Performance Analysis for Overlay Multimedia Multicast on r-ary Tree and m-D Mesh Topologies

Wanqing Tu, *Member, IEEE*, Xing Jin, and Peter S. Excell, *Senior Member, IEEE*

***Abstract*—Without requiring multicast support from the underlying networks, overlay multicast has the advantage of implementing inter-domain multimedia multicast communications. Usually, overlay multicast protocols employ two different topologies: r-ary tree and m-D mesh. In this paper, we study the influence of topology selection on multimedia multicast performance. We present a set of theoretical results on the worst performance, the average performance, and the performance difference along the link stress, the number of overlay hops, and the number of shortest paths for r-ary tree-based and m-D mesh-based multicast, respectively. Furthermore, through simulations in NS2, we observe and compare tree and mesh topologies along the metrics analyzed theoretically. Simulation results match our theoretical analyses. Finally we give our evaluations of and insights into these two kinds of multicast when used to transmit multimedia streams. The selection of overlay topology is application dependent. To the best of our knowledge, this is the first evaluation of multimedia multicast performances in different overlay topologies. We believe that this study is useful for protocol design of target multimedia applications and for investigating multicast functions.**

***Index Terms*—m-D mesh-based overlay multicast, overlay multimedia multicast, performance analysis, r-ary tree-based overlay multicast.**

I. Introduction

HIGH-SPEED multimedia applications (e.g., multimedia conferencing, distance learning, multimedia collaboration, distributed database, online games, distributed simulation, network broadcast and data replication) inherently require group communication to be carried over the commodity Internet. Group communication desires and drives a multicast communication mode because the use of several unicasts is unfeasible in terms of network resources and required processing power of end hosts. Multicast is a transmission mode in which the information can be efficiently sent to one or more receivers at the same time (but not to all end hosts in the network) without passing through a given link several times. Inter-domain multimedia multicast applications have clearly demonstrated a promising market but their deployment through using IP multicast has been delayed for more than 15 years. However, research and implementation of overlay multicast has become popular recently: this is a data structure that provides multicast functionality for high-rate multimedia transmission in the overlay networks through unicast connections among group members in the underlying layers. In contrast to IP multicast, the basic difference in overlay multicast is that packet replication and forwarding takes place at end hosts instead of network routers. This difference enables overlay multicast to address the problems of the deployment complexity, QoS control and routing scalability relating to IP multicast technology.

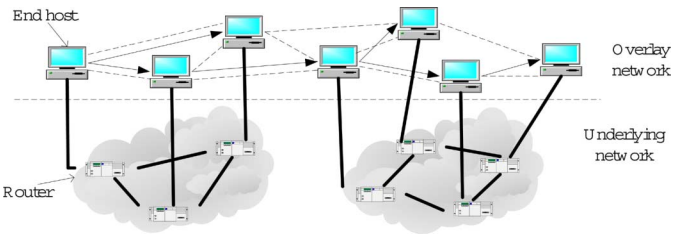

Fig. 1. Overlay network. The black straight lines show the realistic link connections among different network components. The arrow lines illustrate the overlay paths of a tree topology. The dotted lines illustrate the overlay paths of an m-D mesh topology.

Multicast functionality in overlay multicast is implemented in the application layer through constructing an overlay multicast topology that is on top of, but different from, the underlying physical topology. An overlay topology is composed of a set of virtual overlay paths: each path covers several underlying physical links and connects two different end hosts directly in the application layer. The arrow lines in Fig. 1 show several overlay paths which usually form two topologies: r-ary tree and m-D mesh. An r-ary tree is a rooted tree in which each end host has no more than r child members. A full r-ary tree is an r-ary tree on which each end host has either 0 or r child members (n.b. we use the terms group member and end host interchangeably in this paper). The topology constructed by the arrow lines in Fig. 1 is a 2-ary tree topology. An m-D mesh is generated by partitioning an m-D Cartesian space among all end hosts such that every end host "owns" its individual, distinct zone within the overall space. The topology constructed by the dotted lines in Fig. 1 is an m-D mesh topology. The r-ary tree and m-D mesh are two distinctly different topologies that directly impact the performances of overlay multimedia multicast established on them.

As will be introduced in Section II, current studies on overlay multicast [1]–[17] mostly focus on designing protocols to achieve expected performances (e.g., short multicast delays). Few studies consider overlay multicast in theory. References [18]–[20] studied IP multicasting cost in terms of the number

Manuscript received April 17, 2008; revised January 04, 2009. First published May 05, 2009; current version published May 15, 2009. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Ling Guan.

W. Tu and P. S. Excell are with the School of Computing and Communications Technology, Glyndwr University, Wrexham, North East Wales, U.K. (e-mail: w.tu@glyndwr.ac.uk; p.excell@glyndwr.ac.uk).

X. Jin is with the Systems Technology Group, Oracle USA. Inc., Redwood Shores, CA 94065 USA (e-mail: xing.jin@oracle.com).

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/TMM.2009.2017623

of links used to connect group members in a multicast delivery tree. Reference [21] characterized the performance penalty for the overlay multicast tree as compared to IP multicast. In [22], a theoretical study of the influence of topology selection on the overlay multicast was presented for the first time. We analyzed the worst performance of several important metrics for two specific topology models: *hierarchical cluster NICE tree* and *m-D CAN-based mesh*, and presented theorems that observe the performance difference between NICE and CAN-based multicast through an analysis based on overlay networks. The present paper extensively develops the topic, and fundamentally presents the following differences and contributions: 1) use of more general tree and mesh topologies to investigate multimedia overlay multicast; 2) studies of the average performance and the performance difference, compared with the worst performance; 3) consideration not only of overlay networks but also of underlying networks in our theoretical analysis; and 4) evaluation of more direct performance (e.g., multicast delays, multimedia quality, packet loss) of multimedia applications based on our analysis and simulation observations.

Additionally, we evaluate the suitability of different multicast topologies for different multimedia applications, and also present ideas for improving "bad" performance for each kind of overlay topologies through trade-off studies. To the best of our knowledge, this is one of the first explorations and evaluations of the influence of different overlay topologies on multimedia multicast applications. It is believed that this work is useful for protocol design of target multimedia applications and further investigations of multimedia multicast functions.

In [9], *stress*, *stretch*, and *comparable robustness* were defined as the intuitive criteria to evaluate the quality of overlay multicast. *Stress* counts the number of identical packets sent by the protocol. *Stretch* is the ratio of the overlay path length to the direct unicast path length between two end hosts. *Robustness* is the ability to restore normal multicast when node/link failure occurs. This paper evaluates *stress* on a per-link basis that is usually called *link stress*; it simplifies the comparison of *stretch* between tree-based and mesh-based multicast on the same underlying network, by reducing it to the comparison of the overlay path length that generates the metric of *the number of overlay hops*; it observes robustness by *the number of shortest paths* showing the possibility and efficiency to restore communication failure. In accordance with these metrics, we analyze *the worst performance*, *the average performance*, and *the performance difference*. *worst performance* shows the mostly "bad" quality that a multicast group can cause. *Average performance* statically observes the behavior of each member in the group. *Performance difference*, newly defined to observe the difference between *the worst* and *the average performance* of multimedia multicast, shows how much the performance varies among different group members. Based on the proposed metrics and performance, we implement a set of theoretical analyses which achieves the following contributions for multimedia multicasting in a group of n members.

- The worst number of overlay hops of r-ary tree-based multicast is $\log_{\check{r}}^{[1+n\check{r}-n]}-1$, where $\check{r} \in [1,r]$ is a variable that represents the out degree of each group member; the worst number of overlay hops of m-D mesh-based multicast is $mn^{(1/m)} - m$; The average number of overlay hops of r-ary tree-based multicast is $\log_{\check{r}}^{[1+n\check{r}-n]}+\log_{\check{r}}^{[1+n\check{r}-n]}-n\check{r}/n(\check{r}-1)$; the average number of overlay hops of m-D mesh-based multicast is $m(n^{1/m}-1)/2$.
- The worst link stress of r-ary tree-based multicast is $\sqrt{1+nr-n}/r$; the worst link stress of m-D mesh-based multicast is $\sqrt{4mn(m-1)}/(2m-1)^2$; The average link stress of r-ary tree-based multicast is upper bounded by $r+4r^{-\theta}/2$, where $0 < \theta \leq 1$; the average link stress of m-D mesh-based multicast is upper bounded by $m+4m^{-\theta}/2$.
- The number of shortest paths of r-ary tree-based multicast is d^l, where d is the number of child routers that a router can connect in a physical network and l is the number of hops between the two routers who are the closest to the two end hosts respectively; the number of shortest paths of m-D mesh-based multicast is $d^l[m(m-1)+\sum_{i=1}^{m} H_i \sum_{j=1,j\neq i}^{m}(H_j-1)]$.

Apart from the theoretical analyses, we simulate video multicasting in NICE [9], CAN-based multicast [7], SDEM [23], and DSM [17] to evaluate and prove our theoretical results. In general, both the theoretical analyses and the simulation results show that tree-based multicast is good for real-time and interactive multimedia multicast with a single source and a large number of receivers; mesh-based multicast, due to its even and multi-path distribution, suits to reliable multi-source multimedia multicast in which network users are distributed densely.

The paper is organized as follows. Section II introduces the related work. Section III introduces the analysis models and the performance models. Section IV presents theorems on *the number of overlay hops* for r-ary tree and m-D mesh overlay multicast. Section V presents theorems on *the link stress*. Section VI analyzes *the number of shortest paths* for the two types of overlay multicast. Simulation observations are implemented in Section VII. Section VIII presents our evaluation and insights into tree-based and mesh-based multimedia multicast. Section IX concludes the paper.

II. Related Work

Current studies on overlay multicast mostly focus on protocol design to achieve expected performance (e.g., short multicast delays). Based on the overlay topology constructed by routing schemes, protocols can be classified as tree-based [2], [4]–[6], [9]–[16] and mesh-based [1], [3], [7], [8], [17] overlay multicast. These studies give little justification for their particular choices of multicast topologies. We are going to study the influence of tree and mesh topologies on multimedia multicast applications in theory.

Few studies considered multicast in theory. Chuang *et al.* [18] presented the well-known Chuang-Sirbu scaling law that is a cost-based approach to multicast pricing, to facilitate efficient and equatable resource allocation between traffic types. Through calculation and extensive simulations over a range of networks, the number of links in a multicast delivery tree connecting a random source to m random and distinct network sites is $L(m) \propto m^{0.8}$. Phillips *et al.* [19] examined approximately $L(m)$ for a k-ary complete tree topology, and concluded that

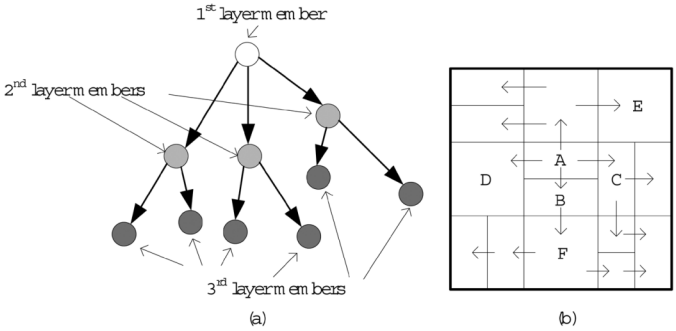

Fig. 2. Architectures of (a) an r-ary tree routing and (b) a 2-D mesh-based flooding.

$L(m)$ grows nearly linearly with m. Adjih *et al.* [20] re-examined the analysis of Phillips *et al.* through replacing the k-ary complete tree topology by a k-ary self-similar tree with similarity factor $0 < \theta \leq 1$. They proved $L(m) \sim m^{(1-\theta)}$ which provides a plausible explanation of the multicast power law. Also, [20] proved that the power law holds in more general conditions for general trees. In [21], Fahmy *et al.* characterized the performance penalty for overlay multicast tree as compared to IP multicast. The overlay tree cost was quantified as $H(n)/U(n) \propto n^{0.9}$ for small n, where $H(n)$ is the total number of hops in all overlay links, $U(n)$ is the average number of hops on the source to receiver unicast paths, and n is the number of members in the overlay multicast session. In [22], we studied the influence of topology selection to the overlay multicast performance. Two popular overlay multicast protocols: NICE tree and CAN-based mesh were analyzed, compared and evaluated through theorems on the worst performance of several important multicast metrics.

III. Analysis Models

A. Topology Models

r-ary tree-based multicast arranges group members into an overlay tree in which each member has at most r direct child members. As shown in Fig. 2(a), each forwarder connects with $\check{r}$ ($\check{r} \in [1, r]$) group members who are the closest to the forwarder in physical locations. Group members who have no child member are called leaf nodes. Assume that the tree height is h and each group member has a layer number based on its overlay hop distance from the root. Then as illustrated in Fig. 2(a), the layer number of the root node is 1, the root's children are in the layer 2, ..., the leaf members, being $h-1$ overlay hops from the root, are in the hth layer. A full r-ary tree is a tree structure on which each node has either 0 or r direct child members. The arrow lines in Fig. 2(a) show an example of tree multicasting. Each forwarder sends the received packets to its upper-layer neighbors, who then forward the packets to their child members in the immediate higher layer. Multicasting stops when all members receive the packets. Overlay multicast trees are usually source trees. The tree root is an end host that has limited capacities and therefore easily suffers from bottleneck when serving multiple high-rate multimedia streams if it is on a shared tree.

m-D mesh-based multicast arranges group members into a m-D Cartesian space in which each member "owns" an individual, distinct zone within the overall space. An example of 2-D mesh-based multicast is shown in Fig. 2(b). The location of each member is usually labeled by the Cartesian coordinates of the center point of the zone where the member locates. Group members who are closer to each other locate in the mesh zones that have smaller Euclidean distances. To deal with a newly joined node, the closest group member splits its mesh zone into two halves and leaves one half space for the new member. But if a current member wants to leave the multicast group, its mesh zone will be taken over by its closest group member. For example, in Fig. 2(b), if A wants to leave, B (A's closest neighbor in the space) takes over A's zone and changes its location to $(X_B, Y_A + Y_B/2)$. When A joins in the space again, B generates a half-zone space to A with the locations (X_A, Y_A) to A and (X_B, Y_B) to B. The arrow lines in Fig. 2(b) illustrate the paths of flooding that is an intuitive and classical routing without refining mesh topologies. Packets transmit to neighboring nodes who then forward packets to their neighbors until all of the group members have received the packets.

B. Performance Models

The worst performance is defined as the worst-case quality that a multicast group can cause. Denote a metric as M. Use $\ddot{M}$ to represent M's *worst performance* caused by overlay multimedia multicasting:

$$\ddot{M} = \begin{cases} \max\{M_i, i \in [0, n-1]\}, & \text{when a small M is expected} \\ \min\{M_i, i \in [0, n-1]\}, & \text{when a large M is expected} \end{cases}$$

where M_i is the ith member's M value, and n is the group size.

The average performance of M, denoted as $\bar{M}$, is defined as the average value of the summation of all group members' M values. That is

$$\bar{M} = \frac{\sum_{i=0}^{n-1} M_i}{n}. \tag{1}$$

$\bar{M}$ statistically reflects a multicast group's performance along M.

The performance difference of M, denoted as ΔM, is defined as the difference between M's *worst and average* performance. Namely

$$\Delta M = \ddot{M} - \bar{M}. \tag{2}$$

ΔM observes the performance variation at different group members. By calculating ΔM, we are able to know how balanced performance that a multimedia multicast system can achieve.

IV. Number of Overlay Hops

This section studies the *stretch* performance of tree-based and mesh-based multicast. An example of *stretch* is in Fig. 3(a). The multimedia transmission from E to D through the overlay topology generates a *stretch* of $8/3 \approx 2.67$, where 8 and 3 are the overlay path length and the unicast path length from

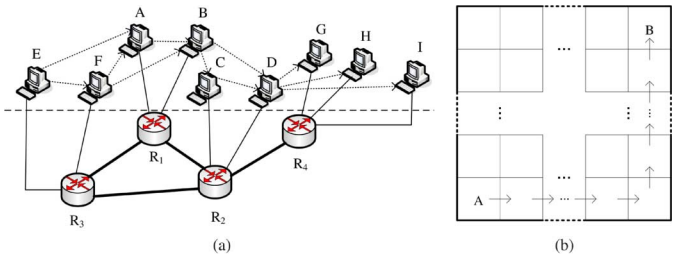

Fig. 3. (a) Overlay multicast routing with the dotted arrow lines. (b) Path with the worst number of overlay hops in a 2-D mesh.

E to D, respectively. For multimedia multicasting in different overlay topologies but based on the same underlying network, the comparison of the *stretch* can be indicated by the comparison of the overlay path length since the calculations of *stretch* for tree-based and mesh-based multicast have the same denominators. A popular metric to measure the length of overlay paths is *the number of overlay hops*. Hence, we analyze *the number of overlay hops* here.

Theorem 1: For a group of n members who construct an r-ary overlay tree, when n is large enough, the worst number of overlay hops in r-ary multimedia multicasting is $\ddot{OH}_{tree} = \log_{\check{r}}^{[1+n\check{r}-n]} - 1$, where $\check{r} \in [1, r]$ represents the out degree of each on-tree forwarder.

Proof: Denote the out degree of each forwarder on the r-ary tree as $\check{r}$ ($\check{r} \in [1, r]$). Since overlay multicast employs source tree structure, multimedia packets experience 0 hop to the root, 1 hop to the end hosts in the 2nd layer, ..., $(i-1)$ numbers of overlay hops to the end hosts in the ith layer, ..., and $(h-1)$ numbers of overlay hops to the end hosts in the hth layer (i.e., the highest layer). It means that *the worst number of overlay hops* is experienced by the end hosts in the $(h-1)$th layer. To calculate h, based on the number of group members in each layer, we have $n = 1 + \check{r} + \check{r}^2 + \cdots + \check{r}^{i-1} + \cdots + \check{r}^{h-1}$. It can be inferred that $h = \log_{\check{r}}^{[1+(\check{r}-1)n]}$. Therefore, *the worst number of overlay hops* is $\ddot{OH}_{tree} = \log_{\check{r}}^{[1+n\check{r}-n]} - 1$. Q.E.D.

Theorem 2: For a group of n members who construct r-ary tree multicasting, when n is large enough, the average number of overlay hops in the r-ary multimedia multicasting is $\bar{OH}_{tree} = \log_{\check{r}}^{[1+n\check{r}-n]} + \log_{\check{r}}^{[1+n\check{r}-n]} - n\check{r}/n(\check{r}-1)$.

Proof: Based on (1), we first need to calculate the total numbers of overlay hops from each end host to the root. According to the proof of Theorem 1, an member in the ith layer has $(i-1)$ numbers of overlay hops to the tree root and there are $\check{r}^{(i-1)}$ members in the ith layer. Hence, the total numbers of overlay hops from each member in the ith layer to the root is $(i-1)\check{r}^{(i-1)}$. Then the total numbers of overlay hops from each member to the root is $S = \sum_{i=1}^{h}(i-1)\check{r}^{(i-1)} = 0 \cdot \check{r}^0 + \cdots + (h-1)\cdot\check{r}^{(h-1)} = \check{r} - \check{r}^h/(1-\check{r})^2 - (h-1)\check{r}^h/(1-\check{r})$. Hence, $\bar{OH}_{tree} = S/n = \log_{\check{r}}^{[1+n\check{r}-n]} + \log_{\check{r}}^{[1+n\check{r}-n]} - n\check{r}/n(\check{r}-1)$.[1] Q.E.D.

Corollary 1: For a multicast group with n members, the performance difference of the number of overlay hops in an r-ary tree is upper bounded by $\Delta OH_{tree} = 1/\check{r} - 1 - 1/n(\check{r}-1) - (\log_{\check{r}}^{[1+n\check{r}-n]} - 1)/n(\check{r}-1)$.

Proof: Based on Theorems 1 and 2, the performance difference of *the number of overlay hops* is $\Delta OH_{tree} = 1/\check{r} - 1 - 1/n(\check{r}-1) - (\log_{\check{r}}^{[1+n\check{r}-n]} - 1)/n(\check{r}-1)$. Q.E.D.

We now analyze *the number of overlay hops* for m-D mesh overlay multimedia multicast.

Theorem 3: For a multicast group with n members that are mapped into an m-D mesh $d_1 \times d_2 \times \ldots \times d_m$, where $d_i (i \in [1, m])$ is the number of mesh zones along the ith dimension, if $d_1 = d_2 = \cdots = d_m$, the worst number of overlay hops is $mn^{1/m} - m$.

Proof:

1) We first consider a 2-D mesh. Since each member only sends packets to the members in its adjacent zones, as illustrated in Fig. 3(b), the worst number of overlay hops is generated when packets transmit from A to B (the arrow lines show one of the longest overlay paths). The number of overlay hops on this longest path is $(d_1 - 1) + (d_2 - 1)$. Considering the theorem condition, we have $d_1 = d_2 = d$. It infers that $(d_1 - 1) + (d_2 - 1) = 2d - 2$. For a large group, $d_1 \cdot d_2 \approx n$. The number of overlay hops is $2n^{1/2} - 2$.
2) For an m-D mesh, similar to a 2-D mesh, packets starting from the bottom-left zone and ending at the top-right zone experience the most numbers of overlay hops $\ddot{OH}_{mesh}$. We have $\ddot{OH}_{mesh} = (d_1 - 1) + (d_2 - 1) + \cdots + (d_m - 1)$. Since $d^m = n$, $\ddot{OH}_{mesh} = mn^{1/m} - m$.

Q.E.D.

Theorem 4: For a multicast group with n members that are mapped into an m-D mesh described in Theorem 3, the average number of overlay hops generated by mesh-based multimedia multicasting is $m(n^{1/m} - 1)/2$.

Proof: We first calculate the total number of overlay hops from the multimedia sending source to all other group members. Without loss generality, suppose the sender locates in $(0, 0, \ldots, 0)$. Any group member in $(x_1, x_2, \ldots, x_m)$ has $(x_1 + x_2 + \cdots + x_m)$ numbers of overlay hops to the sender. Hence, the total number of overlay hops from the source to all other members is $(TOH)_{mesh} = \sum_{x_1=0}^{(n^{1/m})-1} \sum_{x_2=0}^{(n^{1/m})-1} \cdots \sum_{x_m=0}^{(n^{1/m})-1} (x_1 + x_2 + \cdots + x_m)$.

To calculate $(TOH)_{mesh}$, we know each coordinate $x_i (i \in [1, m])$ has the same chance to select a number in $[0, n^{1/m} - 1]$ as its value. Namely, in the view of all dimension coordinates of all end hosts, each number in $[0, n^{1/m} - 1]$ is used as coordinates by the same times as other numbers in this range. Let the times that each number is used as coordinates by all end hosts be F. Then $(TOH)_{mesh} = F\sum_{i=0}^{(n^{1/m})-1} i = Fn^{1/m}(n^{1/m} - 1)/2$. According to the knowledge of probability and statistics, since each number has equal chance to be used as coordinates, we have $F = mn^{m-1/m}$. It infers that *the average number of overlay hops* is $\bar{OH}_{mesh} = mn^{m-1/m}n^{1/m}(n^{1/m} - 1)/2/n = m(n^{1/m} - 1)/2$. Q.E.D.

Corollary 2: For a group of n members mapped into an m-D mesh described in Theorem 3, the overlay hop difference of mesh-based multimedia multicasting is $m(n^{1/m} - 1)/2$.

[1] $\bar{OH}_{tree} = S/n = \check{r} - \check{r}^h/(\check{r}-1)^2 n + h - 1/n[1/\check{r} - 1 + n] = h - 1 + 1/n(\check{r}-1) - 1/\check{r} - 1 + (\log_{\check{r}}^{[1+n\check{r}-n]} - 1)/n(\check{r}-1) = \log_{\check{r}}^{[1+n\check{r}-n]} + 1/(\check{r}-1)n + (\log_{\check{r}}^{[1+n\check{r}-n]} - 1)/n(\check{r}-1) - \check{r}/\check{r} - 1$.

Proof: According to Theorems 3 and 4, $\Delta OH_{mesh} = \ddot{OH}_{mesh} - \bar{OH}_{mesh} = mn^{1/m} - m - m(n^{1/m}-1)/2 = m(n^{1/m}-1)/2$. Q.E.D.

Remark 1: Based on Theorems 1 and 3, when n is large enough, $\ddot{OH}_{tree} = \log_{\check{r}}^{[1+nr-n]} - 1$ and $\ddot{OH}_{mesh} = mn^{1/m} - m$. $\ddot{OH}_{tree}$ logarithmically increases with the increment of $n^{1/m}$, while $\ddot{OH}_{mesh}$ linearly increases with the increment of $n^{1/m}$. Since $\ddot{OH}_{tree}/\ddot{OH}_{mesh} = \log_{\check{r}}^{[1+nr-n]} - 1/mn^{1/m} - m = \log_{\check{r}}^{(1+n\check{r}-n)^{1/m}} - 1/m/n^{1/m} - 1 \ll 1$, the worst *stretch* performance in tree-based multicast is much less than that in mesh-based multicast when multicasting multimedia. This comparison becomes more obvious when the group size increases. For *the average number of overlay hops*, when n is large enough, $\bar{OH}_{tree} = \log_{\check{r}}^{[1+n\check{r}-n]} + 1/(\check{r}-1)n + (\log_{\check{r}}^{[1+n\check{r}-n]} - 1)/n(\check{r}-1) - \check{r}/\check{r}-1$ and $\bar{OH}_{mesh} = mn(n^{1/m}-1)/2$. To compare them, we have $\bar{OH}_{tree}/\bar{OH}_{mesh} = 2\log_{\check{r}}^{[1+n\check{r}-n]^{1/m}}/n(n^{1/m}-1) + 2/mn^2(n^{1/m}-1)(\check{r}-1) + 2\log_{\check{r}}^{[1+n\check{r}-n]^{1/m}}/n^2(n^{1/m}-1)(\check{r}-1) - 2/mn^2(n^{1/m}-1)(\check{r}-1) - 2\check{r}/mn(n^{1/m}-1)(\check{r}-1) \ll 1$. Combining the comparison of *the worst number of overlay hops*, we conclude that r-ary overlay tree multimedia multicasting is good for multimedia applications that have stringent delay requirement such as live conference, online games, etc. The results in Corollaries 1 and 2 show $\Delta OH_{tree}/\Delta OH_{mesh} = 21/\check{r}-1 - 1/n(\check{r}-1) - (\log_{\check{r}}^{[1+n\check{r}-n]} - 1)/n(\check{r}-1)/m(n^{1/m}-1) \ll 1$. r-ary tree generates smaller delay difference than m-D mesh which indicates that tree topology holds the promise to implement interactive multimedia applications.

V. Link Stress

Link stress is a metric to evaluate the *stress* performance based on per link for overlay multimedia multicast. An example of link stress is illustrated in Fig. 3(a). If E is a multicast sender, packets need to pass through the underlying link $\langle E \rightarrow R_3 \rangle$ twice to reach A and F respectively. Essentially, *link stress* unnecessarily consumes network resources but is unavoidable in overlay multicast due to packet replication and forwarding at end hosts. This section compares and evaluates *the link stress* for r-ary tree and m-D mesh overlay multicast.

Theorem 5: For a group of n members who construct an r-ary overlay tree to multicast multimedia streams, when n is large enough, the worst link stress of physical links is upper bounded by $\sqrt{1+nr-n}/r$.

Proof: Physical links used to carry multimedia multicast traffic can be classified into *local links* and *backbone links*. *Local links* connect each end host to its closest network component (e.g., a hub). The number of identical packet copies that each end host replicates and then forwards into its *local link* is decided by the number of direct child members that the end host has in the tree topology. In an r-ary overlay tree, the most number of direct child members that an end host can have is r. It infers that *the worst link stress* among all *local links* is r.

Backbone links connect intermediate network components such as routers in IP networks. For each backbone link in the tree multicast system, its link stress is decided by the number of members who receive the same packets from the backbone link. For an r-ary tree, end hosts in a higher layer use less common backbone links to receive packets because the *backbone links* that they are using are broadly distributed in different areas close to these different end hosts. On the other hand, end hosts in a lower layer have less numbers of direct child members to send the identical packets. Therefore, *the worst link stress* should appear in a backbone link which connects end hosts located in some middle layer. Reference [21] proves that $r^{(h-i)\theta} - 2$ underlying links exist between two end hosts in the $(i-1)$th and the ith layers, respectively, where θ is a positive number not greater than 1. It shows that at most $r^{(i-1)}[r^{(h-i)\theta} - 2]$ underlying links connect the $(i-1)$th and the ith layers. Meanwhile, there are totally $r^{(i-1)}$ identical packets sent from the $(i-1)$th layer to the ith layer. Therefore, *the worst link stress* should appear in a link connected to a layer which contributes the maximum value for the function $y(i) = k^{2(i-1)}[k^{(h-i)\theta} - 2]$. We have $y(i)' = (2-\theta)k^{(h-i)\theta}k^{2(i-1)}ln^k - 4k^{2(i-1)}ln^k$ which shows $i = h - 1/\theta log_k^{[4/2-\theta]}$ is the number of the layer at which the *worst link stress* is caused. Since $r^{h/4\theta} - 2 \geq 0$ for an inter-domain tree, we obtain $1/\theta log_k^{[4/2-\theta]} \leq h/2$. Therefore, $i \geq \lceil h/2 \rceil$. Based on the analysis, we have that *the worst link stress* in the backbone network is $r^{\lceil h/2 \rceil - 1}$.

Considering both the *local links* and the *backbone links*, we have $\ddot{LS}_{tree} = \max\{r, r^{(\lceil h/2 \rceil - 1)}\} = \sqrt{1+nr-n}/r$.[2] Q.E.D.

Theorem 6: For a group of n members who construct an r-ary multicast tree, when n is large enough, the average link stress of physical links is upper bounded by $r + 4r^{-\theta}/2$, where $0 < \theta \leq 1$.

Proof: We first analyze *the average link stress* for *local links*. There are $r^{(i-1)}$ end hosts in the ith layer and the tree height is $h = log_r^{(1-n+nr)}$ if the constructed tree is a full r-ary tree. It infers that the number of leaf nodes is $r^{h-1} = r^{log_r^{(1-n+rn)}}/r = 1 - n + rn/r$ and then the number of non-leaf nodes is $n - 1/r$. Since each non-leaf node has r direct child members, *the average link stress* of *local links* is $r \cdot n - 1/r/n - 1/r = r$.

We analyze *the average link stress* of *backbone links* now. As we known that the end hosts in the $(i-1)$th layer need to send $r^{(i-1)}$ identical packet copies to the end hosts in the ith layer, then *the average link stress* of *backbone links* in the transmission area between the $(i-1)$th layer and the ith layer is r^{i-1}/L_i, where L_i is the number of physical links connecting the $(i-1)$th and the ith layers. Further, it has been analyzed in the proof of Theorem 5 that, according to [21], $L_i = r^{(h-i)\theta} - 2$ and the most number of underlying links between end hosts in the $(i-1)$th layer and the ith layer is $r^{i-1}(r^{(h-i)\theta} - 2)$. With these results, we have $\sum_{i=1}^{h-1} r^{i-1}/L_i = \sum_{i=1}^{h-1} r^{i-1}/r^{i-1}[r^{(h-i)\theta} - 2] \leq 3/r^{\theta} - 1$.[3]

Consider both the *local links* and the *backbone links*, we obtain $\bar{LS}_{tree} \leq r + 4r^{-\theta}/2$. Q.E.D.

[2]For a full r-ary tree, the number of end hosts in the ith layer is $r^{(i-1)}$. It shows that the tree height is $h = log_r^{(1-n+rn)}$. When n is large enough, $h = log_r^{(r-1)n}$.

[3]$\sum_{i=1}^{h-1} r^{i-1}/r^{i-1}[r^{(h-i)\theta} - 2] \leq \sum_{i=1}^{h-1} 3/r^{(h-i)\theta} = 3/r^{h\theta} \cdot r^{h\theta} - r^{\theta}/r^{\theta} - 1$. Since $h = \log_r^{(r-1)n}$, it can be inferred that $\sum_{i=1}^{h-1} r^{i-1}/L_i \leq 3/(r-1)^{\theta}n^{\theta} \cdot (r-1)^{\theta}n^{\theta} - r^{\theta}/r^{\theta} - 1 \approx 3/r^{\theta} - 1$.

Corollary 3: For a group of n members who construct a full r-ary multicast tree, when n is large enough, we have $\Delta LS_{tree} \geq (n/r)^{1/2} - r/2$.

Proof: According to Theorems 5 and 6, $\Delta LS_{tree} = \ddot{LS}_{tree} - \bar{LS}_{tree} \geq \sqrt{1+nr-n}/r - r + 4r^{-\theta}/2 \approx (n/r)^{1/2} - r/2$.[4]

Theorem 7: For a group of n members who construct an m-D mesh overlay topology, when n is large enough, the worst link stress in the multicast system is $\ddot{LS}_{mesh} = \sqrt{4mn(m-1)}/(2m-1)^2$.

Proof: We analyze *the worst link stress* in *local links* and *backbone links*, respectively. For *local link*, *the worst link stress* is $2m$ due to each end host thinks of its topology neighbors as its downstream neighbors, except for the neighbor from whom packets come. For *backbone links*, similar as we have analyzed in Theorem 5, *the worst link stress* in the underlying network should appear when the end hosts who have $\lceil p/2 \rceil$ numbers of hops to the sending source receive multimedia packets multimedia packets from their upstream forwarders. It indicates that *the worst link stress* of *backbone links* is $2m(2m-1)^{\lceil p/2 \rceil - 1}$.

Considering both of the *local links* and the *backbone links*, we have $\ddot{LS}_{mesh} = \max\{2m, 2m(2m-1)^{\lceil p/2 \rceil - 1}\} = \sqrt{4mn(m-1)}/(2m-1)^2$.[5] Q.E.D.

Theorem 8: For a group of n members in m-D mesh multimedia multicasting, when n is large enough, the average link stress is $\bar{LS}_{mesh} \leq m + 4m^{-\theta}/2$, where $0 < \theta \leq 1$.

Proof: For the *local links*, *the average link stress* is $\sum_{i=1}^{2m}(i \times 1/2m + 1) = m$ because the number of downstream members that each end host can connect in the overlay topology is uniformly distributed in $[0, 2m]$. For the *backbone links*, similar to the analysis of r-ary tree routing, *the average link stress* is $\sum_{i=1}^{p-1} m^{i-1}/L_i \leq \sum_{i=1}^{p-1} m^{i-1}/m^{i-1}(m^{(p-i)\theta} - 2) \leq 3/m^{\theta} - 1$.

Considering the two types of links, *the average link stress* of m-D mesh multicasting is $\bar{LS}_{mesh} = m + 3/m^{\theta} - 1/2 \leq m + 4m^{-\theta}/2$. Q.E.D.

Corollary 4: For a group of n members who construct an m-D mesh overlay topology, when n is large enough, the link stress performance difference is lower bounded by $\Delta LS_{mesh} \geq m^{\theta}\sqrt{n} - m^{2+\theta} - 4m/2m^{(1+\theta)}$.

Proof: Based on the results of Theorems 7 and 8, we have $\Delta LS_{mesh} \geq \sqrt{4mn(m-1)}/(2m-1)^2 - m + 4m^{-\theta}/2 \approx m^{\theta}\sqrt{n} - m^{2+\theta} - 4m/2m^{(1+\theta)}$.[6] Q.E.D.

Remark 2: Based on Theorems 5 and 7, when n is large enough, $\ddot{LS}_{tree}/\ddot{LS}_{mesh} = \sqrt{1+nr-n}/r(2m-1)^2/\sqrt{4m(m-1)} = \sqrt{4m(m-1)r - 1/r^2} + \sqrt{r-1}/r\sqrt{4m(m-1)} \approx 2m/\sqrt{r} + \sqrt{r-1}/r\sqrt{4m(m-1)} > 1$. It indicates that tree multicasting is easier to suffer from bottleneck than mesh multicasting. From theorems 6 and 8, $\bar{LS}_{tree}$ is upper bounded by $r + 4r^{-\theta}/2$ and $\bar{LS}_{mesh}$ is upper bounded by $m + 4m^{-\theta}/2$. To compare them, we analyze the function $f(x) = x + 4x^{-\theta}/2$. It can be inferred that $f(x)' = 1/2 - 2\theta/x^{(1+\theta)}$. We have $2/x^{1+\theta} < 1/2$ because $x \geq 2$, which proves that the function $f(x)$ is monotonously increased with the variable x. The comparison between $\bar{LS}_{tree}$ and $\bar{LS}_{mesh}$ depends on the comparison between r and m. In modern networks, $r > m$.[7] Thus, $\bar{LS}_{tree} > \bar{LS}_{mesh}$. Based on corollaries 3 and 4, $\Delta LS_{tree} - \Delta LS_{mesh} = (n/r)^{1/2} - r/2 - m^{\theta}\sqrt{n} - m^{2+\theta} - 4m/2m^{(1+\theta)} = \sqrt{n}/2m\sqrt{r}(2m - \sqrt{r}) - (r-m)/2 + 2/m^{\theta}$. Usually, in practical systems, $2m > \sqrt{r}$.[8] Hence, we have $\sqrt{n}/2m\sqrt{r}(2m - \sqrt{r}) - r - m/2 > 0$ and then $\Delta LS_{tree} > \Delta LS_{mesh}$. The results show that multimedia multicast along m-D mesh topology provides more evenly throughput distribution than multimedia multicast along r-ary tree topology does. At the same time, since $(\ddot{LS}_{tree} - \bar{LS}_{mesh}) < (\ddot{LS}_{tree} - \ddot{LS}_{mesh})$, the group members not affected by bottleneck in tree multicast experience much less traffic load burden than the group members in mesh multicast.

VI. Number of Shortest Paths

In multimedia multicast communications, group members dynamically change their sate which may interrupt the shortest delay transmission. The number of shortest paths observes the robustness of routing protocols in maintaining the short delay performance when dynamic path alterations take place. It is known that mesh-based multicast has more numbers of shortest paths than tree-based multicast. This section quantitatively develops the understanding of the shortest path advantage in mesh-based multicasting, and fundamentally investigates the number of shortest paths between end hosts with different distances. We believe that the analysis is useful for planning reliable multicast while efficiently employing network resources since multimedia transmission always requires high network capacities.

Theorem 9: For a group of n members that construct an r-ary overlay multicast tree, the most number of the shortest paths between any two members is d^l, where l is the number of hops between the two routers who are the closest to the two members, respectively, and d is the number of the child routers that a router can connect in the physical network.

Proof: As illustrated by the arrow lines in Fig. 2(a), each pair of end hosts in the r-ary tree has only 1 shortest overlay path. We now analyze how many shortest paths in the physical network can represent the only 1 shortest overlay path between the two end hosts through using the example in Fig. 4. As labeled in Fig. 4, we classify the routers that can be used to serve the overlay path from E_1 (E_1 is the upstream forwarder) to E_2 into different categories based on their hop distances to the E_1's closest router R_1. The black lines show the links composing of

[4] $\Delta LS_{tree} \geq \sqrt{1+nr-n}/r - r + 3r^{-\theta} + [(r-1)n]^{1-\theta}/2 \approx (n/r)^{1/2} - r/2$. Q.E.D.

[5] It is easy to prove that $p = log^{n(m-1)/m}_{(2m-1)}$ when n is large enough.

[6] $\sqrt{4mn(m-1)}/(2m-1)^2 - m + 4m^{-\theta}/2 \approx \sqrt{mn(m-1)} - m(m^2-m)/2m^2 - 2m - 2/m^{\theta} \approx \sqrt{n} - m^2/2m - 2/m^{\theta}$.

[7] The bandwidth of *local links* is usually 100 Mbps and the video transmission rates are around 1.5 Mbps. An end host should be able to connect up to ten direct child members even when carrying other applications, while $m = 3$ is usually enough for most of mesh multicasting applications.

[8] A 3-D mesh is enough for most of applications. In this situation, only a tree with $r > 36$ makes $2m < \sqrt{r}$. $r > 36$ puts too much traffic load burden to one end host who needs to output high-rate multimedia streams especially in multi-stream applications.

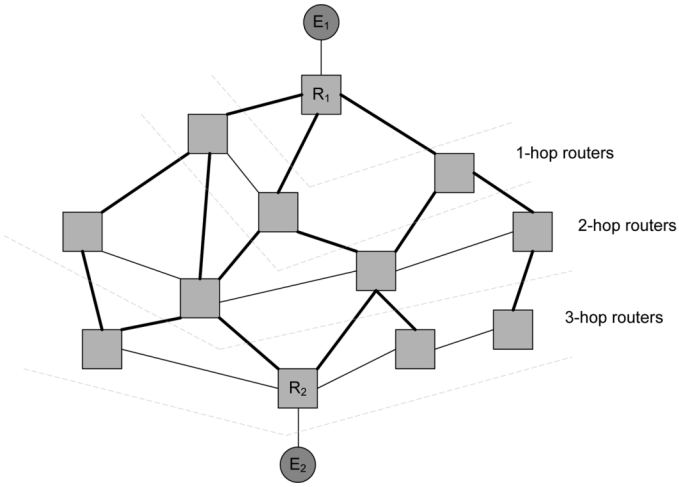

Fig. 4. Analysis of the number of physical links covered by one overlay path.

all of the shortest paths between E_1 and E_2. The figure illustrates that there is one shortest underlying path to R_1 for each 1-hop end host, at most d shortest underlying paths to R_1 for a 2-hop end host. Based on this observation, through using the recursion and induction, it can be inferred that there are $d^{(l-1)}$ shortest underlying unicast paths that can be utilized by the only one shortest overlay path between two end hosts when there are l numbers of hops between the two end hosts' closest routers, respectively. That is to say, the number of shortest paths between any two end hosts in the r-ary tree is $NSP_{tree} = d^{(l-1)}$ if the two end hosts have l numbers of underlying hops. Q.E.D.

Theorem 10: Given a group of n members that construct an m-D mesh $d_1 \times d_2 \times \ldots \times d_m$, for two hosts locating in the zones A and B, if $d_1^B = d_1^A + H_1, d_2^B = d_2^A + H_2, \ldots, d_m^B = d_m^A + H_m$, the most number of shortest paths between A and B is $d^{(l-1)} \times [m(m-1) + \sum_{i=1}^{m} H_i \sum_{j=1, j\neq i}^{m} (H_j - 1)]$.

Proof:

1) For simplicity, we first prove that the number of shortest overlay paths between A and B in a 2-D mesh is $2 + H_1(H_2 - 1) + H_2(H_1 - 1)$.
   (1.1) If A takes the paths along the first dimension first, the shortest paths to B that turn up at the zone $(d_1^A + 1, d_2^B)$ are illustrated by the bunch of arrow lines in Fig. 5(a). Obviously, the number of this bunch of shortest paths is H_2. Similarly, if A takes the paths along the first dimension first, the numbers of shortest paths when the paths turn up at the zones $(d_1^A + 2, d_2^A), (d_1^A + 3, d_2^A), \ldots, (d_1^A + H_1 - 1, d_2^A)$ are H_2, respectively. If A takes the path along the first dimension first, the total number of shortest paths turning up at the zone $(d_1^A + H_1, d_2^A)$ is 1. We illustrate this path with the black arrow line in Fig. 5(a). Therefore, the total number of shortest paths that are firstly along the first dimension is $1 + H_2(H_1 - 1)$.
   (1.2) Similarly, when A takes the paths along the second dimension first, as illustrated in Fig. 5(b), the total number of shortest paths that are firstly along the second dimension is $1 + H_1(H_2 - 1)$.
   According to (1.1) and (1.2), the number of shortest paths between A and B through flooding in a 2-D mesh is $2 + H_1(H_2 - 1) + H_2(H_1 - 1)$.
2) We now consider an m-D mesh $d_1 \times d_2 \times \ldots \times d_m$. From the perspective of mesh multicast, the mesh can be regarded as a disjointed 2-D mesh set $\{d_i \times d_j | i \neq j, i, j \in [1, m]\}$. It is because each host transmits packets to its one hop neighbors who have the same coordinates as the host, except for one coordinate. There are $m(m-1)$ items in the disjointed set. It can be inferred that the number of shortest overlay paths between A and B in a m-D mesh is $m(m-1) + H_1 \sum_{i=2}^{m}(H_i - 1) + \cdots + H_{m-1} \sum_{i=1, i\neq m-1}^{m}(H_i - 1) + H_m \sum_{i=1}^{m-1}(H_i - 1) = m(m-1) + \sum_{i=1}^{m} H_i \sum_{j=1, j\neq i}^{m}(H_j - 1)$.

We have proved in Theorem 9 that each overlay shortest path can use at most $d^{(l-1)}$ different shortest underlying paths to implement its transmission. Hence, we have that the number of shortest paths between any two end hosts in the m-D mesh multicast is $d^{(l-1)} \times [m(m-1) + \sum_{i=1}^{m} H_i \sum_{j=1, j\neq i}^{m}(H_j - 1)]$. Q.E.D.

Remark 3: The expressions of the most number of shortest paths present the observations: 1) when the group size increases, mesh-based multicast improves its robustness while tree-based multicast becomes more fragile; and 2) for two longer distance end hosts, mesh-based multicast generates more numbers of shortest paths while tree-based multimedia multicast takes more risk to be interrupted. Hence, mesh topology holds the promise to implement reliable multimedia transmission and to manage the multimedia groups with quite a few temporary user access. Based on Theorems 9 and 10, more insights are presented in Section VIII.

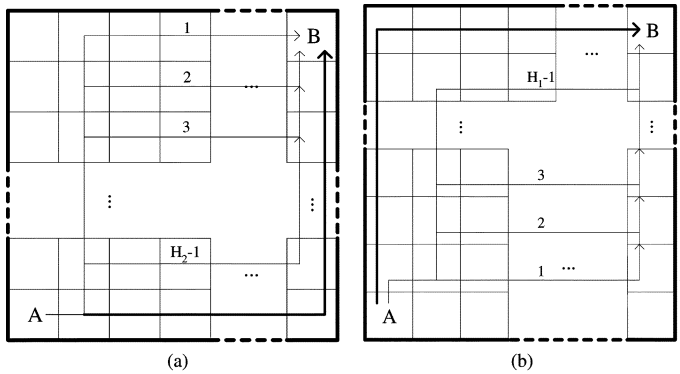

Fig. 5. Examples of the shortest paths in a 2-D mesh. (a) For the paths along the first dimension first. (b) For the paths along the second dimension first.

VII. Simulation Evaluation

This section studies overlay multicast through NS2 simulations. We simulate NICE [9] and SDEM [23] as the example protocols of tree-based multicast, and CAN-based multicast [7] and DSM [17] as the example protocols of mesh-based multicast.

Fig. 6 illustrates the simulation backbone network. All nodes in this topology are routers. Group members connect to the backbone network directly or indirectly through some intermediate network components (e.g., hubs). Links in the backbone

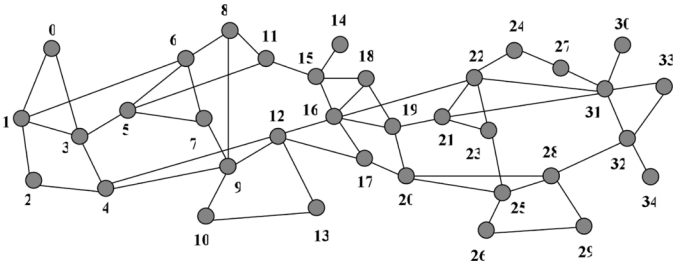

Fig. 6. Experimental backbone network.

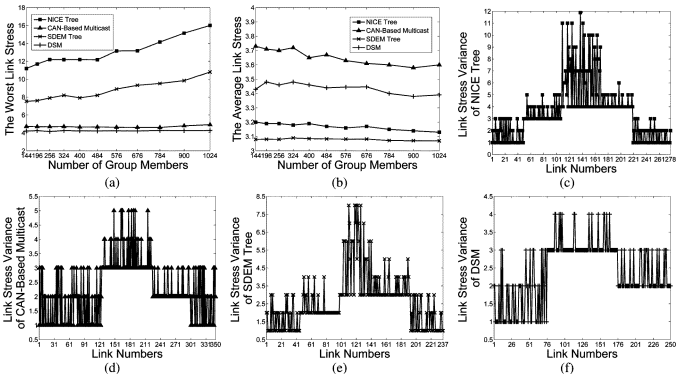

Fig. 7. Link stress performance. (a) Worst link stress. (b) Average link stress. (c)–(f) Variance of link stress at different links for the four protocols, respectively.

(local) network have 1000 Mbps (100 Mbps) bandwidth. The link propagation delays are 2 ms (1 ms) in the backbone (local) network. Simulation traffic is composed of 1.5 Mbps MPEG-1 video streams. The simulation sets $r \in [2, 8]$ which varies with different group sizes of tree-based multicast, and $m \in [2, 3]$ in mesh-based multicast. In order to smooth network instability for statically performance evaluation, each numeric value plotted in result figures is the average values of 20 runs.

Fig. 7 presents the link stress performance in the simulations. According to Fig. 7(a), tree-based multicast protocols generate larger WLS than mesh-based multicast protocols. We proved $\ddot{LS}_{tree} = \sqrt{1 + nr - n}/r$ in theorem 5 which agrees with the correlation between the worst link stress and the group size found in NICE and SDEM simulations. Based on the parameters of the group size n and r (n.b. the simulations change the value of r under different group sizes), we calculate that the discrepancies of the worst link stress between the theoretical and simulation results are in the range of $[23\%, 30\%]$ for NICE tree and $[1.5\%, 20\%]$ for SDEM tree. For m-D CAN-based flooding (DSM), the simulation results are around 4.8 (4.2) which is slightly different from the theoretical result 5. The parameters n (group size) and m (n.b. the simulations change the value of m under different group sizes) present that the worst link stress discrepancies are in the range of $[7\%, 22\%]$ for CAN-based multicast and $[3.4\%, 29\%]$ for DSM.

Fig. 7(b) plots the average link stress performance. The average link stress of tree-based and mesh-based protocols are very close and slightly decreases with the increment of group sizes. The comparison between mesh-based and tree-based protocols actually depends on the network parameters. Our

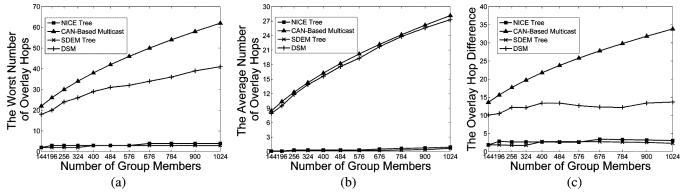

Fig. 8. Number of overlay hop performance. (a) Worst number of overlay hops. (b) Average number of overlay hops. (c) Overlay hop difference.

simulation parameters enable mesh-based protocols to generate slightly larger performance in the average link stress than tree-based protocols. For the discrepancies, they are in the range of $[9.4\%, 25\%]$ for NICE, $[6.3\%, 24\%]$ for SDEM, $[19.2\%, 27\%]$ for CAN-based multicast, and $[14.5\%, 21.3\%]$ for DSM. Fig. 7(c)–(f) presents the variance of link stress at different links generated by the four protocols. Agreeing with our analysis in Remark 2, the results in the four figures show that NICE and SDEM generate larger link stress difference than CAN-based multicast and DSM. This observation indicates that tree-based multicast introduce unbalanced links in terms of traffic load. More specifically, links connecting end hosts in some middle layers have much larger traffic load than links connecting the end hosts in other layers. It proves our idea in Remark 2 that changing the roles of end hosts in multi-source multicast helps to improve multimedia throughput.

Fig. 8 shows the performance of the number of overlay hops. The worst number of overlay hops is given in Fig. 8(a). NICE and SDEM achieve much less in the worst number of overlay hops than CAN-based multicast and DSM. The simulation results closely meet the theoretical prediction. For example, if we put $\check{r} = 6$ and $n = 676$ into $log_{\check{r}}^{[1+n\check{r}-n]}$, we have $O\ddot{H}_{tree} \approx 4.53$ which is almost equal to the simulation results for both NICE and SDEM. Further, the discrepancies between the theoretical and simulation results are in the ranges of $[12\%, 29\%]$ for NICE tree and $[20\%, 30\%]$ for SDEM tree. The numeric values shown in the CAN-based multicast and DSM curves are almost equal to the theoretical result calculated by $mn^{1/m} - m$. The discrepancies are in $[0\%, 5\%]$ for CAN-based multicast and in $[12\%, 23\%]$ for DSM. Fig. 8(b) illustrates the average number of overlay hops. As we predicted in Theorems 2 and 4, when the group size increases from 144 to 1024, the average numbers of overlay hops in mesh-based multicast increase while the average numbers of overlay hops in tree-based multicast change slightly. Furthermore, for the simulation parameters $m \in$ [2,3] and the group size

$$n = \{144,196,256,324,400,484,576,676,784,900,1024\}$$

the theoretical results should be

$$\{11,13,15,17,19,21,23,25,27,29,31\}$$

which match CAN-based multicast simulation results

$$\{8.3,10.5,13,14.8,17,19,20.5,22.5,24,26,28.3\}$$

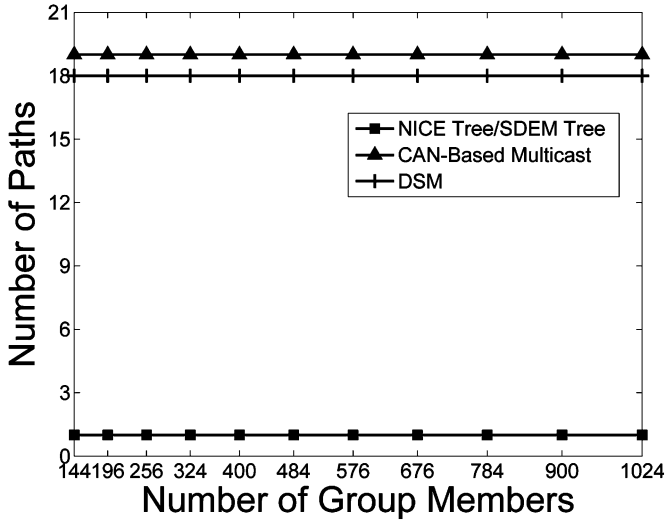

Fig. 9. Relationship between group size and number of paths.

and DSM simulation results

$$\{7.9,9.5,11.8,13.85,15.6,17.6,19.3,21.7,23.78,25.6,27.3\}$$

well. For tree-based multicast, the simulation results present the discrepancies in the ranges of $[16\%, 26\%]$ for NICE and $[18\%, 27\%]$ for SDEM. Fig. 8(c) is the performance difference of the number of overlay hops. The results prove the correctness of our theoretical analysis in Corollaries 1 and 2.

Fig. 9 observes the performance of the shortest paths. The simulation chooses two nodes in the zones (3,6) and (7,3) of a 2-D mesh. Between these two nodes, there are 19 shortest overlay paths in CAN-based multicast and 18 shortest overlay paths in DSM. 19 is also the result of Theorem 10 when these two nodes in a 2-D mesh. For NICE and SDEM, they has one shortest overlay path.

In Fig. 10, we plot multicast delays and overhead traffic that are more direct performance metrics relating to multimedia applications. As shown in Fig. 10(a), the worst delays of mesh-based multicast increase much faster than the worst delays of tree-based multicast when the group size increases. As shown in Fig. 10(b), with the increasing of group size, the average delays of mesh-based multicast increases sharply, while the average delays of tree-based multicast vary smoothly. Fig. 10(c) presents the overhead traffic generated by the four protocols when dynamic group member alterations take place. The simulation period is 100 s and includes two different phases: a join phase and a leave phase. In the join phase, 72 new hosts join in the group uniformly at random between the 30th and 48th seconds. Starting at the 60th second, 180 members leave the group uniformly. The leave procedure lasts 18 s. NICE creates the largest overhead traffic and DSM creates the smallest overhead traffic; CAN-based multicast generates larger overhead traffic than SDEM. While the achieved performance results are partially because of the different group maintenance paths in these four protocols, a general trend is told that tree-based multicast requires larger overhead traffic to maintain dynamic member alterations than mesh-based multicast.

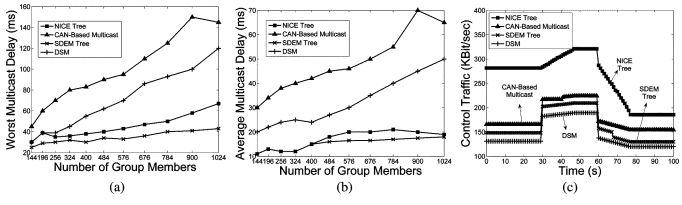

Fig. 10. Performance of packet multicast delays and dynamic network overhead. (a) Worst multicast delays. (b) Average multicast delays. (c) Overhead traffic when dynamic network alterations happen.

In general, the simulation results agree with our theoretical analysis, although the discrepancies (not greater than 30%) have been reported. The major reasons for the discrepancies is that the simulations are case studies but the theorems analyze general cases. At the same time, we believe that the traffic disturbance and the throughput variation in the simulations are likely to be the causes for the observed discrepancies. Further, when dynamic network alternations happen, multimedia traffic may transfer between different links which causes the performance changing in individual links. But the worst and average performance of the whole system can still be predicted by our theorems. For example, our theorems tell the system performance when 240 members leave a 1024-member multicast group if 784 is used by the theorems as the new group size.

VIII. r-ary Tree versus m-D Mesh

Both the theoretical analyses and the simulation observation show that communications on top of r-ary tree and m-D mesh topologies have different advantages and disadvantages for multicasting multimedia streams. Fundamentally, such difference is caused by constructing different overlay transmission paths in different routing schemes. In this section, based on the analysis and observation in the above sections, we evaluate r-ary tree-based multicast and m-D mesh-based multicast when they are used for transmitting multimedia streams.

According to Remark 1, under the same group size n, the increment of the longest overlay path length in mesh-based multicast is much greater than that in tree-based multicast. Since our analysis assumes that tree-based multicast and mesh-based multicast are on top of the same underlying network architecture, the theoretical results reveal that mesh-based routing generates much longer multicast delays than tree-based routing. That is, r-ary tree-based multicast suits to the applications that have requirements for real-time communications. Fundamentally, the longer multicast delays caused by m-D mesh multicast include packet transmission delays in longer overlay paths and packet processing delays at the extra numbers of end hosts who are involved into longer overlay paths. When comparing the results in Corollaries 1 and 2, we have $\Delta OH_{tree}/\Delta OH_{mesh} \ll 1$. As we have analyzed in Remark 1, it indicates that multimedia packets experience less delay differences to arrive at different group members in tree-based multicast than in mesh-based multicast. With these observations, we can conclude that r-ary tree overlay multicasting not only suits for real-time multimedia applications but also suits for interactive multimedia applications.

Video and audio fidelity is another important performance characteristic for multimedia applications. High fidelity is usually taken to mean continuous high resolution video and unin-

terrupted clearly articulated audio. Hence, multimedia multicast schemes require high rate transmission, low delay jitter and low loss rate. Our theoretical analysis on link stress proved that r-ary tree routing is easier to suffer from bottleneck than m-D mesh-based routing which indicates that m-D mesh-based routing is good for transmitting high-rate multimedia streams. For more detailed observation, based on $\ddot{LS}_{tree} = \sqrt{1+nr-n}/r$, the bottleneck possibility in tree-based multicast increases in terms of n's square root with the increment of group size. In order to release bottleneck, the multimedia transmission rates in tree-based multicast have to be decreased in terms of n's square root when the group size increases. Otherwise, bottleneck causes packet loss and longer queueing delays that may generate large delay jitter or even temporary frozen up at the downstream receivers. When it comes to multi-source groups, the higher representation quality advantage of mesh-based multicast over tree-based multicast is more obvious. It is because mesh-based multicast guarantees that each group member has fixed direct child members irrespective of the sources' joining or leaving, while tree-based multicast makes the newly connected direct child members aggregate the load burden of their forwarder. Hence, the trade-off between high transmission rates and real-time continuous transmission becomes more serious. Our analysis on the average link stress and the link stress difference shows that bottleneck in tree-based multicast can be released through balancing the end hosts' roles in different source trees. In an r-ary overlay tree, hot spots are usually in certain areas (e.g., middle-layer areas). Hence, an effective way to avoid heavy burden links/end hosts is to fully utilize the end hosts farther away from the hot spots in one source group as the forwarders closer to the hot spots in another source group.

The study on the number of shortest paths shows that mesh-based flooding is more robust in terms of guaranteeing short delay multimedia communications than tree-based multicast when dynamic network alterations take place. Tree topology only guarantees d^l connections between two group members, while mesh topology generates $d^l[m(m-1)+\sum_{i=1}^{m} H_i \sum_{j=1,j\neq i}^{m}(H_j-1)]$ shortest paths between two group members locating in zones $(d_1^Z, d_2^Z, \ldots, d_m^Z)$ and $(d_1^Z + H_1, d_2^Z + H_2, \ldots, d_m^Z + H_m)$. Moreover, the robustness of mesh-based multicast increases when the end host distance or the group size increases. However, the robustness is in the cost of occupying large amount of network resources and complex control overhead to avoid routing loop generated by the plenty of connectivity in m-D mesh topology. Hence, we believe that mesh-based multicast suits to the multimedia applications that users are distributed densely. As compared to tree-based multicast, in dense area, mesh-based multicast costs less extra resources but provides load-balanced transmission. For large-scale multi-traffic mesh-based multicast, we have introduced to reduce network resource consumption by dividing the shortest paths into different sets with each set assigned to one stream in Remark 3. Network resources are therefore reasonably utilized by mesh-based multicast with satisfied robustness. Tree-based multicast is on the opposite. It is fragile but simple. Only one shortest overlay connection between two group members which makes the group easier to be interrupted but doesn't cause routing looping and network resource waste.

Altogether, the advantages and disadvantages of r-ary tree and m-D mesh are basically complementary. It is difficult to say which topology is better than the other one. Different multimedia applications have different requirements for the acceptable communication qualities. The selection of overlay networks is application dependent. We think that tree-based multicast suits to single-source streaming media multicast applications which have requirements for real time and interactivity, for example network conference, online games, network broadcasting, and network stock and bank, while mesh-based multicast suits to non-real-time and multi-source multimedia multicast applications which more focus on reliable transmission and have dense network distribution, for example text multicast, and video-on-demand within home networking. For multi-source real time multicast applications, tree-based multicast holds advantages as compared to mesh-based multicast but still has the problem of bottleneck which may cause interrupted multimedia transmissions. Fortunately, this problem can be solved by the fact that overlay multicast trees are usually source trees due to overlay multicast properties (e.g., end hosts' low capacities). A promising way to achieve short delay and continuous multimedia transmission in a multi-source environment is to design a tree-based multicast overlay which is rooted at different sources and allows group members to serve in multiple roles. For example, members close to hot spots in one source tree could act as members away from hop spots in another source tree. Additionally, for higher performance overlay multicast that provides rich-content, short delay, low delay jitter, and low packet loss multimedia transmission to receivers, a line of research [1], [11] proposes to combine the advantages of the two topologies, such as designing tree routing on top of m-D control mesh. We are interested in carrying out a performance analysis of this hybrid approach in our next step.

IX. Conclusion

In this paper, we presented our studies on the different multimedia multicast performances caused by different overlay topologies: r-ary tree and m-D mesh. We analyzed *the worst performance*, *the average performance*, and *the performance difference* for the metrics of *the link stress*, *the number of overlay hops*, and *the number of shortest paths*. We presented a set of theorems and used computer simulation to evaluate tree-based and mesh-based multimedia multicast along the analyzed metrics. Simulation results proved our theoretical prediction. In general, tree topologies are efficient in distributing packets, while mesh topologies may incur very long delays; tree topologies are sensitive to the node/link failure, while mesh topologies are robust in facing of dynamic group alteration; tree topologies may scale to large size groups in single source applications, while mesh topologies may scale to large size groups in multi-source applications; tree topologies hold the advantage of implementing interactive communication, while mesh topologies generate inconsistent delay performance among different group members; tree topologies are available for the applications with sparse user distribution, while mesh topologies suits to the applications with dense user distribution. Therefore, we think that the topology selection should be application dependent. We believe that our study is useful for

protocol design of target applications, investigating multimedia multicast functions, and other multimedia network performance analysis.

ACKNOWLEDGMENT

The authors would like to thank the editor and reviewers for their invaluable and helpful comments and Mr. M. O'Brien for discussion of mathematics and help with English.

REFERENCES

[1] Y. H. Chu, S. Rao, S. Seshan, and H. Zhang, "A case for end system multicast. Distance vector multicast routing protocol—DVMRP. RFC 1812, June 1995," in *Proc. ACM SIGMETRICS 2000*, Santa Clara, CA, Jun. 17–21, 2000, pp. 1–12.

[2] P. Francis, Yoid: Extending The Internet Multicast Architecture, 2000. [Online]. Available: http://www.aciri.org/yoid/docs/index.html.

[3] Y. H. Chu, S. G. Rao, S. Seshan, and H. Zhang, "Enabling conferencing applications on the Internet using an overlay multicast architecture," in *Proc. ACM SIGCOMM 2001*, San Diego, CA, Aug. 27–31, 2001, pp. 55–67.

[4] D. Pendarakis, S. Shi, D. Verma, and M. Waldvogel, "ALMI: An application level multicast infrastructure," in *Proc. 3rd Usenix Symp. Internet Technologies and Systems (USITS01)*, San Francisco, CA, Mar. 2001, pp. 49–60.

[5] Y. Chawathe, S. McCanne, and E. A. Brewer, "Rmx: Reliable multicast for heterogeneous networks," in *Proc. IEEE INFOCOM 2000*, Tel Aviv, Israel, Mar. 26–30, 2000, pp. 795–804.

[6] S. Shi and J. Turner, "Routing in overlay multicast networks," in *Proc. IEEE INFOCOM 2002*, New York, Jun. 23–27, 2002, pp. 1200–1208.

[7] S. Ratnasamy, M. Handley, R. Karp, and S. Shenker, "Application-level multicast using content-addressable networks," in *Proc. 3rd Int. Workshop Network Group Communication*, London, U.K., Nov. 7–9, 2001, pp. 14–29.

[8] S. Ratnasamy, P. Francis, M. Handley, and R. Karp, "A scalable content- addressable network," in *Proc. ACM SIGCOMM 2001*, San Diego, CA, Aug. 27–31, 2001, pp. 161–172.

[9] S. Banerjee, B. Bhattacharjee, and C. Kommareddy, "Scalable application layer multicast," in *Proc. ACM SIGCOMM*, Pittsburgh, PA, Aug. 19–23, 2002, pp. 205–217.

[10] B. Zhang, S. Jamin, and L. Zhang, "Host multicast: A framework for delivering multicast to end users," in *Proc. IEEE INFOCOM 2002*, New York, Jun. 23–27, 2002, pp. 1366–1375.

[11] M. Castro, P. Druschel, A. M. Kermarrec, and A. Rowstron, "SCRIBE: A large- scale and decentralized application-level multicast infrastructure," *IEEE J. Select. Areas Commun.*, vol. 20, no. 8, pp. 1489–1499, Oct. 2002.

[12] A. Rowstron and P. Druschel, Pastry: Scalable, Distributed Object Location and Routing for Large-Scale Peer-to-Peer Systems, 2001. [Online]. Available: http://research.mecrosoft.com/antr/PAST/.

[13] J. Jannotti, D. K. Gifford, K. L. Johnson, M. F. Kaashoek, and J. W. O'Toole, Jr., "Overcast: Reliable multicasting with an overlay network," in *Proc. 4th Usenix Symp. Operating Systems Design and Implementation*, San Diego, CA, Oct. 22–25, 2000, pp. 194–210.

[14] S. Banerjee, C. Kommareddy, B. B. K. Kar, and S. K. huller, "Construction of an efficient overlay multicast infrastructure for real-time applications," in *Proc. IEEE INFOCOM 2003*, San Francisco, CA, Mar. 3, 2003, pp. 1521–1531.

[15] A. Riabov and Z. Liu, "Overlay multicast trees of minimal delay," in *Proc. 24th Int. Conf. Distributed Computing Systems (ICDCS 2004)*, Tokyo, Japan, Mar. 23–26, 2004, pp. 654–661.

[16] W. Tu and W. Jia, "A scalable and efficient end host multicast for peer-to-Peer systems—DSCT," in *Proc. IEEE Global Telecommunication Conf. 2004 (Globecom 2004)*, Dallas, TX, Nov.–Dec. 29-3, .

[17] W. Tu and W. Jia, "An end host multicast protocol for peer-to-Peer systems," in *Proc. 30th Annu. IEEE Conf. Local Computer Networks (LCN 2005)*, Nov. 2005.

[18] J. Chuang and M. Sirbu, "Pricing multicast communications: A cost-based approach," in *Proc. Internet Society INET*, Jul. 1998.

[19] G. Phillips, S. Shenker, and H. Tangmunarunkit, "Scaling of multicast trees: Comments on the Chuang-Sirbu scaling law," in *Proc. ACM SIGCOMM*, 1999, pp. 41–51.

[20] C. Adjih, L. Georgiadis, P. Jacquet, and W. Szpankowski, "Multicast tree structure and the power law," in *Proc. SODA*, 2002.

[21] S. Fahmy and M. Kwon, "Characterizing overlay multicast networks," in *Proc. ICNP*, 2003.

[22] W. Tu, "Performance analysis for overlay multicast on tree and m-D mesh topologies," in *Proc. ICC*, 2007.

[23] W. Tu, "Short delay end host multicast," in *End Host Multicast Algorithms for Overlay Networks*. Kowloon, Hong Kong: Ph.D. dissertation, 2005, ch. 3.

Wanqing Tu (M'04) received the Ph.D. degree in computer science from the City University of Hong Kong, Kowloon, in 2006.

She is a Lecturer in the School of Computing and Communications Technology at Glyndwr University, Wrexham, U.K. Her research interests include QoS, multimedia multicast, overlay networks, wireless mesh networks, and distributed computing. She was awarded an Embark Initiative Postdoctoral Research Fellowship from the Irish Research Council for Science, Engineering, and Technology in 2006. She was invited to the University of New South Wales as a Visiting Research Fellow in 2007.

Dr. Tu is a member of the IEEE Computer Society and the IEEE Communications Society. She was awarded one Best Paper Award in ICCNMC 2005.

Xing Jin received the Ph.D. and B.Eng. degrees, both in computer science, from The Hong Kong University of Science and Technology, Clear Water Bay, and Tsinghua University, Beijing, China, in 2007 and 2002, respectively.

He is a member of the technical staff with the Systems Technology Group, Oracle USA, Inc., Redwood Shores, CA. His research interests include distributed information storage and retrieval, peer-to-peer technologies, multimedia networking, and Internet topology inference.

Dr. Jin has been on the editorial board of the *Journal of Multimedia* since 2006 and the *Canadian Journal of Pure and Applied Sciences* since 2007. He was awarded the Microsoft Research Fellowship in 2005. He is a member of the Sigma Xi and the IEEE COMSOC Multimedia Communications Technical Committee.

Peter S. Excell (SM'84) received the B.Sc. degree in engineering science from the University of Reading, Reading, U.K., in 1970 and the Ph.D. degree from the University of Bradford, Bradford, U.K., in 1980.

He is a Professor of communications and Head of the School of Computing and Communications Technology at Glyndwr University, Wrexham, U.K. Until 2007, he was Associate Dean for Research in the School of Informatics at the University of Bradford. His academic interests cover wireless technologies, electromagnetics, and antennas, plus broader interests in mobile communications and their future content and applications. He has published over 300 papers and holds three patents.

Dr. Excell is a Chartered Engineer and Chartered IT Professional; a Fellow of the Institution of Engineering and Technology, the British Computer Society, and the Higher Education Academy; and a member of the Association for Computing Machinery, the Bioelectromagnetics Society, and the Applied Computational Electromagnetics Society.