QoS-Based Manycasting Over Optical Burst-Switched (OBS) Networks

Balagangadhar G. Bathula, *Student Member, IEEE*, and Vinod M. Vokkarane, *Member, IEEE*

***Abstract*—Many distributed applications require a group of destinations to be coordinated with a single source. Multicasting is a communication paradigm to implement these distributed applications. However in multicasting, if at least one of the members in the group cannot satisfy the service requirement of the application, the multicast request is said to be blocked. On the contrary in manycasting, destinations can join or leave the group, depending on whether it satisfies the service requirement or not. This dynamic membership based destination group decreases request blocking. We study the behavior of manycasting over optical burst-switched networks (OBS) based on multiple quality of service (QoS) constraints. These multiple constraints can be in the form of physical-layer impairments, transmission delay, and reliability of the link. Each application requires its own QoS threshold attributes. Destinations qualify only if they satisfy the required QoS constraints set up by the application. We have developed a mathematical model based on lattice algebra for this multiconstraint problem. Due to multiple constraints, burst blocking could be high. We propose two algorithms to minimize request blocking for the multiconstrained manycast (MCM) problem. Using extensive simulation results, we have calculated the average request blocking for the proposed algorithms. Our simulation results show that MCM-shortest path tree (MCM-SPT) algorithm performs better than MCM-dynamic membership (MCM-DM) for delay constrained services and real-time service, where as data services can be better provisioned using MCM-DM algorithm.**

***Index Terms*—BER, constraint-based routing (CBR), manycast, optical burst-switched networks (OBS), quality of service (QoS), QoS routing, WDM.**

I. Introduction

MANYCAST is also called *quorumcast* and was first proposed by [1]. Manycasting is a generalization of multicasting, in which the group of destinations that receive the message are to be selected instead of being given. In manycasting messages are sent to a subset of destinations κ (quorum pool), which are selected from set D_s (quorum group), such that $\kappa \leq |D_s| = m$. A manycast request is said to be successful if any κ of them participate in that session. A quorum pool is a majority group and hence we always require $\kappa \geq \lceil m/2 \rceil$. Manycasting is also a generalization of *anycasting* [2], where the message needs to be delivered to any one of the group. However, in this case $\kappa = 1$ and the above inequality will not be valid for anycasting.

Manycasting has caught the attention of several researchers during the recent past, due to the emergence of many distributed applications [1]–[3]. Distributed applications, such as video conferencing, distributed interactive simulations (DIS), grid computing, storage area network (SAN), and distributed content distribution network (CDN) require large amounts of bandwidths and an effective communication between single source and a set of destinations. Manycasting is also an attractive and viable communication paradigm for providing fault tolerance for the defense information infrastructures in the battlefield [4], [5]. Provisioning of connections based on QoS to these applications is an important issue [4]. QoS can include delays incurred during transmission, reliability, and signal degradation. For example reliability is an important issue in designing SANs. Since SANs are supported over fiber-channel (FC), threat to failure can occur due to cable cuts, physical attacks, and catastrophic effects. Grid applications depends on the QoS that a network can provide to ensure successful completion of the job [6].

To meet the demands of such distributed applications there is an emergence of intelligent optical control plane architectures. WDM networks include optical circuit switching (OCS), optical packet switching (OPS), and optical burst switching (OBS). In OCS a lightpath is set up by the user for the entire duration of the data transfer. In OPS the user data is transmitted in optical packets that are switched entirely in optical domain. In OBS the user data is transmitted all-optically as bursts with the help of an electronic control plane. One of the primary issues with OCS is that the link bandwidth is not utilized efficiently in the presence of bursty traffic. On the other hand, many technological limitations have to be overcome for OPS to be commercially viable. OBS networks overcome the technological constraints imposed by OPS and the bandwidth inefficiency of OCS networks [7]. In this paper, we focus on the optical transport network being OBS. The proposed algorithms can easily be modified to work for OCS and OPS networks. OBS networks are well suited for supporting delay-sensitive computationally intensive grid applications known as Grid OBS (GOBS) [8].

In this paper we propose algorithms that provide QoS-based manycasting over OBS networks. We also develop a mathematical problem formulation for manycast destination selection policies based on QoS constraints as required by certain applications. Our approach can incorporate multiple constraints related to different services. The proposed methods are service-centric and completely decentralized, as they use only local-network state information. The rest of the paper is organized as follows: we first discuss the related work in this topic in Section I-A. In Section II, mathematical problem

Manuscript received June 24, 2008; revised March 02, 2009; approved by IEEE/ACM TRANSACTIONS ON NETWORKING Editor S. Subramaniam. This work was supported in part by the National Science Foundation (NSF) Service-Oriented Optical Networks (SOON) project under Grant CNS-0626798.

B. G. Bathula is with the Department of Computer and Information Science, University of Massachusetts, Dartmouth, MA 02747 USA (e-mail: bgsquare@gmail.com).

V. M. Vokkarane is with the Department of Computer and Information Science, University of Massachusetts, Dartmouth, MA 02747 USA (e-mail: vvokkarane@ieee.org).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TNET.2009.2024498

formulation for ordering destinations based on service constraints is discussed. In Section III, the proposed algorithms are explained with illustrative examples. Section IV discusses the performance evaluation of the proposed algorithms. Finally, Section V concludes the paper.

A. Related Work

Manycasting work was first reported independently by [1] and [9] as the quorumcast problem and the κ-Steiner tree problem. It is defined as an edge cost function $g : E \rightarrow R^{+}$, an integer κ, a source s and the subset of candidate destinations $D_s \subseteq V$, $|D_s| = m \geq \kappa$, find a minimum cost tree spanning κ destinations in D_s. Cost of the tree is the sum of the cost of edges on the tree. Manycast request can be denoted by (s, D_s, κ). The manycast problem is found to be NP-hard in [9]. As IP layer is above the WDM/OBS network, the selection of κ destinations by the IP layer is similar to the random algorithm in [1]. In [10] this random algorithm has been verified using Binomial model and found to provide poor performance. Thus, supporting manycasting in OBS networks is necessary for bandwidth-efficient manycasting [11]. Apart from constructing minimum cost tree that spans from source s to manycast group members, the need for QoS routing has been discussed in [4]. This paper discusses the quality of a tree in terms of source-destination delay constraints imposed by applications that use the tree. As the delay-constrained quorumcast routing problem is NP-complete, an efficient heuristic QoS routing algorithm has been proposed in [4] with cost of the quorumcast tree close to that of optimal routing tree.

Apart from supporting manycasting over optical networks, we also need to provision QoS in OBS networks. This is because QoS provisioning methods in IP will not apply to the optical counterpart, as there is no store-and-forward model [12]. Such mechanisms for QoS provisioning in IP over OBS networks must consider the physical characteristics and limitations of the optical domain. Physical characteristics of the optical domain include optical-signal degradation, propagation delay incurred from source to destination, and link reliability from catastrophic effects. As the optical signal traverses in the transparent optical network, with the absence of electrical regenerators there will be significant loss of power due to many impairments. These impairments can be attenuation loss, multiplexer/demultiplexer loss, optical-cross-connect switch loss (OXC), and split loss (for multicast capable switches) [13]. ASE noise present in the EDFAs decreases the optical-signal-to-noise ratio (OSNR). Decrease in OSNR increases bit error rate (BER) of the signal. Hence, the signal is said to lost if BER is more than the required threshold [14], [15]. 3R regeneration of optical signal resets the effects of nonlinearity, fiber dispersion, and amplifier noise, without introducing any additional noise. This 3R regeneration requires retiming and clock recovery system, which cannot easily be carried all-optically. Hence, O/E/O conversion becomes inevitable. Delay accumulation due to O/E/O conversions can be significant when compared to the propagation delay in OBS networks. Wavelength regeneration can also result in reliability reduction and operational cost increase [16]. Challenges and requirements for introducing impairment-awareness into the management and the control planes in WDM networks have been discussed in [17]. Manycasting (or multicasting) requires the OXC to split the signal. Multicasting over optical networks can be done by the OXC switch incorporating the splitter-and-deliver (SaD) switch [18]. Depending on the fan-out of the switch the input power significantly decreases compared to unicast, thus decreasing OSNR. Multicasting under the optical layer constraint has been discussed in [19]. Power-efficient multicasting for optimizing BER has been studied in [20]. For the first time, impairment-awareness for implementing manycasting over OBS networks has been addressed in [13]. This paper discusses the importance of physical layer awareness and computes the loss due to burst contentions and high BER. Further in [10] performance of different algorithms has been discussed and an analytical model has been proposed for calculating burst loss probability.

Another important QoS parameter is the *reliability* of the links along the end-to-end path between the source and the destination. The work proposed in [21] discusses reliability for SANs. Analytical models are developed for calculation of long-term failures, service availability, and link failures. The reliability factor as a multiplicative constraint has been discussed in [16] and [22].

Performance analysis of end-to-end propagation delay and blocking probability for OBS based grids using anycasting has been presented in [23]. Different types of anycasting algorithms has been compared in [2] with the shortest-path unicast routing, where the destinations has a specific address. Manycasting over OBS networks based on multiple resources has been addressed in [24].

II. Problem Formulation

In this section, we explain the mathematical framework for multiconstrained manycasting. Our work focuses on selecting the best possible destinations that can meet the service demands effectively. Destinations chosen must be able to provide quality of service attributes. A destination is said to qualify as the member of quorum pool if it satisfies the service requirements of the application. The proposed methods are based on distributed routing, where each node individually maintains the network state information and executes the algorithm. Algorithms implemented in the centralized way, may fail due to a single failure and resulting in poor performance.

Our proposed algorithms have the following functionality:

1) Handle multiple constraints with the help of link state information available locally.
2) Service differentiated provisioning of manycast sessions.
3) Find the best possible destinations in terms of service requirements for the manycast sessions.

Notations

(s, D_s, κ) is a manycast request where s is the source node, D_s is the destination set, elements of which are probable candidates for the particular service request, and κ is the minimum number of destinations that are required to participate in the manycast session for the job to be successfully completed. Manycast session is also denoted by m/κ. Manycasting can be understood as the dynamic version of the multicast, where in the members can leave and new members can join, so that κ of them will always participate in the session [25]. The number of ways κ of them can be selected from $|D_s| = m$ is $\binom{m}{\kappa}$. We define a set $\wp_\kappa(m)$ called the power set that contains all the $\binom{m}{\kappa}$ combinations. Our work focuses on selecting the best possible

set $\in \wp_\kappa(m)$ that can meet the service demands effectively. A destination is said to qualify as the member of quorum pool if it satisfies the service requirements of the application, such as BER, reliability, and delay. Notations used for describing lattices can be found in [26].

A. Service Attributes

We define η_j, γ_j, and τ_j as noise factor, reliability factor, and end-to-end propagation delay for the Link j, respectively. Noise factor is defined as ratio of input OSNR ($\mathrm{OSNR}_{i/p} \equiv \mathrm{OSNR}_i$) and output optical signal to noise ratio ($\mathrm{OSNR}_{o/p} \equiv \mathrm{OSNR}_{i+1}$), thus we have

$$\eta_j = \frac{\mathrm{OSNR}_{i/p}}{\mathrm{OSNR}_{o/p}} \tag{1}$$

where OSNR is defined as the ratio of the average signal power received at a node to the average ASE noise power at that node. The OSNR of the link and q-factor are related as

$$q = \frac{2\sqrt{\frac{B_o}{B_e}}\mathrm{OSNR}}{1+\sqrt{1+4\mathrm{OSNR}}} \tag{2}$$

where B_o and B_e are optical and electrical bandwidths, respectively [27]. Bit-error rate is related to the q-factor as follows:

$$\mathrm{BER} = \frac{1}{2}\,\mathrm{erfc}\left(\frac{q}{\sqrt{2}}\right). \tag{3}$$

From (1)–(3), we see that η is a function of q-factor (i.e., BER). If the BER should not exceed a certain threshold (say 10^{-9}), then there exists a corresponding noise factor threshold, η_{th}. The signal is said to be lost due to high BER if the end-to-end η is greater than η_{th} and thus cannot be recovered at the destination. The overall noise factor of a burst that has traversed H hops is given by

$$\eta_H = \prod_{k=1}^{H} \eta_k \tag{4}$$

where in the above equation, the product is performed H times, starting from the initial link.

Reliability is another factor considered for providing services. We define the reliability factor γ of a link, such that $0 \leq \gamma \leq 1$, indicates the percentage of reliability for a particular link. The reliability prediction method involves the calculation of down times contributed to all building blocks required to establish end-to-end network path [21]. In this paper we assign a number generated from a uniformly distributed random variable $\sim\ U[0.6, 1]$ for each link in the network. The end-to-end reliability for the path traversing H-hops is calculated as

$$\gamma_H = \prod_{k=1}^{H} \gamma_k. \tag{5}$$

The last attribute that we consider as an important service parameter for distributed applications is propagation delay. If τ is the propagation delay of a link, then end-to-end delay for H hops, is given by

$$\tau_H = \sum_{k=1}^{H} \tau_k. \tag{6}$$

B. Path Information Vector

The service attributes can be used to maintain the local network information and by properly comparing these vectors, the destinations can be chosen. Comparison of multidimension metrics can be done using the notion of lattices [25]. Lattices are explained using the ordering denoted by $\preccurlyeq$, which has the properties of reflexivity, antisymmetry, and transitivity. We denote the information vector at link j as,

$$\Omega_j = \begin{pmatrix} \eta_j \\ \gamma_j \\ \tau_j \end{pmatrix}. \tag{7}$$

Definition 1: Let Ω_j and Ω_k be the two information vectors for the links j and k, respectively. We define $\Omega_j \preccurlyeq \Omega_k$ and comparable if and only if

$$(\eta_j \leq \eta_k) \wedge (\gamma_j \geq \gamma_k) \wedge (\tau_j \leq \tau_k). \tag{8}$$

Definition 2: Ω_j and Ω_k are not comparable if and only if any one or two of the inequalities in (8) are false. In other words if $(\eta_j > \eta_k)$ or $(\gamma_j < \gamma_k)$ or $(\tau_j > \tau_k)$. We denote them by $\Omega_j \parallel \Omega_k$.

From (4)–(6), we see that the service attributes are either multiplicative (product) or additive (sum). The ordering condition in (8) is chosen such that, noise factor and propagation delay are minimum, and reliability is maximum. Each information vector is a 3-tuple and hence it is a 3-dimensional vector space over the real field $\mathbb{R}$, which is denoted by $\mathbb{R}^3$. The operation over multidimensional vectors is given by

$$\circ : \Omega_j \in \mathbb{R}^3, \quad \Omega_k \in \mathbb{R}^3 \rightarrow \Omega_j \circ \Omega_k \in \mathbb{R}^3. \tag{9}$$

The operation $\circ$ on two vectors Ω_j and Ω_k is given by

$$\Omega_j \circ \Omega_k = \begin{pmatrix} \eta_j\eta_k \\ \gamma_j\gamma_k \\ \tau_j + \tau_k \end{pmatrix}. \tag{10}$$

Definition 3: The path information vector from source s to destination d, is denoted by $\Omega_{\langle s,d\rangle}$ and is given by

$$\Omega_{\langle s,d\rangle} = \Omega_{\langle s,s+1\rangle} \circ \cdots \circ \Omega_{\langle j,j+1\rangle} \circ \cdots \Omega_{\langle d-1,d\rangle} \tag{11}$$

where $\Omega_{\langle j,j+1\rangle}$ is the information vector for the link between the nodes $\langle j, j+1\rangle$ as shown in the Fig. 1.

Thus, using (4)–(6) the previous equation becomes

$$\Omega_{\langle s,d\rangle} = \begin{pmatrix} \prod_{k=s}^{d-1} \eta_k \\ \prod_{k=s}^{d-1} \gamma_k \\ \sum_{k=s}^{d-1} \tau_k \end{pmatrix}. \tag{12}$$

We use the notation $\Omega_{\langle s,d\rangle}$ for path information and Ω_j for the link information vector. However, if the path consists of a single link, then from the (11), we get $\Omega_{\langle s,d\rangle} = \Omega_{\langle j,j+1\rangle} \equiv \Omega_j$.

Definition 4: Consider a manycast request of the form (s, D_s, κ). Let $D_s = \{d_1, d_2, \ldots, d_m\}$. We define the next-hop (or the child nodes) corresponding to source s as $\{u_1, u_2, \ldots, u_r\}$ as shown in Fig. 2, where $r \in \mathbb{Z}^+$. From Definition 3, there exists an information vector $\Omega_{\langle s,u_i\rangle}\ \forall\, 1 \leq i \leq r$.

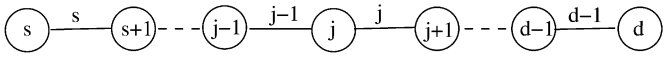


Fig. 1. Notations used in (12).

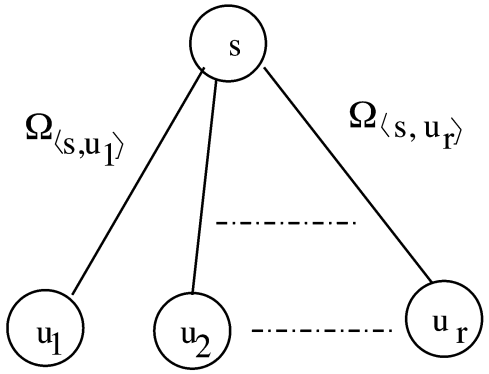


Fig. 2. Child nodes or next-hop nodes are the intermediate nodes toward the destination.

If u_i is any intermediate node, then the overall information vector from source s to the destination is computed using (12) with upper limit replaced by $u_i - 1$

Definition 5: We define differentiated service set as $\Theta = \{\theta_1, \theta_2, \ldots, \theta_S\}$. For each service $\theta_p \in \Theta$ there exists a threshold parameter (or constraint) that is defined as $\top^{(\theta_p)}$ and is given by

$$\top^{(\theta_p)} = \begin{pmatrix} \eta_{\max}^{(\theta_p)} \\ \gamma_{\min}^{(\theta_p)} \\ \tau_{\max}^{(\theta_p)} \end{pmatrix}. \tag{13}$$

For the successful establishment of QoS-based manycast session, the chosen destinations must satisfy the service requirements as defined in (13).

Theorem 1: If $\Omega_{\langle s,d\rangle} \preccurlyeq \top^{(\theta_p)}$, then all the link information vectors $\Omega_{\langle j,j+1\rangle}$'s $\forall\, j \in \{s, s+1, \ldots, d-1\}$, along the path $\langle s, d\rangle$, are comparable to $\top^{(\theta_p)}$, i.e., $\Omega_{\langle j,j+1\rangle} \preccurlyeq \top^{(\theta_p)}$.

Proof: Given $\Omega_{\langle s,d\rangle} \preccurlyeq \top^{(\theta_p)}$, then by Definition 3, we have

$$\begin{aligned} &\Omega_{\langle s,s+1\rangle} \circ \cdots \circ \Omega_{\langle j,j+1\rangle} \circ \cdots \Omega_{\langle d-1,d\rangle} \preccurlyeq \top^{(\theta_p)} \\ &\Rightarrow \Omega_{\langle j,j+1\rangle} \circ \begin{pmatrix} \prod_{\substack{k=s \\ k\neq j}}^{d} \eta_k \\ \prod_{\substack{k=s \\ k\neq j}}^{d} \gamma_k \\ \sum_{\substack{k=s \\ k\neq j}}^{d} \tau_k \end{pmatrix} \preccurlyeq \top^{(\theta_p)}. \end{aligned} \tag{14}$$

From (13) and (14) we get

$$\eta_j \prod_{\substack{k=s \\ k\neq j}}^{d} \eta_k \leq \eta_{\max}^{(\theta_p)}$$

$$\Rightarrow \eta_j \leq \frac{\eta_{\max}^{(\theta_p)}}{\prod_{\substack{k=s \\ k\neq j}}^{d} \eta_k} \leq \eta_{\max}^{(\theta_p)}.$$

Last inequality follows from the fact that $\eta_k > 1$ and, hence, we have $\eta_j \leq \eta_{\max}^{(\theta_p)}$. Similarly for other two service attributes we have

$$\begin{aligned} \gamma_j &\geq \frac{\gamma_{\max}^{(\theta_p)}}{\prod_{\substack{k=s \\ k\neq j}}^{d} \gamma_k} \geq \gamma_{\min}^{(\theta_p)} \\ \Rightarrow \gamma_j &\geq \gamma_{\min}^{(\theta_p)} (\because 0 \leq \gamma_k \leq 1). \\ \tau_j &\leq \left(\tau_{\max}^{(\theta_p)} - \sum_{\substack{k=s \\ k\neq j}}^{d} \tau_k \right) \leq \tau_{\max}^{(\theta_p)} \\ \Rightarrow \tau_j &\leq \tau_{\max}^{(\theta_p)} (\because \tau_k \in \mathbb{R}^+). \end{aligned}$$

Thus, $\forall\, j \in \{s, s+1, \ldots, d-1\}$ we have $\Omega_{\langle j,j+1\rangle} \preccurlyeq \top^{(\theta_p)}$. Hence proved. ■

Definition 6: A path $\langle s, u_i\rangle$, where u_i can be the child node or any intermediate node toward the destination is said to be feasible for service $\top^{(\theta_p)}$, if and only if $\Omega_{\langle s,u_i\rangle} \preccurlyeq \top^{(\theta_p)}$.

In the case of u_i being a child node (or next-hop node), information vector is given by (7) and for an intermediate node it is given by (12) with d replaced with u_i.

Theorem 2: The path $\langle s, d\rangle$ is a feasible path if and only if $\top^{(\theta_p)}$ is the upper bound $\forall\, \Omega_{\langle s,j\rangle}$, where $j \in \{s+1, \ldots, d-1, d\}$.

Proof:

1) *If condition:* Let $\Omega_{\langle s,d\rangle}$ be the feasible path. Then from the Definition 6 we have

$$\Omega_{\langle s,d\rangle} \preccurlyeq \top^{(\theta_p)}. \tag{15}$$

Consider $\Omega_{\langle s,s+1\rangle}$ and $\Omega_{\langle s,s+2\rangle}$, then we have

$$\begin{pmatrix} \eta_s \\ \gamma_s \\ \tau_s \end{pmatrix} \preccurlyeq \begin{pmatrix} \eta_s \eta_{s+1} \\ \gamma_s \gamma_{s+1} \\ \tau_s + \tau_{s+1} \end{pmatrix}.$$

The above inequality follows from the fact that $\eta > 1$, $0 < \gamma < 1$, and $\tau \in \mathbb{R}^+$. Thus, we have

$$\Omega_{\langle s,s+1\rangle} \preccurlyeq \Omega_{\langle s,s+2\rangle} \cdots \preccurlyeq \Omega_{\langle s,d\rangle}. \tag{16}$$

From (15) and (16), we see that $\top^{(\theta_p)}$ is the upper bound for $\Omega_{\langle s,j\rangle}$, where $j \in \{s+1, \ldots, d-1, d\}$.

2) *Only if condition:* Given $\top^{(\theta_p)}$ is the upper bound for $\Omega_{\langle s,j\rangle}$, where $j \in \{s+1, \ldots, d-1, d\}$, then we have

$$\Omega_{\langle s,d\rangle} \preccurlyeq \top^{(\theta_p)}$$

and hence the path $\langle s, d\rangle$ is feasible. ■

Theorem 3: If $\Omega_{\langle s,u_i\rangle}$, the information vector corresponding to the path $\langle s, u_i\rangle$ is not feasible, then all the destinations using node u_i as the intermediate node does not qualify as the members in quorum pool.

Proof: From the Definition 6 we see that, if $\Omega_{\langle s,u_i\rangle}$ is not a feasible path then either $\top^{(\theta_p)} \preccurlyeq \Omega_{\langle s,u_i\rangle}$ or $\Omega_{\langle s,u_i\rangle} \parallel \top^{(\theta_p)}$. If the former condition were true then

$$\begin{pmatrix} \eta_{\max}^{(\theta_p)} \\ \gamma_{\min}^{(\theta_p)} \\ \tau_{\max}^{(\theta_p)} \end{pmatrix} \preccurlyeq \begin{pmatrix} \prod_{k=s}^{u_i-1} \eta_k \\ \prod_{k=s}^{u_i-1} \gamma_k \\ \sum_{k=s}^{u_i-1} \tau_k \end{pmatrix}. \quad (17)$$

From (17) we see that if all the constraints for the service θ_p are not met, the destinations using node u_i as intermediate node disqualify to be in the quorum pool. In a similar manner if at least one of the constraints is not met, then we have $\Omega_{\langle s,u_i\rangle} \parallel \top^{(\theta_p)}$ and thus all the destinations using node u_i as intermediate node do not satisfy the service requirements of θ_p. We define this set as NRD, nonreachable destinations due to insufficient QoS. ∎

Lemma 1: A manycast request (s, D_s, κ) is said to **be** lost (or blocked) due to insufficient QoS, if and only if cardinality of the set that does not satisfy QoS is greater than κ.

Proof: We require any of the κ members in the group for successful completion of the session. But from the Theorem 3 if the number of destinations that cannot be reached through all u_i's is greater that κ, then $|NRD| > \kappa$. As $\kappa \geq \lceil m/2 \rceil$ number of remaining destinations $(m - |NRD|) < \kappa$ and, hence, the request is blocked. ∎

III. Multiconstrained Manycast (MCM) Algorithms

In this section, we explain the proposed MCM algorithms with the help of illustrative examples. We propose two algorithms, MCM-shortest path tree (MCM-SPT) and MCM-dynamic membership (MCM-DM) for evaluating the performance of manycasting with QoS constraints. We first discuss the steps to implement the distributed version of the shortest path tree (SPT) which is given in [10], [11], and [13].

- *Step 1*: Find the shortest path from source s to all the destinations in D_s. Let $D_s = \{d_1, d_2, \ldots d_{|D_s|=m}\}$ and minimum-hop distance from s to d_i, where $1 \leq i \leq m$ is $\mathbb{H}^{(s)} = \{h_1, h_2, \ldots, h_m\}$.
- *Step 2*: All the destinations in D_s are sorted in the non-decreasing order according to the shortest distance from source s to the destinations. Let D'_s be the new set in this order given by $\{d'_1, d'_2, \ldots, d'_m\}$.
- *Step 3*: Select the first κ destinations from D'_s.

Step 1 is implemented using the unicast routing table with the time complexity of $O(E + n\log(n))$ for a network with $n(= |V|)$ nodes. Step 2 sorts the destinations and the time-complexity is given by $O(m\log(m))$. In step 3, we select the first κ from D_s, with time-complexity $O(n)$. Once the first κ of them are selected, the burst header packet (BHP) is sent to corresponding child nodes $\{u_1, u_2, \ldots, u_r\}$ or the next-hop nodes where $1 \leq r \leq \kappa$ from the source s in the manycast request (s, D_s, κ). Construction of routing tree starts from the source s. Once the BHP is received at the corresponding child node(s), the burst is scheduled along the outgoing data channel. OBS is based on one-way reservation protocols, such as just-enough-time (JET) and tell-and-go (TAG) [28], [29], in which data burst is scheduled using BHPs after a certain *offset time* without waiting for the acknowledgment. The primary objective of one-way based signaling techniques is to minimize the end-to-end data transfer delay. However, this can lead to high data loss due to contention of data bursts in the OBS core. Two-way based signaling techniques are acknowledgment-based, where the request for a resource is sent from the source to the destination. The acknowledgment message confirming a successful assignment of requested resources is sent back from the destination to the source. The data burst is transmitted only after a connection is established successfully. The primary objective of the two-way based technique is to minimize packet loss in the core network, but such an objective leads to high data transfer delay due to the round-trip connection setup [30]. Our proposed methods can be modified to work with two-way reservation techniques. However we restrict our study only to one-way signaling techniques.

In this paper, we use JET signaling protocol for the manycasting. Upon receiving the data burst at the corresponding child nodes based on the QoS constraints, the manycast request is updated as $(u_i, D_{u_i}, \kappa_{u_i})$ where u_i is a child node for source s in the previous iteration, D_{u_i} are all the destinations in D_s that can be reached through u_i, and clearly $D_{u_i} \subseteq D_s$. κ_{u_i} is updated accordingly if any of the u_i is a destination and we make sure that $\sum_{i=1}^{p} \kappa_{u_i} \leq \kappa$, where p is the number of next-hop nodes for κ_{u_i}. With u_i as the source node all the above three steps are performed and this iteration proceeds until minimum κ of them are reached. We thus see that SPT works in a distributed manner and each node executes the algorithms based on the local network state information.

In the case of dynamic membership (DM), the above mentioned three steps differ slightly. In DM instead of discarding the rest of $m - \kappa$ destinations, we keep them and are used if any of the first κ are blocked due to the contention or in-sufficient QoS on the link. The former blocking is referred to as *contention blocking* and the later is referred to as *QoS blocking*. Detailed description of the *MCM-SPT* and *MCM-DM* algorithms are given in Sections III-A–C. Before we begin with the description of the proposed algorithms, it is important to know that the OBS control plane can implement this functionality effectively.

A. OBS Control Plane

In the OBS layered architecture, we find two important planes: *data plane* and *control plane*. The *control plane* allows scheduling and reservation protocols to be performed in a domain (electrical) different from data plane (optical). Detailed description of OBS data and control planes can be found in [23], [31]. Examples of control packets are BHP, network control packet (NCP), and burst confirmation packet (BCP). In this work we use BHPs as the control packets and we propose new BHP fields that provide information about QoS. In previous works [10], [13] the BHP was modified to accommodate q-factor (i.e., BER) and burst is scheduled based on the BER threshold. Table I lists possible fields associated with QoS based scheduling of bursts.

We consider the following fields of Table I in BHP as shown in the Fig. 3. In general per-hop processing delay for BHP is

TABLE I
BURST HEADER PACKET FIELDS

BHP Field	Description
Manycast Id	Manycast request identification number
Burst Id	Burst identification number used for sequencing
Source (u)	Initial or starting node of the burst
Quorum members (D_u)	These are the probable destinations to which burst can be reached.
κ_u	Number of members in manycast session
$\top^{(\theta_p)}$	Threshold information vector for service θ_p.
$\Omega_{\langle u-1,u\rangle}$	Link information vector corresponding to the link between $\langle u-1,u\rangle$.
Ingress Channel	Wavelength used for the data burst
Duration	Duration of the data burst in seconds
Offset	Time offset between the control packet the data packet

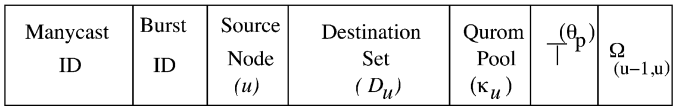


Fig. 3. Burst header packet for QoS-based Manycasting.

more due to additional computation for QoS support. We assume that the source node is aware of such processing delay and is incorporated in the offset time of edge node.

B. Multiconstraint Manycast-Shortest Path Tree Algorithm (MCM-SPT)

In this section, we explain the MCM-SPT algorithm. The pseudocode for this is given in Algorithm 1. When a new burst arrives in the network, it is assigned a unique burst ID, id. A BHP is created for this burst, with all the fields shown in the Fig. 3, where $u = s$, destination set D_s, quorum pool κ_s, threshold parameters $\top^{(\theta_p)}$ for the service θ_p, and the initial information vector $\Omega_{\text{initial}} = [1,1,0]^T$. This is indicated as line 1, in the algorithm. Destinations in D_s are sorted on the shortest path using $SORT.SP[D_s]$. Next-hop nodes from s to $d'_j \in D_s$ are calculated and added to the set N. The loop in lines 7–10 selects the first κ_u destinations from D'_s and next-hop nodes are added to N for each destination d'_j using line 9. Every link to the next-hop node is checked for contention using line 12. If the link is found to be free, the path information vector is calculated using path algebra explained in Section II-C. If $\Omega_{\langle s,n_j\rangle} \preccurlyeq \top^{(\theta_p)}$, then the link $\langle s,n_j\rangle$, qualifies the QoS threshold attributes for the service θ_p. The set of all destinations that use node n_j as the intermediate node is given by the set $DEST[n_j]$. The new destination set is given by D_{n_j} as shown in line 15. The BHP at node n_j is updated with the new values as given by line 17. One must note here that the burst ID remains the same, until it reaches the κ_s destinations. If the condition given in line 14 were false, i.e., $\Omega_{\langle u,n_j\rangle} \parallel \top^{(\theta_p)}$ or $\top^{(\theta_p)} \preccurlyeq \Omega_{\langle u,n_j\rangle}$, then according to Theorem 3 and Lemma 1 the manycast request is said to be blocked, as the minimum number of members in the pool are less then the required κ_u. We refer to this blocking as *QoS Blocking*. Burst is removed from the network due to insufficient QoS parameters. *Contention blocking* occurs when an arriving burst finds all outgoing data channels to be occupied. This algorithm repeats until the κ_s destinations are covered for a burst.

Algorithm 1 Multi-Constraint Manycast-Shortest Path Tree (MCM-SPT)

Input: The manycast request $(id, u, D_u, \kappa_u, \top^{(\theta_p)}, \Omega_{\langle u-1,u\rangle})$ arrives at the source node u with a candidate destination set D_u, along with the κ intended.

Output: Manycast request to the next-hop (or child) node after satisfying QoS parameters for the service θ_p.

1: Initialization: When the burst first enters the network with the request $(s, D_s, \kappa_s, \top^{(\theta_p)})$, we tag the request with a burst ID and $\Omega_{initial}$, where $\Omega_{initial} = [1,1,0]^T$. We therefore have the request as $(id, s, D_s, \kappa_s, \top^{(\theta_p)}, \Omega_{initial})$.
2: **if** $u \in D_u$ **then**
3: $D_u \leftarrow D_u \backslash \{u\}$;
4: $\kappa_u \leftarrow \kappa_u - 1$;
5: **else**
6: $D'_u \leftarrow SORT.SP[D_u]$;
7: **for** $j \leftarrow 1$ to κ_u **do**
8: $n_j \leftarrow NEXT.HOP.NODE.SP[u, d'_j]$;
9: $N = N \cup \{n_j\}$;
10: **end for**
11: **for** $j \leftarrow 1$ to $|N|$ **do**
12: **if** $LINK\langle u, n_j\rangle = FREE$ **then**
13: $\Omega_{\langle u,n_j\rangle} \leftarrow \Omega_{\langle u-1,u\rangle} \circ \Omega_{\langle u,n_j\rangle}$; {QoS parameters computed using the path algebra}
14: **if** $\Omega_{\langle u,n_j\rangle} \preccurlyeq \top^{(\theta_p)}$ **then**
15: $D_{n_j} \leftarrow DEST[n_j]$;
16: $UPDATE.BHP[n_j]$;
17: $(id, n_j, D_{n_j}, \kappa_{n_j}, \top^{(\theta_p)}, \Omega_{\langle u,n_j\rangle})$;
18: **else**
19: $DELETE.BURST[id]$;
20: exit; {QoS Blocking}
21: **end if**
22: **else**
23: $DELETE.BURST[id]$;
24: exit; {Contention Blocking}
25: **end if**
26: **end for**
27: **end if**

Consider a manycast request of the form $(s = 2, D_s = \{6,7,11\}, \kappa = 2)$, this configuration is represented by 3/2. Fig. 4 shows the shortest-path tree for the given manycast request of the NSF network in Fig. 5, with links shown in dotted lines. Let the service threshold be $\top^{(\theta_p)} = [\eta_{\text{th}} = 6, \gamma_{\text{th}} = 0.6, \tau_{\text{th}} = 20\text{ ms}]^T$. In order to guarantee QoS for the service θ_p, our aim is to identify destinations that have overall noise factor $\eta \leq 6$, reliability $\gamma \geq 0.6$ and the propagation delay $\tau \leq 20$ ms. When the burst enters the network at source $s = 2$, burst header fields (in Fig. 3) are updated with the initial values and the path information vector is initialized to $\Omega_{\text{initial}} = [1,1,0]^T$ as given in Step 1 of Algorithm 1. Using *SORT.SP* the destination set D_2 is sorted in the nondecreasing order of the distance and the new set is given by $D'_2 = \{7,6,11\}$. In MCM-SPT we select the first $\kappa_2 = 2$ from D'_2, we have $\{7,6\}$. Next-hop nodes for

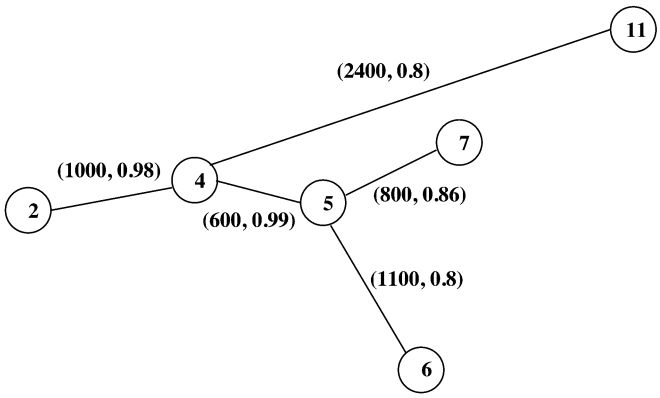


Fig. 4. Illustrative example topology. The weights indicate link distance and reliability. Other nodes and links of the NSF network are not shown for clarity. Each link consists of in-line amplifiers placed 70 km apart.

these destinations are given by $N = \{4\}$. Assuming the link $\langle 2,4\rangle$ is free (no contention), we compute the path information vector $\Omega_{\langle 2,4\rangle} \leftarrow \Omega_{\langle 2,4\rangle} \circ \Omega_{\text{initial}}$ given in line 13. Computation of noise factor is done using the parameter values given in [10], [13]. We assume the input power at node 2 as $P(2) = 1$ mW, with ASE noise as $P_{\text{ase}}(2) = 0.0042$ mW. Taking the ratio of these two powers, we get OSNR at node 2 as $\text{OSNR}(2) = 238$ and based on all the losses mentioned in [10] and [13], we have the OSNR at node 4 as $\text{OSNR}(4) = 56.53$. Using (1) we get $\eta_{\langle 2,4\rangle} = 4.21$. The propagation delay of the burst along the link (ms) is calculated as distance (km) times the velocity of light (250 km/ms). The information vector is given by

$$\Omega_{\langle 2,4\rangle} = \Omega_{\langle 2,4\rangle} \circ \Omega_{\text{initial}} = \begin{pmatrix} 4.21 \\ 0.98 \\ 4.0 \end{pmatrix} \circ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 4.21 \\ 0.98 \\ 4.0 \end{pmatrix}.$$

We thus see that $\Omega_{\langle 2,4\rangle}$ is within the threshold requirement of the service θ_p, satisfying condition in line 14. A new BHP is created at node 4, with same burst ID as, $(id, 4, D_4 = \{6,7\}, \kappa_4 = 2, \top^{(\theta_p)}, \Omega_{\langle 2,4\rangle})$. The algorithm exits in line 18. The same algorithm is repeated, however the Step-1 is skipped as this is the old burst. Lines 2–4 are used when one of the intermediate nodes is a destination. Assuming the link $\langle 4,5\rangle$ is free we have

$$\Omega_{\langle 4,5\rangle} = \Omega_{\langle 2,4\rangle} \circ \Omega_{\langle 4,5\rangle} = \begin{pmatrix} 4.21 \\ 0.98 \\ 4.0 \end{pmatrix} \circ \begin{pmatrix} 1.0702 \\ 0.99 \\ 2.4 \end{pmatrix} = \begin{pmatrix} 4.5056 \\ 0.972 \\ 6.4 \end{pmatrix}.$$

Assuming links from node 5 to nodes 7 and 8 to be free, we have $\Omega_{\langle 5,7\rangle} = [5.44, 0.83, 10.8]^T$ and $\Omega_{\langle 5,6\rangle} = [4.839, 0.776, 9.6]^T$. We thus see that the QoS threshold conditions for the service θ_p are met. The manycast session is successful for the given service. The same manycast request can be blocked for different service threshold conditions like $\top^{(\theta_q)} = [5, 0.8, 10]^T$. If at least one of the destinations is not reachable through the next-hop node, due to contention or insufficient QoS, the entire manycast request is said to be blocked. This is executed by the lines 19 and line 23.

C. Multiconstraint Manycast-Dynamic Membership (MCM-DM) Algorithm

The MCM-DM is given in Algorithm 2. It contains two procedures (1) for calculating the QoS parameters and updating BHP with the new values, defined as *Procedure.QoS( )* and (2) for calculating the number of destinations that can be reached from the next-hop node n is greater than κ_u, defined as *Procedure.Block( )*. Instead of discarding $|D_u| - \kappa_u$ destinations as in MCM-SPT, we keep these destinations as secondary destinations and use them if any of the first κ_u are blocked. Intuitively, one can understand that request blocking could be reduced in the case of MCM-DM as members in the quorum pool are added or removed dynamically. While adding the destinations into the quorum pool the burst traversal can be along a longer path, deteriorating certain QoS parameters. The resulting QoS blocking could be high when compared to MCM-SPT. This algorithm is explained using the same example $(2, \{6,7,11\}, 2)$ for which the manycast tree is shown in Fig. 4. Let the threshold conditions for the service θ_p be $[\eta_{\text{th}} = 6, \gamma_{\text{th}} = 0.6, \tau_{\text{th}} = 20\text{ms}]^T$. BHP is created for this burst and initialized as given in the line 1 of the algorithm.

Destinations are sorted by the shortest-path and $D'_2 = \{7,6,11\}$. The next-hop node(s) is given by the lines 7–10. In this case we have $N = \{4\}$. At node 4, the destination node set is $D_4 = \{6,7\}$. As we select only the first κ_2 of them, destination node 11 is left out. All destinations $|D_u - \kappa_u|$ are added in round-robin to the destination set at the child nodes (or next-hop node). Here as there is only one child node, we have $D_4 = \{6,7\} \cup \{11\}$.[1] The loop in lines 13–22 selects the primary destinations, in this case $\{6,7\}$. The next-hop node for node 6 is node 4, assuming the link $\langle 2,4\rangle$ to be free and as the condition in lines 15–17 is met (since $\kappa_4 = 0$) QoS parameters are calculated using *Procedure.QoS*. BHP at node 4 is updated as $(id, 4, D_4 = \{6,7,11\}, k_4 = 1, \top^{(\theta_p)}, \Omega_{\langle 2,4\rangle})$, with $\Omega_{\langle 2,4\rangle} = [4.21, 0.98, 4]^T$. For the next iteration, i.e., for Destination 7, the next-hop node being same, we have $k_4 = 2$, which is updated in the BHP at node 4. Finally the burst is scheduled and the BHP at node 4 is now $(id, 4, D_4 = \{6,7,11\}, k_4 = 2, \top^{(\theta_p)}, [4.21, 0.98, 4]^T)$. In a similar manner, assuming link $\langle 4,5\rangle$ to be free, the BHP at node 5 is $(id, 5, D_5 = \{6,7,11\}, k_4 = 2, \top^{(\theta_p)}, [4.5056, 0.972, 6.4]^T)$. If the link $\langle 5,7\rangle$ is not free, $k = 1$ and the BHP at node 6 becomes $(id, 6, D_6 = \{6\}, 1, \top^{(\theta_p)}, [4.839, 0.776, 9.6]^T)$. As $k = 1$ the loop in line 24 is executed. The first condition in line 26 ($k \leq |D_5| - 2$), ensures that number of blocked destinations due to contention in not greater than the secondary destinations. The next-hop node to $d'_{2+1} = 11$ from node $u = 5$ is node 4. Thus, the burst at node 5 is updated as $(id, 5, D_5 = 11, 1, \top^{(\theta_p)}, \Omega_{\langle 4,5\rangle})$. The burst at node 5 is routed to node 11 along node 4. We have

$$\Omega_{\langle 5,4\rangle} = \Omega_{\langle 4,5\rangle} \circ \Omega_{\langle 5,4\rangle} = \begin{pmatrix} 4.5056 \\ 0.972 \\ 6.4 \end{pmatrix} \circ \begin{pmatrix} 1.0104 \\ 0.99 \\ 2.4 \end{pmatrix} = \begin{pmatrix} 4.55 \\ 0.96228 \\ 8.8 \end{pmatrix}$$

[1]The procedure for round-robin is described in Appendix B.

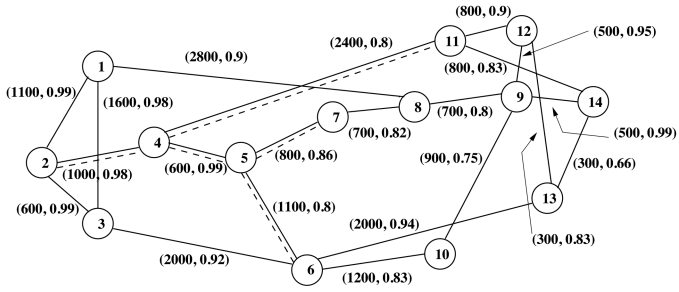


Fig. 5. NSF network with 14 nodes and 21 bidirectional links. The weights represent distance in km and the corresponding reliability factor of the links, respectively.

and the BHP at node 4 is $(id, 4, \{11\}, 1, \top^{(\theta_p)}, \ [4.55, 0.96228, 8.8]^T)$. Finally if the link $\langle 4, 11\rangle$ is free, the BHP is updated with $(id, 11, \{11\}, 1, \top^{(\theta_p)}, [5, 0.76, 18.4]^T)$. We see that the manycast request that would have been blocked using MCM-SPT if one of the destinations is blocked is now satisfied. As MCM-DM adds destinations on the longer path, it is necessary to see whether the route to the destination is within the QoS threshold requirements of the service.

IV. Performance Evaluation

In this section, we present our simulation results. We consider *average request blocking* as the performance metric. We define the average request blocking ratio as given by [11]. Let f be the total number of manycast requests used in the simulation. Consider a manycast request (s, D_s^f, κ). Let $\mathbb{D}$ be the set of destinations which actually receive the data. Then *average request blocking* is given by

$$B_{\text{avg}} = \sum_f \big[1.0 - \min(|\mathbb{D}|, \kappa)/\kappa\big]/f. \tag{18}$$

Derivation of (18) is given in Appendix C. NSF network shown in the Fig. 5 is used for our simulation. All the links are bidirectional and have the same transmission rate of 10 Gb/s. Burst arrivals follow a Poisson process with an arrival rate of λ bursts per second. The length of the burst is exponentially distributed with expected service time of $1/\mu$s.

Assumptions

1) Only one wavelength is considered for analysis. Hence, the dependency of q-factor on the wavelength is ignored.
2) Calculation of noise factor is based on losses due to attenuation, multiplexing/demultiplexing, tapping, and splitting.
3) Only amplified spontaneous emission (ASE) noise can be considered for OSNR. The shot noise and beat noise are ignored.
4) Effects of offset time are ignored.
5) In line amplifiers along the links are placed, with spacing of 70 km between the amplifiers.
6) There are no optical buffers or wavelength converters in the network.
7) The reliability factor is same along both directions of the fiber.

Algorithm 2 Multi-Constraint Manycast-Dynamic Membership (MCM-DM)

Input: The manycast request $(id, u, D_u, \kappa_u, \top^{(\theta_p)}, \Omega_{\langle u-1,u\rangle})$ arrives at the source node u with a candidate destination set D_u, along with the κ intended.

Output: Manycast request to the next hop (or child) node after satisfying QoS parameters for the service θ_p.

Initialization: $(id, s, D_s, \kappa_s, \top^{(\theta_p)}, \Omega_{initial})$.
if $u \in D_u$ **then**
  $D_u \leftarrow D_u \backslash \{u\}$;
  $\kappa_u \leftarrow \kappa_u - 1$;
else
  $D'_u \leftarrow SORT.SP[D_u]$;
  for $j \leftarrow 1$ to κ_u **do**
    $n_j \leftarrow NEXT.HOP.NODE.SP[u, d'_j]$;
    $N = N \cup \{n_j\}$;
  end for
  $ROUND.ROBIN.DEST[|D_u| - \kappa_u]$;
  $k \leftarrow 0$;
  for $j \leftarrow 1$ to κ_u **do**
    if $LINK\langle u, n_j\rangle = FREE$ **then**
      while $\sum_{k=1}^{|N|} \kappa_{n_k} < \kappa_u$ **do**
        $\kappa_{n_j} \leftarrow \kappa_{n_j} + 1$;
      end while
      Procedure.QoS( );
    else
      $k \leftarrow k + 1$;
    end if
  end for
end if
for $i \leftarrow 1$ to k **do**
  $n_k \leftarrow NEXT.HOP.NODE.SP[u, d'_{k+\kappa_u}]$;
  if $(k \le |D_u| - \kappa_u) \& (LINK\langle u, n_k\rangle = FREE)$ **then**
    Procedure.QoS( );
  else
    Procedure.Block( );
  end if
end for
Procedure.QoS( ) {
$\Omega_{\langle u,n\rangle} \leftarrow \Omega_{\langle u-1,u\rangle} \circ \Omega_{\langle u,n\rangle}$;
if $\Omega_{\langle u,n\rangle} \preccurlyeq \top^{(\theta_p)}$ **then**
  $UPDATE.BHP[n]$;
  $(id, n, D_n, \kappa_n, \top^{(\theta_p)}, \Omega_{\langle u,n\rangle})$;
else
  Procedure.Block( );
end if} { `End of Procedure.QoS( )`}
Procedure.Block( ) {
$NOT.REACH.DEST[n] \leftarrow DEST[n]$;
if $|NOT.REACH.DEST[n]| > \kappa_u$ **then**
  $DELETE.BURST[id]$;
else
  `continue`;
end if} { `End of Procedure.Block( )`}

As aforementioned, we have intended destinations, i.e., quorum pool to be majority of the group $\kappa \ge \lceil m/2 \rceil$. The candidate

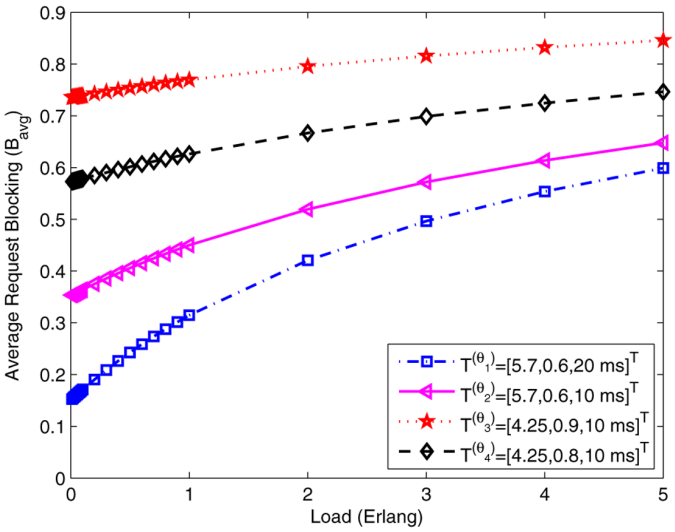


Fig. 6. Blocking probability of MCM-SPT for different service thresholds.

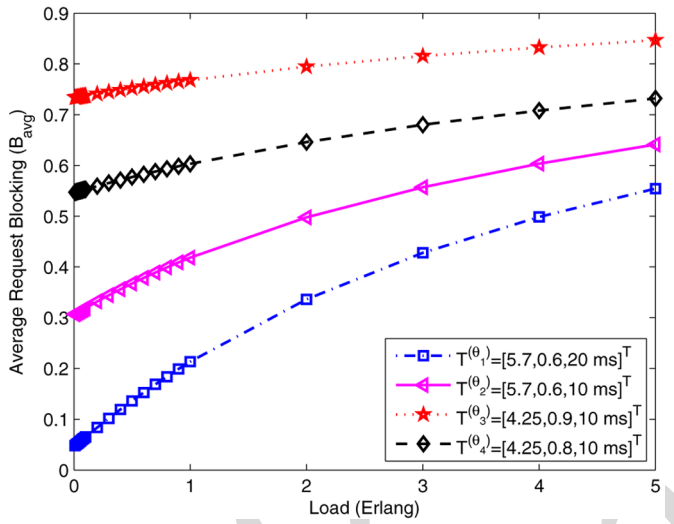


Fig. 7. Blocking probability of MCM-DM for different service thresholds.

destination group (or quorum group) can be small, medium or large. Three typical configurations 3/2, 7/4, and 11/6 are considered for simulations. First we present simulation results for 7/4 manycast configuration. We differentiate among service requirements, i.e., different services have different constraints. Differentiated services considered for simulation are $\top^{(\theta_1)} = [5.7, 0.6, 20]^T$, $\top^{(\theta_2)} = [5.7, 0.6, 10]^T$, $\top^{(\theta_3)} = [4.25, 0.9, 10]^T$, and $\top^{(\theta_4)} = [4.25, 0.8, 10]^T$. The service $\top^{(\theta_1)}$ can represent data service as it has relaxed delay requirements. The service $\top^{(\theta_2)}$ can represent real-time service as it has stringent delay requirement. Service $\top^{(\theta_3)}$ and service $\top^{(\theta_4)}$ have high service threshold requirements.

Fig. 6 shows the performance of the MCM-SPT for different set of services. As the service requirements increase, the blocking probability also increases. MCM-SPT uses shortest-path routing and one can expect to have a lower QoS blocking. But due to random contentions along the links, if any one of the destinations is not reachable, the entire manycast request would be blocked. On the contrary, MCM-DM adds or removes destinations based on the contention in the network. However destinations that are added to the quorum pool can be at a longer distance than the blocked destination. As the result, QoS of this destination can be decreased. In spite of decrease in values, if the path-information vector is within the service threshold, the request can be satisfied. Fig. 7 shows average request blocking for MCM-DM under different service thresholds. At high loads, most of the blocking would be contention blocking and hence the effect of QoS will not be understood much. As our objective is to show the effects of QoS, all the results are simulated under medium network load conditions.

Fig. 8 shows the comparison of average request blocking for the two proposed algorithms. We see that for the data service requirement $\top^{(\theta_1)}$, there is a significant reduction in the request blocking for the network loads between $(0, 1]$. As the network load increase the performance of two algorithms converges and for loads greater than 5, the request blocking is same. Under real-time service requirements like service θ_2, we observe from Fig. 9 that performance benefit of MCM-DM is reduced and the difference in the request loss between MCM-SPT and MCM-DM has decreased. This can be accounted to the fact

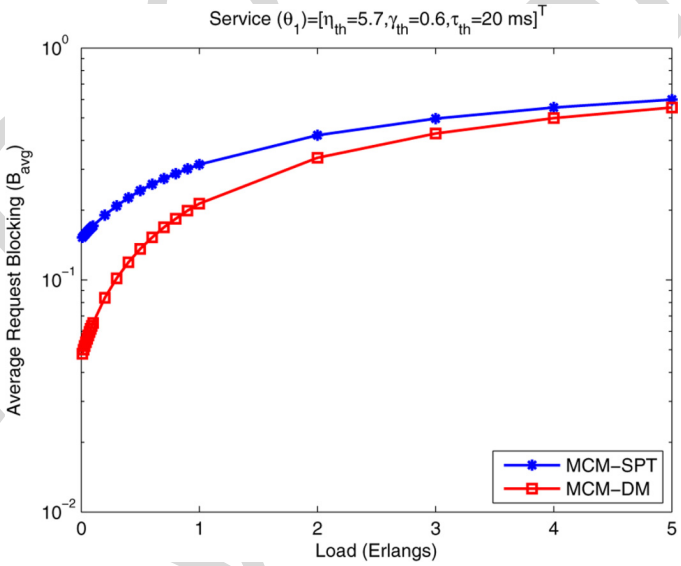


Fig. 8. Blocking probability of MCM-SPT and MCM-DM for 7/4 manycast configuration with service threshold of $\top^{(\theta_1)} = [\eta_{\text{th}} = 5.7, \gamma_{\text{th}} = 0.6, \tau_{\text{th}} = 20 \text{ ms}]^T$.

that, while adding secondary destinations, the longer paths have to be traversed and hence the delay increases, causing a destination to be disqualified based on the delay constraint. We can thus observe that MCM-DM can be chosen for data service applications, where there is no specific upper bound on the propagation delay of the burst. Data service based distributed applications like SAN and CDN have more priority on η and γ rather than τ. We have also simulated the performance of the algorithms for more stringent QoS requirement, like service θ_3. We observe that both the algorithms behave the same for this service requirements. By relaxing the constraint on reliability, we observe a significant decrease in the request loss in Fig. 11.

The same set of services were simulated for the other two manycast configurations, i.e., 3/2 and 11/6. At high loads contentions in the network could be large and hence the effect of the QoS may not be significant. Hence, we restricted our simulation study only to low loads. Fig. 12 shows the performance

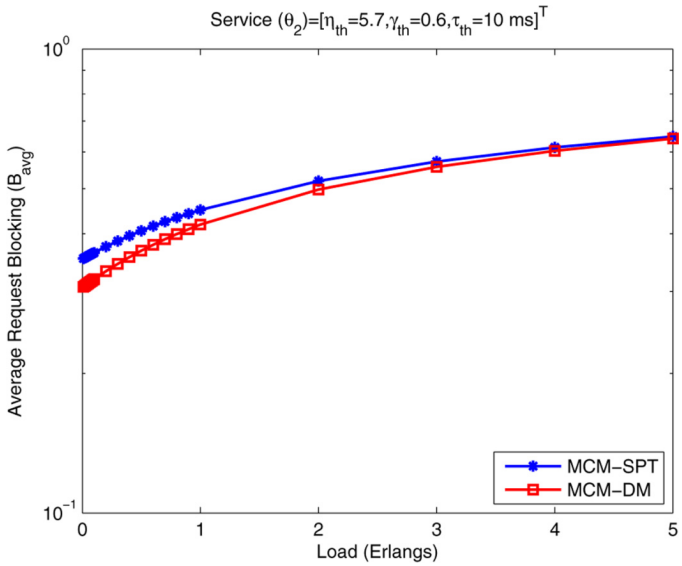
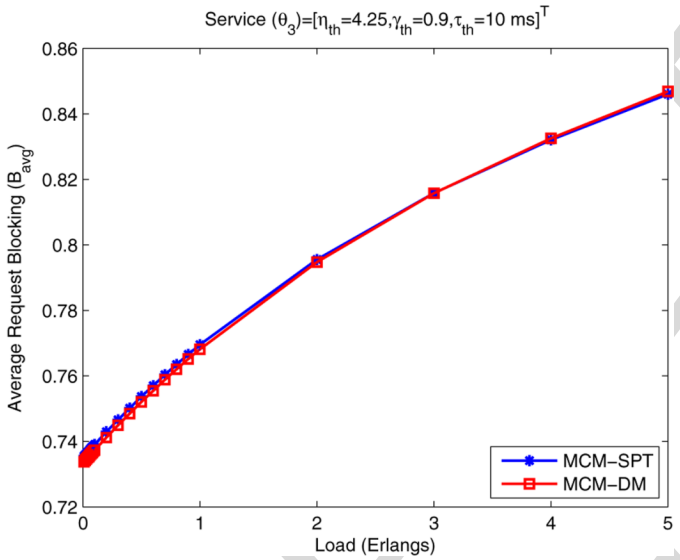


Fig. 9. Blocking probability of MCM-SPT and MCM-DM for 7/4 manycast configuration with service threshold of $\top^{(\theta_2)} = [\eta_{\text{th}} = 5.7, \gamma_{\text{th}} = 0.6, \tau_{\text{th}} = 10 \text{ ms}]^T$.



[Cite figure in text] Fig. 10. Blocking probability of MCM-SPT and MCM-DM for 7/4 manycast configuration with service threshold of $\top^{(\theta_3)} = [\eta_{\text{th}} = 4.25, \gamma_{\text{th}} = 0.9, \tau_{\text{th}} = 10 \text{ ms}]^T$.

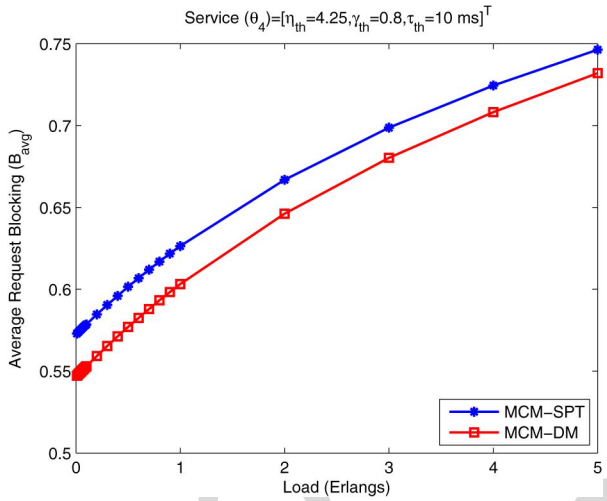


Fig. 11. Blocking probability of MCM-SPT and MCM-DM for 7/4 manycast configuration with service threshold of $\top^{(\theta_4)} = [\eta_{\text{th}} = 4.25, \gamma_{\text{th}} = 0.8, \tau_{\text{th}} = 10 \text{ ms}]^T$.

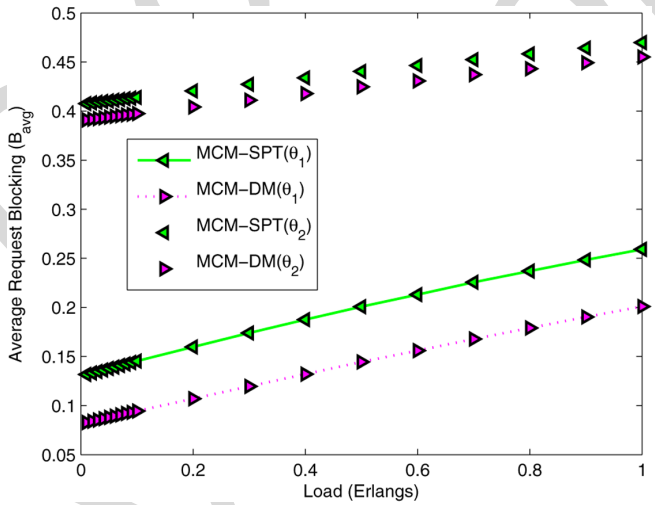


Fig. 12. Blocking probability of MCM-SPT and MCM-DM for $3/2$ manycast configuration for services θ_1 and θ_2.

of 3/2 manycast configuration for services θ_1 and θ_2. Service θ_1 has more relaxed threshold parameters, hence, θ_1 we can improve the blocking marginally using MCM-DM. But in the case of θ_2, where the delay requirement is only 10 ms, we observe that both algorithms offer same performance in terms of request blocking.

We observe an interesting result in Fig. 13. In the case of θ_3 for 3/2 manycast configuration, we find MCM-SPT to offer lower request blocking than MCM-DM. This is because service θ_3 has high QoS requirement (real-time service). In an attempt to decrease the request blocking MCM-DM schedules the burst on longer paths, which causes service attributes to exceed the threshold requirements. Once again we see that MCM-DM can only be used when there are much relaxed QoS parameters (data services).

Finally we also simulate 11/6 manycast configuration for the four services. Figs. 14 and 15 show much similar performance to that of 7/4 and 3/2.

We simulate the impact of each service attribute on the network for 7/4 manycasting. In the other words, only one service attribute (i.e., delay, BER or reliability) is considered, with the remaining two neglected. Thus, the problem reduces to a single-constrained manycast problem. Let service threshold for the delay constrained (DC) service be $\top^{(\theta_d)} = [\infty, 0, 10]^T$. The threshold requirements on noise factor and reliability are removed by keeping $\eta_{\max}^{(\theta_d)} = \infty$ and $\gamma_{\min}^{(\theta_d)} = 0$. Similarly we consider BER constrained (BC) and reliability constrained (RC) with service thresholds given by $\top^{(\theta_b)} = [4.25, 0, \infty]^T$ and $\top^{(\theta_r)} = [\infty, 0.8, \infty]^T$, respectively. Fig. 16 shows the performance of 7/4 manycast configuration for MCM-SPT and MCM-DM. We observe a significant decrease in the request blocking for MCM-DM compared to MCM-SPT in the case

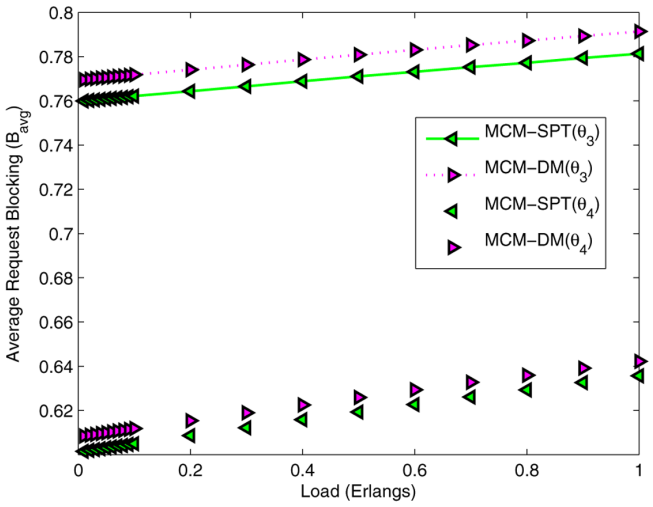


Fig. 13. Blocking probability of MCM-SPT and MCM-DM for 3/2 manycast configuration for services θ_3 and θ_4.

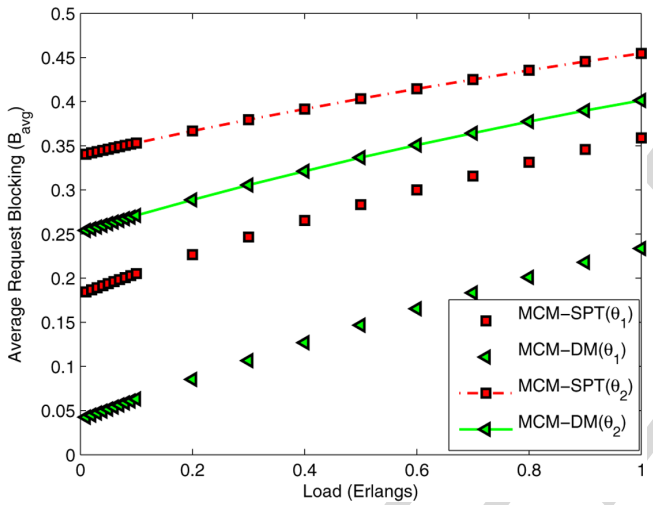


Fig. 14. Blocking probability of MCM-SPT and MCM-DM for 11/6 manycast configuration for services θ_1 and θ_2.

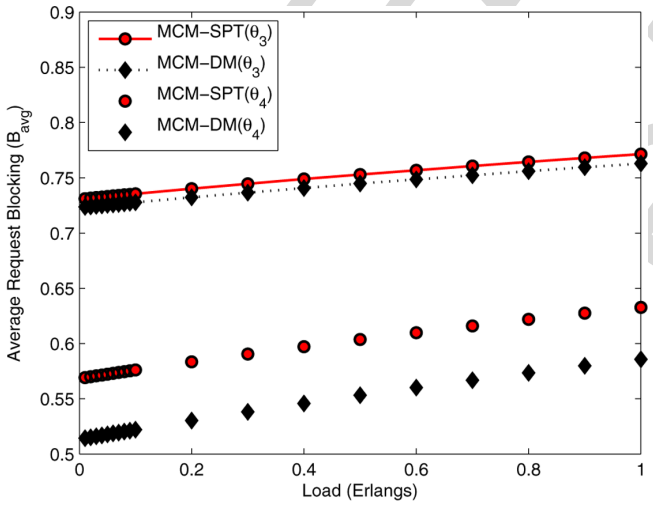


Fig. 15. Blocking probability of MCM-SPT and MCM-DM for 11/6 manycast configuration for services θ_3 and θ_4.

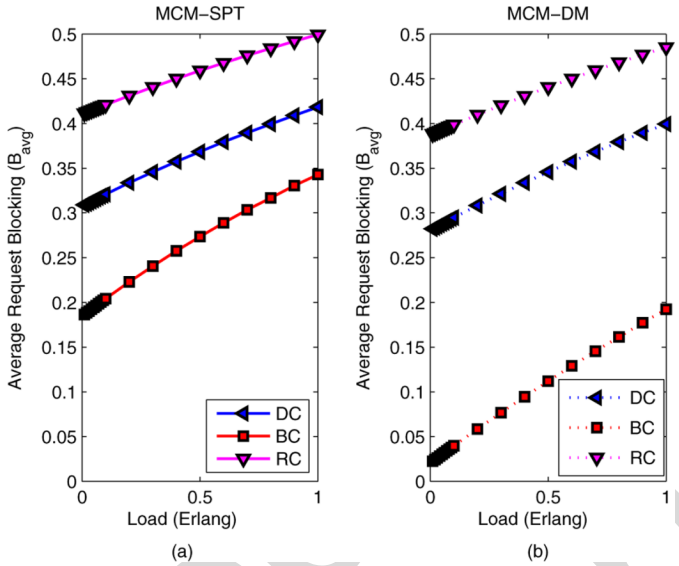


Fig. 16. Performance of attribute specific $7/4$ manycasting for (a) MCM-SPT and (b) MCM-DM.

of BER constrained (BC). Similarly services θ_b and θ_r are observed to have lower average request blocking with MCM-DM.

V. Conclusion

In this paper, we propose algorithms to support QoS-based manycasting over optical burst-switched networks. Our QoS model supports certain service parameters for the transmission of optical bursts, such as physical impairments, reliability, and propagation delay. We have developed a mathematical model based on lattice algebra for the multiconstraint manycast problem. By using distributed scheduling algorithms, bursts are routed to the destinations based on the contentions and QoS conditions. We propose the multiconstrained manycast shortest path tree algorithm and the multiconstrained manycast dynamic membership algorithm to support QoS-based manycasting over OBS. Four types of services were considered to evaluate the performance of QoS-based manycasting algorithms. They represent the service requirements of data service and real-time service. We observe from the simulation results that multiconstrained manycast dynamic membership algorithm is suited for data service and multiconstrained manycast shortest path tree algorithm for real-time service. We also evaluated the performance of our algorithms for different manycast configurations. Our proposed manycasting algorithms can be easily adapted to facilitate other application service requirements. Our work can be further extended by considering sparse wavelength regenerations.

Appendix A
Calculation of Noise Factor Threshold

We calculate the noise factor threshold $\eta_{\max}$ [32] using (1)–(3). From (1), the noise factor of the link j is given by

$$\eta_j = \left(\frac{P(j)}{P_{\text{ase}}(j)}\right)\left(\frac{P_{\text{ase}}(j+1)}{P(j+1)}\right).$$

For the path from source s to the destination d, the overall noise factor is given by

$$\eta_{\langle s,d\rangle} = \left(\frac{P(s)}{P_{\text{ase}}(s)}\right)\left(\frac{P_{\text{ase}}(d)}{P(d)}\right).$$

We assume the transmitting power of the receiver is $P(s) = 1$ mW. The ASE noise power at source node s is given by

$$P_{\text{ase}}(s) = P_{\text{ini}}L_dL_mL_tL_{\text{ins}}(G_{\text{in}}-1)G_{\text{out}} + P_{\text{ini}}L_t[G_{\text{out}}-1] \quad (19)$$

$P_{\text{ini}} = 2n_{sp}hf_cB_o$, where n_{sp} spontaneous-emission factor, h is the Plank's Constant, and f_c is the central frequency of the optical signal. L_d, L_m, L_t, and L_{ins} are demultiplexer, multiplexer, tap, and insertion loss of the optical cross-connect switch, respectively. G_{in} and G_{out} are the input and output gains of the erbium-doped fiber amplifier (EDFA) in switch. Parameter values and the switch architecture can be found in [10], [13], [14]. By using (19) we get the $P_{\text{ase}}(s) = 0.0042$ mW. Thus, the OSNR at source s will be $\text{OSNR}(s) = 238$ a.u. For BER of 10^{-12} we need $q \approx 7$, for which the required $\text{OSNR}(d)$ is obtained by solving (2)

$$7 = \frac{2\sqrt{\frac{B_o}{B_e}}\text{OSNR}(d)}{1+\sqrt{1+4\text{OSNR}(d)}}. \quad (20)$$

For a system operating $B = 10$ Gb/s with $B_o = 70$ GHz and $B_e = 0.7B$, $\text{OSNR}(d) = 56$, which is obtained by solving (20). Hence, if the $\text{OSNR}(d) < 56\ (= \text{OSNR}_{\text{min}})$ then the BER will increase beyond 10^{-12}. Thus, the noise factor threshold $\eta_{\text{max}} = 4.25$ corresponding to $q = 7$. Similarly for $q = 6$, $\eta_{\text{max}} \approx 6$. Thus, we see that as long as the noise-factor of the burst is $\leq \eta_{\text{max}}$, burst can be scheduled for transmission. We derive the relation for noise-factor threshold (η_{max}) and q-factor threshold (q_{th}).

$$\eta_{\text{max}} = \frac{\text{OSNR}(s)}{\text{OSNR}_{\text{min}}}. \quad (21)$$

In order for the BER to be less than the given threshold, the OSNR at the destination should be greater than the OSNR_{min}. Thus, (2) at the threshold conditions is given by

$$q_{\text{th}} = \frac{2\sqrt{\frac{B_o}{B_e}}\text{OSNR}_{\text{min}}}{1+\sqrt{1+4\text{OSNR}_{\text{min}}}}. \quad (22)$$

Solving this equation for OSNR_{min}, we get

$$\text{OSNR}_{\text{min}} = q_{\text{th}}\left(q_{\text{th}} + \sqrt{\frac{B_o}{B_e}}\right)\frac{B_o}{B_e}. \quad (23)$$

Substituting (23) in (21) we get

$$\eta_{\text{max}} = \frac{\text{OSNR}(s)}{q_{\text{th}}\left(\frac{B_o}{B_e}\right)\left(q_{\text{th}} + \sqrt{\frac{B_o}{B_e}}\right)}. \quad (24)$$

APPENDIX B
ROUND-ROBIN PROCEDURE

Details of the round-robin procedure used in Algorithm 2 are given below.

$N_u \leftarrow NEXT.HOP.NODES\{k_u\}$;
$i \leftarrow 1$;
for $l \leftarrow 1$ to $|D_u| - k_u$ **do**
  $D_{N_i} \leftarrow D_{N_i} \cup \{d_{k_u+l}\}$;
  if $i < |N_u|$**then**
    $i \leftarrow i+1$;
  else
    $i = 1$;
  end if
end for

APPENDIX C
DERIVATION OF (18)

Let f be the total number of manycast requests used in the simulation. Every new burst (manycast request) entering the network is assigned a unique ID, the manycast request is given by (s, D_s^f, κ). D_s^f represents the destination set for a burst with manycast request ID f. This ID remains the same until burst reaches its destinations. Let $\mathbb{D}$ be the set of destinations that receive the burst. In MCM, the manycast request is said to be successful only if the burst reaches κ destinations. Thus, if $|\mathbb{D}| = \kappa$, then probability of success is 1. On the other hand, if burst reaches destinations such that $|\mathbb{D}| < \kappa$, then probability of blocking for a burst f is given by $P_f = 1 - |\mathbb{D}|/\kappa$. The ensemble average blocking probability thus becomes

$$B_{\text{avg}} = \frac{\text{Sum of blocking probabilities}}{\text{Total number of bursts}} = \frac{P_1 + P_2 + \cdots + P_f}{f}$$

where

$$P_i = 1 - \min(|\mathbb{D}|,\kappa)/\kappa, 1 \leq i \leq f = \sum_f [1.0 - \min(|\mathbb{D}|,\kappa)/\kappa]/f.$$

ACKNOWLEDGMENT

The authors would like to thank R. R. C. Bikram for working on certain simulations. They would also thank Prof. S. Talabattula, Department of Electrical and Communication Engineering, Indian Institute of Science, for his valuable inputs. Also, the authors would like to thank the reviewers for their valuable comments to improve the quality of the paper.

REFERENCES

[1] S. Y. Cheung and A. Kumar, "Efficient quorumcast routing algorithms," in *Proc. IEEE INFOCOM*, Toronto, ON, Canada, Jun. 1994, pp. 840–847.

[2] M. D. Leenheer, F. Farahmand, K. Lu, T. Zhang, P. Thysebaert, B. Volckaert, F. D. Turck, B. Dhoedt, P. Demeester, and J. P. Jue, "Anycast algorithms supporting optical burst switched grid networks," in *Proc. IEEE ICNS*, Silicon Valley, CA, Jul. 2006, pp. 63–69.
[3] N. Garg, "Saving an epsilon: A 2-approximation for the k-MST problem in graphs," in *Proc. 33rd Ann. ACM Symp. Theory of Comput.*, Baltimore, MD, Jul. 2005, pp. 396–402.
[4] B. Wang and J. C. Hou, "An efficient QoS routing algorithm for quorumcast communication," in *Proc. IEEE Netw. Protocols*, Riverside, CA, Nov. 2001, pp. 110–118.
[5] N. Charbonneau, V. M. Vokkarane, R. Balasubramanian, and D. Silvia, "Mascot: Manycast architecture for service-oriented tactical applications," in *Proc. IEEE Int. Conf. Technol. Homeland Secur.*, Waltham, MA, May 2009, pp. 171–176.
[6] A. Jukan, "Optical control plane for the grid community," *IEEE Commun. Surv. Tutor.*, vol. 9, no. 3, pp. 30–44, 2007.
[7] L. Xu, H. G. Perros, and G. Rouskas, "Techniques for optical packet switching and optical burst switching," *IEEE Commun. Mag.*, vol. 39, no. 1, pp. 136–142, Jan. 2001.
[8] R. Nejabati, "Grid optical burst switched networks (GOBS)," 2006 [Online]. Available: www.ogf.org/Public Comment Docs/Documents/Jan-2007/OGF GHPN GOBS final.pdf
[9] R. Ravi, R. Sundaram, M. V. Marathe, D. J. Rosenkrantz, and S. S. Ravi, "Spanning trees short or small," in *Proc. 5th Annu. ACM-SIAM Symp. Discrete Algor.*, Arlington, VA, Jan. 1994, pp. 546–555.
[10] B. G. Bathula, R. R. C. Bikram, V. M. Vokkarane, and T. Srinivas, "Impairment-aware manycast algorithms over optical burst switched (OBS) networks," in *Proc. IEEE ICCCN*, St. Thomas , U.S. Virgin Islands, Aug. 2008, pp. 1–6.
[11] X. Huang, Q. She, V. M. Vokkarane, and J. P. Jue, "Manycasting over optical burst-switched (OBS) networks," in *Proc. IEEE ICC*, Glasgow, Scotland, May 2007, pp. 2353–2358.
[12] A. Kaheel, T. Khattab, A. Mohamed, and H. Alnuweiri, "Quality-of-service mechanisms in IP-over-WDM networks," *IEEE Commun. Mag.* , vol. 40, no. 12, pp. 38–43, Dec. 2002.
[13] B. G. Bathula, V. M. Vokkarane, and R. R. C. Bikram, "Impairment-aware manycasting over optical burst-switched (OBS) networks," in *Proc. IEEE ICC*, Bejing, China, May 2008, pp. 5234–5238.
[14] B. Ramamurthy, D. Datta, H. Feng, J. P. Heritage, and B. Mukherjee, "Impact of transmission impairments on the teletraffic performance of wavelength-routed optical networks," *J. Lightw. Technol.*, vol. 17, no. 10, pp. 1713–1723, Oct. 1999.
[15] Y. Huang, J. P. Heritage, and B. Mukherjee, "Connection provisioning with transmission impairment consideration in optical WDM networks with high-speed channels," *J. Lightw. Technol.*, vol. 23, no. 3, pp. 982–993, Mar. 2005.
[16] A. Jukan and G. Franzl, "Path selection methods with multiple constraints in service-guaranteed WDM networks," *IEEE/ACM Trans. Netw.*, vol. 12, no. 1, pp. 59–72, Feb. 2004.
[17] R. Martinez, F. Cugini, N. Andriolli, L. Wosinska, and J. Comellas, "Challenges and requirements for introducing impairment-awareness into management and control planes of ASON/GMPLS WDM networks," *IEEE Commun. Mag.*, vol. 44, no. 12, pp. 76–75, Dec. 2007.
[18] W. S. Hu and Q. J. Zeng, "Multicasting optical cross connects employing splitter-and-delivery switch," *IEEE Photon. Technol. Lett.*, vol. 10, no. 7, pp. 970–972, Jul. 1998.
[19] Y. Xin and G. N. Rouskas, "Multicast routing under optical constraint," in *Proc. IEEE INFOCOM*, Hong Kong, China, Mar. 2004, pp. 2731–2742.
[20] M. Ali and J. Deogun, "Power-efficient design of multicast wavelength routed networks," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 10, pp. 1852–1862, Oct. 2000.
[21] X. Qiu, R. Telikepalli, T. Drwiega, and J. Yan, "Reliability and availability assessment of storage area network extension solutions," *IEEE Commun. Mag.*, vol. 43, no. 3, pp. 80–85, Mar. 2005.
[22] A. Jukan and G. Franzl, "Constraint-based path selection methods for on-demand provisioning in WDM networks," in *Proc. IEEE INFOCOM*, New York, Jun. 2002, pp. 827–836.
[23] F. Farahmand, M. D. Leenheer, P. Thysebaert, B. Volckaert, F. D. Turck, B. Dhoedt, P. Demeester, and J. P. Jue, "A multi-layered approach to optical burst-switched based grids," in *Proc. IEEE Int. Conf. Broadband Commun., Netw., Syst.*, Boston, MA, Jul. 2005, pp. 1050–1057.
[24] Q. She, X. Huang, N. Kannasoot, Q. Zhang, and J. P. Jue, "Multi-resource manycast over optical burst switched networks," in *Proc. IEEE ICCCN*, Honolulu, HI, Aug. 2007, pp. 222–227.
[25] A. B. Przygienda, "Link state routing with QoS in ATM LANs," Ph.D. dissertation, Swiss Fed. Inst. Technol. (ETH), Zurich, Switzerland, 1995.
[26] B. A. Davey and H. A. Priestley, *Introduction to Lattices and Order*. Cambridge, U.K.: Cambridge Univ. Press, 2002.
[27] R. Ramaswami and K. N. Sivarajan, *Optical Networks*. San Francisco, CA: Morgan Kaufmann, 2004.
[28] M. Yoo and C. Qiao, "Just-enough-time (JET): A high speed protocol for burst traffic in optical networks," in *Proc.IEEE/LEOS Summer Topical Meet. Dig. Conf. Technol. Global Inf. Infrastructure*, Montreal, Canada, Aug. 1997, pp. 26–27.
[29] C. Qio and M. Yoo, "Optical burst switching OBS-a new paradigm for an optical internet," *J. High Speed Netw.*, vol. 8, no. 1, pp. 69–84, 1999.
[30] V. M. Vokkarane, "Intermediate-node-initiation (INI): A generalized signaling framework for optical burst switched networks," *Opt. Switch. Netw.*, vol. 4, no. 1, pp. 20–32, 2007.
[31] J. P. Jue and V. M. Vokkarane, *Optical Burst-Switched Networks*, ser. Opt. Netw. Ser.. Berlin, Germany: Springer, 2005.
[32] B. G. Bathula, "QoS aware quorumcasting over optical burst switched networks," Ph.D. dissertation, Indian Inst. Sci., Bangalore, India, 2008.

Balagangadhar G. Bathula (S'08) was born in Pithapuram, India. He received the M.Tech. degree in optoelectronics and laser technology from Cochin University of Science and Technology, Cochin, India, in 2004, and the Ph.D. degree in electrical communication engineering the from Indian Institute of Science, Bangalore, India, in 2008.

From August 2007 to May 2008, he was a Visiting Scholar with the Department of Computer and Information Science, University of Massachusetts, Dartmouth. From August 2008 to September 2009, he worked as a Post-Doctoral Research Associate with the School of Electronic and Electrical Engineering, University of Leeds, Leeds, U.K. He is currently working as a Visiting Scholar with the Department of Computer and Information Science, University of Massachusetts, Dartmouth. His research interests include performance modeling of optical networks and grid over optical networks.

Vinod M. Vokkarane (S'02–M'04) received the B.E. degree with Honors in computer science and engineering from the University of Mysore, Mysore, India, in 1999, and the M.S. and Ph.D. degrees in computer science from the University of Texas at Dallas in 2001 and 2004, respectively.

He is currently an Assistant Professor of Computer and Information Science with the University of Massachusetts, Dartmouth. He is co-author of the book *Optical Burst Switched Networks* (New York: Springer, 2005). His primary areas of research include design and analysis of architectures and protocols for optical and wireless networks.

Dr. Vokkarane received the Texas Telecommunication Engineering Consortium Fellowship for 2002–2003 and the University of Texas at Dallas Computer Science Dissertation of the Year Award 2003–2004.