Design of Reliable Communication Networks (DRCN) 2003, Banff, Alberta, Canada, October 19-22, 2003

BEST PAPER AWARD

Role of Optical Network and Spare Router Strategy in Resilient IP Backbone Architecture

Jean-François Labourdette[#], *Senior Member, IEEE*, Eric Bouillet[#], *Member, IEEE*,
Sid Chaudhuri[#], *Member, IEEE*
[#]Tellium, 2 Crescent Place, P.O. Box 901, Oceanport, NJ 07757-0901, USA
E-mail: jlabourdette@tellium.com, {sc, ebouillet}@tellium.com

***Abstract* – 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 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. It has been established that the majority of the traffic in a node is transit traffic and a significant cost reduction can be achieved by siphoning off the transit traffic from the IP layer into the optical layer. In this paper we discuss the current trend in the IP backbone network which is poised to take over other premium services, in addition to best effort IP services, as an integrated transport platform. We discuss several network architecture options with the critical attribute being that it must be as resilient as the current SONET transport network. We propose an innovative architecture option in which a resilient network is built with current router technology. In another option we assume that the router layer can be as resilient as the current SONET layer with the emerging resilient router technology. We perform an economic evaluation and discuss the reliability of these network architectures.**

***Index Terms* – optical network, IP network.**

I. 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 (2.5Gb/s and 10 Gb/s) are directly connected via point-to-point WDM optical-transport systems. 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[1] but low profit margin for IP services there is a challenge as well as an opportunity for a service provider 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, however, must be enhanced to provide the same level of resiliency and service quality well established in the traditional service domains. The evolution to such a reliable and integrated network that is capable of fast, reliable service delivery requires 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 reducing the access cost and providing fast service delivery capability. The second component is the resilient high-capacity core IP routers 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 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.

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 allow significant cost reduction by siphoning off the transit traffic from the router layer to the Optical Transport Network (OTN) layer [2]. In this paper we show that further resiliency is achieved by restoring high-capacity WDM links at the optical layer, building upon the deployment of WDM-based optical networks [3] that support fast and capacity-efficient shared mesh restoration [4,5,6]. The value of an optical layer has been previously addressed [7,8]. For example, previous work has shown how an optical layer allows the network to handle surges in IP traffic automatically [9], or to reroute trunks around a router failure [10].

Focusing on the second and third components of the integrated network evolution, we have analyzed four network architectures. In Architecture 1, which is the current mode of operation, there are two core routers in each node and the access routers are connected 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[2]. Layer 3 or layer 2 (MPLS) rerouting is used for service recovery from all types of failures.

[1] IP traffic has been growing 100% year over year except in 1995-1996 when it grew 100% every 3-4 months [1]. Every year, the amount of new IP traffic is as much as all the existing traffic up to that year.

[2] See [11] for a report on router reliability and on-going work by vendors to improve router robustness.

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 (3 to 4 times that of optical switch ports) 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 public and commercial services. Service restoration by Layer 3 rerouting 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. A recent experimental case study [12] on Internet stability using operational failure logs over a period of twelve months shows that the major catastrophic Internet failures stem from congestion collapse. 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 guaranteed bandwidth requirements via alternate routes in case of a catastrophic failure affecting thousands of traffic flows. Even with MPLS based restoration against high-capacity link failures without bandwidth guarantee it is likely that protocol messages such as KeepAlive may be delayed or lost for significant duration causing other links and possibly the network to collapse. 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.

Incorporating an optical core transport network leads to architecture 2. In this second architecture we thus still have two routers per node but the routers are connected via optical switches. The optical switches provide low-cost bypass of transit traffic and provide restoration against catastrophic network failure using shared backup capacity in the optical layer [5,6]. The router layer remains completely impervious to such catastrophic network failures by instead relying on fast shared mesh restoration (~100 msec) [13] with guaranteed bandwidth at the optical layer. The dual router configuration is used to provide resiliency from router failures as in Architecture 1. While this architecture provides a low-cost and resilient integrated service backbone, further cost reduction and enhanced reliability can be achieved. And this leads to Architectures 3 and 4.

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. 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. With the emergence of high reliability routers with average downtime of only 0.5 minutes per year this architecture can be further simplified by eliminating the shared backup routers as well without sacrificing overall service reliability, which leads us to our last architecture. In architecture 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. This architecture provides the lowest cost and most robust backbone network suitable for mission critical as well as best effort services.

We used a network model of 39 nodes to analyze these four architectures. We assumed a linearly distributed traffic demand with 25% of the node pairs having 500 Mb/s and a small percentage of the node pairs having 2.5 Gb/s. We have allocated shared backup WDM channels for restoration against network failures. Then using typical optical switches, routers and WDM cost we see that Architecture 3 saves 34% capital cost over the current mode of operation. We should note that there is an opportunity to reduce cost even further by segregating best effort services and using the shared backup channels for those services. Based on the results and argumentation presented in this paper we draw the following main conclusions:

- Transit traffic grows much faster than the terminating traffic in a network as the network size as well as the traffic grows. IP-over-OTN architectures that siphon off the transit traffic from the higher layer and routes it in the less expensive and more reliable optical layer provides the lowest cost and a more scalable network for integrated services backbone.
- 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.
- The switched optical layer enables the backbone network resiliency required for mission critical services on an integrated IP backbone. The router layer resiliency is achieved even with the current router technology using the innovative shared spare router strategy proposed in Architecture 3. The shared spare router architecture is further simplified with the availability of non-stop router technology by eliminating dual routers at each node.

II. ARCHITECTURAL COMPARISON

Current IP networks connect core routers directly over WDM. It was shown in earlier work that incorporating an optical core transport network was economical [2]. However, one complicating factor is the presence of redundant routers per node in today's current IP backbone networks. To address this architectural aspect, we propose a new paradigm by which a single (or a few) redundant router is deployed in the network and is used to replace any failed router. Effectively, a single (or a few) shared redundant router replaces a redundant router in each office. Such an architecture requires a 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 and analyze an architecture with a single core router per node, which is becoming a feasible alternative as router availability increases.

We show in this paper that IP-over-OTN architectures are more economical and resilient than the current IP-over-WDM architecture, extending the work in [2] by taking into account the redundant router configuration of current IP networks, or assuming that single router are feasible thanks to their high availability. This conclusion results from the following:

1. Cheaper price per port on OXC than router for transit traffic
2. Optical shared mesh restoration faster than IP rerouting or LSP-based restoration in IP-over-WDM [13]
3. Sharing of spare router cheaper than dedicated redundant router per office

In addition, we describe how relying on a reconfigurable optical network layer for IP enhances the restoration time and availability of IP services, improving on the behavior of the current architecture.

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.

A. *Present Mode of Operations (PMO)*

In Architecture 1 shown in Figure 1, which is the current mode of operation, there are two core routers in each node and the access routers are connected to both core routers. The core routers are then directly connected with each other by point-to-point WDM links. The dual router configuration has been adopted for redundancy because of the historical low reliability of routers and the common occurrence of router failure. Layer 3 or layer 2 (MPLS) rerouting is used for service recovery from all types of failures. In this architecture, 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. IP regrooming thus takes place at every office as needed. 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 do not assume any manual bypass in our architecture and analysis.

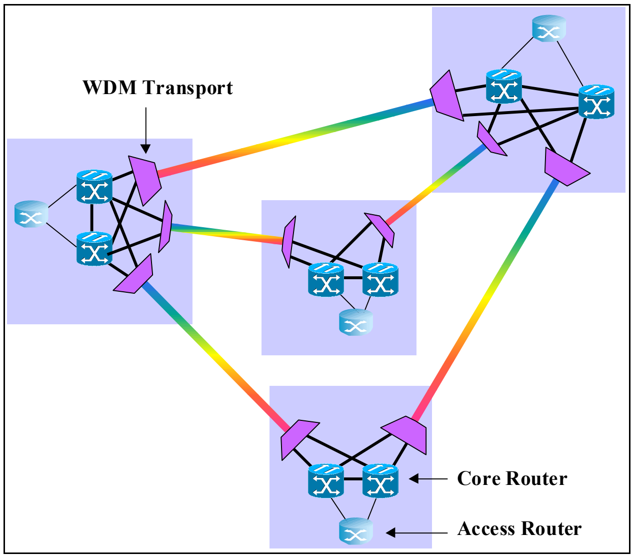


Figure 1 - IP over WDM Architecture.

In architecture 1, the access ports (ports facing the access routers) on the core routers as well as the network ports (ports connected to the WDM equipment) are assumed to operate at less than 50% utilization. This allows all the traffic to be rerouted by the access and core routers after any failure.

We now describe the restoration mode for the following three main failure scenarios. 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 10s 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 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 10s 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 10s 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.

B. *Dual-router architecture with optical network*

In this second architecture we have still two routers per node but the routers are connected via optical switches as shown in Figure 2. The optical switches provide low-cost bypass of 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.

Earlier work [2] showed the benefits of deploying a reconfigurable optical layer, in an architecture where a single router per node was connected to optical switches, with the transit traffic going through intermediate optical switches rather than core routers. This earlier work didn't address dual-router configuration currently used for redundancy [12].

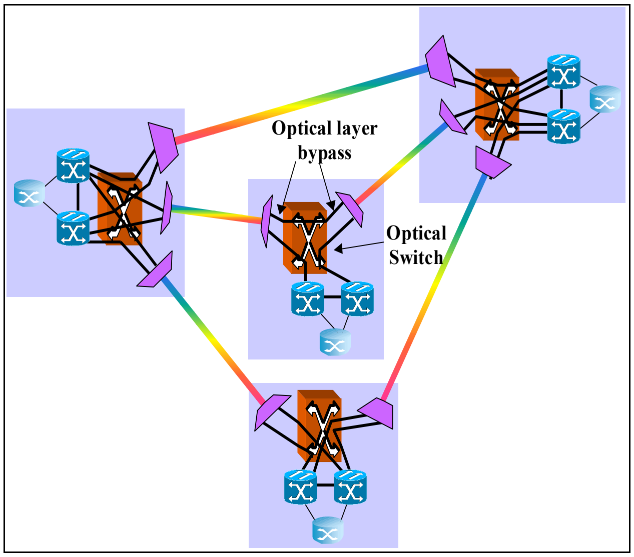


Figure 2 – IP over OTN Architecture with dual router.

An architecture diagram for IP-over-OTN with dual-router configuration is shown in Figure 2, with at least two core routers per node. 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 are assumed to be 2.5Gbps 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 10Gbps, with the OXC providing the grooming of 2.5Gbps ports facing the routers into 10Gbps 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 to significantly reduce the cost of the network due to the much cheaper price per port of OXC equipment compared to router equipment.

We now describe the restoration mode for the three main failure scenarios. 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 10s 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 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 10s 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 optical switch restores all links on the route using shared backup capacity among all services. The restoration takes place in ~ 100 msec, before any attempt to do IP-level rerouting, therefore causing no impact on the router network, and on the traffic. The bandwidth and traffic performance are guaranteed and not impacted.

An alternative architecture is to use a single backbone router per site with or without redundancy in case of router failure provided by a remote backbone router shared among all the routers, as described below. In this architecture, the OTN provides the support for re-homing access routers to the spare shared router as well as any trunking configurations required between the core routers, including the spare shared core router.

C. *Single router architecture with optical network and shared spare router strategy*

In this third architecture shown in Figure 3, we have assumed that the router reliability is still not at per 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 one or two network wide shared backup routers. 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 suitable for mission critical as well as best effort services than architectures 1 and 2. With the emergence of high reliability routers with average downtime of only 0.5 minutes per year this architecture can be further simplified by eliminating the shared backup routers as well without sacrificing overall service reliability as shown in architecture 4 in the next section.

As shown in Figure 3, we have a single router configuration with spare router shared at a remote node. Now, the access ports (towards the access routers) on the core router are utilized at up to 80% as well as the network ports (towards the OXC). The network ports on the core routers are assumed to be 2.5Gbps ports that are directly connected to the optical switch and operated at up to 80% utilization. The network ports on the OXC are connected to the WDM systems at 10Gbps, with the OXC providing the grooming of 2.5Gbps ports facing the routers into 10Gbps 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.

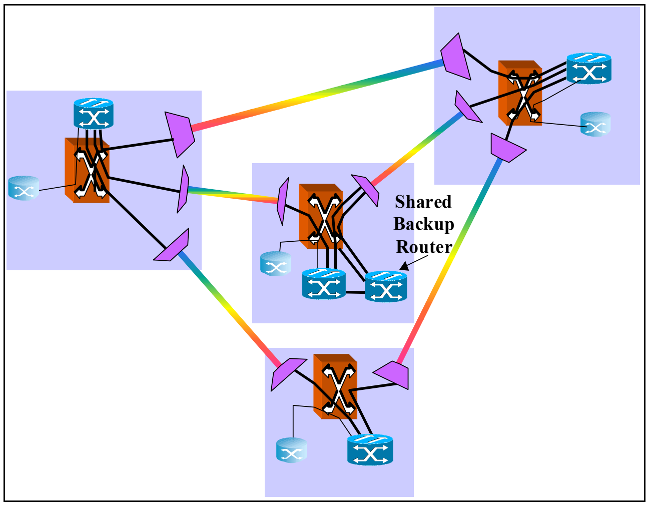


Figure 3 – IP over OTN Architecture with shared back-up router.

We now describe the restoration mode for the three main failure scenarios. 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 10s 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**, the failure is detected (requires a new capability - more on this later), and the access routers are re-homed to one of the spare shared core routers (as shown in Figure 3) using backup shared mesh capacity. The access routers in one office as shown in Figure 3 use layer 3 IP rerouting with OSPF/IS-IS routing table pdates or layer 2 rerouting e.g. MPLS to reroute the IP traffic through the spare shared router router. Again, such routing table updates may take 10s of seconds. After that, there is no service degradation, and no impact on IP-based QoS. In case of **transport link failure**, the optical switch restores all links on the route using shared backup capacity among all services. The restoration takes place in ~ 100 msec, before any attempt to do IP-level rerouting, therefore causing no impact on the router network, and on the traffic. The bandwidth and traffic performance are guaranteed and not impacted.

There are two possible approaches to managing the reconfiguration in architecture 3 (shared spare router) in case of core router failure. In a centralized approach, a traffic and bandwidth management system keeps track of access and core router trunking over the OTN. Such a bandwidth manager would need the ability to be informed of a core router failure, possibly from the router element management system, or to infer such a failure, for example from multiple signal failure received by the OXC and communicated through the OXC element management system. Upon recognition of the failure, the bandwidth manager would rehome the access router to the shared router, using either a pre-planned procedure, or calculating in real-time the best re-homing arrangement, given available capacity (including shared mesh back-up capacity). In a distributed approach, UNI signaling [14] between the router and the OXC could be used, along with a distributed control plane on the OXC, as a mechanism to trigger and carry out the trunking reconfiguration required to rehome the affected access router to a shared spare router.

D. Single router architecture with optical network

In this fourth architecture shown in Figure 4, we have assumed that the router reliability is at par with that of the traditional carrier class systems. The required level of reliability can thus be achieved with just one router per node. This architecture provides the lowest cost and most robust backbone network suitable for mission critical as well as best effort services.

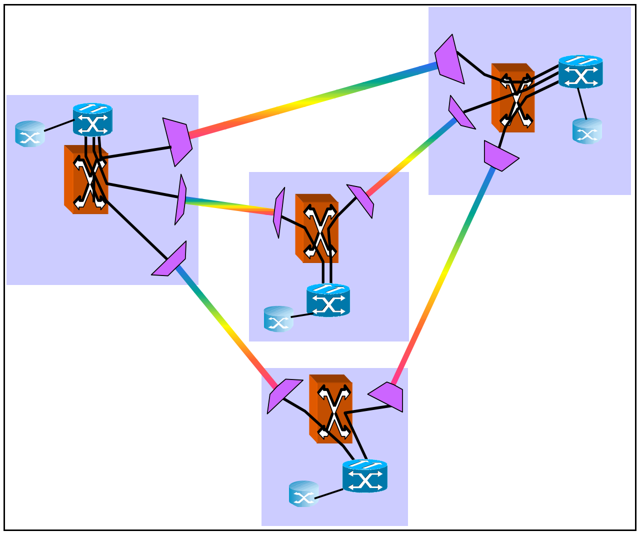


Figure 4 – IP over OTN Architecture with single router.

As shown in Figure 4, we have a single router configuration WITHOUT shared spare router at a remote node. Now, the access ports (towards the access routers) on the core router are utilized at up to 80%. The network ports on the core routers (towards the OXC) are assumed to be 2.5Gbps ports that are directly connected to the optical switch and operated at up to 80% utilization. The network ports on the OXC are connected to the WDM systems at 10Gbps, with the OXC providing the grooming of 2.5Gbps ports facing the routers into 10Gbps ports facing the WDM systems/network. The access routers are directly connected to the core routers, not through the OXC as was the case in architecture 3.

We now describe the restoration mode for the three main failure scenarios. 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 10s of seconds but has no significant network wide impact. There is also no impact on traffic due to capacity constraints. In case of **transport link failure**, the optical switch restores all links on the route using shared backup capacity among all services. The restoration takes place in ~ 100 msec, before any attempt to do IP-level rerouting, therefore causing no impact on the router network, and on the traffic. The bandwidth and traffic performance are guaranteed and not impacted.

III. NETWORK ANALYSIS AND DESIGN

A. *Network model*

We used a network model of 39 nodes (Figure 5) to analyze these three architectures. We assumed pair-wise traffic demand linearly distributed in the range 500 Mb/s to 2.5 Gb/s across all node pairs (see Figure 6).

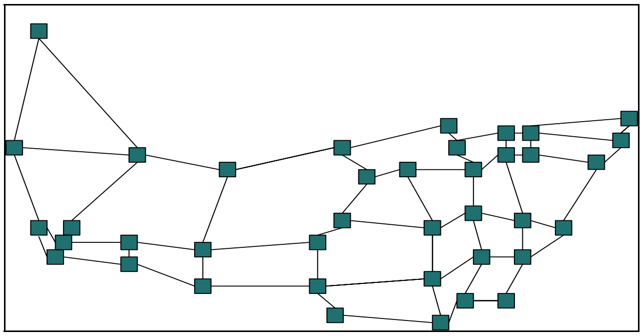


Figure 5 – Network Topology.

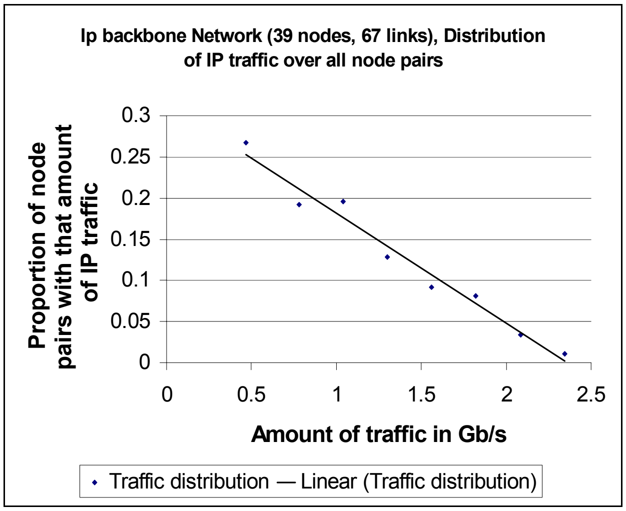
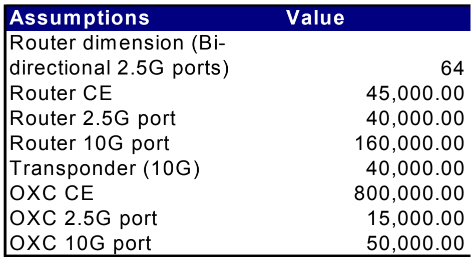


Figure 6 – Traffic Distribution.

The equipment configuration and cost model for the router and OXC is shown in Figure 7. Generic prices of routers, optical cross connects, and lightwave systems are considered. Note that the results are insensitive to limited variations of those prices. Access routers are not considered in the cost analysis because the access router port counts are the same in all architectures.



Assumptions	Value
Router dimension (Bi-directional 2.5G ports)	64
Router CE	45,000.00
Router 2.5G port	40,000.00
Router 10G port	160,000.00
Transponder (10G)	40,000.00
OXC CE	800,000.00
OXC 2.5G port	15,000.00
OXC 10G port	50,000.00

Figure 7 – Equipment Configuration and Cost Model.

B. *Network design procedures*

In this section we describe the procedure used to provide a cost comparison of the four architectures discussed earlier. All the scenarios are constructed from the network model described in Figure 5. The 39 nodes of the model represent the Central Offices (CO), each of which may consist of one or more routers, and if applicable an optical cross-connect (OXC). We assume that the traffic processed within each CO is "fluid" and is uniformly distributed across its routers. In the following the rate of the access ports on the client side is in units of 2.5Gb/s, and the rate of the network ports between pairs of CO is in units of 10Gb/s.

1. IP over WDM – dual routers

For every central office (CO) ***k***, we compute the aggregated IP traffic $\boldsymbol{T_{access\text{-}k}}$ in units of Gbits per seconds terminating in this CO. Assuming that the sum is uniformly distributed across 2.5Gb/s ports without exceeding α_1=50% of their capacity, the number of ports required to access the routers is:

$$R_{access-ports} = \sum_{k \in Nodes} \left\lceil \frac{T_{access-k}}{2.5\alpha_1} \right\rceil$$

Next, we route this traffic using min-hop paths (one that minimizes the number of intermediate COs), and for each link (u,v) connecting two COs, we compute the aggregated load $\boldsymbol{T_{link\text{-}uv}}$ carried on that link. The number of network 10Gb/s ports required to carry this traffic without exceeding α_1=50% of their capacity is thus:

$$R_{network-ports} = \sum_{(u,v) \in Links} \left\lceil \frac{T_{link-uv}}{10\alpha_1} \right\rceil$$

Finally, for every CO ***k***, we measure according to the routing above, the transit traffic $\boldsymbol{T_{transit\text{-}k}}$ through that CO. Assuming that 50% of this traffic traverses a single router, we compute the number of inter-router 2.5Gb/s ports required to transport the remaining β=50% of traffic traversing two routers:

$$R_{inter-router-ports} = \sum_{k \in Nodes} \left\lceil 2\frac{T_{transit-k}}{2.5\alpha_1}\beta \right\rceil$$

There are at least two routers per CO, and possibly more in order to accommodate all the access ports, the network ports, and the inter-router ports within that CO. More specifically, every CO ***k*** requires $\boldsymbol{N_k}$ routers, where $\boldsymbol{N_k}$ is given by:

$$N_k = \max\left(2, \left\lceil \frac{1}{64}\left(\left\lceil \frac{T_{access-k}}{2.5\alpha_1} \right\rceil + \left\lceil \frac{2\beta T_{transit-k}}{2.5\alpha_1} \right\rceil + 4\sum_{v \in Nodes} \left\lceil \frac{T_{link-kv}}{10\alpha_1} \right\rceil \right)\right\rceil\right)$$

From these quantities, we determine the number of routers needed assuming routers of size 64 2.5Gb/s ports, the total number of 2.5Gb/s and 10Gb/s router ports, the number of 10Gb/s transponders, and derive from it the total cost of the architecture.

2. IP over OTN – dual routers

This scenario requires the same number $\boldsymbol{R_{access\text{-}ports}}$ of access-ports as in the IP over WDM case. This traffic is carried across an OTN network that consists of OXC capable to switch shared mesh-protected lightpaths at rates of 2.5Gb/s and 10Gb/s. This solution requires some preprocessing in which the IP traffic is packed at these rates. For every pair of CO we replace as much traffic as possible by 2.5Gb/s lightpaths, while using between 67% and 75% of their bandwidth. Using less than 67% of an end-to-end lightpath is cost-inefficient and avoided. Instead, the residual traffic that must be left-over in order to satisfy this range of utilization is aggregated and packed into "single-hop" lightpaths, still without exceeding 75% percent of their bandwidth. Based on the total number of lightpaths $\boldsymbol{N_{lightpaths}}$, the number of 2.5Gb/s ports between routers and OXCs is:

$$R_{router-to-OXC-ports} = R_{OXC-to-router-ports} = 2N_{lightpaths}$$

The lightpaths are routed in the OTN network and shared mesh protected [5,6]. From this we determine the required number $\boldsymbol{R_{network\text{-}ports}}$ of 10Gb/s ports (and transponders). For each CO ***k*** we also measure the amount of traffic $\boldsymbol{T_{transit\text{-}k}}$ that traverses a router but does not terminate in the CO. Assuming that β=50% of this transit traffic traverses two routers within the CO, and that the ports are used at α_2=75% of their capacity, we compute the required number of 2.5Gb/s inter-router ports:

$$R_{inter-router-ports} = \sum_{k \in Nodes} \left\lceil 2\frac{T_{transit-k}}{2.5\alpha_2}\beta \right\rceil$$

Our results indicate that a single 128 ports OXC is sufficient in each CO to accommodate the prescribed traffic, however the number of routers varies and must be calculated for each CO. Given $\boldsymbol{L_k}$, the number of packed lightpaths terminating at CO ***k***, the number of routers in that CO is given by:

$$N_k = \max\left(2, \left\lceil \frac{1}{64}\left(\left\lceil \frac{T_{access-k}}{2.5\alpha_1} \right\rceil + \left\lceil \frac{2\beta T_{transit-k}}{2.5\alpha_2} \right\rceil + \left\lceil \frac{L_k}{2.5\alpha_2} \right\rceil \right)\right\rceil\right)$$

3. IP over OTN – shared routers

In this architecture each client accesses a single router through the OXC in its adjacent CO, and may use up to α_3=80% of the access ports. According to this, the number of access ports on the router side is:

$$R_{router-access-ports} = \sum_{k \in Nodes} \left\lceil \frac{T_{access-k}}{2.5\alpha_3} \right\rceil$$

And the number of OXC ports to access the routers from the client side is:

$$R_{oxc-access-ports} = 2R_{router-access-ports}$$

Furthermore, we dedicate two shared protection routers at strategic locations in the network, and a number of ports between the shared protection routers and the adjacent OXC which is at least as large as the maximum number of access ports of any single router in any CO. Note that we could also reserve the shared ports on existing routers distributed in the network. We then pack the demands as described in sub-section 2 above for the dual routers, with the exception that we now allow to use up to α_3=80% of the lightpath bandwidth. The number of 2.5Gb/s ports between routers and OXC, and 10Gb/s ports on the network side is then determined as before. We compute the number of routers per CO as follows:

$$N_k = \left\lceil \frac{1}{64}\left(\left\lceil \frac{T_{access-k}}{2.5\alpha_3} \right\rceil + \left\lceil \frac{L_k}{2.5\alpha_3} \right\rceil\right) \right\rceil$$

In CO with two routers or more due to the router size being exceeded, we measure the transit traffic $\boldsymbol{T_{transit-k}}$ that is re-groomed within the CO and derive from it the number of 2.5Gb/s inter-router ports required for the traffic traversing two routers:

$$R_{inter-router-ports} = \sum_{k \in Nodes} \left\lceil 2\frac{T_{transit-k}}{2.5\alpha_3}\beta \right\rceil$$

If necessary we increase $\boldsymbol{N_k}$ to accommodate the inter-router ports. As a consistency check, we then verify that in case of router failure, there is enough shared protection capacity, normally reserved to restore failures in the OTN network, to reroute the traffic from the corresponding access router through one or both of the shared routers.

4. IP over OTN – single router

This scenario requires the same number of 2.5Gb/s ports between the routers and the OXC, and number of 10Gb/s ports on the network side, as compared to the IP over OTN with single router and shared spare routers scenario described in section 2.3.3. The main difference is that the client accesses the network through a single router and may use up to α_3=80% of the 2.5Gb/s router access ports capacity instead of α_1=50%. The number of access router ports and the number of routers per CO are thus respectively:

$$R_{access-ports} = \sum_{k \in Nodes} \left\lceil \frac{T_{access-k}}{2.5\alpha_3} \right\rceil$$

And

$$N_k = \left\lceil \frac{1}{64}\left(\left\lceil \frac{T_{access-k}}{2.5\alpha_3} \right\rceil + \left\lceil \frac{L_k}{2.5\alpha_3} \right\rceil\right) \right\rceil$$

In CO with two routers or more we measure the transit traffic $\boldsymbol{T_{transit-k}}$ that is re-groomed within the CO and derive from it the number of 2.5Gb/s inter-router ports required to accommodate the traffic traversing two routers:

$$R_{inter-router-ports} = \sum_{k \in Nodes} \left\lceil 2\frac{T_{transit-k}}{2.5\alpha_3}\beta \right\rceil$$

If necessary we increase $\boldsymbol{N_k}$ to accommodate the inter-router ports

IV. Results and Discussion

Figure 8 and Figure 9 show side-by-side comparisons of respectively the port requirements, in unit of 2.5Gb/s, and the cost details, for the four scenarios.

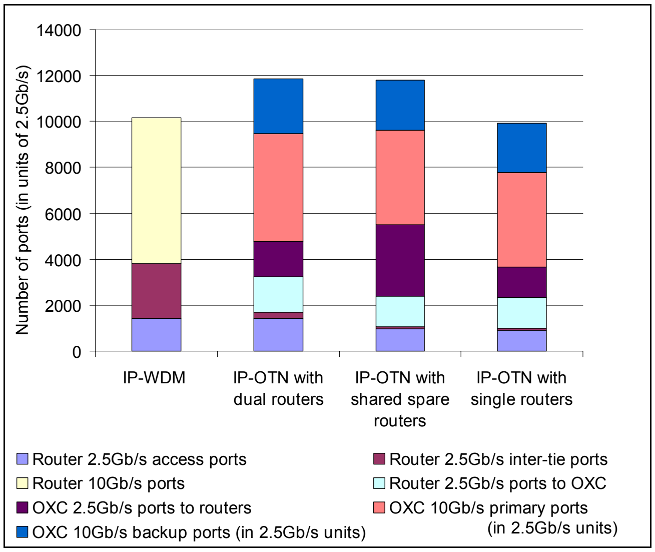


Figure 8 – Bandwidth and Port Comparison.

The results show that all scenarios of IP-over-OTN are cheaper than IP-over-WDM (PMO) solutions. A closer look at the port requirements indicates that the number of 10Gb/s network ports is about the same in all solutions. IP-over-OTN solutions benefits from a 50% better packing of the capacity, which compensates for the 50% to 60% capacity overhead that is required to protect the lightpaths in the OTN. The price per router ports being three times that of an OXC, the cost of transport in the IP-over-OTN scenario is thus approximately one third of the cost of transport in the IP-over-WDM scenario. The IP-over-WDM solutions is also penalized by a much higher transit traffic through the routers, half of which traverses up to two routers within each CO, whereas this traffic is carried more cost-effectively within 2.5Gb/s lightpaths in the IP-over-OTN scenarios. Put together, the cost savings incurred by the IP-over-OTN solution largely compensate for and justify the cost of deploying the OTN. Furthermore this approach has the added advantage to be predictable and robust, with restoration latencies for network failures that are several orders of magnitudes faster than those achieved by a level 3 restoration scheme.

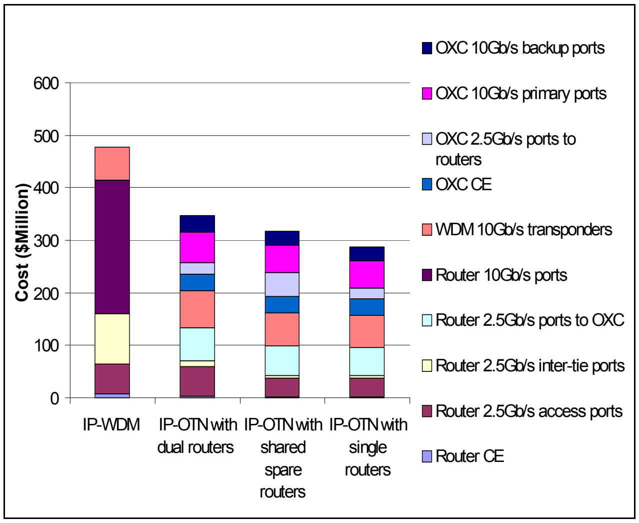


Figure 9 – Cost Comparison.

We also observe some cost differences among the IP-over-OTN strategies, although not as important as between IP-over-WDM and IP-over-OTN. The shared spare router architecture is cheaper than the dual router IP over OTN, in part because it suppresses many of the redundant routers in the CO locations, but also mainly because of a more efficient utilization of the router ports. In particular the utilization of the ports on the access side is not limited to 50% of the port capacity, required in the dual router scenario so that in case of failure of one router, the traffic could resume through the spare capacity of the other router. However some of the gains cancels out with the additional cost incurred for traversing the OXC to access the router from the client premises. They are some trade-off associated with this approach in case of router failures. One is that it requires a longer latency to re-home the traffic from their access router to the spare redundant router. The second is that it needs to re-compute the IP routes from the spare-router to the rest of the network, however the latter can be alleviated by maintaining active connectivity between spare routers and the rest of the network.

Finally, the single router architecture improves the cost effectiveness by trading some protection of the traffic against router failures and thus requires high availability routers to be competitive in terms of resilience. The bandwidth utilization of this architecture is similar to the shared redundant router architecture, except that the clients are connected directly to the routers instead of accessing them through the OXCs. This approach requires less OXC ports; however in case of router failure, the traffic accessing the failed router is unprotected, as it cannot be redirected to redundant routers in the network.

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. Eventually, as routers become even more robust to the point of being fully carrier class, the network can evolve from architecture 3 to 4 by ending reliance on shared spare routers.

V. CONCLUSION

In this paper, we have compared 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 optical reconfiguration; and (4) an architecture where single carrier-class routers are directly connected over an optical transport networks. Based on the results in this paper we draw the following main conclusions:

- 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 provide cost efficiency 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. As a byproduct all these three architectures provide better scalability compared to the currently used 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 backup routers.

REFERENCES

[1] K.G. Coffman and A.M. Odlyzko, “Growth of the Internet”, *Optical Fiber Telecommunications IV B: Systems and Impairments*, I. P. Kaminow and T. Li, eds. Academic Press, 2002, pp. 17-56.

[2] S. Chaudhuri and E. Goldstein, “On the Value of Optical-Layer Reconfigurability in IP-over-WDM Lightwave Networks”, Photonics Technology Letters, 12, August 2000

[3] P. Charalambous et al., “National Mesh Network using Optical Cross-Connect Switches”, Proc. OFC 2003, Atlanta, GA.

[4] T.E. Stern and K. Bala, "Multi-wavelength Optical Networks: A Layered Approach", Reading, MA: Addison Wesley, 1999.

[5] G. Ellinas et. al., "Routing and Restoration in Mesh Optical Networks", Optical Network Magazine, Jan-Feb 2003.

[6] J-F. Labourdette et al., "Routing Strategies for Capacity-Efficient and Fast-restorable Mesh Optical Networks", Photonic Network Communications, June-Dec 2002.

[7] S. Philips, R. Doverspike, and N. Reingold, "Network Studies in IP/Optical Layer Restoration", Proc. OFC 2002, Anaheim, CA, 2002.

[8] Y. Qin, L. Mason, and K. Jia, "Study on a Joint Multiple Layer Restoration Scheme for IP over WDM Networks", IEEE Network Magazine, March/April 2003, Vol. 17, No. 2.

[9] P. Pongpaibool et al., "Handling IP Traffic Surges via Optical Layer Reconfiguration", Proc. OFC 2002, Anahaim, CA, March 2002.

[10] A. Chiu and J. Strand, "An Agile Optical Layer Restoration Method for Router Failures", IEEE Network Magazine, March/April 2003, Vol. 17, No. 2.

[11] Report on IP Reliability, Lightreading, http://www.lightreading.com/document.asp?doc_id=28259.

[12] C. Labovitz, A. Ahuja, and F. Jahanian, "Experimental study of Internet Stability and Wide-Area Backbone Failures", Proceedings of Fault Tolerant Computing Symposium, June 1999.

[13] A. Akyamac, et al., "Ring Speed Restoration and Optical Core Mesh Networks", NOC 2002, June 2002, Darmstadt, Germany.

[14] Implementation Agreement **OIF**-**UNI**-01.0, http://www.oiforum.com/public/documents/OIF-UNI-01.0.pdf.