Dynamic Anycast in Inter-Datacenter Networks over Elastic Optical Infrastructure

Liang Zhang and Zuqing Zhu[†]
School of Information Science and Technology
University of Science and Technology of China, Hefei, China
[†]Email: {zqzhu}@ieee.org

***Abstract*—Since the optical orthogonal frequency-division multiplexing (O-OFDM) technology can facilitate agile spectrum management, the elastic optical infrastructure based on it has become a promising candidate for the physical-layer of inter-datacenter networks. In this paper, we investigate dynamic anycast in such networks. Firstly, we design three metrics to consider computing and bandwidth resources jointly, and propose anycast algorithms with single-path routing based on them. Then, we propose an anycast algorithm with multipath routing to further improve the network performance. Our simulation results indicate that the anycast algorithm with multipath routing can fully utilize the bandwidth resources in the optical infrastructure, make computing resources become the bottleneck, and reduce the bandwidth blocking probability of anycast requests effectively.**

***Index Terms*—Anycast, Cloud Computing, Elastic Optical Networks (EONs), Multipath Routing.**

I. INTRODUCTION

Nowadays, with the boosting up of cloud computing [1], datacenter networks have attracted intensive interests from both academia and industry. Meanwhile, the bandwidth-intensive applications, such as high-definition video services and e-Science [2], make datacenter networks exhibit the characteristics of huge throughput and large traffic burstiness. Recent advances on the optical orthogonal frequency-division multiplexing (O-OFDM) technology have demonstrated optical super-channels with more than 1 Tb/s transmission capacity [3] and flexible bandwidth allocation with a granularity at 12.5 GHz or less [4]. Therefore, it can achieve agile spectrum management and hence facilitate seamless integration of physical transmission and upper-layer applications. To this end, the elastic optical infrastructure based on O-OFDM has become a promising candidate for the physical-layer of datacenter networks, especially for inter-datacenter networks [5].

One attractive attribute of cloud computing is that demands can be served with statistical multiplexing and the servers in datacenters can be utilized more intelligently [6]. This means that by leveraging anycast [7], the optical spectra on fiber links can be utilized more wisely with the idea that the location(s) of the datacenter(s) that a job is executed can be transparent to the customer. More specifically, anycast refers to the communication scheme that the destination(s) of a connection can be flexible, as long as the service-level agreements (e.g., bandwidth and computing demands) are satisfied.

Previous studies have investigated anycast in wavelength-division multiplexing (WDM) networks. In [8], Din studied offline routing and wavelength assignment (RWA) for anycast demands. They assumed that all demands were known *a prior*, and solved the RWA problem in three phases: 1) destination selection, 2) path routing, and 3) wavelength assignment, with the objective of minimizing the total usage of wavelengths. The online RWA for dynamic anycast demands was investigated in [9], where the authors proposed to change the destinations of anycast demands according to traffic load. She *et al.* considered survivable traffic grooming for anycast in WDM networks in [10], the shared backup path protection for anycast flows was studied in [11], and the provisioning of advance reservation anycast demands in WDM grids was addressed in [12]. However, to the best of our knowledge, anycast in flexible-grid O-OFDM networks is still under-explored, even though compared with WDM, O-OFDM can potentially provide better physical-layer support to inter-datacenter networks [13]. Moreover, due to the fine bandwidth allocation granularity, O-OFDM can bring some unique benefits to optical networks. For example, multipath routing can be realized in a more spectrum-efficient way [14].

In this paper, we investigate dynamic anycast in inter-datacenter networks over elastic optical infrastructure based on O-OFDM. To consider computing and bandwidth resources jointly during request serving, we first design three metrics and propose anycast algorithms with single-path routing based on them. Then, in order to further improve the network performance, we propose an anycast algorithm with multipath routing to fully utilize the bandwidth resources in the elastic optical infrastructure. The proposed algorithms are evaluated with numerical simulations with the NSFNET topology and the simulation results indicate that the algorithms can effectively reduce the bandwidth blocking probability.

The rest of this paper is organized as follows. Section II formulates the problem of anycast in inter-datacenter networks over elastic optical infrastructure. The proposed algorithms are discussed in Section III, and we evaluate their performance in Section IV. Finally, Section V summarizes this paper.

II. PROBLEM FORMULATION

In this section, we formulate the problem of anycast in inter-datacenter networks over elastic optical infrastructure, including the design constraints and objective.

978-1-4799-2358-8/14/$31.00 ©2014 IEEE

A. Design Constraints

We use a directed graph $G(V, E)$ to represent the physical topology of the inter-datacenter network, where V and E denote the sets of nodes and fiber links, respectively. Based on the working principle of O-OFDM, we assume that there are B subcarrier frequency slots (FS') to allocate on each fiber link $e \in E$. Within the node set V, we have a subset of nodes that are connected to a datacenter locally, and it is denoted as V_{DC} ($V_{DC} \subset V$). For a node $v \in V_{DC}$, the available computing capacity of its datacenter is C_v. We define an anycast request as $R(s, b, c)$, where $s \in V \setminus V_{DC}$ is the source node, b is the bandwidth demand, and c is the computing demand. The bandwidth demand b is in terms of Gb/s, the computing demand c is in number of servers, and we assume there is a linear relation between b and c,

$$c = \alpha \cdot b, \tag{1}$$

where α is a constant coefficient. For each node pair u-v, where $u \in V \setminus V_{DC}$ and $v \in V_{DC}$, we calculate K shortest paths and denote them as $\{p_{u,v}^{(k)},\ k = 1, \cdots, K\}$. During the online provisioning, we use function $BW(p_{u,v}^{(k)})$ to get the available bandwidth of $p_{u,v}^{(k)}$ in number of FS', and use function $hops(p_{u,v}^{(k)})$ to obtain the hop-count of $p_{u,v}^{(k)}$.

In order to serve the anycast request $R(s, b, c)$, we need to select datacenter node(s) as the destination(s), determine the amounts of computing capacity to allocate, and perform routing and spectrum assignment (RSA) to set up lightpath(s). We define c_v as the computing capacity allocated on $v \in V_{DC}$ for the request, and $c_v = 0$ if v is not selected as a destination. We also define $b_{s,v}^{(k)}$ as the bandwidth allocated on $p_{s,v}^{(k)}$ for the request, and $b_{s,v}^{(k)} = 0$ if $p_{s,v}^{(k)}$ is not used.

Therefore, in addition to the spectrum non-overlapping, continuity and contiguous constraints [15] for normal RSA problems, we have the following design constraints,

- Path capacity constraint:

$$b_{s,v}^{(k)} \le BW(p_{s,v}^{(k)}), \quad \forall v \in V_{DC}, k \in [1, K]. \tag{2}$$

  Eq. (2) ensures that the spectrum resource allocated on each path does not exceed its available bandwidth.
- Path bandwidth allocation granularity constraint (for multipath routing only):

$$b_{s,v}^{(k)} \ge g, \quad \{v, k : b_{s,v}^{(k)} > 0\}. \tag{3}$$

  Eq. (3) ensures that the minimum number of FS' allocated on a path is not smaller than the granularity g, when the multipath scheme is used.
- Computing capacity constraint:

$$c_v \le C_v, \quad \forall v \in V_{DC}. \tag{4}$$

  Eq. (4) ensures that the computing resource allocated on each datacenter node does not exceed its available computing capacity.
- Bandwidth-computing relation constraint:

$$c_v = \sum_{k=1}^{K} \alpha \cdot b_{s,v}^{(k)}. \tag{5}$$

  Eq. (5) ensures that there is a linear relation between the bandwidth and computing resource allocations, as defined in Eq. (1).
- Service-level agreement constraint:

$$c = \sum_{v \in V_{DC}} c_v. \tag{6}$$

  Eq. (6) ensures that the required resources are allocated to the request.

B. Objective

In online provisioning, each dynamic anycast request associates with two time parameters, the arrival time and the holding period, since they are time-variant and can arrive and leave on-the-fly. If sufficient resources (both bandwidth and computing) cannot be allocated at an anycast request's arrival time, it is blocked. In this work, we aim to minimizing the bandwidth blocking probability (BBP),

$$Minimize \qquad p_b = \lim_{T \to \infty} \frac{N_b(T)}{N(T)}, \tag{7}$$

where $N_b(T)$ and $N(T)$ are the numbers of blocked and total requested FS' in time duration $[0, T]$.

III. PROPOSED ALGORITHMS

In this section, we explain the detailed procedures of the proposed algorithms for dynamic anycast in inter-datacenter networks over elastic optical infrastructure.

A. Dynamic Anycast with Single-Path Routing

We first consider anycast with single-path routing. In order to minimize BBP, we need to balance the bandwidth and computing resource utilizations in the network, and therefore design three metrics for a path candidate $p_{s,v}^{(k)}$, as follows

$$m_1(p_{s,v}^{(k)}) = \frac{BW(p_{s,v}^{(k)}) \cdot \sqrt{C_v}}{\sqrt{hops(p_{s,v}^{(k)})}}, \tag{8}$$

$$m_2(p_{s,v}^{(k)}) = \frac{BW(p_{s,v}^{(k)}) \cdot C_v}{\sqrt{hops(p_{s,v}^{(k)})}}, \tag{9}$$

$$m_3(p_{s,v}^{(k)}) = BW(p_{s,v}^{(k)}) \cdot C_v. \tag{10}$$

Algorithm 1 shows how to realize anycast with single-path routing by leveraging the assistance of the metrics defined in Eqs. (8-10). It is a two-phase algorithm. In the first phase (*Lines* 5-9), we collect all possible routing paths and select the one that has the largest metric ($m_1(\cdot)$ or $m_2(\cdot)$ or $m_3(\cdot)$) to carry the request. The actual allocation of the bandwidth and computing resources is performed in the second phase (*Lines* 10-16). Since the destination can be determined with the path, we try to allocate the required resources on the path and the destination. If no sufficient resources can be found, we mark the request as blocked, otherwise, it is provisioned successfully. In the rest of the paper, we refer to the above algorithms that use metrics $m_1(\cdot)$, $m_2(\cdot)$ and $m_3(\cdot)$ as Anycast-SPR-1, Anycast-SPR-2, and Anycast-SPR-3, respectively.

Algorithm 1 Anycast with Single-Path Routing

```
while the network is operational do
    get current network status;
    collect anycast request $R(s, b, c)$;
    release the resources of expired requests;
    build path set $P = \{p_{s,v}^{(k)}, \forall k, v \in V_{DC}\}$;
    for all routing paths in path set $P$ do
        calculate metric $m_1(\cdot)$ with Eq. (8), or $m_2(\cdot)$ with Eq. (9), or $m_3(\cdot)$ with Eq. (10);
    end for
    select the routing path whose metric is the largest;
    try to allocate $b$ FS' on the path;
    try to allocate $c$ servers on the path's destination;
    if both allocations are successful then
        update network status;
    else
        mark $R$ as blocked;
    end if
end while
```

Algorithm 2 Anycast with Multipath Routing

```
1:  while the network is operational do
2:      get current network status;
3:      collect anycast request $R(s, b, c)$;
4:      release the resources of expired requests;
5:      build path set $P = \{p_{s,v}^{(k)}, \forall k, v \in V_{DC}\}$;
6:      $n_p = 0$;
7:      while $b \geq g$ OR $n_p = 0$ do
8:          for all routing paths in path set $P$ do
9:              calculate metric $m_3(\cdot)$ with Eq. (10);
10:         end for
11:         select the routing path whose metric is the largest;
12:         find the largest block of available contiguous FS' in the path's spectrum;
13:         obtain the size of the block as $b_p$;
14:         if $b_p < g$ then
15:             break;
16:         else
17:             allocate $\min(b_p, b)$ FS' on the path;
18:             try to allocate $\alpha \cdot \min(b_p, b)$ servers on the path's destination;
19:             if server allocation is not successful then
20:                 break;
21:             else
22:                 $b = b - \min(b_p, b)$;
23:                 $n_p = n_p + 1$;
24:             end if
25:         end if
26:     end while
27:     if $b = 0$ then
28:         update network status;
29:     else
30:         mark $R$ as blocked;
31:     end if
32: end while
```

B. Dynamic Anycast with Multipath Routing

Since single-path routing may not be able to fully utilized the bandwidth resources in the network, we design an algorithm for anycast with multipath routing to reduce BBP further. *Algorithm* 2 shows the detailed procedures. It is a greedy algorithm that tries to allocate the largest block of contiguous FS' on a path in each loop (*Lines* 7-26). To avoid an anycast request from being split over too many paths, we define a path bandwidth allocation granularity and implement the constraint in Eq. (3). Basically, when a request is provisioned over more than one routing paths (i.e., $n_p > 1$), the minimum number of FS' to allocate on each path is g. From the viewpoint of network operation, g is the path-splitting granularity that the operator is willing to offer. In the rest of the paper, we refer to this algorithm of anycast with multipath routing as Anycast-MPR.

IV. PERFORMANCE EVALUATIONS

In this section, we discuss simulations to evaluate the performance of the proposed algorithms.

Fig. 1 shows the network topology, NSFNET, which we used in the simulations for performance evaluations. The topology consists of 14 nodes, and we have $V_{DC} = \{3, 5, 8, 10, 12\}$. In each datacenter node, we assume there are 960 servers and thus the initial value of C_v is 960. The bandwidth of each FS is 12.5 GHz and its capacity is 12.5 Gb/s. On each fiber link, there are 260 FS' to allocate initially (i.e., $B = 260$), when the network is empty. For each anycast request $R(s, b, c)$, the source node s is randomly selected from $V \setminus V_{DC} = \{1, 2, 4, 6, 7, 9, 11, 13, 14\}$, the bandwidth demand b is uniformly distributed within $[12.5, 200]$ Gb/s, and the computing demand is then calculated with Eq. (1) with $\alpha = 0.2$. Here, we determine α according to the assumption in [6]. The dynamic anycast requests arrive according to a Poisson process with an average arrival rate of λ requests per time unit, and the holding period of each request follows the negative exponential distribution with an average of $\frac{1}{\mu}$ time units. Hence, the traffic load can be quantified with $\frac{\lambda}{\mu}$ in Erlangs. Each request can be served with either single-path routing (i.e., Anycast-SPR-1(2 or 3)) or multipath routing (i.e., Anycast-MPR). We refer to the connection over a path as a lightpath, and for each lightpath, one additional FS is allocated as the guard-band to prevent crosstalk. Table I summarizes the parameters we use in the simulations.

Fig. 2 shows the comparisons on the simulation results on bandwidth blocking probability (BBP) from different anycast algorithms. For Anycast-MPR, we set g as 1 FS. Note that we also modify the three-phase algorithm in [8] (i.e., Anycast-3-Phase in Fig. 2) and implement it as the benchmark algorithm. The simulation results in Fig. 2 indicate that Anycast-3-Phase provides the highest BBP among all anycast algorithms. This is due to the fact that it only considers the computing resources for destination selection. Even though choosing the datacenter

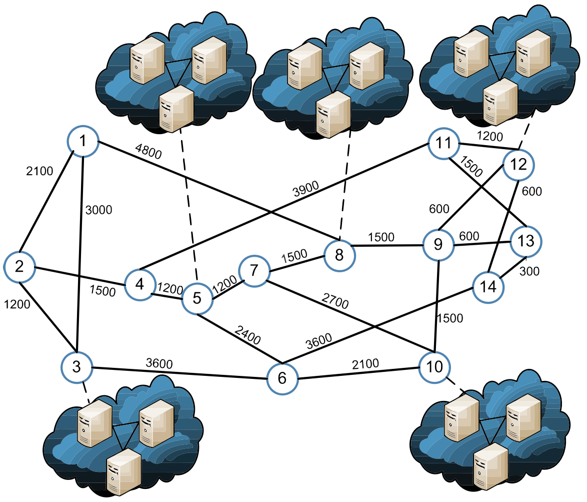

Fig. 1. NSFNET topology with datacenters.

TABLE I
SIMULATION PARAMETERS

Network topology	NSFNET
Average node degree	3.21
B, Fiber link capacity	260 FS'
Capacity of an FS	12.5 Gb/s
V_{DC}, Datacenter node set	{3,5,8,10,12}
C_v, Computing capacity of a datacenter	960 servers
Bandwidth demand range of a request	$[12.5, 200]$ Gb/s
Guard-band for a lightpath	1 FS
α, Coefficient to map bandwidth to computing	0.2 servers/Gb/s
K, Number of shortest paths pre-calculated for each node pair	5
g, Path bandwidth allocation granularity	$[1, 9]$ FS'

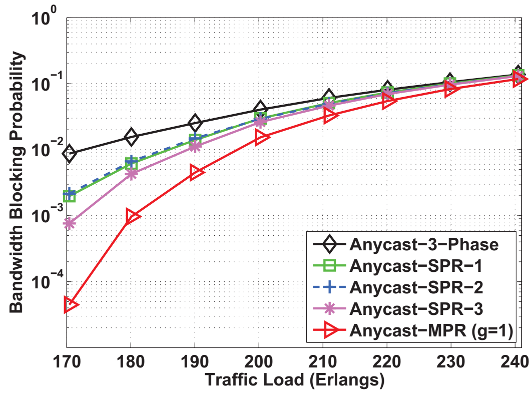

Fig. 2. Bandwidth blocking probability.

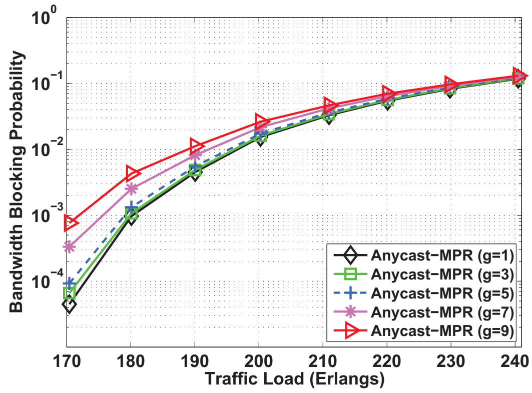

Fig. 3. Bandwidth blocking probability of Anycast-MPR with different g.

node that has the largest available computing capacity in its first phase can make computing loads be distributed evenly among datacenters, a valid routing path may not be found in consequent phases due to insufficient bandwidth. On the other hand, our proposed algorithms consider computing and bandwidth resources jointly and therefore can provide better results on BBP. Among the anycast algorithms that use single-path routing, Anycast-SPR-3 has the best performance on BBP, which suggests that when choosing the routing path for an anycast request, we should pay more attention to the path's available bandwidth, while its hop-count should be weighted less. This relation among the BBP results is also the reason why we only utilize metric $m_3(\cdot)$ in Anycast-MPR. As expected, we can see that Anycast-MPR provides the lowest BBP among all algorithms when it uses the smallest path bandwidth allocation granularity (i.e. $g = 1$). These results verify that Anycast-MRP can make better utilization of both computing and bandwidth resources in the network.

In order to investigate the effect of g on Anycast-MRP's performance, we run simulations for $g \in [1, 9]$ and plot the BBP of Anycast-MRP in Fig. 3. It can be seen that with a larger g, Anycast-MRP yields higher BBP. This is due to the fact that with a larger g, the constraint in Eq. (3) discourages path-splitting for a larger portion of the requests. By comparing the results in Figs. 2 and 3, we can also see that when $g = 9$, the BBP results of Anycast-MRP is comparable to those of Anycast-SPR-3. The results in Fig. 3 also indicate that for serving dynamic anycast requests, there is a tradeoff between BBP and operation complexity. More specifically, in order to obtain a lower BBP, we need to implement a smaller g to allow more lightpaths to be set up for a request, while to set up more lightpaths, more transponders need to be allocated, which increases the operation complexity.

Note that in the simulations, an anycast request can be blocked in three cases: 1) the bandwidth resources on the path(s) are sufficient, but the computing resources in the destination(s) are not (*Computing Blocking*), 2) the computing resources are sufficient, but the bandwidth resources are not (*Bandwidth Blocking*), and 3) both resources are insufficient (*Combinational Blocking*). We analyze the percentages of these three blocking cases for Anycast-3-Phase, Anycast-SPR-3, and Anycast-MPR, and plot the results in Figs. 4-6, respectively. In Fig. 4, we can see that the majority of the request blockings are due to *Bandwidth Blocking*, which is a clear indication that this algorithm cannot utilize the bandwidth resources in the network wisely. The situation gets improved in Fig. 5 with Anycast-SPR-3, and we can see that with the increase of the traffic load, more and more blockings are due to *Computing Blocking*. Fig. 6 shows that the majority of the request blockings are due to *Computing Blocking* when we use Anycast-MPR. Therefore, the bandwidth resources are fully utilized, and this makes the computing resources become the bottleneck. Note that in datacenter networks, upgrading the datacenters by putting in more servers is much easier and less expensive than upgrading the optical infrastructure.

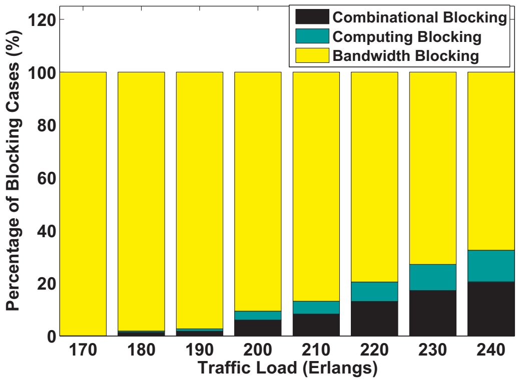

Fig. 4. Percentages of blocking cases when using Anycast-3-Phase.

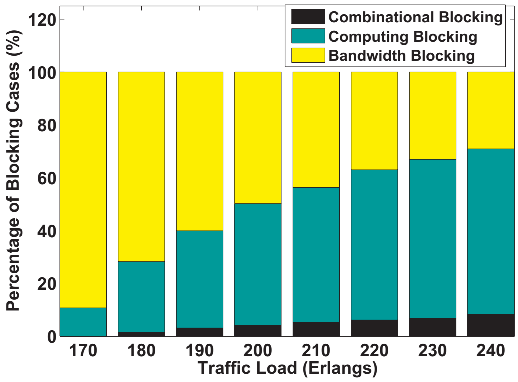

Fig. 5. Percentages of blocking cases when using Anycast-SPR-3.

V. CONCLUSION

In this paper, we investigated dynamic anycast in inter-datacenter networks over elastic optical infrastructure based on O-OFDM. Firstly, we designed three metrics to consider computing and bandwidth resources jointly, and proposed anycast algorithms with single-path routing based on them. Then, we proposed an anycast algorithm with multipath routing to further improve the network performance. Our simulation results indicated that the anycast algorithm with multipath routing could fully utilize the bandwidth resources in the optical infrastructure, make computing resources become the bottleneck, and reduce the bandwidth blocking probability of anycast requests effectively.

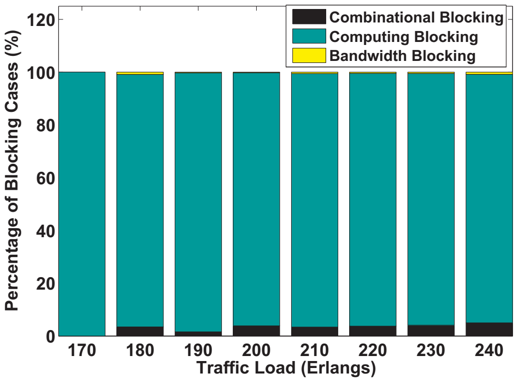

Fig. 6. Percentages of blocking cases when using Anycast-MPR ($g = 1$).

ACKNOWLEDGMENTS

This work was supported by in part the NCET program under Project NCET-11-0884, the National Natural Science Foundation of China (NSFC) under Project 61371117, the Fundamental Research Funds for the Central Universities (WK2100060010), and the Strategic Priority Research Program of the Chinese Academy of Sciences (XDA06010301).

REFERENCES

[1] B. Hayes, “Cloud computing,” *Commun. ACM*, vol. 51, pp. 9–11, Jul. 2008.
[2] W. Johnston, “Networking for the future of DOE science,” Available: https://es.net/assets/Uploads/ESnet4-Networking-for-the-Future-of-Science-2008-05-05.NP.v1.pdf.
[3] T. Xia *et al.*, “High capacity field trials of 40.5 Tb/s for LH distance of 1,822 km and 54.2 Tb/s for regional distance of 634 km,” in *Proc. of OFC 2013, paper PDP5A.4*, pp. 1–3, Mar. 2013.
[4] J. Armstrong, “OFDM for optical communications,” *J. Lightw. Technol.*, vol. 27, pp. 189–204, Feb. 2009.
[5] C. Develder *et al.*, “Optical networks for grid and cloud computing applications,” *Proc. IEEE*, vol. 100, pp. 1149–1167, May 2012.
[6] J. Buysse *et al.*, “Energy-efficient resource-provisioning algorithms for optical clouds,” *J. Opt. Commun. Netw.*, vol. 5, pp. 226–239, Mar. 2013.
[7] W. Jia, D. Xuan, and W. Zhao, “Integrated routing algorithms for anycast messages,” *IEEE Commun. Mag.*, vol. 38, pp. 48–53, Jan. 2000.
[8] D. Din, “Anycast routing and wavelength assignment problem on WDM network,” *IEICE Trans. Commun.*, vol. EE88-B, pp. 3941–3951, Oct. 2005.
[9] K. Bhaskaran, J. Triay, and V. Vokkarane, “Dynamic anycast routing and wavelength assignment in WDM networks using ant colony optimization,” in *Proc. of ICC 2011*, pp. 1–6, Jun. 2011.
[10] Q. She *et al.*, “Survivable traffic grooming for anycasting in WDM mesh networks,” in *Proc. of GLOBECOM 2007*, pp. 2253–2257, Nov. 2007.
[11] K. Walkowiak and J. Rak, “Shared backup path protection for anycast and unicast flows using the node-link notation,” in *Proc. of ICC 2011*, pp. 1–6, Jun. 2011.
[12] H. Kulkarni, A. Gadkar, and V. Vokkarane, “Deadline-aware co-scheduling using anycast advance reservations in wavelength routed lambda grids,” in *Proc. of ICNC 2013*, pp. 257–262, Jan. 2013.
[13] O. Gerstel *et al.*, “Elastic optical networking: a new dawn for the optical layer?” *IEEE Commun. Mag.*, vol. 50, pp. S12–S20, Apr. 2012.
[14] Z. Zhu, W. Lu, L. Zhang, and N. Ansari, “Dynamic service provisioning in elastic optical networks with hybrid single-/multi-path routing,” *J. Lightw. Technol.*, vol. 31, pp. 15–22, Jan. 2013.
[15] K. Christodoulopoulos, I. Tomkos, and E. Varvarigos, “Elastic bandwidth allocation in flexible OFDM-based optical networks,” *J. Lightw. Technol.*, vol. 29, pp. 1354–1366, May 2011.