Interworking of L3/2 Traffic Engineering and Protection Schemes and L1 Adaptive-Mesh

Introduction - *Product and Process Innovations*

There are technological innovations that aim at change; and then there are technological innovations that don't aim at change, but rather at efficiency improvement without requiring a change in a way the external world uses them compared to how the existing technology alternatives were used. We shall call the former type of new technologies as *"product"* technologies, and the latter type as *"process"* technologies.

Both of these types of new technologies can produce economic value, and it is important to notice that the latter type of technologies, while not often perceived as "disruptive", can in deed unlock much greater net-value, due to the lack of, or minimized, transition cost of adopting such new technological innovations.

In the field of communications networks, specifically network technologies involving new network hardware, most new technologies have been introduced as products rather than processes. This is mainly due to the history that the network hardware equipment developers were also the equipment manufacturers. As long as that was true, as it was too complicated for a user of network equipment, i.e., for a service provider, to handle the full range of functions from equipment manufacturing to end-customer telecom services, companies involved with networking and telecom business had to specialize as either equipment manufacturers or as service providers.

However, nowadays the actual equipment manufacturing is routinely performed by specialized electronics manufacturing service providers rather than by the developers of network hardware technologies. Moreover, network hardware equipment can presently be designed based on very straightforward physical hardware implementation, e.g. just a single programmable logic chip and protocol-transparent line transceivers on single, generic circuit board. Thus, at present time, the key function of a developer of networking technologies is the design of the *functionality*, while implementation and design of the physical hardware can be outsourced. This creates opportunities for innovating new process technologies, to be utilized by the developer of such innovative process technologies in producing existing standard compatible services -- just more efficiently.

Tapping into these new opportunities enabled by the evolution of network technologies and service supply chain, a network service provider startup, Optimum Communications Services, Inc. (OCS), has developed a new *process* technology, called Adaptive-Mesh Layer 1.5 VPN (A-M for short), that provides existing, common network protocol standard compatible wholesale network connectivity services at multiple times improved cost-efficiency.

However, while Adaptive-Mesh based network services seamlessly work with the existing standard (MPLS, SDH/SONET, Ethernet) equipment, the fact that the way that Adaptive-Mesh implements the commonly understood, standard network connectivity models (e.g. functionally equal to regular hub-and-spoke architectures) is new, can cause a level of confusion among network engineers accustomed to evaluating product, rather than process, based technology innovations.

It should be observed that an important benefit of process innovations, versus product based innovations, is that **no matter how different the way of implementing a given customer-visible functionality via a new process technology, the customers can continue to interface with such new process technology the same way they have interfaced with the prevailing process technologies and the services produced by them**.

Accordingly, unlike with product innovations, with process innovation the customer can simply validate, via its existing test patterns, whether a service produced by a new process technology in deed meets or exceeds the existing standard (e.g. Service Level Agreement, SLA) requirements and inter-operates with their existing standard systems using the service.

Thus, it should be understood that, unlike product innovations that aim to get a customer to buy something new, process innovations generally simply aim at (radically) reducing the unit cost of what the customers are already buying. That does not mean that process innovations would be limited to just an incremental improvement. Rather, the opposite may be true: due to allowing transition-cost-free adoption, the most groundbreaking innovations are more straightforwardly commercialized as process technologies.

Adaptive-Mesh: Process Technology Producing Standard Services *at New Standard of Cost-Efficiency*

Since Adaptive-Mesh L1.5 VPN (A-M) is used by OCS as a process technology, using which OCS produces standard network connectivity services, the service produced can be explained very simply:

> ***A-M appears to the customer as a dedicated, protected hub-and-spoke network with industry standard SLA term based service availability and throughput, latency and jitter guarantees between the set of customer sites connected to the A-M.***

Thus, as a wholesale network service, A-M appears to its customer no different (and specifically, no worse in performance) than if the wholesale carrier provided a physical, per-customer-dedicated, doubled hub-and-spoke network for the customer of the wholesale network service contract.

Which brings us to the question: S*o why would Adaptive-Mesh then be needed if the service produced through it is no different than what can be done with existing technologies?*

The answer: *The Adaptive-Mesh enables producing the same given SLA based service contract with a small fraction of the cost of providing the same contract with any alternative technologies.*

But *how* does A-M allow providing the same service capacity with lower cost without compromising quality?

In case of OCS' A-M, the techniques enabling the up to 20 times higher cost-efficiency are published without restrictions via the patent filings[1] by the company regarding the technology, and the company's website[2], and thus OCS does not need to maintain as any secret the reasons for the architectural cost-efficiency gain of its A-M process technology. Moreover, the operation of an A-M test network is demonstrable, and tested to work in physical hardware as is expectable by the theory of its operation per above referred technical literature.

Thus, there are the below key considerations regarding any risk for a customer to utilize the cost-efficiency gains -- *which realistically could be equal to in the order of 50% revenue increase in their impact on the customers' business bottomline* -- achievable via transitioning from traditional wholesale network services to using A-M L1.5VPN service:

- The service contracts based on A-M can be understood as if they were implemented using a conventional, dedicated and doubled-for-protection hub-and-spoke network, the behavior of which architecture is well understood.
- The service contracts include industry standard SLA term based availability and performance guarantees.
- The theory of operation, architectural and reference implementation specifications of A-M networks are publicly available.
- An A-M test network, involving 3rd party protocol test equipment as well as the A-M network nodes, emulating a reference customer network, is demonstrable.
- The customer can simply test and validate, on a simple pass/fail basis, whether or not an A-M network implementation for a given contract is performing per the SLA.

Given the above considerations, there are no rational reasons for not realizing the cost-efficiency improvement through (non-disruptively) transitioning to network connectivity service contracts implemented by A-M L1.5VPNs.

In the following, we will cover the functioning of A-M L1.5VPN regarding its packet forwarding, traffic engineering and traffic protection aspects.

---

[1] US patent application #12/363,667 and US patents #7,986,713, #7,333,511 and #7,254,138.

[2] www.ocsipholding.com

Why is Adaptive-Mesh Functionally No Different (or at least No Worse) Than a Physical Hub-and-Spoke?

Short answer: because it is designed, tested, and demonstrable to be equal, or at least no worse; A-M L1.5VPN is merely a more cost-efficient implementation of the functionality provided by the well-understood physical hub-and-spoke network architecture.

Let's study the intended behavior of a physical hub-and-spoke network. The behavior of the network is properly defined in terms of what the customer nodes can receive through it from the sources that given customer nodes are desired to receive traffic (rather than what any given node can send to the network, as it is the reception by the indented destination that completes transmission of given data being carried by the network). So what are the policies for what traffic a given customer router should receive from the other routers of the customer through our ideal hub-and-spoke architected, assumed third party wholesale carrier managed network? Since that is a problem regularly solved by L3/2 Traffic Engineering (TE), this brings us to the question of *does the data load (rather than TE policy) driven bandwidth allocation of A-M conflict the TE based inter-router LSP bandwidth quotas of a standard MPLS network?* So let us compare how a physical hub-and-spoke and A-M behave regarding TE .

Comparison of Hub-and-Spoke and Adaptive-Mesh Regarding Traffic Engineering

It is desirable, and often necessary, to know the throughputs between the customer nodes through the (assumed wholesale carrier provided) core network. A brute-force way to accomplish such deterministic performance connectivity is to provide a full mesh of dedicated, constant bandwidth L1 connection between each customer router pair. Such architecture naturally however would soon become prohibitively expensive as the customer node count increases; even just 10 routers to be connected that way would require a total 10x9=90 dedicated L1 connection pairs.

Let's assume that each of the 10 routers can exchange traffic with the other 9 routers at capacity of e.g. 10Gbps (full duplex), in any breakdown of bandwidth allocation between the nine other routers. In that case, there would be 9 10Gbps connections per each (destination) router in the network, for a total capacity of 90Gbps, while any one of the routers (as destination) can forward traffic only at 10Gbps, i.e., the network can only be utilized at most at 10/90 = 11% utilization level -- and even lower rates as the node count increases. Such static mesh network architecture, even though deterministic, thus is overly inefficient and non-scalable.

The common solution to the above low capacity utilization and scalability problems is to perform packet level switching in the core network. With packet-switched interconnect network, each router needs only a single, stat-muxed connection over which to exchange its traffic with all other routers connected to that packet-switched network cloud, allowing these reduced number of connections to be utilized to their maximum capacity. So the utilization level and physical port count based scalability problems are solved, and everything is fine -- right?

But through such packet-switched, stat-muxed cloud, what are the inter-router throughput, latency and jitter level guarantees? Well, that can depend on multiple factors, such as what the other routers are sending and to where at a given time on the network. Thus, without implementing packet-layer coordination of data transmission between the nodes interconnected by the packet-switched cloud, that network is non-deterministic.

Which takes us back to beginning of this section. Since there is no going back to static, physical (or sub-physical) layer mesh networks in particular as the network node counts keep on increasing, it appears that a method is needed to enable network-scale coordinated data transmissions between the routers across a packet-switched inter-connect network.

Of course, there already exist several methods for Traffic Engineering and QoS policing that aim at achieving more deterministic source-destination performance over packet-switched network clouds. However, the interesting question here is:

Is the way that the packet-layer bandwidth reservation techniques seek to achieve deterministic throughputs among all the nodes able to exchange traffic across the given shared pool of packet-switched network capacity really fundamentally different than providing L1 circuits between the nodes?

The laws of physics dictate that when bandwidth across the shared network is reserved for a given source-destination router path (e.g. LSP), that same bandwidth is not available for any other traffic, even when the intended traffic did not materialize to the capacity reserved for the LSP. Effectively, the bandwidth-reservation based paths, e.g. Traffic Engineered MPLS LSPs, are soft-fixed bandwidth “L2 circuits”, that collectively consume essentially the same amount of physical network capacity as if they were actual L1 circuits of the same bandwidth. *So did we really improve on regular TDM, or did we just make achieving the intent of it more complex?*

There certainly are advantages of using L2 circuits rather than traditional L1/L0 circuits, primarily that L2 circuits can have more flexible bandwidth allocated to them. However, they also have disadvantages: since L2 circuits go through a number of instances of packet-level processing, buffering and switching, each of which adds a variable amount of delay, the L2 circuits cause more latencies and jitter than actual L1 circuits, i.e., they provide inferior QoS. Also, the L2 circuits are administratively more complex, as well as require significantly more advanced processing power at the network nodes -- in that way making the networks less scalable.

Thus, the packet layer TE at best provides a soft-circuit equal of the L1 mesh network; other than more granular capacity allocation than traditional TDM or WDM, they do not provide any actual net benefits.

So it is far from problem solved, in particular as the service provider networks at the same time need to support an ever increasing number of nodes, and provide realtime interactive service type strict QoS guarantees for an increasing portion of the traffic.

That is not to say that TE would not be needed; quite the opposite, TE will very likely be an essential portion of the solution.

But in addition to TE, what is needed is an innovation allowing to combine the desired aspects of both “L2 circuits” and traditional L1 circuits, i.e., soft bandwidth and hard QoS, while simplifying the implementation and administration of the networks and service contracts, thus enabling more cost-effectively scalable, high quality networks. Benefits of such adaptive bandwidth L1 mesh network include that it allows providing actual, direct L1 circuit based, strictly deterministic QoS, and eliminates the need for any packet-layer processing at the intermediate nodes, via allowing L1 based traffic pass-through at any nodes where a given packet is not entering or leaving the network interconnecting the customer’s network access points.

But, with a packet traffic load adaptive bandwidth L1 mesh, even assuming that it was technically possible to implement it, how are its adaptive bandwidth L1 connections sized to match the TE-policy based inter-router bandwidth allocation? Doesn’t the control of the bandwidths of such flexible bandwidth L1 connections cause at least as much complexity as there would be with a mesh of L2 circuits? And can the L1 connections capacities ever be as granular as the bandwidths of L2 circuits?

Strikingly enough, in the so far only known working implementation of realtime dynamic, traffic-load-adaptive bandwidth L1 connectivity, i.e., in OCS’ A-M, there is no (explicit) control by the routers that it inter-connects on the capacities of its mesh of dynamic bandwidth STS-X circuits. Instead, each destination node in A-M, for every new STS (VC-3) row period (i.e. at rate of 72000 times per second), autonomously re-optimizes the allocation of STS-1 timeslots on a STS-N bus through the network carrying packets from the source routers to the destination routers served by the bus. In the process, A-M achieves source-destination router specific bandwidth granularity of roughly 50Mbps/72000 = 700 bits per second, i.e., much finer granularity than is realistically achievable with any packet-switched techniques.

But is there any way for TE-capable routers to control the bandwidth allocation of the STS-X circuits between A-M interface modules, and is there thus a risk that the automatic bandwidth allocation of A-M may conflict with the TE policies? To study these questions, a highly important factor to notice here is that:

> ***By A-M, as it operates as a transparent courier between TE-capable routers, an arrival of a packet from a customer router that is destined toward a given destination router of the same customer is taken as a proxy that the destination is allowing such a packet to be sent to it.***

That is because A-M allows the customer's TE-capable (MPLS) routers to negotiate their TE based source-destination LSP bandwidth allocations directly among themselves, and therefore, instead of re-policing customers' traffic, A-M assumes, as it should, that packets received by it from customer's router and destined toward another router of the customer is valid traffic according the to the TE policies.

Thus, it is seen that L3/2 Traffic Engineering and the adaptive-bandwidth L1 based Adaptive-Mesh are not just compatible techniques; they are highly ***complementary*** techniques: While TE is needed for macro-level (e.g. in units of Mbps) and non non-realtime (policy re-configuration/enforcement at time scales of one second and higher), A-M handles the packet-by-packet, network byte-timeslot accurate maximization of network traffic throughput, with 72,000 timeslot allocation optimization cycles/second.

But even then - L3/2 switches/router already work fine with existing L1/0 transport networks, even without any adaptive bandwidth L1 connectivity, why would that A-M be needed in the first place, as what it claims to do between a set of routers already can be done without it?

Because, A-M can provide the required inter-router connectivity with up to 20:1 reduction in physical network capacity requirements[3], i.e., at accordingly lower cost per Gbps of network throughput. Here, it is noted that A-M is truly a *process* technology; essentially the only impact the customer will unavoidably notice if switched to using A-M based wide area network connectivity services is the financial impact, i.e., multiplication of Gbps of network service capacity per unit cost vs. the alternatives, in many cases allowing the customer to get increased network capacity and reduced cost.

So, we saw that A-M, instead of conflicting with TE policy based inter-router bandwidth allocation quotas, or just not being TE-aware, actually implements the ultimate intent of the TE based LSP bandwidth scaling. And it does that in its realtime automatic and transparent manner.

Comparison of Hub-and-Spoke and Adaptive-Mesh Regarding Traffic Protection

As far as the physical network design is considered, the situation again is the same for both the physical hub-and-spoke and A-M implementations: for end-to-end protection, double the entire network, including the customer routers, organize the customer routers as mutually protecting pairs, in each case. Of course, with regular physical hub-and-spoke, there would be multiple times more physical network capacity that needs to be doubled, making protection of a regular hub-and-spoke network that much more costly.

Another issue to consider regarding the relative robustness of regular, doubled physical hub-and-spoke and A-M network implementations is the impact of any single or multiple failures in the network. With hub-and-spoke, failure of any of the core switches would have the same impact on available network capacity than the exponentially less likely event of concurrent failure of all A-M interface units of the A-M network segment corresponding to the hub-switch that it could replace. Moreover, in case of simultaneous concurrent failure of both of the mutually protecting core switches, no communication is possible through the hub-and-spoke network, while with A-M a good portion of the inter-router links will still be working.

---

[3] http://www.ocsipholding.com/OCS_White_Paper_200Gbps_of_Non-oversubscribed_MPLS_access_over_40Gbps_wavelength_ring.pdf
Note that the inter-connect networks for any given customer contract can have any number of levels of hierarchical core switches or Adaptive-Mesh groups.

Also, ideally the customer nodes should be able to monitor their mutual reachability status directly through the core network, e.g. via link-state advertisement mechanism. A-M provides transparent PPP links between each customer site connected, so it enables the customer routers to directly check their mutual reachability. A disadvantage of regular physical hub-and-spoke is that it does not provide direct, transparent L2 links between the customer routers, so they need to transact through the hub switches/routers.

Furthermore, even before the customer routers could react to a change in their mutual reachability, the A-M based interconnect network would already have performed the necessary packet-layer traffic protection[4], e.g. re-directing the packets to an alternative destination if so allowed and so indicated by the packet forwarding instructions for packet primarily destined on the given failed route or destination, or detour routing the packet through an intermediate A-M node (on the alternative fiber ring direction) that re-forwards the detour routed packets to their intended primary destination.

In brief, as with traffic engineering, also with traffic protection, A-M, does what the customer wanted the inter-connect network to do, based on clear, documented rules and SLA-guaranteed response times, and in its signature style -- automatically and transparently.

Summary

Adaptive-Mesh L1.5 VPN is an intelligent server layer network that does exactly what the customer routers exchanging traffic through it intended it to do, according to the implicit control by the forwarding instructions on the packets transmitted through it by the customer routers, and through the inter-router traffic volumes, as well as realtime destination reachability and route load status.

At packet layer, A-M is directly controllable via the packet forwarding instructions inserted in the packets by the customer routers and by the agreed-on, well-documented rules for A-M to interpret such forwarding instructions[5]. Moreover, the packet-layer transparent A-M allows the L3/2 nodes that it interconnects to interact with each others directly at all packet layer protocol levels, including any and all packet layer traffic engineering or protection protocols.

Thus, A-M, for instance, allows MPLS routers connected through it to directly negotiate their mutual L3/2 packet-layer Traffic Engineering based bandwidth allocation quotas. As the routers then send traffic to each others through the A-M according to the inter-router TE policies, A-M automatically optimizes the bandwidth allocation among its internal mesh of dynamic bandwidth physical layer (L1) circuits according to the realtime inter-router traffic load variations. Via its use of adaptive-bandwidth L1, A-M allows providing given inter-router throughput requirements, with direct L1 circuit type high, deterministic QoS, and with an up to 20 times architectural cost-efficiency gain compared to any other type of networks, since any other known network technology is limited to non-adaptive L1.

It is thus seen that L3/2 TE and L1 A-M are directly complementary techniques: While TE should generally seek to avoid network congestion (e.g. by coordinating transmissions from a set of sources toward a given destination in order to prevent multiple sources sending more traffic toward the destination than what it can process), since packet traffic however is inherently bursty (each packet even alone is a burst) and since TE is a software technique without data-plane synchronous network control capabilities, there is a need for realtime dynamic capability to optimize network physical layer bandwidth allocation with the packet to packet as well as sub-packet (in case of interleaved transmissions from multiple sources) time granularity. OCS' A-M is designed, and tested, to provide that capability.

---

[4] US patent applications #61/060,905, #61/075,108.

[5] US patent application # 12/390,387 (see in particular Table 1 and related descriptions); http://www.ocsipholding.com/MPLS-TP_for_L1VPNs.pdf

Similarly, for traffic protection, A-M automatically performs, with sub millisecond processing delays, any and all traffic protection re-routing actions desired by the customer in the event of any network line or node failures, while allowing the customers' routers to transact among each others directly at all packet-layer protocols, e.g., to carry out any less time critical changes to their mutual operation due to any possible changes in their mutual reachability.

Accordingly, Adaptive-Mesh automatically carries out the actual intent of any control that could be applied to it by the set of L3/2 nodes it interconnects.

However, instead of requiring any explicit control from the routers that it interconnects, Adaptive-Mesh L1.5VPN performs its packet forwarding directly based on the packet forwarding instructions inserted to the packets by the customer routers (as well as based on the agreed rules for interpreting such instructions, and on realtime reachability of destinations), and continuously re-optimized its capacity allocation via the implicit control of the TE-controlled inter-router traffic volumes.

Contacts:

info@ocsipholding.com

http://www.ocsipholding.com