Dimensioning Optical Clouds with Shared-Path Shared-Computing (SPSC) Protection

(Invited Paper)

Carlos Natalino*, Paolo Monti‡, Luis França†, Marija Furdek‡, Lena Wosinska‡, Carlos R. Francês* and João W. Costa*
*Laboratory of Planning of High Performance Networks (LPRAD), Federal University of Pará, Belém, Pará, Brazil
Telephone: +55 (91) 3201-8112 Email: {cns,rfrances,jweyl}@ufpa.br
‡Optical Network Laboratory (ONLab), KTH Royal Institute of Technology, Stockholm, Sweden
Email: {pmonti,marifur,wosinska}@kth.se
†National Institute for Space Research, São José dos Campos, São Paulo, Brazil
Email: lfamorim@gmail.com

***Abstract*—Service relocation represents a promising strategy to provide flexible and resource efficient resiliency from link failures in the optical cloud environment. However, when a failure affects a node hosting a datacenter (DC), service relocation from the affected DC is not possible. One alternative to protect against DC failures relies on using design strategies that duplicate the IT (i.e., storage and processing) resources in a backup DC at the expense of increasing resource overbuild (i.e., cost) of the network. This work proposes a dimensioning strategy based on the shared-path shared-computing (SPSC) concept able to protect against any single link, server, or DC failure scenario with minimal resource overbuild for the network and IT infrastructures. SPSC is based on the intuition that only storage units need complete replication in backup DC, while processing units can be instantiated only after the occurrence of a failure, leaving the design strategy some leeway to minimize their number. As result, the proposed SPSC design shows a considerable reduction in the amount of backup resources when compared to the dedicated protection strategies.**

I. Introduction

Network resilience is very crucial in optical networks, since a huge amount of data can be affected by a single failure (i.e., node and/or fiber link failure). Therefore, resiliency needs to be considered in network design and provisioning strategies. Resilience becomes an even more critical in the optical cloud scenario [1] where failures may affect both network and/or IT resources (i.e., storage and processing) [2]. For this reason the survivability of optical could services has been the objective of extensive studies.

Some of the resiliency strategies available in the literature take advantage of the anycast nature of cloud services and allow for the use of backup datacenter (DC) nodes that are different from the primary ones. This specific concept is known as service relocation and it allows for greater flexibility in the choice of the backup path and for better utilization of backup resources with respect to conventional (i.e., with no relocation capabilities) protection strategies. In [3] an example of a resilient scheme for optical cloud services that is based on the service relocation concept is proposed. Service relocation can also be used in conjunction with restoration strategies to improve the network performance in terms of average availability and restorability of the cloud services [2] [4].

Aside from the described benefits, service relocation has also a number of drawbacks. First, service relocation requires live migration of virtual machines (VMs). A VM is defined in terms of a specific number of storage units (SUs) and processing units (PUs). During a VM migration, new processing units need to be instantiated in the destination DC and all storage units used by a cloud service need to be transferred from the source DC to the destination DC. This latter operation has the biggest impact on the cloud service downtime. It depends on the size and on the number of storage units that need to be transferred as well as on the capacity of the fiber links used to connect the source and the destination DCs [5]. Furthermore, although service relocation has shown its effectiveness against link failure scenarios, it cannot be used to recover cloud services when storage units or entire source DC node become unreachable of failed (i.e., when a storage unit is malfunctioning, or when the optical node to which a DC is connected is down, or when the DC is destroyed as a consequence of a disastrous event) [6]. The main reason is that in such scenario the VM migration process cannot be performed, because the data to be migrated cannot be accessed.

Since restoration-based strategies are not effective in recovering disrupted cloud services when storage units or entire DC nodes are down, the only way to ensure service survivability is to employ protection strategies that duplicate both network and IT resources before any failures occur, i.e., strategies with dedicated protection (DP). The duplication of content in different DC nodes (belonging to different disaster zones) has been studied in the literature to solve the content placement problem in content delivery networks when node (or disaster) resiliency was required [7]. In such a scenario the challenge is to keep multiple copies of the same content synchronized among different DC nodes. This problem can be solved by using the continual migration approach [8].

However, in the case of optical clouds not only content resources (i.e., storage), but also the processing resources (i.e., servers) need to be duplicated. This means that using a DP

978-1-4799-9871-5/15/$31.00 ©2015 IEEE

based approach may significantly increase the cost. A possible way to reduce the resource overbuild is to allow for sharing of backup network and IT resources among optical cloud services. In terms of network resources, shared path protection (SPP) can be used to allow cloud services that have disjoint primary paths to share the same backup network resources. In terms of IT resources, a further distinction between storage and processing units is needed. Backup storage units can not be shared among different cloud services since each one of them comes with its own content that needs to be duplicated [7]. On the other hand, it is possible to share the backup processing units among cloud services that run on different primary DC nodes. This is because the backup servers need to be instantiated only after the primary VM is failed [8] [3].

The above considerations allow for the definition of hybrid optical cloud protection strategies where only storage resources are always duplicated on backup DC nodes while network and processing units used for protection purposes can be shared among cloud services whose primary paths are disjoint and their VMs run on different primary DC nodes. Such strategies have the potential to guarantee complete recovery against DC node failures at a lower cost (in terms of resource overbuild) than the ones offered by DP-based strategies.

This paper presents an optical cloud design strategy called Shared-Path Shared-Computing (SPSC) able to provide 100% survivability against single link or DC failures. The proposed strategy is formulated as an Integer Linear Program (ILP) with the objective of dimensioning both the transport and the IT resources such that the resulting resilient infrastructure has a minimum network and DC resource overbuild. Compared to a conventional DP-based approach, SPSC shows an average 43% reduction of the number of backup network resources and a 48% reduction of the number of backup processing units, while guaranteeing the same level of resiliency as DP.

II. The Shared-Path Shared-Computing (SPSC) Approach

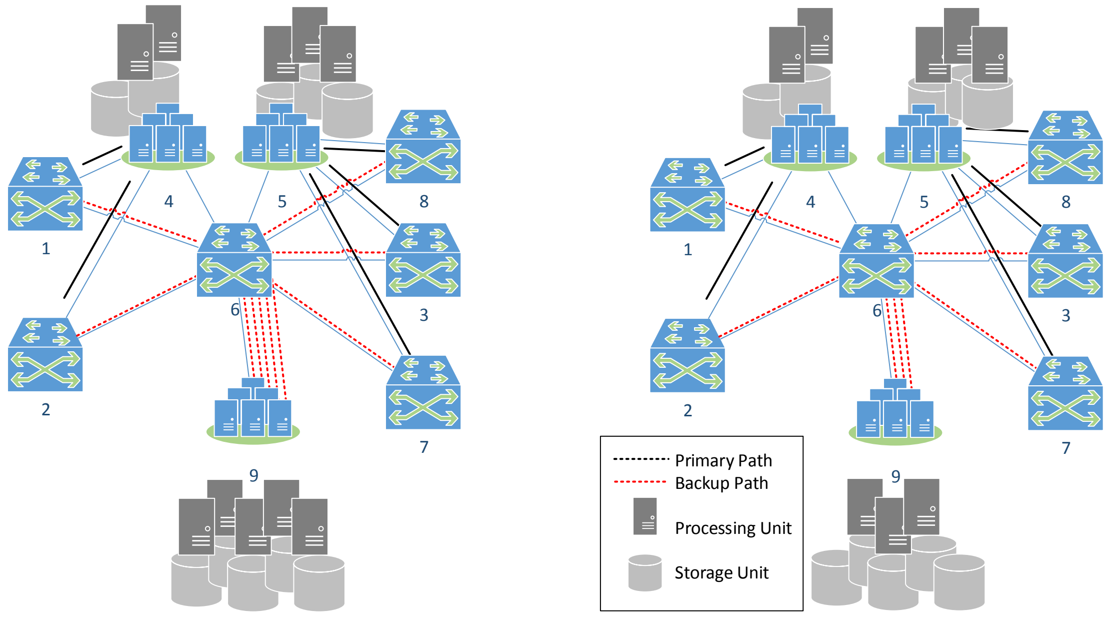
The problem addressed in this paper aims at dimensioning an optical could network such that each given cloud service is protected against a single link, server or DC failure occurrence. The inputs of the problem are the following: *(a)* optical network topology comprising a set of optical nodes (i.e., optical cross connects - OXCs) and fiber links interconnecting the nodes; *(b)* a set of DCs hosted by a subset of network nodes; and *(c)* a set of cloud services to be provisioned (i.e., the demand matrix), where each service is specified by the source optical node and the required amount of IT resources (i.e., storage and processing units). The given set of cloud services is expressed in terms of steady-state traffic estimates (i.e., long-time averages). For simplicity and without loss of generality it is assumed that a DC resides in the vicinity of the OXC to which it is connected using a fiber link with unlimited wavelength resources. Also, the impact of the communication between optical source nodes and service end users, as well as the inter-DC communication required to implement the continual migration between primary and backup storage units is neglected.

The objective of the problem is to accommodate all cloud services presented in the traffic matrix while minimizing the amount of necessary backup network resources (i.e., wavelength), as well as storage and processing resources. To explain the idea behind the SPSC strategy, an illustrative example is shown in Fig. 1 is used. The figure presents an optical cloud network comprising 9 nodes and 13 bidirectional fiber links. As it can be seen three DCs are connected to the network. These DCs are hosted by nodes 4, 5 and 9, respectively. Let us assume a traffic matrix comprising 5 cloud service requests originating at nodes 1, 2, 3, 7 and 8, respectively. These 5 cloud services require one network resource unit (i.e., one wavelength), as well as one storage and one processing unit each and they need to be 100% survivable against any single link or DC failure.

Depending on the design strategy, the result of the network dimensioning process can be different. For example, if a DPbased strategy is used (e.g., such as the one adapted in [9]), the 5 cloud services can be allocated as follows. The DC at node 4 can host the cloud services originating at nodes 1 and 2, while the DC at node 5 hosts the remaining ones. The DC at node 9 is used to allocate the IT resources needed for protection purposes. In this way, upon the failure of the DC at node 4 or the DC at node 5, the affected cloud services can be rerouted via node 6 to the DC at node 9. The overall dimensioning result would then be the following. The DC at node 4 needs to be equipped with 2 storage and 2 processing units, the DC at node 5 with 3 storage and 3 processing units, while the DC at node 9 requires a total of 5 storage and 5 processing units. In terms of network resources, each source node will be connected to its primary DC node using a direct link (i.e., 1-4, 2-4, 3-5, 7-5, 8-5), while each protection path will go through node 6 (i.e., 1-6-9, 2-6-9, 3-6-9, 7-6-9, 8-6-9), requiring a total of 15 network units in the optical cloud. Fig. 1a illustrates this solution, where the black (continuous) lines represent the primary paths and the red (dashed) lines represent the backup paths.

Looking at the results of the DP-based dimensioning strategy in Fig. 1a, and remembering that we are operating under the single DC or link failure assumption, two observations can be made. First, it is clear that it is not strictly necessary to duplicate at the DC in node 9 all server resources used in the DCs at nodes 4 and 5. This is because processing units need to be allocated only after the occurrence of a failure. Second, some of the backup network resources on link 6-9 can be shared among cloud services whose primary paths and primary DC nodes are disjoint. This is exactly the intuition behind the Shared-Path Shared-Computing (SPSC) approach. If SPSC were to be applied to dimension the optical cloud network in the example, the solution would require 10 storage units, 8 processing units, and 13 network units (Fig. 1b). The results of the dimensioning exercise using both strategies are summarized in Table I.

A dimensioning strategy based on the SPSC concept is



(a) (b)

Fig. 1. Example of a DP-based vs. the SPSC strategy: (a) DP-based dimensioning results and, (b) SPSC dimensioning result

TABLE I
THE NUMBER OF PRIMARY (P) AND BACKUP (B) NETWORK (NU), PROCESSING (PU) AND STORAGE UNITS (SU) REQUIRED BY EACH DIMENSIONING STRATEGY

Strategy	NUs (P+B)	PUs (P+B)	SUs (P+B)
DP	5+10	5+5	5+5
SPSC	5+8	5+3	5+5

also advantageous in terms of cost. Aside from the obvious savings in terms of wavelength resources, an important cost saving stems from a reduced number of backup processing units. To illustrate this benefit, let us compute the monthly cost of IT resources for a standard cloud service requiring 10 GB of storage and 1 processor [5]. According to Amazon Web Services (AWS) the average monthly price for processing and storage is \$30.24 and \$1.25, respectively. Applying these numbers on the simple example from Fig. 1, the dimensioning based on the DP strategy would have a monthly cost of the backup IT resources equal to \$157.45, while the one based on SPSC amounts to only \$96.97 per month, offering a 38% cost reduction.

Given all the potential benefits brought by the SPSC concept, the next section presents an ILP that can be used for dimensioning optical clouds based on the shared-path shared-computing approach.

III. Optical Clouds Dimensioning with Shared Path Shared Computing

In SPSC, each service request is assigned to one primary and one (different) backup DC, as well as one primary and one (disjoint) backup path. The model assumes that the IT resources required for each service encompass one processing and one storage unit, which must be available on both the primary and the backup DC. Each service also requires one network unit (e.g., a lightpath with a wavelength capacity) from its source node to both the primary and the backup DC. The formulation can be easily modified to consider different (and independent) number of required resources for each cloud service. The model assumes that all nodes offer full wavelength conversion capabilities, i.e., the wavelength continuity constraint is not enforced. Formally we have:

- $N = N_{DC} \cup N_{src}$: set of all nodes on the topology, comprising source nodes (N_{src}) and datacenter nodes (N_{DC});
- E: set of unidirectional links on the topology, each one ($(i,j) \in E$) with $i \in N$ as the source node and $j \in N$ as the destination node of the link;
- R: set of service requests (s,c), where s is the requesting node and c represents the service that needs to be deployed, with $s \in N_{src}$;

Variables:

- $x_{(i,j)}^{(s,c)} \in \{0,1\}$ is equal to 1 when request (s,c) uses link (i,j) on its primary path;
- $y_{(i,j)}^{(s,c)} \in \{0,1\}$ is equal to 1 when link (i,j) is used on the backup path for request (s,c);
- $p_d^{(s,c)} \in \{0,1\}$ is equal to 1 when DC $d \in N_{DC}$ is used as primary DC for request (s,c);
- $b_d^{(s,c)} \in \{0,1\}$ is equal to 1 when DC $d \in N_{DC}$ is used as backup DC for request (s,c);
- $z_{(i,j),(k,l),a,b}^{(s,c)} \in \{0,1\}$ is equal to 1 when link (i,j) is used on the backup path for the request (s,c) which uses link (k,l) on its primary path, and the request (s,c) has a as primary DC and b as backup DC, with a and $b \in N_{DC}$;

- $B_{(i,j)}$ is an integer variable which accounts the number of wavelengths needed for the backup paths on the link $(i,j) \in E$;
- sto_d is an integer variable which accounts the number of storage units needed on DC $d \in N_{DC}$;
- cpu_d is an integer variable which accounts the number of CPUs needed on DC $d \in N_{DC}$;

Objective function:

$$min \left(\sum_{(i,j)} \sum_{(s,c)} x_{(i,j)}^{(s,c)} + \sum_{(i,j)} B_{(i,j)} + \alpha \sum_{(i,d) \in E | d \in N_{DC}} B_{(i,d)} \right) \quad (1)$$

Subject to:

$$\sum_{j:(i,j)\in E} x_{(i,j)}^{(s,c)} - \sum_{j:(j,i)\in E} x_{(j,i)}^{(s,c)} = \begin{cases} 1, \text{ if } i = s, i \in N_{src} \\ -p_d^{(s,c)}, \text{ if } i \in N_{DC} \\ 0, \text{ otherwise} \end{cases}, \forall (s,c) \in R, \forall i \in N \quad (2)$$

$$\sum_{j:(i,j)\in E} y_{(i,j)}^{(s,c)} - \sum_{j:(j,i)\in E} y_{(j,i)}^{(s,c)} = \begin{cases} 1, \text{ if } i = s, i \in N_{src} \\ -b_d^{(s,c)}, \text{ if } i \in N_{DC} \\ 0, \text{ otherwise} \end{cases}, \forall (s,c) \in R, \forall i \in N \quad (3)$$

$$x_{(i,j)}^{(s,c)} + y_{(i,j)}^{(s,c)} \leq 1, \forall (i,j) \in E, \forall (s,c) \in R \quad (4)$$

$$\sum_{d \in N_{DC}} p_d^{(s,c)} = 1, \forall (s,c) \in R \quad (5)$$

$$\sum_{d \in N_{DC}} b_d^{(s,c)} = 1, \forall (s,c) \in R \quad (6)$$

$$p_d^{(s,c)} + b_d^{(s,c)} \leq 1, \forall d \in N_{DC}, \forall (s,c) \in R \quad (7)$$

$$z_{(i,j),(k,l),a,b}^{(s,c)} + 3 \geq x_{(k,l)}^{(s,c)} + y_{(i,j)}^{(s,c)} + p_a^{(s,c)} + b_b^{(s,c)} \quad , \quad \forall (i,j),(k,l) \in E, (k,l) \neq (i,j), \forall a,b \in N_{DC}, b \neq a \quad (8)$$

$$B_{(i,j)} \geq \sum_{(k,l)\neq(i,j)\in E} \sum_{(s,c)\in R} z_{(i,j),(k,l),a,b}^{(s,c)} \quad , \quad \forall a,b \in N_{DC}, b \neq a, \forall (i,j) \in E \quad (9)$$

The objective of the SPSC formulation given in Eq. (1) is to minimize the total wavelength usage in the network and the DC resource usage needed for protection. The first term in the objective function accounts for the total number of wavelengths used on the primary paths for all cloud services; the second term accounts for the number of wavelengths used to protect the primary paths under shared path protection; and the third term accounts for the number of backup PUs needed in the datacenters. In the last term, is used to establish the cost relation between one network unit (first and second terms) and one PU. For instance, if is set to 0.1, it means that the cost of one PU equals 10% of a network unit value.

The constraints in Eqs (2) and (3) model the flow conservation constraint under anycast routing. Equation (4) ensures that the primary and backup paths for each service are disjoint. Equations (5) and (6) set one primary and one backup datacenter for each service, respectively, which must be different according to Eq. (7).

Constraints (8) and (9) represent the main difference between the proposed SPSC formulation and the existing approaches found in the literature. Equation (8) identifies whether two services may share both the network and the DC protection resources, which is only possible if their corresponding working resources are not susceptible to a simultaneous failure. Therefore, Eq. (8) indicates whether link (i,j) and backup DC b are used to protect service (s,c) from failures of link (k,l) and DC a. Equation (9) then uses this information to count the number of network resources required on each link for service protection.

Once a solution is found, Eqs (10) and (11) can be used to evaluate the processing and storage resource usage. The number of PUs required in each DC is found by summarizing the (deterministic) number of primary services assigned to that DC and the (minimized) number of backup network resources connected to this DC, according to Eq. (10). Accordingly, Eq. (11) accounts for the number of storage units needed in each DC.

$$cpu_d = \sum_{(s,c)\in R} p_d^{(s,c)} + \sum_{i:(i,d)\in E} B_{(i,d)}, \forall d \in N_{DC} \quad (10)$$

$$sto_d = \sum_{(s,c)\in R} \left(p_d^{(s,c)} + b_d^{(s,c)} \right), \forall d \in N_{DC} \quad (11)$$

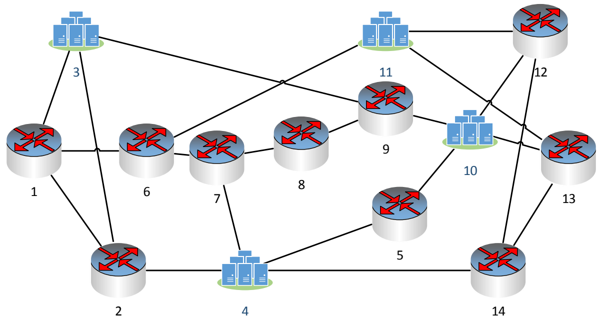


Fig. 2. Reference Network Topology

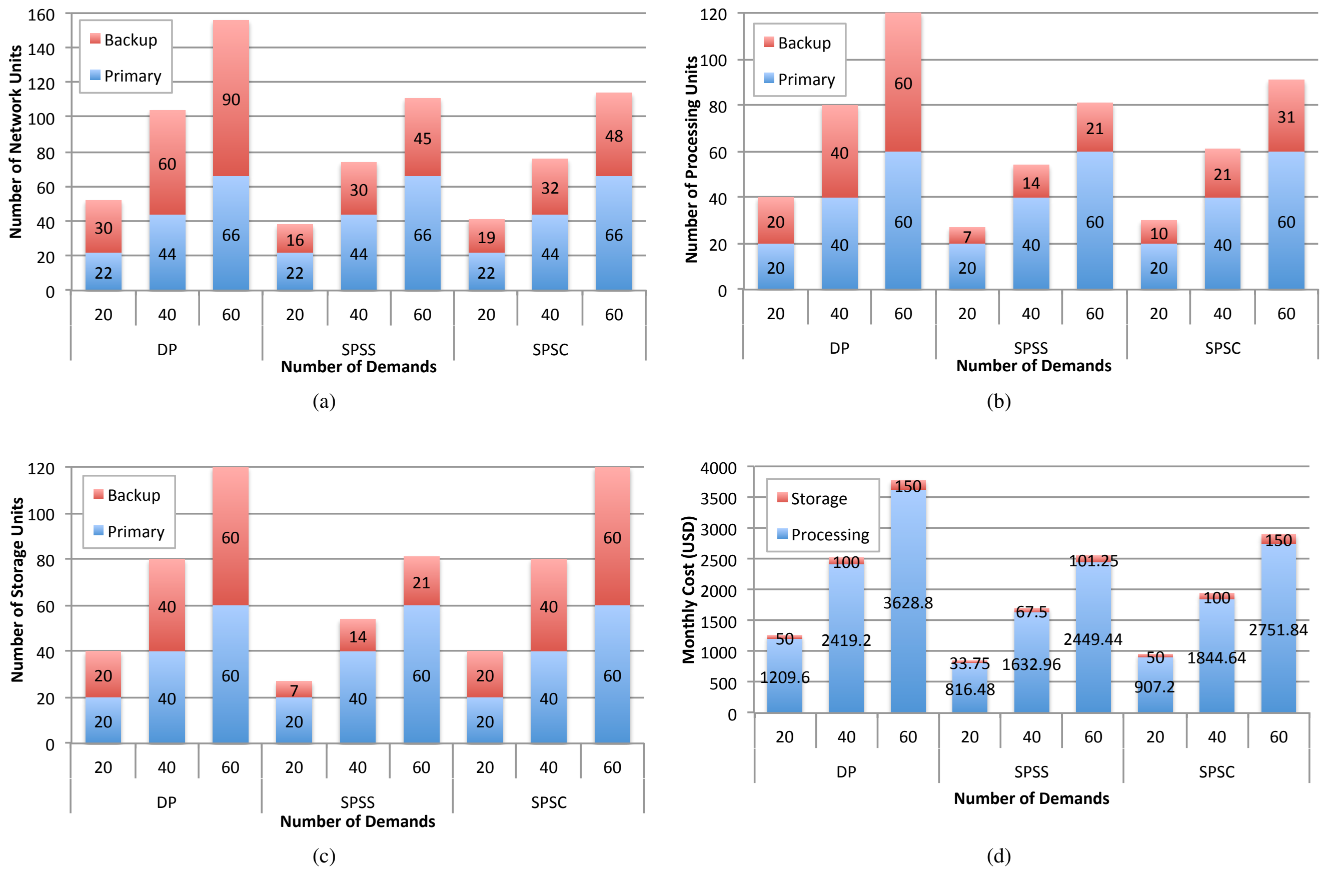


Fig. 3. Results for the NSF topology in function of number of cloud services: (a) number of network units; (b) number of processing units; (c) number of storage units and; (d) monthly cost.

IV. Numerical Results

This section presents a performance assessment study of the proposed dimensioning strategy based on the SPSC concept. The evaluation is done in terms of the number of network, processing, and storage units required to guarantee 100% survivability against any single link or DC failure for all cloud services in the traffic matrix. Two strategies are chosen as a benchmark. The first one is a DP-based dimensioning strategy (adapted from [9]) where cloud services have dedicated network and IT resources. This strategy is chosen because it guarantees 100% survivability in the considered failure scenario. The second dimensioning strategy is the shared path and shared server (SPSS) strategy prosed in [3]. This strategy offers fiber link and server failure resilience, but does not protect against storage or DC failures.

The performance of the described approaches is evaluated on the NSF topology shown in Fig. 2. It has a total of 14 nodes ($|N| = 14$) interconnected by 21 bidirectional links ($|E| = 42$). As it can be seen from the figure, 4 nodes host a DC, i.e., we have a total of 4 DCs ($|N_{DC}| = 4$) plus 10 source nodes ($|N_{src}| = 10$). The 4 DCs are placed at the nodes with the highest value of nodal degree.

Three traffic matrices where considered in the experiments, i.e., with 2, 4, and 6 cloud services originating at each of the 10 source nodes, respectively. This translates to having $|R| = 20$, $|R| = 40$ and $|R| = 60$, respectively. Each cloud service is assumed to require one network, one processing, and one storage unit. In all experiments α is set to 0.1.

The total number of required network units for each strategy is shown in Fig. 3a. All three strategies require the same amount of network units for the primary paths, while for the backup paths the strategies using shared protection (SPSS and SPSC) reduce their usage by 46%, compared to DP. The proposed SPSC strategy needs, on average, 4% more network units than SPSS. This slightly higher resource consumption can be justified by the greater resiliency level provided by SPSC (i.e., protection against DC failure). However, compared to DP, SPSC offers the same resiliency level while reducing network resource usage by up to 46%.

Figure 3b shows the number of processing units (PUs) needed for each strategy. As expected, in DP the amount of backup PUs is equal to the number of cloud services to be protected. In the case of the strategies with sharing capability, SPSC needs on average 48% less backup PUs than DP, and a slightly higher number of backup PUs than SPSS. The higher number of backup PUs needed by SPSC can be explained by the fact that this strategy is able to provide 100% survivability against DC failure.

The required number of storage units (SUs) is shown in Fig. 3c. As expected, DP and SPSC require the same amount, as there is no sharing of storage resources for these strategies. Notwithstanding, SPSS needs 65% less backup SUs as it allows for sharing of storage resources, but it does not provide protection against DC failures.

Figure 3d shows the monthly cost for the IT resources based on the cost values presented at the end of Section II. The cost increase incurred by having dedicated storage backup has a slight impact on the total monthly IT cost, given that SUs cost only a fraction of PUs. Hence, the overall extra cost of SPSC compared to SPSS might be acceptable considering the increase resiliency level provided.

V. Conclusion

This paper presented a Shared-Path Shared-Computing (SPSC) approach that can be used to dimension optical clouds to guarantee 100% survivability against single link, server, or DC failure scenario. In the SPSC concept, cloud services that have disjoint primary paths and run on different DC nodes can share backup network and processing unit resources.

The SPSC problem was modeled as an ILP and it was found that an optical cloud infrastructure dimensioned with SPSC offers the same protection level as DP, with a 50% reduction of the number of backup network units and a 48% reduction of the number of backup processing units.

Acknowledgment

This work has been supported by FAPESPA (The Amazon Research Foundation) under Grant no. 005/2011, by CNPq-CISB-SAAB under Grant no. 55/2012 and by ICT DISCUS (FP7 program under Grant agreement no. 318137).

References

[1] 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, no. 5, pp. 1149–1167, May 2012.

[2] J. Ahmed, P. Monti, L. Wosinska, and S. Spadaro, "Enhancing Restoration Performance Using Service Relocation in PCE-based Resilient Optical Clouds," in *Optical Fiber Communication Conference*. Washington, D.C.: OSA, 2014, p. Th3B.3.

[3] C. Develder, J. Buysse, B. Dhoedt, and B. Jaumard, "Joint Dimensioning of Server and Network Infrastructure for Resilient Optical Grids/Clouds," *IEEE/ACM Transactions on Networking*, vol. PP, no. 99, pp. 1–16, 2013.

[4] C. Natalino, J. Ahmed, P. Monti, L. Wosinska, and R. Frances, "A relocation-based heuristic for restoring optical cloud services," in *Optical Communications and Networks (ICOCN), 2014 13th International Conference on*, Nov 2014, pp. 1–4.

[5] T. Wood, K. K. Ramakrishnan, P. Shenoy, J. Van der Merwe, J. Hwang, G. Liu, and L. Chaufournier, "CloudNet: Dynamic Pooling of Cloud Resources by Live WAN Migration of Virtual Machines," *IEEE/ACM Transactions on Networking*, vol. PP, no. 99, pp. 1–1, 2014.

[6] B. Mukherjee, M. Habib, and F. Dikbiyik, "Network adaptability from disaster disruptions and cascading failures," *IEEE Communications Magazine*, vol. 52, no. 5, pp. 230–238, May 2014. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6815917

[7] M. F. Habib, M. Tornatore, M. De Leenheer, F. Dikbiyik, and B. Mukherjee, "Design of Disaster-Resilient Optical Datacenter Networks," *Journal of Lightwave Technology*, vol. 30, no. 16, pp. 2563–2573, Aug. 2012. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6208805

[8] W. Cui, D. Ma, T. Wo, and Q. Li, "Enhancing Reliability for Virtual Machines via Continual Migration," in *2009 15th International Conference on Parallel and Distributed Systems*. IEEE, 2009, pp. 937–942.

[9] J. Buysse, M. De Leenheer, B. Dhoedt, and C. Develder, "On the impact of relocation on network dimensions in resilient optical Grids," in *2010 14th Conference on Optical Network Design and Modeling (ONDM)*. IEEE, 2010, pp. 1–6.