Les Cahiers du GERAD ISSN: 0711–2440

A Heuristic for the Design Optimization of a Next Generation Internet Network

A. Reinert
B. Sansò

G–2004–87

December 2004

A Heuristic for the Design Optimization of a Next Generation Internet Network

Anne Reinert

Brunilde Sansò

GERAD and Département de mathématiques et de génie industriel
École Polytechnique de Montréal
C.P. 6079, Succ. Centre-ville
Montréal (Québec) Canada H3C 3A7
{anne.reinert;brunilde.sanso}@polymtl.ca

December 2004

Les Cahiers du GERAD

G–2004–87

Abstract

This paper proposes an efficient heuristic to solve 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. The network design problem leads to a very hard combinatorial problem that is very difficult to solve for large-sized instances. The heuristic we have developed is based on repeated matchings. Computational results will be presented and discussed.

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

Résumé

Cet article propose une heuristique efficace pour résoudre le design topologique d'un réseau optique de prochaine génération qui produit des chemins directs entre des noeuds d'accès électroniques. Une telle architecture, 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 originale dans laquelle les noeuds centraux sont complètement disconnectés. Ce problème de design de réseau conduit à un problème combinatoire très lourd qu'il est très difficile de résoudre pour de grands réseaux. L'heuristique que nous avons développée est basée sur des couplages répétés. Des résultats comparatifs sont discutés.

Acknowledgments: This work was supported by cooperative Grant CRDPJ 248035-01 between NSERC and Nortel Networks and by Strategic Grant STPGP 246159.

1 Introduction

The PetaWeb is a new network structure that offers a total capacity of several petabits per second (10^{15} bit/s) and that has been proposed for a next generation Internet [1, 2, 3]. The structure provides fully meshed connectivity with direct optical paths between some electronic edge nodes. It is composed of several OXCs (Optical Cross-Connectors), also named core nodes, that commute the traffic exchanged by the edge nodes. One particular feature is that each optical core node is connected to all edge nodes. Thus, another way of interpreting the PetaWeb is as a superposition of star structures as shown in Figure 1. Such a structure greatly simplifies some network functionalities such as routing, addressing and scheduling. It also leads to higher reliability.

The PetaWeb is based on WDM technology (Wavelength Division Multiplexing). The fiber is composed of a fixed number of channels, each channel corresponding to one wavelength. When the fiber enters a core node, it is demultiplexed in its channels and each channel is connected to its associated switching plane. In Figure 2, W switching planes corresponding to the wavelengths λ_1, λ_2, ..., λ_W are shown. The channels that are sent to the same destination edge node are re-multiplexed together. Note that, to ease the figure interpretation, only the channels from and to edge node 1 are pictured.

The relevance of the PetaWeb network was previously studied by Blouin and al. in [3]. They compared the PetaWeb with an optical multi-hop network. Although the PetaWeb requires a higher fiber length, it needs much fewer ports and no wavelength conversion as the traffic is carried out in one hop.

The most advanced version of the PetaWeb network uses TDM technology (Time Division Multiplexing). Connexions between edge nodes pairs are temporally multiplexed and sent over optical fiber. This technology is very common for electronic switches but not for optical switches. since it needs very fast switches. Using TDM in the PetaWeb is

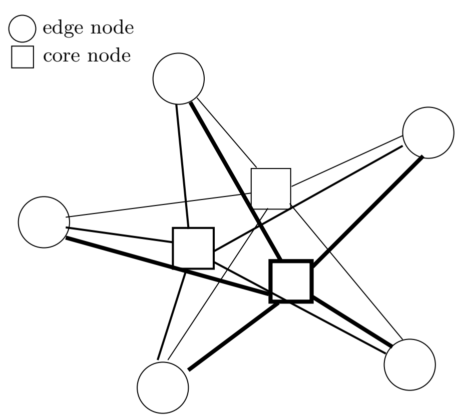


Figure 1: The PetaWeb architecture : a composite-star structure

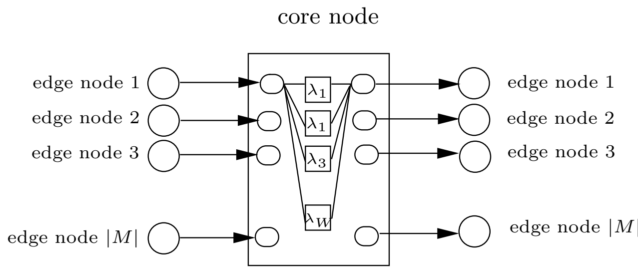


Figure 2: Connection between the edge nodes and a core node

very interesting since the traffic of one pair of edge nodes can fill only a small part of the channel bandwidth, thus greatly improving the network granularity.

Now, in order to construct a PetaWeb, it is necessary to efficiently tackle the network design problem. This is particularly important given that the PetaWeb may be one of the largest networks ever designed and that it has been proposed as a building block for the YottaWeb, a mega-network with aggregated capacities in the order of yottabits per second (10^{24} bits/s) [4, 5].

The PetaWeb design problem, however, is very difficult to classify given that it consists of a structure with no backbone [6]. Indeed, the composite star topology greatly differs from the classical concept of a network made up of access nodes that connect to backbone switches interconnected between themselves.

In mathematical terms, the PetaWeb design can be seen as a particular location problem since it has some similarities with the Capacitated Facility Location Problem where a set of capacitated plants send a product to a set of customers [7]. In [8], we use this approach to propose a mathematical formulation of the PetaWeb design. However the design leads to a very hard combinatorial problem that is very difficult to solve for large-sized instances.

The objective of this article is to present a completely new resolution approach to be able to tackle large instances of the PetaWeb design problem. For this, a method based on an adaptation of the repeated matching heuristic for the capacitated facility location problem by Rönnqvist [9, 10] will be presented.

This article is divided as follows. In the next section we present the mathematical formulation of the PetaWeb design. The heuristic is developed in Section 3. Computational results are presented in Section 4, followed by conclusions and recommendations for further work.

2 Mathematical formulation

For the sake of completeness, we now give a mathematical formulation of the PetaWeb design problem. Details can be found in [8]. The problem consists in determining 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 : 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, $T \in M \times M$,

$V =$ 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 cores.

$C_{channel}$ = 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),

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

$f_r =$ 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)

$\delta_{ij} =$ 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 on node pair p, then $d_{ip} = \delta_{ij} + \delta_{ik}$.

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 which are proportional to the distance between the edge node pairs and the amount of exchanged traffic.

Then we have the following formulation :

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

subject to :

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

$$x_{ire,p} \leq y_{ire}, \forall i \in N, \forall r \in V, \forall e \in E, \forall p \in T \tag{4}$$

$$\sum_{p\in T} Q_p x_{ire,p} \leq K_r y_{ire}, \forall i \in N, \forall r \in V, \forall e \in E \tag{5}$$

$$C_{channel} \times W \times \sum_{i\in N}\sum_{r\in V}\sum_{e\in E} s_r\,y_{ire} \leq C_j, \forall j \in M \tag{6}$$

$$\sum_{p\in T\,origin\,j} Q_p x_{ire,p} \leq C_{channel} \times W \times s_r\,y_{ire}, \quad \forall j \in M, \forall i \in N, \forall r \in V, \forall e \in E \tag{7}$$

$$\sum_{p\in T\,destination\,k} Q_p x_{ire,p} \leq C_{channel} \times W \times s_r\,y_{ire}, \quad \forall k \in M, \forall i \in N, \forall r \in V, \forall e \in E \tag{8}$$

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

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

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.

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

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

(9) and (10) respectively indicate that y_{ire} and $x_{ire,p}$ are binary variables.

3 The heuristic

The mathematical model 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 heuristic for solving the PetaWeb design.

3.1 Reformulation of the PetaWeb design problem

Our heuristic is based on a series of matching problems. Thus, we first reformulate the problem of the PetaWeb design to adapt it to a matching problem.

We use the same notation as in the mathematical formulation in Section 2. An index number associated with a set indicates which exemplar of the set is meant.

An edge node pair is designated by the letter p, $p \in T$.

A subset k of edge node pairs is designated by D_k so that $D_k \in T$. For example, with three edge nodes, we could have : $T = \{(1,2),(1,3),(2,1),(2,3),(3,1),(3,2)\}$, $D_1 = \{(1,2),(1,3),(2,3)\}$ and $D_2 = \{(1,3)\}$.

A core node is designated by the triplet (i, r, e), $i \in N, r \in V, e \in E$. i indicates the site of the core node, r indicates the type of the core node and e indicates which exemplar of the core node of type r at site i we are dealing with.

We now define a "kit". A kit is composed of a core node (i, r, e), $i \in N, r \in V, e \in E$, and a subset D_k of edge node pairs. A kit has a special signification. It means that the edge node pairs of D_k are assigned to the core node (i, r, e), i.e. each edge node pair of D_k commutes its traffic through the core node (i, r, e). The core node (i, r, e) and its assigned edge node pairs D_k will be noted : $((i, r, e), D_k)$.

A kit $((i, r, e), D_k)$ is said to be feasible if : the capacity constraint of the core node (i, r, e), the capacity constraints of the links between each origin edge node of D_k and the core node (i, r, e), and the capacity constraints of the links between the core node (i, r, e) and each destination edge node of D_k are satisfied.

We define a "packing" as a union of feasible kits. Let $((i_1, r_1, e_1), D_1)$ and $((i_2, r_2, e_2), D_2)$ be two feasible kits. $((i_1, r_1, e_1), D_1)$ is composed of the core node (i_1, r_1, e_1) and the edge node pairs of D_1. $((i_2, r_2, e_2), D_2)$ is composed of the core node (i_2, r_2, e_2) and the edge node pairs of D_2.

These two kits form a packing Π if the following is true :

$$((i_1, r_1, e_1), D_1), ((i_2, r_2, e_2), D_2) \in \Pi \Rightarrow (i_1, r_1, e_1) \neq (i_2, r_2, e_2) \text{ and } D_1 \cap D_2 = \emptyset.$$

Given a packing Π, we define :

L_1 : the set of core nodes that are not active, i. e. that do not commute traffic,

L_2 : the set of edge node pairs that are not assigned to a core node,

L_3 : the set of active core nodes with their associated edge node pairs, i. e. the set of feasible kits.

These three sets can be mathematically described by :

$L_1 = \{(i, r, e) \mid ((i, r, e), D) \notin \Pi\}$

$L_2 = \bigcup_{((i,r,e),D) \notin \Pi} D$

$L_3 = \Pi$.

Figure 3 illustrates the three sets L_1, L_2 and L_3. An edge node pair is represented by a circle. A core node is represented by a square. A line binds an edge node pair and a core node if the pair is assigned to this core node, i. e. if the traffic of this edge node pair is commuted by this core node.

Let us assume that L_1 has n_1 elements, L_2 has n_2 elements, and L_3 has n_3 elements. For example, in Figure 3, $n_1 = 2$, $n_2 = 3$ and $n_3 = 2$.

Figure 3 shows a packing Π whose cost can be determined as the sum of :

1. the cost of the active core nodes of L_3 and the cost of the fiber between each active core node of L_3 and each edge node in the network :
   $\sum_{\{(i,r,e) \mid ((i,r,e),D) \in L_3\}} ((2\, M\, W\, s_r\, \gamma^{(s_r-1)}\, P + f_r)$
   $+2\, W\, F\, s_r \sum_{j \in M} \delta_{ij})$

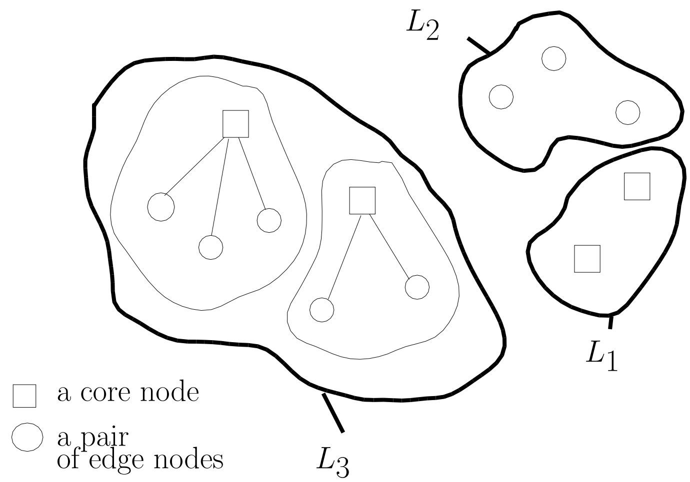


Figure 3: The sets L_1, L_2 et L_3 associated with a packing Π

2. the cost of the propagation delay for the assigned edge node pairs of L_3 :
   $\sum_{\{((i,r,e),p)|((i,r,e),D)\in L_3, p\in D\}} \beta\, d_{ip}\, Q_p$
3. the cost of the edge node pairs of L_2 because unassigned pairs must be penalized so as to make them progressively disappear :
   $\mathcal{M} * n_2$, where $\mathcal{M}$ is a very large number.

In a repeated matching approach, we want to match elements of L_1, L_2 and L_3 so as to generate new sets L'_1, L'_2 and L'_3 that have a lower total cost. The cost of the packing is reduced at each iteration, details will be given in Section 3.2 and in Appendix A.

The value of $\mathcal{M}$ is chosen very high. Consequently, the matching aims at reducing the number of elements in L_2. At last, all edge node pairs will be assigned to a core node. Then the path for each edge node connection will be known and the active core nodes will be identified. The problem of the PetaWeb design will be solved.

3.2 The matching problem

The classical matching problem can be described as following :

Let A be a set of q elements $h_1, h_2, ..., h_q$. A matching over A is so that each $h_i \in A$ can be matched with only one $h_j \in A$. An element can be matched with itself, which means that it remains unmatched.

Let c_{ij} be the cost of matching h_i with h_j. We have $c_{ij} = c_{ji}$.

We introduce the variable :

$$z_{ij} = \begin{cases} 1 & \text{if } h_i \text{ is matched with } h_j, \\ 0 & \text{else} \end{cases}$$

The matching problem consists in finding the matching over A that minimizes the total cost of matched pairs :

$$\min \sum_{i=1}^{q} \sum_{j=1}^{q} c_{ij} z_{ij} \tag{11}$$

subject to :

$$\sum_{j=1}^{q} z_{ij} = 1, \qquad i = 1, ..., q \tag{12}$$

$$\sum_{i=1}^{q} z_{ij} = 1, \qquad j = 1, ..., q \tag{13}$$

$$z_{ij} = z_{ji}, \qquad i, j = 1, ..., q \tag{14}$$

$$z_{ij} \in 0, 1, \qquad i, j = 1, ..., q \tag{15}$$

Constraints (12) and (13) ensure that each element is exactly matched with another one.

Constraint (14) imposes that if h_i is matched with h_j, then h_j is matched with h_i.

Constraint (15) indicates that variable z_{ij} is binary.

In our heuristic, one matching problem is solved at each iteration between the elements of L_1, the elements of L_2 and the elements of L_3. At each iteration, the number of elements to be matched is $n_1 + n_2 + n_3$, where n_1, n_2 and n_3 are the current cardinalities of the sets L_1, L_2 and L_3.

For each matching problem, the costs c_{ij} have to be evaluated. The cost c_{ij} is the cost of the resulting packing after having matched element h_i of L_1, L_2 or L_3 with element h_j of L_1, L_2 or L_3.

The costs c_{ij} are stored in a matrix C. The dimension of cost matrix C is $(n_1 * n_2 * n_3) * (n_1 * n_2 * n_3)$. Note that this dimension changes at each iteration.

C is a symmetric matrix composed of nine sub-matrices. Given the symmetry, only six blocks must be considered. The notation $[L_i - L_j]$ is used to indicate the matching between the elements of L_i and the elements of L_j.

$$C = \begin{pmatrix} \text{[L1-L1]} & \text{[-]} & \text{[-]} \\ \text{[L2-L1]} & \text{[L2-L2]} & \text{[-]} \\ \text{[L3-L1]} & \text{[L3-L2]} & \text{[L3-L3]} \end{pmatrix}$$

$$= \begin{pmatrix} [1] & [-] & [-] \\ [2] & [3] & [-] \\ [4] & [5] & [6] \end{pmatrix}$$

Calculating the matching costs for blocks 1 and 3 is a very easy task. For the other blocks, it requires great attention since the feasibility of the matching result has to be verified. Furthermore, a matching between two elements can produce several results. In such a case, the result with minimal cost is chosen. We develop the matching costs for each block in the Appendix A.

Once the cost matrix is calculated, the matching problem (11)–(15) is solved. The resolution is not easy because of the symmetry constraint (14). We have implemented the algorithm of Forbes [11] that is based on the method of Engquist [12]. The starting point for Forbes' algorithm is the solution of the matching problem without the symmetry constraint (14)). Such a starting solution is obtained with the algorithm of Jonker and Volgenant [13] which was chosen for its speed performance.

Figure 4 illustrates a possible solution of the matching problem.

The solution of the matching problem is then analysed. Some matchings result in new elements in L'_1, L'_2 and L'_3 whereas other elements disappear. For example, the matching between an inactive core node (i, r, e) of L_1 and an unassigned edge node pair p of L_2 results in the new element $((i, r, e), D = \{p\})$ of L_3.

3.3 The repeated matching heuristic for the PetaWeb design

A global chart of the heuristic is given in Figure 5.

The algorithm starts with a feasible packing. We choose a packing where no core node is opened and no edge node pair is assigned:

$L_1 = \{$all potential core nodes$\}$,

$L_2 = \{$all origin/destination edge node pairs exchanging traffic$\}$,

$L_3 = \emptyset$.

A series of feasible packings with decreasing cost is formed.

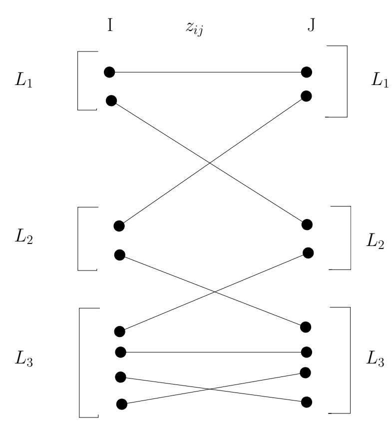


Figure 4: The solution of the matching problem

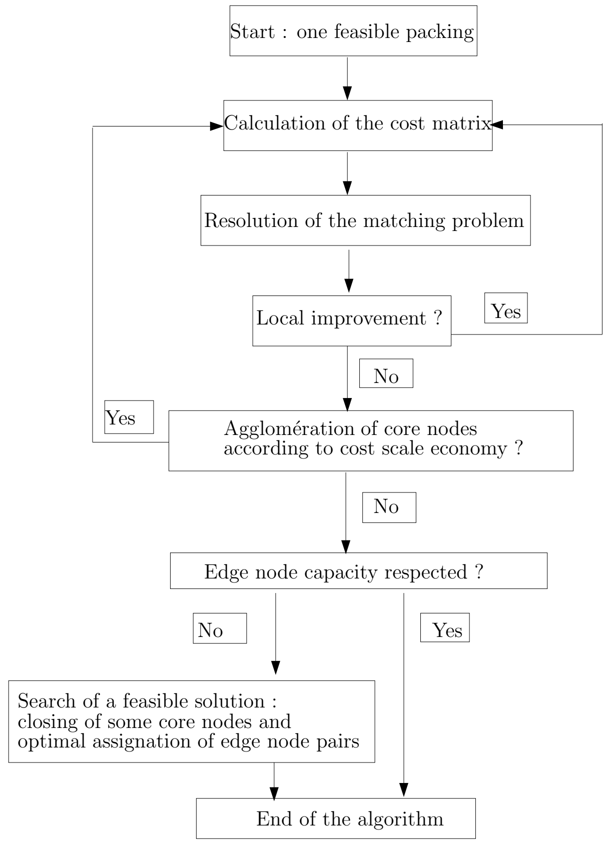


Figure 5: Chart of the repeated matching heuristic for the PetaWeb design, figure inspired by the article of Rönnqvist [9]

When the cost of the packing can not be reduced any more, active core nodes are agglomerated so as to take into account the scale economy in the core node cost. Indeed, a core node of type 2 opened at a site presents the same capacity but it is less expensive than two core nodes of type 1. The same can be said for one type 3 compared with two type 2 core nodes. We underline that the heuristic could not do these agglomerations while building packings with lower cost.

If one agglomeration at least is relevant, a new series of feasible packings is generated. Such a process is repeated until no progress can be done. Finally, one constraint must yet be verified : the edge node capacity constraint. This constraint has been omitted by now in order to allow multiple little kits to be built at the beginning of the algorithm and then be agglomerated. Knowing the active core nodes in the current best solution, we verify if :

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

  - If the constraint is respected, the heuristic stops.
  - If the constraint is not respected, the heuristic searches for a feasible solution in restricting the number of active core nodes.

- If one edge node capacity is exceeded by one link, a core node of type 1 or the equivalent capacity must be closed in the network. Step by step, at each site, the equivalent of a core node of type 1 is closed and the optimal assignment of all edge node pairs to the core nodes remaining active is calculated. This assignment must verify the capacity of each core node still active and the link capacity between each edge node and each active core node. The optimal assignment is solved by the software CPLEX. Whenever the equivalent of a core node of type 1 is closed at one site, the total cost of the network with optimal assignment of the pairs is calculated. Finally, we choose the solution with the lowest total network cost.
- If one edge node capacity is exceeded by two links, a core node of type 2 or the equivalent capacity must be closed in the network. Each combination is tried to close the equivalent of a core node of type 2 in the network.
- If one edge node capacity is exceeded by more than two links, we sporadically choose the core nodes that will be reduced in capacity or entirely closed.

4 Computational results

The proposed heuristic was tested using two networks, respectively composed of 10 and 34 edge nodes. The locations of the edge nodes are specific cities of the United States.

Two matrices were used :

- Traffic matrix A, which is a sparse matrix provided by industrial data,
- Traffic matrix B, that is calculated using a gravity model based on urban populations and distances between cities. The urban populations were found in [14]. Note that this matrix does not include any zeros, except on its diagonal entry.

Note that the distance matrix between edge nodes was calculated as follows. To work with realistic distances, geographical coordinates were first found in an American national atlas [15] and a formula to assess the distance between two points on a sphere [16] was used. The calculated distances were later compared and validated with a few air distances estimated at the University of Minnesota [17].

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 : $v = 3$,
Maximal number of core nodes of one type at one site : $e = 3$, except for the 34 edge nodes network, traffic matrix B, where $e = 4$ for the core node of type 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.

4.1 Results with default parameters

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 number of iterations and the solution time are given at the end. The actual solutions obtained for all instances treated are presented in Figure 7 and in Figure 8. The legend used is portrayed in Figure 6.

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, the number of iterations is quite small. The solution time is really short : it is about 10 seconds for the 10 edge nodes network and about 5 minutes for the 34 edge nodes network.

4.2 Relevance of the heuristic

It is interesting to compare the results obtained by the heuristic with the results obtained by CPLEX for the mathematical model previously presented in[8] so as to assess the relevance of the heuristic. The results are presented in Table 2 for the 10 edge nodes network and in Table 3 for the 34 edge nodes network. The gap in the next to last line is the optimality gap given by the CPLEX solution. The gap in the last line is the discrepancy in percentage between the total network cost found by the heuristic and the total network cost found by CPLEX for the mathematical model.

We note that the gaps between the total network cost found by the heuristic and the one found by CPLEX are very tightened. They lie below 0.4% for the 10 edge nodes network and below 5.5% for the 34 edge nodes network. The opened sites vary between the two results but no trend can be clearly identified. The percentages of the several costs in the objective function are quite identical between the heuristic solution and the CPLEX solution.

The great difference between the two results concerns the solution time : the solution time of the heuristic is insignificant in comparison to CPLEX's. For example, for the 10 edge nodes network with traffic matrix A, the solution time of the heuristic is about 6 seconds while it is about 6.6 hours for CPLEX. For the 34 edge nodes network with traffic matrix B, the solution time of the heuristic is about 6 minutes while it is about 448 hours for CPLEX.
We can conclude that the heuristic finds solutions similar to CPLEX's but with great time economy. The use of the heuristic is strongly relevant.

4.3 Sensitivity of the heuristic to the propagation delay cost

In the heuristic solution for the 34 edge nodes network with traffic matrix B, all active core nodes are located in Topeka. They are all opened at the geographical barycenter of the country, as if the propagation delay cost has not been taken into account. Therefore we can wonder if our heuristic is sensitive to the propagation delay weight.

Some tests were made for the 34 edge nodes network in order to measure the heuristic sensitivity to the propagation delay cost. The parameter β that represents the propagation delay cost was progressively increased. The results for the traffic matrix A are presented in Table 4 and in Figure 9.

It clearly appears that the heuristic is sensitive to the propagation delay cost. When the coefficient β increases, the active core nodes are more spread in the country. The total cost of the network increases, as well as the proportion of the delay cost in the total cost. These results are coherent.
We can notice that the percentages of the core node cost and of the fiber cost in the total cost decrease. In fact, the number and the type of the active core nodes are constant when β increases. The fiber cost increases with the weight of the propagation delay because the core nodes are more spread. However, as the delay cost increases faster, the proportion of the fiber cost in the total cost decreases.

4.4 Limits of the heuristic

The heuristic has given very good results for the 10 and 34 edge nodes networks. We are now looking for the limits of the heuristic when the size of the network increases.

Some tests were made adding at each time some cities of the United States according to their decreasing population importance. For each test, a full traffic matrix was elaborated using the gravity model. The sum of the total exchanged traffic was the same for all cases.

The values of the parameters were default values. The parameter representing the propagation delay was increased to $\beta = 1$. The maximum core nodes of one type that could be opened at one site was 4 and the maximum edge node capacity was $C_j = 3000$ Gbit/s.

The results given by the heuristic within one week for 40 to 136 edge nodes are given in Tables 5 and 6. Three cases are illustrated in Figure 10.

We note that the heuristic finds a solution for all cases within one week except for the 90 edge nodes network that will be explained later on. The total network cost increases when the network size is growing. The proportions of the several costs in the total cost keep quite constant. The fiber cost predominates with a percentage of 60% to 70% of the total cost. The delay cost comes next with a percentage of 25% to 35% of the total cost. The core node cost is the lowest with a percentage of 4% to 5.5% of the total cost. Opened sites vary slightly from one case to another.

As expected, the resolution time increases with the network size. However, it appears that some cases are quite more difficult to solve. For example, the 90 edge nodes network calculation needs more than one week. The difficulty is located in the matching problem resolution. Other tests were triggered to better characterize the solution time. Figure 11 illustrates the solution time diagram for 10 to 130 edge nodes networks with all default parameters. The time limit was fixed at one day.

5 Conclusions and further work

In this paper, we developed a specialized method to treat large-scale instances of the PetaWeb design problem. The problem was previously defined as a unique network design problem that is equivalent to the design of composed stars with specific capacity constraints. The proposed resolution approach is a heuristic based on repeated matchings.

We included in the design three types of costs : core, fiber and delay-related costs. We found that the cost distribution in both formulations were quite similar. In particular, in both cases up to 80% of the costs were due to fiber costs. We verified that the heuristic was sensitive to the weight of the propagation delay cost.

We also proved the great relevance of the heuristic. We noticed that the gap between the network cost found by the heuristic and the network cost found by a solver for the mathematical model is kept within 5.5%.

In terms of computational complexity, it appeared that the solution time of the heuristic is amazingly short in comparison to CPLEX time. The heuristic lasts some minutes, while the resolution of the mathematical model by a solver may need many days.

The proposed method is very robust and scalable. In fact, networks with more than hundred edge nodes were treated. As expected, the solution time increases exponentially with the network size but some cases are more difficult to solve than others. We have observed that the difficulty lays on the resolution of the matching problem. This provides an avenue to improve our resolution method.

Table 1: Results obtained with the default parameters by the heuristic

Network	10 edge nodes traffic A	10 edge nodes traffic B	34 edge nodes traffic A	34 edge nodes traffic B
Objective function	2289564	2153868	31940857	44757016
Percentage of the core node cost	11.4%	11.9%	5.5%	5.4%
Percentage of the fiber cost	77.1%	83.3%	82%	82.2%
Percentage of the delay cost	11.5%	4.8%	12.6%	12.5%
Number of iterations	9	15	17	12
Solution time	$6s$	$11s$	$217s$	$322s$

Table 2: Results obtained with the default parameters for the 10 edge nodes network

Network	Traffic A heuristic	Traffic A math. model CPLEX	Traffic B heuristic	Traffic B math. model CPLEX
Objective function	2289564	2280980	2153868	2152920
Percentage of the core node cost	11.4%	11.2%	11.9%	12.1%
Percentage of the fiber cost	77.1%	77.8%	83.3%	83.8%
Percentage of the delay cost	11.5%	11%	4.8%	4.1%
Solution time	$6s$	$23650s$	$11s$	$232s$
Optimality gap (CPLEX)	N/A	0.01%	N/A	0%
Gap between heuristic and math. model	0.38%	N/A	0.04%	N/A

Table 3: Results obtained with the default parameters for the 34 edge nodes network

Network	Traffic A heuristic	Traffic A math. model CPLEX	Traffic B heuristic	Traffic B math. model CPLEX
Objective function	31940857	31837547	44757016	42406000
Percentage of the core node cost	5.5%	5.3%	5.4%	5.3%
Percentage of the fiber cost	82%	81.7%	82.2%	81.6%
Percentage of the delay cost	12.6%	13%	12.5%	13.1%
Solution time	$217s$	$579998s$	$322s$	$1614383s$
Optimality gap (CPLEX)	N/A	7.22%	N/A	0%
Gap between heuristic and math. model	0.32%	N/A	5.5%	N/A

Table 4: Influence of the propagation delay weight for the 34 edge nodes network, traffic matrix A, with default parameters (heuristic)

β value	0.1	0.5	1	1.5
Objective function	31940857	46864904	61244206	74650622
Percentage of the core node cost	5.5%	3.7%	2.9%	2.4%
Percentage of the fiber cost	82%	60.3%	46.6%	41.7%
Percentage of the delay cost	12.5%	36.0%	50.5%	55.9%
Solution time	$217s$	$234s$	$349s$	$386s$

Table 5: Results for scalable networks with $\beta = 1$ and the default other parameters (heuristic)

Network	40 nodes traffic B	50 nodes traffic B	60 nodes traffic B	70 nodes traffic B	80 nodes traffic B
Total cost	66790998	66213371	71377149	72220022	83079134
Percentage of the core node cost	3.9%	4.6%	4.6%	5.0%	5.5%
Percentage of the fiber cost	60.0%	59.3%	62.8%	64.7%	70.5%
Percentage of the delay cost	36.1%	36.2%	32.7%	30.3%	24.0%
Number of iterations	22	19	26	33	23
Solution Time	$981s$	$2109s$	$6226s$	$13497s$	$13016s$

Table 6: Results for scalable networks with $\beta = 1$ and the default other parameters (heuristic)

Network	100 nodes traffic B	110 nodes traffic B	120 nodes traffic B	130 nodes traffic B	136 nodes traffic B
Total cost	88692651	89806499	103524125	102858855	104526279
Percentage of the core node cost	5.3%	5.2%	5.4%	5.4%	5.8%
Percentage of the fiber cost	65.4%	66.4%	67.4%	69.3%	70.6%
Percentage of the delay cost	29.3%	28.4%	27.2%	25.3%	23.6%
Number of iterations	28	27	27	30	38
Solution Time	$90369s$	$71619s$	$555416s$	$155500s$	$505115s$

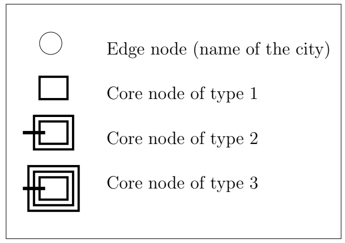
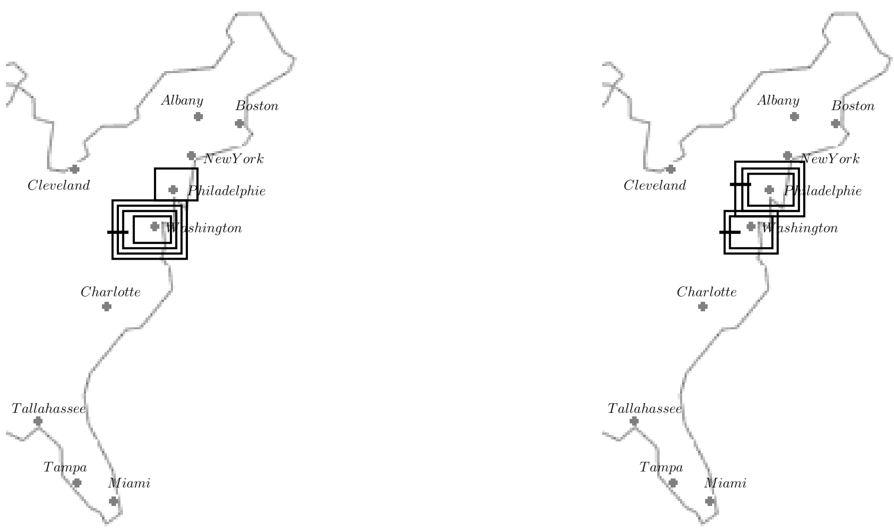
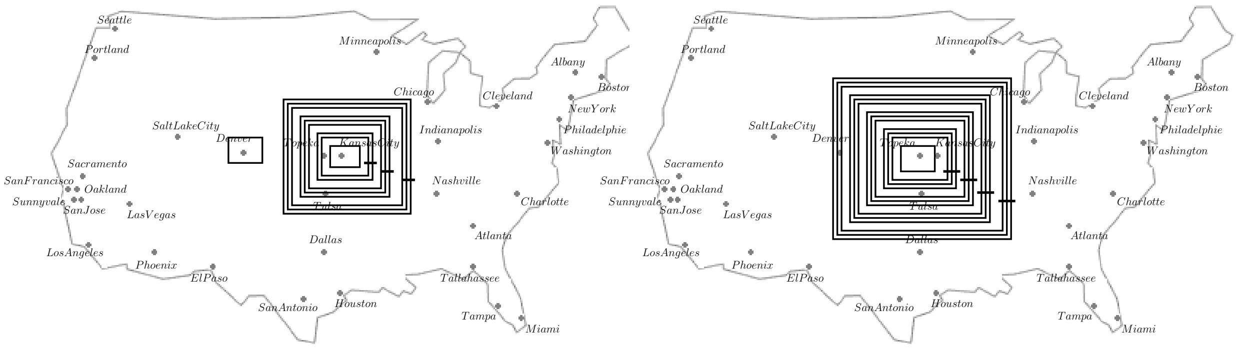


Figure 6: Legend



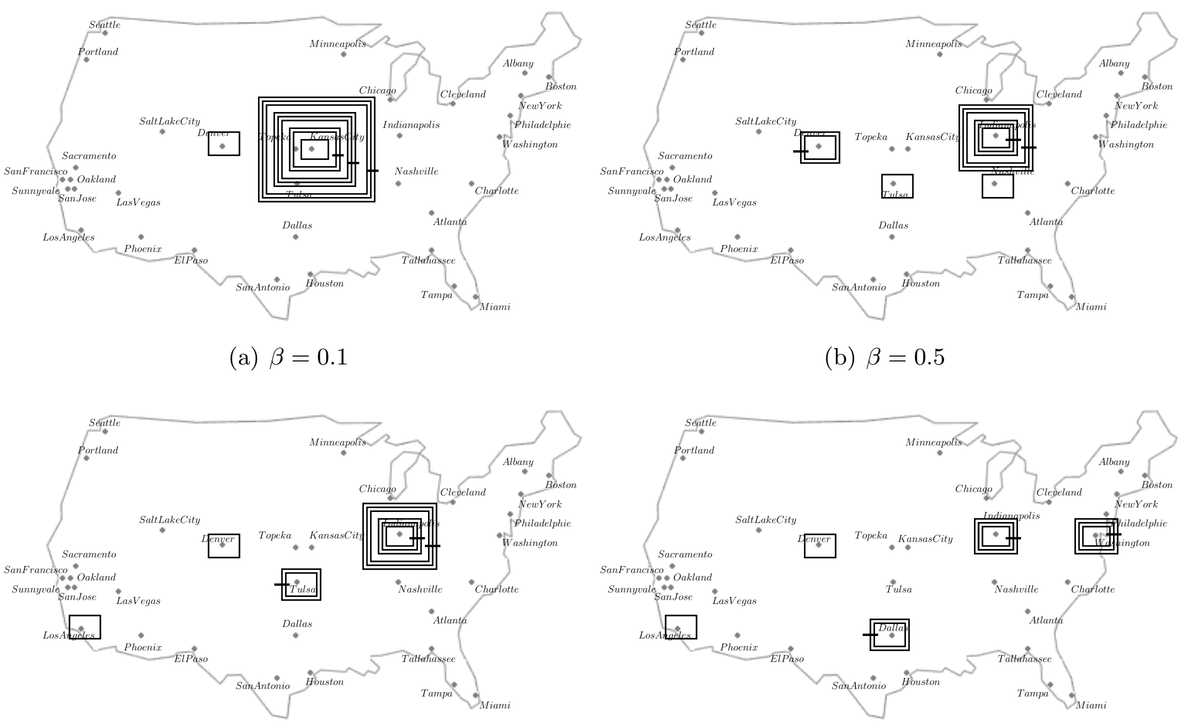
(a) Traffic matrix A (b) Traffic matrix B

Figure 7: 10 edge nodes network with default parameters (heuristic)



(a) Traffic matrix A (b) Traffic matrix B

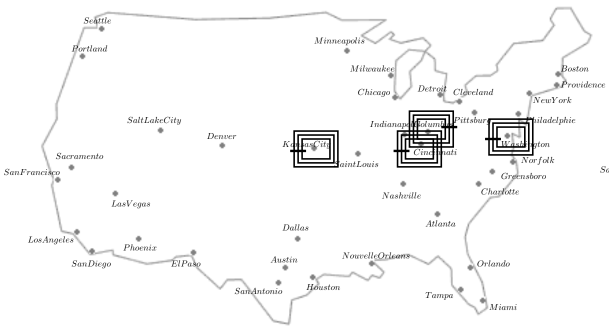
Figure 8: 34 edge nodes network with default parameters (heuristic)



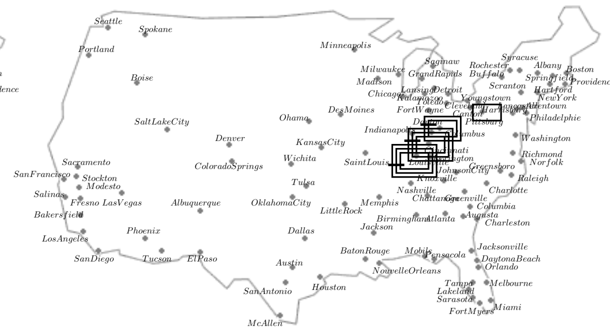
(a) $\beta = 0.1$ (b) $\beta = 0.5$

(c) $\beta = 1$ (d) $\beta = 1.5$

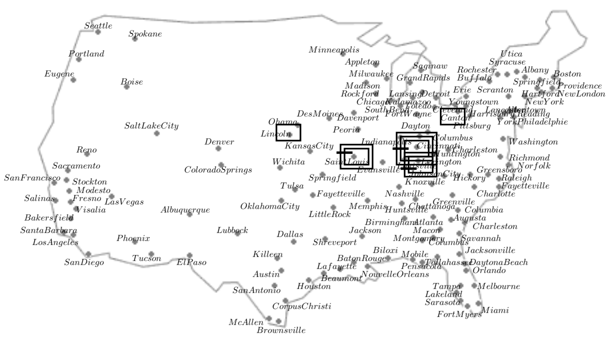
Figure 9: 34 edge nodes network, traffic matrix A, for several propagation delay weights (heuristic)



(a) 40 edge nodes



(b) 100 edge nodes



(c) 136 edge nodes

Figure 10: Scalable networks with traffic matrix B, $\beta = 1$ and the other default parameters (heuristic)

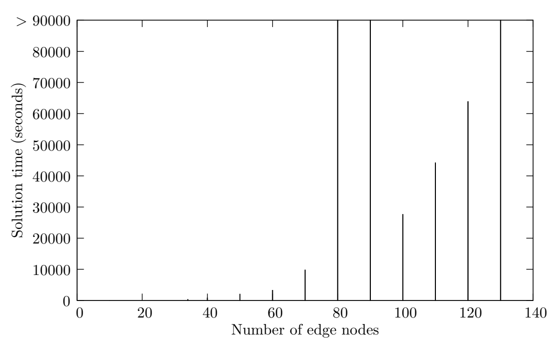


Figure 11: Solution time for scalable networks with all default parameters (heuristic)

Appendix A Matching costs

In this section, we develop the matching costs for each block of the symmetric cost matrix presented in Section 3.2.

Block 1

Matching two inactive core nodes.

Let (i_1, r_1, e_1) be the s^{th} core node of L_1 and (i_2, r_2, e_2) be the t^{th} core node of L_1.

The matching cost for block 1 is :

$$c_{s,t} = \begin{cases} \infty & \text{if } s \neq t, \\ 0 & \text{if } s = t \end{cases}$$

Block 2

Matching an unassigned edge node pair with an inactive core node.

Let p be the s^{th} pair of L_2 with origin/destination (j,k) and (i, r, e) be the t^{th} core node of L_1.

The matching is allowed if :

1. the capacity of core node (i, r, e) is respected :
   $Q_p \leq C_{canal} * M * W * s_r$,
2. the link capacity between the origin j of the pair p and the core node (i, r, e) on the one hand, and the link capacity between the core node (i, r, e) and the destination k of the pair p on the other hand, are respected :
   $Q_p \leq C_{canal} * W * s_r$.

If the capacity constraints are verified, the matching results in a new element $((i, r, e), D = \{p\})$ of L_3 whose cost is the sum of :

- the cost of the core node (i, r, e) : $2\,|M|\,W\,s_r\,\gamma^{(s_r-1)}\,P + f_r$
- the cost of the fiber between the core node (i, r, e) and all edge nodes in the network : $2\,W\,F\,s_r\,\sum_{j\in M}\delta_{ij}$
- the cost of the propagation delay of the pair p traffic via the core node (i, r, e) : $\beta d_{ip} Q_p$.

The matching cost for the block 2 is finally :

$$c_{n_1+s,t} = \begin{cases} 2\,|M|\,W\,s_r\,\gamma^{(s_r-1)}\,P + f_r \\ +2\,W\,F\,s_r\,\sum_{j\in M}\delta_{ij} + \beta d_{ip} Q_p \\ \qquad \text{if } Q_p \leq C_{canal} * W * s_r, \\ \infty \quad \text{otherwise} \end{cases}$$

Block 3

Matching two unassigned edge node pairs.

If the two pairs are different, the matching is impossible and the cost is set to infinity.

If a pair is matched with itself, it remains unmatched. The cost is twice the cost of one unassigned pair because each matching cost must appear twice in the objective function.

Let p_1 be the s^{th} unassigned edge node pair of L_2 and let p_2 be the t^{th} unassigned edge node pair of L_2.

The matching cost for the block 3 is :

$$c_{n_1+s,n_1+t} = \begin{cases} \infty & \text{if } s \neq t, \\ 2\mathcal{M} & \text{if } s = t \end{cases}$$

Block 6

Matching two kits of L_3.

Let $((i_1, r_1, e_1), D_1)$ be the s^{th} kit of L_3 and $((i_2, r_2, e_2), D_2)$ be the t^{th} kit of L_3.

If $s = t$, the element is matched with itself. The matching cost is twice the cost of one element as explained above. We remind the reader that the cost of the kit $((i_1, r_1, e_1), D_1)$ is composed of :

- the cost of the core node (i_1, r_1, e_1) : $2\,|M|\,W\,s_{r_1}\,\gamma^{(s_{r_1}-1)}\,P + f_{r_1}$
- the cost of the fiber between the core node (i_1, r_1, e_1) and all edge nodes : $2\,W\,F\,s_{r_1} \sum_{j\in M} \delta_{i_1 j}$
- the cost of the propagation delay of the D_1 traffic pairs via the core node (i_1, r_1, e_1) : $\beta \sum_{p\in D_1} d_{i_1 p}\, Q_p$.

The self-matching cost is then :
$2 * (2\,M\,W\,s_{r_1}\gamma^{(r_1-1)}\,P + f_{r_1} + 2\,W\,F \sum_{j\in M} \delta_{i_1 j}$
$+ \beta \sum_{p\in D_1} d_{i_1 p}\, Q_p)$

If $s \neq t$, three cases must be considered :

Case I : All edge node pairs of D_1 and D_2 are assigned to the core node (i_1, r_1, e_1).

This case is possible if :

- the capacity of the core node (i_1, r_1, e_1) is respected, that is : $\sum_{p\in D_1} Q_p + \sum_{p\in D_2} Q_p \leq C_{canal}\,M\,W\,s_{r_1}$.
- the link capacity between each origin edge node of D_1 and D_2 and the core node (i_1, r_1, e_1) on the one hand, and the link capacity between the core node (i_1, r_1, e_1) and each destination edge node of D_1 and D_2 on the other hand, are respected : $\sum_{p\in(D_1\cup D_2)\in Orig_j} Q_p \leq C_{canal}\,W\,s_{r_1}$, $\forall origin\, j \in (D_1 \cup D_2)$ and $\sum_{p\in(D_1\cup D_2)\in Dest_k} Q_p \leq C_{canal}\,W\,s_{r_1}$, $\forall destination\, k \in (D_1 \cup D_2)$.

The matching cost for this case is then :

$$v_I = \begin{cases} 2\,M\,W\,s_{r_1}\gamma^{(s_{r_1}-1)}\,P + f_{r_1} \\ +2\,W\,F\,s_r\sum_{j\in M}\delta_{i_1 j} + \beta\sum_{p\in(D_1\cup D_2)} d_{i_1 p}Q_p \\ \quad \text{if } \sum_{p\in D_1} Q_p + \sum_{p\in D_2} Q_p \le C_{canal}\,W\,M\,s_{r_1} \\ \quad \text{and } \sum_{p\in(D_1\cup D_2)\in Orig_j} Q_p \le C_{canal}\,W\,s_{r_1}, \\ \quad \forall\, origin\, j \in (D_1\cup D_2), \\ \quad \text{and } \sum_{p\in(D_1\cup D_2)\in Dest_k} Q_p \le C_{canal}\,W\,s_{r_1}, \\ \quad \forall\, destination\, k \in (D_1\cup D_2), \\ \infty \quad \text{otherwise} \end{cases} \tag{16}$$

Case II : All edge node pairs of D_1 and D_2 are assigned to the core node (i_2, r_2, e_2).

This case is the same as the one before if we reverse the roles of the core nodes.

The matching cost for this case is then :

$$v_{II} = \begin{cases} 2\,M\,W\,s_{r_2}\gamma^{(s_{r_2}-1)}\,P + f_{r_2} \\ +2\,W\,F\,s_r\sum_{j\in M}\delta_{i_2 j} + \beta\sum_{p\in(D_1\cup D_2)} d_{i_2 p}Q_p \\ \quad \text{if } \sum_{p\in D_1} Q_p + \sum_{p\in D_2} Q_p \le C_{canal}\,W\,M\,s_{r_2} \\ \quad \text{and } \sum_{p\in(D_1\cup D_2)\in Orig_j} Q_p \le C_{canal}\,W\,s_{r_2}, \\ \quad \forall\, origin\, j \in (D_1\cup D_2), \\ \quad \text{and } \sum_{p\in(D_1\cup D_2)\in Dest_k} Q_p \le C_{canal}\,W\,s_{r_2}, \\ \quad \forall\, destination\, k \in (D_1\cup D_2), \\ \infty \quad \text{otherwise} \end{cases} \tag{17}$$

Case III : The core nodes (i_1, r_1, e_1) and (i_2, r_2, e_2) are both active.

This is a difficult case because the core nodes may exchange some edge node pairs. We then need to find the optimal assignment of the pairs to the two core nodes. A mathematical formulation of this integer problem must be given.

Let us introduce some additional variables :

$$w_p = \begin{cases} 1 & \text{if the pair } p \in D_1 \text{ switches its core node} \\ & \text{and is assigned to the core node } (i_2, r_2, e_2), \\ 0 & \text{otherwise} \end{cases}$$

$$z_p = \begin{cases} 1 & \text{if the pair } p \in D_2 \text{ switches its core node} \\ & \text{and is assigned to the core node } (i_1, r_1, e_1), \\ 0 & \text{otherwise} \end{cases}$$

The pair exchange problem can be formulated as :

$$v* = \min \sum_{p\in D_1} g_p w_p + \sum_{p\in D_2} h_p z_p \tag{18}$$

subject to :

$$\sum_{p \in D_1} -Q_p w_p + \sum_{p \in D_2} Q_p z_p \leq \delta_w, \tag{19}$$

$$\sum_{p \in D_1} Q_p w_p - \sum_{p \in D_2} Q_p z_p \leq \delta_z, \tag{20}$$

$$\sum_{p \in D_1 \in Orig_j} -Q_p w_p + \sum_{p \in D_2 \in Orig_j} Q_p z_p \leq \epsilon_{wj}$$
$$\forall\, origin\, j \in (D_1 \cup D_2), \tag{21}$$

$$\sum_{p \in D_1 \in Orig_j} Q_p w_p - \sum_{p \in D_2 \in Orig_j} Q_p z_p \leq \epsilon_{zj}$$
$$\forall\, origin\, j \in (D_1 \cup D_2), \tag{22}$$

$$\sum_{p \in D_1 \in Dest_k} -Q_p w_p + \sum_{p \in D_2 \in Dest_k} Q_p z_p \leq \eta_{wj}$$
$$\forall\, destination\, k \in (D_1 \cup D_2), \tag{23}$$

$$\sum_{p \in D_1 \in Dest_k} Q_p w_p - \sum_{p \in D_2 \in Dest_k} Q_p z_p \leq \eta_{zj}$$
$$\forall\, destination\, k \in (D_1 \cup D_2), \tag{24}$$

$$w_p, z_p \in 0, 1, \quad \forall p \in (D_1 \cup D_2). \tag{25}$$

g_p and h_p are the marginal costs if the edge node pair p switches its core node.
δ_w and δ_z are the surplus capacity of the core nodes (i_1, r_1, e_1) and (i_2, r_2, e_2) respectively.
ϵ_{wj} is the surplus capacity of the links between the origin edge node j and the core node (i_1, r_1, e_1).
ϵ_{zj} is the surplus capacity of the links between the origin edge node j and the core node (i_2, r_2, e_2).
η_{wk} is the surplus capacity of the links between the core node (i_1, r_1, e_1) and the destination edge node k.
η_{zk} is the surplus capacity of the links between the core node (i_2, r_2, e_2) and the destination edge node k.

$$\begin{aligned} g_p &= \beta d_{i_2 p} Q_p - \beta d_{i_1 p} Q_p, \forall p \in D_1, \\ h_p &= \beta d_{i_1 p} Q_p - \beta d_{i_2 p} Q_p, \forall p \in D_2, \end{aligned}$$

$$
\begin{aligned}
\delta_w &= C_{canal} * M * W * s_{r_1} - \sum_{p \in D_1} Q_p, \\
\delta_z &= C_{canal} * M * W * s_{r_2} - \sum_{p \in D_2} Q_p, \\
\epsilon_{wj} &= C_{canal} * W * s_{r_1} - \sum_{p \in D_1 \in Orig_j} Q_p, \\
& \quad \forall\, origin\, j \in (D_1 \cup D_2), \\
\epsilon_{zj} &= C_{canal} * W * s_{r_2} - \sum_{p \in D_2 \in Orig_j} Q_p, \\
& \quad \forall\, origin\, j \in (D_1 \cup D_2), \\
\eta_{wk} &= C_{canal} * W * s_{r_1} - \sum_{p \in D_1 \in Dest_k} Q_p, \\
& \quad \forall\, destination\, k \in (D_1 \cup D_2), \\
\eta_{zk} &= C_{canal} * W * s_{r_2} - \sum_{p \in D_2 \in Dest_k} Q_p, \\
& \quad \forall\, destination\, k (D_1 \cup D_2).
\end{aligned}
$$

The objective (18) is to minimize the cost of the packing.

(19) and (20) impose that the capacity surplus in the core node be respected.

(21) and (22) are surplus capacity constraints for the links between each origin edge node j and the core nodes (i_1, r_1, e_1) and (i_2, r_2, e_2) respectively.

(23) and (24) are surplus capacity constraints for the links between the core nodes (i_1, r_1, e_1) and (i_2, r_2, e_2) respectively and each destination edge node k.

(25) indicates the variables w_p and z_p are binary.

The matching cost for this case is finally :

$$
\begin{aligned}
v_{III} &= 2\, M\, W\, s_{r_1} \gamma^{(r_1 - 1)}\, P + f_{r_1} + 2\, W\, F \sum_{j \in M} \delta_{i1j} \\
&+ \beta \sum_{p \in D_1} d_{i1p} Q_p + 2\, M\, W\, s_{r_2} \gamma^{(r_2 - 1)}\, P + f_{r_2} \\
&+ 2\, W\, F \sum_{j \in M} d_{i_2 j} Q_p + \beta \sum_{p \in D_2} d_{i_2 p} Q_p + v *
\end{aligned}
\tag{26}
$$

Among the three cases whenever $s \neq t$, we choose the solution with minimal cost : $\min\{v_I, v_{II}, v_{III}\}$.

At last, the matching cost for the block 6 is :

$$c_{n_1+n_2+s,n_1+n_2+t} = \begin{cases} 2(2\,M\,W\,s_{r_1}\gamma^{(r_1-1)}\,P + f_{r_1} \\ +2\,W\,F\sum_{j\in M}\delta_{i_1 j} \\ +\beta\sum_{p\in D_1} d_{i_1 p}\,Q_p) \\ \quad \text{if } t = s \\ \min\{v_I, v_{II}, v_{III}\} \quad \text{otherwise} \end{cases}$$

where v_I, v_{II} and v_{III} are given by the expressions (16), (17) and (26) respectively.

Block 4

Matching a kit of L_3 with an inactive core node of L_1. Note that this is a particular case of block 6.

Let $((i_1, r_1, e_1)$,$D_1)$ be the s^{th} kit of L_3 and (i, r, e) be the t^{th} core node of L_1. The inactive core node (i, r, e) can be seen as an active core node with no assigned pair : $(i, r, e) = ((i_2, r_2, e_2), \emptyset) \in L_3$.

The matching cost for the block 4 is then :

$$c_{n_1+n_2+s,t} = \min\{v_I, v_{II}, v_{III}\},$$

where v_I, v_{II} and v_{III} are given by the equations (16), (17) and (26).

Block 5

Matching a kit of L_3 with an unassigned pair of L_2.

Let $((i_1, r_1, e_1)$,$D_1)$ be the s^{th} kit of L_3 and q be the t^{th} pair of L_2 with origin/destination $(j$,$k)$.

Two cases must be considered :

Case I : The unassigned edge node pair can be assigned to the core node (i_1, r_1, e_1). Then D_1 becomes $D_1 \cup q$.
This case is possible if :

- the capacity of the core node (i_1, r_1, e_1) is respected.
- the link capacity between the origin edge node j and the core node (i_1, r_1, e_1) on the one hand, and the link capacity between the core node (i_1, r_1, e_1) and the destination edge node k on the other hand, are respected.

The matching cost for this case is (D_1 now contains the pair q) :

$$c_{n_1+n_2+s,n_1+t} = \begin{cases} 2\,M\,W\,s_{r_1}\gamma^{(s_{r_1}-1)}\,P + f_{r_1} \\ +2\,W\,F\,s_{r_1}\,\sum_{j\in M}\delta_{ij} + \beta\sum_{p\in D_1} d_{ip}Q_p \\ \quad \text{if } \sum_{p\in D_1} Q_q \leq C_{canal}\,M\,W\,s_{r_1} \\ \quad \text{and } \sum_{p\in D\in Orig_j} Q_p \leq C_{canal}\,W\,s_{r_1}, \\ \quad \forall\, origin\, j \in D, \\ \quad \text{and } \sum_{p\in D\in Dest_k} Q_p \leq C_{canal}\,W\,s_{r_1}, \\ \quad \forall\, destination\, k \in D. \end{cases}$$

Case II : The unassigned edge node pair can not be assigned to the core node (i_1, r_1, e_1). If one capacity constraint is not respected, one pair or more have to be removed from the kit. A problem of pair exchange is then solved as for the block 6.

The edge node pair q is inserted in the set D_1. D_2 is built as an empty set.
$D_1 = D_1 \cup q$
$D_2 = \emptyset$.

We solve the problem (18)-(25) without considering the equations (20), (22) and (24) where :

$$w_p = \begin{cases} 1 & \text{if the pair } p \in D_1 \text{ removes from the} \\ & \text{core node } (i_1, r_1, e_1) \\ & \text{and becomes an unassigned pair of } L_2, \\ 0 & \text{else} \end{cases}$$

$z_p = 0$,
$g_p = \mathcal{M} - \beta\, d_{i1p}\, Q_p, \forall p \in D_1 \backslash \{paire\, q\}$,
$g_{|D_1|+1} = \mathcal{M} - \beta\, d_{i1q} Q_q$,
and $h_p = 0$.
Note that the surplus capacity δ_w, ϵ_{wj} and η_{wk} can be negative.

The set $\overline{D_1} \subset D_1$ corresponds to the edge node pairs assigned to the core node (i_1, r_1, e_1) in the exchange problem solution. Let $\overline{n_1}$ be the number of elements in $\overline{D_1}$.

The matching cost for this case is then :

$$\begin{aligned} c_{n_1+n_2+s,n_1+t} &= 2\,M\,W\,s_{r_1}\gamma^{(s_{r_1}-1)}\,P + f_{r_1} \\ &+ 2\,W\,F\,s_{r_1}\sum_{j\in M}\delta_{i_1j} \\ &+ \beta\sum_{p\in\overline{D_1}} d_{i_1p}Q_p + (n_1 - \overline{n_1})\mathcal{M} \end{aligned}$$

At last, the matching cost for the block 5 is :
(the reader is reminded that D_1 contains the pair of L_2)

$$c_{n_1+n_2+s,n_1+t} = \begin{cases} 2\,M\,W\,s_{r_1}\gamma^{(s_{r_1}-1)}\,P + f_{r_1} \\ +2\,W\,F\,s_{r_1}\,\sum_{j\in M}\delta_{i_1 j} + \beta\sum_{p\in D_1} d_{i_1 p}Q_p \\ \quad \text{if } \sum_{p\in D_1} Q_q \le C_{canal}\,M\,W\,s_{r_1} \\ \quad \text{and } \sum_{p\in D_1\in Orig_j} Q_p \le C_{canal}\,W\,s_{r_1}, \\ \quad \forall origin\, j\in D_1, \\ \quad \text{and } \sum_{p\in D\in Dest_k} Q_p \le C_{canal}\,W\,s_{r_1}, \\ \quad \forall destination\, k\in D_1, \\ 2\,M\,W\,s_{r_1}\gamma^{(s_{r_1}-1)}\,P + f_{r_1} \\ +2\,W\,F\,s_{r_1}\,\sum_{j\in M}\delta_{i_1 j} + \beta\sum_{p\in\overline{D_1}} d_{i_1 p}Q_p \\ +(n_1-\overline{n_1})\mathcal{M} \quad \text{otherwise} \end{cases}$$

References

[1] M. Beshai and R. Vickers, "Petaweb architecture," in *Networks 2000, Towards Natural Networks* 9^{th} *International Telecommunication Network Planning Symposium*, Toronto, 2000.

[2] M. Beshai *et al.*, "Petaweb-building block for a yottabit-per-second network," private communication.

[3] F. Blouin, A. Lee, A. Lee, and M. Beshai, "A comparison of two optical-core networks," *Journal of Optical Networking, Optical Society of America*, vol. 1, no. 1, 2002.

[4] J. Dégila and B. Sansò, "Design optimization of a next generation yottabit-per-second network," in *Proc. IEEE International Conference on Communications*, vol. 2, no. 20-24, 2004, pp. 1227–1231.

[5] ——, "Topological optimization of a yottabit-per-second lattice network," *Optical Communications and Networking of IEEE JSAC*, Nov 2004, to be published.

[6] J. Klincewicz, "Hub location in backbone/tributary network design : a review," in *Location Science*, 1998, vol. 6, pp. 307–335.

[7] M. Balinski, "On finding integer solutions to linear programs," in *Mathematica*, May 1964.

[8] A. Reinert and B. Sansò, "Design optimization of a new network structure for the next generation internet," *Les Cahiers du Gerad*, 2004, to appear.

[9] M. Rönnqvist, S. Tragantalerngsak, and J. Holt, "A repeated matching heuristic for the single-source capacitated facility location problem," *European Journal of Operational Research*, vol. 116, pp. 51–68, 1999.

[10] M. Rönnqvist, K. Holmberg, and D. Yuan, "An exact algorithm for the capacitated facility location problems with single sourcing," *European Journal of Operational Research*, vol. 113, pp. 544–559, 1999.

[11] M. Forbes, J. Holt, P. Kilby, and A. Watts, "A matching algorithm with application to bus operations," *Australian Journal of Combinatorics*, vol. 4, pp. 71–85, 1991.

[12] M. Engquist, "A successive shortest path algorithm for the assignment problem," in *INFOR*, 1982, vol. 20, pp. 370–384.

[13] R. Jonker and A. Volgenant, "A shortest augmenting path algorithm for dense and sparse linear assignment problems," *Computing*, vol. 38, pp. 325–340, 1986.

[14] E. US Department of Commerce and S. Administration. (2002) United states census 2000 population and housing. [Online]. Available: `http://www.census.gov/main/www/cen2000.html`

[15] *The National Atlas of the United States of America*, United States Department of the interior geological Survey, Washington, DC, 1970.

[16] F. o. E. S. University of Waterloo, Department of Geography. (2003) Geographic co-ordinates. [Online]. Available: `http://www.fes.uwaterloo.ca/crs/geog165/gcoords.htm`

[17] *Airmaps Inc*, Cartographic Laboratory Department of Geography, University of Minnesota, 1989.