# Selecting the Best Locations for Data Centers in Resilient Optical Grid/Cloud Dimensioning

## (Invited paper)

**Brigitte Jaumard**[*], **Ali Shaikh**[*], **Chris Develder**[†]

[*]*CSE, Concordia University, Montreal (Qc) H3G 1M8 Canada*
[†]*Dept. of Information Technology - IBCN, Ghent University – IBBT, Ghent, Belgium*

**ABSTRACT**

For optical grid/cloud scenarios, the dimensioning problem comprises not only deciding on the network dimensions (i.e., link bandwidths), but also choosing appropriate locations to install server infrastructure (i.e., data centers), as well as determining the amount of required server resources (for storage and/or processing). Given that users of such grid/cloud systems in general do not care about the exact physical locations of the server resources, a degree of freedom arises in choosing for each of their requests the most appropriate server location. We will exploit this anycast routing principle (i.e., source of traffic is given, but destination can be chosen rather freely) also to provide resilience: traffic may be relocated to alternate destinations in case of network/server failures. In this study, we propose to jointly optimize the link dimensioning and the location of the servers in an optical grid/cloud, where the anycast principle is applied for resiliency against either link or server node failures. While the data center location problem has some resemblance with either the classical p-center or k-means location problems, the anycast principle makes it much more difficult due to the requirement of link disjoint paths for ensuring grid resiliency.
**Keywords:** Optical Grids, Dimensioning Networks, ILP, Column Generation, Cloud Computing

## 1. INTRODUCTION

A strong motivation for an optical grid/cloud network is to share resources among multiple users distributed geographically but linked by high speed communication links. Optical grid/cloud networks emerge from WDM optical networks and grids/distributed computing. Research communities in different fields of e-Science are already using optical grid networks. But also more commercially oriented service providers (offering licensing software, storage-systems, e-commerce, and bandwidth-intensive multimedia services, etc.) can benefit from such distributed computing approaches to attain reliable and cost-effective infrastructure. As such, cloud computing evolved from grid computing, aiming at on-demand provisioning of resources. A well-known example showcasing the benefits for customers is that of *The New York Times* using Amazon infrastructure through to process terabytes data in hours [1]. For a comprehensive recent survey on optical grids and clouds, see, e.g., [2].

Below, we address a dimensioning problem, expanded for optical grid/cloud networks: given a set of traffic requests, determine (*i*) the link transport capacities, and (*ii*) the location of server resources.

## 2. RELATED WORK

Traffic engineering and the dimensioning of optical networks rely on Routing and Wavelength Assignment (RWA) with the objective of minimizing (min-RWA) the resources or maximizing (max-RWA) the Grade of Service (GoS). Different RWA approaches have been studied in the literature, and their ILP formulation comparison is presented in [3] and [4]. For optical grid/cloud environments, where the anycast routing principle — jobs can be executed on any designated server, and the location selection is often transparent to the end users — applies, the problem becomes that of anycast RWA (ARWA).

Buysse *et al.* [5] solve the optical grid network dimensioning problem in two steps: (*i*) first the server location is found, based on node origin of the job request, and its arrival rate; (*ii*) then the routing and wavelength assignment is optimized for working and backup paths. Other multiple-step approaches have been proposed: in [6], first the server locations are chosen, then the amount of computing resources, subsequently the inter-site traffic volumes, and finally the routing and resulting link transport capacities.

The optimization of the server location in an optical grid environment has some resemblance with classical facility location problems, namely the *p*-median and *p*-center problems, which have been much investigated in the literature, see, e.g., [7] and [8]. Both problems deal with the location of *p* facilities. The *p*-median problem searches the facility locations such that the sum of the shortest demand weighted distance between "customers" and *p* "facilities" is minimized. On the other hand, the *p*-center problem identifies the facility locations in order to minimize the maximal distance for all demand points. While the optical grid/cloud network dimensioning problem shares some features of the *p*-median problem, the

former is more complex due to: ($i$) the distance function (expressed in terms of, e.g., optical hops), and ($ii$) the additional requirements of backup paths in order to ensure network survivability. These last two features makes the joint optimization of finding server locations and dimensioning both working and backup paths, much more complex than the facility locations in a $p$-center or $p$-median context.

We next propose two new decomposition ILP models for the joint optimization of the location of servers and the link dimensioning catering for both primary and backup paths. Some numerical results follow.

## 3. JOINT OPTIMIZATION MODELS

The two proposed ILP models rely on a decomposition (column generation) scheme that extends our previous work on link dimensioning in an optical grid network [9]. The decomposition is around so-called configurations which refer to a set of potential primary and backup paths for a given source node. Details are given below for each model.

The optical network is assumed to be described by a directed graph $G = (V, L)$, where $V$ is the node set ($n = |V|$) and $L$ is the set of (directed) links. $D_v$ is number of job requests originating from source node $v \in V$, $n_{\mathrm{S}}$ is the maximum number of server nodes.

In both models, binary parameters $p_\ell^{\mathrm{W},c}$ (resp. $p_\ell^{\mathrm{B},c}$) equal 1 if link $\ell$ is used by the working (resp. backup) path in configuration $c$, 0 otherwise. Similarly, $a_v^c=1$ if node $v$ is selected as a server location either by the working or the backup path in configuration $c$, and 0 otherwise. Note that provisioning of job requests originating from $v$ is considered only if $v$ is not selected as a server node.

### 3.1. Model I

In Model I, a configuration $c$ is defined for a given source node $v \in V_{\mathrm{D}}$ and consists of a pair of working and backup paths from $v$ to a server node, see Fig. 1(a) for an illustration of two possible configurations for source node $v_1$ (the first one has a working path towards $v_7$, and one backup towards $v_6$, the second one has a working path towards $v_7$ and a backup towards $v_4$ as well as via $v_4$) and one potential one for source node $v_2$. The subset of configurations associated with source node $v$ is denoted by $C_v$, and $C = \bigcup_{v \in V_{\mathrm{D}}} C_v$. Model I comprises three sets of variables: $z_c \in \mathbb{Z}^+$ is equal to the number of selected copies of configuration $c$, $b_\ell^{\mathrm{B}} \in \mathbb{Z}^+$ is equal to the number of shared backup bandwidth units on link $\ell$, and $y_v \in \{0,1\}$ is a decision variable equal to 1 if node $v$ is chosen as a server location, 0 otherwise.

The objective function, which aims at minimizing the overall bandwidth requirements is as follows:

$$\min \quad \sum_{\ell \in L} \left( b_\ell^{\mathrm{B}} + \sum_{c \in C} p_\ell^{\mathrm{W},c} \cdot z_c \right) \tag{1}$$

$$\text{subject to:} \quad D_v \cdot y_v + \sum_{c \in C_v} z_c \geq D_v \qquad v \in V_{\mathrm{D}} \tag{2}$$

$$\sum_{c \in C} p_\ell^{\mathrm{W},c} \cdot p_{\ell'}^{\mathrm{B},c} \cdot z_c \leq b_{\ell'}^{\mathrm{B}} \qquad \ell, \ell' \in L : \ell \neq \ell' \tag{3}$$

$$\sum_{c \in C} a_v^c \cdot z_c \leq M \cdot y_v \qquad v \in V \tag{4}$$

$$\sum_{v \in V} y_v \leq n_{\mathrm{S}} \tag{5}$$

Constraints (2) are the demand constraints. If node $v$ hosts a server, then all the job requests originating from $v$ are readily served by $v$. Otherwise, we need some configurations associated with $v$ to satisfy the demand. Constraints (3) are used to compute the bandwidth requirement for a link $\ell'$ in a shared backup path. Indeed, if $\ell'$ protects link $\ell$, with $\ell$ belonging to several working paths (modeled here through the various configurations $c$ associated with working paths containing $\ell$), we must ensure $\ell'$ to have a large enough transport capacity. Constraints (4) prevents from selecting a configuration in which node $v$ has been selected as a server location when $y_v = 0$. The maximum number of server locations ($n_{\mathrm{S}}$) is controlled in (5).

### 3.2. Model II

In Model II, a configuration $c$ is defined for a given source node $v \in V_{\mathrm{D}}$ and consists of a set of working paths and a single backup path from $v$ to a server node, see Fig. 1(b) for an illustration with one potential configuration for node $v_1$ with two working path towards two different servers, $v_6$ and $v_7$, and one backup path towards $v_7$, and another possible configuration $v_2$ with two working paths towards the same server node $v_7$ and one backup towards $v_6$. At most one configuration per source node can be selected. Its

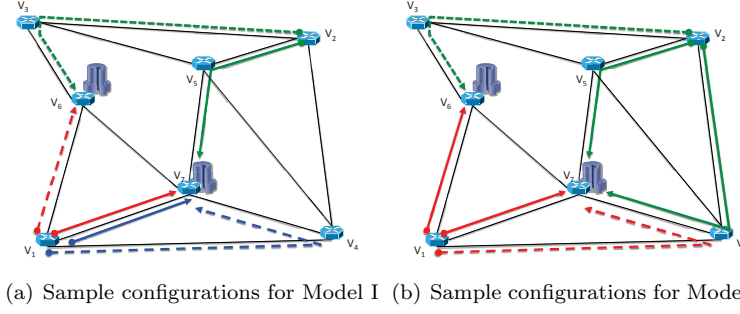(a) Sample configurations for Model I  (b) Sample configurations for Model II

Fig. 1: In Model I, a configuration comprises a single working and single backup path, where multiple such configurations can be used for a single source. In Model II, only a single configuration per source node will be used, but it consists of multiple working paths and a single backup. (Note that plain arrows represent working paths, dashed arrows are backup paths.)

mathematical expression can be rather easily deduced from that of Model I, but is not detailed here due to the lack of space.

### 3.3. Solving Models I and II

Both Models I and II are solved using the combination of a column generation algorithm with two different ILP heuristic solution schemes. The first one, referred below as the classical one, is very similar to the one in [9], i.e., corresponds to the use of an ILP solver on the constraint matrix associated with the set of generated columns in order to reach the optimal solution of the linear relaxation. The second, referred below as MILP, comprises the following four steps: (*i*) optimally solve the linear relaxation of the model using a column generation (CG) algorithm, (*ii*) solve the MILP problem obtained with an integer requirement only for the server variables ($y_v$), (*iii*) fix the server variables either to 0 or to 1 based on the MILP solution, and re-optimize the resulting linear relaxation, again with a CG algorithm, (*iv*) solve the ILP associated with the columns generated by the linear relaxation of Step *iii*.

### 4. NUMERICAL RESULTS

Both models were tested on a European network topology (28 nodes and 41 bidirectional links, see e.g., [9]) with $n_s = 5$ server (cluster) nodes, whose locations had to be optimized. We generated uniform and non-uniform incremental traffic instances from 50 to 500 (discrete) job requests. In the non-uniform traffic instances, the number of jobs is proportional to the population of the node (city). Models were solved using the IBM ILOG CPLEX Optimizer software.



(a) Link Protection
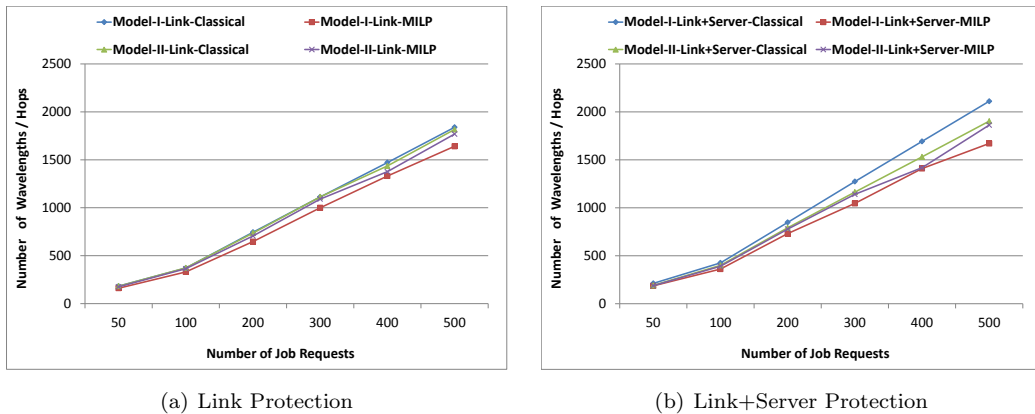
(b) Link+Server Protection

Fig. 2: Uniform traffic: Model I vs. Model II and Classical vs. MILP solution approach.

Evolution of bandwidth requirements with an increasing number of job requests is illustrated in Fig. 2 for uniform traffic. In Fig. 2(a) (resp. Fig. 2(b)), results are provided for protection against single link failures (resp. single link or single server node failures). We show the total number of wavelengths summed over all links, where each data point represents an overage over 5 traffic instances, for a given number of job requests. We observe, as intuitively expected (cf. Model I has more flexibility by possibly using

3

multiple different backup paths for a given source demand), that Model I requires lower network capacity than Model II in the MILP approach, under both protection schemes. The MILP approach takes less bandwidth than the classical, i.e., 12% in Model I and 5% in Model II. Although Model II solution computing times are 66% higher than for Model I, it still offers a reasonable solution time for the off-line algorithms to be used in planning or advance reservation systems. We note that the MILP approach is more efficient than the classical one for both models, both in terms of computing times and in terms of solution accuracy.

Looking at the cost of link+server protection, which also covers failures of a complete server node, we find (as intuitively expected) that this requires more network resources. Yet, the additional network capacity is reasonably limited: the link+server protection scheme requires 10% (resp. 6%) more bandwidth units than only link protection for Model I (resp. Model II).

We also assessed how the different models and solution approaches differ for the selection of the locations of the server nodes. We observed that in more than 98% of the studied cases, the server nodes are selected among the same five nodes. In addition, even though Model I allows more than one backup path (Model II does not), in practice there are very few cases where Model I selects more than one backup path.

The same experiments were conducted with the non-uniform traffic instances, and the results are shown in Fig. 3. The qualitative conclusions as drawn from the uniform demands still apply in terms of relative performance of the various models (I vs II) and solution approaches (Classical vs MILP).
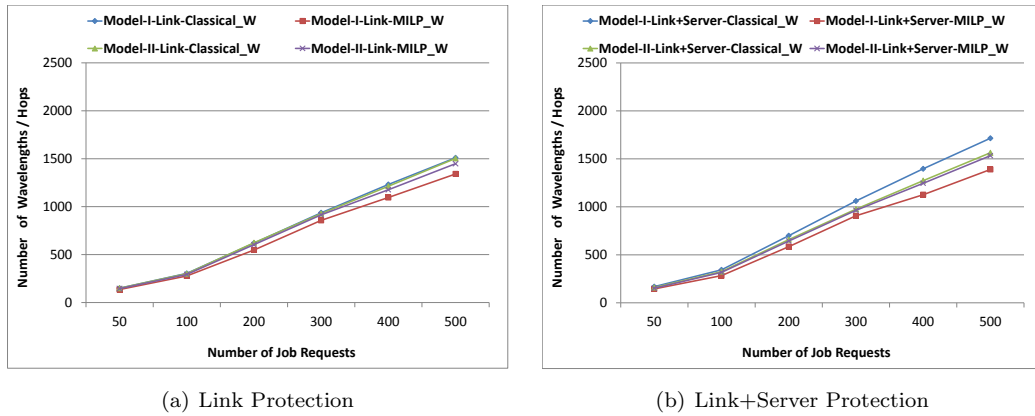


| (a) Link Protection | (b) Link+Server Protection |

Fig. 3: Non-uniform traffic: Model I vs. Model II and Classical vs. MILP solution approach.

## REFERENCES

[1] D. Gottfrid, "Self-service, prorated supercomputing fun." http://open.blogs.nytimes.com/2007/11/01/self-service-prorated-super-computing-fun/, 1 Nov. 2007.

[2] C. Develder, M. De Leenheer, B. Dhoedt, M. Pickavet, D. Colle, F. De Turck, and P. Demeester, "Optical networks for grid and cloud computing applications," *Proceedings of the IEEE*, vol. 100, pp. 1149–1167, May 2012.

[3] B. Jaumard, C. Meyer, and B. Thiongane, "Comparison of ILP formulations for the RWA problem," *Optical Switching and Networking*, vol. 4, pp. 157–172, November 2007.

[4] B. Jaumard, C. Meyer, and B. Thiongane, "On column generation formulations for the RWA problem," *Discrete Applied Mathematics*, vol. 157, pp. 1291–1308, March 2009.

[5] J. Buysse, M. De Leenheer, C. Develder, and B. Dhoedt, "Exploiting relocation to reduce network dimensions of resilient optical grids," in *Proceedings of IEEE/VDE Workshop on Design of Reliable Communication Networks - DRCN*, pp. 100–106, October 2009.

[6] C. Develder, B. Mukherjee, B. Dhoedt, and P. Demeester, "On dimensioning optical grids and the impact of scheduling," *Photonic Network Communications*, vol. 17, pp. 255–265, June 2009.

[7] J. Reese, "Solution methods for the *p*-median problem: An annotated bibliography," *Networks*, vol. 48, no. 3, pp. 125–142, 2006.

[8] S. H. Owen and M. S. Daskin, "Strategic facility location: A review," *European Journal of Operational Research*, vol. 111, no. 3, pp. 423 – 447, 1998.

[9] A. Shaikh, J. Buysse, B. Jaumard, and C. Develder, "Anycast routing for survivable optical grids: Scalable solution methods and the impact of relocation," *Journal of Optical Communications and Networking*, vol. 3, pp. 767–779, 2011.