Integrated Traffic Grooming in Converged Data-Optical Networks

N. Ghani*, C. Assi**, A. Shami*, M. Ali*

*Tennessee Tech University, **Concordia University
†University of Western Ontario, City University of New York

Abstract: Optical dense wavelength division multiplexing (DWDM) has yielded unprecedented levels of bandwidth scalability. In order to exploit these gains, new converged multi-service transport setups have been evolved, most notably under the multi-protocol label switching (MPLS) and generalized MPLS (GMPLS) frameworks. These paradigms offer very efficient data-optical integration and enable a host of new service capabilities. As operators deploy these new technologies, the provisioning of “sub-wavelength” demands over wavelengths has become a crucial requirement, i.e., traffic engineering/grooming. This paper addresses data-optical grooming in converged GMPLS networks. Here, novel integrated constraint-based routing paradigms are developed to provision sub-wavelength demands at both packet-switching and lightpath routing levels. Simulations indicate notable performance gains and resource efficiencies with the proposed schemes.

1. Introduction

Advances in optical dense wavelength division multiplexing (DWDM) technology and wavelength routing paradigms [1] have delivered tremendous gains in bandwidth scalability and service flexibility. At the same time, data traffic has grown steadily, largely usurping legacy voice traffic as the dominant form on most carrier networks. Collectively, these developments have had profound impacts on the design of transport architectures [2][3]. Earlier, data services support over legacy transport networks required complicated “multi-layering” setups which comprised up to four protocol layers, e.g., IP/Ethernet, asynchronous transfer mode (ATM) or frame relay (FR), SONET/SDH, and optical fiber or DWDM. However, as is well-known, these setups proved very inefficient and costly for fielding increasing data traffic demands [2]. Consequently, network designers evolved towards improved “converged” paradigms for streamlined data transport capabilities.

In particular, there have been notable developments in new, collapsed data-optical transport architectures. The key catalyst here has been the multi-protocol label switching (MPLS) [2] framework, which was originally conceived to provide quality of service (QoS) capabilities in packet switching networks. MPLS provides a comprehensive protocols suite for “soft-circuit” provisioning, e.g., resource discovery, signaling, routing, etc. Since many of these broader functionalities are also required in other emerging network layers, particularly DWDM, designers gradually enhanced MPLS to handle “non-packet” layer (i.e., circuit-switching) technologies. Termed as generalized MPLS (GMPLS) [2][4], this expanded framework offers a unified, intelligent control solution for optical domains. Hence, data transport infrastructures are rapidly moving towards a “two-tiered” setup comprising high-speed IP/MPLS routers interconnected via dynamic GMPLS-based DWDM switches [2][5]. The former elements implement finer-granularity packet-based control whereas the latter deliver scalable wavelength interconnectivity. Overall, converged MPLS/GMPLS paradigms yield a much simpler and lower-cost data transport solution and effectively subsume key “multi-layering” functionalities, e.g., protection in optical layer, traffic engineering in IP/MPLS layer, etc.

As carriers deploy new MPLS/GMPLS networking technologies, some crucial provisioning concerns need to be addressed. In particular, consider the fact that DWDM nodes can only provision capacity in very large increments, i.e., multi-gigabit wavelengths (usually 2.5 or 10 Gbps), yet end-user packet demands predominantly comprise smaller “sub-wavelength” increments. For example, many packet service offerings provision capacity in 50 Mbps increments or less. This spread yields a sizeable “granularity gap” between the data routing and optical DWDM layers and mandates the need for efficient traffic grooming/engineering [3][5][6] solutions to boost resource/bandwidth efficiencies. Clearly, packet-based statistical multiplexing, as performed by IP/MPLS routers, is the best form of grooming within these domains.

Traffic grooming algorithms have become a key research focus today [3][5][7]. Ideally, these algorithms must deliver notable throughput gains and achieve a good balance between resource utilization (i.e., cost) at the data and optical layers. In this study, the problem of sub-wavelength (packet flow) grooming in “two-tiered” data-optical domains is treated and novel integrated schemes are proposed for routing demands at the packet-switching (IP/MPLS) and lightpath routing (GMPLS) layers. In addition, various combined metrics are also derived in order to minimize network resource consumption and balance network-wide loadings. Note that that many
other studies have also looked at sub-rate SONET/SDH traffic grooming over DWDM layers, see [3],[7] for surveys. However, this work either pertains to legacy SONET/SDH and/or multi-protocol support, as facilitated by ongoing evolutions in dynamic "next-generation" SONET/SDH technologies [3]. Since these scenarios represent more specialized technologies, the focus here is strictly on packet-based data-optical grooming.

The remainder of paper is organized as follows. Section 2 briefly reviews the GMPLS framework and then Section 3 tables an integrated wavelength/sub-wavelength provisioning framework for converged data-optical networks. Detailed simulation and performance analysis results are then presented in Section 4 for a variety of realistic testcase scenarios. Finally, conclusion and future research directions are discussed in Section 5.

2. GMPLS Overview

GMPLS is an evolution of MPLS in which labels are abstracted to represent a wider range of bandwidth entities, e.g., wavelengths, bands, SONET/SDH timeslots [2]. GMPLS also allows the control plane to be physically separate from the data plane. As such, this flexibility enables the concept of a generalized LSP (GLS) which can be used to represent a wide range of circuit connections, e.g., lightpaths, TDM circuits, etc. From a protocols perspective, GMPLS also extends advanced MPLS provisions—resource discovery, rapid signaling, constraint-based routing—by defining "technology-specific" extensions to MPLS protocols, e.g., RSVP, CR-LDP, OSPF [2],[4]. For example, routing enhancements carry optical resource information such as available wavelengths, switching ports, converter resources, etc. Based upon these additions, several data-optical integration models have been proposed, i.e., overlay, augmented, and peer (the key differences between these types pertain to the specifics of the protocol instances running at the respective layers [2]).

Since this study focuses on connection provisioning, it is instructive to take a closer look at MPLS/GMPLS signaling and routing. Here, connection provisioning—packet LSP or wavelength lightpath—entails two key functions, namely route computation via constraint-based routing (CBR) algorithms and setup via distributed signaling. Now depending upon the network layer involved, the related CBR algorithms will vary significantly. These algorithms are designed to specifically optimize various performance goals (e.g., maximizing utilization, minimize cost, etc) and the algorithmic details are treated in Section 3. Meanwhile, once a route has been (partially) determined, related signaling procedures are invoked to reserve the bandwidth resources, e.g., router port capacities, OXC wavelengths/ports, etc. Here, only RSVP-TE signaling [2] is considered as it is most commonplace. RSVP-TE basically uses a "downstream-on-demand" mode in which upstream nodes make explicit label requests (Path messages) and downstream nodes assign particular labels in response (Resv messages). Thus, connection setup consists of a label request traversing hop-by-hop from the source to the destination, followed by a label assignment traversing in the reverse direction. Note that RSVP-TE allows source nodes to specify the CBR-computed route in the Path message via explicit routing. In addition, upstream nodes can also restrict the set of labels that a downstream node can choose from—particularly important in optical domains where wavelength conversion limitations may affect channel selections.

3. Integrated IP-DWDM Routing

Sub-wavelength provisioning in converged data-optical domains is a key issue, particularly since most client demands only amount to a fraction of wavelength capacity. Clearly, assigning full lightpath connection(s) between all possible router nodes, i.e., N^2 mesh, is not feasible and may easily lead to "wavelength exhaust" [3]. Alternatively, provisioning lightpath(s) between adjacent routers and relegating all switching to the packet layer will yield excessive electronic complexity (i.e., IP hop-by-hop, [6]). Hence, the respective CBR algorithms at each layer have to be carefully integrated to yield acceptable resource efficiencies (lower costs). This can be achieved via multi-layer traffic grooming [3],[5],[6] techniques, which exploit intermediate packet routing nodes to selectively aggregate sub-wavelength demands onto wavelengths, Figure 1. Overall, grooming can yield notable throughput gains yet still lower wavelength requirements and electronic packet-switching costs [5].
Packet grooming setups assume that router nodes are interconnected via DWDM lightpath entities. These architectures essentially induce a logical (i.e., virtual) topology over the physical topology. Hence an end-to-end path at the IP/MPLS layer is either a set of consecutive lightpaths traversing intermediate routing nodes and/or a direct lightpath between the source and destination routers. In the more common case of multiple lightpaths, packet traffic will have to be terminated and processed by intermediate routing nodes (attached to the optical switches where the lightpath is terminated). This is shown in Figure 1, where router C terminates all local packets and further aggregates (i.e., grooms) "transit" packets onto output lightpath(s). Note that packet label swapping procedures are required here.

The proposed grooming setup assumes that each network layer (data, optical) runs its own independent set of MPLS/GMPLS-based protocols, i.e., overlay model [2],[6]. Namely, the DWDM layer is provisioned via GMPLS protocols and related lightpath connections are essentially G-LSP entities. Similarly, the packet layer is provisioned using MPLS protocols and related LSP paths are encapsulated in G-LSP entities. Hence each node—IP/MPLS router or DWDM switch—runs separate instances of the respective signaling (RSVP-TE) and IGP routing (OSFP, IS-IS) protocols. Each node therefore maintains detailed information on connectivity/resource levels at its respective layer, and this information is used by its CBR algorithms. For the DWDM layer, this information commonly comprises wavelength resources, cross-connection capacities, wavelength converters, etc. Meanwhile, at the IP/MPLS packet layer, related information includes available capacity on logical links (lightpaths) and router switching capacities, etc. Overall, this decoupling between the data routing and optical protocols is well-suited for practical settings, since most carriers maintain clear operational delineations between large scale routing and transport domains.

Now a crucial component in CBR/traffic engineering is the connection routing algorithm. Here a novel integrated scheme is proposed to coordinate resource allocation between the data and optical layers at setup time, Figure 2. Namely, incoming sub-wavelength (packet) requests first initiate LSP setup signaling between IP/MPLS routers. Assuming RSVP-TE Path/Resv messaging, these procedures follow a hop-by-hop "downstream-on-demand" mode where an upstream IP/MPLS routers makes explicit label requests and downstream nodes assign labels. Note that intermediate nodes can only assign a label to their upstream neighbors if they have received a label from their respective downstream nodes, i.e., next hop. If there are adequate resources to field the request and the signaling procedures are successful, the connection request is granted. However, under heavily loaded scenarios, connection blocking can easily occur due to lack of available switching and/or logical link capacities. In this case, the source routing node must invoke new lightpath setup requests in order to generate more logical-layer interconnection capacity, e.g., GMPLS signaling at optical layer (Figure 2). Overall, this approach maximizes resource utilization at the packet routing layer before resorting to adding wavelength capacity. Broadly, this is inline with general grooming principles that focus on conserving scarce optical (wavelength) resources [3].

Now consider the actual CBR provisioning algorithms for routing connection requests at the respective network layers. Here, shortest path routing techniques have been extensively used for traffic engineering in distributed networks—at the wavelength and sub-wavelength layers [1],[5]-[9]. These schemes define a representative graph model of the network and run appropriate shortest path algorithms, e.g., Djikstra. In dynamic connection scenarios it is assumed that resource databases are continually updated with the latest changes, e.g., IGP updates. Clearly, the many intricacies of the respective DWDM optical and IP/MPLS packet-switching layers will mandate specialized renditions of these algorithms. These are now considered.

Figure 2: Joint wavelength/sub-wavelength provisioning

3.1 Lightpath Circuit Routing

DWDM networks use routing and wavelength assignment (RWA) [1],[2] algorithms to determine lightpath routes (Figure 4). In essence RWA is a specialized case of circuit-routing, in which wavelength (i.e., lightpath color) limitations have to be taken into account. In particular, RWA algorithms entail two key...
steps, lightpath route resolution and wavelength allocation. To date, the issue of RWA has been studied extensively, see [1],[8],[9],[11]. Moreover, many route resolution schemes are based upon modified shortest-path heuristics, as these algorithms are very amenable to distributed signaling setups (Section 3). Herein, various cost metrics have been considered, including hop counts, available wavelengths, switching ports, converter resources, etc. Meanwhile, a host of matching wavelength selection strategies have also been coupled with the above path selection algorithms, including first fit (FF), random wavelength selection (RWS), most/least loaded wavelength, etc (see [8]).

Given that the focus herein is on the broader issue of integrated wavelength/sub-wavelength provisioning, the detailed study of optical RWA schemes is clearly out of scope. Instead, it much more beneficial to simply apply existing, mature RWA schemes in conjunction with sub-wavelength (i.e., packet) routing algorithms. As a result, modified shortest path RWA heuristics are chosen using a simple hop count metric, i.e., minimize number of wavelengths used. Meanwhile, wavelength selection is done using FF and RWS. Namely, the former simply picks the first available wavelength whereas the latter randomly selects a wavelength.

3.2 Sub-Wavelength Packet Routing

A different set of constraints are proposed for sub-wavelength path CBR at the IP/MPLS level (Figure 4). Ideally, these schemes must select a feasible path in order to minimize resource consumption and/or balance loads, etc. Now typically resource usage is reduced by restricting hop counts, whereas load balancing can be achieved by selecting less loaded paths [11]. However, since these two constraints can conflict, a novel hybrid constraint is proposed here to incorporate both types. Specifically, consider a path at the IP/MPLS layer, \( p = (i_1, \ldots, i_k) \), comprising a series of routers. The maximal reservable bandwidth, i.e., \( mrb_p \), for this path is defined as the minimum residual (reservable) bandwidth across all logical links (lightpaths) traversed by this path, i.e., \( mrb_p = \min \{ R_{ij} | i,j \in p \} \), where \( R_{ij} \) is the residual bandwidth on link \( ij \). Here, \( R_{ij} \) is also defined as the cost associated with link \( ij \). Clearly, \( p \) is feasible if \( mrb_p \) is no less than the requested bandwidth. Applying the above to sub-wavelength routing, the following shortest path heuristics are examined:

**Modified Shortest Path (MSP):** This scheme simply calculates the minimum hop path across all possible feasible paths. If several paths exist, then the one with the largest \( mrb \) value is chosen.

**Maximal Reservable Bandwidth Path (MRBP):** This scheme computes a path with the maximum reservable bandwidth \( mrb_p \) value across all feasible paths. If there are several such paths, the one with the minimum hop count is selected.

**Minimal Cost Path (MCP):** This scheme searches for a feasible path which yields minimal cost as defined as:

\[
 f(p) = \sum_{j=1}^{k} \frac{1}{R_i} \tag{1}
\]

where \( k \) is the number of nodes in the path \( p \), i.e., \( k=|p| \). Note that the above cost function increases for larger hop counts and also heavier link loads. This contrasts with [6] which only incorporates hop counts. Overall, the MSP scheme is geared towards limiting hop counts, whereas the MRBP scheme is more focused on balancing network loads. Meanwhile, the MCP scheme strikes a balance between hop count and the path loading constraints. Note that the above formulation assumes that sub-wavelength requests are characterized by a bandwidth quantity, either the mean or peak rate. The inclusion of more specific packet-level parameters (e.g., burst sizes, delay parameters) is left for future study.

4. Simulation Results

The performance of the integrated wavelength/sub-wavelength provisioning scheme is now studied using discrete event simulation for the NSF topology. This topology is very reflective of many real-world settings since larger core domains typically feature gigabit-speed IP routers interconnected with DWDM switching cores. In particular, the following assumptions are made:

- Node message processing time (IP/MPLS routers, DWDM switches), \( P_1 \), is 10 \(
\mu\text{s}\)
- DWDM switch configuration/setup time 500 \(
\mu\text{s}\)
- Number of wavelengths on each fiber link, \( W=4\)
- Poisson arrivals, mean \( \lambda \in [0.05,0.2] \) arrivals/\( \text{ms} \)
- Connection requests are uniformly distributed amongst all source-destination pairs
- Connection-holding time is exponentially distributed with mean \( 1/\mu = 100 \text{ ms} \)
- No wavelength conversion is performed at the DWDM switching nodes (i.e., transparent)
- Signaling uses RSVP backwards reservation (IP/MPLS data, GMPLS optical layers)

From the above, nodal traffic loadings are estimated as \( \lambda/\mu \) (Erlangs). Meanwhile, two key metrics are used to gauge performance levels, namely connection setup time and connection blocking probability. The former measures the time required to establish a connection from
source to destination and includes path computation and resource allocation/configuration delays at all intermediate nodes (IP/MPLS routers and DWDM switches). Meanwhile, the latter metric gauges overall resource efficiency and contention resolution during setup (e.g., lightpath route, packet LSP).

Meanwhile, Figure 4 plots the setup time versus the traffic loading for DWDM lightpath requests, and it can be seen that the minimal achievable value is about 22 ms. Furthermore, it is also interesting to note that higher loadings actually yield lower setup times. The reason here is that shorter paths are usually more feasible (i.e., likely) than longer paths under such conditions, and hence setup times tend to be smaller (similar results were also noted in [9]). In all, these results confirm the feasibility of shortest path RWA via distributed RSVP-TE signaling.

The performance of integrated lightpath-packet LSP provisioning algorithms is now tested using discrete event simulation. Here it is assumed that sub-wavelength connection requests are uniformly distributed over the wavelength capacity range, i.e., [0, 1.0], and underlying RWA uses random wavelength selection. Blocking performance is shown in Figure 5 for varying input loadings, and it can be seen that lower loadings, typically below 8 Erlangs, yield minimal variation amongst the schemes. This is primarily because abundant resource levels can effectively handle almost all incoming sub-rate requests. However, the MRBP algorithm consistently yields the lowest performance, particularly at higher loadings, as it tends to allocate longer path routes as compared to the other two algorithms. Meanwhile, the relative performance of the MSP and MCP schemes is more intricate. Namely, at lighter loadings (under 10 Erlangs), the MSP scheme performs slightly better than the MCP scheme, as it effectively limits the number of hops (i.e., preserves resources and increases the likelihood of accepting future requests). However, at higher loadings, the new MCP scheme outperforms the MSP scheme, since Eq. (1) implements a careful balance between limiting hop count and balancing network loads.

Meanwhile, the key resources at the packet-level are link capacities and packet switching fabric speeds. Now in “two-tiered” data-optical domains, the former entities are effectively pre-determined by wavelength granularities. However, the latter are much more dependent upon the capabilities of the IP/MPLS routers. Hence, it is important to measure their effect on performance. Figure 6 shows the blocking performances for a range of fabric speeds as supported by large commercial platforms (10-100 Gbps). Carefully note that the MCP cost function, Eq. (1), has to be modified to incorporate any residual forwarding capacity in packet fabrics. Hence, a modified cost function is used, as defined by the following expression:
\[ f(p) = \sum_{j} \frac{1}{R_j} + \sum_{j} \frac{1}{C_j} \quad \text{Eq.(2)}, \]

where \( C_j \) is the residual switching capacity of IP router \( j \) along path \( p \). Overall, the results indicate a strong correlation between increased packet switching capacity and lower blocking probability. In fact, initial increases in switching capacity yield the highest gains. Moreover, both MSP and MCP exhibit similar performance at lower switching throughputs. However, the MCP scheme notably outperforms the MSP scheme for increasing both MSP and MCP exhibit similar performance at lower loads, as it avoids congested links and nodes, and also limits hop counts. As such, MCP is better suited for heavily loaded networks. Note that the above results are for relatively small wavelength counts, i.e., \( W=4 \), although limited runs with increased wavelength counts (\( W=8,16 \)) also yielded similar relative results. However, due to excessive compute time requirements, larger multi-test case evaluations prove to be prohibitive.

5. Conclusions

Developments in MPLS/GMPLS frameworks have totally revamped older legacy data transport architectures. Today, data-optical integration is evolving towards streamlined "two-tiered" setups in which IP/MPLS routing nodes are interconnected over dynamic DWDM switching domains. These new architectures offer much improved scalability and rapid service creation capabilities. As carriers look to deploy these new technologies, the effective integration between data and optical layer provisioning schemes is becoming a key focus issue. This paper studies sub-wavelength demand provisioning in converged MPLS/GMPLS data-optical domains. Namely, novel integrated RSVP-TE constraint-based routing algorithms are designed to achieve a careful balance between resource utilization and network-level load balancing. Detailed simulations results indicate high resource efficiencies and good load balancing behaviors. Future studies will look at provisioning multiple sub-wavelength priority levels along with sub-connection survivability issues.

![Figure 5: Blocking probability vs. load](image5)

![Figure 6: Blocking probability vs. forwarding speed](image6)

6. References