Dynamic configuration of virtual topology through joint operation of packet-routing and wavelength-routing network

Kai Wu*, Qingji Zeng, Yizhi Xiong, Xudong Yang

R&D Center for Broadband Optical Networking Techonology, Shanghai Jiao Tong University
1716 Haoran Bldg, 1954 Huashan Rd, Shanghai 200030, P. R. China

ABSTRACT

In addressing the integration of IP and WDM networks, we propose a joint operation approach. Through reconfiguration or real-time setup of lightpaths, joint operation focuses on optimizing network performance or guaranteeing it to a defined level under current network configuration and dynamic data traffic pattern at the network operational stage. Our model features a separation of control plane and data plane. The data plane Formulations have been studied in building the model, and the initial findings help in understanding the behavior of data plane and designing the routing and signaling protocols within control plane.

Keywords: Joint operation, IP over WDM, dynamic virtual topology, routing and signaling, traffic grooming, cost function

1. INTRODUCTION

The past decade sees two trends driving revolutionary changes to today's network infrastructure: one is that various data services are developed and provided over internet protocol (IP) networks; the other is that a new optical network layer leveraging wavelength division multiplexing (WDM) is emerging to boost the bandwidth capacity and provisioning flexibility. New architectures of IP directly over WDM [1,2,3] have been proposed to eliminate the bandwidth inefficiency caused by such intermediate layers as ATM, SDH/SONET, etc., to enhance the total network performance and together to lower the total network cost. However, current and emerging WDM networks are in nature connection-oriented, circuit-switched, and with coarse granularity, while the IP networks are packet-switched, connectionless with finer granularity[4]. To bridge the gaps between them, we need to introduce an interaction mechanism. The main approach in addressing this issue considered here is joint operation of packet and optical network rather than a joint design at network planning stage.[5,6]

In IP over WDM networks, lightpaths are setup to connect IP routers. These lightpaths form a virtual topology seen from the upper IP layer. In operation, lightpaths are setup on demand and released when traffic demand terminates, hence a dynamic virtual topology. The main aim of joint operation is to configure the virtual topology in real time so that a changing traffic is supported without reducing network performance and resource efficiency.

In our proposal, the framework for joint operation features a separation of control plane and data plane. Within data plane, virtual topology together with upper served IP layer and lower provisioning WDM layer resembles a multi-layer model. The control plane takes a multi-domain model with defined interfaces at domain edges. In network operation, the virtual topology changes dynamically under the control of control plane. Routing and signaling protocols are running among control entities to exchange resource and topological information from both the electronic and optical layers, and to achieve quality network performance. These protocols should also solve the issues in granularity gap and switching timescale gap between packet and WDM networks.

In the next section we discuss the two framework of joint operation, and propose a generic model for integration of IP and WDM networks in operation. Section 3 focuses on dynamic virtual topology, and gives more detailed descriptions of interfaces between IP and WDM networks. Section 4 presents a group of formulations in defining the routing and grooming issues to bridge the gap between IP and WDM networks. Section 5 proposes plans for cost function in designing routing and signaling protocols. Section 6 draws conclusion.

2. MODEL FOR JOINT OPERATION

In our proposed framework the considered network is shown on figure 1. This is a horizontal view of joint operation. In this view, the IP routers and switches are attached to an optical core network composed of optical crossconnectors (OXC's). Notice that the key here is the separation of control from data manipulation (transportation). The data plane is layered to

---

*Correspondence: Email: kwu816@mail1.sjtu.edu.cn; Telephone: (86 21) 6293-2166; Fax: (86 21) 6282-0892

support the idea of multi-layer switching. The control plane does not necessarily resemble the layered model in data plane. It can be a simple IP network or separate networks interconnected with defined interfaces at edges. Either way, the optical part and the electronic part (part outside the optical core network) of the control plane can talk to each other in some language at the network layer (or layer 3) boundary. These talks are enabled by certain routing and switching protocols.[5]

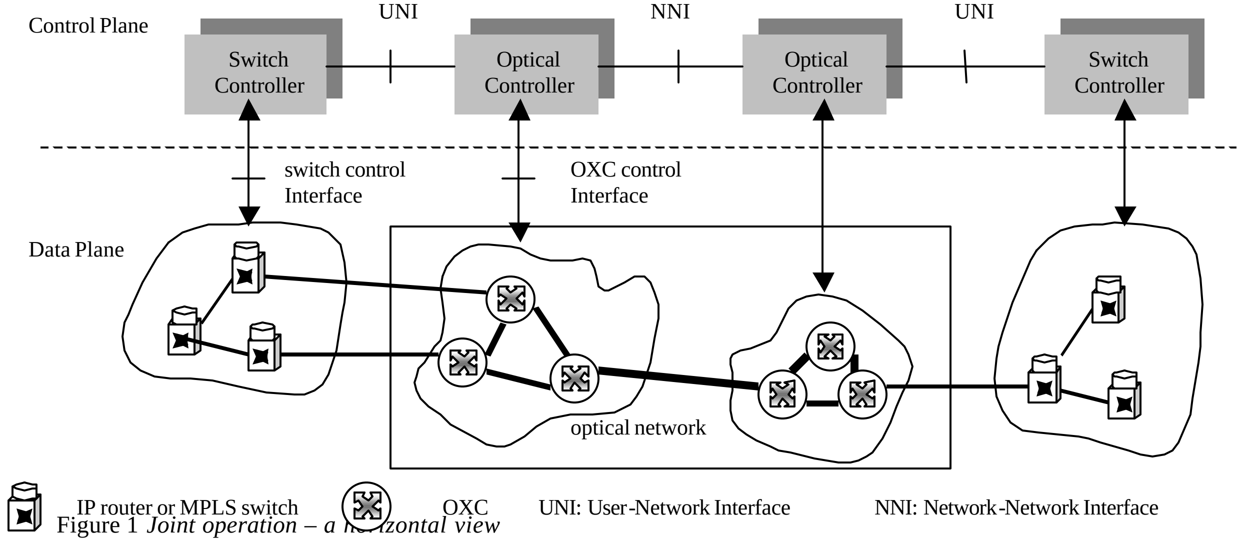

Figure 1 *Joint operation – a horizontal view*

Figure 2. *Joint operation – a vertical view*

In joint operation, virtual topology is a good abstraction of resources in both optical and electronic layers. By maintaining a dynamic virtual topology, the optical part of the control plane as shown in Figure 2 can lead the whole network into an efficient state in real time.

To the upper e-layer, virtual topology is a set of connections. Data ports are the abstraction of interfaces for electronic routers or switches (we call them e-switches hereafter if not specified) attached to the terminals of light paths. These data ports feature fixed or discrete series of bit rates, smart framing and flexible bandwidth aggregation. A connection corresponds to a data link (DL) connecting two data ports from any two nodes. Note that it is possible for a node pair to have multiple links, with diverse or common routes. To the lower o-layer, virtual topology is a set of lightpaths (LPs).

In the data plane of optical networks overlaid by an e-layer, DL's act as logical links between switches from different sites, while light paths are the physical carrier for these DL's. When a light path is setup or released, the corresponding DL is added or reduced. The corresponding exchanges of the pertaining status and command information are initiated at the UNI defined at the rightmost of figure 3. In the control plane, the change of DL's is reflected by the change of connectivity or available bandwidth, and stored in the DL information base, while change of lightpaths stored in the LP information base. New protocols or extensions of existing routing protocols (OSPF etc.) must be employed to dispatch the alteration information as soon to make these two databases consistent and initiate rerouting of data traffic. Thus DLs and LPs define well the interfaces for e-layer and o-layer within data plane, and also distinguish the UNI and NNI within control plane. Data traffic routing and wavelength routing run in two worlds demarcated by these two interfaces.

The control entity of smart framing and grooming acts as a "bridge and cussion" in solving the problem of granularity gap between e-layer and o-layer. Since the each LP is with a fixed or discrete series of bandwidth, which is deemed with coarse granularity, while data flows or data packets at the e-layer are with an arbitrary and much less bandwidth. A smart framing means that the data ports are capable of aggregating flows of various bandwidths without causing problems like contention or interleaving. Grooming indicates an efficient multiplexing is employed[1], and possibly reduces the electronic processing of e-layer (for most cases, the add/drop operations). The detailed physical implementation is not discussed here. However, in a computational view, this entity needs such information as connections and usage of the connections (e.g. ratio of spare to used bandwidth). Such information is fetched from electronic control entities across the UNI.

3. PROBLEM FORMULATION

By adopting the concept of dynamic virtual topology, we can re-use some formulations previously used in joint planning or joint configuration[6,7]. These formulations are useful in designing control algorithm in the control plane and in establishing analytical models. We present these formulations according to the model illustrated in last section. We mention e-layer and o-layer to match the layered model in the data plane.

In this section, the following are defined:

N	Number of nodes, a node is an oxc (o-switch) overlaid by one or a group of e-switches.
N_{sd}	Number of source-destination node pairs, $N_{sd} = N*(N-1)/2$. For simplicity, we assume all connections bi-directional and symmetric.
$\mathbf{D} = \{D_n\}$	Number of data ports on node n, where $n = 1, 2, \ldots, N$ We assume a uniform capacity for all data ports and take it as 1 unit. So D_n also represents the maximum capacity for node n.
$\mathbf{d} = \{d_s\}$	Number of data links over node pair s, also the total capacity for connections over node pair s, where $s=1,2, \ldots, N_{sd}$. The value of 0 means no connectivity.
$\mathbf{t} = \{t_s\}$	Aggregated traffic over node pair s in terms of requested bandwidth, where $s=1,2, \ldots, N_{sd}$
$\mathbf{R} = \{r\}$	Set or routes r, where r is a concatenation of node pairs, over each of which there is at least one data link.
$\mathbf{a} = \{a_r\}$	Bandwidth allocated to route r. $r \in \mathbf{R}$
$\mathbf{A}=\{A_{s,r}\}$	Traffic routes matrix in which $A_{s,r}$ takes on the value of 1 if route r belongs to s-d pair s, 0 otherwise.
$\mathbf{B}=\{B_{s,r}\}$	Traffic routing matrix in which $B_{s,r}$ takes on the value of 1 if route r goes through a DL over node pair s, 0 otherwise.

First of all, the allocated bandwidth for all the routes added up should meet the bandwidth demand. We have

(1)

The internet traffic has been characterized as bursty and fractal in nature. In (1) we use the sign of inequality instead of equality to spare room for sudden change of traffic demand. The term bandwidth works as a common word for mean traffic bitrate, peak bitrate, or equivalent bitrate, depending on which metric is of concern.

Secondly, over each s-d pair, the aggregated bandwidth should not exceed the total capacity of the data links , i.e.

(2)

Expression (1) and (2) regulate the determinations of routes **R** and allocated bandwidth **a**. We define **routing** as find or computation of the routes **R**, and **grooming** as allocation of bandwidth **a** for these routes. To support grooming, the data ports must be capable of smart framing as described in last section.

Expression (1) relates routing and grooming (RAG) with traffic matrix **t**, while (2) connects them with data link matrix **d**. Matrix **d** could be considered the demand matrix of light paths for o-layer. By properly altering **R** and **a**, we could accommodate some traffic changes without affecting o-layer configurations. When o-layer reconfiguration is unavoidable due to bigger traffic variations, a smart RAG could also minimize such reconfiguration or maximize resource efficiency.

There is a special group of routes, each of which passes over only one data link. They are also called one-hop routes. These routes are important since traffic taking on them do not incur any electrical forwarding, which is a key factor affecting the network performance. In fact, if there are enough resources for all traffic to be carried on one-hop routes, no routing and grooming is needed, traffic are transported on the fly. However, such situation could hardly happen because of the obvious scalability problem and limitations of both optical resources like wavelength and electrical resources like data ports or transceivers.

One-hop routes correspond to node pairs with at least one data link. We assume there is a potential data link and hence a one-hop route over every node pair. We denote the allocated bandwidth for one-hop routes as $\mathbf{a}^{(1)}$. (If there is no data link for some node pair s at a certain network state, the allocated bandwidth $a_s^{(1)}$ is 0.) When one-hop route is not available or bandwidth along with it is congested, a detouring multi-hop (more than one hops) route is chosen for demanded traffic. Detoured traffic is very common in today's e-layer networks. Typically a minimum hop or a shortest path routing algorithm is running. If load balancing or traffic engineering is leveraged, several edge-disjoint routes may be chosen for traffic detouring. The total allocated bandwidth for multi-hop routes per node pair is denoted by $\mathbf{a}^{m}$. Thus, the anticipated traffic could b carried by two parts, one-hop routes and multi-hop routes, i.e.

(3)

For da links per node pair, the bandwidth constraint still holds. (2) could be rewritten as

(4)

In expression (4), a multi-hop grooming matrix is denoted by $\mathbf{C}=\{C_{s,s'}\}$. Here the value of $C_{s,s'}$ indicates the allocated bandwidth on data links over node pair s for multi-hop traffic of node pair s'. If more than one multi-hop routes are chosen and bifurcation of traffic is allowed (an implicit assumption in this paper), the value of $C_{s,s'}$ can be a fraction other than 1 or 0. Note that $C_{s,s}$=0, since **C** refers only to bandwidth distribution of multi-hop traffic.

Now we can turn (3) ,(4) into matrix forms:

(5)

Where **E** denotes an N_{sd}-by-N_{sd} identity matrix, $\mathbf{a}^{(1)}$, $\mathbf{a}^{\mathrm{m}}$, **t** and **d** are all N_{sd}-by-1 column vectors. From (5) we can see how $\mathbf{a}^{(1)}$, $\mathbf{a}^{\mathrm{m}}$, **C**, **t** and **d** are coupled together.

In order to have a closer look at the RAG issue, let's define some derived variables:

(6)

D refer to the difference between provisioned bandwidth and the demanded bandwidth over each node pair;

(7)

S refer to the spare capacity left for grooming of multi-hop traffic;

(8)

R refer to the residue traffic other than one-hop traffic, these traffic must be multi-hop routed.

The nonnegativeness of **R** and **S** sets an upper bound for one-hop traffic $\mathbf{a}^{(1)}$. Finally we define

(9)

G refer to the bandwidth groomed over each node pair.

From the meaning of **D**, **S**, **R** and **G**, we also have

(10)

(11)

(12)

From these equations we can define three distinguished states of connections over each node pair, determined by bandwidth difference **D** and above equations. They are listed on the table 1.

In case we could find a grooming matrix **C** satisfying (10)-(12), the network is in a stable state where each node pair remains in one of the three states. We call these states stable since there is no need to initiate the reconfiguration at o-layer or re-routing at e-layer. Otherwise, at least one node pair is departed from the stable states. By changing R or S (adjusting the one hop traffic bandwidth) for each node pair, we can expect a new **C** generated. This changing refers to re-routing (of multi-hop traffic) or re-grooming (adjusting the allocated bandwidth). When it does not success, a VT reconfiguration is initiated. Usually, grooming adjusting is again evoked after a VT change. Therefore special care should be taken to avoid a circular initiation of grooming adjusting and VT change. And that's why we put grooming and lightpath control entities in optical control domain.

State	I	II	III
Bandwidth difference One-hop traffic carried	$D >> 0$ $a^{(1)} > t$	$D \approx 0$ $a^{(1)} = d - S$	$D << 0$ $a^{(1)} = d$
Multi-hop traffic groomed	$a^{m} = 0$	$a^{m} > R$	$a^{m} > R$
Residue traffic for multi-hop routed	$R = 0$	$R = S - D$	$R = -D$
Spare space for traffic grooming	$S = D$	$S > 0$	$S = 0$

Table 1 *Bandwidth usage states for connections over one node pair*

4. CONCLUSION

In this paper, we study the joint operation approach in integration of IP over WDM networks. We propose a network model established on a mesh-connected optical (WDM) core network with electronic switches/routers attached to it. A separation of control plane and data plane is defined. The data plane resembles a layered model, and dynamic virtual topology works fine in this model. We then put virtual topology into formulations and define the grooming and routing in mathematical expressions. Three stable states of network connections are identified through these formulations. These findings help to identify sequence of actions in reconfiguration, i.e. reconfiguration algorithm.

ACKNOLEDGMENT

This work was supported by National 863 Plan of China, National Natural Science Foundation of China, and National Key Lab of Broadband Optical Fiber Transmission and Communication system of China.

REFERENCE

1. Eytan Modiano, "WDM-based packet networks", IEEE Comm. Mag., pp. 130-135, March, 1999
2. Nasir Ghani et al., "On IP over WDM integration", IEEE Comm. Mag., pp. 72-84, March, 2000
3. Kai Wu et al., "Joint operation in multi-hop WDM networks under dynamic data traffic", ICCT 2000, Beijing, August, 2000
4. M. Veeraraghavan, M. Karol, "Internetworking connectionless and connection-oriented networks"
5. Kai Wu, Qingji Zeng, "Joint operation of packet network and optical network", submitted to Photonic East 2000
6. Tibor Cinkler et al., "Heuristic algorithms for joint configuration of the optical and electrical layer in multi-hop wavelength routing networks", Infocom 2000, March, 2000
7. R. Ramaswami and K. N. Sivarajan, *Optical networks: a practical perspective.* Morgan Kaufmann Publishers, 1998