PAPER

Effectiveness of overlay routing based on delay and bandwidth information

Go HASEGAWA†, *Member*, **Yuichiro HIRAOKA**††, *Nonmember*, *and* **Masayuki MURATA**†, *Member*

SUMMARY

Recent research on overlay networks has revealed that user-perceived network performance, such as end-to-end delay performance, could be improved by an overlay routing mechanism. However, these studies consider only end-to-end delay, and few studies have focused on bandwidth-related information, such as available bandwidth and TCP throughput, which are important performance metrics especially for long-lived data transmission. In the present paper, we investigate the effect of overlay routing both delay and bandwidth-related information, based on the measurement results of network paths between PlanetLab nodes. We consider three metrics for selecting the overlay route: end-to-end delay, available bandwidth, and TCP throughput. We then show that the available bandwidth-based overlay routing provides significant gain, as compared with delay-based routing. We further reveal the correlation between the latency and available bandwidth of the overlay paths and propose several guidelines for selecting an overlay path.

key words: *Overlay networks, Overlay routing, Available bandwidth, end-to-end delay, TCP throughput*

1. Introduction

As the Internet increasingly diversifies and the user population grows rapidly, new and varied types of service-oriented networks are emerging. Service overlay networks [1] are defined as upper-layer networks that provide special-purpose services that are built on the lower-layer IP network, and include P2P networks, anonymous file-sharing services, audio and video conferencing services, and Content Delivery/Distribution Networks (CDNs). Therefore, the performance of service overlay networks depends primarily on how well the networks take advantage of the characteristics and resources of the underlying IP network.

In overlay networks, the endhosts and servers that run the applications become overlay nodes that form the upper-layer logical network with logical links between the nodes, as depicted in Figure 1. Some of the overlay networks select a route for data transmission according to network conditions such as link speed, delay, packet loss ratio, hop count, and TCP throughput between overlay nodes. In WinMX, an endhost can report the type of network link used to connect to the Internet when joining the network. CDNs such as NetLightning [2] and Akamai [3] distribute overlay nodes

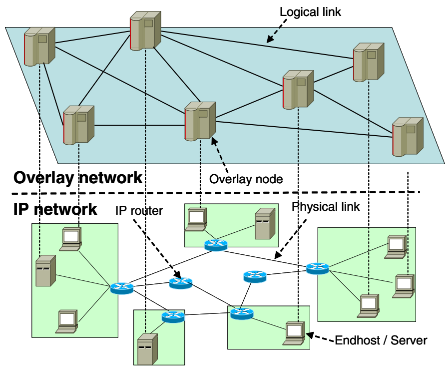


Fig. 1 Overlay network

(content servers) over the entire Internet and select appropriate source and destination hosts according to the network condition when the contents are moved, duplicated, or cached.

Some overlay networks do not assume specific upper-layer applications and concentrate only on the routing of overlay network traffic. We call such application-level traffic routing *overlay routing*, and overlay networks for traffic routing are referred to as *routing overlay networks*, as depicted in Figure 2. In Resilient Overlay Networks (RON) [4], for example, each overlay node measures the end-to-end latency and packet loss ratio of the network path to other nodes, and determines the route for the overlay network traffic originating from the node, which can be a direct route from the node to the destination node or a relay route that traverses other node(s) before reaching the destination node. In [4], the authors reported that RON can provide an effective traffic transmission path compared with lower-layer IP routing. Furthermore, RON can detect network failures (link and node failures, and mis-configured routing settings) and can provide an alternate route faster than IP routing convergence.

Several studies have examined the effectiveness of overlay routing with respect to IP routing [5]–[11]. For example, in [8], the authors used actual measurement data of the transmission latency among several geographically-distributed hosts in two Internet Service Providers (ISPs) in Japan, and showed that the transmission latency of approximately 28% of end-to-end paths can be reduced by relaying another host, as com-

†The authors are with Graduate School of Infomation Science and Technology, Osaka University

††The author is with NEC System Technologies, Ltd.

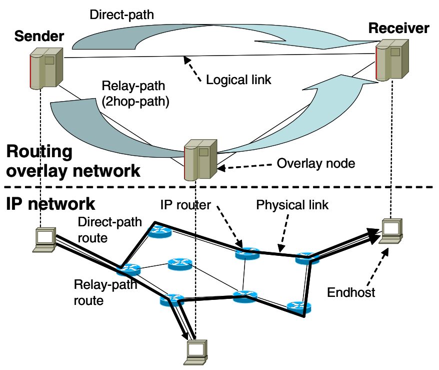


Fig. 2 Overlay routing and routing overlay network

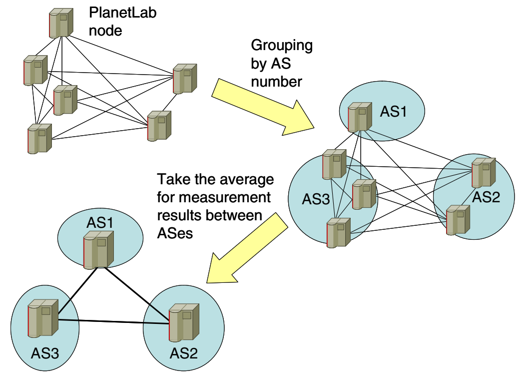


Fig. 3 Grouping PlanetLab nodes

pared to using the direct path. In [9], the authors investigated the effectiveness of the reactive overlay routing by using the measurement data on four ISPs in United States, and confirm its effectiveness compared to the IP routing and proactive overlay routing. However, most of these studies focus on end-to-end delay performance, and few studies have focused on bandwidth-related information, such as available bandwidth and TCP throughput, which are an important performance metric, especially for long-lived data transmission.

In the present paper, we investigate the effectiveness of overlay routing, based on both delay and bandwidth information. We assume that PlanetLab [12] nodes construct a routing overlay network and use the measurement results obtained from the Scalable Sensing Service (S^3) [13], which measures various properties of network paths between PlanetLab nodes. We use the following three metrics in selecting an overlay route: end-to-end delay, available bandwidth, and TCP throughput. In the present study, we investigate the effectiveness of the three-hop relay overlay path, whereas most of the previous studies on overlay routing focused on the two-hop relay overlay path. Another interesting result in this paper is the correlation between transmission latency and available bandwidth of the end-to-end path. In addition, we investigated whether a network path with a larger available bandwidth has a smaller transmission latency, and vice versa.

The remainder of this paper is organized as follows. In Section 2, we explain the methodology and performance metrics. We then present the investigation results for evaluating the effectiveness of overlay routing in Section 3. Section 4 summarizes the conclusions of the present study and discusses areas for future consideration.

2. Methodologies

2.1 Dataset used for evaluation

We investigate the effectiveness of overlay routing based on delay and bandwidth information under the assumption that PlanetLab nodes construct a routing overlay network. For this purpose, we utilize the measurement results obtained from S^3. S^3 measures various properties of end-to-end paths between PlanetLab nodes, including physical capacity, available bandwidth, end-to-end delay, and packet loss ratio. The measurement results are provided every four hours via a Web site. In this paper, we use one snap-shot data obtained on Oct. 25th, 2006. Note that we have investigated with the datasets on other dates and obtained the similar results to those in this paper.

There exist 588 PlanetLab nodes in the measurement data utilized herein. However, a number of nodes are located in the same subnetwork, as estimated from the IP address and the host name of the nodes. In evaluating the effectiveness of overlay routing, we should avoid using the nodes in the same subnetwork as relay nodes for the following four reasons: (1) Most planetlab nodes in an AS seem to locate at the same subnetwork, that is estimated from their hostnames. (2) The measurement results of end-to-end delay and available bandwidth between nodes in the same subnetwork may be quite small for delay and quite large for available bandwidth, which may overestimate the effectiveness of overlay routing. (3) The measurement results between nodes in the same subnetwork may include large errors especially for available bandwidth. ()4) There is almost no meaning in using a relay node in the same subnetwork as the source and destination nodes.

Therefore, we divide the PlanetLab nodes into groups according to their AS number and assume that there is only one overlay node in each AS. We obtain the AS number of PlanetLab nodes by tracerouting from a route server in traceroute.org [14] to the PlanetLab nodes. As a result, the number of overlay nodes decreases to 179, which is equal to the number of ASes of PlanetLab nodes. In grouping, we take the average for measurement results when we have more than one measurement result between the overlay nodes (ASes). Figure 3 depicts this process for node grouping.

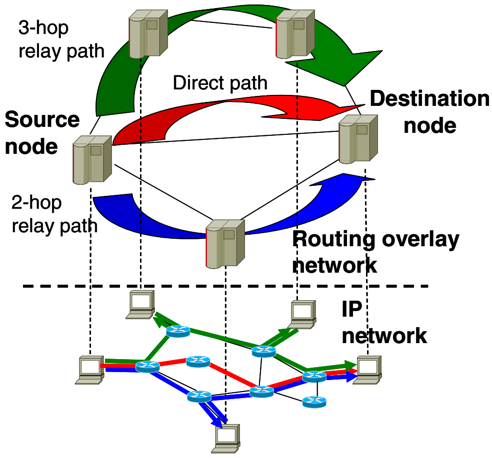


Fig. 4 Definition of overlay path

2.2 Overlay path candidates

When one node (source node) selects the transmission path to another node (destination node), we compare the end-to-end latency and available bandwidth of the following three candidates (Figure 4):

- Direct path: the source node to the destination node
- Two-hop relay path: the source node to the destination node via a relay node
- Three-hop relay path: the source node to the destination node via two relay nodes

2.3 Metrics

In this subsection, we explain the metrics utilized for selecting overlay paths.

2.3.1 End-to-end latency

Overlay routing based on end-to-end latency would be adapted for applications, including voice chat applications such as Skype [15] that need quick response, rather than bandwidth-related resources. We utilize the measurement results from S^3 for the end-to-end latency of the direct path between nodes. We define the end-to-end latency of a relay path as the sum of the latencies of direct paths constructing the relay path. We assume that the number of overlay nodes is M and that the measured results of the end-to-end delay of the network path between nodes N_i and N_j is τ_{ij} $(1 \leq i, j \leq M)$. Then, we can describe the latencies of the direct path, the two-hop relay path, and the three-hop relay path, as follows:

$$D^1_{ij} = \tau_{ij} \quad (1)$$

$$D^2_{ikj} = \tau_{ik} + \tau_{kj} \quad (2)$$

$$D^3_{iklj} = \tau_{ik} + \tau_{kl} + \tau_{lj} \quad (3)$$

We denote the relay node for the two-hop relay path as N_k and the relay nodes for the three-hop relay path as N_k and N_l $(1 \leq k, l \leq M, \quad k \neq l, \quad k, l \neq i, j)$. Furthermore, we define the latency-optimized path as the relay path that has the smallest end-to-end latency. We can then obtain the respective latencies of the two-hop and three-hop latency-optimized paths as follows:

$$\hat{D}^2_{ij} = \min_{k \neq i,j} (D^2_{ikj}) \quad (4)$$

$$\hat{D}^3_{ij} = \min_{k \neq l,\ k,l \neq i,j} (D^3_{iklj}) \quad (5)$$

In this paper, we compare the performance of the direct path and the relay path for each node pair. We therefore define the *improvement ratio* of the relay path with respect to the direct path as follows:

$$I(D^2_{ikj}) = \frac{D^1_{ij}}{D^2_{ikj}}$$

$$I(D^3_{iklj}) = \frac{D^1_{ij}}{D^3_{iklj}}$$

When the above ratio is larger than 1, we can say that the relay path is effective compared with the direct path.

2.3.2 Available bandwidth

Available bandwidth is an important performance metric for audio video streaming services such as YouTube [16] and GyaO [17]. We simply use the measurement results of available bandwidth in S^3 for the available bandwidth of direct paths. We define the available bandwidth of a relay path as the minimum available bandwidth of direct paths constructing the relay path. Denoting the available bandwidth of the network path between node N_i and N_j as ρ_{ij}, we can describe the available bandwidths of the direct path, the two-hop relay path, and the three-hop relay path, as follows:

$$B^1_{ij} = \rho_{ij} \quad (6)$$

$$B^2_{ikj} = \min(\rho_{ik}, \rho_{kj}) \quad (7)$$

$$B^3_{iklj} = \min(\rho_{ik}, \rho_{kl}, \rho_{lj}) \quad (8)$$

We also define the bandwidth-optimized path as the relay path that has the largest available bandwidth among all possible relay paths. We can then obtain the respective available bandwidths of the two-hop and three-hop bandwidth-optimized paths as follows:

$$\hat{B}^2_{ij} = \max_{k \neq i,j} (B^2_{ikj}) \quad (9)$$

$$\hat{B}^3_{ij} = \max_{k \neq l,\ k,l \neq i,j} (B^3_{iklj}) \quad (10)$$

Furthermore, we define the improvement ratio of the relay path with respect to the direct path as follows:

$$I(B^2_{ikj}) = \frac{B^2_{ikj}}{B^1_{ij}}$$
$$I(B^3_{iklj}) = \frac{B^3_{iklj}}{B^1_{ij}}$$

2.3.3 TCP throughput

Overlay routing based on TCP throughput would be adapted for file sharing applications like Bittorrent [18] and WinMX. In [4], RON utilizes TCP throughput as a performance metric for overlay routing, which is calculated from the end-to-end delay and the packet loss ratio as follows:

$$T = \frac{\sqrt{1.5}}{RTT\sqrt{Loss}} \text{ (packet/sec)} \quad (11)$$

Equation (11) is based on the formula for the average throughput of the long-lived TCP connection in [19]. This metric increases as the packet loss ratio and RTT decrease, but never exceeds the available bandwidth of the path in the actual situation. In the present study, we use Eq. (11) as a performance metric of TCP throughput. This equation includes the packet loss ratio of the path, and S^3 has the measurement results of packet loss ratios of the network path between the PlanetLab nodes. However, we do not use them, since they are obtained by sending only 100 probe packets. Note that from the measurement results of the packet loss ratio obtained from S^3, we found that the packet loss ratio is not related to the AS-lelvel hop count of the path. Generally, the correlation between the hop count and packet loss ratio can not be determined easily. Therefore, we utilize the following two extreme models for packet loss ratio of the direct path:

(A) AS-hop-count-base loss ratio (AS): the packet loss ratio of the path is determined in proportion to the AS-level hop count of the path.
(B) Overlay-hop-count-base loss ratio (OL): the packet loss ratio of the direct path is constant value regardless of the other characteristics of the path. The relay path has a packet loss ratio proportionally that is proportional to its overlay-level hop count.

In the actual network environment. we expect that we would have moderate results between AS and OL cases.

We define P^1_{ij} as the TCP throughput of the direct path between node N_i and node N_j, and we can describe P^1_{ij} as follows:

$$P^1_{ij} = \min\left(\frac{(8\cdot MSS)\sqrt{1.5}}{D^1_{ij}\sqrt{L_1}}\ ,\ B^1_{ij}\right) \text{ (bps)} \quad (12)$$

$$L_1 = \begin{cases} n_{ij}\cdot L_\mathrm{A} & \text{(case(A))} \\ L_\mathrm{B} & \text{(case(B))} \end{cases}$$

L_A and L_B are parameters that determine the packet loss ratio per AS-level hop and the packet loss ratio per overlay-level hop, respectively. In addition, we denote that the maximum segment size as MSS and the AS-level hop count between node N_i and node N_j as n_{ij}.

The definition of the TCP throughput of the relay path is different depending on whether the TCP connection is terminated at each relay node, which means that we utilize the TCP proxy mechanism [20] at the relay node. When we do not use the TCP proxy mechanism, meaning that we utilize an end-to-end TCP connection, we calculate the TCP throughput of the relay path from the end-to-end latency and the available bandwidth of the relay path as follows:

$$P^2_{ikj}(e2e) = \min\left(\frac{(8\cdot MSS)\sqrt{1.5}}{D^2_{ikj}\sqrt{L_2}}\ ,\ B^2_{ikj}\right) \quad (13)$$
$$P^3_{iklj}(e2e) = \min\left(\frac{(8\cdot MSS)\sqrt{1.5}}{D^3_{iklj}\sqrt{L_3}}\ ,\ B^3_{iklj}\right) \quad (14)$$
$$L_2 = \begin{cases} (n_{ik}+n_{kj})\cdot L_\mathrm{A} & \text{(case(A))} \\ 2L_\mathrm{B} & \text{(case(B))} \end{cases}$$
$$L_3 = \begin{cases} (n_{ik}+n_{kl}+n_{lj})\cdot L_\mathrm{A} & \text{(case(A))} \\ 3L_\mathrm{B} & \text{(case(B))} \end{cases}$$

On the other hand, when the TCP proxy mechanism is deployed, we determine the TCP throughput of the relay path as the minimum TCP throughput of direct paths constructing the relay path:

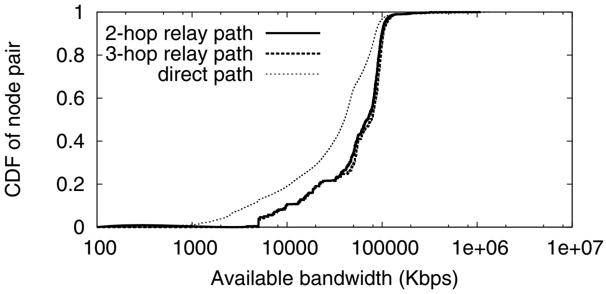
$$P^2_{ikj}(pxy) = \min\left(P^1_{ik},\ P^1_{kj}\right) \quad (15)$$
$$P^3_{iklj}(pxy) = \min\left(P^1_{ik},\ P^1_{kl},\ P^1_{lj}\right) \quad (16)$$
$$(17)$$

As in the case of the available bandwidth, we can define the throughput-optimized relay paths as follows:

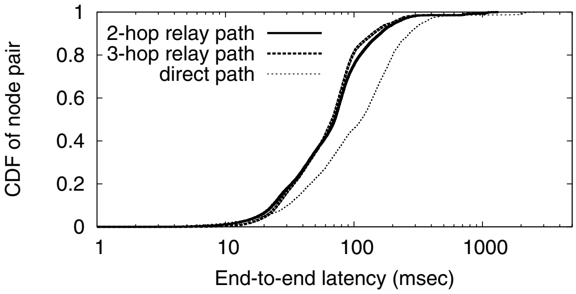
$$\hat{P}^2_{ij}(e2e) = \max_{k\neq i,j}\left(P^2_{ikj}(e2e)\right) \quad (18)$$
$$\hat{P}^3_{ij}(e2e) = \max_{k\neq l,\quad k,l\neq i,j}\left(P^3_{iklj}(e2e)\right) \quad (19)$$
$$\hat{P}^2_{ij}(pxy) = \max_{k\neq i,j}\left(P^2_{ikj}(pxy)\right) \quad (20)$$
$$\hat{P}^3_{ij}(pxy) = \max_{k\neq l,\quad k,l\neq i,j}\left(P^3_{iklj}(pxy)\right) \quad (21)$$

Furthermore, the improvement ratio of the relay path with respect to the direct path can be described as follows:

$$I(P^2_{ikj}(e2e)) = \frac{P^2_{ikj}(e2e)}{P^1_{ij}}$$
$$I(P^3_{iklj}(e2e)) = \frac{P^3_{iklj}(e2e)}{P^1_{ij}}$$
$$I(P^2_{ikj}(pxy)) = \frac{P^2_{ikj}(pxy)}{P^1_{ij}}$$
$$I(P^3_{iklj}(pxy)) = \frac{P^3_{iklj}(pxy)}{P^1_{ij}}$$



(a) Available bandwidth



(b) End-to-end latency

Fig. 5 Distribution of end-to-end latency and available bandwidth

3. Evaluation results and discussions

3.1 Performance distribution of overlay path

In Figure 5, we show the distributions of end-to-end latency and available bandwidth of direct paths and relay paths for all node pairs. We can observe from Figure 5(a) that 80% of the direct paths have an available bandwidth of from 10 Mbps to 100 Mbps. However, using the relay path, the ratio increases to 90%. For end-to-end latency (Figure 5(b)), roughly half of the direct paths the end-to-end latency from 10 ms to 100 ms, and it increases to 80% by using relay paths. Furthermore, the degree of improvement is quite large, especially when the performance of the direct path is poor: less than 10 Mbps for available bandwidth and greater than 20 msec for end-to-end latency. Based on these results, we can expect to find a relay path with a better performance than that of the direct path both in terms of end-to-end latency and available bandwidth, especially when the performance of the direct path is poor.

In Figure 6, we show the distributions of TCP throughput of direct paths for four combinations of packet loss ratio and TCP connection setting in calculating TCP throughput. The four cases are as follows:

AS/e2e The packet loss ratio is proportional to the AS-level hop count of the path (AS), and the end-to-end TCP connection is utilized (e2e).

OL/e2e The packet loss ratio is proportional to the overlay-level hop count of the path (OL), and the end-to-end TCP connection is utilized (e2e).

AS/pxy The packet loss ratio is proportional to the AS-level hop count of the path (AS), and the TCP proxy mechanism is deployed (pxy).

OL/pxy The packet loss ratio is proportional to the overlay-level hop count of the path (OL), and the TCP proxy mechanism is deployed (pxy).

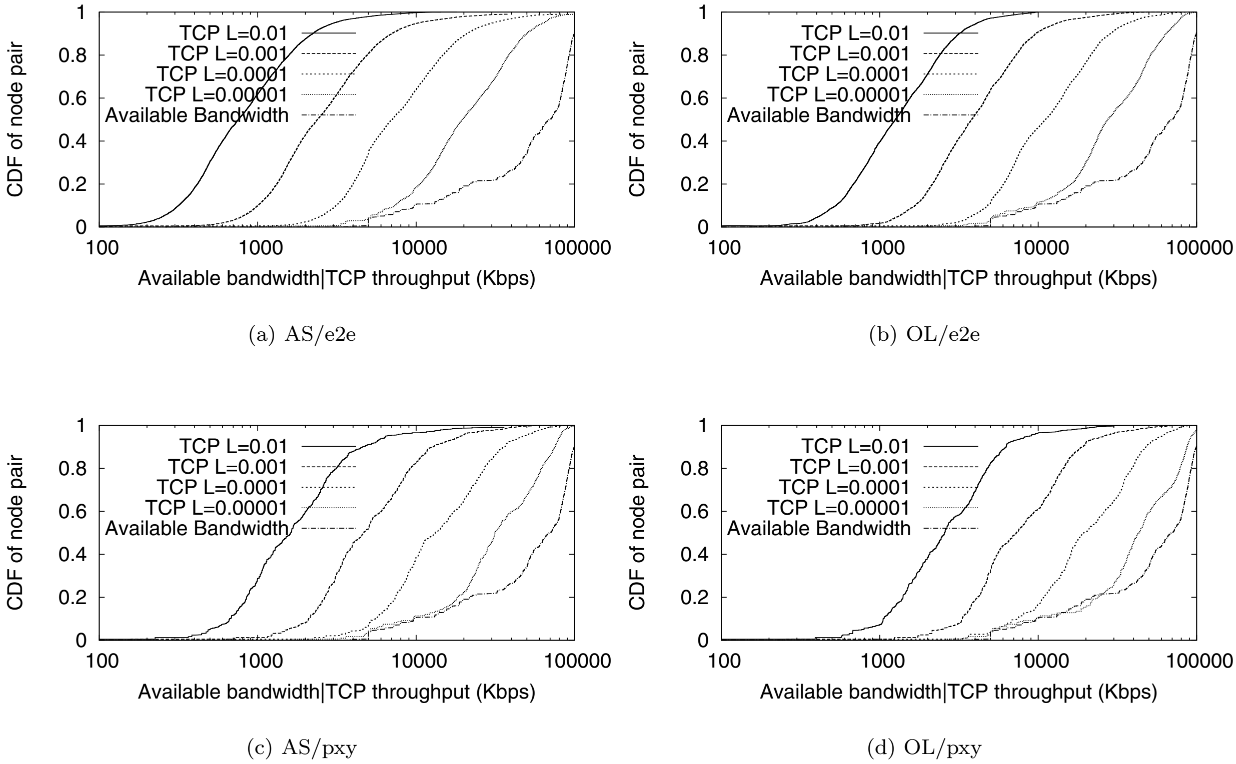
Note that the variable L in the following figures and explanations means L_A and L_B explained in Subsection 2.3.3.

Figure 6 shows that the TCP throughput in OL cases (Figure 6 (b) and (d)) is better than that in AS cases (Figure 6 (a) and (c)) and that the TCP throughput in pxy cases (Figure 6 (c) and (d)) is better than that in e2e cases (Figure 6 (a) and (b)). The reason for this is that the packet loss ratio of the OL case does not depend on the AS-hop-count. Therefore, the packet loss ratio of the OL case generally becomes smaller than that of the AS case, and the TCP proxy mechanism can isolate the effect of packet loss, and, as a result, the TCP throughput remains unaffected by the packet losses at other parts of the overlay path. We also observe that the TCP throughput does not reach the available bandwidth even when $L = 0.00001$ for all cases. This result may indicate that the bandwidth resource of PlanetLab nodes is sufficiently large, and the available bandwidth is far larger than the TCP throughput to be achieved.

3.2 Characteristics of relay path

In Figure 7, we present the distribution of the relationship between the available bandwidth of the direct path and that of the bandwidth-optimized relay path for each node pair, for two-hop relay paths (Figure 7(a)) and three-hop relay paths (Figure 7(b)), respectively. Figure 8 shows similar plots for end-to-end latency. For 96.6% of all node pairs, we can find a two-hop relay path that has a larger available bandwidth than the direct path. When we compare the direct path and the 3-hop relay path, for 97.7% of all node pairs, we can find a three-hop relay path that has a larger available bandwidth. For end-to-end latency, these percentages decrease to 87.5% and 85.4%, respectively.

Furthermore, with respect to available bandwidth, 46.9% of node pairs for which a better two-hop relay path cannot be found, a three-hop relay path having a larger available bandwidth than the direct path can be found. In addition, for 51.6% of the node pairs that has a larger available bandwidth than the direct path, we can find a better three-hop relay path than



(a) AS/e2e

(b) OL/e2e

(c) AS/pxy

(d) OL/pxy

Fig. 6 Distribution of TCP throughput

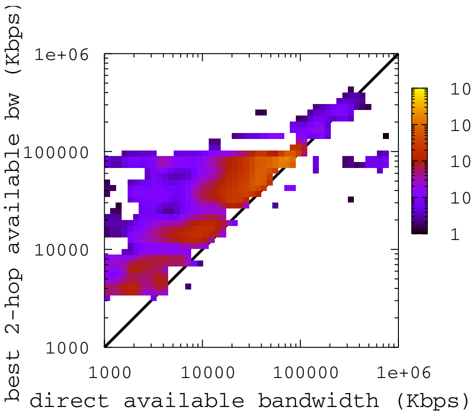
the bandwidth-optimized two-hop relay path. With respect to end-to-end latency, these percentages decrease to 17.8% and 47.3%, respectively.

The above results indicate that the effectiveness of the latency-based relay path is smaller than that of the available bandwidth-based relay path. A reasonable explanation for this is that the underlying IP routing is configured based on router-level and AS-level hop count, which have some degree of correlation with the end-to-end delay.

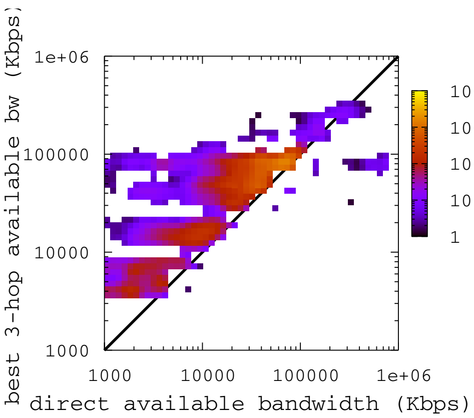
In Figure 9, we present the distribution of the relationship between the TCP throughput of the direct path and that of the throughput-optimized two-hop relay path for each node pair. We plot four combinations of packet loss ratio and TCP connection setting in calculating TCP throughput. We set $L = 0.00001$ in this figure. From Figure 9(a), for 45.9% of all node pairs, we can find a two-hop relay path that has a larger TCP throughput than the direct path, and 47.8% for Figure 9(b). For Figures 9(c) and (d), these percentages increase 95.2% and 95.8%, respectively. The reasons for this can be explained as follows.

When we use TCP proxy, the TCP throughput does not degrade significantly due to the effect of the TCP proxy mechanism described in the previous subsection. Therefore, the effectiveness of the relay path becomes similar to that of available bandwidth shown in Figure 7. On the other hand, by comparing Figures 9 (a) and (c) and Figures 9 (b) and (d), the effect of packet loss model does not affect the effectiveness of the relay path, compared with that using TCP proxy. Based on these results, if we use TCP throughput for the metric in the overlay routing, introducing the TCP proxy mechanism is key to improving the performance.

Next, we present the distribution of the improvement ratio of the bandwidth-optimized two-hop and three-hop relay paths with respect to the direct path in Figure 10(a). In the figure, we also plot the improvement ratio of the bandwidth-optimized three-hop relay path with respect to the bandwidth-optimized two-hop relay path. In Figure 10(b), we present similar results for end-to-end latency. These figures indicate that by using the relay path, we can obtain a significant improvement in terms of both available bandwidth and end-to-end latency. However, the effectiveness of three-hop relay path is quite limited when compared to two-hop relay path. Thus, seeking the three-hop relay path has a limited effect for overlay routing when we consider normal data transmission using a single path. However, when we consider multipath data transmission, three-hop relay paths may become possible candidates for path selection. The effectiveness of the three-hop relay path for multipath data transmission is discussed in Subsection 3.3.

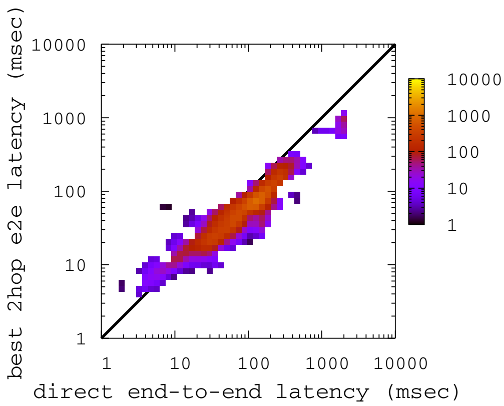


(a) Two-hop relay path

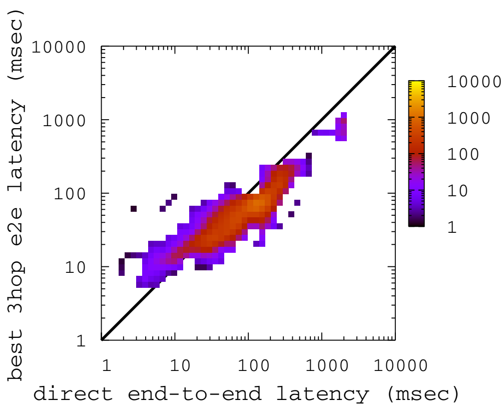


(b) Three-hop relay path

Fig. 7 Available bandwidths for the direct path and the bandwidth-optimized relay path



(a) Two-hop relay path



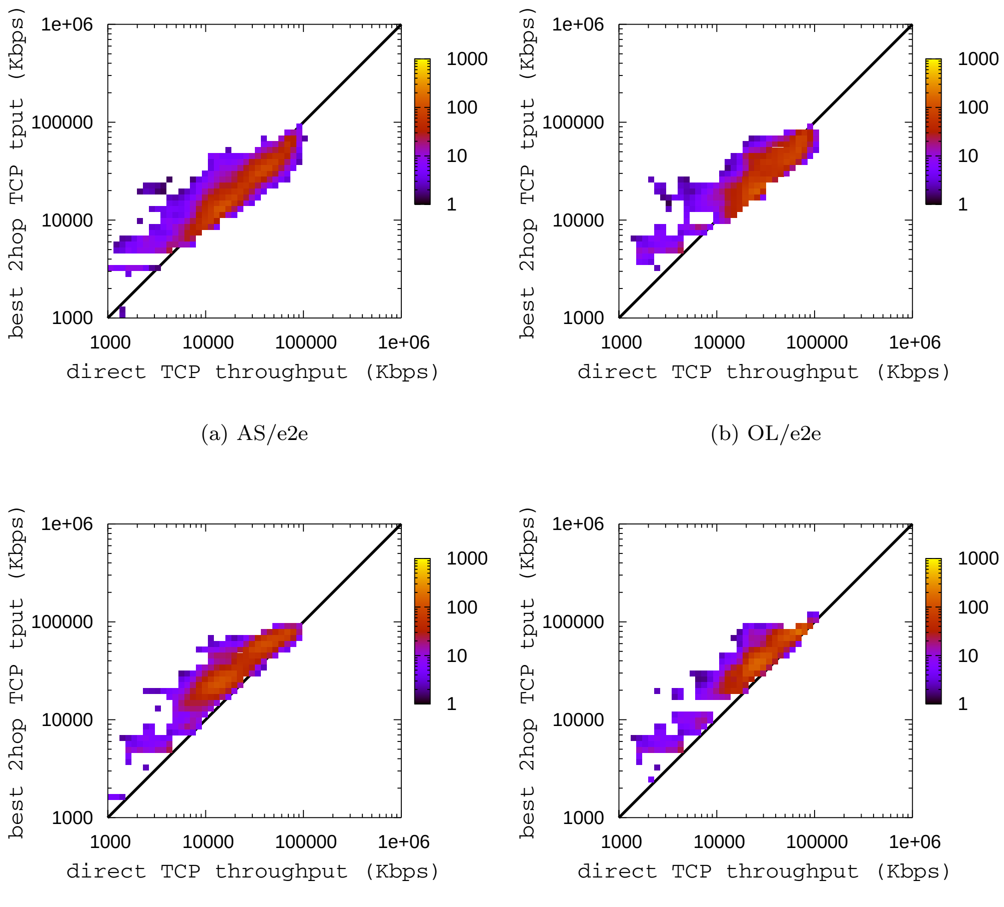
(b) Three-hop relay path

Fig. 8 End-to-end latencies for the direct path and the latency-optimized relay path

We present the distribution of the improvement ratio of the throughput-optimized two-hop path with respect to the corresponding direct path in Figure 11. The results of the cases of end-to-end latency and available bandwidth are also plotted in these graphs. Figure 11 shows that, for $50-70\%$ of all node pairs, we cannot find any two-hop relay path that has better performance than the corresponding direct path in the AS/e2e case and OL/e2e case. By using TCP proxy, however, we can obtain a performance gain similar to available bandwidth, when the improvement ratio falls between 1 and 2. However, in the region where the improvement ratio is larger than 2, the effectiveness of using TCP throughput does not reach that of available bandwidth. This is the same reason as in Figure 6. Namely, since the TCP throughput is affected by the packet loss ratio of the network, the performance gain of the relay path degrades.

3.3 Effectiveness in multipath transmission

We next investigate the effectiveness of seeking the three-hop relay path in multipath transmission. Here, we define multipath transmission as data transmission using multiple paths for one data transmission between source and destination nodes. We assume that we choose the multiple paths in the best order of available bandwidth or end-to-end latency from all of the direct, two-hop, and three-hop paths while considering the path disjointedness of selected paths. Note that considering the overlay-level disjointness is the first step in multipath transmission in overlay routing. For more precise evaluation, we need to know the physical-level disjointness of the overlay path. However, to do that, we must collect the information on the physical topology by additional mechanism such as full-mesh tracerouting. So, in this paper, we only consider the overlay-level disjointness.



(a) AS/e2e

(b) OL/e2e

(c) AS/pxy

(d) OL/pxy

Fig. 9 TCP throughputs for the direct path and the throughput-optimized relay path

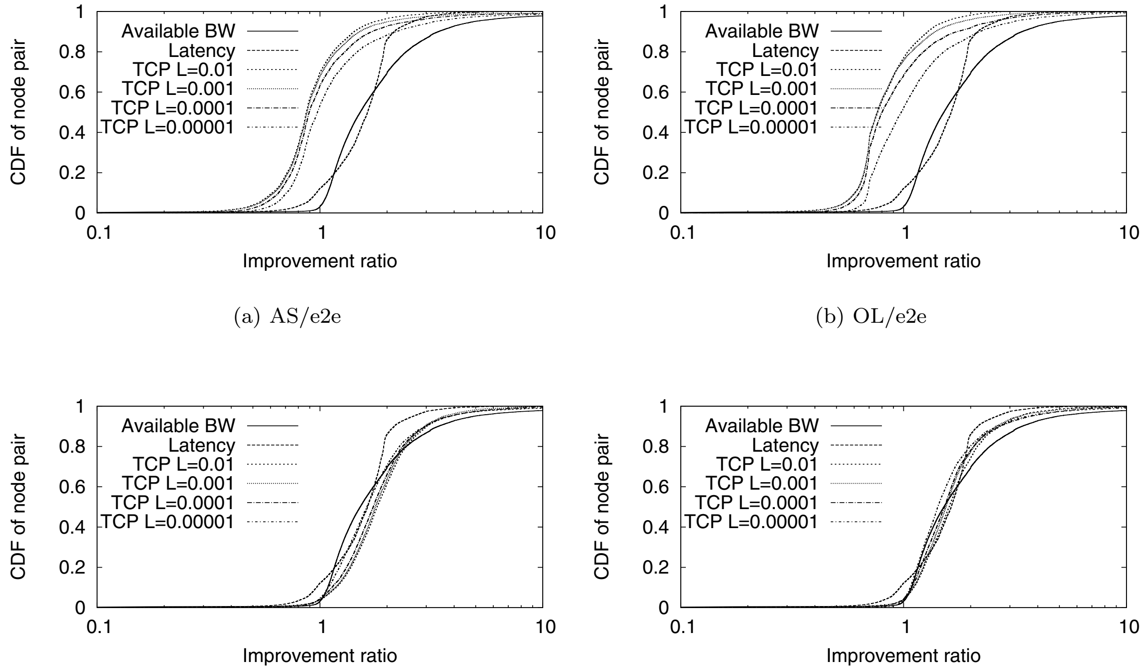
Figure 12 shows the average ratio of the number of direct, two-hop, and three-hop paths in the selected multiple paths, as a function of the total number of paths used in multipath transmission, when we use end-to-end latency and available bandwidth as a performance metric. This figure shows that seeking three-hop relay paths is meaningful in multipath transmission with a few paths, but its effectiveness gradually decreases as the number of total paths used in multipath transmission increases. This is because the number of available disjoint paths decreases when the number of paths used in multipath transmission.

3.4 Correlation between available bandwidth and end-to-end latency

Finally, we investigate the correlation between the improvement ratio in available bandwidth and the end-to-end latency, in order to clarify whether a relay overlay path that is "good" for available bandwidth is also good for end-to-end latency, and vise versa. Note that we do not consider TCP throughput, because this metric is calculated from end-to-end latency and available bandwidth. In Figure 13(a), we plot the relationship between the improvement ratio of the bandwidth-optimized two-hop relay path and the improvement ratio of the path in end-to-end latency. Figure 13(b) shows a similar graph for the bandwidth-optimized three-hop relay path.

These figures indicate that when we can find a multi-hop relay path that has a larger available bandwidth than the direct path, such path has a larger end-to-end latency than the direct path. That is, when we select the overlay path based on the available bandwidth, the selected path generally has a large end-to-end latency. Therefore, we should carefully choose the metric in selecting overlay paths according to the characteristics of upper-layer applications.

We also note that, when we cannot find a relay path that has a larger available bandwidth than the direct path ($x < 1.0$ in Figures 13 (a) and (b), such relay paths have a significantly larger end-to-end latency. In such cases, simply choosing the direct path is reasonable, regardless of the type of upper-layer applications.



(a) AS/e2e (b) OL/e2e

(c) AS/pxy (d) OL/pxy

Fig. 11 Distribution of improvement ratio for TCP throughput

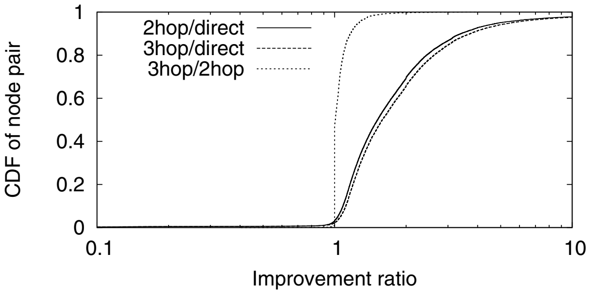
Figure 14(a), which is the converse graph to Figure 13(a), plots the relationships between the improvement ratio of the latency-optimized two-hop relay path and the improvement ratio of the path in available bandwidth. Figure 14(b) is a similar graph for the latency-optimized three-hop relay path. In contrast to the results shown in Figure 13, these figures indicate that when we choose the latency-optimized relay path, it is likely that the path also has a larger available bandwidth than the direct path. One possible reason for this is that when the latency of the overlay path decreases, the number of networks the path traverses is likely to decrease. It bring the decrease of the probability in which the path traverses the network with tight (narrow) link. This means that when we choose the path based on end-to-end latency, the path generally has a larger available bandwidth than the direct path.

One can imagine from these results that it is sufficient to select the overlay path based only on end-to-end latency and that it is meaningless to observe the available bandwidth. However, Figure 15, which plots the distribution of the ratio of the available bandwidth of the latency-optimized relay path with respect to the available bandwidth of the bandwidth-optimized relay path for all node pairs, clearly shows that the available bandwidth of the latency-optimized relay path is significantly smaller than that of the bandwidth-optimized relay path. That is, when we want to find a data transmission path with sufficiently large available bandwidth, we should measure the available bandwidths of the overlay network paths directly.

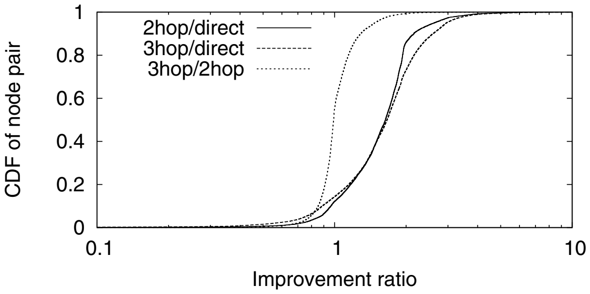
However, since a larger number of packets is required for measuring the available bandwidth than for measuring end-to-end latency, we propose one possible guideline for selecting the data transmission path in routing overlay networks for the bandwidth-centric applications. When we transmit the data to a destination for which there is insufficient information on the available bandwidth, we select the path based on end-to-end latency. When we have sufficient and accurate information on the available bandwidth, we choose the path based on available bandwidth.

4. Conclusions

In this paper, we focused on overlay routing based on delay and bandwidth information. We considered three metrics in selecting overlay route: available bandwidth, end-to-end latency, and TCP throughput. By investigating the effectiveness of overlay routing based on the assumption that the PlanetLab nodes make up the routing overlay network, the following results were obtained. When we select the bandwidth-optimized relay path, for most of node pairs, we could find a two-hop relay path that has higher available bandwidth than the corresponding direct path. When we select TCP's throughput-optimized relay path, introducing TCP proxy mechanism at relay nodes is key for ob-

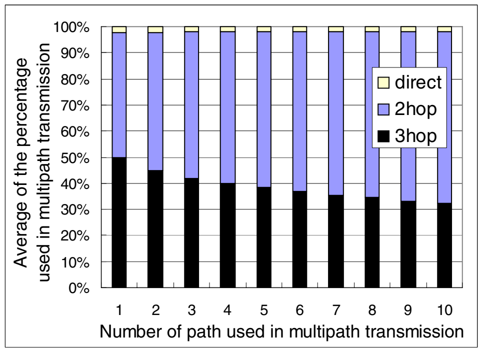


(a) Available bandwidth

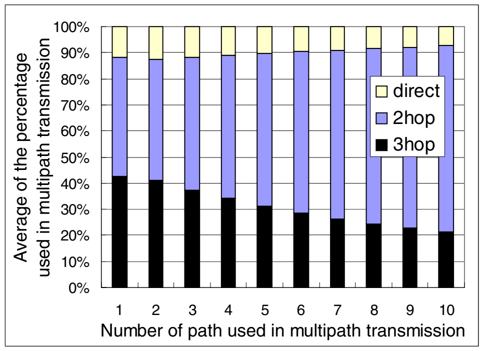


(b) End-to-end latency

Fig. 10 Distribution of improvement ratio for available bandwidth and end-to-end latency



(a) Available bandwidth



(b) End-to-end latency

Fig. 12 Breakdown of paths used in multipath transmission

taining a performance gain by overlay routing. We also found that the three-hop relay path becomes effective particularly when we deploy multipath data transmission. Furthermore, the latency-optimized relay path is likely to have larger available bandwidth than the direct path.

For future work, we plan to investigate the effectiveness of the multipath overlay routing with consideration of physical-level disjointeness.

References

[1] Z. Duan, Z.L. Zhang, and Y.T. Hou, "Service overlay networks: SLAs, QoS, and bandwidth provisioning," Proceedings of IEEE ICNP 2002, Nov. 2002.

[2] NetLightning Web Page. available at `http://www.netli.com/services/netlightning/`.

[3] Akamai Home Page. available at `http://www.akamai.com/`.

[4] D.G. Andersen, H. Balakrishnan, M.F. Kaashoek, and R. Morris, "Resilient overlay networks," Proceedings of 18th ACM Symposium on Operating Systems Principles, Oct. 2001.

[5] C.L.T. Man, G. Hasegawa, and M. Murata, "Inferring available bandwidth of overlay network paths based on inline network measurement," Proceedings of ICIMP 2007, July 2007.

[6] G. Hasegawa, M. Kobayashi, M. Murata, and T. Murase, "Free-riding traffic problem in routing overlay network," Proceedings of ICON 2007, Nov. 2007.

[7] M. Uchida, S. Kamei, and R. Kawahara, "Performance evaluation of QoS-aware routing in overlay network," Proceedings of ICOIN 2006, Jan. 2006.

[8] S. Kamei, "Applicability of overlay routing in Japan using inter-domain measurement data," Overlay Network Workshop, Dec. 2006.

[9] Y. Zhu, C. Dovrolis, and M. Ammar, "Dynamic overlay routing based on available bandwidth estimation: A simulation study," Computer Networks Journal, vol.50, pp.739–876, April 2006.

[10] D.G. Andersen, A.C. Snoeren, and H. Balakrishnan, "Best-path vs. multi-path overlay routing," Proceedings of ACM SIGCOMM conference on Internet measurement, Oct. 2003.

[11] S. Banerjee, C. Kommareddy, K. Kar, B. Bhattacharjee, and S. Khuller, "Construction of an efficient overlay multicast infrastructure for real-time applications," Proceedings of IEEE INFOCOM 2003, April 2003.

[12] PlanetLab Web Page. available at `http://www.planet-lab.org/`.

[13] Scalable Sensing Service. available at `http://networking.hpl.hp.com/s-cube/`.

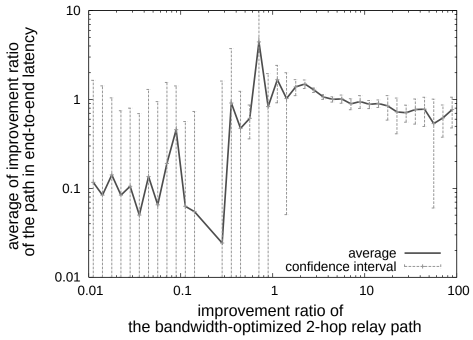
[14] traceroute.org. available at `http://www.traceroute.org/`.

[15] Skype Home Page. available at `http://www.skype.com/`.

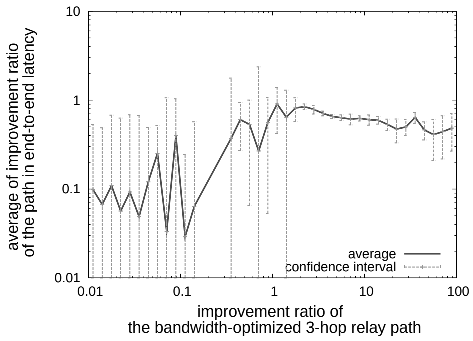
[16] YouTube Home Page. available at `http://www.youtube.com/`.

[17] GyaO Home Page. available at `http://www.gyao.jp/`.

[18] Bittorrent Home Page. available at `http://www.bittorrent.com/`.

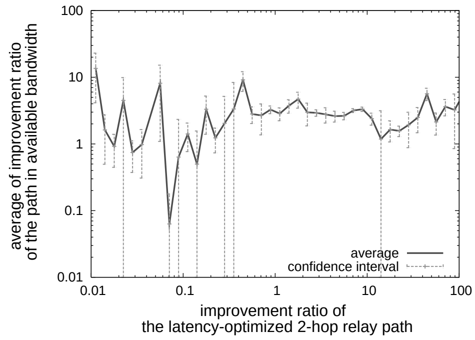


(a) Two-hop relay path

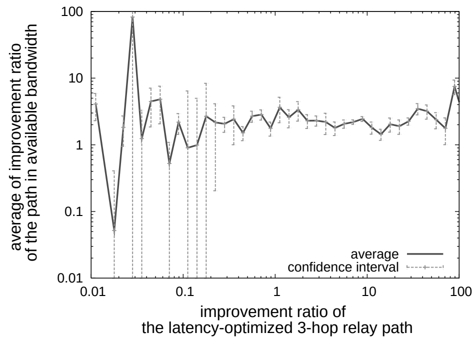


(b) Three-hop relay path

Fig. 13 Correlation between end-to-end latency and available bandwidth of overlay paths (1)



(a) Two-hop relay path



(b) Three-hop relay path

Fig. 14 Correlation between end-to-end latency and available bandwidth of overlay paths (2)

[19] M. Mathis, J. Semke, and J. Mahdavi, "The macroscopic behavior of the TCP congestion avoidance algorithm," ACM SIGCOMM Computer Communication Review, vol.27, No.3, July 1997.
[20] I. Maki, G. Hasegawa, M. Murata, and T. Murase, "Performance analysis and improvement of TCP proxy mechanism in TCP overlay networks," Proceedings of ICC 2005, May 2005.

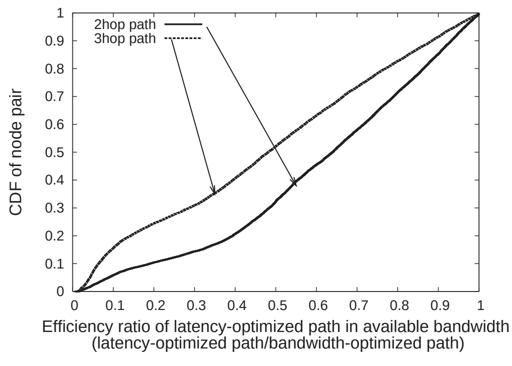


Fig. 15 Distribution of ratio of available bandwidth of latency-optimized relay path to that of bandwidth-optimized relay path

Go Hasegawa received the ME. and DE. degrees in Information and Computer Sciences from Osaka University, Osaka, Japan, in 1997 and 2000, respectively. From July 1997 to June 2000, he was a Research Assistant of Graduate School of Economics, Osaka University. He is now an Associate Professor of Cybermedia Center, Osaka University. His research work is in the area of transport architecture for future high-speed networks and overlay networks. He is a member of the IEEE and IEICE.

Yuichiro Hiraoka received the ME degree in Information Science and Technology from Osaka University in 2008. He is now at NEC System Technologies, Ltd.

Masayuki Murata received the ME and DE degrees in Information and Computer Sciences from Osaka University, Japan, in 1984 and 1988, respectively. In April 1984, he joined Tokyo Research Laboratory, IBM Japan, as a Researcher. From September 1987 to January 1989, he was an Assistant Professor with Computation Center, Osaka University. In February 1989, he moved to the Department of Information and Computer Sciences, Faculty of Engineering Science, Osaka University. From 1992 to 1999, he was an Associate Professor in the Graduate School of Engineering Science, Osaka University, and from April 1999, he has been a Professor of Osaka University. He moved to Advanced Networked Environment Division, Cybermedia Center, Osaka University in 2000, and moved to Graduate School of Information Science and Technology, Osaka University in April 2004. He has more than two hundred papers of international and domestic journals and conferences. His research interests include computer communication networks, performance modeling and evaluation. He is a member of IEEE, ACM, The Internet Society, IEICE and IPSJ.