process modelling. Data modelling is split into four types: *elements and users*, *QoS characterization*, *lambda services* and *monitoring points*. While process modelling handles all the process requirements having into account the input and the output parameters and messages described above, the data model defines how information is stored and retrieved. It is important to design an open data model that ensures integrity and simplify the processes; using entity-relationship techniques, we defined entities based on tables and fields in each table with different properties (number, char, unique key, etc.), as well as relationships between the entities. Table II lists and describes the entities created, which are briefly described at the following.

NETWORK includes optical nodes and links, *NODE* includes active optical elements, *LINK* lists the interconnection of nodes, bandwidth available at each link, etc., and *USERS* contains a list of the network users. In order to configure QoS, *SLA*, *SLA_PARAMS* and *LRM* store SLA and link-state thresholds as input requirements. As for the monitoring points, *MONITORS* contains the devices that retrieve or capture performance information, *EVENT_FIELDS* describes the fields of events for each monitoring point in the table *MONITORS*, and *EVENTS* contains events related to

performance. Finally, lambda services encompass the entities *LSP_CNX_ENTRY*, which contains lambda services running in the network, *TUNNEL_USER*, which lists registered users with active services, *TUNNEL_NODES*, which includes the nodes involved in each active service, and *TUNNEL_LINKS*, which includes the links involved in each active service.

TABLE II
ENTITIES OF THE INIM SYSTEM

Type	Entity name	Description
Elements and users	<i>NETWORK</i>	Nodes, links and topology
	<i>NODE</i>	Network elements
	<i>LINK</i>	Node adjacencies
	<i>USERS</i>	Network users
QoS characterization	<i>SLA</i>	Service Level Agreements
	<i>SLA_PARAMS</i>	Parameters for each SLA
	<i>LRM</i>	Link Resource Management
Monitoring points	<i>MONITORS</i>	Spectral OPM and IP probes
	<i>EVENT_FIELDS</i>	Fields of events sent by monitors
	<i>EVENT</i>	Events received from monitors
Lambda services	<i>LSP_CNX_ENTRY</i>	New lambda service (historic)
	<i>TUNNEL_USER</i>	Users with active services
	<i>TUNNEL_LINKS</i>	Lightpath links of active services
	<i>TUNNEL_NODES</i>	Lightpath nodes (active services)

The process modelling is split in three categories: *network configuration*, which involves the data models of elements and users, as well as QoS configuration, *service management*, which mainly uses data from lambda services, and *QoS verification*, which contains SLA and link-state verification processes, and used data of monitoring points type. The first process is performed when initializing the INIM system, and populates the system's database with the required input parameters. The service management process registers and unregisters user connections based on setup and teardown notifications, and extracts the nodes and links traversed by lightpaths, as well as the SLAs associated to them. The SLA and LRM Verify processes receive events from the network, filter and aggregate them. Finally, QoS verification correlates information from 'elements and users' and 'QoS characterization' to check whether SLAs are violated (end-to-end for each connection), and if link-state information has been altered. At this stage, we must combine the data and process models described above to form the architecture of the INIM system. This architecture is the responsible of system data flow from the reception of a performance event to the logs and alarms of SLA violation and link-state information. Therefore, the system architecture defines the data flow between processes. We designed an architecture based on three pillars: the *monitoring points*, the *gatherers*, and the *event manager*. A brief description of each of these elements follows. For a detailed description, the interested reader is referred to [15].

This work considers a dynamic IP/WDM network, in which optical connections are established on demand with certain QoS parameters, captured in the SLA (Table I), and where active optical elements that may fail. However, the approach is also valid in static environments. Different monitoring points are needed throughout the network in order to generate events (alarms and notifications) about the status of optical resources

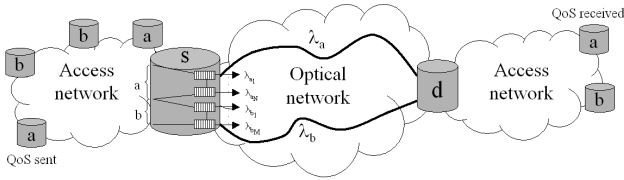


Fig. 1. Scenario of transparent lambda services.

and services. The status of resources and services encompass the ‘health’ of the optical channel carrying a service, which in this work is obtained through spectral monitoring techniques (optical channel power, OSNR and frequency drift), and the status of the information carried in the channel, which is transparent to the optical network and therefore can only be retrieved at the edges, where opto-electronic conversions take place. Moreover, both degradations in optical channels and failures in optical elements may occur, affecting service availability. Major component failures may take place in the control plane (faults in the GMPLS engines), and in the transport plane (lasers, amplifiers, switches, etc.). Therefore, these active elements are monitoring points for component failures. Monitoring points need to be placed conveniently in the transport plane in order to obtain status information of optical resources that will be communicate with the INIM system in a timely and optimized manner [16]. To estimate the status of data, IP meters are embedded in the nodes to retrieve packet loss and delays by using probe packets, for what we assume that the optical network ‘knows’ the traffic type carried on the wavelength channels.

This is achieved by considering that the edge elements are responsible for aggregating traffic flows of a given type or class so that a given tributary of an edge element corresponds to a traffic type. Given this, the optical network allocates sets of channels to each tributary, and QoS is based on wavelength allocation. Figure 1 illustrates this scenario; we may observe that the ingress edge router (source node s as seen by the transparent optical network) aggregates traffic flows a and b so that the optical network can allocate up to N WDM channels in the range $[\lambda_{a_1}, \lambda_{a_N}]$ to the traffic flow a and up to M channels in the range $[\lambda_{b_1}, \lambda_{b_M}]$ to the traffic flow b . The lambda services are accessed through an interface located at the ingress point in the optical network (OADM or OXC), where the end points of the transparent lightpath are accessible. Then, the provisioning process is based on setting up and tearing down optical connections (lightpaths) taking into account QoS: wavelength from a set according to the traffic type (service class) and path with bounded accumulation of impairments. The traffic type is indicated as framing in order to reserve resources taking into account the capabilities of the tributary ports, and QoS is directly related to tributaries and wavelength allocation.

The *gatherers* are entities that collect performance data from the monitoring points across the optical network. They are responsible as well for performing a first filtering of the information, and aggregating it prior to sending it to the event manager, in order to unify the data format from

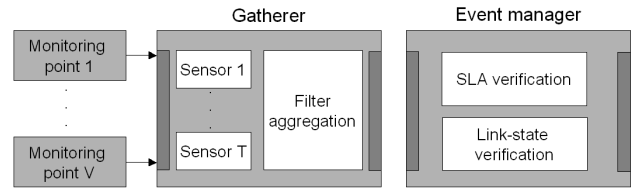


Fig. 2. Architecture of the performance monitoring system.

monitoring points. This allows the event manager to receive events in a unified manner. The gatherers are depicted in the center of Figure 2. The input messages of the gatherers are received or polled from V monitoring points distributed in the optical network, such as power meters embedded in the optical nodes, spectral monitors or IP meters, in T protocols and/or languages. The gatherers receive information about spectral (optical) and IP performance events, as well as notifications of setup and teardown of optical connections (lambda services).

The *event manager* (illustrated on the right of Figure 2) receives and processes the events related to performance of the optical services and resources. It is composed of a repository and a real-time information processor. This processor evaluates and updates changes in the status of resources, handles alarms (previously, it classifies events in notifications and alarms) and verifies SLAs and link-state information. If an SLA is violated or if a link exceeds its predefined thresholds for IRWA, the event manager raises the appropriate alarms. Functions related to performance management, security and policy, as well as billing and accounting are out of the scope of this work. The event manager is the core of the INIM system and supports three types of processing: lambda service setup, lambda service teardown, and monitoring event, which triggers checking of SLAs and link-state information, and updates network status information in its database. To do so, it requires all input parameters defined above, uses the processes of service management and QoS verification.

B. Processes and data flow

Starting from the system architecture described in the previous section, we developed a prototype of the INIM system, which is composed of six software processes that are based on several programming languages depending of the component criticism and self evolution while in development phase. Basically, critical time processes were developed in java, medium critical processes in Java/FESI and non time critical in Shell Script (ksh), as illustrated in Figure 3. The *gatherer* (trapreceiver) process receives events from monitoring points. It stores all events (for logging purposes) and forwards them to the verification and alarm handling processes described below. This process can receive two kinds of events; SNMP traps from the OPM monitors, and SNMP traps from the IP Meter. The OPM traps are related to events (*EventTypeMask*), links where events occur (*LinkIndex*) and lambda services affected by events (*ChannelIndex*). The IP Meter traps are proprietary and have the following objects and values:

- *DestinationNode* contains the IP address of the node that was ‘pinged’. For example, IP metrics from node 1 to node 8

of the ADRENALINE testbed would have a value of 10.0.50.8 in this OID.

- *SourceNode* contains the IP address of the node that performed the ‘ping’. Continuing with the previous example, the value of this object would be 10.0.50.1.

- *MetricTypeMask*, with types 1 (Round Trip Time, RTT) or 2 (PLR) and the numeric value of the IP metric type.

The *network characterization* process is part of the event manager, and its purpose is to configure the initial values of the monitoring system according to the topology, resources and policies of the optical network. To configure the network, it uses all the entities of elements and users (Table II). To configure the monitoring points, it uses the entities *MONITORS* and *EVENT_FIELDS* of Table II. The entities of QoS characterization (Table II) are used by this process to configure the SLA values (*SLA* and *SLA_PARAMS*) link-state thresholds (*LRM*). This process is launched when the INIM system is launched, and the values of the entities involved are entered manually.

The *service management* (*EM_UserCnx*) process, also part of the event manager, receives notifications of setup and teardown of optical services. Setup notifications contain all the connection-related parameters [17], from which this process captures the source and destination nodes, the path of the connection (which results in a relation of intermediate nodes and links), the unique connection identifier (*cnx_ID*), the wavelength channel/s that carry the lambda service, and the *class* of the lambda service, which will be used to verify the SLA. The outputs of this process are the files *lsp_cnx_entry*, which stores the parameters of the lambda service [17], *tunnel_links*, which stores the links traversed by the connection, and *tunnel_nodes*, which stores the IP addresses of the nodes traversed by the connection. The *QoS verification* process is split in three parallel processes, illustrated in Figure 3 as *EM_SLAVerify*, *EM_LinkUpdate* and *EM_LRMVerify*. All QoS verification processes receive filtered and aggregated events from the gatherers, and correlate them to update verify SLAs, update link status and validate link-state thresholds, respectively. To this end, the processes correlate information from ‘elements and users’ and ‘QoS characterization’ in Table II to check whether SLAs are violated (end-to-end for each connection), and if link-state information has been altered. To verify SLAs, this process is given the associated *cnx_ID* of each connection, from which it derives the SLA parameters and thresholds to verify, based on the *class* associated to the *cnx_ID* (retrieved by the *service management* process). An example of threshold verification (for the BER) follows:

```

Get  $OSNR_{dest}$  for cnx_ID
Get  $B_o$  and  $B_e$ 
Compute  $OSNR$  value for target  $BER$  (from eq. 3)
if  $OSNR \leq OSNR_{targetBER} + \Delta OSNR_{BERest}$ 
    return (BER verified)
else return (Raise alarm SLA violation for cnx_ID)

```

Note that the BER is estimated from the OSNR value at the destination of the connection, as described in Section

III. Both *EM_SLAVerify* and *EM_LRMVerify* need information from *EM_LinkUpdate*, which retrieves OMS- and OCh-layer information from the OWM and the IP Meters. Note that the QoS verification processes need prior categorization of events from monitoring points, which is illustrated in Figure 3 as *EM_OWM* and *EM_IPMeter*. Finally, *AlarmHandling* is a high-level process that evaluates the events received from the gatherers and, if necessary, triggers alarms. In other words, this process classifies events into notifications and alarms. Processes dealing with security and policy, scheduling services, billing/accounting are not taken into account in this work. There is, however, a common process for the above-listed ones that performs correlation between performance information of the different OSI layers (1 to 3, in this work) because troubleshooting is difficult between the layers.

The data flow of the INIM system is illustrated in Figure 3. Step *A* is the reception of an SNMP trap by the *gatherer* process. This event can be a notification of a connection setup/teardown or a performance monitoring event. Therefore, it is filtered accordingly (step *B* in Figure 3). Step *C* is the reception and categorization of unified information by the event manager (*EM_OWM* and *EM_IPMeter* processes). If the event received is a notification of a service setup or teardown, step *D* takes place (*EM_UserCnx* process). Otherwise, steps *E* to *G*, running in parallel, update the status of resources, and check potential SLA violation and/or link-state variation beyond the predefined thresholds. Step *E* receives events relevant to services, issues an alarm if the SLA is violated, and logs the SLA verification information. Inputs to SLA verification are events and SLA information:

```

FILE: data/events
      20050411.092030.037:add:04SF:4:7
FILE: data/sla
      # A+SLA_ID:[ALARM1:ALARM2:..:ALARMn]
      A0:01SD:02SD:03SD:04SD:05SD

```

The result of the SLA verification is a log and a possible alarm. The basic function (step *G*) of the link-state validation is to update information about resources to the GMPLS control plane, when an alarm is received:

```

FILE: data/events
      20050202.123454.232:add:04SF:4:7
      20050202.123454.232:del:04SF:4:7
FILE: data/lrm
      # L+LINK_ID:[ALARM1:ALARM2:..:ALARMn]
      L47:01SF:02SF:03SF:04SF:05SF

```

This update is done by raising an alarm that is processed by the management controllers hosted in each GMPLS node, and forwarded to the LRM module. A more advanced function of link-state validation is to aggregate information on a per link basis and as a function of the parameters considered by the IRWA performed by the GMPLS control plane. To this end, the INIM system integrates link models in its link-state validation function (Section V).

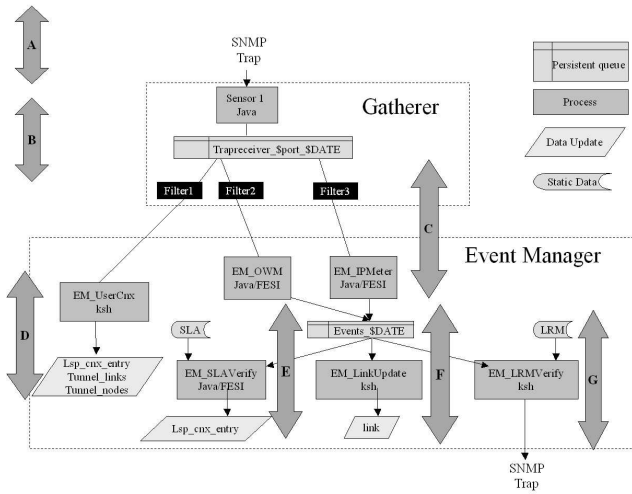


Fig. 3. Data flow of the INIM system processes.

V. EVALUATION OF THE INIM SYSTEM

In this section we evaluate the INIM system through outlining different time-related performance aspects. These aspects serve for obtaining the periodicity of verification. Moreover, an example of link-state information aggregation for IRWA is provided. The experimental implementation and evaluation of the INIM system is done in the ADRENALINE testbed [18], which is a hybrid platform, developed at the Centre Tecnològic de Telecomunicacions de Catalunya labs, which combines both real and emulated transparent nodes and links based on a distributed GMPLS-based control plane and a distributed management plane that integrates the INIM system. Nine Linux-based routers which emulate the GMPLS-based control plane, and three 1.5 GHz PCs emulate the INIM system. ADRENALINE is equipped with three 35-km fiber links (up to eight DWDM channels of 1-2.5 Gbps) and three spectral monitors that measure channel power, wavelength drift and OSNR, as well as aggregate power and OSNR. Monitors are SNMP enabled. Moreover, a router tester is employed to generate IP traffic and measure packet statistics. ADRENALINE is an intensity modulation direct detection system. The testbed has the peculiarity of combining both real (fiber) and emulated links, allowing the dynamic configuration of network topology; in this work, a ring configuration is used.

A. Experimental characterization of INIM delays

The primary application of the INIM system is the certification of SLAs between network operators and their clients. The study of the verification periodicity of the INIM system is crucial to ensure the certification of SLAs in a periodic manner (T_{SLA}). This study can be also applied to informing the control plane about any significant degradation in the optical links to be used in path computation for future service requests. Starting from the monitoring points, which inform of significant variations in the status of resources and services, we analyze the processing delays of all the elements of the system, until alarms are raised to the service management system or to a node's LRM module. Figure 3 illustrates the delays of

TABLE III
EXAMPLES OF VALUES FOR D_{event}

Parameter	D_{event}
Power	15 msec (sampling and SNMP Trap)
OSNR	110 msec (sampling and SNMP Trap)
Packet delay	5 msec (<i>ping</i> , RTT of 8-link path and SNMP Trap)

the INIM system. For any change or sample of the status of a lambda service or optical resource, the gatherers receive an event message after a delay named D_{event} (step A). The processing in the gatherers (filtering and aggregation) will take D_{not} time (step B). D_{event} depends on the monitoring point (sampling rate, interworking with the gatherers, etc.), whereas D_{not} may vary for each type of performance parameter only in what concerns the sensor process in the gatherer. Once in the event manager, D_{verify} can be either for SLA (depending on number of users, step E) or link/channel values (step G). In the latter, an alarm may be raised to the LRM of the/an affected node (D_{al}). In the control plane, information will be updated globally through computing a suitable algorithm (D_{LRM}) and flooding the aggregated information (D_{fl}). Then, the delay for verifying an SLA can be expressed as:

$$D_{SLA} = D_{event} + D_{not} + D_{verSLA} \quad (4)$$

where $D_{verSLA} = D_{verify}|_{action=SLA \text{ verification}}$, and the delay for informing the control plane about significant impairments on a channel/link is:

$$D_{l-s} = D_{event} + D_{not} + D_{verLRM} + D_{al} + D_{LRM} + D_{fl} \quad (5)$$

where $D_{verLRM} = D_{verify}|_{action=link-state \text{ validation}}$. Note that D_{event} is the time elapsed between the occurrence of an event and the alarm generated by the monitoring point that detected the event. An event can be a failure, or just the periodic obtention of measurements by a monitoring point. Failures are asynchronous if the alarms sent by the monitors are autonomous. Many monitors obtain samples of performance data at a given rate, which makes events (both notifications of resources' status and alarms) synchronous. Table III illustrates some values of D_{event} depending on the performance parameters, which have been obtained from the implementation performed in the ADRENALINE testbed. For channel power and OSNR, D_{event} is worst-case, that is, a change occurs right after sampling, which results in $D_{event} = T_{sampling} + D_{SNMPtrap}$.

After this, we analyze the impact of the different processes in the event manager in D_{verify} . Basically, D_{verify} does SLA and link-state verification in parallel. Link-state information verification is straightforward because the monitoring information basically comes from optical channel (OCh) and multiplex (OMS) levels at each link of the network. As for the SLA verification, IP metrics are end-to-end, and therefore also straightforward to correlate with SLA tables (see Table I). Finally, we analyze the delays of raising alarms of SLAs and link-state information. In both cases, the event manager takes D_{al} to notify the interested parties. In the latter, once

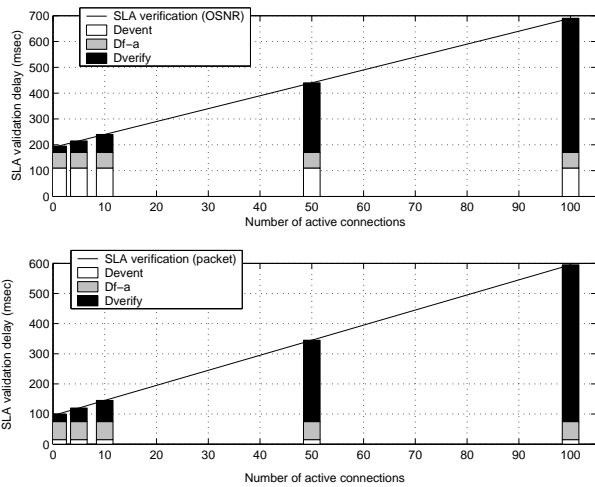


Fig. 4. Sample delays for SLA verification.

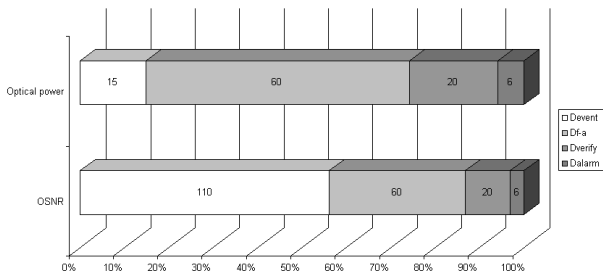


Fig. 5. Sample delays for link-state information update (OCh).

the alarm reaches an LRM, local tables are updated (D_{LRM}) and the updated information is flooded to the remaining nodes (D_{fl}). Sample values of these delays are given in Figure 4; SLA verification delays and range from 235 msec for verifying spectral and IP parameters (one service) to 730 msec for 100 services. Note that D_{verify} for SLA is a function of the number of optical connections (approximately $20 + 5S$ msec, where S is the number of active lambda services). Link validation delays, including flooding, are around 400 msec for OPM, and around 300 msec for IP metrics (Figure 5).

Note that the temporal resolution of monitoring delays is 2 msec (accuracy of the event manager’s OS), and that end-to-end delays are heavily influenced by D_{event} and the sequential validation of SLAs. This can be improved by choosing faster OPM monitors (commercially available equipment have scan times of 0.1 msec), implementing the event manager in real-time technologies and OS, and verifying SLAs in parallel (robust database). This way the INIM system would be rapid enough to be used for fault management, and also scalable for high numbers of active services.

B. Strategies to estimate link-state parameters from real-time monitoring information provided by INIM

Recently, IRWA is gaining interest because it allows to set up connection taking into account the status of physical resources that will be used for establishing the lightpath, hence helping assure QoS. There are two main alternatives to obtain

information about physical impairments; modelling and real-time monitoring. In the literature we may find models for some performance parameters that aim at capturing the most dominant impairments in high-speed networks, such as the polarization-mode dispersion (PMD) and OSNR constraint models proposed by Huang *et al.* [19], which require a centralized database with detailed performance monitoring information and iterative computation in the IRWA algorithm. Throughout this paper we have seen that recent advances in optical monitoring techniques make it possible to obtain spectral parameters in milliseconds and in a non-intrusive way, i.e. by extracting a portion of WDM signals by tapping optical fibers. The novelty of the INIM approach is the use of up-to-date performance monitoring information to build a link-state parameter that is used by IRWA in a distributed way and with no iterative computation.

This approach is interesting because the constraint models found in the literature, such as [19], make use of a centralized entity that has detailed information on the physical infrastructure, so that it can compute OSNR, PMD penalty, Q factor or BER from the ingress to the egress. Of course, this approach is the most exact, and also the most complex to implement in on-line IRWA decisions, especially in distributed IRWA: cost-efficient monitoring is incompatible with dynamic physical-layer parameter databases that contain information about all components and all wavelengths. The amount of monitoring points needed to update the values of the parameters needed by these models results in increased expenses. In fact, almost each network element along the path should report on its performance status. For this reason, in the literature performance parameters for IRWA are assumed to be static. Moreover, such a database would prevent distributed IRWA, which needs information update and forwarding, from being scalable. Of course, INIM may also provide channel-state information to IRWA (Figure 5).

Since lightpath computation is usually done on a per hop (link) basis, it is important to aggregate physical impairments in one or more link-state parameters. IRWA assumes that the quality of an optical signal traversing H hops can be estimated as a function (maximum, minimum, addition, etc.) of the impairments introduced by the optical elements along the path. A description of the estimation functions of the INIM system follows for OSNR and delay constraint models that are to be embedded in the GMPLS nodes follows. Note that PMD influences system performance in high-speed transmissions, i.e. 10 Gbps and above, and therefore it is not an issue in the ADRENALINE testbed.

To illustrate these estimation functions, we assume the setup depicted in Figure 6, which is a generic scenario with several services and changes in an extended ADRENALINE network. Four lambda services are established; the channels λ_{30} , λ_{31} are allocated to VoIP-like services, λ_{32} to IPTV-like, and λ_{33} to Internet+. An 8-port spectral monitor taps the input and output DWDM fibers of nodes 1 to 4, to monitor links 1-2, 2-3, 3-4 and 4-5. Moreover, IP meters located in nodes 1 to 4 measure packet loss and delays for each service (for example, the VoIP-like service carried on λ_{30} has an IP meter instance associated that pings node 5 from node 1 to get the above-mentioned

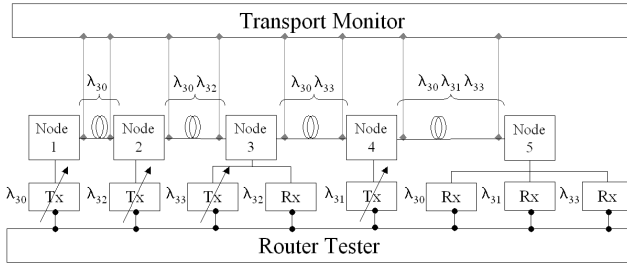


Fig. 6. Setting of the INIM system.

IP metrics). The delay constraint model is straightforward, because the INIM system can estimate the delay in any link of the network by combining the packet delays measured by IP meters and the knowledge of the lambda services' paths. For example, in the setup depicted in Figure 6, the INIM system obtains RTTs of 1 hop (λ_{31} and λ_{32}), 2 hops (λ_{33}) and 4 hops (λ_{30}). Therefore, it has four samples to estimate the delay in a link. This procedure can be also applied if links have different lengths.

The OSNR constraint model is more complex. A dominant impairment at any bit rate is noise (basically from amplified spontaneous emission, ASE). So, the OSNR level at the destination of a lightpath (Figure 7) can be expressed as:

$$OSNR_{d_{monitored}}(\lambda) \simeq \frac{P_{sig,d_{monitored}}(\lambda)}{P_{ASE,d}} \quad (6)$$

However, IRWA serves for estimating the QoS of future lambda services, which may be established over a lightpath that does not exist in the present, and distributed IRWA needs link-state information for scalability reasons. Therefore, the INIM system must provide a value of OSNR-related value for each link of the network, so that the IRWA algorithm can use a suitable expression to estimate OSNR from a given signal level at the source ($P_{sig,s}$), passing through the link spans of a target route, and verify whether the OSNR estimated is within the acceptable range for the class of the lambda service to be set up, according to Table I.

In this work, the approach taken is an aggregate link parameter, which shall integrate the status of a link in a reduced set of status values. In the literature we may find some examples of this parameter, which are based on the maximum distance (D_{max}). For example, RFC 4054 [20] proposes a translation of D_{max} into a transparent domain, in which the link parameter is the equivalent length of link k , which equals the distance sum of all fiber spans on the link and the equivalent length of fiber for the network element/s on the link. This proposal is only viable if non-OSNR constraints (PMD, ASE, etc.) are not binding factors as long as the maximum distance constraint is met, if appropriate network design is done and if no update on link status is envisaged.

In the case of OSNR, and assuming the link model depicted in Figure 7 and that the noise level is caused by ASE, on each link the INIM system collects channel power and OSNR values of active services and infers a single parameter for the link, with a limited set of values related to the minimum OSNR supported at OMS layer. For example, if the measured

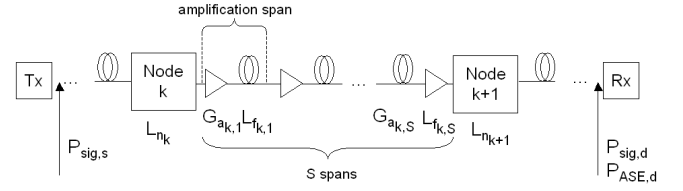


Fig. 7. Optical link model.

signal level of λ_{30} in the receiving end of node 5 (Figure 6) is -11 dBm, and the OSNR is 10 dB, by applying the appropriate OSNR threshold of Table I we obtain the minimum OSNR level supported by the channel. If we do the same with the rest of active channels on the link, we obtain a set of minimum thresholds for OSNR, which we may aggregate with control and analysis techniques such as fuzzy logic. Then, a route of H hops computed for a service class m accomplishes the OSNR constraint imposed by IRWA if it verifies $\min(OSNR_{agg1}, \dots, OSNR_{agg1}) \geq OSNR_{th}[m]$, where $OSNR_{agg}$ is the link-level value of OSNR obtained from active channels and $OSNR_{th} = OSNR_{targetBER} + \Delta OSNR_{BERest}$.

VI. CONCLUSIONS AND FURTHER WORK

We have presented an in-service non-intrusive performance monitoring system that is capable of certifying 'all-optical' SLAs and to infer on-demand performance information on a per link basis, which can be used for IRWA. The system is distributed, modular and low-complexity, and performs validation procedures in milliseconds. For example, it can validate 100 SLAs in less than 800 msec, and link-state information in less than 400 msec. The system monitors layer 1 (spectral) and 3 (IP) metrics in a non-intrusive way with respect to optical signals in a dynamic environment. Future work will focus on the computing and provision of updated link parameters for impairment-aware RWA. Key issues to be considered are the avoidance of unnecessary flooding of impairment-related information, for which thresholds for notifying the control plane of significant changes in the resources are crucial; and the aggregation of link-state information (currently on a per wavelength basis) to preserve scalability.

VII. ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers, whose valuable comments and suggestions helped to improve the quality of a previous version of the manuscript, as well as Rodrigo Jiménez for his collaboration in the software design and development of the INIM system. This work is part of the NetCat and EUREKA/CELTIC PROMISE projects, supported by the Centre Tecnològic de Telecomunicacions de Catalunya.

REFERENCES

- [1] E. Mannie/Editor, "Generalized Multi-Protocol Label Switching (GMPLS) Architecture," IETF RFC 3945, October 2004.
- [2] C. J. Youn and K. J. Park and J. H. Lee and Y. C. Chung, "OSNR monitoring technique based on orthogonal delayed-homodyne method," *IEEE Photonics Technology Letters*, vol. 14, no. 10, pp. 1469-1471, October 2002.

- [3] M. Petersson and H. Sunnerud and M. Karlsson and B.-E. Olsson, "Performance monitoring in optical networks using Stokes parameters," *IEEE Photonics Technology Letters*, vol. 16, no. 2, pp. 686–688, February 2004.
- [4] G. Rossi, T. E. Dimmick, and D. J. Blumenthal, "Optical performance monitoring in reconfigurable WDM optical networks using subcarrier multiplexing," *IEEE Journal of Lightwave Technology*, vol. 18, no. 12, pp. 1639–1648, December 2000.
- [5] Z. Tao, Z. Chen, L. Fu, D. Wu, and A. Xu, "Monitoring of OSNR by using a Mach-Zehnder interferometer," *Microwave and Optical Technology Letters*, vol. 30, no. 1, pp. 63–65, July 2001.
- [6] N. S. Bergano, F. W. Kerfoot, and C. R. Davidson, "Margin measurements in optical amplifier systems," *IEEE Photonic Technology Letters*, vol. 5, no. 3, pp. 304–306, March 1993.
- [7] Y. Hamazumi and M. Koga, "Transmission capacity of optical path overhead transfer scheme using pilot tone for optical path network," *Journal of Lightwave Technology*, vol. 15, no. 12, pp. 2197–2205, December 1997.
- [8] L. E. Braten and J. Karstad, "Internet traffic performance over a fixed radio communication link," Telenor R&D Report 36/2001 ISBN 82-423-0408-4, December 2001.
- [9] D17, "QoS monitoring, QoS requirements and QoS classes - including Service Level Specification," FP6-506760 NOBEL, annex of deliverable D17, February 2005.
- [10] D. Marcuse, "Derivation of analytical expressions for the bit-error probability in lightwave systems with optical amplifiers," *IEEE Journal of Lightwave Technology*, vol. 8, no. 12, pp. 1816–1823, December 1990.
- [11] M. A. P. A. Humblet, "On the bit error rate of lightwave systems with optical amplifiers," *Journal of Lightwave Technology*, vol. 9, no. 11, pp. 1576–1582, November 1991.
- [12] P. C. Becker, N. A. Olsson, and J. R. Simpson, "Erbium-doped fiber amplifiers fundamentals and technology," in *Optics and Photonics*. New York Academic Press, 1999.
- [13] G.8080/Y.1304, "Architecture for the Automatically Switched Optical Network (ASON)," ITU-T Recommendation, November 2001.
- [14] R. Martinez, C. Pinart, J. Comellas, and G. Junyent, "On-line icbr in a transparent gmpls network: a reality check," in *Proc. V Workshop in G/MPLS networks*, March 2006.
- [15] C. Pinart, A. Amrani, and G. Junyent, "Design and experimental implementation of a hybrid optical performance monitoring system for in-service SLA guarantee," in *Proc. 9th IFIP/IEEE International Symposium on Integrated Network Management (IM)*, May 2005.
- [16] —, "Monitoring service 'health' in intelligent, transparent optical networks," in *Proc. IFIP Optical Networks & Technologies Conference (OpNeTec)*, October 2004.
- [17] C. Pinart and G. Junyent, "Experimental test of management integration in GMPLS enabled metro WDM networks for service provisioning," in *Proc. 30th European Conference on Optical Communication (ECOC)*, September 2004.
- [18] R. Munoz, C. Pinart, R. Martinez, J. Sorribes, M. Maier, A. Amrani, and G. Junyent, "The ADRENALINE Test Bed: Integrating GMPLS, XML and SNMP in transparent DWDM networks," *IEEE Communications Magazine*, vol. 43, no. 8, pp. s40–s48, August 2005.
- [19] Y. Huang, J. P. Heritage, and B. Mukherjee, "Connection provisioning with transmission impairment consideration in optical WDM networks with high-speed channels," *Journal of Lightwave Technology*, vol. 23, no. 3, pp. 982–993, March 2005.
- [20] A. Chiu/Editor, "Impairments and other constraints on optical layer routing," IETF RFC4054, May 2005.



Carolina Pinart (SM'02-'05) was born in Barcelona in 1975. She is a graduate in Telecommunications Engineering and PhD of the Universitat Politècnica de Catalunya (UPC, Barcelona, Spain, April 1999 and December 2005, respectively). She joined the Centre Tecnològic de Telecomunicacions de Catalunya (CTTC, Barcelona) in November 2001 as a Director of Institutional Relations, collaborating as well with CTTC's Optical Networking Area. She is member of SPIE and recipient of a Fundació Agrupació Mútua 2003 Graduate Prize. Prior to joining CTTC, she served as undergraduate researcher (Center for Research Siemens-Nixdorf, Munich, 1998-1999) and as consultant in technology (Altran Group, Paris, 1999-2001), holding positions of research engineer (Ericsson France), project leader (FERMA and Mobinil) and ITS project manager (Renault R&D). Since 2000, she has been participating in 9 Spanish and EU R&D projects (ESPRIT, EUREKA and IST programmes) in wireless and optical networking. She has published 4 papers in journals and more than 20 papers in conferences about optical management and monitoring.



Gabriel Junyent (1950) is Telecommunications Engineer (UPM, Madrid, 1973), and holds a PhD degree in Communications (UPC, 1979, Barcelona). He has been teaching assistant (UPC, 1973-1977), adjunt (UPC, 1977-1983), associate (UPC, 1983-1985), professor (UPC, 1985-1989), and he is full professor since 1989. In the last 15 years he has participated in more than 34 national and international R&D projects, both public (CICYT, DURSI, European Commission or CIRIT) and private funded. Since 1982 he has published more than 25 journal papers and book chapters (IEEE/OSA Journal of Lightwave Technology, Optical Communications, IEEE Photonics Technology Letters, IEEE Lasers and Electro-Optics Society, among others). In the last years he has published more than 30 conference papers (N&OC, LEOS, CLEO, OFC, ECOC and SPIE international conferences, among others), more than 50 conference papers in national conferences, and has made several presentations as invited speaker.