Effectiveness of overlay routing based on delay and bandwidth information

Yuichiro Hiraoka, Go Hasegawa, Masayuki Murata
Graduate School of Information Science and Technology, Osaka University
1-5 Yamadaoka Suita, Osaka 565-0871, Japan
Email: {y-hiraok, hasegawa, murata}@ist.osaka-u.ac.jp

***Abstract*—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 only consider end-to-end delay, and there are few works focusing on bandwidth-related information, such as available bandwidth and physical capacity. In the present paper, we use the measurement results of delay and available bandwidth of network paths between PlanetLab nodes and investigate the effect of overlay routing using both delay and bandwidth information. We further reveal the correlation between the latency and available bandwidth of the overlay paths and propose several guidelines for selecting an overlay path.**

I. Introduction

As the Internet increasingly diversifies and the user population grows rapidly, new and varied types of service-oriented networks are emerging. Called service overlay networks include P2P networks, anonymous file-sharing services, audio and video conferencing services, and Content Delivery/Distribution Networks (CDNs). Service overlay networks are defined as upper-layer networks providing special-purpose services that are built on the lower-layer IP network. Therefore, their performances depend primarily on how well they take advantage of the characteristics and resources of the underlying IP network. To improve their performances, service overlay networks need fast and accurate information concerning the resource availability in the IP network to realize adaptive control mechanisms. Some examples of these control mechanisms are as follows:

- P2P networks. When a resource discovery mechanism finds multiple peers having the same requested contents, this information is used to determine which peer should transmit the contents.
- Grid networks. When multiple sites contain the same data, this information is used to determine from which site data will be copied or read.
- CDNs. When backup data or cached data is transmitted, this information can be used to prevent other network traffic from being deprived of resources during transmission.

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 Fig. 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 Akamai [1] distribute overlay nodes (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*. In Resilient Overlay Networks (RON) [2], 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 [2], 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 provide an alternate route faster than IP routing convergence.

The primary reason why overlay routing mechanisms can improve throughput and transmission time in data transfer is that the traditional IP routing operated by Internet Service Providers (ISPs) does not always determine the route according to user-perceived performance. In IP routing, the metrics determining the route are hop count and link loads, and the end-to-end delay and bandwidth-related information, which affect the data transmission throughput for short- and long-lived TCP connections, are not taken into account. In addition, inter-domain routing by Border Gateway Protocol (BGP) is based on autonomous system-level (AS-level) network topology and AS-level hop count, which are more abstracted than router-level IP network topology and hop count. Furthermore, most ISP-driven IP routings are configured by political and financial factors: the billing mechanism of transit links to upper-layer ISPs, the relationships between the ISP and other ISPs interconnected by public or private peering links, and the amount of traffic traversing transit and peering links. Therefore, the resulting IP routing policy cannot maximize network performance and user demand.

Several studies have examined the effectiveness of overlay routing with respect to IP routing [3–9]. For example, in [6], the authors used actual measurement data of the transmission latency among several geographically-distributed hosts in two ISPs in Japan and showed that a transmission latency of approximately 28% of end-to-end paths can be reduced by relaying another host, as compared to using the direct path. However, most of these studies are based only on the end-to-end delay between overlay nodes, whereas the bandwidth-

1-4244-1557-8/07/$25.00 ©2007 IEEE

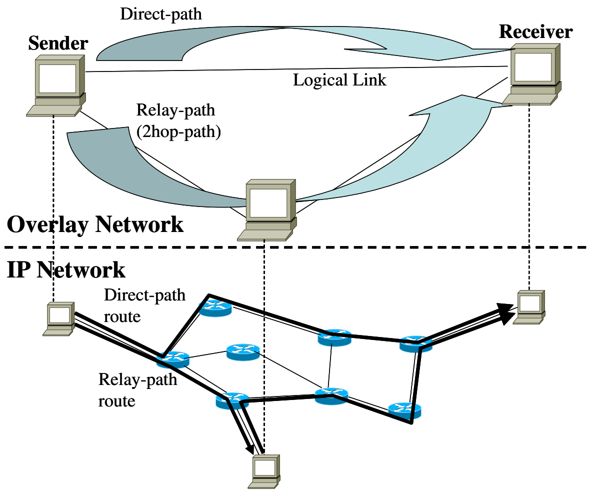


Fig. 1. Routing overlay network and overlay routing

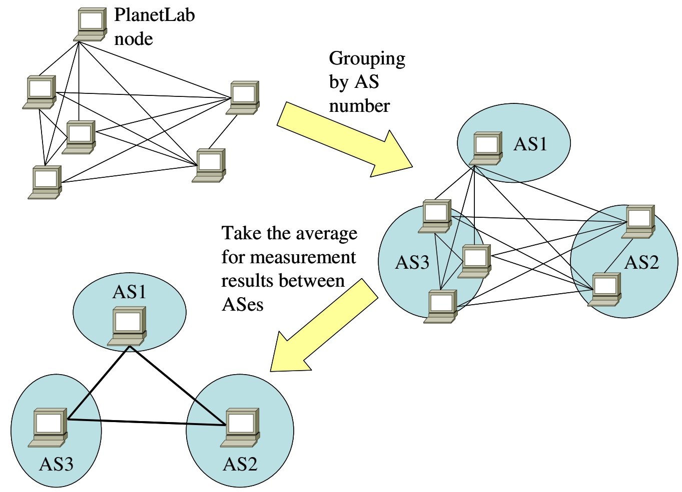


Fig. 2. Grouping nodes

related information such as physical capacity and available bandwidth is more important, especially for long-lived data transmission. As far as we know, there has been no previous work on the effectiveness of overlay routing based on the actual measurement data of bandwidth-related information.

In the present paper, we investigate the effectiveness of overlay routing, assuming that PlanetLab [10] nodes construct a routing overlay network. We utilize the measurement results obtained from Scalable Sensing Service (S^3) [11], which measures delay, bandwidth, and loss-related properties of network paths between PlanetLab nodes. In particular, we focus on the effectiveness of overlay routing when we use both delay and bandwidth information for selecting data transmission paths on the overlay network. One important result of the present study is the confirmation of the effectiveness of 3-hop relay overlay path, whereas almost all of the previous studies on overlay routing focused on the 2-hop relay overlay path, at most. Another interesting result is the correlation between transmission latency and available bandwidth of the end-to-end path. We revealed whether or not a network path with larger available bandwidth has smaller transmission latency, and vice versa.

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

II. Methodology

A. Measurement data

We investigate the effectiveness of overlay routing utilizing 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 the present paper, we use the data obtained on Oct. 25th, 2006 and evaluate the effectiveness of a routing overlay network constructed with PlanetLab nodes.

There exist 588 PlanetLab nodes in the measurement data utilized herein. However, some nodes are located in the same subnetwork, as estimated from the IP address and 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 three reasons: (1) The measurement results of transmission latency and available bandwidth between nodes in the same subnetwork may be quite small for latency and quite large for available bandwidth, which may overestimate the effectiveness of overlay routing. (2) The measurement results between nodes in the same subnetwork may include large errors, especially for available bandwidth. (3) There is almost no meaning in using a relay node in the same subnetwork as the source and destination nodes.

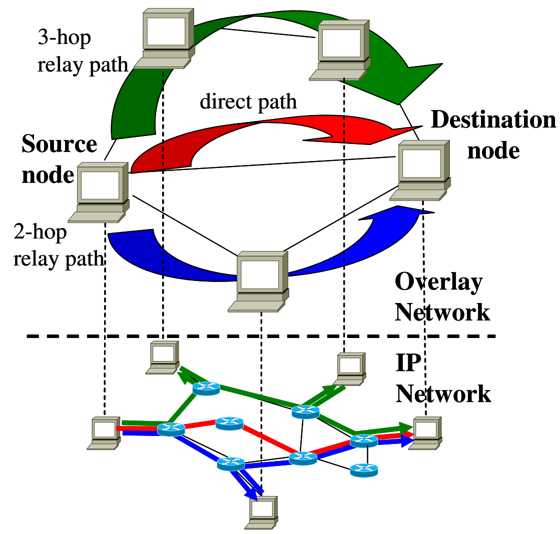


Fig. 3. Definition of overlay path

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 [12] 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). Fig. 2 depicts this process for node grouping.

B. Performance metrics

When one node (source node) selects the transmission path to another node (destination node), we compare the latency and available bandwidth of the following three candidates (Fig. 3):

- Direct path: the source node to the destination node
- 2-hop relay path: the source node to the destination node via a relay node
- 3-hop relay path: the source node to the destination node via two relay nodes

We set the latency and the available bandwidth for the direct path using the measurement results of end-to-end delay and available bandwidth, respectively. We further define the latency of a relay path as the sum of the latencies of direct paths constructing the relay path, and the available bandwidth of the relay path as the minimum of the available bandwidth of direct paths constructing the relay path. We assume the number of overlay nodes is M, and the latency and available bandwidth of the network path between node N_i and N_j is τ_{ij} and ρ_{ij} $(1 \leq i, j \leq M)$, respectively. Then, we can describe the latencies and available bandwidths of the direct path, the 2-hop relay path, and the 3-hop relay path, as follows:

$$D^1_{ij} = \tau_{ij} \tag{1}$$
$$D^2_{ikj} = \tau_{ik} + \tau_{kj} \tag{2}$$
$$D^3_{iklj} = \tau_{ik} + \tau_{kl} + \tau_{lj} \tag{3}$$
$$B^1_{ij} = \rho_{ij} \tag{4}$$
$$B^2_{ikj} = \min(\rho_{ik}, \rho_{kj}) \tag{5}$$
$$B^3_{iklj} = \min(\rho_{ik}, \rho_{kl}, \rho_{lj}) \tag{6}$$

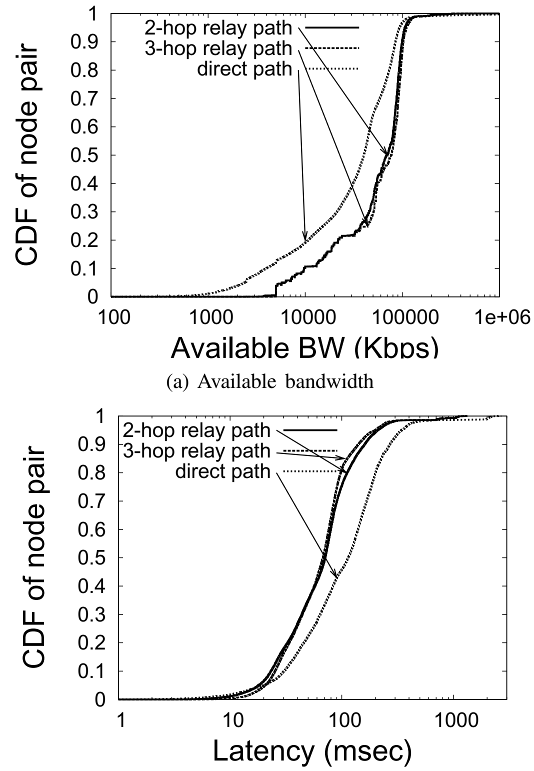
We denote that the relay node for the 2-hop relay path as N_k and the relay nodes for the 3-hop relay path as N_k and N_l $(1 \leq k, l \leq M,\ \ k \neq l,\ \ k, l \neq i, j)$.

Furthermore, we define the bandwidth-optimized path as the relay path that has the largest available bandwidth among all possible relay paths, and the latency-optimized path as the relay path that has the smallest latency. We can then obtain the respective latencies of the 2-hop and 3-hop latency-optimized paths and the available bandwidths of the 2-hop and 3-hop bandwidth-optimized paths as follows:

$$\hat{D}^2_{ij} = \min_{k \neq i,j}(D^2_{ikj}) \tag{7}$$
$$\hat{D}^3_{ij} = \min_{k \neq l,\ k,l \neq i,j}(D^3_{iklj}) \tag{8}$$
$$\hat{B}^2_{ij} = \max_{k \neq i,j}(B^2_{ikj}) \tag{9}$$
$$\hat{B}^3_{ij} = \max_{k \neq l,\ k,l \neq i,j}(B^3_{iklj}) \tag{10}$$

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

$$I(D^2_{ij}) = \frac{D^1_{ij}}{\hat{D}^2_{ij}} \tag{11}$$
$$I(D^3_{ij}) = \frac{D^1_{ij}}{\hat{D}^2_{ij}} \tag{12}$$
$$I(B^2_{ij}) = \frac{\hat{B}^2_{ij}}{B^1_{ij}} \tag{13}$$



(a) Available bandwidth

(b) Latency

Fig. 4. Distribution of latency and available bandwidth

$$I(B^3_{ij}) = \frac{\hat{B}^3_{ij}}{B^1_{ij}} \tag{14}$$

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

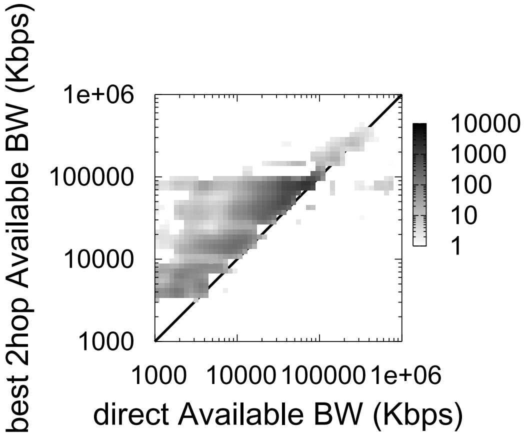
III. Evaluation results

A. Distribution of latency and available bandwidth

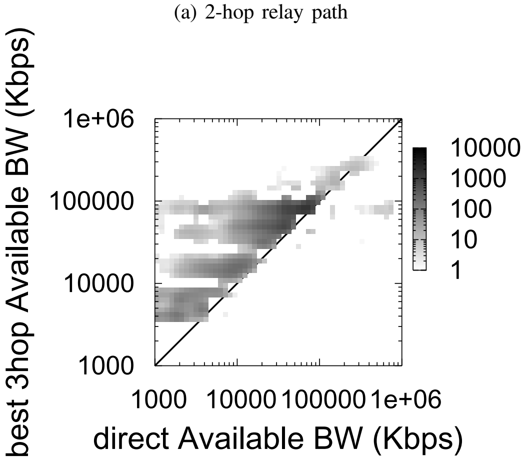
In Fig. 4, we show the distributions of latency and available bandwidth of direct paths and relay paths for all node pairs. We can observe from Fig. 4(a) that 80% of the direct paths have an available bandwidth of from 10 Mbps to 100 Mbps. However, using the relay path, the ratio increases to 90%. For latency (Fig. 4(b)), roughly half of the direct paths the 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 not good: less than 10 Mbps for available bandwidth and greater than 20 msec for latency. From these results, we can expect to find a relay path that has better performance than that of the direct path in both terms of latency and available bandwidth, especially when the performance of the direct path is not so good.

B. Characteristics of relay path

In Fig. 5, 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 2-hop relay paths (Fig. 5(a)) and 3-hop relay paths (Fig. 5(b)), respectively. Fig. 6 shows similar plots for latency. For 96.6% of all node pairs, we can find a 2-hop relay path that has



(a) 2-hop relay path



(b) 3-hop relay path

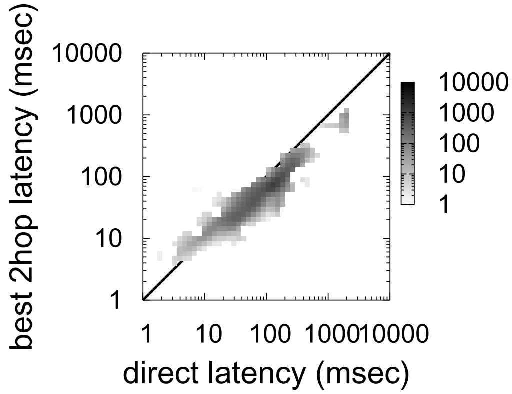
Fig. 5. Available bandwidths for the direct path and the bandwidth-optimized relay path

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 3-hop relay path that has a larger available bandwidth. For latency, these percentages decrease to 87.5% and 85.4%, respectively.

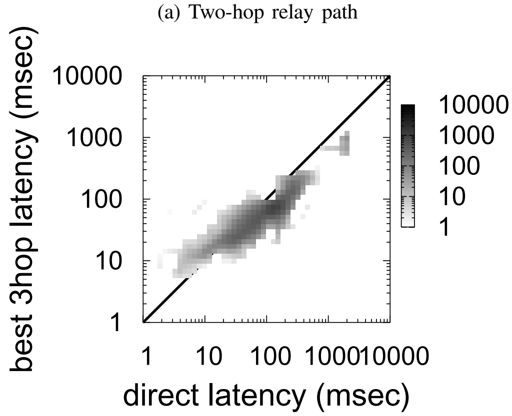
Furthermore, with respect to available bandwidth, 46.9% of node pairs for which a better 2-hop relay path cannot be found, a 3-hop relay path having a larger available bandwidth than the direct path can be found. In addition, for 51.6% of the node pairs that has a larger available bandwidth than the direct path, we can find a better 3-hop relay path than the bandwidth-optimized 2-hop relay path. With respect to 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.

Next, we present the distribution of the improvement ratio of the bandwidth-optimized 2-hop and 3-hop relay paths with respect to the direct path in Fig. 7(a). In the figure, we also plot the improvement ratio of the bandwidth-optimized 3-hop relay path with respect to the bandwidth-optimized 2-hop relay path. In Fig. 7(b), we present similar results for latency. These figures indicate that by using the relay path, we can obtain a significant improvement in terms of both available bandwidth and latency. However, the effectiveness of the 3-hop relay



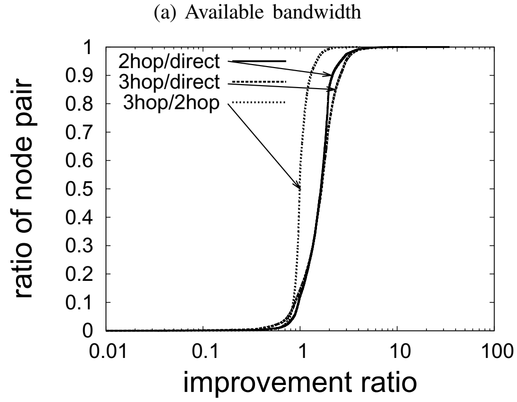
(a) Two-hop relay path



(b) Three-hop relay path

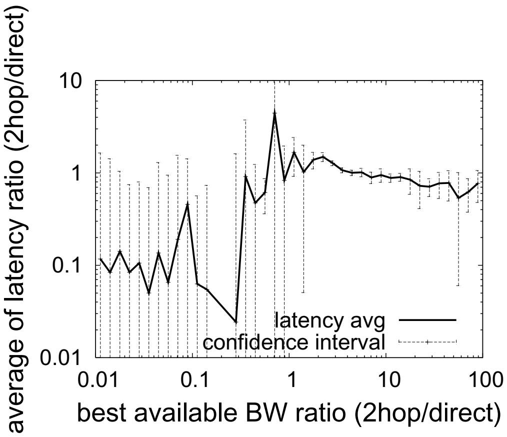
Fig. 6. Latencies for the direct path and the latency-optimized relay path

(a) Available bandwidth

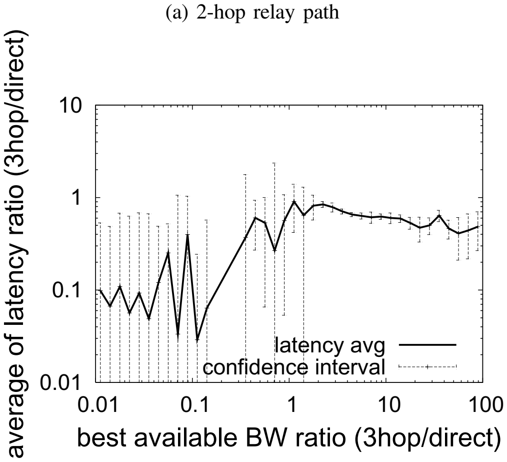


(b) Latency

Fig. 7. Distribution of improvement ratio



(a) 2-hop relay path



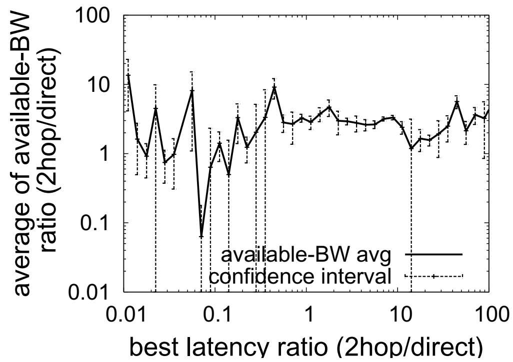
(b) 3-hop relay path

Fig. 8. Correlation between latency and available bandwidth of overlay paths (1)

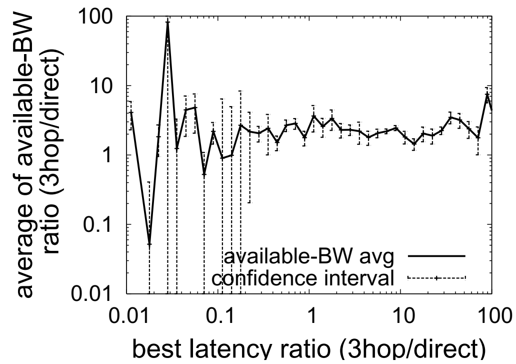
path is quite limited when compared to the 2-hop relay path. Thus, seeking the 3-hop relay path has a limited effect for overlay routing when we consider normal data transmission using a single path. However, when we consider multipath data transmission, 3-hop relay paths may become possible candidates for path selection. The effectiveness of the 3-hop relay path for multipath data transmission is discussed in Subsection III-D.

C. Correlation between available bandwidth and latency

We next investigate the correlation between improvement ratio in available bandwidth and latency, in order to clarify whether a "good" relay overlay path for available bandwidth is also good for latency, and vise versa. In Fig. 8(a), we plot the relationship between the improvement ratio of the bandwidth-optimized 2-hop relay path and the improvement ratio of the path in latency. Fig. 8(b) shows a similar graph for the bandwidth-optimized 3-hop relay path. From these figures, we observe the following: when we can find a multi-hop relay path that has a larger available bandwidth than the direct path, such a path has a larger 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 latency. Therefore, we should carefully choose the metric in selecting overlay paths according to the characteristics of upper-layer applications.



(a) 2-hop relay path



(b) 3-hop relay path

Fig. 9. Correlation between latency and available bandwidth of overlay paths (2)

From Fig. 8, when we cannot find a relay path that has a larger available bandwidth than the direct path ($x < 1.0$ in Fig. 8(a) and 8(b)), such relay paths have a significantly larger latency. In such cases, simply choosing the direct path is reasonable, regardless of the type of upper-layer applications.

Fig. 9(a) shows plots of the relationships between the improvement ratio of the latency-optimized 2-hop relay path and the improvement ratio of the path in available bandwidth. Fig. 9(b) is a similar graph for the latency-optimized 3-hop relay path. In contrast to the previous results (Fig. 8), 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. This means that when we choose the path based on latency, the path generally has a larger available bandwidth than the direct path.

These results may appear to indicate that it is sufficient to select the overlay path based only on latency and that it is meaningless to observe the available bandwidth. However, Fig. 10, which plots the distribution of the ratio of the available bandwidth of latency-optimized relay path to the available bandwidth of 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 directly measure the available bandwidths of the overlay network paths.

However, since we generally require a larger number of packets for measuring the available bandwidth than for measuring latency, we propose the following guideline for selecting the data transmission path in routing overlay networks for

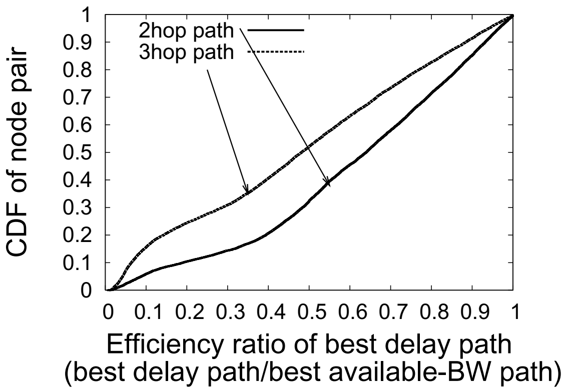
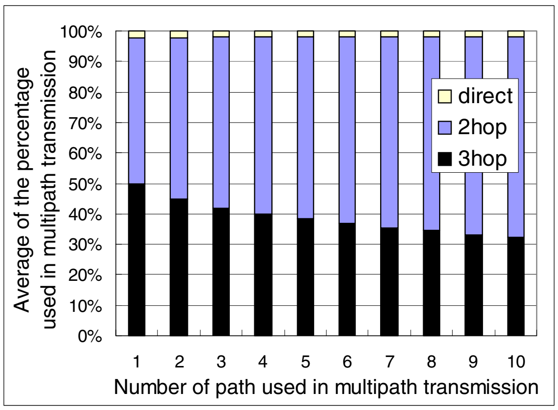
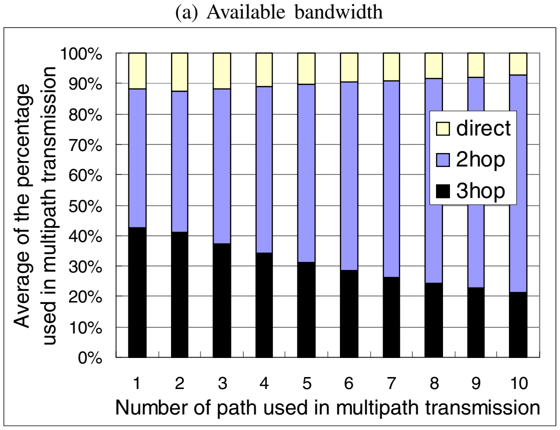


Fig. 10. Distribution of ratio of available bandwidth of latency-based best relay path to available bandwidth of available bandwidth-based best relay path



(a) Available bandwidth



(b) Delay

Fig. 11. Percentage of paths used in multipath transmission

the bandwidth-centric applications. When we transmit the data to a destination where we do not have sufficient information on the available bandwidth, we select the path based on latency. When we have sufficient and accurate information on the available bandwidth, we choose the path based on available bandwidth.

D. *Effectiveness in multipath transmission*

Finally, we investigate the effectiveness of seeking the 3-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 latency from all of the direct, 2-hop, and 3-hop paths with considering the path disjointness of selected paths. Fig. 11 shows the average ratio of the number of direct, 2-hop, and 3-hop paths as a function of the total number of using paths in multipath transmission. This figure shows that seeking 3-hop relay paths is meaningful in multipath transmission with a few paths, but its effectiveness decreases as the number of total using paths in multipath transmission increases.

IV. Conclusion

In the present paper, we evaluated the effectiveness of overlay routing based on both latency and available bandwidth, by using the measurement results in PlanetLab. The main results are as follows. (1) The available bandwidth-based overlay routing provided significant gain, compared with latency-based routing. (2) The effectiveness of the 3-hop relay path is limited in a single transmission, but would be effective in multipath transmission. (3) Small-latency paths generally have large available bandwidth, but large-available-bandwidth paths do not always have small latency.

In the future, we intend to evaluate the effectiveness of the path selection guideline proposed in Subsection III-C.

References

[1] Akamai Home Page. available at http://www.akamai.com/.
[2] D. G. Andersen, H. Balakrishnan, M. F. Kaashoek, and R. Morris, “Resilient overlay networks,” in *Proceedings of 18th ACM Symposium on Operating Systems Principles*, Oct. 2001.
[3] C. L. T. Man, G. Hasegawa, and M. Murata, “Inferring available bandwidth of overlay network paths based on inline network measurement,” in *Proceedings of ICIMP 2007*, July 2007.
[4] G.Hasegawa, M.Kobayashi, M.Murata, and T. Murase, “Free-riding traffic problem in routing overlay network,” in *Proceedings of ICON 2007*, Nov. 2007.
[5] M. Uchida, S. Kamei, and R. Kawahara, “Performance evaluation of QoS-aware routing in overlay network,” in *Proceedings of ICOIN 2006*, Jan. 2006.
[6] S. Kamei, “Applicability of overlay routing in Japan using inter-domain measurement data,” *Overlay Network Workshop*, Dec. 2006.
[7] 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, Apr. 2006.
[8] D. G. Andersen, A. C. Snoeren, and H. Balakrishnan, “Best-path vs. multi-path overlay routing,” in *Proceedings of ACM SIGCOMM conference on Internet measurement*, Oct. 2003.
[9] S. Banerjee, C. Kommareddy, K. Kar, B. Bhattacharjee, and S. Khuller, “Construction of an efficient overlay multicast infrastructure for real-time applications,” in *Proceedings of IEEE INFOCOM 2003*, Apr. 2003.
[10] PlanetLab Web Page. available at http://www.planet-lab.org/.
[11] Scalable Sensing Service. available at http://networking.hpl.hp.com/s-cube/.
[12] traceroute.org. available at http://www.traceroute.org/.