

Composite Links™

A Tool for Bandwidth Optimization

Abstract

The IP Service Provider market is going through an unprecedented period of bandwidth growth and services transformation. The carriers best positioned to thrive in this environment have low-cost networks with the flexibility to efficiently support a broad range of services.

For carriers seeking to reduce costs, one of the best opportunities lies in increased network utilization levels. Higher utilization levels enable more revenue to be supported across the network without triggering increased costs. The challenge is how to increase network utilization without adversely impacting service levels and revenue.

Avici's Composite Links technology provides carriers with a powerful proven tool to reduce bandwidth costs, increase network reliability, and deliver greater operational flexibility. This document outlines the applications and benefits of Composite Links, and illustrates them using a real-world carrier case study.

Author

*G. Hudson Gilmer
hgilmer@avici.com*

Introduction

Efficient bandwidth management is fundamental to the survival and success of IP Service Providers. There are tremendous financial and competitive pressures to reduce cost levels and increase network reliability. As carriers seek strategies to “sweat” their assets to reduce costs, improving network bandwidth utilization is a prime candidate. A significant percentage of carrier assets are tied up in their networks, with a large IP network costing hundreds of millions of dollars to build and operate. These networks are often highly underutilized, with average utilization levels at 20% or less. Several factors contribute to this low level of utilization, including resiliency requirements, poor interface granularity, inflexible provisioning processes, and costly link protection mechanisms.

The challenge for carriers is how to simultaneously increase network utilization levels and network reliability. Cost reduction is critical, but the growth of high-margin real-time IP services such as voice, video and gaming are placing higher reliability demands on networks originally designed for forgiving best-effort services. New tools and strategies are required to improve bandwidth efficiency and raise network utilization levels above current levels.

Avici's Composite Links provides a powerful operational tool by delivering cost-effective local link protection, bandwidth flexibility, and the ability to gracefully scale to higher capacities.

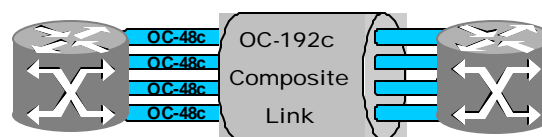
Technology Description

Avici's Composite Links technology extends many of the benefits of Ethernet link aggregation to SONET/SDH links. Composite Links enables up to 64 separate physical SONET/SDH network interfaces to be aggregated into a single logical trunk at the network or IP layer. Each Composite Link member is a SONET/SDH interface, using Packet over SONET (POS) encapsulation. Each member is established using the standard PPP protocol, thus offering multi-protocol support. The resulting Composite Link is a

single logical, concatenated interface for the IGP forwarding table with a single IP address. Composite Links technology is protocol and service agnostic, supporting unicast, multicast, MPLS and IPv6 traffic.

Figure 1 shows a simple example of a virtual OC-192c Composite Link consisting of four OC-48c/STM-16c member links. Upper layer protocols (IP and MPLS) treat the Composite Link as a single logical entity with a capacity of the aggregate Composite Link member links.

Figure 1 – Virtual OC-192c Composite Link



Support for Various Link Speeds

Composite Links offer the unique ability to combine member links with a 4:1 speed differential. For example, one OC-48 and five OC-12s can be combined into a 5.25 Gbps Composite Link. This capability enables non-disruptive link expansion and improved bandwidth granularity.

No Packet Re-Ordering

Another important feature of Composite Links is the ability to maintain packet ordering for individual IP and MPLS flows. Packet ordering is maintained at the micro-flow level by applying a hash value to the IP source and destination address or MPLS label to select a Composite Link member. This hashing mechanism load balances traffic based on the bandwidth of each member link, ensuring that individual micro-flows are not split across multiple physical links. This hashing technique evenly distributes traffic across all member links without latency-inducing buffers.

Benefit 1 - Local 1:N Link Protection

As carriers strive to improve the reliability and stability of their IP networks, improved link protection mechanisms are critical. Optical link failures are responsible for over 30% of downtime in IP networks, and current link

protection mechanisms are too slow, cumbersome and expensive for widespread deployment.

IGP Convergence & ECMP

Most IP networks today rely primarily on IGP route convergence to circumvent link failures in the network. Peak non-failure IP network utilization levels are generally maintained at or below 50% so there is sufficient capacity to accommodate rerouted traffic without triggering major congestion. A significant drawback of this approach is that IGP route convergence typically takes 10 or more seconds, disrupting sensitive services such as VoIP, ATM and Interactive Gaming. As these services continue to grow, carriers are recognizing the need for faster less disruptive link protection mechanisms.

Equal Cost Multi-Path (ECMP) represents a partial improvement over simple IGP convergence, by enabling link aggregation and faster link protection. Because the local route table must be updated in the event of member link failure, the recovery time for ECMP is significantly longer than for Composite Links and above the 50ms threshold for local protection. Unlike Composite Links, ECMP does not aggregate all links into a single logical link, rather each ECMP member link is installed as a next-hop in the routing table. This causes the IGP forwarding table to grow as ECMP members are added. Perhaps the largest drawback of ECMP is the operational burden and complexity of manipulating, documenting and managing individual ECMP member link traffic metrics to achieve the desired routing behavior.

Local Link Protection Alternatives

Using only IGP convergence or ECMP for link protection, every link or linecard failure disrupts service for all traffic on the link. When the majority of traffic on IP networks was best-effort applications such as web browsing, e-mail or file transfers, such disruptions were rarely noticed by users. With the growth of real-time IP services and converged legacy services carried over IP, carriers are looking for

mechanisms to further improve network stability and achieve guaranteed quality of service.

To improve recovery speeds and prevent service disruption, alternative “local protection” mechanisms such as MPLS Fast Reroute (FRR) and SONET Automatic Protection Switching (APS) have emerged. These protection mechanisms are described as “local” because they fail over within 50 milliseconds, do not disrupt services, and do not trigger route convergence. The drawback of these approaches is cost and complexity. SONET APS and MPLS FRR (detour) require 1+1 protection, meaning that a standby/protect link must be provisioned for each active link. The cost of this 100% protection “tax” prevents carriers from deploying such protection mechanisms widely throughout the network. MPLS FRR carries the additional disadvantage of operational complexity. MPLS FRR requires that Traffic Engineering be implemented, and then requires manual provisioning of each backup LSP – a burden which most carriers would prefer to avoid.

1:N Link Protection with Composite Links

Avici’s Composite Links technology also supports local link protection with sub 45 ms failover, but does so more cost effectively and with less complexity than the methods described above. Composite Links employs a 1:N loadsharing redundancy, enabling carriers to spread N links of traffic across N+1 member links. In the event that any link fails, traffic is rapidly redistributed across the N surviving links. This 1:N protection behavior can dramatically reduce the cost of protection, making it more affordable to deploy protection widely. For example, the cost of protection for a 4-member Composite Link is 1/3 of the cost of protection using SONET APS. In addition to protecting against link failure, Composite Link technology protects against failure of a linecard module or an interface component (i.e. optical transceiver) on a module.

Most large carrier networks employ a combination of protection mechanisms for different parts of the network, driven by the objective of minimizing protection costs and complexity while maximizing scalability and

speed. The following table (figure 2) summarizes the scalability, cost, complexity and speed of several link protection mechanisms:

Figure 2 – Link Protection Mechanism Summary

Link Protection Mechanism	Scalability (# of Addresses)	Cost of Protection	Complexity	Speed of Failover
IP Reroute	1	up to 50% for resilience	Low	10-30 sec.
MPLS Fast Reroute (0:1) protection	1+1	100% unless shared	High	sub 45 ms
ECMP (1:N) Loadsharing Redundancy	N	100% trans.	High	100+ ms
SONET APS (1:1) Active/Standby	1	100%	Low	sub 45 ms
Composite Links (1:N) Loadsharing Redundancy	1	1/N%	Low	sub 45 ms

Benefit 2: Bandwidth Flexibility & Efficiency

Carriers with large and growing IP networks face a number of significant issues in planning, maintaining and upgrading the links on their networks. Issues include service disruptions due to link expansion, poor utilization due to non-granular link speeds, and the high cost of backbone link upgrades.

Non-Disruptive Link Expansion

Traditionally, expansion of a link involves a disruptive and cumbersome process of provisioning a parallel link, manipulating administrative distance parameters to send traffic over the new link, then disconnecting the old link. This process creates two “convergence events” which can each disrupt traffic for up to 10 seconds while the routing tables reconverge. Such a disruption may not noticeably impact customers using applications such as e-mail and file transfer, but can cause sensitive Interactive Gaming, Video and VoIP sessions to be lost. These disruptions can be very costly for carriers, triggering customer service calls, maintenance activities, SLA penalties and even customer churn. As the number of real-time IP applications increases, the cost of such disruptions to carriers will continue to rise.

Composite Links enables carriers to non-disruptively expand and reduce link capacity by leveraging the capability to include different

speed links in a single Composite Link. For example, to non-disruptively upgrade from an OC-3 to an OC-12 link an OC-12 link is added via Composite Links to the existing OC-3 link. Once the combined link has been tested and is stable, the old OC-3 member link can be removed from the Composite Link. With Composite Links, there is zero impact to the layer 3 protocols, and no service disruption. In addition, by eliminating the need to adjust administrative distances, the risk of human error and configuration-related outages during the link upgrade process is significantly reduced. This same non-disruptive approach can also be used to downsize a link, add member(s) to an existing link, or replace /upgrade linecards.

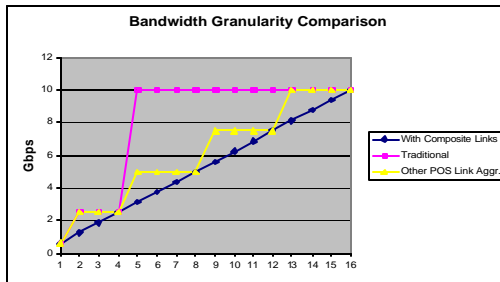
Bandwidth Granularity for Higher Utilization

Composite Links also offers another important advantage to carriers seeking to improve network costs – increased bandwidth granularity. Traditionally, carriers face significant delays and constraints to turning up new or higher-speed fiber in their network. Carriers must plan months ahead for new capacity, face significant financial penalties if ordered fiber is cancelled, and limit the number of links per route to avoid scaling issues and network complexities. Carriers typically avoid provisioning more than 2 parallel links per route. After a second link is added, additional bandwidth requires a 400% jump in link capacity (i.e. from OC-12 to OC-48). Such a link upgrade typically requires costly changes to the underlying physical layer technology. Furthermore, the jump to 4x larger links causes carriers to make a difficult tradeoff between congested or severely underutilized links.

In contrast, Composite Links offers fine link scaling granularity because the number of physical members does not increase logical network complexity or cause scaling problems. As illustrated in figure 3, a carrier looking to scale an OC-12/STM-3 link would have only two steps without ECMP or four with ECMP. However, using Composite Links the carrier can optimize the link utilization level using up to 16 incremental steps. This “just in time” bandwidth management enables fine-tuned

traffic engineering for higher overall network utilization and improved ROI.

Figure 3 – Bandwidth Granularity Comparison



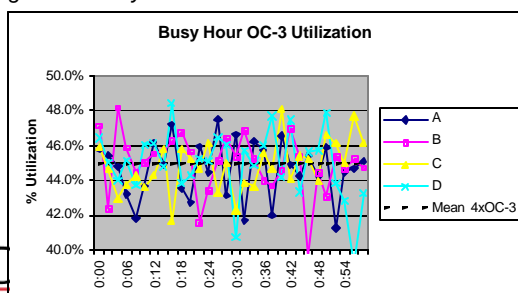
With the ability to dynamically scale link bandwidth up or down, carriers gain a significant competitive advantage in provisioning speed, which directly impacts revenues. Changing capacity of an individual link can be done transparently to the rest of the network. This greatly simplifies and speeds the provisioning process, and positioning carriers to rapidly respond to wholesale bandwidth buying and selling opportunities.

Higher Network Utilization & Performance

A basic principle of traffic engineering is that a single large capacity link can achieve higher utilization levels than multiple smaller links with the same aggregate capacity for a constant level of performance/blocking. This is due to the fact that a larger capacity link aggregates more traffic, resulting in less burstiness and more predictable traffic peaks.

Figures 4 and 5 illustrate the utilization improvement resulting from Composite Links, by comparing utilization rates of 4 standalone OC-3s with utilization of a virtual OC-12. Figure 4 illustrates the normal variability of traffic utilization on each of four OC-3s. Due to this variability, a carrier who is engineering links on the network for 50% maximum utilization must maintain average utilization at 45%.

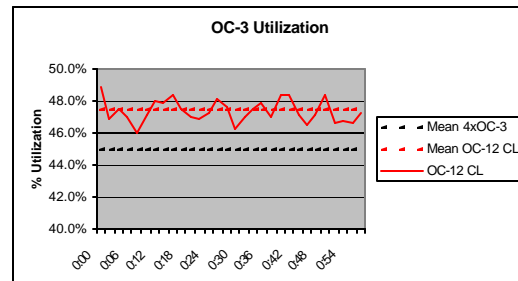
Figure 4 – Busy Hour Utilization



By aggregating the traffic of 4 OC-3s into a virtual OC-12 using Composite Links, the variability of utilization is significantly reduced (Standard Deviation of traffic reduces by almost 60% from .017 to 0.0069)

The net result, as illustrated in Figure 5, is that with Composite Links, carriers can increase effective network utilization and reduce cost per Gigabit by nearly 6%, while staying below 50% maximum utilization.

Figure 5 – Comparison of Mean Utilization



Network Simplification for OpEx Savings

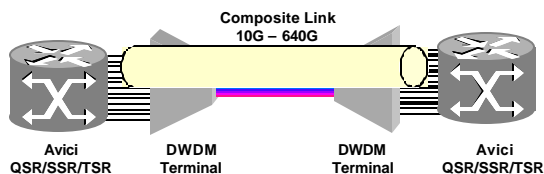
In addition to the direct savings by increasing effective utilization on network links, carriers can reduce OpEx costs by reducing network complexity. By reducing the number of logical links in the network, Composite Links simplifies the network management function. Fewer links means fewer routes to traffic engineer, less difficulty managing growth, and faster maintenance/fault isolation. Network complexity and operational costs are further reduced because link protection does not have to be separately provisioned. Because Composite Links can protect against failure of a member link in less than 50 milliseconds, the need to provision and maintain local protection/backup LSPs can be eliminated. Composite Links can also simplify the implementation of Quality of Service (QoS) because QoS is supported on the logical Composite Link, rather than separately on each physical member link. This simplifies the implementation of QoS on a network, and enables higher-priority traffic to be insulated from outages affecting individual member links.

Benefit 3 – Composite Links and DWDM: A New Model for Network Scaling

Traditionally, a carrier seeking to increase capacity on a backbone route was forced to choose between two costly options: lay new fiber or purchase wholesale capacity. The advent of DWDM has fundamentally altered the economics of adding network capacity, enabling order of magnitude increases in optical capacity simply by provisioning additional lambdas on an existing fiber pair. As a result of the widespread deployment of DWDM equipment in core networks, carriers will choose to exhaust the capacity available to them in their existing fiber before laying new fiber or upgrading networks to higher link speeds. This new scaling model will result in networks consisting of relatively large numbers of parallel OC-48/STM-16 or OC-192/STM-64 links. The disadvantage of such a model is that the increased number of links can lead to serious network scaling, instability, and management complexity problems.

Composite Links is ideally suited to controlling the scaling, instability and management issues associated with DWDM-driven Nx10G/Nx2.5G networks. As illustrated in Figure 6, Backbone networks can scale from a single 10G link to an N x 10G Composite Link without increasing network complexity, triggering route convergence, or encountering scaling inefficiencies. With Composite Links, the network topology is unchanged regardless of the number of physical member links added.

Figure 6 – Nx10G Scaling w. Composite Links



By keeping the number of logical links in the network constant, the size of the Interior Gateway Protocol (IGP) remains constant. This conserves valuable router memory resources, improving protocol scaling, network stability, and route convergence time.

IP-Optical Integration with Composite Links

Leveraging industry standards such as the Optical Internetworking Forum's (OIF) Universal Network Interface (UNI)¹, Composite Links can be extended to dynamically scale links across a switched optical backbone. For example, a carrier can use the "Open Threshold" feature of Composite Links to pre-define a utilization threshold, which triggers an additional lambda to be added to an existing Composite Link if exceeded. These integrated IP-Optical mechanisms for dynamic bandwidth scaling have been proven in multi-vendor laboratory environments, and promise to improve utilization levels and reduce provisioning costs in future production networks.

Summary

Leading carriers have made tremendous investments in their IP networks and recognize that they must improve the return on those assets.

Avici's Composite Links is a powerful tool for maximizing the ROI of network assets by improving utilization levels throughout the network, delivering cost-effective local link protection for improved reliability, and gracefully scaling to manage years of profitable growth.

¹ visit <http://www.oiforum.com/> for the OIF signaling specification, details of multi-vendor testing, etc..

Case Study – Optimizing Transatlantic bandwidth through Composite Links

This case study is based on a real-world scenario of a large carrier looking for solutions to improve unit cost and reliability on costly transatlantic IP network links. Total direct costs of transatlantic capacity over 3 years will be compared using a traditional architecture (IGP convergence only) and a Composite Links-based architecture.

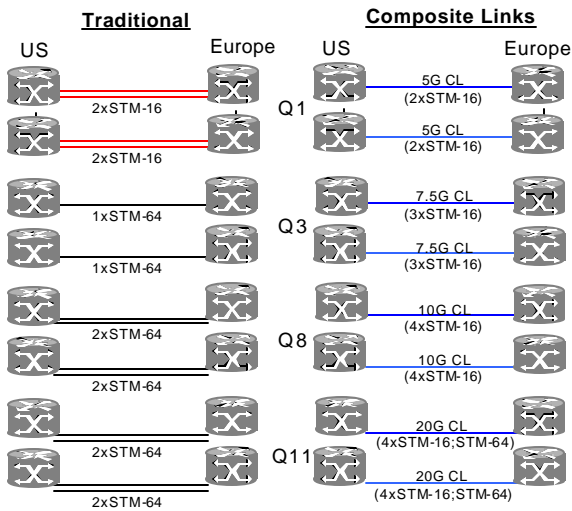
Background:

Carrier XYZ Telecom currently has two transatlantic routes, each with 5Gbps of capacity, and is experiencing 75% traffic growth per annum. XYZ Telecom does not own transatlantic fiber, and must lease wholesale capacity. Leased capacity requires one-year minimum term contracts.

To satisfy network reliability SLAs, XYX Telecom manages the network to a 50% maximum utilization rate, which equates to a 25% average utilization rate. To manage complexity and prevent scaling issues, XYX Telecom never runs more than 2 parallel links between city-pairs.

As shown in Figure 7 below, the constraints of traditional network design result in underutilized links and disruptive link upgrades.

Figure 7 – XYZ Telecom Transatlantic Scaling Options



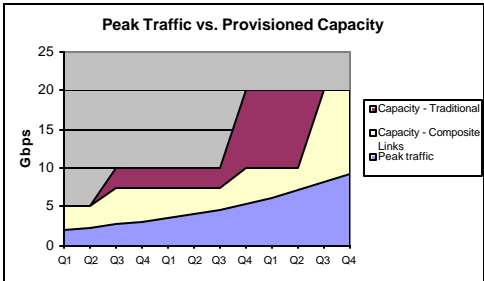
Benefit 1: 1:N Local Protection

By grouping the transatlantic links into a Composite Link, XYZ Telecom can simultaneously improve utilization on the aggregate links and reduce downtime. The traditional configuration relies on IGP rerouting for resiliency, necessitating 50% maximum utilization to maintain 50% spare capacity for (disruptive) rerouting in the event of link/module failure. By implementing Composite Links, XYZ Telecom is able to leverage 1:N loadsharing redundancy to increase average network utilization. Furthermore, with the deployment of Composite Links, XYZ Telecom achieves local link protection against link and linecard failures. (cost benefit of improved reliability is not quantified)

Benefit 2: Bandwidth Flexibility

Composite Links frees XYZ Telecom from traditional bandwidth constraints, resulting in greater granularity and less overprovisioning. As shown in the Figure 8 below, XYZ Telecom can use Composite Links to more closely match provisioned capacity to peak traffic requirements, minimizing overprovisioning. The ability to combine links with a 4:1 speed differential further enhances this flexibility. Regardless of the number of member links, the Composite Link appears as a single logical link, preventing network scaling and complexity issues.

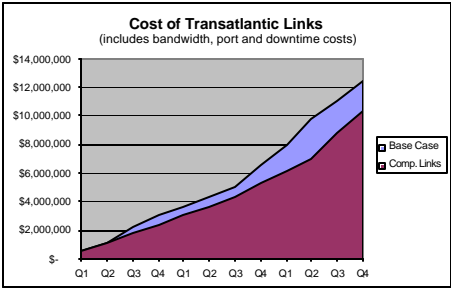
Figure 8 – Traffic vs. Provisioned Capacity



Benefit 3: In-Service Link Upgrades

In the base case scenario, XYX Telecom must perform two disruptive link upgrades. The first upgrade involves migrating from two STM-16 links to a single STM-64 on each route. The second involves adding a second STM-64. While these upgrades are typically scheduled for maintenance windows to minimize service disruption, some impact on traffic over the

affected route and wider network impact of the resulting route convergence is inevitable.



Summary of Cost Benefits

By deploying Composite Links, XYZ Telecom can improve network reliability and achieve a \$2.1M (16.8%) cost reduction on transatlantic link equipment, bandwidth and downtime costs. The equipment and bandwidth savings are achieved through improved utilization, which is made possible by a combination of greater bandwidth granularity and 1:N loadsharing redundancy. The downtime cost reductions are based on eliminating disruptive upgrades in Q3 and Q4.

Additional Benefits

This case study addresses only the costs of bandwidth, router port costs and disruption from upgrades. Additional factors such as improved reliability through local link protection, optical equipment costs and operational costs are not quantified in this study, but would further increase the TCO advantage of using Composite Links.

Furthermore, this case study assumes that XYZ Telecom has perfect visibility on future traffic requirements and orders bandwidth accordingly. In reality, visibility is less than perfect and one of the greatest advantages of Composite Links is the ability to avoid costly over/under provisioning by rapidly adapting to market requirements.

Key Inputs & Assumptions:	
Annual Growth Rate	75%
Transatlantic Bandwidth Cost*	
OC-3/STM-1	\$ 10,000
OC-12/STM-4	\$ 32,500
OC-48/STM-16	\$ 105,625
OC-192/STM-64	\$ 343,281
* source Bandwidthmarket.com/T.Soja & Assoc. 11-02	
Router Interface CapEx Cost **	
OC-48/STM-16	\$ 50,000
OC-192/STM-64	\$ 180,000
** source Avici/Cisco/Juniper avg. port prices @20% discount	
Cost of Svc Disruption***	
Cost per incident of disruptive upgrade	\$ 100,000
*** source NSP Reliable IP Nodes: A Prerequisite for Profitable IP Services; Used cost of downtime based on minor outage	