Les Cahiers du GERAD ISSN: 0711–2440

Design Optimization of a New Network Structure for the Next Generation Internet

A. Reinert
B. Sansò

G–2004–25

March 2004
Revised: November 2004

Design Optimization of a New Network Structure for the Next Generation Internet

Anne Reinert

Brunilde Sansò

GERAD, LORLAB and Department of Electrical Engineering
École Polytechnique de Montréal
Montréal (Québec) Canada, H3C 3A7

anne.reinert@polymtl.ca
brunilde.sanso@polymtl.ca

March 2004
Revised: November 2004

Les Cahiers du GERAD

G–2004–25

Abstract

This paper deals with the topological design of a next generation optical network that provides fully meshed connectivity between electronic edge nodes. Such an architecture, nicknamed "PetaWeb", is simple to manage and offers a total capacity of several petabits per second. From the topology standpoint, the PetaWeb presents a very unusual structure as the backbone nodes are totally disconnected. In this paper, we present for the first time a formal mathematical formulation of the network design problem involved in a PetaWeb deployment. We also want to assess the different design parameters of interest in the design of the PetaWeb. Computational results will be presented and discussed.

Key Words: PetaWeb, composite-star network, topological design, dimensioning, capacitated location problem.

Résumé

Ce papier traite du design topologique d'un réseau optique de prochaine génération qui produit des connexions directes entre des noeuds d'accès électroniques. Une telle architecture de réseau, appelée "PetaWeb", est facile à gérer et offre une capacité totale de plusieurs petabits par seconde. D'un point de vue topologique, le PetaWeb présente une structure particulière dans laquelle les noeuds coeur ne sont pas reliés les uns avec les autres. Dans cet article, nous présentons pour la première fois une formulation mathématique du problème de design du réseau PetaWeb. Nous voulons également évaluer les différents paramètres du PetaWeb. Des résultats comparatifs sont discutés.

Mots clés : PetaWeb, réseau en superposition d'étoiles, design topologique, dimensionnement, problème de localisation avec contraintes de capacité.

1 Introduction

One interesting proposal to deal with the great progression of the Internet is to build a wide high-capacity network structure, called the PetaWeb [1, 2, 3], that will have a total capacity of several petabits per second (10^{15}bit/s). The PetaWeb uses WDM fibers (Wavelength Division Multiplexed) and OXCs (Optical Cross-Connectors) and will be completely edge-controlled, with an agile core that can be rapidly reconfigured so as to adapt to traffic variations. Moreover, the PetaWeb can be viewed as a building bloc for a multi-dimensional structure called the YottaWeb, whose external capacity reaches the yottabit per second (10^{24}bit/s) [4, 5].

Besides the obvious capacity increase of the PetaWeb with respect to the current Internet, there is a very particular feature that has not explicitly been exploited before in the area of network design. The PetaWeb proposes not only a high capacitated network, but also a structure that would greatly simplify other networking functionalities, such as routing and addressing. The proposed topology, a composite star structure, is quite unique since it would lead to a network where there is no traditional "backbone" as the core nodes are not connected among themselves. To our knowledge, there have not been any previous attempts at formally modeling the design of such a network structure. Thus, the main objective of this article is to present a mathematical formulation for the design of the proposed PetaWeb. We want to assess the particular modeling features that should be included in the design and make an analysis of the different network design trade-offs. We also want to draw the attention of the network design community on this unique kind of problem.

This paper is divided as follows. The PetaWeb architecture is described in Section 2. Section 3 is dedicated to a literature review. The problem formulation is presented in Section 4, followed by some computational results in Section 5. A second formulation is given in Section 6. Section 7 is devoted to the analysis of the results for the second formulation, and conclusions and suggestions for further research are presented in Section 8.

2 The PetaWeb Architecture

The PetaWeb is composed of several electronic edge nodes and a few optical bufferless core nodes. Fig. 1 represents its proposed architecture. Each edge node is connected to all core nodes and each core node is connected to all edge nodes. In the figure, core nodes 1 and 2 are connected to all edge nodes and are the centers of star structures. Thus, another way of interpreting the PetaWeb is as a composite-star, i.e. a superposition of star structures.

In Fig. 2, we present the connection between the edge nodes and one core node. Each edge node is connected to one core node through one or several optical links. Each optical link is composed of several channels. We consider the regular case where each optical link has the same number of channels. The core node consists of several superposed planes to produce a so-called multiple-plans core node. The superposed planes are also called parallel

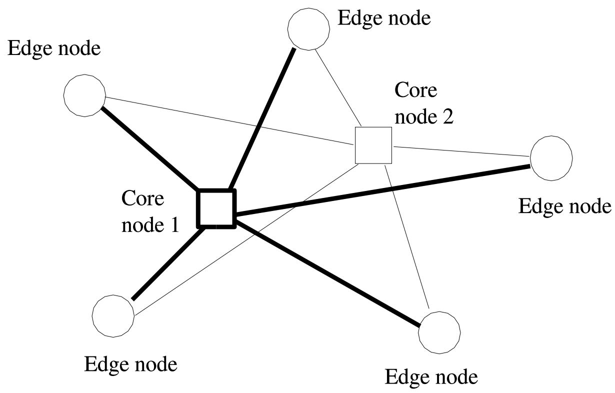


Figure 1: The PetaWeb architecture: a composite-star structure

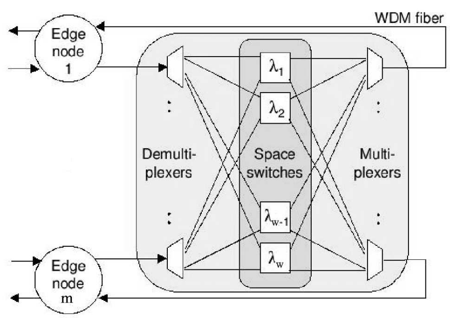


Figure 2: Connection between the edge nodes and one core node - Fig. extracted from [2]

space switches as indicated in Fig. 2. Each space switch corresponds to one wavelength and receives all channels modulated at this wavelength. The fiber entering the core node is demultiplexed in its channels. Each channel is connected to its associated space switch. In the figure, the space switch named λ_W receives the m channels modulated at wavelength λ_W, one channel from each edge node. Each space switch commutes its traffic and sends it to the multiplexers.

In each multiplexer, the new entering W channels are remultiplexed in order to form a link to a destination edge node. There are as many space switches as the number of wavelengths W, each space switch corresponding to one wavelength. For bigger core nodes, the number of space switches can be a multiple of the number of wavelengths. For example, with $W = 16$ channels per link, a core node can have 16, 32, 48 or 64,... space switches.

Let us classify each core node by its type r, which represents the size of the core node. Let s_r be the number of groups of W space switches for the core node of type r. Then, a core node of type r has $s_r * W$ space switches. For example, a core node of type $r = 1$ has

$s_1 * W = 1 * 16 = 16$ space switches (see Fig. 3) whereas a core node of type $r = 2$ has $s_2 * W = 2 * 16 = 32$ space switches (see Fig. 4). The understanding of these features will be important for the forthcoming problem formulation.

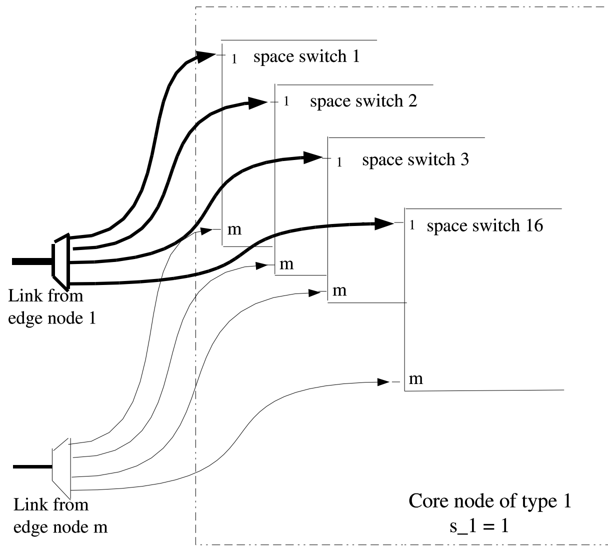


Figure 3: Connection between edge nodes and a core node of type 1

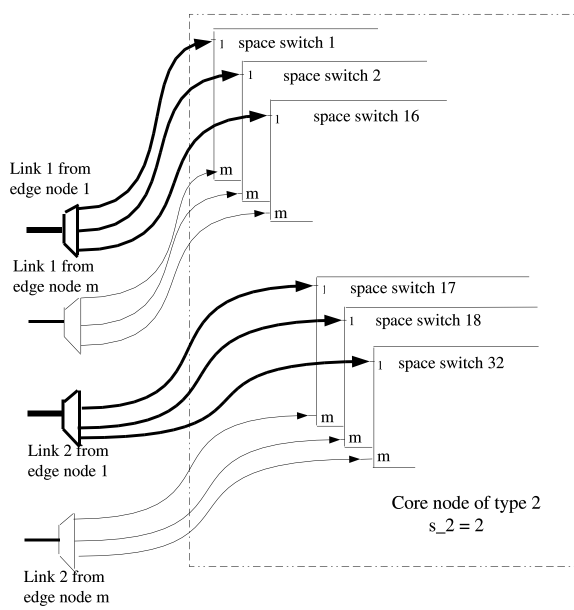


Figure 4: Connection between edge nodes and a core node of type 2

3 Literature Review

Many articles in the Operations Research and Telecommunications literature have dealt with the problem of network design. Klincewicz [6] classified the problems based on the access and the backbone topology and described typical design structures such as mesh/mesh, star/star, tree/star, etc. Most articles have dealt with separate backbone and access design. Some authors, on the other hand, have proposed a unified formulation for the access and backbone design of network that provides important gains in terms of global optimization [7, 8]. Such methodology, however, cannot be applied to the PetaWeb since this structure has no backbone (the core nodes are not connected to each other).

In [2], Blouin et al. compared an optical multi-hop optical network (with backbone) with the PetaWeb. Although the composite-star structure requires a higher fiber length, it needs much fewer ports and no wavelength conversion as the traffic is carried out in one hop. They concluded that the simpler architecture of the PetaWeb results in higher reliability, higher protection against unpredictable traffic behavior and a straightforward routing and traffic engineering.

As previously discussed, the PetaWeb design problem has not been formally classified. In mathematical terms, it can be seen as a particular location problem since it has some similarities with the Plant Location Problem (PLP) where a set of plants or sources send one product to a set of customers (see [9, 10, 11, 12, 13]). In the PetaWeb, core nodes are similar to plants and edge nodes are similar to customers. However, given the capacity constraints at the core nodes, it also present some similarities to the Capacitated Facility Location Problem (CFLP) (see [14, 15]).

4 Mathematical Formulation

We want to determine both the number and the optimal location of the core nodes given a traffic matrix. In other words, we want to know which core nodes should be opened and through which core node each traffic connection should be switched.

We assume that the location of edge nodes, the matrix of traffic between the edge nodes and the potential locations for the core nodes are given. Furthermore, it is also assumed that the potential locations for the core nodes are the sites of the edge nodes.

The restrictions are all related to the maximum capacity supported by the equipment. Thus, we take into account in our modeling framework the maximal capacities for the core nodes, the maximal capacity for the edge nodes and the maximal capacities for the links. The objective is to minimize the total cost of the network.

Let us introduce some useful notation.

$M =$ the edge node set,

$N =$ the set of potential core node locations,

$T =$ the set of edge node pairs, $T \in M * M$,

$V =$ the set of core node types,

$E =$ the number of core nodes of one type that can be opened at one site, $E \subset \mathbf{N}$. In practice, this number can be kept quite small by analyzing the cost structure of the core nodes.

$C_{channel} =$ the channel capacity (in Gbit/s),

$W =$ the number of channels per link,

$s_r =$ the number of groups of W switching planes for the core node of type r, $r \in V$,

$C_j =$ the capacity of edge node j, $j \in M$, (in Gbit/s),

$K_r =$ the total capacity of a core node of type r, $r \in V$, (in Gbit/s),

$$K_r = s_r * W * |M| * C_{channel}, r \in V \tag{1}$$

$f_r =$ the cost of one core node of type r, $r \in V$,

$P =$ the cost of one port in a core node,

$\gamma =$ the scale factor for the cost of the ports in a core node,

$F =$ the fiber cost per length and wavelength unit,

$\beta =$ the cost representing the propagation delay, per length and traffic unit,

$Q_p =$ the traffic of the origin/destination pair p, $p \in T$, (in Gbit/s),

$d_{ij} =$ the distance between the site i, $i \in N$, and the edge node j, $j \in M$,

$d_{ip} =$ the sum of the distance between the origin edge node of the pair p, $p \in T$, and the site i, $i \in N$, and the distance between the site i and the destination edge node of the pair p. For instance, if j and k are, respectively, the origin and the destination of node pair p, then $d_{ip} = d_{ij} + d_{ik}$.

In this model, we use two types of variables: location and traffic variables denoted by y and x respectively.

The objective (2) is to minimize the total network cost. We have three cost terms : the costs of the core nodes which are composed of a fixed cost and the cost of the ports, the costs of the links between edge and core nodes which are proportional to the distance between edge and core nodes and the costs representing the propagation delay. The last is an additional cost that we have added because we do not want that a high amount of traffic be routed through a core node located too far away from the edge nodes origin and destination since the propagation delay would be significant. Consequently we built a cost representing the propagation delay which is proportional to the distance between the edge node pairs and the amount of exchanged traffic.

Then we have the following formulation:

$$
\begin{aligned}
\min\ F(y_{ire}, x_{ire,p}) \quad = & \quad \sum_{i\in N}\sum_{r\in V}\sum_{e\in E}(2\,|M|\,W\,s_r\,\gamma^{(s_r-1)}\,P + f_r)\,y_{ire} \\
+ & \quad \sum_{i\in N}\sum_{r\in V}\sum_{e\in E} 2\,W\,F\,s_r\,(\sum_{j\in M} d_{ij})y_{ire} \\
+ & \quad \sum_{i\in N}\sum_{r\in V}\sum_{e\in E}\sum_{p\in T}\beta d_{ip} Q_p x_{ire,p} \qquad (2)
\end{aligned}
$$

subject to:

$$\sum_{i\in N}\sum_{r\in V}\sum_{e\in E} x_{ire,p} = 1, \forall p \in T \qquad (3)$$

$$x_{ire,p} \le y_{ire}, \forall i \in N, \forall r \in V, \forall e \in E, \forall p \in T \qquad (4)$$

$$\sum_{p\in T} Q_p x_{ire,p} \le K_r y_{ire}, \forall i \in N, \forall r \in V, \forall e \in E \qquad (5)$$

$$C_{channel} * W * \sum_{i\in N}\sum_{r\in V}\sum_{e\in E} s_r\, y_{ire} \le C_j, \forall j \in M \qquad (6)$$

$$
\begin{aligned}
\sum_{p\in T\, origin\, j} Q_p x_{ire,p} \le C_{channel} * W * s_r\, y_{ire}, \qquad (7) \\
\forall j \in M, \forall i \in N, \forall r \in V, \forall e \in E
\end{aligned}
$$

$$
\begin{aligned}
\sum_{p\in T\, destination\, k} Q_p x_{ire,p} \le C_{channel} * W * s_r\, y_{ire}, \qquad (8) \\
\forall k \in M, \forall i \in N, \forall r \in V, \forall e \in E
\end{aligned}
$$

$$y_{ire} = \begin{cases} 1 & \text{if the } e^{th} \text{ core node of type } r \text{ located at } i \text{ is opened,} \\ 0 & \text{else} \end{cases} \qquad (9)$$

$$x_{ire,p} = \begin{cases} 1 & \text{if the traffic } Q_p \text{ is switched by the } e^{th} \text{ core node of type } r \\ & \text{ located at } i, \\ 0 & \text{else} \end{cases} \qquad (10)$$

We now describe the constraints of the problem.

(3) indicates that the total traffic exchanged by a pair of edge nodes must be routed through a core node.

(4) specifies that the traffic can be routed through the e^{th} core node of type r located at site i only if this core node is active.

(5) states that the capacity of each core node must be respected.

(6) indicates that the edge node capacity must be respected.

(8) is a link capacity constraint for all the links between each origin edge node and each core node.

(9) ensures that the link capacity is respected for all the links between each core node and each destination edge node.

(9) indicates that y_{ire} is a binary variable.

(10) indicates that $x_{ire,p}$ is a binary variable.

5 Computational Results

The proposed mathematical model was tested using the CPLEX Mixed Integer Optimizer. The tested networks were respectively composed of 10 and 34 edge nodes. The locations of the edge nodes are specific cities of the United States.

The distance matrix between edge nodes were calculated as follow. To work with realistic distances, geographical coordinates were first found in an American national atlas [16] and a formula to assess the distance between two points on a sphere [17] was used. The calculated distances were later compared and validated with a few air distances estimated at the University of Minnesota [18].

We tried to construct traffic matrices that would represent realistic values. Two matrices were used:

- Traffic matrix A, which is a very sparse matrix provided by industrial data,
- Traffic matrix B, that is calculated using a gravity model based upon urban populations and distances between cities. The urban populations were found in [19]. Note that this matrix does not include any zeros, except on its diagonal entry.

The following default values were used.
$W = 16$ channels per link,
$C_{channel} = 10$ Gbit/s for the channel capacity.
Number of types of core nodes: $v = 3$,
Number of space switches for the core node of type 1: $s_1 = 1 * W$,
Number of space switches for the core node of type 2: $s_2 = 2 * W$,
Number of space switches for the core node of type 3: $s_3 = 4 * W$,
Scale ratio for the cost of the ports: $\gamma = 0.95$,
Ratio cost of a core node port divided by the fiber unitary cost: $P/F = 150$,

Ratio propagation delay unitary cost divided by the the fiber unitary cost: $\beta/F = 0.1$,
Ratio cost of a core node of type 1 divided by the fiber unitary cost: $f_1/F = 20$,
Ratio cost of a core node of type 2 divided by the fiber unitary cost: $f_2/F = 50$,
Ratio cost of a core node of type 3 divided by the fiber unitary cost: $f_3/F = 100$,
Edge node capacity in the 10 edge nodes network, traffic matrices A and B: $C_j = 1000$ Gbit/s,
Edge node capacity in the 34 edge nodes network, traffic matrix A: $C_j = 2000$ Gbit/s,
Edge node capacity in the 34 edge nodes network, traffic matrix B: $C_j = 2800$ Gbit/s.

The results are presented in Table 1 where we portray the total cost of the design for each case and traffic matrix as well as the percentage of the core, fiber and delay costs that compose the solution. The solution time and the optimality gap are given at the end. The actual solutions obtained for all instances treated are presented in Fig. 6 and in Fig. 7. The legend used is portrayed in Fig. 5.

Table 1: Results obtained with default parameters for the first mathematical model

Network	10 edge nodes traffic A	10 edge nodes traffic B	34 edge nodes traffic A	34 edge nodes traffic B
Objective function	2280980	2152920	31837547	42406000
Percentage of the core nodes cost	11.2%	12.1%	5.3%	5.3%
Percentage of the fiber cost	77.8%	83.8%	81.7%	81.6%
Percentage of the delay cost	11%	4.1%	13%	13.1%
Solution time	23650s	232s	579998s	1614383s
Gap	0.01%	0%	7.22%	0%

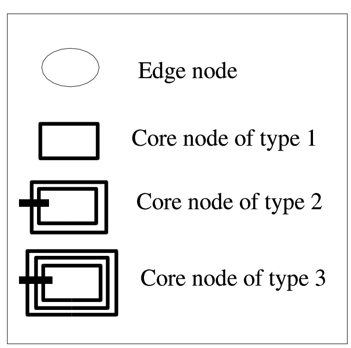
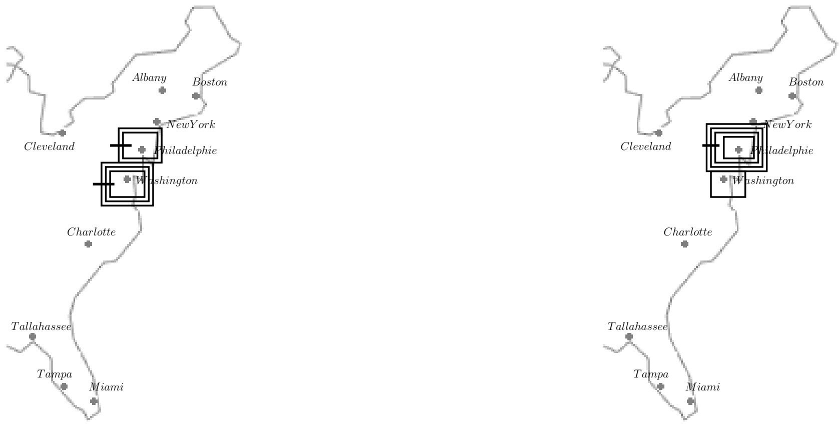
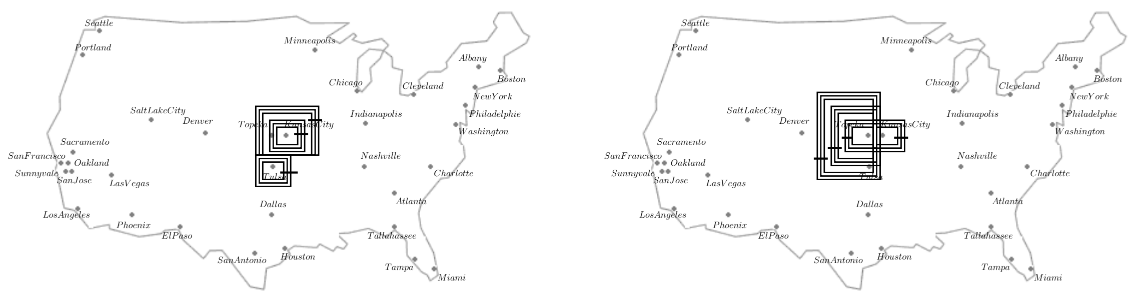


Figure 5: Legend



(a) traffic matrix A (b) traffic matrix B

Figure 6: Optimal 10 edge nodes network with default parameters (model 1)



(a) traffic matrix A (b) traffic matrix B

Figure 7: Optimal 34 edge nodes network with default parameters (model 1)

We can see that, in all instances, the fiber costs predominate over all the other network costs, representing about 80 % of the total cost.

In terms of computational complexity, we expected to see a high increase of the resolution time for the 34 edge nodes network, which is confirmed by the results. For the 34 edge nodes network with a full traffic matrix, the calculation lasted 19 days.

6 An Alternative Mathematical Formulation

The previous model allows us to know exactly through which core node each traffic connection should be switched. This leads to a very hard combinatorial problem that cannot be solved by a general purpose solver such as CPLEX for large networks. Therefore, we now propose a more compact formulation that would let us know how many core nodes of each type should be opened at each site while verifying the maximum equipment capacity per site.

Moreover, in this model, we allow the traffic exchanged by one edge node pair to be split in the edge node origin and pass in parallel through different core nodes before reaching the destination node. This is a feature that can be technically added to the proposed architecture. The reader is referred to Fig. 8 where the traffic from edge node A to edge node C is split and switched into two different core nodes.

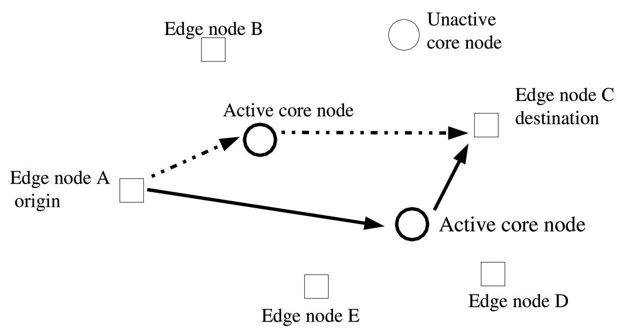


Figure 8: Traffic from one origin/destination switched into several core nodes

We use the same notation as before and we keep the y and x variables for location and traffic, respectively.

Let y_{ir}, $i \in N$, $r \in V$, represent the number of core nodes of type r active at location i and x_{ip}, $i \in N$, $p \in T$, be the fraction of the traffic Q_p that is routed through core nodes located at site i.

Then we have the following formulation:

$$\begin{aligned} \min\ F(y_{ir}, x_{ip}) = & \sum_{r \in V} \sum_{i \in N} (2\,|M|\,W\,s_r\,\gamma^{(s_r-1)}\,P + f_r)\,y_{ir} \\ & + \sum_{r \in V} \sum_{i \in N} \sum_{j \in M} 2\,W\,F\,s_r\,d_{ij} y_{ir} \\ & + \sum_{i \in N} \sum_{p \in T} \beta d_{ip} Q_p x_{ip} \end{aligned} \tag{11}$$

subject to:

$$\sum_{i \in N} x_{ip} = 1, \forall p \in T \tag{12}$$

$$0 \leq x_{ip} \leq \sum_{r \in V} y_{ir}, \forall i \in N, \forall p \in T \tag{13}$$

$$\sum_{p \in T} Q_p x_{ip} \leq \sum_{r \in V} K_r y_{ir}, \forall i \in N \tag{14}$$

$$\sum_{r \in V} \sum_{i \in N} s_r \, y_{ir} * C_{channel} * W \leq C_j, \forall j \in M \tag{15}$$

$$\sum_{p \in T \, origin \, j} Q_p x_{ip} \leq \sum_{r \in V} s_r \, y_{ir} * C_{channel} * W, \forall j \in M, \forall i \in N \tag{16}$$

$$\sum_{p \in T \, destination \, k} Q_p x_{ip} \leq \sum_{r \in V} s_r \, y_{ir} * C_{channel} * W, \forall k \in M, \forall i \in N \tag{17}$$

$$y_{ir} \in Z+, \forall i \in N, \forall r \in V \tag{18}$$

$$0 \leq x_{ip} \leq 1, \forall i \in N, \forall p \in T \tag{19}$$

The objective (11) is, again, to minimize the total network cost, i.e. the costs of the core nodes, the costs of the links between edge and core nodes, and the costs representing the propagation delay, which are proportional to the distance between the edge node pairs and the amount of exchanged traffic.

(12) indicates that the total traffic exchanged by a pair of edge nodes must be routed through a set of core nodes.

(13) specifies that the traffic can be routed through a core node located at site i only if one core node is active at site i.

(14) states that the global core node capacity at site i must be respected.

(15) indicates that the edge node capacity must be respected.

(16) is a link capacity constraint for all the links between each origin edge node and the core nodes at site i.

(17) ensures that the link capacity is respected for all the links between the core nodes at site i and each destination edge node.

(18) indicates that y_{ir} is a positive integer variable.

(19) indicates that x_{ip} is a fractional variable.

7 Computational Results

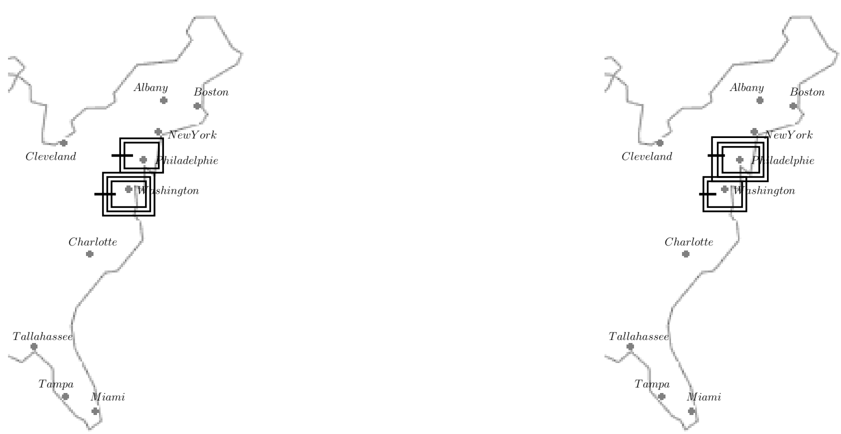
7.1 Results with Default Parameters

The results presented in Table 2 were found using the default parameters. The actual solutions obtained for all instances treated are presented in Fig. 9 and 10.

As expected, this model produces computational results with a slightly lower objective function cost and a decreased resolution time than the first formulation. In fact, this formulation can be seen as a relaxation of the first one, as the capacity constraints have been widened and the flow can be bifurcated.

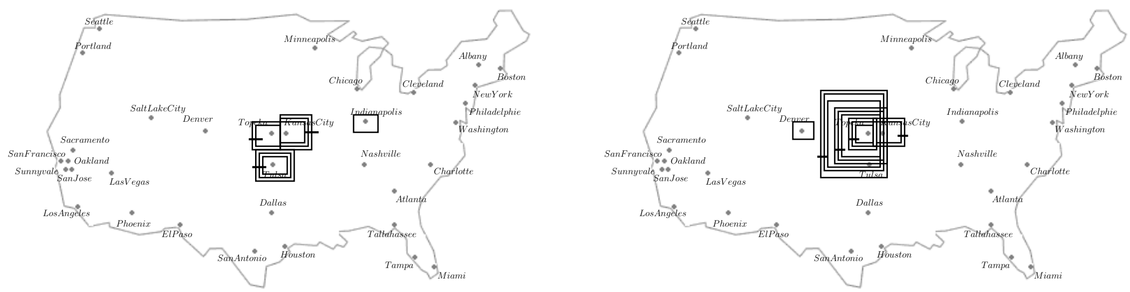
Table 2: Results obtained with default parameters for the second mathematical model

Network	10 edge nodes traffic A	10 edge nodes traffic B	34 edge nodes traffic A	34 edge nodes traffic B
Objective function	2279215	2150101	29544869	40146729
Percentage of the core nodes cost	11.2%	11.9%	5.4%	5.4%
Percentage of the fiber cost	77.9%	83.5%	81.5%	81.3%
Percentage of the delay cost	10.9%	4.6%	13.1%	13.3%
Solution time	36s	7.3s	517.7s	252885s
Gap	0	0	0	3.0%



(a) traffic matrix A (b) traffic matrix B

Figure 9: Optimal 10 edge nodes network with default parameters (model 2)



(a) traffic matrix A (b) traffic matrix B

Figure 10: Optimal 34 edge nodes network with default parameters (model 2)

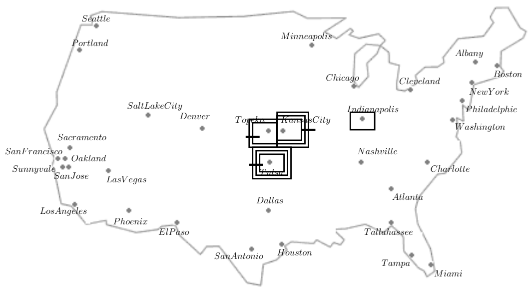
As with the previous formulation, the leading costs in the network are the fiber costs. Moreover, we can observe that the cost distribution is almost identical to the distribution found for the solved instances in the first formulation. This suggests that the flow and capacity relaxation does not have a great impact in how the total amount of money is actually used. Now, if we compare the network solutions for the available instances, we see that Fig. 9(a) and 6(a) present exactly the same solution. On the other hand, Fig. 9(b) presents a slightly different and more equilibrated solution than Fig. 6(b) in which some of the large Philadelphia capacity is deviated to Washington. The same trend can be seen when we compare Fig. 10(a) with Fig. 7(a), and Fig. 10(b) with Fig. 7(b). We can observe that for the second model the total capacity is more spread among several cities. This is probably the effect of allowing the traffic flow to be bifurcated and suggests that the small increase in network cost very well justify a more spread solution.

7.2 Sensitivity Analysis

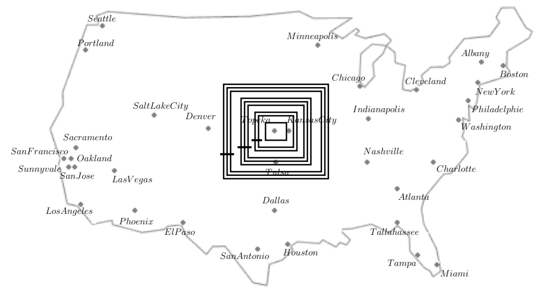
We now want to understand the importance of the different costs in the objective function. For this, we carried out a series of tests with the second formulation for which some of the costs were assumed to be zero. We refer the reader to Fig. 11 where the results for the 34 edge nodes network with traffic matrix A are presented.

The reader can appreciate that when all costs are taken into account, several core node locations are opened (see Fig. 11(a)). On the other hand, when delay costs are not considered, a few core nodes are opened and are all concentrated at the barycenter (see Fig. 11(b)).

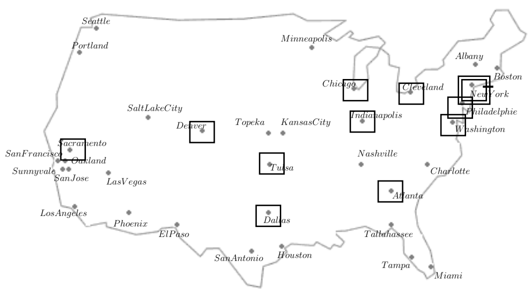
In the solution with optimal delay without considering the core nodes or the fiber costs, several core nodes are opened and they are all widely dispersed (see Fig. 11(c)). Now in the solution seeking an optimal delay and core node costs but without considering the fiber cost, some core nodes are opened and they are widely dispersed (see Fig. 11(d)). These results are coherent. Indeed, delay costs encourage more core nodes to be widely opened



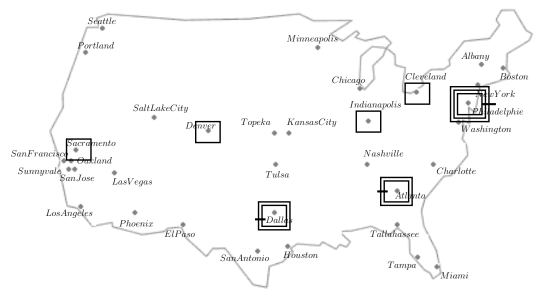
(a) Optimal network with default parameters (model 2)



(b) Without considering the delay (model 2)



(c) Without considering the core nodes and fiber costs (model 2)



(d) Without considering the fiber cost (model 2)

Figure 11: Optimal 34 edge nodes network with traffic matrix A (model 2)

whereas core node and fiber costs drive not many core nodes to be active, these core nodes locating at the barycenter.

7.3 Results on "Suitable" Networks

We now take a different approach in the problem analysis. Let us define a "suitable" network as a network that respects some specific industrial criteria. For instance, a suitable network could be defined by the two following criteria:

- Criterion 1: its total cost should not exceed 20 % of the optimal cost obtained without considering the delay;
- Criterion 2: its delay cost should not exceed its optimal delay cost by more than 55%.

  For each network, we have tested if the two criteria were fulfilled.

Results are given in Table 3.

To examine criterion 1, we have calculated the cost of the optimal network built without considering the delay. The percentage of the total cost exceeding the total cost of the network with default parameters is tabulated in the "cost exceeding" entry of Table 3. To examine criterion 2, we have calculated the cost of the optimal network built without considering the core node and the fiber costs. This is the cost of the network with optimal delay. The percentage of the delay cost exceeding the delay cost of the optimal network with default parameters is presented in the "delay exceeding" of Table 3. We use the letters V and O, respectively to indicate if the criterion was fulfilled or not.

Table 3: Evaluation of the solution networks

Network (nodes)	Traffic matrix	Cost exceeding	Crit. 1	Delay exceeding	Crit. 2
10	A	13%	V	61%	O
10	B	6.9%	V	27%	V
34	A	16%	V	53%	V
34	B	16%	V	61%	O

We can see that the 10 edge nodes network with traffic matrix B and the 34 edge nodes network with traffic matrix A fulfill the two criteria. For the two other networks, the criterion 2 is not fulfilled which indicates that the delay in the network with the default parameters is too important in comparison with optimal delay. This result leads us to adjust the weight of the delay cost in the objective function. The ratio of the propagation delay unitary cost divided by the fiber unitary cost should then be increased for these instances so as to obtain suitable networks.

8 Conclusions and Further Work

In this paper, we formally formulated for the first time the network design problem for the proposed PetaWeb architecture. This is particularly interesting given that the topological structure presented by the PetaWeb has not been previously studied in optimization network design.

Two models were presented to assess the structure: a very detailed one that gave rise to an extremely hard combinatorial problem and a modified formulation that captured most of the design features while keeping the problem tractable. We included in the design three types of costs: core, fiber and delay-related costs. Interestingly enough, the cost distributions in both formulations were quite similar, however, the second formulation "opened" more sites and produced less extremist results in terms of capacity concentration. This is particularly interesting if the network planner is concerned with survivability issues. We also found in both cases that up to 80% of the costs were due to fiber costs. This can suggests some avenues of mathematical methods to solve large instances of the problem.

We have also studied the effect of the different costs in the type of design solution obtained. We found that the delay costs have a definite importance since their inclusion leads to a more spread assignment of cores whereas their total omission would lead to a barycenter solution.

Our final analysis dealt with the notion of "suitable networks", that is, networks that were not strictly optimal but answered to industrial criteria of suitability. This notion should be further explored and could lead to new types of formulations for the PetaWeb design.

In terms of resolution methods, we are currently working on a specialized heuristic to treat large-scale instances of the problem.

Acknowledgments

The authors would like to express their gratitude to François Blouin and Maged Beshai of Nortel Networks for suggesting the problem and for stimulating discussions on this topic. This work was supported by cooperative Grant CRDPJ248035-01 between NSERC and Nortel Networks.

References

[1] M. Beshai, R. Vickers, "PetaWeb Architecture," in *Networks 2000, Towards Natural Networks* 9^{th} *International Telecommunication Network Planning Symposium*, Toronto (2000).

[2] F. J. Blouin, A. W. Lee, A. J. Lee, and M. Beshai, "Comparison of two optical-core networks," *J. Opt. Netw.*, 1, 56–65 (2001), `http://www.osa-jon.org/abstract.cfm?URI=JON-1-1-56`.

[3] M. Beshai, F. J. Blouin, "The PetaWeb: a Building Block for a Yottabit-per-second Network," personal communication.

[4] J. Dégila, B. Sansò, "Design Optimization of a Next Generation Yottabit-per-second Network," in *Proc. IEEE International Conference on Communications 2004*, Vol. 2, 20-24, pp. 1227–1231 (2004).

[5] J. Dégila, B. Sansò, "Topological Design Optimization of a Yottabit-per-second Lattice Network," to appear in *Optical Communications and Networking* from *IEEE Journal on Selected Areas in Communications (IEEE-JSAC)*, Nov. 2004.

[6] J. G. Klincewicz, "Hub Location in backbone/tributary Network Design: a Review," *Location Science*, 6, 307–335 (1998).

[7] S. Chamberland, O. Marcotte, and B. Sansò, "On the Joint Topological, Dimensioning and Location Problem for Broadband Networks," in *Proceedings of the International IFIP-IEEE Conference on Broadband Communications 1996*, pp. 525–536 (1996).

[8] S. Chamberland, B. Sansò, and O. Marcotte, "Topological Design of Two-Level Telecommunication Networks with Modular Switches," *Operations Research*, 48-4, 745–760 (2000).

[9] M. L. Balinski, "On Finding Integer Solutions to Linear Programs," *Mathematica*, (1964).

[10] M. A. Efroymson, T. L. Ray, "A Branch-Bound Algorithm for Plant Location," *Operations Research*, 14, 361–368 (1966).

[11] O. Bilde, J. Krarup, "Sharp Lower Bounds and Efficient Algorithms for the Simple Plant Location Problem," *Annals of Discrete Mathematics*, 1, 79–97 (1977).

[12] T. B. Boffey, J. Karkazis, "The Multi-commodity Facilities Location Problem," *Journal of the Operational Research Society*, 32, 803–814 (1981).

[13] K. G. Murty, "Solving the Fixed Charge Problem by Ranking the Extreme Points," *Operations Research*, 16, 268–279 (1967).

[14] H. Pirkul, "Efficient Algorithms for the Capacitated Concentrator Location Problem," *Computers and Operations Research*, 14(3), 197–208 (1987).

[15] M. Rönnqvist, S. Tragantalerngsak, and J. Holt, "A Repeated Matching Heuristic for the Single-source Capacitated Facility Location Problem," *European Journal of Operational Research*, 116, 51–68 (1999).

[16] United States Department of the interior geological Survey, Washington, DC, "The National Atlas of the United States of America" (1970).

[17] Department of Geography, Faculty of Environmental Studies, University of Waterloo, "Geographic Co-ordinates," `http://www.fes.uwaterloo.ca/crs/geog165/gcoords.htm`.

[18] Cartographic Laboratory Department of Geography, University of Minnesota, "Air Distances," Airmaps Inc. (1989).

[19] US Department of Commerce, Economics, and Statistics Administration, "United States Census 2000 Population and Housing," `http://www.census.gov/main/www/cen2000.html`.