

# Direct Optimal Design of a Quasi-Regular Composite-Star Core Network

Stefano Secci<sup>1,2</sup>, Alberto Ceselli<sup>3</sup>, Federico Malucelli<sup>1</sup>, Achille Pattavina<sup>1</sup>, Brunilde Sansò<sup>4</sup>

<sup>1</sup>Dipartimento di Elettronica e Informazione, Politecnico di Milano, Piazza Leonardo da Vinci, 32, 20133 Milano, Italy

<sup>2</sup>Département Informatique et Réseaux, Ecole Nationale Supérieure des Télécommunications (ENST), 37/39, rue Dareau, 75014 Paris, France

<sup>3</sup>Dipartimento di Tecnologie dell'Informazione, Università degli Studi di Milano, via Bramante 65, 26013 Crema, Italy

<sup>4</sup>Département de Génie Electrique, Ecole Polytechnique de Montréal, C.P. 6079, Succ. Centre-ville, Montréal (Québec) Canada H3C 3A7

E-mail: secci@enst.fr, ceselli@dti.unimi.it, {malucelli, pattavina}@elet.polimi.it, brunilde.sanso@polymtl.ca

**Abstract**—In this paper we consider the optimization of a Petaweb backbone, a core-based composite-star infrastructure capable of accommodating petabit-per-second traffic volumes. The problems of designing both regular and quasi-regular Petaweb structures have recently been discussed in the literature.

Our contribution is twofold. First, we propose a new model to optimally plan the regular Petaweb structure: using general purpose optimization software with our formulation allowed us to obtain provably optimal solutions even for large networks. Second, we tackle the problem of directly optimizing the quasi-regular topology: we introduce an Integer Linear Programming (ILP) formulation and devise an ad-hoc heuristic algorithm.

We illustrate through experimental results that we can obtain in this way very good configurations for the original architecture.

## I. INTRODUCTION

For many years operators and carriers have been moving towards wide area backbones capable of offering terabit-per-second PoP-to-PoP data connections. Services like video-on-demand, layer-1 circuits for VLAN and interprovider connections require high resource availability between the PoPs and connections with better-than-best-effort constraints, fast resources allocation and restoration.

To meet those requirements, a composite-star optical transport architecture, called the Petaweb, has been proposed by Nortel Networks [1]. It is a high-capacity core-based network structure capable of operating at a global rate of the order of the petabit-per-second ( $10^{15}$  bps). A set of 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. The idea behind such infrastructure is to offer a very large transport capacity through a regular topology so as to facilitate upgrades and simplify traffic engineering issues. Difficult design problems arise, as the problem of locating and dimensioning core nodes, determining the switching schemes and the optimal resource assignment.

Important work has been conducted on the subject. In [2] an emulation of the Petaweb is explained whereas in [3] it is compared to typical mesh networks. A routing algorithm for the Petaweb is proposed in [4] and in [5] IP/Petaweb scenarios are studied.

The Petaweb design problem was formulated in [6]. The design considering TDM channel partitioning was presented in [7], and [8] introduced the notion of dedicated path protection;

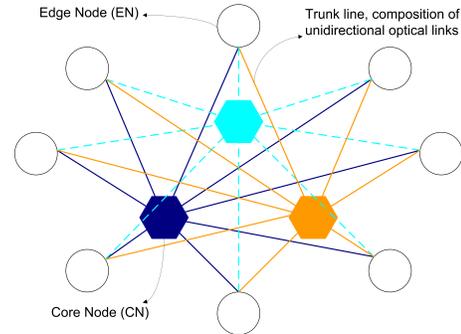


Fig. 1. Composite-star Petaweb structure

in the two last papers an inexpensive quasi-regular Petaweb architecture was also proposed and studied: it can offer more than 55% savings in network cost yielding significant improvements in network utilization. Encouraged by those results, in this paper we propose methods for the direct optimal design of a quasi-regular Petaweb, showing how we can further improve those results. We introduce an alternative formulation from those presented in [7] and [8] that will allow the resolution of larger instances by relaxing some modelling features, to support the direct optimization of the quasi-regular topology.

The paper is organized as follows. The Petaweb architecture and the network model are described in Section II and Section III respectively. In particular, in Section II we highlight advantages and disadvantages of the proposed architecture, and in Section III we characterize the quasi-regular topology. In Section IV the new ILP model for the optimal design of regular networks is described. In Section V we tackle the direct design of a quasi-regular topology: an ILP formulation to be used by general purpose solvers and a three-step heuristic are presented. Results are analyzed and compared in Section VI. In Section VII we draw some conclusions.

## II. A CORE-BASED COMPOSITE-STAR ARCHITECTURE

The Petaweb architecture is composed of Edge Nodes (ENs) connected through Core Nodes (CNs) as shown in Fig.1. Every EN is connected to every CN, and neither ENs nor CNs are directly connected with one another. Such a configuration allows a two-hop optical circuit between ENs through a CN.

An *edge node* is an electronic node that asks the transport

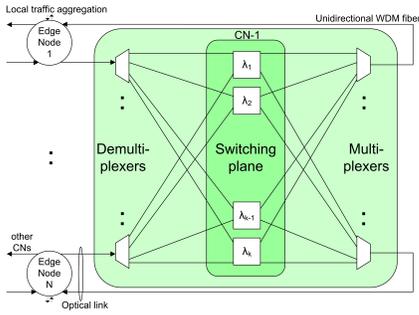


Fig. 2. Parallel-planes optical core node

network to supply media for its set of source-destination static connection requests with other edge nodes. The connection between  $N$  edge nodes and a core node is shown in Figure 2. Every edge node is connected to a core node through one *optical link*, composed of one or more optical fibers. We suppose the use of one or more optical fibers that we assume to be unidirectional, so that an edge node has one optical link incoming and one outgoing for every core node. We suppose that all the fibers carry the same number of wavelengths.

A *core node* is a set of arrays of parallel space switches, also called switching planes. We use three types of core nodes. The number of switching planes identifies the type of a CN: a CN of a low type has half the switching planes of a CN of a higher type. Every switching plane is connected to every EN with one fiber per direction. All the incoming WDM fibers are demultiplexed into their different lambda-channels, each of which is connected to the corresponding space switch of the respective array. Given the regularity of the architecture, core nodes do not require wavelength converters, and no continuity constraint needs to be considered. Which renders the Petaweb an all-optical network. 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 an all-optical parallel-plane structure increases the reliability of CNs because a hypothetical failure in a switching plane would affect only the connections on that plane. In [3] Blouin supposed the use of Time Division Multiplexing (TDM) in the Petaweb to realize sub-channels within a lambda-channel. To use the TDM we need to extend the functionalities of switching core nodes; in [7] authors proposed the replacement of the switching plane with the compatible all-optical TDM Wavelength Space Routers of Huang [9], recently evaluated in [10], which multiplexes in a time-slot basis remaining in the optical domain and, thus, without any buffering operation. In what follows, as suggested in [9], we call a time-slotted lightpath a *ts-lightpath* (TLP).

The edge node locations define the set of network sites. Multiple CNs can be installed in the same site. A network site in which CNs are installed becomes a *switching site*. The physical connection between the site of an EN and a switching site can be composed of many optical links because more CNs can be installed in a network site. From now on we call this physical connection *optical trunk line*.

### A. Advantages vs disadvantages

The advantages stem from the simplicity of the architecture:

- The 2-hop edge connectivity guarantees fully utilization of the available capacity: the bottlenecks at intermediate links that characterize mesh networks are not possible.
- Fast resource allocation, thanks to the absence of resource negotiation between core nodes.
- The architecture based on independent core nodes offers simple routing and traffic engineering [1].
- Due to dedicated path protection, the network is robust to site, node and link failures [8].
- Easy and inexpensive upgrades. In [11] it is shown that the upgrade of an existing architecture can efficiently and cheaply exploit the existing capacity and equipment.
- The worrisome queueing delay in the core is entirely eliminated because the core is bufferless.

These characteristics yield an architecture particularly suitable for carrier backbones: edge and core nodes may be backbone IP routers and WDM switches at different distant PoPs; trunk lines may represent long distance optical links that practically would be a concatenation of physical links subject to optical regenerations (our model is transparent to such additive costs).

The main disadvantages of adopting a regular Petaweb architecture are a high fiber-consuming infrastructure and an a-priori big propagation delay.

In [3] it is shown that the Petaweb architecture requires roughly 17% more fiber than an equivalent mesh network; nevertheless, the authors considered the descending trend for fiber cost encouraging. To bypass the high cost of a regular Petaweb, authors in [7] showed how a quasi-regular structure, obtained by downgrading an optimized regular one, can be dimensioned with a much smaller cost while preserving the regularity (i.e. still allowing to upgrade a quasi-regular network to the regular one by simple activation of CN ports and addition of fibers).

Considering the propagation delay as a virtual cost in the dimensioning of a Petaweb architecture, in [7] and [8] short lightpaths are created by minimizing a unique network cost function. Nevertheless, fiber length or propagation delay are usually not considered in the dimensioning of optical networks due to the high transmission speed.

Another issue is the survivability: since a core node could commute terabit-per-second data flow, a failure to a whole core node or the isolation of a switching site may be disastrous. This aspect was tackled in [8]. As discussed in the next session, our methods guarantee against both path and node failures by applying dedicated path protection.

Cost, delay and survivability are thus surmountable drawbacks of this particular composite-star architecture.

It is worth mentioning that the Petaweb is a scale-free network architecture: it contains few hub sites and most sites have just few connections. An edge node prefers a current switching site instead of requiring a core node in a new site. Resilience properties of scale-free networks are inherited. While in a mesh network the accidental failure of a number of

links can fracture the system into non communicating islands, this can not happen with a composite-star network, which is, however, highly vulnerable to coordinated attacks against its hubs. We will show how, by designing directly a quasi-regular topology, the hubs number increases reducing considerably the network vulnerability.

### III. NETWORK MODEL

In what follows, we first provide an overall view of the model presented in [7], then we characterize the quasi-regular topology and show the effects of applying path protection.

A set of permanent connection requests with static traffic is served upon the Petaweb physical topology through a set of ts-lightpaths (TLPs). Three classes of TLPs are employed: the first (TLP-1) uses the transport capacity of a time-slot, the second (TLP-2) the capacity of a whole wavelength while the third class (TLP-3) uses a whole optical fiber. Let us indicate by  $Z_h$  the transport capacity of a ts-lightpath of class  $h$ , by  $C_{ch}$  the capacity of a lambda-channel, and by  $W$  the number of wavelength per fiber. We have, thus,  $Z_2 = C_{ch}$  Gb/s and  $Z_3 = WC_{ch}$ .  $Z_1$  should be chosen with respect to the customer requirements and to the equipment to be employed.

The network design problem consists in the minimization of a network cost composed of the core node cost, the fiber cost and the propagation delay cost. The *CN cost* is composed of a fixed cost and a variable cost. Let us indicate by  $s_r$  the number of switching planes identifying the type  $r$  of a CN (indicated with CN- $r$ ): the optical link connecting an EN to a CN- $r$  has  $s_r$  unidirectional optical fibers, one for every switching plane. As previously mentioned,  $s_r$  is such that  $s_r = 2s_{r-1}$ . The fixed cost  $h_r$  depends on the type of the CN and is such that  $h_r > h_{r-1} > \dots > h_1$ . The variable cost is determined by the number of active ports. An active port has a cost  $P$  scaled for higher types. Let us indicate by  $N_{en}$  the number of ENs and by  $\gamma$  the scale factor for  $P$ , the global cost of a CN- $r$  is  $K_r = h_r + 2N_{en}W s_r P \gamma^{(s_r-1)}$ . Typically  $K_r < 2K_{r-1}$ .

The *fiber cost* is indicated as  $F$  and is considered per unit of length. It is the cost of a reference fiber type, which is then scaled by a discrete function  $\phi(W)$  that depends on the number of wavelengths. Let us indicate by  $\Delta_{ij}$  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  $\beta$ , is proportional to the distance traveled and to the lightpath bitrate. The reason for bitrate dependence is that we can assure higher QoS to high bitrate lightpaths, which are supposed to transport aggregates requiring low propagation delay. We add this virtual cost to real network costs to account for the common thought that the propagation delay will become one of the most important pay-back factors in offering end-to-end circuits, especially in a multi-provider scope [13].

The design method has to deal with the capacity bounds of edge nodes and optical links. The resources can be allocated and incremented only by discrete quantities: the link capacity

can be increased by a multiple of the capacity of  $W$  lambda-channels at a time; the capacity of an EN depends on the number of optical fibers connected to it. Furthermore, to control buffering delay all the TLPs of a Connection Request (CR) should be transported on the same optical trunk line.

The design problem consists in finding the best physical topology for the given traffic matrix, and in assigning the corresponding TLPs to communication media (wavelengths and time-slots). Differing from mesh networks design, in the Petaweb design problem the physical topology is not pre-assigned. The set of possible routes is not established a priori. Indeed, the optimization consists in dimensioning the physical topology, choosing the core nodes location and their size, and allocating the requested fiber resources.

#### A. The quasi-regular topology

In the dimensioned network, only a portion of the transport capacity would be assigned and many optical fibers or links may be totally unused. This may occur mainly if the traffic matrices contain few peaks of traffic between two sites, and a lot of medium-low values; the peaks would cause high load on the optical links connected to the related sites, while other optical links, interesting sites with low CRs, would remain under-used. Indeed, the activation of a switching plane requires the installation of one fiber for every edge node.

The quasi-regular topology is a variant to the composite-star regular topology: we activate only fibers which are actually needed, neglecting the others. For every optical link if the number of fibers needed by the assigned TLPs is inferior to  $s_r$ , the superfluous fibers are no longer considered. Therefore this operation does not imply an undifferentiated grooming of TLPs of different optical links on the same fibers. Regularity is preserved and will remain attainable through re-installation of the disabled fibers (consider that the initial requirement of Nortel Networks was the dimensioning of a regular Petaweb [7]). 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 with the corresponding CN ports, the network cost decreases. Additional physical hypothesis should, however, be assumed with quasi-regular topologies: the switching planes of the same core node must be able to cooperate in order to multiplex lightpaths on the same fiber, and switches should be equipped with wavelength converters.

#### B. Dedicated Path Protection

Fig. 3 reports optimized Petaweb topologies obtained by using the procedure proposed in [7]; we indicate over every trunk line the number of fibers of which it is composed. The quasi-regular topology is there determined heuristically as reduction of an optimal regular Petaweb topology, removing unused fibers and ports without any change in lightpath routes and switching schemes. The drawback is the absence of reliability. The reader is referred to the Tallahassee and Albany edge nodes in Fig. 3b, which are connected to the network

through only one trunk line. In case of trunk failure those edge nodes would remain totally isolated from the network and their outgoing and incoming traffic could not be restored at all. This would happen even when all core nodes are located in the same switching site (a reasonable possibility for small networks): all the edge nodes would be connected to the transport network through only one trunk line<sup>1</sup>.

Dedicated Path Protection (DPP) was the protection strategy applied in [8]. For every working ts-lightpath (wTLP) a protection ts-lightpath (pTLP) is allocated. In case of one trunk line failure all the wTLPs are recovered from the allocated pTLPs. In the 1 + 1 protection, the signal is always split in the protection path, with no signaling disrupting the data reception; in the 1:1 modality, the signal is split only in case of failure, little signaling is required, and it makes sense to attribute shorter paths to the working lightpaths. The DPP strategy requires the use of a *protection constraint*: every pTLP should be multiplexed on trunk lines different than those of the corresponding wTLP; this corresponds to forcing a pTLP to be switched to a different network site than that of its wTLP.

The protection constraint guarantees even *node protection* and *site protection*: if a core node fails, all its TLPs can be recovered; if a whole site is disconnected pursuant to a disaster, all the traffic switched by its core nodes can be recovered elsewhere. While in mesh networks the application of node or site protection constraints create irregularities that worsen the network utility, in the Petaweb new equipment is added without generating bottlenecks and inefficiencies.

To illustrate the differences with and without DPP, the reader is referred to Fig.4, which shows the 10-node quasi-regular topology with DPP [8], to compare with the topologies in Fig.3 [7]: while before many edge nodes were connected to the network through only one trunk line, this does not happen in the second case, even if a larger amount of fibers is needed. A reliable quasi-regular structure can thus be planned.

#### IV. OPTIMIZATION OF A REGULAR ARCHITECTURE

It has been demonstrated that the Petaweb design problem is NP-hard [6], by reduction from a Single Source Facility Location Problem. In [7] and [8] the optimal design of a regular architecture is tackled: each connection request is initially split into one or more TLPs, and afterwards the optimization problem of locating CNs and assigning TLPs to CNs is solved by ILP. In this section, we propose to design the regular topology with an alternative approach, by relaxing some modeling features: we start by designing the least expensive network which can satisfy the CRs; then, when the CNs are located and the CRs are assigned to the switching sites, we decide how to split each CR into the minimum number of TLPs. We consider the case involving dedicated path protection.

We complete the notation introduced in Sect.III as follows:

- $M$ : the set of sites

<sup>1</sup>The trunk lines connecting two switching sites are not links between core nodes, but links between edge nodes and core nodes, and vice-versa

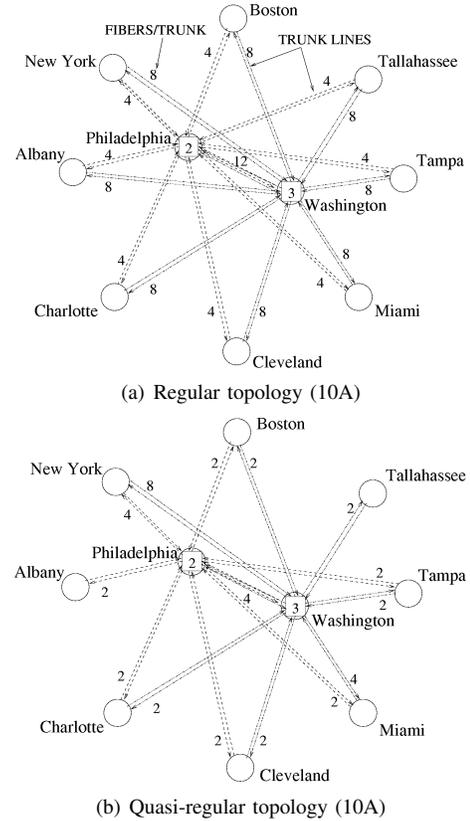


Fig. 3. Optimized 10-node Petaweb

- $V$ : the set of types of core nodes
  - $Q$ : the set of connection requests
  - $C_{\min}$ : the maximum edge node capacity
  - $w(q)$ : the bandwidth consumption of CR  $q$
  - $\delta < 1$ : the scaling factor for pTLPs delay costs
  - $d_{iq}$ : the distance from the origin of CR  $q$  to site  $i$  plus the distance from site  $i$  to the destination of CR  $q$
  - $C_r = (Z_2/Z_1)W_{s,r}$ : capacity of a core node of type  $r$
- and we consider the following decision variables:
- $a_q^i$ : a binary variable equal to 1 if CR  $q$  is switched at site  $i$
  - $\tilde{a}_q^i$ : a binary variable equal to 1 if the protection flow for CR  $q$  is switched at site  $i$
  - $y_{ir}$ : integer variable equal to the number of core nodes of type  $r$  to install at site  $i$

Capacities  $C_r$  and  $C_{\min}$  are measured in time slots; these terms are computed by dividing each value by  $Z_1$  and rounding down to the nearest integer. In a similar way, each bandwidth consumption value  $w(q)$  is measured in time slots, dividing by  $Z_1$  and rounding up to the nearest integer.

The problem of finding the optimal regular Petaweb can be stated as follows:

$$\min G(y, a, \tilde{a}) = \sum_{i \in M} \sum_{r \in V} (K_r + F_{ir}) y_{ir} + \sum_{i \in M} \sum_{q \in Q} \beta d_{iq} w(q) Z_1 a_q^i + \sum_{i \in M} \sum_{q \in Q} \delta \beta d_{iq} w(q) Z_1 \tilde{a}_q^i \quad (1)$$

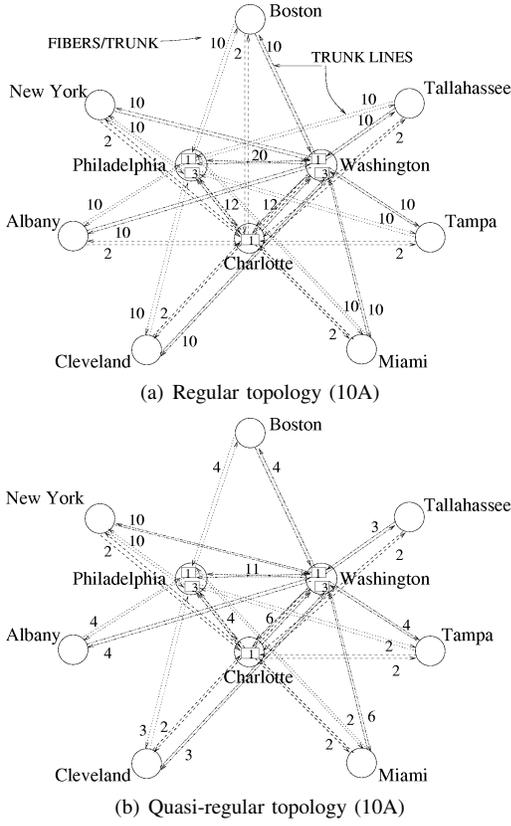


Fig. 4. Optimized 10-node Petaweb with Dedicated Path Protection.

$$\text{s.t. } \sum_{i \in M} a_q^i = 1 \quad \forall q \in Q \quad (2)$$

$$\sum_{i \in M} \tilde{a}_q^i = 1 \quad \forall q \in Q \quad (3)$$

$$a_q^i + \tilde{a}_q^i \leq 1 \quad \forall q \in Q, \forall i \in M \quad (4)$$

$$\sum_{q | \text{head of } q=j} w(q)(a_q^i + \tilde{a}_q^i) \leq \sum_{r \in V} C_r y_{ir} \quad \forall i \in M, \forall j \in M \quad (5)$$

$$\sum_{q | \text{tail of } q=j} w(q)(a_q^i + \tilde{a}_q^i) \leq \sum_{r \in V} C_r y_{ir} \quad \forall i \in M, \forall j \in M \quad (6)$$

$$\sum_{i \in M} \sum_{r \in V} C_r y_{ir} \leq C_{\min} \quad (7)$$

$$y_{ir} \in Z_+, a_q^i, \tilde{a}_q^i \in \{0, 1\} \quad \forall i \in M, \forall q \in Q, \forall r \in V \quad (8)$$

The objective (1) is to minimize the network cost; this is composed of the core node cost, the fiber cost, and a term accounting for the propagation delay of wTLPs and pTLPs, respectively. Constraints (2) and (3) impose a unique switching and protection site on each CR  $q$ , while constraints (4) ensure that working and protection paths of each CR are disjoint. Constraints (5) and (6) limit the number of optical links ending to and starting from a given switching site to the number of installed ports. Inequality (7) imposes a maximum capacity on each edge node, while conditions (8) restrict the variables to take integer and binary values.

By dimensioning a regular network with model (1)-(8) we obtain the number of core nodes to install, their location and their size (as depicted in Fig.4a). Moreover, the number of optical links and fibers per trunk line can easily be computed by the number  $s_r$  of ports in each core node. Because of variable aggregation, the complexity of this model is significantly lower than that reported in [7]. This allowed us to obtain solutions in reasonable time even for large networks (see Sect.VI).

Finally, we remark that this new method requires that once the optimal design is obtained, the problem of splitting the traffic of each CR into TLPs has to be solved, and these TLPs have to be assigned to time-slots, wavelengths and fibers; the optimal solution may require to split more TLP-3s and TLP-2s than [7] in order to fully exploit the available capacity. From the computational point of view, this involves solving a set of *bin packing problems with item fragmentation* [15]. However, this proved not to be an issue. In fact, the size of the instances for these subproblems is typically very small, and their optimization can easily be tackled by general purpose solvers. Therefore, we do not analyze this last step in the subsequent sections.

## V. DIRECT DESIGN OF A QUASI-REGULAR PETAWEB

After the optimal dimensioning of a regular Petaweb, a quasi-regular topology can be retrieved by simply dropping unused fibers and ports, as displayed in Fig.4b; however, this approach does not guarantee the optimality of the obtained quasi-regular topology. In this section we tackle the problem of optimizing a quasi-regular Petaweb directly. We model the problem and propose a three-step heuristic.

### A. Modeling the quasi-regular design problem

We introduce three new sets of variables:

- $f_{q(ire)}$  is the number of timeslots uploaded by the source node for CR  $q$  to the  $e^{th}$  CN- $r$  at site  $i$  (which is also the number of timeslots downloaded by the destination);
- $l_{j(ire)}^u$  (resp.  $l_{(ire)j}^d$ ) is the number of fibers connecting EN  $j$  and CN  $(ire)$  (resp. CN  $(ire)$  and EN  $j$ ).

The exact optimization of the quasi-regular topology can be pursued by solving the following ILP:

$$\begin{aligned} \min G(y, l, a, \tilde{a}) = & \sum_{i \in M} \sum_{r \in V} h_r y_{ir} + \\ & + \sum_{j \in M} \sum_{i \in M} \sum_{r, e} (\phi(W)\Delta_{ij}F + W\gamma^{s_r-1}P)l_{j(ire)}^u \\ & + \sum_{j \in M} \sum_{i \in M} \sum_{r, e} (\phi(W)\Delta_{ij}F + W\gamma^{s_r-1}P)l_{(ire)j}^d \\ & + \sum_{i \in M} \sum_{q \in Q} \beta d_{iq} w(q) a_q^i + \sum_{i \in M} \sum_{q \in Q} \delta \beta d_{iq} w(q) \tilde{a}_q^i \end{aligned} \quad (9)$$

$$\text{s.t. } \sum_{i \in M} a_q^i = 1 \quad \forall q \in Q \quad (10)$$

$$\sum_{i \in M} \tilde{a}_q^i = 1 \quad \forall q \in Q \quad (11)$$

$$a_q^i + \tilde{a}_q^i \leq 1 \quad \forall q \in Q, \forall i \in M \quad (12)$$

$$\sum_{re} f_{q(ire)} = w(q)(a_q^i + \tilde{a}_q^i) \quad \forall q \in Q, \forall i \in M \quad (13)$$

$$\sum_{q|\text{head of } q=j} f_{q(ire)} \leq C_{ch} W l_{j(ire)}^u \quad \forall j \in M, \forall (ire) \quad (14)$$

$$\sum_{q|\text{tail of } q=j} f_{q(ire)} \leq C_{ch} W l_{(ire)j}^d \quad \forall j \in M, \forall (ire) \quad (15)$$

$$\sum_{j \in M} \sum_{r,e} l_{j(ire)}^u \leq \sum_{r \in V} s_r y_{ir} \quad \forall i \in M \quad (16)$$

$$\sum_{j \in M} \sum_{r,e} l_{(ire)j}^d \leq \sum_{r \in V} s_r y_{ir} \quad \forall i \in M \quad (17)$$

$$\sum_{i \in M} \sum_{r \in V} C_r y_{ir} \leq C_{\min} \quad (18)$$

$$y_{ir} \in Z_+, a_q^i, \tilde{a}_q^i \in \{0, 1\}, l_{ij} \in Z_+, f_{q(ire)} \in Z_+ \quad (19)$$

The objective (9) is composed of the fixed cost of enabled CNs, the cost of single unidirectional fibers going to and coming from core nodes, and the propagation delay cost of working and protection data flow. Constraints (10), (11), (12), (18) and (19) have the same meaning as (2), (3), (4), (7) and (8) of the previous formulation. Constraints (13), together with integrality conditions on the  $a_q^i$  and  $\tilde{a}_q^i$  variables, ensure that the whole traffic of each connection request is switched in the same site. Constraints (14) and (15) model the capacity of installed fibers. (16) and (17) ensure that the number of fibers entering and exiting each site is coherent with the number of available switching planes.

### B. A three-step heuristic

Unlike the regular network model, this formulation is hard to optimize for a general purpose solver. Therefore, we devised a special-purpose heuristic algorithm.

*a) First step:* We compute the optimal regular network by solving model (1)–(8). As discussed, this can be done with general purpose optimization software. We indicate by  $Q^w(i) \subseteq Q$  the set of connection requests having wTLPs assigned to a core node in site  $i$  in an optimal regular network. Similarly, we indicate by  $Q^p(i) \subseteq Q$  the set of connection requests having pTLPs assigned to a core node in site  $i$ . Then, as in the seminal work of Cooper [14] on location problems, we proceed to an iterative two-step local search.

*b) Second step:* Each site hosting core nodes is optimized independently. This requires, for each site  $i$ , fixing all the variables  $a_q^i$  and  $\tilde{a}_q^i$  as follows:

$$a_q^i = \begin{cases} 1 & \text{if } q \in Q^w(i) \\ 0 & \text{otherwise} \end{cases} \quad \tilde{a}_q^i = \begin{cases} 1 & \text{if } q \in Q^p(i) \\ 0 & \text{otherwise} \end{cases}$$

and to solve a set of Single-Site restrictions of the previous Optimization Problem (SSOP):

$$\begin{aligned} \text{SSOP}(i) : \min s(y, l) = & \sum_{r \in V} h_r y_{ir} + \\ & + \sum_{j \in M} \sum_{r,e} (\phi(W) \Delta_{ij} F + W \gamma^{s_r-1} P) l_{j(ire)}^u \\ & + \sum_{j \in M} \sum_{r,e} (\phi(W) \Delta_{ij} F + W \gamma^{s_r-1} P) l_{(ire)j}^d \end{aligned}$$

subject to (13)–(19).

*c) Third step:* Each switch installed in the second step is shifted to the location that minimizes the fiber and delay costs. The best location  $i^*$  of each core node during the second step is found by solving a set of (Weighted) Median Problems (WMP), one for each core node  $e$  of type  $r$  at site  $i$ :

$$\text{WMP}(ire) : i^* = \operatorname{argmin}_{k \in J} \sum_{j \in J} (l_{j(ire)}^u \cdot d_{jk} + l_{(ire)j}^d \cdot d_{kj})$$

each WMP can easily be solved by inspection. Whenever a core node  $(ire)$  is moved during this step ( $i^* \neq i$ ), the wTLPs of certain connection requests may be removed from  $i$

$$Q^w(i) := Q^w(i) \setminus \{q \in Q^w(i) | f_{q(ire)} > 0\}$$

and assigned to  $i^*$

$$Q^w(i^*) := Q^w(i^*) \cup \{q \in Q^w(i) | f_{q(ire)} > 0\}.$$

An analogous update is performed for the pTLPs; however, in order to prevent from using the same switching site for both the wTLPs and pTLPs of certain connection requests, when  $q \in Q^p(i)$  and  $q \in Q^w(i^*)$  the shift of the pTLPs is forbidden. In a similar way, if  $q \in Q^w(i)$  and  $q \in Q^p(i)$  the real and protection paths are swapped.

$$\begin{aligned} Q^w(i) &:= Q^w(i) \setminus \{q\}; Q^p(i) := Q^p(i) \cup \{q\} \\ Q^p(i^*) &:= Q^p(i^*) \setminus \{q\}; Q^w(i^*) := Q^w(i^*) \cup \{q\} \end{aligned}$$

Therefore, even connection requests with a small fraction of traffic switched by  $(ire)$  are shifted to  $i^*$ . Even if in a single iteration the solution cost may increase, this strategy helps the algorithm to escape local minima. Due to constraint (18), this shifting may even yield infeasible subproblems.

The second and third steps are iterated, until no more changes to the solution are made, or the algorithm encounters a previously visited solution.

## VI. RESULTS

In this section we show the results obtained by implementing the resolution algorithms for the quasi-regular Petaweb optimization. The algorithms have been implemented in C++, and CPLEX 10.0 was used to optimize the ILP formulations. The simulations run on a workstation equipped with a AMD Opteron 64bit 2.4Ghz CPU, 1MB cache, 16GB RAM.

We considered two starting networks, having 10 and 34 nodes respectively. Moreover, for each network we considered two types of traffic matrix: type **A** matrices represent real traffic data, given by Nortel Networks; since these matrices are sparse, we considered also dense type **B** matrices, whose entries are drawn from a gravity model [7]. In this way we obtained four test instances, named 10A, 10B, 34A and 34B.

We considered  $s_1 = 1$  and, thus,  $s_2 = 2$  and  $s_3 = 4$ . We set  $C_{ch} = 10$  Gbps and  $W = 16$ , and thus we have  $Z_2 = 10$  Gb/s and  $Z_3 = 160$  Gb/s;  $Z_1$  had been chosen equal to 0.625 Gb/s to optimize the capacity efficiency. Using these bit-rate classes we can efficiently serve tributaries from SDH and OTN interfaces [12]. Furthermore, we set  $\gamma = 0.95$ ,  $P/F = 150$ ,  $\beta/F = 0.1$  [ $Km$  Gbit/s] $^{-1}$ ,  $h_1/F = 20$ ,  $h_2/F = 50$ ,  $h_3/F = 100$  (the network cost is normalized to  $F$ ), for 10A and 10B  $C_j = 1000$  Gbit/s, for 34A  $C_j = 2000$  Gbit/s, for

TABLE I  
RESULTS COMPARISON FOR THE REGULAR PETAWEB DESIGN

Instance	Old model exec. time (s)	New model exec. time (s)
10A	2768	37
10B	2330	5
34A	214524	50834
34B	223560	3961

34B  $C_j = 2800$  Gbit/s, and  $\phi(W) = W$  (approximating the cost of a fiber as proportional to the number of wavelengths).

*Regular topology optimization.* In Table I we report the CPU time needed for optimizing the regular topology with the formulation proposed in [8] (column 'Old model') and with the model presented in this paper (column 'New model'), for each instance (whose id is indicated in the first entry of each row). As expected and mentioned in the previous section, the computational effort required with the new procedure drastically decreases thanks to the TLPs post-selection.

*Quasi-regular topology optimization.* In Table II we compare the results obtained by removing unused ports and fibers from the optimal regular Petaweb obtained in the previous experiment (removal), by using CPLEX to optimize formulation (9)–(19), and by running the three-step heuristic.

In order to speed-up the computation, in the heuristic we limited CPLEX to explore at most 100 nodes of the branch-and-bound tree and to perform at most 100 cutting planes iterations at each node. Furthermore, we emphasized the generation of useful cuts by fine-tuning CPLEX internal parameters (their complete setting is available on request).

We report our results in Table II, which consists of three horizontal blocks, one for each solution method. For each method we report the solution value, the computing time required to obtain the solution for each instance (whose id is indicated in the first column). To complete the analysis, we report also the lower bounds on the optimal solution values obtained by CPLEX, and for the 3-step heuristic we indicate the total number of iterations performed (Tot. iter.) and the iteration in which the best solution was found (Best iter.).

While on the smaller networks CPLEX was able to provide good solutions within two days, it was unable to optimize the larger instances: we aborted the execution after two and four days of computation for respectively instance 34A and instance 34B, and no feasible integer solution was found. The heuristic showed good computational results: on the smaller instances the quality of its solutions is the same as those of CPLEX, but the computing time is significantly smaller; moreover the heuristic substantially improved the best known solutions on larger instances with still affordable computing resources.

Both CPLEX and the heuristic methods showed that directly optimizing the quasi-regular topology actually yields substantial savings in terms of network cost.

*A-posteriori analysis of the solutions.* Looking at the geographical distribution of core nodes in Fig.5a and Fig.5b for the 10A instance, it can be noticed that the direct optimization of the quasi-regular topology allows a fairer distribution of

core nodes. In fact, they are no longer concentrated in the weighted baricenter of the network. This may also offer an additional protection against large-scale disasters.

The quasi-regular topology obtained for instance 10A by downgrading the optimal regular is plotted in Fig.4b, while the quasi-regular topology optimized through (9)–(19) is plotted in Fig.6. By directly optimizing a quasi-regular topology, a core node is not forced to be connected to every site. Then, it becomes appealing to install core nodes in the regions of the network having isolated peaks of traffic. For example, some core nodes in Fig.6 are installed far from the high traffic axes New York - Washington, New York - Philadelphia and Washington - Philadelphia. By looking at the solutions proposed in the literature, several core nodes were concentrated in these sites, and these axes required 10 or more fibers each (see Fig.4b). In our solutions a big fraction of the traffic is switched to sites closer to some boundary CRs, like Tampa and Albany. In fact, we observed a significant reduction of fiber costs.

Fig.7a shows the fiber, delay and core node contributions to the solution cost for the 10A instance. We compared the optimal regular topology, a quasi-regular topology obtained by removal, and the optimal quasi-regular topology obtained with CPLEX. The removal of unused equipment from an optimal regular network implies a significant gain in fiber cost, 10% in cost ratio and 56% in absolute value [8]. By directly optimizing the quasi-regular we obtain a solution in which the fiber cost decreased by 40% (from 1446067 to 863045), the delay cost increased by 7% (from 482022 to 516670), and the core cost decreased by 38% (from 268132 to 165115); with the number of switching planes unchanged at 11 (see Fig.5), the 38% CN cost reduction is due to a lower number of installed CN ports. The impact of the core node costs in this solution remains almost the same, the delay costs are slightly higher, and the fiber costs dropped by 10%. Therefore, we observed that the direct optimization of a quasi-regular topology, besides lowering the fiber needed, may improve the CN configuration in terms of activated ports, and may refine lightpaths routes at the price of slightly higher propagation delays.

Finally, we stress how the direct design of a quasi-regular topology yields a network which is 65-75% cheaper than the optimal regular one (see Fig. 7b). The largest improvements can be observed on networks with high traffic loads (B scenarios): our methods allowed us to obtain a quasi-regular Petaweb 23-45% less expensive than a quasi-regular structure obtained by simple equipment removal.

## VII. CONCLUSIONS

In this paper we tackled the design optimization of an innovative architecture proposed for next generation transport network. We proposed new models and methods for the direct optimization of both the regular and the quasi-regular Petaweb. The experimental results showed that the regular topology can be designed with a significantly lower computational effort, and that the new proposed methods for the quasi-regular Petaweb optimization can improve the network cost by roughly 70% with respect to the regular structure, and by roughly

TABLE II  
RESULTS COMPARISON FOR QUASI-REGULAR PETAWEB DESIGN

Instance	Value	Time (s)		
<b>Quasi-regular by removal</b>				
10A	2196002	37		
10B	1829473	5		
34A	27086832	50834		
34B	36407731	3961		
<b>Quasi-regular by CPLEX</b>			<b>Lower bound</b>	
10A	1456683	> 2 days	1211505	
10B	1038975	> 2 days	1031920	
34A	NA	> 2 days	10399274	
34B	NA	> 4 days	14962700	
<b>Quasi-regular by heuristic</b>			<b>Tot. iter.</b>	<b>Best iter.</b>
10A	1544830	242	8	8
10B	1034150	176	3	2
34A	21541800	29194	11	11
34B	21656300	240063	31	27

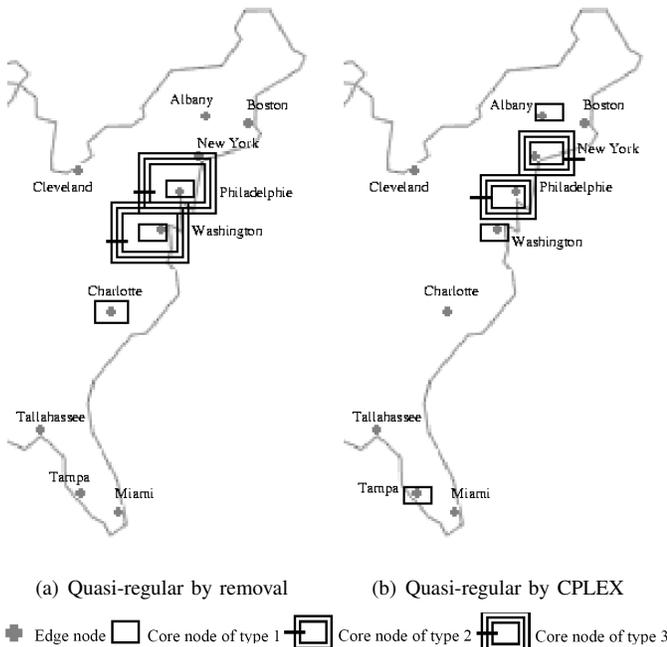


Fig. 5. Geographical distribution of core nodes (10A)

30% with respect to a quasi-regular structure determined as proposed in previous works. The proposed heuristic for the quasi-regular optimization showed good computational results even for large networks.

## REFERENCES

[1] R. Vickers, M. Beshai, *Petaweb Architecture*, 9<sup>th</sup> International Telecommunication Network Planning Symposium, Toronto (2000)  
 [2] F. J. Blouin, S. Yazid, B. Bou-Diab, *Emulation of a Vast Adaptive Network*, 9<sup>th</sup> Int. Tel. Network Planning Symposium, Toronto (2000)  
 [3] F. J. Blouin, A. W. Lee, A. J. M. Lee, M. Beshai, *A comparison of two Optical-Core Networks*, J. of Optical Networking, V.1, N.1:56-65 (2001)  
 [4] M. Beshai, F. J. Blouin, *Courteous routing*, 9<sup>th</sup> Int. Tel. Network Planning Symposium, Toronto (2000)  
 [5] J. P. G. Sterbenz, L. Chapin, R. Krishnan, *Routing issues in interconnecting ip networks with the Petaweb*, 9<sup>th</sup> Int. Tel. Network Planning Symposium, Toronto (2000)

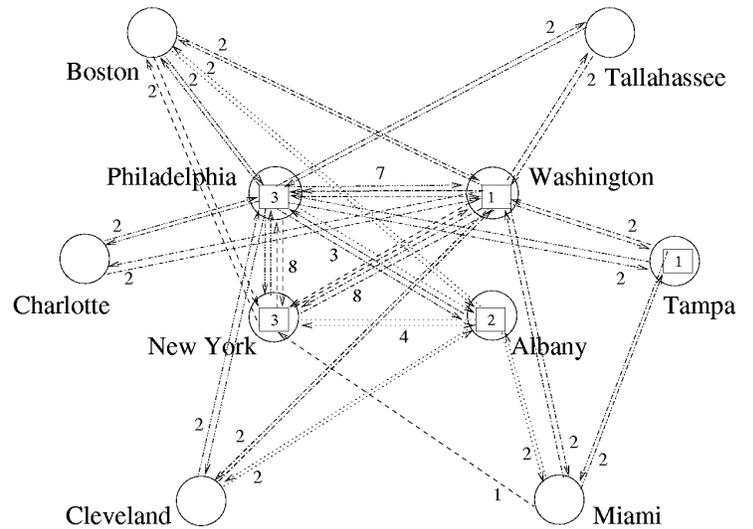


Fig. 6. Optimal quasi-regular Petaweb topology (10A), with path protection

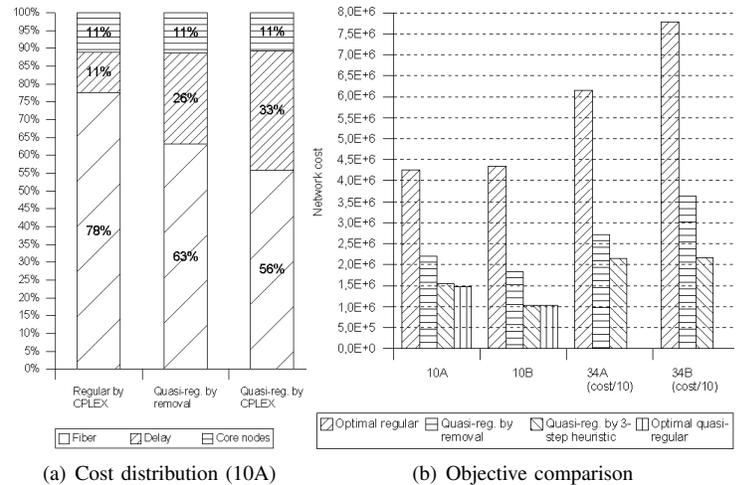


Fig. 7. Cost distribution and objective comparison

[6] A. Reinert, B. Sansò, S. Secci, "Design Optimization of the Petaweb Architecture", Cahier du GERAD 2006-86. Submitted to *IEEE-ACM Transactions on Networking*  
 [7] S. Secci, B. Sansò, "Design and dimensioning of a novel composite-star WDM network with TDM channel partitioning", in Proc. of IEEE BROADNETS'06, Oct. 1-5, 2006, San José (CA).  
 [8] S. Secci, B. Sansò, "Optimization of a dedicated path protected TDM/WDM Petaweb architecture", in Proc. of NAEC 2006  
 [9] N. F. Huang, G. H. Liaw, C. P. Wang, *A Novel All-Optical Transport Network with Time-Shared Wavelength Channels*, IEEE J-SAC, V.18 Is.10:1863-1875 (2000)  
 [10] Y. C. Hueti, P. H. Keng, "Framework for shared time-slot TDM wavelength optical WDM networks", *J. Opt. Netw.* 5, 554-567 (2006).  
 [11] S. Secci, B. Sansò, "Upgrade of a Composite-Star Optical Network", *Lecture Notes on Computer Science, Proc. of NTMS 2007*.  
 [12] *Interfaces for the Optical Transport Networks (OTN)*.G.709/Y.1331, ITU-T International Communication Union, Mar. 2003  
 [13] S. Secci, J.L. Rougier, A. Pattavina, *On the Selection of Optimal Diverse AS-Paths for Inter-Domain Tunnel Provisioning*, submitted to Globecom'07  
 [14] L. Cooper, "The transportation-location problem", *Operations Research*, Vol. 20, pp. 94-108 (1972).  
 [15] N. Menakerman, R. Rom, "Bin Packing with Item Fragmentation", in Proc. of the 7<sup>th</sup> Int. Workshop on Algorithms and Data Structures(2001)