Traffic Engineering for Ethernet over SONET/SDH: Advances and Frontiers

Chongyang Xie, Nasir Ghani, Qing Liu, Rolande Kouatang, and Wei Wennie Shu, University of New Mexico
Yan Qiao and Minyou Wu, Shanghai Jiao Tong University
Shuping Peng, Peking University

Abstract

Continued growth in data services coupled with the entrenchment of existing carrier infrastructures have inevitably forced a convergence of SONET/SDH and Ethernet technologies. As a result, the very notion of Ethernet connectivity has expanded beyond traditional local area domains into broader metro/wide area settings. The key enabling advances have come in the form of improved SONET/SDH provisioning features as well as new carrier Ethernet services standards. These provisions have allowed carriers to build and deploy much improved multigrade point-to-point Ethernet line services and have also opened up new challenges for extended multipoint offerings. This article looks at this evolving area and summarizes recent advances in EoS traffic engineering. It also highlights some key future challenges and presents a novel scheme for multipoint services provisioning in EoS settings.

Carriers today face increased challenges arising from evolving client data service demands. Most notably, the ubiquity of Ethernet data interfaces on most computing gears has generated immense commercial interest in extending Ethernet services beyond the traditional local area network (LAN) space and on into larger metro and wide area domains [1]. Furthermore, many applications are also demanding much improved carrier-grade service qualities for these evolved data offerings. As a result, the concept of carrier Ethernet has evolved [1], a leading proponent of which has been the Metro Ethernet Forum (MEF), which has developed new standards for both point-to-point and multipoint services [2].

The above evolutions pose notable challenges for carriers seeking to expand their base metro/core infrastructures with Ethernet-ready capabilities. At the same time, however, there is a strong desire to preserve existing legacy infrastructure investments in order to maintain traditional revenue streams (e.g., voice, leased line). Hence the extension of Ethernet services over ubiquitous synchronous optical networking/synchronous digital hierarchy (SONET/SDH) time-division multiplexing (TDM) networks has become a major focus area for the research and industrial communities (i.e., Ethernet-over-SONET/SDH (EoS)). This trend has been given impetus by improved multiservice SONET/SDH standards that resolved many of the earlier bandwidth efficiency and scalability concerns of older legacy TDM systems. Widely termed next-generation SONET/SDH (NGS), these technologies include generic framing procedure (GFP), virtual concatenation (VCAT), inverse multiplexing, and link capacity adjustment scheme (LCAS) [3–5], as shown in Fig. 1. Overall, NGS enables carriers to extend genuine carrier-class service reliability and operation, administration, maintenance, and provisioning (OAM&Ps) capabilities to new data service offerings — a major saliency. Moreover, even though many metro/core backbone networks are moving toward layer 1 dense wavelength-division multiplexing (DWDM) setups [6], most of these networks still rely on TDM framing at the wavelength level with trends toward SONET/SDH-compatible International Telecommunication Union — Telecommunication Standardization Sector (ITU-T) optical transport network (OTN) formats. Hence, revamped NGS technologies can readily serve as much-needed subrate grooming solutions for larger-granularity wavelength pipes (i.e., layer 1.5).

Given the above developments, the focus for many carriers has shifted toward reusing/reapplying SONET/SDH technologies to build higher value-added Ethernet services [1]. Indeed, many new EoS studies have looked at incorporating VCAT-based inverse multiplexing techniques to improve multipath routing efficiency and resiliency levels [7–14]. Meanwhile, others have also studied EoS grooming over SONET-DWDM networks [15–18]. However, the majority of studies in this area have primarily focused on building point-to-point Ethernet services with little or no consideration given to more expansive multipoint-to-multipoint Ethernet services. Clearly these latter types represent more lucrative revenue streams, and require further investigation and design in the context of EoS settings.

This article focuses on traffic engineering (TE) for EoS services and is organized as follows. First, brief overviews of relevant standards developments in NGS and carrier Ethernet are presented. Subsequently, a detailed survey of the latest research work in EoS TE provisioning is presented, and future
open areas and challenges are identified. Finally, a novel scheme is presented for multipoint Ethernet LAN services extension along with associated performance evaluation results. Final thoughts and conclusions are then presented.

Background

New and improved ITU-T standards for streamlining non-TDM support have played a vital role in improving data-over-SONET integration. Concurrently, related developments in MEF carrier Ethernet standards have also pushed service requirements across wider geographic domains. These advances are briefly overviewed here, and interested readers are referred to the related references for more details.

Advances in NGS

The ITU-T has been evolving a comprehensive set of multi-service SONET/SDH features over the last decade to improve protocol transparency and enhance traditional services support (Fig. 1). For example, GFP (ITU-T G.7041) defines highly efficient client interface mappings onto byte-synchronous TDM timeslots for Ethernet, Fibre Channel, ESCON, InfiniBand, and so on. [4]. This scheme uses robust error-controlled frame delineation, like asynchronous transfer mode (ATM), and defines two client signal adaptation modes, frame (GFP-F) and transparent (GFP-T). The former yields deterministic overheads for Ethernet medium access control (MAC) or IP packets, whereas the latter transparently maps 8b/10b-encoded blocks (Fibre Channel, ESCON, FICON) with minimal packetization and buffering delays. Recently the ITU-T has also published amendments to G.7041 to address larger maximum transmission unit (MTU) sizes for newer Ethernet 802.3 interfaces. Meanwhile, VCAT (ITU-T G.707) enables customizing of tributary channels for end-user demands (STS-1, VT1.5 levels). A key provision here is inverse multiplexing [5], which allows splitting payloads into individual channels (subconnections). These channels can then be routed separately and then recombined into a contiguous data signal at the sink side. Finally, LCAS (ITU-T G.7042) reinforces VCAT by increasing/decreasing channel counts in a concatenated trail in a hitless manner. This has notable applications for dynamic TE and service survivability.

A major advantage of the above NGS features is that they are backward compatible with legacy SONET/SDH as well as newer ITU-T OTN standards. Hence carriers need only deploy it selectively as needed (e.g., at client edge-access points) [1]. Note that the Optical Internetworking Forum (OIF) showcased joint VCAT and LCAS operation in a multi-vendor demo at IEC SUPERCOMM 2004 [5]. This setup ran across a global interconnection of seven carrier networks, and successfully achieved dynamic end-to-end connection management between client devices and transport nodes via intelligent control plane mechanisms, such as the OIF user-network interface (UNI) 2.0 and network-network interface (NNI) 1.0.

Carrier Ethernet Standards

The key difference between carrier Ethernet and regular LAN Ethernet is that the former transfers layer 2 frames across service provider networks. Carrier Ethernet literally creates a single network spanning across layer 1 (SONET/SDH, DWDM) or layer 3 (IP/MPLS) networks. Along these lines, the MEF has tabled five carrier Ethernet attributes to define a robust deterministic client experience: standardized services, scalability, reliability, quality of service (QoS), and service management. Furthermore both point-to-point and multipoint service renditions have been defined under three broad categories: Ethernet line (E-Line), Ethernet LAN (E-LAN), and Ethernet tree (E-Tree) [2]. For each of these, both private and virtual private service versions are also specified, yielding a total of six types (Fig. 2). These are briefly overviewed here.

Any Ethernet service using a point-to-point Ethernet virtual connection (EVC) is designated an E-Line Ethernet private line (EPL) type (Fig. 2). This service supports bidirectional connectivity with or without performance guarantees, such as committed information rate (CIR) with an associated committed burst size (CBS), excess information rate (EIR) with an associated excess burst size (EBS), delay, delay variation, loss, and availability for a given class of service (CoS) instance. Now all E-Line services implement all-to-one bundling at the client UNI port (i.e., name port-based) [2]. Conversely, the virtual private extension of this service allows UNI multiplexing, thereby permitting more than one point-to-point EVC to share the same physical port, referred to as Ethernet virtual private line (EVPL), as shown in Fig. 2.

Meanwhile, E-LAN services provide LAN extension via multipoint-to-multipoint EVC connectivity between data sites (Fig. 2). Nevertheless, the detailed mechanisms for implementing such connectivity over layer 1/3 carrier networks are not specified. The latest MEF service, E-Tree (MEF 10.1), also uses a rooted-multipoint EVC definition. The logical overlay topology is defined here (similar to star topology) where the root UNI acts as the hub, and all leaf UNI nodes only communicate with the root. For both of these services, virtual extension is also defined via UNI multiplexing across multiple sites, similar to EVPL.
EoS TE: A Survey

NGS inverse multiplexing and dynamic bandwidth adjustment offer many new avenues for improving multipath protection/restoration capabilities vs. legacy SONET/SDH. Along these lines, researchers have developed various EoS survivability schemes. For example, [7] outlines several low-overhead protection for Ethernet over SONET (PESO) solutions: α, β, γ. The idea here is to ensure adequate immunity (to single link failures) by exploiting multipath diversity between subconnections. Subconnection route overlaps are minimized to limit outages from single link failures (albeit no results are presented). More recently [8] tables similar strategies for degraded-service-aware provisioning. Specifically, subconnections are routed to ensure that no one path carries more than a given fraction of the total flow, and multipath load distribution is also used to minimize the maximum incremental link utilization via integer linear programming (ILP).

Although the results here show good improvements in blocking and load balancing, these strategies can be susceptible to limited topological connectivities, i.e., multi-path disjointness. Finally, [9] also develops a new effective multi-path bandwidth metric to account for link bandwidth and availability constraints. Here two multi-path routing heuristics are developed to achieve desired availability levels, with both showing significant improvements vs. single path strategies.

To better address topological concerns, other studies have looked at direct subconnection protection/restoration strategies. For example, [10] proposes a tiered (partial) protection scheme that implements dedicated protection for a subset of working subconnections. Simulation results show notable efficiency gains and very high recovery rates with the use of post-fault restoration (even for lower protection thresholds). Meanwhile, [11] proposes two inverse multiplexing shared protection schemes, protecting individual virtual connection group members (PIVM) and provisioning a fast restorable virtual connection group (PREV). The former allows backup capacity sharing between link-disjoint subconnections, whereas the latter limits sharing to link-disjoint subconnections with the same source-destination. As expected, PIVM gives much higher efficiency while PREV gives much faster recovery. However, since the PIVM scheme requires complex per-link conflict state, it is only amenable to centralized control. Meanwhile, the PREV scheme is more susceptible to networks with reduced connectivity/path diversity.

Various studies have also looked at differential delays for multipath routing. For example, [12] develops the cumulative differential delay routing (CDDR) problem, focusing on resolving an integral number of subconnection paths to limit destination (sink) memory requirement below a given threshold. Note that sink-side buffering is directly proportional to bandwidth-delay products along end-to-end paths. The CDDR problem is shown to be both NP-complete and provably hard to estimate within a constant factor. Hence two heuristics are proposed and their effectiveness in reducing sink-side buffering demonstrated via simulation. Meanwhile [13] develops a modified link-weight k-shortest path algorithm that takes into account the lower bounds on differential delay. The cost of each link is computed by combining the original link weight and the inverse of the weight. Simulation results show that by bounding the cost function with upper (shortest path) and lower (longest path) costs, the scheme can outperform more traditional algorithms. Various other differential delay distribution schemes have also been proposed to lower sink-side buffering. For example, [14] proposes smaller levels of distributed buffering along intermediate path nodes to lower sink-side memory needs. Specifically, the scheme artificially increases the delays along shorter subconnection paths to help reduce the total differential delay. Simulation results show significantly reduced network-wide buffering overheads.
Due to close association of SONET/SDH and optical technologies in the metro/core space, researchers have also studied EoS grooming over DWDM networks. For example, [15, 16] propose basic schemes to resolve incoming requests to the STS-1 level and table modified shortest-path heuristics for multipath routing. Specifically, [15] proposes a methodology for dynamic routing of fractional-wavelength traffic in WDM grooming networks and also develops a new performance metric that indicates the number of links in which a particular wavelength is not used. Performance evaluation is carried out for shortest-widest path, widest-shortest path, and other available shortest path routing algorithms. The findings here reveal a counterintuitive result: increasing grooming capability in a network can result in degraded performance for widest-shortest algorithms. Meanwhile, [16] proposes several iterative graph heuristic algorithms for multipath subconnection routing: shortest path first (SPF), widest path first (WPF), and maximum flow (MF). Findings here show that in most cases, the MF algorithm gives the best performance. However, SPF can outperform WPF under high traffic loads. In addition, it is found that inverse multiplexing yields good blocking reduction (about 10–15 percent lower than regular non-inverse multiplexed routing) and can compensate for reduced grooming intermediates in multicast tree core. Also, [17] studies Ethernet grooming in optical networks with mixed line rate (MLR) links. The authors argue that MLR networks can yield better performance over single line rate (SLR) networks, and present and analyze an efficient routing algorithm for such settings. Note that the layering of Ethernet-SONET/SDH-DWDM technologies may introduce additional processings delays and require specific differential line treatments (discussed earlier). Finally, survivable grooming of subrate VCAT streams over DWDM networks is studied in [19] and two graph-heuristic methodologies are tabled, protection at connection (PAC) and protection at lightpath (PAL). The results here show that the former is most efficient for high grooming port counts, whereas the latter is most efficient for low–moderate grooming port counts.

**Open Challenges**

Although the above contributions represent many notable innovations for EoS, many further challenges still remain. Foremost, MEF services such as E-LAN or E-Tree (Fig. 2) require multipoint-to-multipoint connectivity over metro/wide-area sites [2]. However, the inherent lack of time slot multiplexing in SONET/SDH poses some notable complexities here. Therefore, new solutions are required to achieve coordinated setup of multiple point-to-point TDM connections to emulate multipoint-to-multipoint connectivity: connection groups or topology overlays. Note that the very nature of these overlays will depend on higher-layer layer 2 switching capabilities, either within the carrier network or at the client side.

To date, however, few studies have looked at overlay design (connection group provisioning) for Ethernet LAN services in the context of SONET/SDH networks. Perhaps the only related work is in the area of overlay networks [18, 20–22], studying the design of specialized topologies over physical substrates to support end-to-end applications such as voice over IP (VoIP), video on demand (VoD), multicasting, and peer-to-peer file sharing. For example, the resilient overlay network (RON) [20] project builds an Internet overlay to improve routing resiliency and overcome border gateway protocol (BGP) convergence limitations. Meanwhile, the service overlay network (SON) [21] study addresses QoS support using queuing theory and optimization to handle static partitioning and oversubscription. However, the above efforts have only addressed longer-term preplanned overlay designs. Alternate provisioning of more dynamic/online overlays is more complex, and some studies have also emerged here. For example, [18] proposes a virtual network (VN) assignment solution for node/link selection with and without reconfiguration. A novel stress ratio concept is introduced, and heuristic and adaptive optimization techniques are proposed. Findings here show notable improvements with moderate reconfiguration, particularly with sparse connectivity. Although the above studies present some key innovations, additional avenues exist for new carrier Ethernet LAN services. In particular, there is a need to study commensurate overlay provisioning schemes for multipoint services within the context of demand splitting (inverse multiplexing) to achieve multiplexed full and fractional resiliency. This issue is addressed further in the next section.

Furthermore, given the expanding footprint of carrier Ethernet services, client data will likely start to traverse across multiple distributed carrier domains. In general, these settings will preclude the availability of global resource state for obvious scalability and (intercarrier) trust reasons. Hence, there is an emergent need to provision services over multiple Ethernet (as well as SONET/SDH and DWDM) domains. This is a largely unexplored area, and it brings to the fore some critical challenges (resource abstraction, policy enforcement, path computation (limited partial state), and survivability (protection, restoration) [6]. Although many studies have looked at multidomain IP and even DWDM networks (see survey in [6]), EoS setups introduce an additional distributed grooming dimension. However, very few studies have addressed this space. For example, [22] presents a Ethernet-over-DWDM scheme that uses integrated multidomain signaling to simultaneously provision lightpaths and groom point-to-point Ethernet connections. However, underlying DWDM topologies are treated as infrastructural and hence assumed to be static (i.e., precluding the need for dynamic multilayer routing). Although this may suffice in limited DWDM settings, the inherent flexibility of SONET/SDH and NGs domains will mandate more active global state exchange.

Overall, much work remains to be done here.

Finally, post-fault subconnection reconconfigurability can be considered to improve resource efficiencies and lower blocking rate in EoS settings. Related studies on lightpath reprovisioning in DWDM networks have shown good results, and these paradigms can also be applied in EoS settings. However, few EoS schemes make adequate use of LCAS capabilities for post-fault restoration as of now.

**Multipoint EoS Services**

A novel provisioning scheme is now presented for multipoint Ethernet extension in EoS settings (see also [23]). This solution uses graph-theoretic algorithms and can readily be integrated into centralized or distributed (generalized multiprotocol label switching, GMPLS) constraint-based routing (CIR) engines [3]. A key innovation here is the concept of a connection group overlay to implement multipoint connectivity between LAN sites, either mesh or star-based (Fig. 3). The scheme also implements tiered survivability by using both prefaul path protection and optional post-fault path restoration mechanisms. Consider the notation first. A SONET/SDH network is represented by a graph $G(V,L)$, where $V$ is the set of nodes and $L$ is the set of links: $V = \{v_1, v_2, \ldots, v_N\}$ and $L = \{l_{ij}, l_{ij}, \ldots, l_{ij} \}$. Link $l_{ij}$ connects nodes $i$ and $j$ and has $c_{ij} \leq C$ STS-1 units of available capacity (since DWDM links are bidirectional, there is also an equivalent link $l_{ij}$). Now consider the $i$th multipoint Ethernet LAN request for $x_i$ STS-1 units to interconnect a subset of nodes, $v_{i} \subseteq V$. The
proposed scheme first transforms this request into a connection group overlay comprising a set of (overlay link) connections between nodes in $v_i$. This group comprises $n_i$ bidirectional connections, $(s_i, d_i)$, where $s_i = (s_{i1}, s_{i2}, \ldots) \subseteq v_i$ and $d_i = (d_{i1}, d_{i2}, \ldots) \subseteq v_i$ represent the endpoint nodes, and the individual connections are $s_{i1}-d_{i1}$, $s_{i2}-d_{i2}$, and so on. Each connection in the overlay is split into VCAT subconnections using an inverse multiplexing factor $K$. Although various splitting strategies are possible, herein an even distribution approach is used over $K$ subconnections (see [10] for details). Finally, the resolved working subconnections are provisioned using iterative shortest path routing techniques. A LAN protection factor, $p$, $0 \leq p \leq 1$, is also defined here to provide tiered (full/partial/none) protection against single link failures. Details are now presented.

**EoS Overlay Selection**

Connection group overlays define the interconnectivity between LAN nodes in $v_i$ and are very dependent on layer 2 switching functionalities in network or client nodes. In the most basic case with insufficient layer 2 Ethernet switching capabilities, carriers can resort to mesh overlays (Fig. 3) to set up TDM connections between all nodes in $v_i$ to deliver LAN-like connectivity. This overlay approach is brute force and resource-intensive (and susceptible to low topological node degrees), generating $O(|v|^2)$ connections (or $O(K|v|^2)$ VCAT subconnections) and yielding higher LAN request blocking rates. Alternatively, if carrier and/or client nodes possess adequate Ethernet switching capabilities, more resource-efficient star overlays are more germane (E-tree service, Fig. 2). A designated Layer 2 hub site, $h_i$, can provide connectivity to all other LAN nodes, yielding $O(|v|^1)$ overlay connections (or $O(K|v|^1)$ VCAT subconnections), as described recently in [23]. However, hub selection is a major concern here, and it is generally desirable to choose a site that minimizes resource utilization and lowers blocking for future requests. Hence, three different hub selection schemes are proposed here.

**Random Hub Selection** — This basic scheme (RS) randomly selects a hub site from the subset of LAN connection nodes (sites) that have Ethernet switching capabilities. Thus, $h_i$ is chosen using a uniform distribution between 1 and $|v_i|$.

**Minimum Average Hops** — This scheme chooses the hub site with minimum average hop (MAH) count distance to all other LAN nodes in $v_i$. The goal here is to minimize LAN overlay resource utilization as follows:

$$h_i = \{v_j\}_j$$ where $j^*$

$$j^* = \min \left\{ \sum_{k=1}^{K} \text{hop}(v_{ij}, v_{ik}) \right\} \forall j, 1 \leq j \leq |v_i|$$

(1)

Since this scheme only uses static topology information, computational speedup can be attained by simply precomputing and storing the hop counts between all possible nodes. $O(|v_i|(|v_i| - 1)) = O(|v_i|^2)$ shortest path computations are required prior to startup along with $O(|v_i|^2) = O(|v_i|^2)$ storage overheads.

**Minimum Average Cost** — This scheme chooses the hub site with minimum average cost (MAC) to all other nodes in the LAN group, where the link cost is defined as inversely proportional to the available link capacity, so for $h_i$

$$\omega_y = \frac{1}{c_y + \epsilon}$$

where $\epsilon$ is a small quantity. Hence, hub selection is given as

$$h_i = \{v_j\}_j$$ where $j^*$

$$j^* = \min \left\{ \sum_{k=1}^{K} \text{dist}(v_{ij}, v_{ik}) \right\} \forall j, 1 \leq j \leq |v_i|$$

(3)

This scheme uses dynamic resource state to avoid congested hubs and links (albeit yielding longer routes). In general, the computational complexity is significantly higher: $O(|v_i|(|v_i| - 1)) = O(|v_i|^2)$ shortest path computations needed for per request between $v_i$ nodes, yielding $O(K^2 \cdot |v|^1 \log |v|)$ complexity. However, note that in sparsely connected graphs, the shortest path from a node to all other nodes can be computed in $O(|v|^1 \log |v|^1)$, yielding a reduced complexity of $O(|v|^1 \log |v|^1)$.

Note that the formulations in Eqs. 2 and 3 generally assume that all sites have layer 2 switching capabilities, and these can easily be tailored for more sparse switching scenarios. Furthermore, alternate overlay topologies are also possible, but are not considered here as they are more costly and will require layer 2 switching and QoS functionalities at all LAN sites; for example, ring overlays for Ethernet resilient packet ring (IEEE P802.17), packet add-drop, and generalized minimum spanning tree overlays.

**Multilayer Overlay Survivability**

Subsequent to overlay generation, detailed connection group provisioning is done. This step comprises two key phases, working and protection setup, both of which use iterative graph-theoretic algorithms. First, the working phase loops through each of the $n_i$ connections in an overlay group and resolves it into $K$ subconnections. Next, successive Dijkstra’s shortest path iterations are used to compute individual subconnection routes (only considering links with sufficient capacity, i.e., $C_k \geq x_i$). If any subconnection cannot be routed, the LAN
request is dropped. Furthermore, two different link cost metrics are used here, hop count and load balancing (i.e., minimum cost, Eq. 2), akin to those used in star hub selection. Specifically, the former pursues resource minimization, while the latter pursues route diversity (across lightly loaded links). Although delay-based metrics are also possible, they are not considered here as most VCAT implementations provide receiver-side buffering for up to 256 ms, more than adequate for large backbone settings.

After successful provisioning of all working group connections, the protection stage iterates over all the overlay connections to implement dedicated protection on a per subconnection basis, similar to [10]. Link-disjoint protection paths are computed for a minimal subset of working subconnections until the desired (LAN) protection threshold of $\rho_x$ STS-1 units is achieved. This approach places no restrictions on the topological overlap between connections (apart from interconnection establishment costs for connection pairings), thereby reducing the need for sufficient multipath diversity in the network. Moreover, it also simplifies protection switchovers during link failures as working and protection subconnections are of equal size (obviating the need for complex edge buffering). Finally, since protection is done on a per subconnection basis, protection granularity is inversely proportional to $K$. Overall, the above algorithms yield $O((V^3)\log(V))$ complexity, assuming Dijkstra complexity of $O((V^2)\log(V))$ and a maximum of $O(V^2)$ connections in an overlay group (mesh).

Performance Analysis
A multipoint LAN provisioning scheme is evaluated using discrete simulation with OPNET Modeler. All tests are done over the 16-node NSFNET topology (average node degree of 3.125) with OC-48/STM-16 (2.544 Gb/s) link rates. LAN group sizes range from 3–5 nodes and are randomly chosen using uniform distributions. Furthermore, LAN requests have exponentially distributed holding and interarrival times with means $\mu$ and $\lambda$, respectively. Specifically, a scaled value of $\mu = 600$ s is used here, and $\lambda$ is adjusted according to desired load. Additionally, LAN bandwidth sizes are varied from 200–1000 Mb/s in 200 Mb/s increments (4 STS-1 units) to model fractional Ethernet demands. Modified Erlang load metrics are also defined to account for differing LAN overlay group sizes:

$$\text{Modified load (mesh overlay)} = \sum_{n=x_1}^{x_2} \frac{n(n-1)}{2} \frac{\mu}{\lambda}$$  \hspace{1cm} (4)$$

$$\text{Modified load (star overlay)} = \sum_{n=x_1}^{x_2} (n-1) \frac{\mu}{\lambda}$$  \hspace{1cm} (5)

where the LAN groups range in size from $x_1 = 3$ to $x_2 = 5$ nodes, and the $1/\lambda$ represents the mean interarrival rate (i.e., inverse of interarrival time). Finally, all runs are averaged over 500,000 randomly generated LAN requests.

Initial tests (Figs. 4a and 4b) plot the carried load for nonprotected LAN scenarios (i.e., $\rho = 0$) for increased inverse multiplexing factors given a nominal request blocking rate of 2 percent. These scenarios are very relevant to carriers as they indicate the true load carrying (revenue generation) capacity of the network at a typical operating point. As expected, the star overlays yield the highest carried load due to smaller connection group sizes. Furthermore, the load balancing (minimum distance) routing approach also yields much higher carried loads for mesh overlays, about 50–70 percent higher (Fig. 4a). Commensurate carried load gains with star overlays (and load balancing) are lower, however, averaging about 7–12 percent. Another key finding is the sizeable improvement with increased levels of inverse multiplexing, about 30–50 percent (25 percent) higher carried load for mesh (star) overlays. These results also show that the more intelligent hub selection strategies (for star overlays) yield the best overall gains. For example, the MAH scheme gives the highest carried load, averaging 17 percent higher than the minimum distance-matching average cost scheme and over 40 percent higher than RS.

Next, LAN blocking is tested for the more effective star overlays and load balancing routing. The results for varying protection thresholds ($\rho = 0, 0.25, 0.5$) and inverse multiplexing factors ($K = 1, 2, 4$) are shown (Fig. 5a). Here it is seen that increased protection factors significantly increase request blocking. In addition, it is seen that increased inverse multiplexing (demand splitting) yields notable gains for equivalent
protection factors; for example, $K = 4, \rho = 0.5$ gives about 10–30 percent lower blocking than $K = 2, \rho = 0.5$ (Fig. 5a).

Furthermore, a detailed look at the individual blocking rates for different LAN sizes ($K = 4, \rho = 0.25$) shows that larger 5-node LAN groups experience almost an order magnitude higher blocking than smaller 3-node LAN groups (Fig. 5b).

This discrepancy is clearly due to the difficulties in routing a larger number of working (and diverse protection) connections at a hub site (i.e., limited by topological connectivity/node degree). However, the hop-based MAH selection scheme still gives the lowest blocking for all LAN sizes (at least 40 percent lower at most loads) and even achieves under 1 percent blocking for 5-node LAN requests at low loads. Indeed, this blocking level is sufficient for operational settings and cannot be achieved with the other star and mesh overlays.

Finally, tiered LAN protection is tested in conjunction with post-fault restoration of nonprotected working VCAT subconnections on a failed link. The goal here is to achieve full recovery for partially protected (lower tier) LAN services. All link failures have exponentially distributed interarrival and mean time to repair (MTTR) values with mean 600 s. Note that the MTTR values represent times for truck roll repairs and are typically much larger than protection switchovers.

Restoration performance is gauged by measuring recovery rates for carried loads up to 20 percent blocking (i.e., high loading). The restoration rate is defined as the percentage of LAN groups that recover all failed nonprotected connections, that is, regain full 100 percent post-fault throughput.

LAN restoration for star overlays with varying protection thresholds and inverse multiplexing factors is also tested (MAH selection and load balancing routing, Fig. 6). $K = 1, \rho = 1$ corresponds to regular noninverse multiplexed protection (e.g., 1+1, 1:1), whereas $\rho = 0.25$ and 0.5 correspond to partial protection.

The results here reveal some important findings. First, as expected, full protection provides 100 percent recovery, although the carried load in this case is much lower. More important, for partial protection it is seen that increased levels of resource overprovisioning (protection) do not necessarily translate into higher levels of post-fault restoration. For example, $K = 4, \rho = 0.25$ gives higher recovery rates than $K = 2, \rho = 0.5$. Meanwhile, the maximum carried load (for up to 20 percent LAN request blocking) here is also almost two times higher. These findings indicate that post-fault subconnection restoration is highly beneficial and will allow carriers to offer full (100 percent) LAN recovery for well over 95 percent of single link failure events. In turn, these gains will lower pre-reservation levels and allow operators to design new services with increased revenue potential.

Conclusions

Advances in NGS coupled with growing client data networking demands have generated intense interest in Ethernet services extension over larger metro/core distances. To address
these challenges, comprehensive new carrier Ethernet standards have emerged, defining new services for point-to-point and varied multipoint services. Concurrently, many researchers have studied EoS provisioning, with a focus on leveraging the unique inverse multiplexing capabilities for improved TE resource efficiency and service resiliency. However, most studies have only focused on point-to-point Ethernet offerings and the more challenging problem of multipoint connectivity support for EoS remains unaddressed. This article surveys this crucial area and also presents results from a novel multipoint EoS provisioning scheme. Findings show the best blocking and carried load performances for more efficient star overlays and load balancing routing.

Acknowledgments

This research has been supported in part by the U.S. National Science Foundation (NSF) Computer and Network Systems (CNS) division under award number CNS 0806637 and the UNM ECE Department. The authors are very grateful to these sponsors for their generous support.

References


Biographies

QING LIU received a B.E. degree in telecommunications engineering and an M.E. degree in computer science from Nanjing University of Post and Telecommunication, China, in 2001 and 2004, and a Ph.D. degree in electrical engineering from the University of New Mexico in 2008. In the past he has worked as a network engineer at Softbank Telecom, Beijing, China, and also at China Mobile, Shenzhen. Recently, he also worked as a research and development engineer intern at Innovasic Semiconductor, Shanghai, China, where he is involved in various funded research projects. He has a wide range of industrial and academic experience in the networking area, and has held senior positions at Nokia, IBM, Motorola, Sorrento Networks, and Tenessee Tech University. He is a recipient of the NSF CAREER Award and has authored over 100 publications (including journal and conference papers, book chapters, industry magazine articles, and standardization drafts). He is a co-chair of the IEEE INFOCOM High-Speed Network (HSN) series, and has served as a co-chair for symposia at IEEE ICC 2006 and IEEE GLOBECOM 2006 and 2009. He is also a program committee member for IEEE/OSA OFC and numerous other conferences, and has served regularly on NSF, DOE, and international panels. He is an associate editor of IEEE Communications Letters and has guest edited IEEE Network, IEEE Communications Magazine, and Cluster Computing.

CHONGYANG XIE received B.E and M.E degrees, both in electronic engineering, from South China University of Technology, Guangzhou, China, in 2004 and 2007, respectively. He is currently pursuing a Ph.D. degree in the Department of Electrical and Computer Engineering at the University of New Mexico. He has held a number of key roles in key projects in the UK, China, and Canada. His recent projects include traffic engineering, advanced reservation, and multilayer and multidomain networking.

NASIR GHANI [SM] received a Bachelors degree in computer engineering from the University of Waterloo, Canada, in 1997, a Masters degree in electrical engineering from McMaster University, Canada, in 1992, and a Ph.D. degree in electrical and computer engineering from the University of Waterloo in 1997. Currently, he is an associate professor in electrical and computer engineering at the University of New Mexico, where he is involved in various funded research projects. He has a wide range of industrial and academic experience in the networking area, and has held senior positions at Nokia, IBM, Motorola, Sorrento Networks, and Tennessee Tech University. He is a recipient of the NSF CAREER Award and has authored over 100 publications (including journal and conference papers, book chapters, industry magazine articles, and standardization drafts). He is a co-chair of the IEEE INFOCOM High-Speed Network (HSN) series, and has served as a co-chair for symposia at IEEE ICC 2006 and IEEE GLOBECOM 2006 and 2009. He is also a program committee member for IEEE/OSA OFC and numerous other conferences, and has served regularly on NSF, DOE, and international panels. He is an associate editor of IEEE Communications Letters and has guest edited IEEE Network, IEEE Communications Magazine, and Cluster Computing.

ROLANDE KOUSANGIAN received her Laurea in information engineering from the University of Genova, Italy, in 2005 and a Master’s degree in computer engineering from the University of New Mexico in 2008. From May 2006 to May 2008 she was a research assistant at the University of New Mexico. Her research interests are in the general area of optical networks (physical layer) and include carrier Ethernet services, next-generation SONET networks, routing, traffic engineering, multilayer networks, IP/ Ethernet networks, ATM, MPLS/GMPLS, path computation, survivability, simulation, and performance analysis.

WEI WEINHE SHU [SM] received her Ph.D. degree from the University of Illinois at Urbana-Champaign. Since then she has worked at Yale University, the State University of New York at Buffalo, and the University of Colorado at Boulder. Currently she is an associate professor in the Department of Electrical and Computer Engineering at the University of New Mexico. Her current interests include multimedia networks, distributed systems, and wireless networks.

YAN QIAO will receive her B.S. degree in computer science from Shanghai Jiao Tong University in the summer of 2009. She worked as a summer intern researcher at the University of New Mexico in 2008 and is currently planning to attend graduate school. Her interests include computer networks and algorithm design.

MENYOU WU [SM] is a professor in the Department of Computer Science and Engineering at Shanghai Jiao Tong University and a research professor with the University of New Mexico. His research interests include wireless and sensor networks, multimedia networking, and parallel and distributed systems. He has published over 180 journal and conference papers in the above areas. His research has been supported by the National Science Foundation, DoD, DoE, DARPA, China 863 and 973 programs, and Natural Science Foundation of China. He is an editor of the Communication of China Computer Federation.

SHUHRUH ZHAO received her B.S. degree in physics from Jilin University, Changchun, China, in 2005. Subsequently, she was recommended for admission to directly pursue a Ph.D. degree at Peking University in the same year. Currently, she is a doctoral candidate in the State Key Laboratory of Advanced Optical Communication Systems and Networks, Peking University, Beijing, China. Her research interests include optical networks, optical networking, and TCP over optical networks. Her papers have been accepted and published in the IEEE/OSA Journal of Lightwave Technology, IEEE/OSA OFC, and IEEE ICC, among others.

IEEE Network • May/June 2009

Authorized licensed use limited to UNIVERSITY OF NEW MEXICO. Downloaded on June 1, 2009 at 18:01 from IEEE Xplore. Restrictions apply.