Design and Dimensioning of a Novel composite-star WDM Network with TDM Channel Partitioning

Stefano Secci
Department of Network and Computer Science
ENST -Paris
37/39, rue Dareau, 75014 Paris, France
Email: stefano.secci@enst.fr

Brunilde Sansò
GERAD and Department of Electrical Engineering
Ecole Polytechnique de Montréal
C.P. 6079, Succ. Centre-ville
Montréal (Québec) Canada H3C 3A7
Email: brunilde.sanso@polymtl.ca

***Abstract*— This paper presents the design and dimensioning optimization of a novel optical network structure, called the Petaweb, having a total capacity of several Pb/s (10^{15} bit/s). Its topology is a superimposition of stars that drastically eases signaling and switching operations. Firstly, we deal with the network model, focusing on the insertion of the time-sharing of the optical channel in network components and in lightpath provisioning. The design problem is jointly a network dimensioning and an assignment problem; we propose for the dimensioning an integer linear programming formulation and a linear resolution algorithm for the assignment. We also propose the use of a quasi-regular topology extracted from the optimized regular topology to reduce costs and improve the network utilization.**

I. INTRODUCTION

To accommodate the needs of the future Internet, the Petaweb [4] [3] [5], a high-capacity network structure capable of operating at a global rate of the order of the petabits per second (10^{15} bps), was proposed by Nortel Networks. The Petaweb presents a composite-star architecture and an optical core based network infrastructure. This novel structure is composed of edge nodes connected through core nodes as shown in Fig. 1. Every edge node is connected to every core node but neither the core nodes nor the edge nodes are directly connected to each other. This kind of configuration allows a one-hop optical path between two edge nodes through a core node.

The idea behind such an optical infrastructure is to offer a very large transport capacity through a regular topology to facilitate upgrades and expansions. Furthermore, essential network functionalities such as routing, addressing and labeling are drastically simplified because of the unique Petaweb architecture. A set of permanent static traffic connection requests defines a pre-assigned virtual topology. Many time-divisioned lightpaths serve these connection requests, and their traffic is transported upon the physical topology. In this paper, we aim at the joint design and dimensioning problem of this structure. We want to locate and dimension the core nodes while determining the switching schemes and the optimal resource assignment of lightpaths, taking into account the modeling specificities of the TDM/WDM nature of the network.

This is the first time that this problem is tackled in all its modeling complexity. In [3] an emulation of the Petaweb was explained whereas in [5] the Petaweb was compared to typical optical mesh networks. A routing algorithm for the Petaweb was proposed in [6] and in [7] IP/Petaweb scenarios were studied. The first attempt to formally define the mathematical problem behind the Petaweb design can be found in [8], where an efficient resolution approach is proposed. The Petaweb has been also considered as a construction block for a structure capable to operate at the rate of Yotta bits per seconds (10^{24} bps) called the YottaWeb. Architectural rules and optimization procedures for regular and irregular topologies of the YottaWeb have been presented in [9].

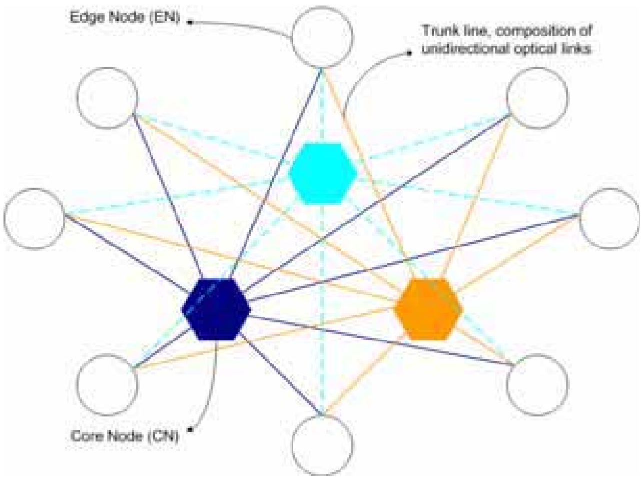


Fig. 1. Composite-star Petaweb structure.

One key element that has not been taken into account in any of the previous studies, and that is the main objective of this article, is how the TDM features can be integrated into the optimization model and how to guarantee the optimal assignment of connections.

The paper is divided as follows. The network model is introduced in Section II. A design problem formulation is presented in Section III; the problem is divided into two sub-problems: the first tackles the resource allocation while the second deals with the assignment of wavelength and time-slot to the allocated resources. We describe an ILP formulation and a linear algorithm for their resolution. Results are analyzed in Section IV whereas Section V concludes the paper and presents ideas for further research.

1-4244-0425-8/06/$20.00 ©2006 IEEE

II. NETWORK MODEL

As previously described in the Introduction, the Petaweb is essentially composed of edge nodes that interconnect between themselves by core nodes. The architecture that results from this topology is very peculiar, since it is, to our knowledge, the only type of network in which the *backbone* nodes are not interconnected.

Physically, an *edge node* (EN) is an electronic node that can interface with any device requesting end-to-end connections. An EN asks the transport network to supply media for its set of source-destination connection requests with other ENs; in this paper, we assume that a Connection Request (CR) consists of traffic that is coming from a SONET/SDH interface.

The connection between the N edge nodes and a core node is shown in Fig. 2. Every EN is connected to a core node through one *optical link*, composed by one or more optical fibers. We suppose the use of unidirectional optical fibers, so that an EN has one optical link incoming and one outgoing for every core node. We suppose that all the fibers of the network carry the same number of wavelenghts.

A *core node* (CN) is a set of arrays of parallel space switches, also called *switching planes*. We suppose that there are different types of CNs, each type having a different number of switching planes s_r. From now on, a core node of type r will be denoted as CN-r. Thus, the optical link connecting an EN to a CN-r has s_r unidirectional optical fibers, one for every switching plane. In this paper, we assume three types of core nodes: a CN with one array of space switches is of type 1 (as the one illustrated in Fig. 2); a CN with two arrays is of type 2 whereas the third type of core node presents four arrays. All the incoming WDM fibers are demultiplexed into their different lambda-channels, each of which is connected to the associated space switch of the respective array. These switches are Optical Cross Connects able to perform full wavelength conversion, rendering the Petaweb an all-Optical network without any O/E/O conversion. Each space switch handles channels of the same wavelength; those referred to the same EN are then multiplexed into the optical link going back to that EN. Such parallel-planes structure increases the reliability of CNs because a hypothetical failure in a switching plane would affect only the connections on that plane. Moreover, this modular CN structure allows to keep every single switching plane directly connected to all the ENs through two unidirectional fibers, and a single fiber send or receive traffic only to or from its switching plane; an upgrade of a switching site would thus not affect the existing interconnections and switching planes, that makes such core network easily scalable [4].

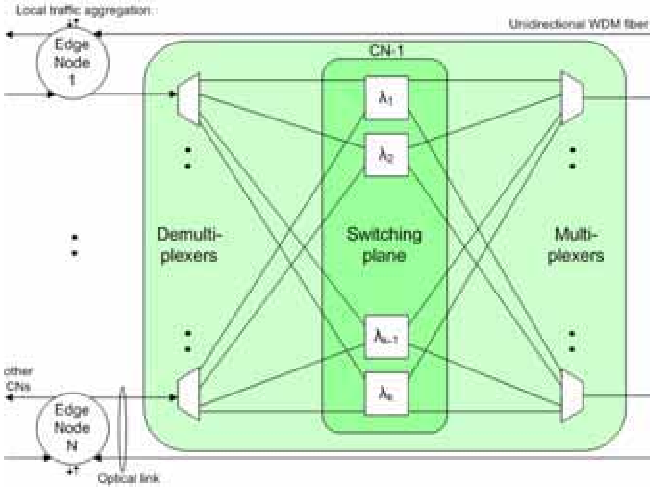


Fig. 2. Parallel-planes optical core node in the Petaweb.

Since in our modelling we will try to assess the best CN location, we must first define what are the potential sites where the CNs can be located. We assume that the potential sites are the ENs locations, that are known in advance. Thus, an EN site becomes a *switching site* if a CN is installed in it. Note that several CNs can be installed in the same site; and, if a particular CN is chosen to switch the traffic of the EN in the same site, the traffic is directly added and dropped. Also note that, since several CNs can be installed in the same site, the physical connection between an EN and a switching site can be composed of many optical links. From now on we call this physical connection an *optical trunk line*.

The network design consists in the minimization of the total cost of the network, determined as the sum of the costs of the network components. We have three kinds of costs in our model: the core node cost, the fiber cost and the propagation delay cost. The *core node cost* is composed of a fixed cost f_r, that depends on the type r of the core and such that $f_r > f_{r-1} > ... > f_1$. Furthermore, the number of switching planes is such that $s_r = 2s_{r-1}$: a CN-2 is capable to switch the equivalent of traffic of two CN-1 with a minor cost, and idem for a CN-3. The variable cost is determined by the active ports. An active port has a cost P, to be scaled for higher types. Let us indicate by W the number of wavelengths per fiber, with M the set of ENs sites, and with γ the scale factor for P, the global cost of a CN-r is $K_r = f_r + 2N_{en}W s_r P \gamma^{(s_r-1)}$, such that $K_r < 2K_{r-1}$ because the second cost term of K_r is scaled by γ for CN-2 and by γ^3 for CN-3.

The *fiber cost* is indicated as F and is in unit of length. It is the cost of a single-wavelength fiber, which is than scaled by a discrete function $\phi(W)$ that depends on the number of wavelengths. Let us indicate with Δ_{ij} the distance between the sites i and j; the installation of a CN-r in the site i requires the installation of s_r fibers per direction for every EN, with, thus, a global cost of $F_{i,r} = 2\,\phi(W)\,F\,s_r \sum_j \Delta_{ij}$.

Every lightpath is transported through only two optical links; the *propagation delay cost*, indicated by β, is proportional to the travelled distance (between the origin EN and the switching CN and between the CN and the destination EN) and to the lightpath traffic. The propagation delay cost expresses the need for fast communications in the network. The minimization of the global cost part due to propagation delays allows to assign shortest possible paths to lightpaths.

What about lightpath provisioning? How the requested traffic volume is accommodated in the physical channels? The answers to these questions lead us to the heart of our network model, that is the setting up of different classes of service. This will be treated in the next section.

A. Classes of service

In [5] Blouin et al. assumed the use of TDM in the Petaweb to realize sub-channels within a lambda-channel. They proposed the use of electronic switches. In order to model the TDM features, we assume the switch architecture with time-slot interchanging functionalities that has been proposed by Huang et al. [10]. They introduced an all-optical TDM Wavelength Space Routers (TWSR) (e.g. resumed in [11]) where the time-slot switching is implemented using Optical Time-Division Space Switches (OTDSSs); the alignment of the time-slots can be done by the schemes of input synchronizer described in [15]. An OTDSS can reconfigure itself with the granularity of a time-slot and multiplexes in a time-slot basis. No buffering operations are required and a local access unit guides the alignment of the incoming time-slots. The TWSR is composed of a number of OTDSS equal to the number of used wavelengths and each OTDSS manages the time-slots of the same wavelength; therefore this kind of switch can easily replace the wavelength-driven space switch already discussed for the Petaweb (Fig. 2), and we will thus assume the TWSR as the switching plane unit. In what follows, as suggested in [10], we call a time-slotted lightpath, a *ts-lightpath* (TLP).

As it emerges from specification of G.709 of the ITU-T [2], the use of TDM in WDM networks is useful for two main reasons: fractionating the lambda-channel into more sub-channels improves the network capacity utilization, the requested resources are smaller and the network cost is more competitive; many transport levels can be used, e.g. ITU-T G.709 defines, for the OTN interfaces, three different traffic levels. Therefore, we have to choose the number of TLP classes and the transport capacity for every class. In [12] the authors face this problem in the design of a WDM network with static traffic load and TDM channel partitioning. Referring to the Optical Transport Unit (OTU) hierarchy specifications, they chose three transport classes that correspond to the three OTU-rates (2.5 Gb/s, 10 Gb/s and 40 Gb/s). This choice is justified with the fact that, for the moment, operators use transmission systems with fixed transmission rates.

We introduce here a three-level hierarchy for a ts-lightpath inspired from the OTU hierarchy. However, we will not limit our choice of bit-rate values to the OTU rates. Let Z_h denote the transport capacity of a TLP of class h (TLP-h). A CR is best served using the minimum number of TLPs. This means that the traffic volume of a CR is rounded up to a value that is the sum of the transport capacity of several TLPs that can even be of different classes; the number and the classes of these TLPs is such that the rounding up value is as close as possible to the traffic of the CR. Let us assume in this article that the rules to set the transport capacity Z_h of a TLP-h are the following:

- $Z_1 = \frac{1}{2^n} C_{ch},\ n \in \mathbf{N}$, represents the transport capacity of a time-slot; it is a fraction, multiple of 2, of the transport capacity of a wavelength (C_{ch});
- $Z_2 = C_{ch}$, that is the transport capacity of a wavelength
- $Z_3 = W\, C_{ch}$, that is the transport capacity of a fiber

The above capacity values have been chosen so that switching operations, already simplified by the Petaweb architecture, are further simplified. An optical switch can commute a TLP-1 in a time-slot basis, and a TLP-2 switching the whole wavelength without any time-slot alignment; to switch a TLP-3 one may simply interconnect the incoming fiber with an outgoing fiber without any demultiplexing/multiplexing. The TLPs are then switched and transported independently and the original data flow is then recomposed at the destination EN.

The association to a TLP of a propagation delay cost directly proportional to its bit-rate, besides that to traveled path length, is an important novelty for the design of backbones with QoS (Quality of Service) guarantee: the minimization of the global network cost will give priority to high bit-rate classes in getting the bandwidth over short paths. Indeed, high bit-rate lightpaths may accommodate traffic, belonging to video streaming and voice-over-ip services, which need the lowest possible propagation delay. For example, a modern operator offering Video on Demand services, and having few video pump station in its geographical network, will present high bit-rate connection between the PoPs (Points of Presence) of the pump stations and the PoPs of the clients. Similarly, the gateway to public switched telephone networks may be located not in all the PoPs and the calls would be aggregated in high bit-rate flows. Therefore, with our model we can guaranteed to high bit-rate connection requests the priority for the attainment of short path.

B. Constraints

The Petaweb design has to preserve the characteristics of network components and to simplify routing operations. For these reasons, mathematical formulations and algorithms must consider what we call Capacity and Coherency Constraints.

Edge nodes and optical links have maximal capacities to respect. The *Capacity Constraints* are modular, the capacities can be allocated and incremented only through discrete quantities: the optical link capacity can be increased by a multiple of the capacity of W lambda-channels at a time and must be verified for both directions; the capacity of an EN depends on the number of optical fibers connected to it.

Furthermore, we have to consider additional constraints to satisfy basic communication system requirements in terms of delay and buffering operations; we call these constraints *Coherency Constraints*:

1) All the TLPs of a CR should be transported on the same optical trunk line;
2) All the time-slots associated to a TLP should be transported on the same optical link;
3) All the TLPs of a CR should be transported contiguously in the time and in the frequency domains.

The 1^{st} Coherency Constraint assures that the traffic between two ENs is switched in the same site. Without this constraint one may loose too much time at the destination EN because of out-of-order buffering operations; two TLPs of the same connection may be switched in different sites cumulating different propagation delays. The 2^{nd} Coherency Constraint imposes that the time-slots used by a TLP must be switched in the same core node. The 3^{rd} has been introduced for three main reasons: to ease multiplexing/demultiplexing operations; to lighten buffering operations at destination ENs; to relate lost data in case of damage of a single switching plane to the minimum possible number of CRs.

III. THE DESIGN PROBLEM

The optimization problem consists in finding the best composite-star physical topology for the given set of TLPs, respecting the network model and the peculiar composite-star architecture. A pre-processing phase produces the optimal set of TLPs for the assigned traffic matrix; the resulting set drives the dimensioning of the physical topology and the assignment of its resources to the TLPs. This means that the route of every TLP through a CN must be chosen in parallel with the number, the size and the position of CNs, and that wavelengths and time slots must be assigned to the routed TLPs.

Dealing with the optimization of WDM networks with TDM channel partitioning upon a pre-assigned physical topology and with static CRs, the authors in [12] call their design problem RFWTA (Route, Fiber, Wavelength and Time-slot Assignment), keeping as the objective the minimization of the total number of fibers, which is the only variable in the physical topology. In our design problem we also need to determine an efficient routing and assignment of wavelengths and time-slots for a set of pre-assigned CRs. Nevertheless, we need also to dimension the physical topology: the nodes location and their size is an additional variable because the physical topology is not pre-assigned. The set of possible routes is not established a priori. Thus, we must allocate the resources for routes and fibers to transport the requested traffic volume. The virtual topology drives the choice for the physical routes and affects the dimensioning of the physical network.

The Petaweb network design is jointly a dimensioning and an assignment problem. To underline the connection with the RFWTA problem we call it Route, Fiber, Wavelength, Time-slot Allocation and Assignment (RFWTAA); we divide our problem into two sub-problems: Route and Fiber Allocation (RFA) problem, which treats the allocation of the resources guaranteeing an efficient routing, and Wavelength and Time-slot Assignment (WTA) problem, that concerns the assignment of the allocated resources.

A. RFA resolution

The task of the RFA problem is to find the optimal location of network components in order to efficiently switch all the TLPs of the virtual topology; every TLP has to be assigned to its switching CN so that its route and optical links are decided. In mathematical terms, the RFA can be seen as a particular Location Problem since it has many similarities with the Plant Location Problem [13]: a set of clients is given and every client has a specific demand for a product; the goal is to optimally locate the plants, which send products to the clients, thus minimizing the global cost expressed by the fixed cost of plants and by the transport cost; the potential sites for plants are known and the demand of a client is satisfied by one plant. In the Petaweb, core nodes are similar to plants and edge nodes are similar to customers. The propagation delay cost is a transport cost. Nevertheless, while in the PLP the product is transported from a plant to a client, in the RFA the product use the plant as transit; moreover, because of the capacity constraints on optical links, ENs and CNs, it also presents some similarities to the Capacitated Facility Location Problem [14].

We propose an ILP formulation for the RFA problem resolution. Let us add the notations in Table I to the already described ones. The ILP formulation for resources allocation follows.

TABLE I

NOTATIONS

M	set of sites
T	set of pairs of sites ($M \times M$), $p \in T$ represents a CR
O_j	a subset of T with a fixed origin site j
D_k	a subset of T with a fixed destination site k
V	set of types of core nodes
E_r	number of CN-r specimens that can be enabled in a site
(i,r,e)	triple representing a CN specimen, $i \in M, r \in V, 1 \leq e \leq E_r$
C_j	capacity, in Gb/s, of the edge node in site $j, j \in M$
H	set of TLPs classes
L_h	maximal number of TLP-h specimens for a CR, $h \in H$
(p,h,l)	triple representing a TLP specimen, $p \in T, h \in H, 1 \leq l \leq L_h$
d_{ip}	distance traveled going from the origin j to the destination k of the CR p passing by the site i: $d_{ip} = \Delta_{ij} + \Delta_{ik}$
y_{ire}	indicates if the e^{th} CN-r specimen is enabled in the site i
x_{phl}^{ire}	indicates if l^{th} TLP-h specimen of CR p exists and is switched by the CN (i,r,e)

$$\min G(\overline{y},\overline{x}) = \sum_{(i,r,e)} (K_r + F_{i,r})\, y_{ire} + \sum_{(i,r,e)} \sum_{(p,h,l)} \beta\, d_{ip}\, Z_h\, x_{phl}^{ire} \qquad (1)$$

$$s.t. \qquad \sum_{r\in V} \sum_{e=1}^{E_r} x_{ph_1l_1}^{ire} = \sum_{r\in V} \sum_{e=1}^{E_r} x_{ph_2l_2}^{ire} \qquad (2)$$

$\forall i \in M$, $\forall p \in T$, $\forall h_1, h_2 \in H$, $\forall l_1 | 1 \leq l_1 \leq L_{h_1}$, $\forall l_2 | 1 \leq l_2 \leq L_{h_2}$, $(h_1, l_1) \neq (h_2, l_2)$

$$\sum_{(i,r,e)} x_{phl}^{ire} = 1 \qquad \forall (p,h,l) \qquad (3)$$

$$\sum_{(i,r,e)} C_{ch} W s_r\, y_{ire} \leq C_j \qquad \forall j \in M \qquad (4)$$

$$\sum_{(p\in O_j,h,l)} Z_h\, x_{phl}^{ire} \leq C_{ch} W s_r\, y_{ire} \qquad \forall j \in M \forall (i,r,e) \qquad (5)$$

$$\sum_{(p\in D_k,h,l)} Z_h\, x_{phl}^{ire} \leq C_{ch} W s_r\, y_{ire} \qquad \forall k \in M \forall (i,r,e) \qquad (6)$$

$$y_{ire} \in \{0,1\}\,,\; x_{phl}^{ire} \in \{0,1\} \qquad (7)$$

The objective (1) is to minimize the total network cost, sum of the costs of all the enabled core nodes, of the optical fibers between the edge nodes and the enabled core nodes, and of the propagation delays cumulated by the TLPs;
(2) expresses the 1^{st} Coherency Constraint: two TLPs associated to the same CR must be switched by CNs in the same site and, thus, transported on the same optical trunk line. Given a switching site and a CR, for every couple (TLP_1,TLP_2) of the CR, with $TLP_1 \neq TLP_2$, the number of CNs in that site switching TLP_1 must be equal to the number of CNs switching TLP_2;
(3) expresses the 2^{nd} Coherency Constraint: the traffic transported through a TLP, from an origin EN to a destination EN, must be entirely switched in the same CN and, thus, transported in the same optical links. For every TLP, the number of CNs switching that TLP must be equal to 1;
(4) imposes the respect of the capacity constraint for every EN;
(5), (6) impose the respect of the capacity constraint for the optical links between every origin EN and every CN, and between every destination EN and every CN;
(7) indicates that x_{phl}^{ire} and y_{ire} are binary variables.

This formulation presents some fundamental differences from the one proposed in [8]; the first two terms of the objective (1) remain unchanged but the third determines the propagation delay cost considering the discrete capacity Z_h of a ts-lightpath.

The resulting complexity is high, but not prohibiting for the assigned instances; we could control the number of TLPs, and thus the number of variables and constraints, by grooming end-to-end TLP-1s and TLP-2s of the same CR in sub-classes consisting of, respectively, 4 time-slots and 4 wavelengths.

B. WTA resolution

The result of the RFA problem is the optimal location of the CNs to switch all the TLPs at minimum cost. The results contain, thus, the set of TLPs for every CN and, consequently, the assignment of every TLP to two optical links, one between the origin edge node and the CN, and one between the CN and the destination edge node. The task of the WTA problem is to assign to every TLP a subset of the wavelengths and the time-slots of its optical links. Let us remember that a TLP-3 needs one fiber, a TLP-2 one wavelength and a TLP-1 one time-slot. The only constraint to respect is the 3^{rd} Coherency Constraint.

The flow chart of the WTA resolution algorithm is showed in Fig. 3. The input are the sets of enabled CNs, optical links CN-ENs for every CN, fibers for every optical link, and TLPs assigned to every optical link (grouped according to their class and their CR). The algorithm starts considering one optical link at a time and assigns time-slots, wavelengths and fibers to the TLPs. As it can be noticed from the flow chart, the assignment of whole fibers to TLP-3s is done independently of their CR.

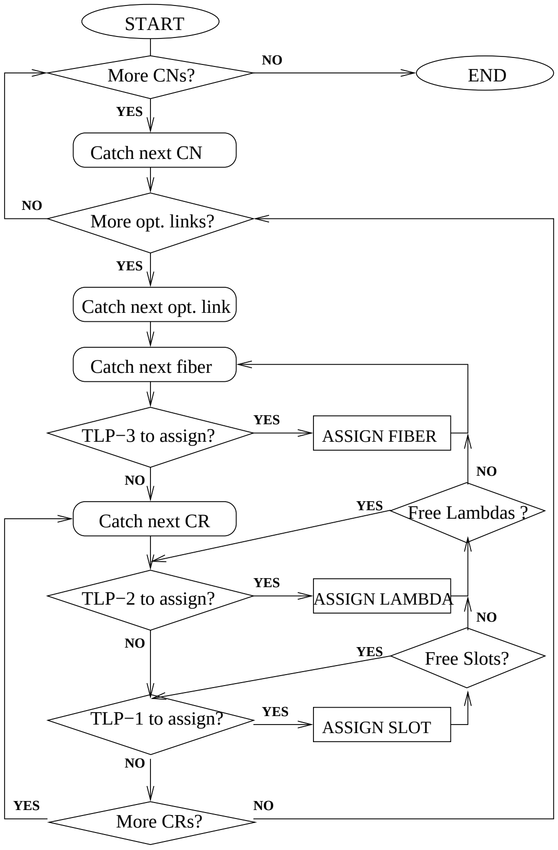


Fig. 3. Flow chart for WTA resolution algorithm

C. Extraction of a quasi-regular topology

In the RFA solution only a portion of the transport capacity will be reasonably assigned and many optical fibers or links may be totally unused. This may occur because the traffic matrixes contain few peaks of traffic between two sites, and a lot of medium-low values; the peaks will cause high exploitation of the optical links connected to the related sites, while other optical links, interesting sites with low CRs, will be under-used. Indeed, the activation of a switching plane in the network requires the installation of one fiber for every edge node.

What we propose is a variant to the composite-star regular topology: we pass to the WTA algorithm only the fibers really exploitable, the others are ignored. How to do so? For every optical link we determine the number of fibers needed by the assigned TLPs, and if this number is inferior to s_r the abundant fibers are no longer considered. Therefore this operation does not imply an undifferentiated grooming of TLPs of different optical links on the same fibers. The regularity is preserved and will be reachable with future re-installations of the disabled fibers (please consider that the main requirement remains the dimensioning of a regular Petaweb). Thus, the final composite-star topology becomes irregular, or, better, quasi-regular. The physical connection between an EN and a CN may become partial, but sufficient for the traffic reserved on them.

When the unused fibers are disabled, in conjunction to the corresponding CN ports, the network cost decreases. The extraction of the quasi-regular topology does not require changes to the WTA algorithm, because one just change the fibers number for the optical links. An additional physical hypothesis must, however, be assumed with quasi-regular topologies: the switching planes of a same core node must be able to communicate each others in order to multiplex lightpaths on the same fiber (if required).

IV. RESULTS

In this section we show the results obtained implementing the resolution algorithms for the Petaweb design problem. The algorithms have been implemented in C++, and we used the callable library of CPLEX 9.0.1 to solve the ILP formulation of the RFA problem. The simulations run on a CPU AMD Opteron 64bit 2.4Ghz, 1MB cache, 16GB RAM.

The traffic introduced by access interfaces is extracted from traffic matrixes. Every EN has CRs of the order of the Gbit/s. The total requested transport capacity is in the order of the Pb/s. We use in this paper four traffic matrixes, two for 10-node networks (10A, 10B) and two for 34-nodes networks (34A, 34B). The suffix indicates the traffic type: **A** matrixes were provided by the industry and present many zero values; **B** matrixes are dense and are obtained from a gravity model estimating the traffic between two cities as being directly proportional to the product of the populations and inversely proportional to the square of the distance between the sites.

In this paper we adopt $C_{ch} = 10$ Gbit/s and $W = 16$ (as in [4] [5] [8]) and, thus, $Z_2 = 10$ Gbit/s and $Z_3 = 160$ Gbit/s. Having to choose a value for Z_1, we verified experimentally that an appropriate value guaranteeing an acceptable rounding up value for our traffic models is Z_1=0.625 Gbit/s. Furthermore, being Z_1 a fraction multiple of 2 of the capacity of a lambda-channel of 10 Gbit/s, the OTU rates 2.5 Gbit/s, 10 Gbit/s and 40 Gbit/s [2] are reachable through a composition of TLP-1 and TLP-2.

Underlining that the final network cost is normalized to F, the parameters values (some furnished by an industrial partner) not yet explicited are: $\gamma = 0.95$, $P/F = 150$, $\beta/F = 0.1$ $[Km\, Gbit/s]^{-1}$, $f_1/F = 20$, $f_2/F = 50$, $f_3/F = 100$, for 10A and 10B $C_j = 1000$ Gbit/s, for 34A $C_j = 2000$ Gbit/s, for 34B $C_j = 2800$ Gbit/s, $L_1 = L_2 = 12$, $L_3 = 20$; moreover, we set $E_1 = E_2 = 1$ to avoid the generation of equivalent solutions (remember that $K_r < K_{r-1}$). We employed $\phi(W) = W$ considering, thus, that the cost of a fiber is proportional to the number of wavelengths.

A. RFA results

Table II reports the final results for the 10A, 10B, 34A and 34B cases, Table III shows the changes obtained extracting a quasi-regular topology from the solution; the objectives are normalized to F. Figures 5 and 6 illustrate the optimized regular and quasi-regular topologies for 10-node architectures. Fig. 7 displays the CNs geographical distribution for the optimized 34-node networks. The solutions for 10-node networks have an affordable execution time. For 34-node networks we contain the execution time setting the CPLEX upper cut-off value exploiting the objective values obtained through a heuristic tool [8] opportunely modified to adapt it to our network model (we do not describe it for page limits).

TABLE II

RFA SOLUTION

Case	10A	10B	34A	34B
Objective	2281804	2155353	31995440	42596082
Time (s)	169	100	43452	1685
Fibers cost	77.8%	83.27%	82.46%	81.27%
CNs cost	11.22%	11.88%	5.44%	5.26%
Delays cost	10.98%	4.86%	12.1%	13.47%
μ_R	17.91%	15.15%	16.83%	12.39%

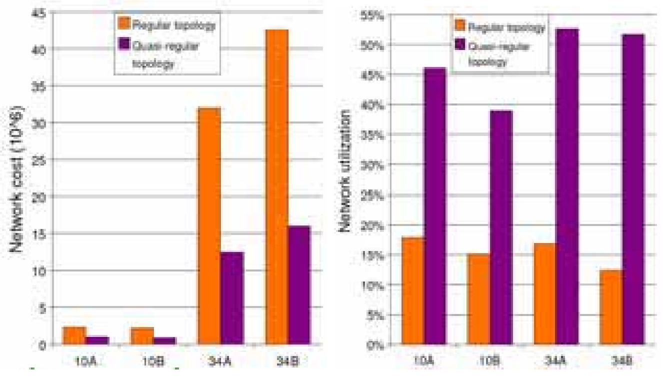
TABLE III

RFA SOLUTIONS CHANGES USING A QUASI-REGULAR TOPOLOGY

Case	10A	10B	34A	34B
Cost	982492	840006	12406718	15976542
Fibers cost	63.13%	73.65%	63.81%	60.41%
CNs cost	11.36%	13.88%	4.97%	3.67%
Delays cost	25.51%	12.46%	31.21%	35.91%
μ_R	48.36%	38.97%	52.69%	51.73%

The results with the A matrixes provide better values of the network utilization coefficient (μ_R). The explanation is that the B traffic matrixes are dense and present many CRs with low traffic demands; thus, the links used by these CRs are under-used. What are the changes if we can use the quasi-regular topology? As it can be assessed by Table III, the network cost is dramatically reduced, more than 50% (Fig. 4a), and the network utilization is doubled (Fig. 4b). This is due to the fact that the regular topology demands the allocation of too many unused fibers. Blouin et al. [5] estimated the quantity of fiber requested by a regular Petaweb as roughly the 17% more than an equivalent mesh architecture; now, we reduced considerably the quantity of fiber to install; for example, the km of installed fibers with the quasi-regular topology reduced of 65% for 10A, and of 72% for 34B. A quasi-regular Petaweb can, thus, compete more concretely with the classical proposed mesh topologies.

In Fig.5 and Fig.6 for every connection between the EN and the switching site we indicate upon it the number of fibers to install. We now analyze in detail the regular topology for the 10A model (Fig. 5a): we have two enabled core nodes, a CN-2 in Philadelphia and a CN-3 in Washington. The optical trunk lines between every switching site and every EN are composed of two optical links in opposite directions. An optical link is composed by s_r fibers; thus, the ones connected to the CN-2 in Philadelphia are composed of 2 fibers, and the ones connected to the CN-3 in Washington are composed of 4 fibers. In the quasi-regular topology (Fig. 5b) the unused fibers were disabled and not considered in the solution. Entire trunk lines have been disabled because their optical links were totally unused: the Washington-Albany one and the Philadelphia-Tallahassee one. Furthermore, the number of fibers per trunk line has generally decreased: the optical links connected to the



(a) Network cost (b) Network utilization

Fig. 4. Comparison for a quasi-regular and a regular topology

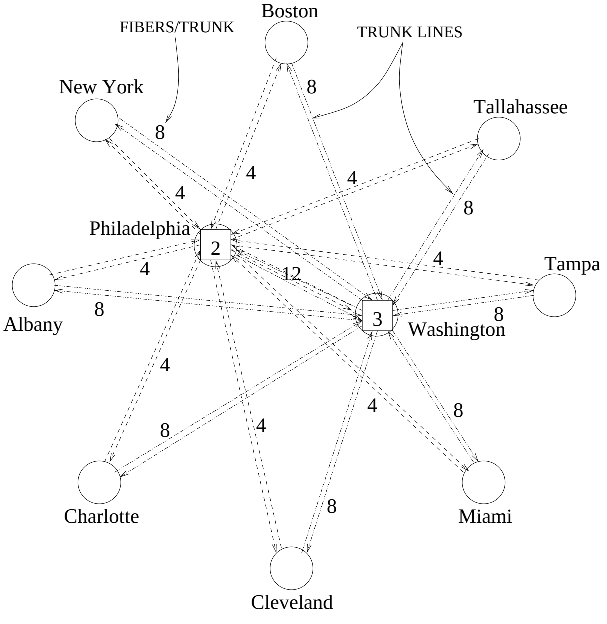
CN-2 and the CN-3 are now composed of only one and two fibers (instead of 2 and 4), except the one with Miami, and the one with New York; the EN in New York asks more than 800 Gbit/s to be switched, more than any other EN, and forces the opening of high type CNs.

For the model 10B (Fig. 6), on the contrary, one can observe that there is not any optical trunk line disabled. Why? Because the traffic matrix of 10B is dense, every EN is fully connected with the others and there are more TLPs to be switched.

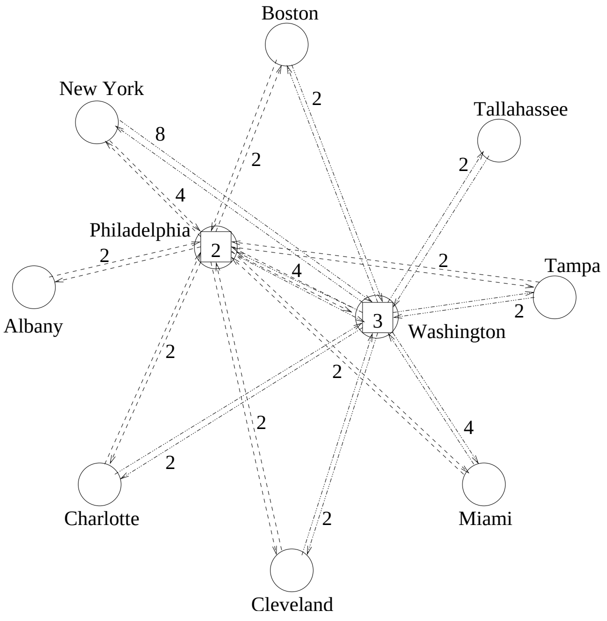
One has to pay attention to the optical trunk line Philadelphia-Washington that is not a direct connection between the CNs, but a trunk composed by optical links connecting the two ENs with the two CNs.

Therefore the cost reduction concerns unused fibers cost and disabled ports cost; consequently the cost allocation changes. In Fig. 8 we compare the cost allocation assuming the two topologies for 10-node cases. Adopting a quasi-regular topology, the weight of fibers cost decreases more than 10 percentage points and, thus, the weight of delays and CNs costs increases; the delays cost assumes a weight of more than 10 percentage points than in the case with the regular topology where the fibers cost is over-estimated; the CNs cost increases even if its absolute value decreases because the cost reduction due to fibers is more important than that due to ports cost. In the case of 34-node networks it is evident that the CNs cost becomes unimportant at the expense of the delay propagation cost. And this is even more evident with a quasi-regular topology. The presence of high order CNs allows to assign the TLPs to few trunk lines and to decrease to total CNs number.

As a consequence of the deletion of unused fibers and links, with the quasi-regular topology the network utilization increases very significantly and this indicates that, on average, the network fibers are used almost at the 50% of their transport capacity. This is a good result, considering that, thanks to the time-sharing we have better exploited the lambda-channels; the idle capacity is available for traffic restoration and for fornishing bandwidth to other licenced operators.



(a) 10A regular



(b) 10A quasi-regular

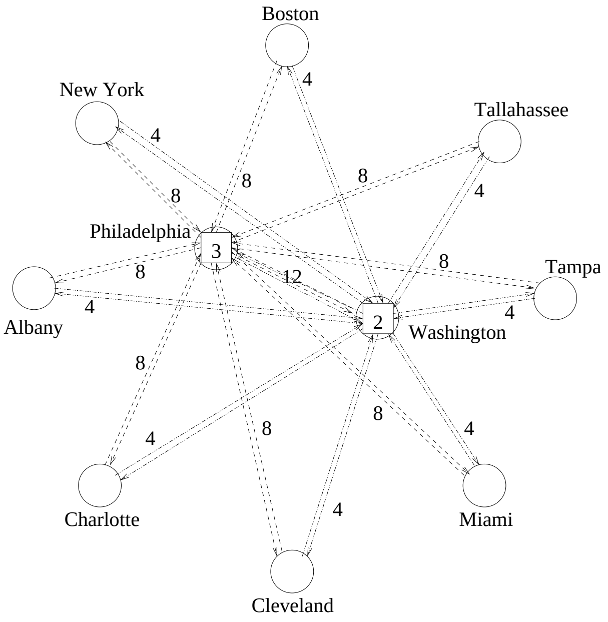
Fig. 5. Route and Fiber Allocation solution for 10A case

Nevertheless, one can observe a problem in the solutions with a quasi-regular topology: there are ENs that are connected through only one trunk line to the transport network (Fig.5). The same problem may appear in an optimized regular topology where all the CNs are installed in the same site. There is no reliability in such networks. To tackle this problem, we explore a reliable Petaweb architecture in [1].

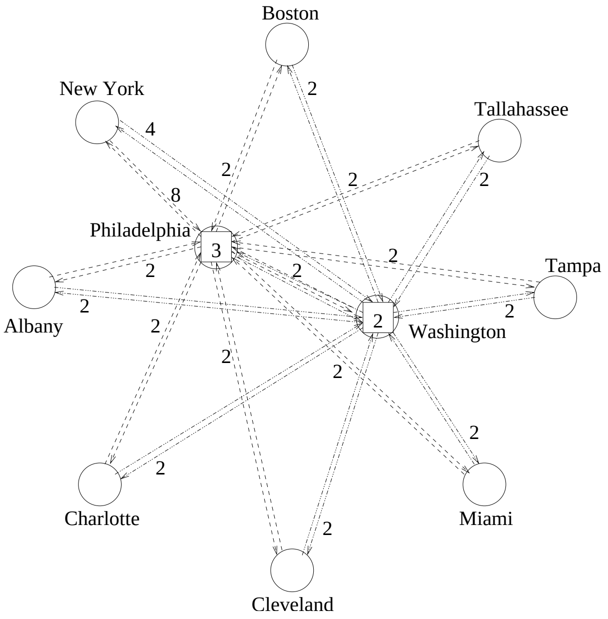
B. WTA results

The WTA task is solved with linear complexity. In a 10-node network with two enabled CNs we have a total of 40 optical links. We report a part of the 10A WTA solution concerning only two optical links.

We have to assign time-slots and wavelengths of every

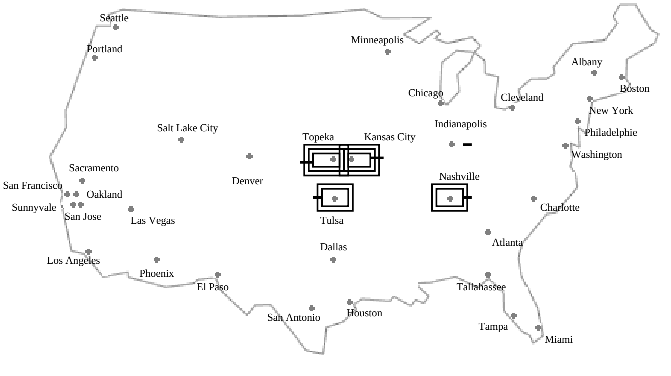


(a) 10B regular

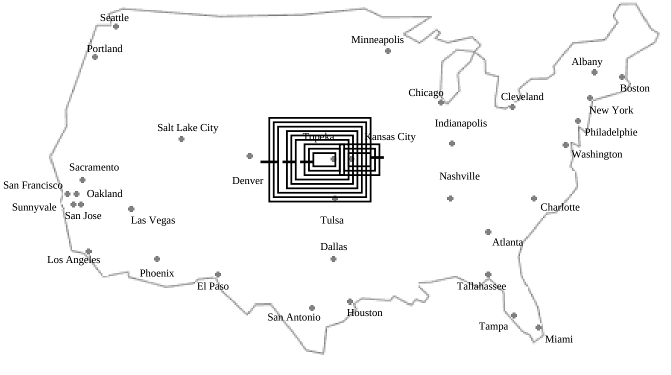


(b) 10B quasi-regular

Fig. 6. Route and Fiber Allocation solution for 10B case

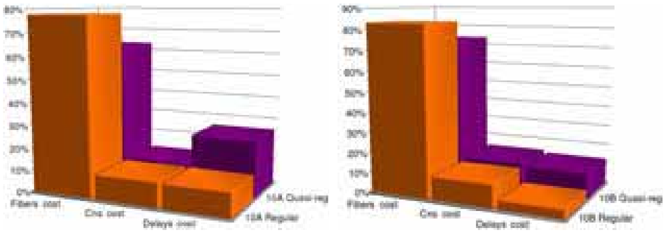


(a) 34A



(b) 34B

Fig. 7. Core nodes geographical distribution



(a) 10A (b) 10B

Fig. 8. Cost allocation for regular and quasi-regular topologies

optical link to the TLPs whose traffic has been allocated on that link. Here we consider the optical link from Tallahassee to Washington and the one from Washington to Tampa, following the route of the outgoing TLPs of the EN in Tallahassee. Tallahassee and Tampa are connected to the CN-3 in Washington through only one optical link per direction. The connection requests are $CR_{9,8}$=1.6 Gbit/s and $CR_{9,10}$=0.2 Gbit/s; 3 TLP-1s serve $CR_{9,8}$, and 1 TLP-1 serves $CR_{9,10}$. The RFA solution indicates that the optical link Tallahassee-Washington has to transport only those TLP-1s (then the TLP-1 of $CR_{9,10}$ is directly dropped in Washington), while the optical link Washington-Tampa serve even TLPs of others CRs. Fig. 9 illustrates the assignment of time-slots and wavelengths.

Following the WTA algorithm one can notice that the TLP-1s of $CR_{9,8}$ and $CR_{9,10}$ have been assigned to the first four time-slots on λ_1 of fiber 1 on the optical link Tallahassee-Washington. The other 3 fibers, 15 wavelengths and 12 time-slots of that optical link remain unused. Then the TLP-1s of $CR_{9,8}$ are routed on the optical link between Washington and the destination EN in Tampa, where they occupy the third, the forth and the fifth time-slots on λ_2 of the fiber 1. On this optical link three fibers and one wavelength remain unused. The other wavelengths and time-slots are assigned to the TLPs of other CRs having as destination the EN in Tampa.

What about the unused fibers? If one considers a regular

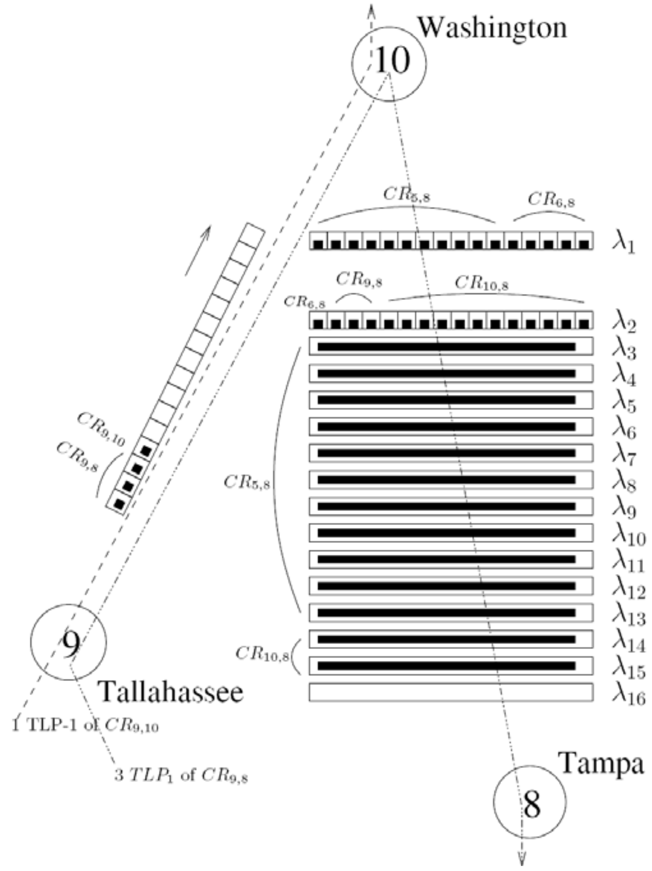


Fig. 9. WTA for two optical links of the 10A solution

topology the Tallahassee-Washington trunk line would have an utilization of 0.39% and the Washington-Tampa one of 23.44%; otherwise with a quasi-regular topology one would have an utilization of, respectively, 1.56% and 93.75%.

Fig. 9 illustrates how ts-lightpaths of different classes are assigned to different medium. While the outgoing TLPs of Tallahassee are of the lower class, and are assigned to a time-slot each, others TLPs of class 2 are routed on the used fiber between Washington and Tampa. The benefits of using many hierarchically ordered classes of service is evident, because we can use efficiently the link capacity thanks to the TDM over WDM; in the link Washington-Tampa we could set up 45 independently manageable ts-lightpaths on a fiber with only 16 wavelengths.

V. CONCLUSIONS

The results showed how a Petaweb network architecture can be planned. The proposed planning procedures are the first ones that tackle a realistic design for this innovative architecture. We showed two possible topologies for an optimized network, a regular one with high fiber cost, and a quasi-regular one that costs less and presents higher network utilization. Currently we are working on an upgrade procedure with regular and quasi-regular topologies considering scalability issues for the Petaweb network architecture.

The complexity of the RFA problem remains high for large instances, but using a modified version of the heuristic in [8] we succeeded in obtaining fastly results very close to the optimum. Actually, we are defining other ILP approaches for the design of a quasi-regular topology directly and not heuristically. Further works may analyze the planning of a Petaweb architecture adapting it partially to an existing infrastructure. Other suggestions are contained in [1] where we describe the issues that relate to the design and optimization of a reliable Petaweb architecture.

ACKNOWLEDGMENTS

The authors would like to express their gratitude to A. Pattavina for having allowed the collaboration that produced this work, to M. Beshai and F. Blouin of Nortel Networks for proposing the problem, to A. Reinert for her availability, and to A. Ceselli for his suggestions. This work was supported by NSERC strategic grant nb. STPG 246/59.

REFERENCES

[1] S. Secci, B. Sansò, "Optimization of a Dedicated Path Protected TDM/WDM Petaweb Architecture", in *Proceedings of Networking and Electronic Commerce Research Conference (NAEC 2006)*, to appear, Oct. 19-22, 2006, Riva Del Garda, Italy.

[2] "Interfaces for the Optical Transport Networks (OTN)", G.709/Y.1331, ITU-T International Communication Union, Mar. 2003.

[3] F. J. Blouin, S. Yazid, B. Bou-Diab., "Emulation of a vast adaptative network", presented at *Networks 2000 - Toward Natural Network: 9^{th} International Telecommunication Network Planning Symposium*, Toronto (2000).

[4] R. Vickers, M. Beshai, "Petaweb architecture", presented at *Networks 2000 - Toward Natural Network: 9^{th} International Telecommunication Network Planning Symposium*, Toronto (2000).

[5] F. J. Blouin, A. W. Lee, A. J. Lee, M. Beshai, "A comparison of two optical-core networks", *Journal of Optical Networking*, vol. 1, no. 1, pp. 56-65 (2001).

[6] M. Beshai, F. J. Blouin, "Courteous routing", presented at *Networks 2000 - Toward Natural Network: 9^{th} International Telecommunication Network Planning Symposium*, Toronto (2000).

[7] J. P. G. Sterbenz, L. Chapin, R. Krishnan, "Routing issues in interconnecting IP Networks with the Petaweb", presented at *Networks 2000 - Toward Natural Network: 9^{th} International Telecommunication Network Planning Symposium*, Toronto (2000).

[8] A. Reinert, B. Sansò, S. Secci, "Design Optimization of the Petaweb Architecture", GERAD 2004-87, submitted to IEEE Transactions on Networking.

[9] J. Dégila, B. Sansò, "Topological design optimisation of a yottabit-per-second lattice network", *Journal on Selected Areas in Communications*, vol.22, no. 9, pp. 1613-1625 (2004).

[10] N. F. Huang, G. H. Liaw, C. P. Wang, "A novel all-optical transport network with time-shared wavelength channels", *Journal on Selected Areas in Communications*, vol. 18, no. 10, pp. 1863-1875 (2000).

[11] M. Sivakumar, S. Subramannian, "A performance evaluation of time-switching in TDM wavelength routing networks", in *Proceedings of Broadband Networks, 2004. BroadNets 2004. First International Conference on*, San Jose, CA, Oct. 2004.

[12] A. Pattavina, M. Tornatore, A. De Fazio, G. Maier, M. Martinelli, "Static WDM network planning with TDM channel partitioning", in *Proceedings of Networking 2004*, Athens, Greece, 2004.

[13] O. Bilde, J. Krarup, "Sharp lower bounds and efficient algorithms for the simple plant location problem", *Annals of Discrete Mathematics*, vol. 1, pp. 79-97 (1977).

[14] H. Pirkul, "Efficient algorithms for the capacitated concentrator location problem", *Computers and Operations Research*, vol. 14, no.3, pp. 197-208 (1987).

[15] P. Gambini et al., "Transparent optical packet switching: Network architecture and demonstrators in the KEOPS project", *Journal on Selected Areas in Communications*, vol. 16, no. 7, pp. 1245-1259 (1998).