Operator use cases that benefit from multi-layer optimization and application awareness

Victor Lopez[(1)], Dimitrios Konidis[(2)], Domenico Siracusa[(3)], Juan Pedro Fernandez-Palacios[(1)]

[(1)] Telefónica I+D/GCTO, c/Ronda de la Comunicación s/n, 28050 Madrid, Spain.
victor.lopezalvarez@telefonica.com
[(2)] Athens Information Technology (AIT), Athens, Greece
[(3)] CREATE-NET, Trento, Italy

Abstract *Multi-layer optimization enables the operators to optimize their packet and transport resources. Application awareness will provide potential savings as well as offer a better adaptation of network services to applications.*

Planning in multi-layer environments

During the last decades, backbone networks have experienced a clear trend towards simplification. IP/MPLS technology became the current standard in packet core networks, while WDM/OTN networks with GMPLS control plane emerges as the next generation transport network. It combines the scalability of WDM technologies with the dynamicity provided by a control plane. Besides, even though IP/MPLS and WDM/OTN still represent two significantly different domains, previous work[1] demonstrates that significant CAPEX savings could be obtained by a rational combination of packet and optical resources.

When deploying a packet/optical network, the network operator must consider which is the virtual topology to deploy (Fig. 1a). Services between the routers can be offered using different alternatives. The operator may decide to purchase more IP interfaces, as depicted in Fig. 1b, or to deploy new fibers between the routers to use grey interfaces or use direct fibers between the routers (Fig. 1c). The fourth alternative is to invest on transponders and ROADMs to create direct connections between the routers, as shown in Fig. 1d. In real deployments, such process is a mixture of the previous alternatives, aimed at minimizing CAPEX while ensuring that enough bandwidth to serve customers' demands is provided.

Beyond realizing a cost-effective network, operators must cater to the needs of customers' applications. Indeed, many applications may be better served by services that are not necessarily constrained by just bandwidth, but also by other network parameters such as latency, jitter, reliability, data location, etc., thus requiring more complex transport services.

In this paper, we present relevant operator use cases in which the introduction of application-awareness can produce benefits in the offline planning and the online operations of multi-layer transport networks.

Application-aware planning

The application-aware planning concept considers the optimum allocation of IP and Optical layer resources according to the requested connectivity specifications, imposed by the applications. These specifications are denoted as *service requirements* and include the set of networking characteristics that must be met through a joint IP-Optical multi-layer resource allocation process. The adopted process considers 4 general options for the allocation of resources to service requirements:

- Opt1 – the use of an existing single-hop IP link (over an established end-to-end optical path)
- Opt2 – the use of an existing multi-hop IP link (over multiple optical paths)
- Opt3 – the set-up of a single-hop IP link (by setting-up a direct end-to-end optical path)
- Opt4 – the set-up of a multi-hop IP link (by setting up an optical path that extends an already established IP link)

The concept is based on the work presented here[2]. However, the options are prioritised differently for each service requirement. So for

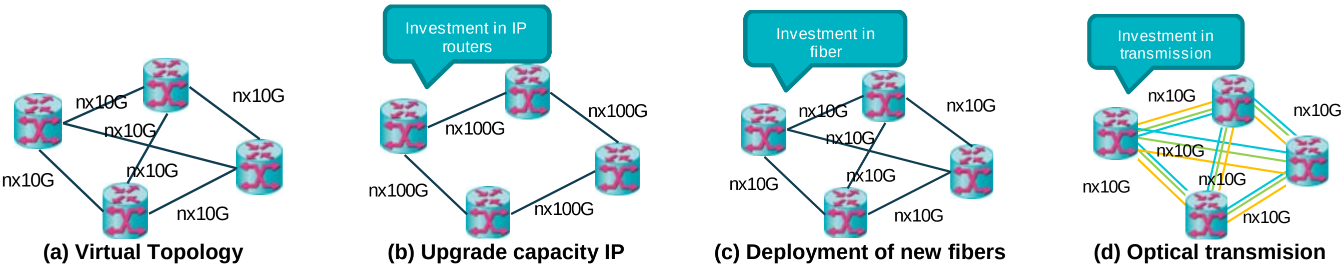

Fig. 1: Alternatives to increase network capacity

example, if a latency sensitive connection is requested, Opt1 is checked first and if no single-hop link is available then Opt3 is activated to establish one. On contrary, for a best effort type of request, Opt2 is examined first so as to reserve any direct connections for more demanding applications.

In addition to the overall application-driven allocation options, a set of allocation strategies is defined depending also on the type of the application use cases. A typical strategy is the one that separates between IP and Optical protection type of services in conjunction with other requirements like latency, high bandwidth or any special polices. Also, the k-shortest paths are calculated for each service requirement in the IP and Optical layer (typically with K=3 for national networks). This allows specific requests to be allocated over longer IP links or even optical paths, thus offering both load balancing options (by avoiding congested parts of the network) and optimised resource provisioning, for demanding large bandwidth or low latency service requirements. This strategy essentially activates Opt4 in the allocation process allowing remote optical paths to be set-up to accommodate longer shortest path connections in IP.

The interplay between the IP and Optical layer in the application aware allocation process and for different type of service requirements (i.e. application requests) is illustrated in the simple 4-node example of Fig. 2. The initial state is shown in Fig. 2a, assuming that 3 bidirectional optical links are already established between IP ports in nodes A-B, B-D and C-D. The arrival of a latency sensitive service requirement request between A and D (Fig. 2b) dictates the establishment of an optical path ABD between these nodes, thus generating a direct A-D IP link in the topology. On the contrary, for a best effort connection between A-C (Fig. 2c), the multi-hop IP route A-B-D-C is preferred as first option in order to reuse the already established optical links and IP ports. In Fig. 2d a latency sensitive request between A-C arrives that also requires optical protection. In this case, a direct A-C link is established while also an optical path through ABDC nodes is established thus generating the working and protection single hop links between A-C. If instead this request was a high bandwidth connection request between A and C, which could not be served on the existing capacity over the multi-hop A-B-D-C IP link (Fig. 2e), then again the establishment of a direct A-C link is requested. However, in order to provide IP level protection for this service, the already established A-B-D-C multi-hop link can be

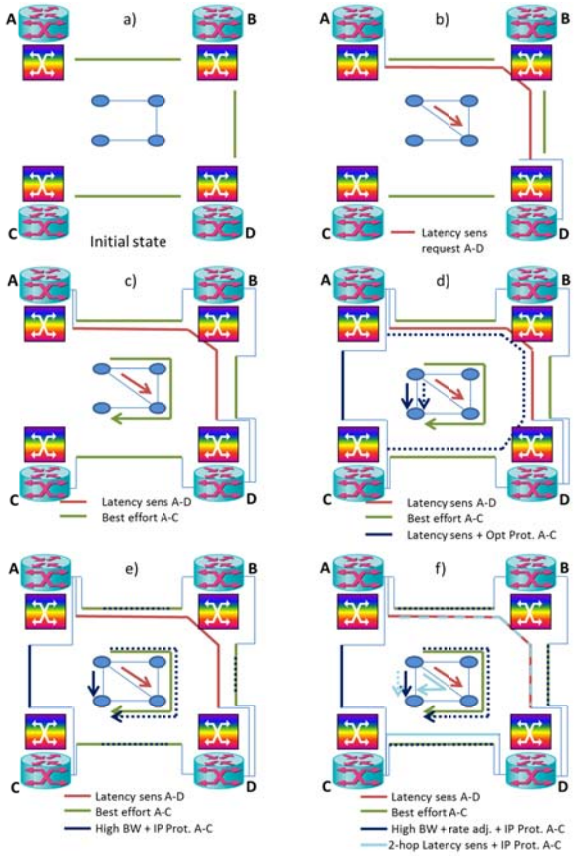

Fig. 2: Application aware planning example

considered that carries only best-effort service. In case of a A-C link failure, the high bandwidth IP protected service will switch to the A-B-D-C path while a new request for the best effort service will be generated and examined. A final example (Fig. 2f) considers a service requirement with moderate latency request of maximum 2 IP hops between A-C. In this case either the direct A-C or the 2-hop A-D-C IP link could be considered which have been established though over the AC and ABDC paths respectively in the optical layer (provided that in each case there is available capacity for this request). If, however, the high bandwidth A-C connection also has rate-adjustable features (e.g. supporting a datacentre migration request), then, the longer A-D-C IP route can be preferred, thus supporting the extra requested feature for the previous request.

Application-aware in-operation planning

In-operation network planning[3] refers to the ability of automatically reconfiguring or re-optimizing the network in near real time. In other words, it is an online process, unlike traditional planning processes that are performed offline over long time-scales. In this process, the knowledge of the network topology and status is exploited to reactively adapt to network changes, such as network failures (re-configuration), or to proactively re-arrange IP and optical network connections according to operators' policies (re-optimization). The main benefit on in-operation network planning is the ability to better tune the state of the network to the real short-term needs of its traffic, rather than aiming to be ready to serve a higher amount of traffic which may not materialize in several months.

Current approaches[3] are bandwidth-driven, i.e., they focus only on the capacity needed to support short-term traffic demands. This is sufficient to achieve CAPEX and OPEX savings in best-effort network scenarios, but also limiting in terms of the characteristics of the services being offered to the applications that ultimately drive the network traffic.

An example of the introduction of application-awareness in the in-operation planning process is shown in Fig. 3, where we consider the latency requirement only. The initial status of the network is presented in Fig. 3a, where 5x100Gb/s bidirectional lightpaths are already provisioned on the shortest optical paths: R1-R2 (red), R1-R3 (orange), R1-R4 (blue), R2-R3 (green), R2-R4 (grey). Let us assume that each optical link adds a latency value of 10 (the unit of measure is not relevant here) to the whole end-to-end delay, while the delay added by the routers is not considered for simplicity. Applications APP1 and APP2 have two different latency requirements (30 and 100, respectively) and their traffic is initially served over the R1-R2 (red) lightpath, so both experience a latency equal to 20 (their requirements are satisfied). APP1 and APP2 require 10Gbit/s each and are the only ones using that lightpath. Let us also assume that the capacity on the other established lightpaths is not saturated (e.g. 40% each, serving further applications' traffic). If, given its policies, the operator needs to save energy or to spare R1-R2 capacity for future high-priority demands, then, the red lightpath can be torn down and the traffic from APP1 and APP2 can be redirected through a different IP interface.

From an application-unaware perspective, at the IP level, both R1-R3-R2 (orange+green) and R1-R4-R2 (blue+grey) are viable options (equal cost). However, APP1 latency requirement (30) would not be satisfied on R1-R3-R2 (40), while the one of APP2 (100) would be satisfied on both. As shown in Fig. 3b, an application-aware mechanism should select R1-R4-R2 for APP1, thus ensuring a latency of 20, and in order to save the low latency path for future use, it could reroute the APP2 traffic on R1-R3-R2.

Given the above considerations, an application-aware approach to in-operation planning, where explicit application requirements are known to

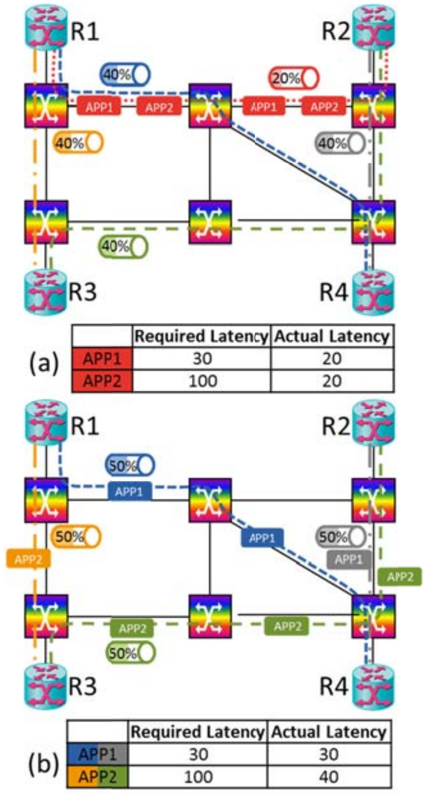

Fig. 3: Example of in-operation network re-optimization for a latency sensitive application.

the control plane of the network, can be useful to safeguard the service characteristics that are really important to applications, while still achieving significant savings for operators. The previous example can be extended to more complex cases where multiple requirements are considered or to reconfiguration scenarios.

Conclusions

Current approaches to the optimization of multi-layer networks are not sufficient to effectively satisfy the heterogeneous requirements of applications. This paper presents operator use cases that will benefit from the introduction of application awareness in the optimization of transport networks.

Acknowledgements

This research was partially funded by the European Commission H2020 Research and Innovation program, ACINO project, Grant Number 645127.

References

[1] O. Gerstel, et al. "Multi-Layer Capacity Planning for IP-Optical Networks", in IEEE Communications Magazine, Vol. 52, Issue. 1, pp. 44-51 (2014).

[2] S. Martinez, et al., "Assessing the Performance of Multi-Layer Path Computation Algorithms for different PCE Architectures," in OFC/NFOEC 2013, OSA Technical Digest, 2013, paper OW4G.1.

[3] L. Velasco et al., "In-operation network planning," in IEEE Communications Magazine, vol. 52, no. 1, p. 52 (2014).