Jean-François Labourdette
& Zhensheng Zhang
Editors

Role of Optical Network in Resilient IP Backbone Architecture

Jean-François Labourdette

In this column we discuss the current trend in IP backbone networks, which are poised to take over mission-critical services in addition to best-effort IP services as an integrated transport platform. We discuss several network architecture options that introduce an optical network layer, with the critical attribute being that they must be as resilient as the current SONET/SDH transport network.

Introduction

At the heart of IP backbone networks are the core IP routers with throughput of hundreds of Gb/s. These routers with interfaces operating at the per-wavelength bit rates (2.5 Gb/s and 10 Gb/s) are directly connected via point-to-point WDM optical-transport systems. For acceptable service reliability, even for best effort services, typically two interconnected routers are used for redundancy in each backbone node. This *de facto* IP backbone architecture is in tune with the current network environment in which each service — ATM, Frame Relay, Private Lines, Voice — is essentially delivered over its own overlay network. Given the high growth but low profit margin of IP services there is a challenge as well as an opportunity for service providers to deliver most of these services over a unified IP network to reduce capital and operations cost. Current IP networks, which have been perfectly suitable for best effort services, must however be enhanced to provide the same level of resiliency and service quality well established in the traditional enterprise service domains. The evolution to such an integrated network that is capable of fast, reliable service delivery will require three basic components. The emergence of multi-service data aware access and edge platforms with intelligent network functions (such as automatic topology discovery, routing, and signaling) will enable the integration of multiple services closer to the customer, thereby reducing the access cost and providing fast service delivery capability. The second component is the resilient high-capacity core IP router forming the backbone of the integrated service network. At the lowest layer is the third component, the optical switch, which interconnects the core routers via a switched optical network layer over WDM links. While new high-availability (so-called "non-stop routing") core routers provide service resiliency at the packet layer, the optical layer provides lowest cost and highest level of resiliency at the physical layer against catastrophic network events such as optical amplifier failure and fiber cuts.

Focusing on the second and third components of the integrated network evolution, there are four network architectures of interest. Current IP networks connect core routers directly over WDM in a dual-router configuration. This is architecture 1 and the Present Mode of Operations (PMO). Incorporating an optical core transport network leads to architecture 2. As the number of nodes in a network grows, the transit traffic in a node grows exponentially. Since optical switch port costs are only a fraction of the router port costs, the optical switches[1] allow significant cost reduction by siphoning off the transit traffic from the router layer to the optical layer. In addition, further resiliency is achieved by restoring high-capacity WDM links at the optical layer, building upon the deployment of WDM-based optical networks that support fast and capacity-efficient shared mesh restoration. Finally, an optical layer can allow the network to handle surges in IP traffic automatically, or to reroute trunks around a router failure. Common to architectures 1 and 2 is the presence of redundant routers per node, as in today's current IP backbone networks. To address this architectural aspect, and in conjunction with deploying an optical layer, we propose a new paradigm by which a single (or a few) redundant routers are deployed in the network and used to replace any failed router. This is architecture 3. Effectively, a single (or a few) shared redundant routers replace a second router in each office. This architecture leverages the rearrangeable optical layer to re-home access routers into the remote shared redundant router in case of a core router failure, as well as to appropriately re-trunk the spare router now in use to the rest of the IP network. We also consider with architecture 4 a single core router per node, along with the optical layer,

[1] *The optical switch is assumed to be an OEO switch, which terminates the optical signal and switches it in the electrical domain.*

which is becoming a feasible alternative as router availability increases.

We argue in this column that IP-over-OTN architectures are more economical and resilient than the current IP-over-WDM architecture, taking into account the redundant router configuration of current IP networks, or assuming that single routers are feasible thanks to their high availability. This assertion is based on the following:

1. Cheaper price per port on OXC than router for transit traffic
2. Optical shared mesh restoration for network failure faster than IP rerouting or MPLE LSP-based restoration in IP-over-WDM
3. Single high-availability router configuration, with or without shared spare router, cheaper than dedicated redundant router per office

Finally, deploying a reconfigurable optical layer for both IP and TDM traffic benefit from cross-sharing of protection bandwidth across both types of traffic and further minimizes the total network cost across both IP and non-IP services.

Architecture 1—IP-over-WDM Present Mode of Operations (PMO)

In architecture 1, which is the current mode of operation, there are two core routers in each node and the access routers are dual-homed to both core routers. The core routers are then directly connected with each other by point-to-point WDM links. The dual router architecture has been adopted because of the low reliability of routers. The traffic transiting through an office is terminated on one of the core routers in that office, and leaves from the same or the other core router towards the final destination[2]. IP regrooming thus takes place at every office as needed. The access ports (ports facing the access routers) on the core routers as well as the network ports (ports connected to the WDM equipment) would be typically operated at less than 50% utilization. This allows all the traffic to be rerouted by the access and/or core routers after any failure.

Layer 3 or layer 2 (MPLS) rerouting is used for service recovery from all types of failures. In case of core router **port failure**, the edge and/or core routers rely on layer 3 IP rerouting with OSPF/IS-IS routing table updates or layer 2 rerouting, e.g. MPLS to reroute the IP traffic around the failure. Such rerouting may take tens of seconds but has no significant network-wide impact. There is also no impact on traffic due to capacity constraints. In case of **core router failure**, access routers

[2] *While manual bypass of intermediate routers via patch panel is a possibility when traffic is small, it is not an operationally scalable solution as traffic increases, and we thus assume no manual bypass*

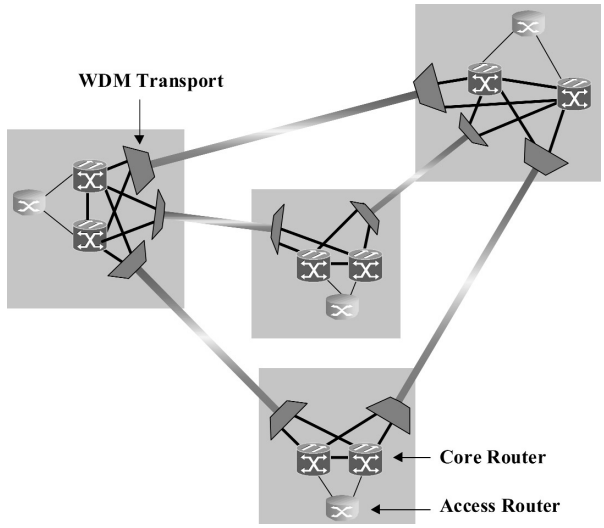

Figure 1: IP over WDM architecture.

use layer 3 IP rerouting with OSPF/IS-IS routing table updates or layer 2 rerouting, e.g. MPLS to reroute the IP traffic around the failed router. Again, such rerouting may take tens of seconds and have some moderate network-wide impact. Without complex traffic engineering, the network may incur packet loss. In case of **transport link failure**, the edge and/or core routers use layer 3 IP rerouting with OSPF or IS-IS routing table update or layer 2 rerouting, e.g. MPLS to reroute the traffic around the failure. It may take tens of seconds, and can have a huge network-wide impact. For example, the network may encounter routing table non-convergence leading to possible network-wide instability. In spite of enough capacity left in the network, IP routing may not be able to use it leading to potentially severe congestion.

There are two fundamental problems with this architecture. First, it is the most expensive and least scalable. As the traffic and the number of nodes in a network grow, the traffic transiting intermediate routers grows exponentially. Since router port costs are high (three to four times that of optical switch port) and most router ports are consumed to simply route transit traffic, the network cost also grows exponentially. Second, while this IP backbone architecture may be suitable for Internet traffic, it is not so for delivering mission-critical services. Service restoration by rerouting at Layer 3 for catastrophic failures is simply not amenable to such services. There is no bandwidth guarantee in the rerouted paths, routing table updates could take minutes and a huge network-wide routing table update could lead to network instability. Congestion collapse is likely to occur when backbone routers are overwhelmed due to multi-wavelength link failure. Arguably restoration using layer 2 rerouting such as MPLS may provide better restoration performance than Layer 3. However, it is still not suitable for mission-critical services because it is almost an impossible task to do traffic engineering for

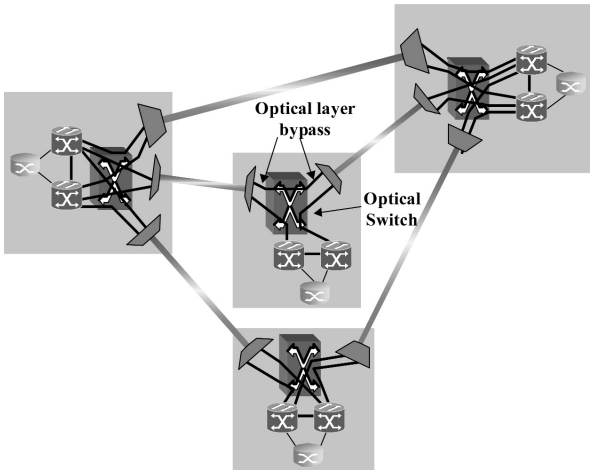

Figure 2: IP over OTN architecture with dual router.

guaranteed bandwidth requirements via alternate routes in case of a catastrophic failure for thousands of traffic flows. MPLS-based restoration against high-capacity link failures is an option being explored but still faces many difficulties. We believe that it is imperative to enhance the overall stability and reliability of the IP backbone network before the enterprise customers would agree to transport their mission-critical services over a unified IP backbone network. As more robust and high-availability routers are becoming available, the weakest link in the network resiliency will be congestion failure caused by high-capacity link failures, which this architecture cannot address.

Architecture 2—Dual-Router Architecture with Optical Network

In the second architecture shown in Figure 2, we still have two routers per node but the routers are connected via optical switches. The optical switches provide low-cost bypass for transit traffic. They also provide fast restoration (~100 msec) against catastrophic network failure using shared backup capacity, with guaranteed availability in case of single failure, in the optical layer. The router layer thus remains completely impervious to such catastrophic network failures. The dual router configuration is used to provide resiliency from router failures as in architecture 1. The traffic is equally divided between the two routers, with the access ports on the core routers operated at less than 50% utilization. The network ports on the core routers would typically be 2.5 Gbps ports that are directly connected to the optical switch and operated at up to 75% utilization. The network ports on the OXC are connected to the WDM systems at 10 Gbps, with the OXC providing the grooming of 2.5 Gbps ports facing the routers into 10 Gbps ports facing the WDM systems/network. In the IP-over-OTN architecture, transit traffic goes mostly through the OXC and not the core routers, unless it is determined that terminating at the core router for re-grooming the IP traffic is beneficial and economical. Predominantly transiting through the OXC rather than the router allows significant reduction of the cost of the network due to the much cheaper price per port of OXC equipment compared to router equipment.

In case of core router **port failure** or **core router failure**, combination of edge and/or core router layer 3 IP rerouting or MPLS layer 2 rerouting would take place as in architecture 1, with analogous network and traffic impact. However, in case of **transport link failure**, the optical switch restores all links on the route using backup capacity shared among all services. The restoration takes place in ~100 msec, before any attempt to do IP-level rerouting or layer 2 MPLS rerouting, therefore causing no impact on the router network, and on the traffic. The bandwidth and traffic performance are guaranteed and not impacted.

Architecture 3—Single Router with Optical Network and Shared Spare Router Strategy

In architecture 3 we have assumed that the router reliability is still not at par with that of the traditional carrier class systems. In spite of this assumption the same level of reliability can be achieved with just one router per node and a few network-wide shared backup routers as shown in Figure 3. In this configuration if a router fails, the optical switches reconnect the associated access routers to the shared backup router. This architecture provides a lower cost and more robust backbone network than architectures 1 and 2 that is suitable for mission-critical as well as best effort services. The access ports (towards the access routers) on the core router could be utilized at up to 75%. The network ports on the core routers (towards the OXC) are assumed to be 2.5 Gbps ports that are directly connected to the optical switch

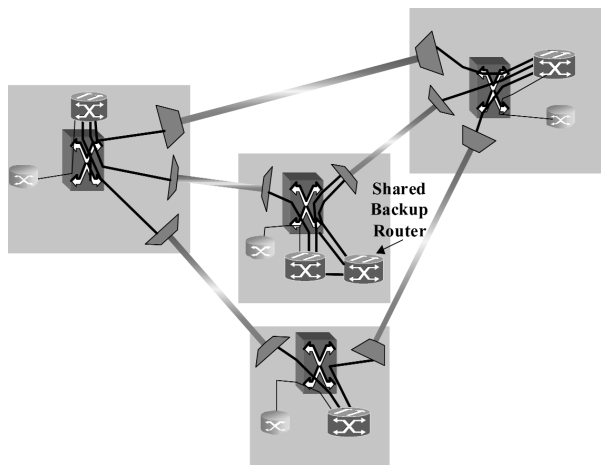

Figure 3: IP over OTN architecture with shared spare router.

and operated at up to 75% utilization. The network ports on the OXC are connected to the WDM systems at 10 Gbps, with the OXC providing the grooming of 2.5 Gbps ports facing the routers into 10 Gbps ports facing the WDM systems/network. The access routers are connected to the core routers through the OXC so that the access lines can be re-homed in an automated way to a shared spare router following a core router failure. The network trunking used to handle the re-homing as well as the trunking from the selected shared spare router to the rest of the network is a combination of shared mesh protection trunking, trunking capacity left available from the failed router, as well as any spare trunking available.

In case of core router **port failure**, combination of edge and/or core router layer 3 IP rerouting or MPLS layer 2 rerouting would take place as in architectures 1 and 2, with analogous network and traffic impact. In case of **core router failure**, after the failure is detected, the access routers are re-homed to one of the spare shared core routers using shared backup capacity. The access routers use layer 3 IP rerouting with OSPF/IS-IS routing table updates or layer 2 rerouting, e.g. MPLS to reroute the IP traffic through the spare shared router. Again, such routing table updates may take tens of seconds. After that, there is no service degradation, and no impact on IP-based QoS. In case of **transport link failure**, optical layer restoration is performed before any layer 3 IP rerouting or layer 2 MPLS rerouting as in architecture 2, thereby causing no impact on the router network and on the traffic.

Architecture 4—Single Router Architecture with Optical Network

In architecture 4 shown in Figure 4, we have assumed that the router reliability is at par with that of the traditional carrier class systems. With this assumption the same level of reliability can be achieved with just one router per node, and without shared spare routers at remote nodes. This architecture provides the lowest cost and most robust backbone network suitable for mission-critical as well as best-effort services. Now, the access ports (towards the access routers) on the core router are utilized at up to 75% as well as the network ports (towards the OXC). The network ports on the core routers are assumed to be 2.5 Gbps ports that are directly connected to the optical switch and operated at 75% utilization. The network ports on the OXC are connected to the WDM systems at 10 Gbps, with the OXC providing the grooming of 2.5 Gbps ports facing the routers into 10 Gbps ports facing the WDM systems/network. The access routers are directly connected to the core routers, but not through the OXC as was the case in architecture 3.

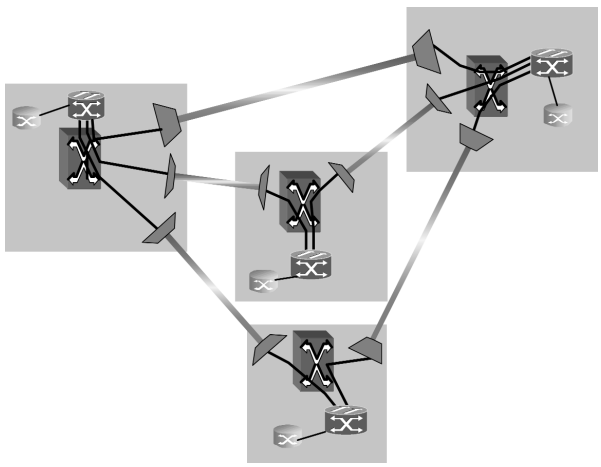

Figure 4: IP over OTN architecture with single router.

In case of core router **port failure**, combination of edge and/or core router layer 3 IP rerouting or MPLS layer 2 rerouting would take place as in architectures 1, 2, and 3, with analogous network and traffic impact. In case of **transport link failure**, optical layer restoration is again performed before any layer 3 IP rerouting or layer 2 MPLS rerouting as in architectures 2 and 3, thereby causing no impact on the router network, and on the traffic.

Conclusion

In this column, we have discussed four different architectures: (1) the PMO where routers are connected directly to WDM systems; (2) an architecture where routers in a dual configuration are connected over an optical transport network; (3) an architecture where single routers are connected over an optical transport network with shared redundant routers providing redundancy for router failure through rehoming of access routers and reconfiguration of the optical layer; and (4) an architecture where single carrier-class routers are directly connected over optical transport networks. We have made the following observations:

- Transit traffic grows much faster than the terminating traffic in a network as the network size, as well as the traffic, grows. Architectures 2, 3, and 4 would provide cost efficiency over the current architecture 1 by siphoning off the transit traffic from the router layer into the optical layer and additional cost savings and higher reliability by providing network restoration against catastrophic failures. An additional byproduct of all three IP-over-OTN architectures is to provide better scalability compared to the present mode of operations of architecture 1.
- The switched optical layer with fast shared mesh restoration completely shields the router layer from catastrophic network failures and thus provides highest level of reliability at lowest cost for mission-critical services as well as best-effort services.

- With shared backup router architecture the router layer resiliency is achieved even with the current router technology. The shared backup router architecture is further simplified with the availability of non-stop router technology by eliminating the dual router at each node.

An interesting observation is the relationship between the four architectures. There is a clear evolution path from architecture 1 to 2 by introducing an optical layer capable of fast shared-mesh restoration and moving transit traffic off the routers and relying on optical layer restoration for network failures. From architecture 2, one would evolve towards architecture 3 by relying on a few shared spare routers rather than dual routers per office to address the risk of core router failure. This will happen as routers become more reliable. Eventually, as routers become even more robust to the point of being fully carrier class, the network can evolve from architecture 3 to architecture 4 by further ending reliance on shared spare routers.

Acknowledgments

I would like to acknowledge the contribution of Sid Chaudhuri and Eric Bouillet to this column and to the work on which it is based.