Quality of Provisioning as an OPEX-related Issue in Research Networks

Guido Maier\textsuperscript{A}, Lorenzo Colombo, Dario Giuseppe Costantino, Achille Pattavina\textsuperscript{A}, Péter Szegedi\textsuperscript{B}

\textsuperscript{A} Politecnico di Milano, Dept. of Electronics and Information, Piazza Leonardo Da Vinci 32, 20133 Milano, ITALY
Tel.: +39-02-2399.3580 – Fax: +39-02-2399.3413
\{maier, pattavina\}@elet.polimi.it

\textsuperscript{B} Magyar Telekom - PKI Telecommunications Development Institute, 9 Magyar tudosok krt, 1117-H Budapest, HUNGARY
Tel.: +36 1 481 7560 – Fax: +36 1 481 7505
szegedi.peter3@t-com.hu

Abstract — National Research and Education Networks (NRENs) are investigating the option to migrate their infrastructure from IP over static transport network to IP over control-plane based transport network, possibly deploying GMPLS or ASON technology. In the NREN-participated MUPBED project, quality of provisioning of bandwidth on-demand services has been studied. In this paper the main results are presented showing that without an automatic control-plane, connection provisioning is impaired by setup and release delays. In managed networks, performance under dynamic traffic may be severely limited if an insufficient number of operators are active in the network operations center. The relation between number of operators and network performance is proved by theoretical models and simulations. The reduction of network-configuration workload achievable by an ASON/GMPLS control plane may be an important motivation for NRENs because of their specific operational-expenditure model.

Index Terms — On-demand connection provisioning; setup and release delays; network operations centre; managed networks; control plane

I. INTRODUCTION

Bandwidth on demand (BoD) is the new service of the future that all operators are looking at. Delay in provisioning of requested connections (or tearing them down when not used any more) becomes obviously an important limiting factor to the delivery of a satisfactory BoD service. The quality of provisioning, as we may call it, is relevant to the operator as well, since network efficiency is dependent on it.

In this paper we are presenting a study aimed at finding out the relation between connection setup and release delays and network efficiency in terms of blocking probability in dynamic traffic conditions (i.e. all BoD requests). We will investigate in particular what happens when an attempt is made to offer BoD service with a network that is not controlled by an ASON/GMPLS control plane, but is configured by the Network Operations Centres (NOCs). In such a network provisioning delays are so relevant that by comparison the same delays in a control-plane based network are negligible. A large component of provisioning delay is in fact due to queuing of connection requests at the NOC waiting to be served. This delay is strongly dependent on the number of operators available at the NOC and we will show that, for any given dynamic traffic load, there is a minimum number of NOC operators under which network performance become unacceptable. This fact is particularly important for the National Research and Education Network (NREN) operators, due to their specific operational-expenditure (OPEX) model.

The following sections of the paper are dedicated to answer to the second and third basic questions that surely have already came to the mind of the reader: why the scenario in which BoD is offered by a NOC-managed network should be considered in the NREN context (Sec. II); why the number of NOC operators is important in the NREN OPEX model (Sec. III).

The first question that the reader most probably has in his mind, i.e. why focusing the NREN case at all, is readily answered. A general reason is that NRENs are important as class of network operators needing to support advanced applications and usually open to technological innovation due to their agile structure. As a more specific reason, this study has been carried out in the framework of the IST FP6 European research project MUPBED [1], which is specifically dedicated to investigate the application and deployment of ASON/GMPLS technology in the Pan-European research infrastructure.

The other sections IV and V are specifically dedicated to the main technical topic of the paper that is quality of provisioning in an IP-over-OTN.

II. EVOLUTION OF THE NREN INFRASTRUCTURE

Many applications are today requiring guaranteed bandwidth and QoS to network operators in general. This is particularly true for NRENs as they support the most advanced scientific applications and interconnect a scientific community having more and more the need to distribute information: not just final results to be shared but nowadays also data to be processed (or stored) in different locations. During MUPBED [2] the following examples of advanced application were identified: VLBSI (Very Large Base Space Interferometry: interconnection of remote radio-telescopes), grid computing, high-definition video production, video transmission for teleme-
cine or distance-learning, etc. Most of these requirements are today satisfied by offering fixed connectivity. But this seems not to be a future-proof solution, taking into account that users need mobility and also that the requests for QoS/bandwidth demanding applications are going to increase. Bandwidth on Demand (BoD) service seems the most promising solution to both satisfy an increasing request volume and at the same time allowing an efficient network utilization by the operator.

Implementing BoD service in most cases requires upgrading network technology currently deployed in most NRENs. Up to very recent past, NRENs main services have not been QoS-critical, but basically best effort IP (at most adding to that one single undifferentiated “premium-IP” class). Traditionally, NRENs have followed the top-down approach building their own network infrastructure based on IP routers, since this is a cost-effective and less investment demanding. The infrastructure for point-to-point interconnecting the routers was mainly loaned or leased from one or more transport service provider. The “pure IP” model (Fig. 1) [3] contains a big-fat router with traffic ports towards the leased DWDM equipments.

Fig. 1. IP over transport-network reference models: pure IP, IP over OTN, IP over ASON/GMPLS.

An enhancement currently adopted in many NRENs is the transport of traffic flows requiring special QoS guarantees over dedicated connections. In most cases this is implemented by creating layer-2 pipes (commonly Ethernet connections) carried on configured MPLS Label Switched Paths (LSPs). This has proven so far to be a successful solution, but MPLS has been primarily used for management purposes and not to provide actual QoS guarantees. As more and more NRENs are acquiring their own fibre infrastructure by investing in WDM transmission systems as the technology of the future, the MPLS LSPs could be gradually substituted by configured layer-one transport circuits of an Optical Transport Network (OTN), dedicating e.g. SDH channels or even wavelength paths to carry QoS-guaranteed flows.

For our scope, the specific implementation used to provide transport, whether MPLS LSPs or layer-one circuits, is not important: the two solutions both have in common that connections are configured by exploiting the network management system in a centralized way. Given this commonality we will refer for simplicity to both the cases with the term “IP over OTN” architecture [3] (Fig. 1).

For many technical reasons, including the one that we will present in this paper, neither the pure-IP nor the IP over OTN models seems to be satisfactory in effectively providing BoD service. Thus there is a keen interest of NRENs in investigating the control-plane based model in order to find possible migration strategies. This interest is testified by several research projects (as MUPBED) participated or promoted by NRENs and dedicated to control-plane technologies such as GMPLS and ASON. In what we refer to with the term “IP over ASON/GMPLS” model [3] (Fig. 1) an intelligent control plane distributed on the multi-layer nodes is introduced. This enables to set-up and tear-down transport-plane connections automatically, based on the actual characteristics of the BoD traffic. The dynamic behaviour of the optical connections leads to statistical gain under dynamic traffic and hence cost efficiency also in the transport layer.

III. OPEX: GENERAL MODEL AND NREN-SPECIFIC ISSUES

When designing a new network infrastructure or choosing to adopt a new network technology, not only the initial investment (Capital Expenditure - CAPEX) shall be taken into account, but also the economical resources needed to maintain the network working after the deployment phase. The Operational Expenditure (OPEX) can in some cases be as relevant as or even more important than CAPEX.

Based on literature studies [4, 5, 6], the total OPEX can be partitioned in several sub-costs, as shown in Fig. 2. In comparing the aforementioned network scenarios in terms of OPEX, some components (administration, planning, supply, etc.) can be assumed constant in the three scenarios and thus neglected. The simplified OPEX is then given by the sum of maintenance and reparation costs, and provisioning and service management costs.

In our simplified model, provisioning and service management are reduced to the simple provisioning, which includes the costs for setting up or modifying the connections in the network. Maintenance and reparation are the costs for maintaining the network, keeping it operational. We will not consider here such OPEX contribution, in order to focus just the provisioning component (including it would lead to same conclusions [3]). OPEX for provisioning can be assumed proportional to the workload (person-months in a year) dedicated to network management and reconfiguration operations.

In the pure-IP scenario, this workload is negligible, since classical IP routing protocols (e.g. OSPF, BGP) are usually adopted. Human operator intervention is only occasionally needed to configure routing parameters and manage network elements. However it should be noted that the pure-IP model is not able to guarantee per-connection QoS. In the IP over OTN case QoS is guaranteed by the transport network, but connection provisioning requires the intervention of the network-management plane, with most functions manually carried out by the network operators. Provisioning-OPEX is expected to be high. The IP over ASON/GMPLS architectural scenario is equally able to provide QoS, and in addition the automatic
setup and tear-down capabilities of the distributed control plane allow large savings in OPEX by drastically reducing network configuration workload. In conclusion, the IP over OTN scenario is the most OPEX-critical of the three.

Let us now map the general OPEX model defined above to the NREN reality. It is important to underline the non-profit nature of the research environment. This can explain the fact (apparent by comparing different versions of the yearly TERENA Compendium [7]) that NREN budgets tend to be stable over time. Thus NRENs will probably accept to introduce the BoD service provided that their OPEX remains roughly the same [3].

An NREN interested in providing BoD with guaranteed QoS can choose whether adopting an “IP over OTN” solution or migrating to the ASON/GMPLS scenario. The first option is surely CAPEX saving because it does not require great changes to the legacy network infrastructure of most NRENs.

From an OPEX point of view, however, there is a significant benefit in implementing ASON/GMPLS functionality (compared to static legacy solutions); automatic provisioning can dramatically decrease provisioning workload. A particular aspect here should be considered. OPEX is fundamental for NRENs, but it is not usually seen as OPEX but workload instead: more or less people needed. In addition, NRENs usually have limited personal budget. For this reason, ASON/GMPLS features would be desirable for relative OPEX decrease. The advantages of automatic provisioning would ease the network operation and make it feasible for NRENs to manage the new layer without painfully increasing Personnel costs. In other words, an ASON/GMPLS enabled network with automated UNI and NNI interfaces is particularly appealing for NRENs as it is managed with limited human resources, due to the fact that most of the tasks are performed by the control plane.

IV. PROVISIONING DELAYS IN DYNAMIC TRAFFIC CONDITIONS

As anticipated, one of the most important advantages of introducing ASON/GMPLS control-plane technology is the drastic reduction of delays related to connection provisioning operations. The ASON/GMPLS control plane automatically performs these operations in times of the order of seconds or microseconds [9, 10]. In non-ASON/GMPLS transport networks (as in the IP over OTN case) connection setup/release require instead the human intervention to reconfigure the network via the management system, implying much larger delays, of the order of hours or even days. While ASON/GMPLS control plane is fully distributed, a limited number of network operators who are located in a Network Operations Centre (NOC) manage an IP over OTN: control is thus centralized. In this section we will show how the centralized nature of the managed networks [11, 12], conjugated to the high provisioning delays, can severely limit an IP-over-OTN network performance in case of a dynamic-traffic demand, justifying the migration to ASON/GMPLS as a future-proof solution. Provisioning delays in the IP-over-OTN case will be evaluated, also taking NOC effects into account. The IP-over-ASON/GMPLS case will be considered for comparisons, obtaining it by the general model in which setup and release delays are set to zero. According to the MUPBED framework, we will refer to European NRENs when estimating model parameters and selecting case-study topology.

A. Setup and release delays and general NOC model

The process of connection setup (release) implies the following main steps [13]:

a) a request for the new connection is issued by the user to the network provider;

b) the provider forwards the client’s request to the NOC;

c) a NOC operator computes the route for the connection taking into account the QoS constraints requested by the user (e.g. bandwidth, delay) and the availability of resources in the network;

d) if enough resources are available, the NOC remotely configures the network elements (e.g. SDH switches, optical cross connects, MPLS routers) in order to reserve resources for the connection. If network resources are not sufficient, the NOC notifies the provider and an expansion of network hardware may be decided, returning to step c afterwards. Otherwise, the connection request is refused (or blocked);

e) the outcome of the setup process (either activation or refusal) is sent by the provider to the user, who can start transferring data through connection in case of a positive outcome.

The exchange of messages between the various entities involved may occur using the network itself with suitable signaling protocols (e.g. RSVP) or completely “out-of-band” (e.g. via e-mail, chat-line, phone, faxes, etc.). In the general case, there may be more than one NOC if the network is partitioned into different operational areas. More NOCs would help to speed up provisioning by sharing the workload by more operators. On the other hand, coordination between far-apart operators may complicate the sequence, especially if the NOCs fall within different network administrative domains. In this paper we have considered only the single-domain and single-NOC case for simplicity, leaving the multi-NOC case for further study.

A connection is characterized by a set of parameters: \( r, (s,d,t_a,h) \), where \( s \) and \( d \) identify the source and destination, \( t_a \) is the instant when the request is issued and \( h \) is the duration of the connection (holding time) desired by the user. Moreover \( t_s \) is the actual starting instant of the connection, \( d_s \) and \( d_r \) are setup and release (provisioning) delays and \( t_s \) and \( t_r \) represent the starting and the ending instants of the process. All these parameters are related each other as in the time-diagram shown in Fig. 4a. In our simplified model all requests are for OTN-layer optical connections (lightpaths) which need a single unit of bandwidth (a WDM channel) on each crossed link.

The general model of NOC we are going to adopt is illustrated in Fig. 4b. Assuming a single-NOC scenario, the centre collects all setup requests from all the nodes of the network and manages each request providing necessary configuration operations to initiate and close. Resources in the NOC dedicated to network operation and management are limited: re-

---

2 TERENA is the association of European NRENs [8].
quests that arrive when all \( m \) NOC operators are busy have to be queued. We have assumed a double queuing system with a queue for tear down events and one for setup requests (both FIFO-based) sharing the same pool of servers (NOC operators).

![Fig. 4. (a) Time diagram of the connection provisioning process. (b) Single-NOC double-queue model.](image)

Each operator can carry out one setup or release task at a time and is therefore busy during a time equal to the duration of the setup or release sequence. We identify the operator busy periods with the term setup/release service time, in order to avoid confusion with the setup/release delays that takes into account also queuing delays. In our model we will assume for simplicity equal setup and the release service time.

Let us now consider the average behaviour of the NOC when many requests arrive as result of a dynamic BoD traffic in the network. The setup queue receives connection requests from all the \( N \) nodes of the network at a rate \( \Lambda \); the average total setup delay is \( d_s = D_{\text{setup}} + W_r \), where \( D_{\text{setup}} \) is the average service time for setup and \( W_r \) is the average queuing time at setup queue. We have assumed the setup queue capacity is limited to hold only up to a maximum number \( q_s \) of requests, so that the NOC refuses new demands in overload conditions. The release queue must not instead be limited to guarantee that sooner or later all connections active in the network will be terminated (else, network resources would remain allocated to unused connections forever).

Provided a request can be queued in the setup queue (otherwise it is refused with probability \( \Pi_Q \)), the requested lightpath is actually setup if enough resources are available in the OTN at the time of request processing; otherwise, it is lost (with a blocking probability \( \Pi_B \)). An accepted connection enters the release queue after its holding time \( h \) is expired. The average total release delay is \( d_r = D_{\text{release}} + W_r \), where \( D_{\text{release}} \) is the average queuing time at release queue. The arrival rate of requests at the release queue is equal to \( \Lambda' = \Lambda'(1 - \Pi_Q) \), where \( \Lambda' = \Lambda(1 - \Pi_Q) \) is the arrival rate of connection requests to the network.

**B. Markovian simple model**

Let us show how the NOC-based provisioning process can be modelled. It should be mentioned that similar studies already presented in literature \([9, 14]\) assume setup and release operations to be performed by different servers. We are proposing for the first time a model that considers a common pool of servers taking care of both setup and release operations in time-sharing. This vision is more in-line with the embodiment of the server-pool as a set of NOC operators.

Be \( m \) the total number of servers in the NOC. For simplicity we consider the network to be loaded by a dynamic Poisson traffic with assigned offered load \( A_s = \Lambda_s/\eta \) (Erlang), where \( 1/\eta = h \) is the mean connection holding time, assumed to have a neg-exp distribution. Also setup and release service times are assumed to have neg-exp distribution with equal mean value \( D_s = 1/\mu \). Under these assumptions the setup queue can be regarded as an \( M / M / m / m' + q_s \) \([15]\) Markovian system with \( q_s \) the maximum number of setup requests that can be queued and \( m' \) the number of operators that are available for the setup process (i.e. not engaged in release operations). \( m' \) is in principle variable, but it is expected to converge to an equilibrium value if \( A_s \) remains constant.

The output of this queue (arrival rate \( \Lambda' \)) determines the traffic effectively offered to the network. The network load is \( A_o = \Lambda'/\eta \). The effective average holding time of connections \( h = 1/\eta \) is somewhat complicated to compute. A new connection request is not processed until it reaches the top of the setup queue. Thus during its waiting time in queue it does not impact on network resources. We assumed that resource engagement starts only at the end of the setup operation (after the service time \( D_{\text{setup}} \) – see Fig. 4a). On the other hand waiting time in the delay queue is detrimental to network performance since resources remains allocated to a connection until the release operation is over. Thus we have: \( h = h + d_r = h + D + W_r \).

The network blocking probability \( \Pi_Q \) depends on \( \Lambda_o \), on the amount of resources (WDM channels) and on their distribution on the topology of the specific OTN considered. As connection holding-time distribution is neg-exp, it can be proved that the arrival process of release requests at the release queue is still Poissonian. Thus the release queue can be modelled as an \( M / M / (m - m') + \infty \) Markovian system with unlimited capacity (neglecting that the “user population” – i.e. the established connections needing termination – is not infinite, but limited at least by \( m' + q_s \)).

The total offered load (setup + release requests) to the servers is \( \rho = (\Lambda_s + \Lambda_r)/\mu \). If \( \rho < m \) then \( m' = m - \text{int}(\Lambda_s/\mu) \).

When \( \rho \geq m \) the setup queue starts to quickly saturate: \( \Lambda' \) (and consequently \( \Lambda_r \)) remains practically constant as \( \Pi_Q \) grows up. We have: \( m' = m - \text{int}(\Lambda_r/\mu) \).

Given the above relations and the input load of the setup queue \( \sigma = \Lambda_s/\mu \), the blocking probability of the setup queue
is the probability of having \( m^i + q_i \) requests in the system:

\[
\Pi_Q = \left( \frac{\sigma}{m^i} \right)^q \frac{\sigma^q}{m^i!} \left[ \sum_{i=0}^{\infty} \frac{\sigma^i}{m^i} \right]^{-1}
\]  

(1)

The average waiting time in the setup queue is:

\[
W_s = \frac{1}{\Lambda_{c} (1 - \Pi_Q)} \frac{p_0}{m^i!} \left( \sum_{i=0}^{\infty} \frac{\sigma^i}{m^i} \right)^{-1}
\]  

(2)

where \( p_0 \) is the probability of empty queue

\[
p_0 = \left[ \sum_{i=0}^{m^i} \frac{\sigma^i}{m^i!} \right]^{-1}
\]

Once \( \Pi_Q \) is known, \( W_s \) can be easily computed solving the simple \( M/M/(m-m')/\infty \) model [15] (equation not reported for brevity). Eq. 1 is not easily solvable, since \( m^i \) depends on \( \Pi_Q \) itself and on \( \Xi_N \), which in turn depends on \( W_s \). We have used an iterative numerical solution method, jointly with simulation to compute \( \Xi_N \) as a function of \( A_s \), once a specific OTN is considered (see model validation results in Sec. V).

C. Optimized NOC model

As we stated, the average effective connection holding time in the OTN is \( h = h + D \cdot W_s \). The NOC model can be optimized trying to decrease \( W_s \) in order to reduce the load on the OTN. To do this we can assign to release operations a (static) priority over setup operations. Obviously, the setup service will get worse, but an overall performance improvement (by means of a lower \( \Pi_Q \)) can be expected.

In general when \( M \) static priority classes are defined in a Markov queue system, the average waiting time in queue for a generic class \( i \) is [15]:

\[
W_s^i = \frac{W_s}{(1 - a_i)(1 - \Lambda_i / \mu)}
\]

where \( W_s \) is the waiting time provided by Eq. 2 (without priority) and \( a_i = \sigma + \Lambda_i / \mu \). Obviously, \( W_s^c = W_s \) (the top-priority traffic is unaffected). We have validated the model with simulations also in this priority case (comparison not reported for brevity) achieving good accordance.

V. SIMULATION RESULTS AND ANALYSIS

In order to simulate the IP over OTN scenario with specific reference to NREN operators, we have selected as case-study network GÉANT2* (Fig. 5), a simplified version of the GÉANT2 [16], the research network interconnecting NRENs in Europe. Connection request are assumed to be (10Gbit/s-modulated) lightpaths and the network is assumed to be opaque (wavelength conversion capability in each node). Connections are all routed on the shortest available path. Each one of the \( N = 18 \) nodes is source and destination of connection requests, as it represents the access point of an NREN to GÉANT2. Traffic volume generated by each node is modulated according to realistic traffic measurements [3].

For the simulations, we have assumed the service time for a given connection to be proportional to the connection length (number of crossed links). This is more realistic than the neg-exp distribution: setting-up connections crossing less network hops requires to reconfigure fewer nodes and to reserve fewer resources, and thus less processing and signalling efforts.

Given a connection \( c \), its service times are:

\[
D_{\text{setup}} = D_{\text{release}} = D_{\text{hop}} \cdot \bar{L},
\]

where \( D_{\text{hop}} \) is a specific per-hop service time and \( \bar{L} \) is the number of network hops crossed by \( c \) according to the path it has been routed on. Averaging for all connections:

\[
D_{\text{setup}} = D = D_{\text{hop}} \cdot L,
\]

where \( L \) is the mean connection length. After evaluating \( L \) of GÉANT2*, we have set \( D_{\text{hop}} = 2.5 \) hours taken as basic time-unit. Thus: \( D_{\text{setup}} = 1 \text{ and } h_{\text{setup}} = 10 \). The setup queue of the NOC has been limited to \( q_c = 18 \) positions (except when otherwise mentioned).

First, the network has been simulated assuming the simple NOC model without priority. Fig. 6a and 6b show, respectively, the total blocking probability \( \Pi_r = \Pi_Q + \Pi_N \) and the network throughput \( A_r = A_s(1 - \Pi_N) \) as a function of the offered load \( A_s \) for different numbers \( m \) of operators in the NOC. The ASON/GMPLS case (negligible provisioning delays – briefly named “ASON”) is reported by comparison. The graphs clearly show how \( m \) clamps the network throughput.

---

3 \( L \) is evaluated, by approximation, only taking into account connections that have been actually setup (non-blocked)
preventing a full utilization of OTN resources due to provisioning configuration operations at the NOC. The average setup delay perceived by the user $d_s=D_{setup} + W_s$ is reported in Fig. 6c: delay starts to build up as soon as the network throughput is clamped by the NOC and rapidly reaches high values (limited only by the finiteness of the setup queue). Fig. 6d compares $\Pi_s$ values obtained by simulation$^4$ with the corresponding values computed by numerically solving Eq. 1 and 2, showing good accordance of our model with simulations (both in terms of $\Pi_s$ and of $m$). 

Then we have simulated the network considering the optimized NOC model with prioritized release operations. The throughput is displayed in Fig. 6e, while Fig. 6f compares $\Pi_s$ in the two cases with and without priority, showing the expected advantages attained by the NOC optimization in terms of network blocking probability.

We have so far assumed the behaviour of the customers of the BoD service to be “persistent”. It is probably more realistic to set a threshold on the waiting time after which the connection is lost as the customer gives up his request (for example because it is switched to another network operator), abandoning the system. We have finally simulated GEANT2$^*$ introducing user defection ($q_c = \infty$ in this case). The defection time has been set to $T_{max} = 9$ hours ($T_{max,norm} = 3.5$). Fig. 6g and Fig. 6h display the resulting network throughput and setup delay.

Fig. 6k reports the conclusive most important outcome of this work. It concerns the two scenarios: optimized NOC with priority and with user defection (the case without priority is very similar to the case with priority, and is thus omitted). The maximum offered load $A_{O,max}$ and throughput $A_{T,max}$ are plotted as functions of the number of NOC operators $m$. The attribute “maximum” here means the greatest possible under the condition of preserving good network performance (i.e. high resource utilization) and acceptable quality of provisioning towards the users (i.e. limited setup delay). The $A_{O,max}$ and $A_{T,max}$ values plotted correspond to the values of $A_O$ and $A_T$, for different $m$, when the throughput curves start deviating from the ASON/GMPLS curve (knee-points) in graphs like those in Fig. 6b and 4g; or – equivalently – when $d_s = 1.5 \cdot D \Rightarrow W_s = D/2$ in the delay graphs (Figs. 4c and 4h). In an OPEX context, Fig. 6k should be read as follows: if an IP-over-OTN operator wants to offer BoD service and expects a given maximum offered load $A_{O,max}$ he can satisfy his customers (avoiding them to wait too long for connection setup) exploiting the network infrastructure at its best only provided that $m$ operators will be present at the NOC, where $m$ can be derived from the graph. Actually, he will be able to guarantee only $A_{T,max}$ as throughput, but the mismatch between $A_{O,max}$ and $A_{T,max}$ only depends on the network capacity and not on the management plane (i.e. it can be improved only by installing new capacity). $m$ gives indication to the operator about workload and thus OPEX that has to be sustained. If $h$ is constant, $m$ increases almost linearly with the arrival rate of BoD requests $A_s$ (i.e. traffic dynamicity). It is remarkable that in our example 4 NOC operators are sufficient if each node receives 1 setup request per day, but at least 12 operators are needed if demands for BoD service increase to 4 per day per node.

VI. CONCLUSIONS

From this study we can conclude that an OTN not equipped by an automatic control plane is limited in performance by the number of operators of the centralized management system. If the network has to provide a BoD service with good effectiveness, an OPEX growth must be sustained to increase the number of NOC operators: this penalty gets more severe as traffic becomes more dynamic. If the number of NOC operators is too scarce, the network remains under-utilized and the waiting time increases dramatically, generating defections of the customers. Since OPEX, especially in terms of workload, is very important for NRENs, this study seems to indicate that migrating to an ASON/GMPLS network is a preferable choice for these operators compared to trying to offer BoD service with a NOC-managed infrastructure.
Fig. 6. GÉANT2* case-study results (see the text for the caption of each graph)