Exploiting AS Hierarchy for Scalable Route Selection in Multi-Homed Stub Networks

[Extended Abstract]

Sanghwan Lee
Computer Sci. and Engg.
University of Minnesota
sanghwan@cs.umn.edu

Zhi-Li Zhang
Computer Sci. and Engg.
University of Minnesota
zhzhang@cs.umn.edu

Srihari Nelakuditi
Computer Sci. and Engg.
University of South Carolina
srihari@cse.sc.edu

ABSTRACT

Multi-homing is a common practice among many (especially large) customer (or stub) networks. Although the purpose of multi-homing is primarily for enhanced reliability, it has also increasingly been used for load balancing and latency reduction. In this paper, we address the problem of how to perform *scalable* route selection in a multi-homed *stub* network to optimize network latency to various destinations as measured by round-trip-time (RTT). A straightforward method is to simply perform RTT measurements (e.g., using ping) to each destination via each provider and select the one with the minimum RTT as the "best" next-hop to the destination. Is there a more scalable alternative?

To answer this question, we carry out a measurement-based study to analyze the differences of RTTs in using two different providers in a multi-homed stub network to reach a large number of randomly selected destinations. Our study reveals that because of the AS hierarchy, for a large fraction of the network prefixes, the two AS paths through two providers merge in the core of the Internet. Furthermore, the router at which the two router level paths merge is actually in the AS at which the AS level paths merge. This phenomenon causes the RTT difference between the two paths through the two providers to be determined by the *non-shared* portion of the paths. Our study reveals that most of the two router level paths through the two upstream providers merge at the AS at which the two AS level paths merge. Based on this finding, we devise a scalable route (next-hop provider) selection algorithm using BGP information in a multi-homed stub network. We also present a preliminary evaluation.

Categories and Subject Descriptors

C.2 [**Computer Systems Organization**]: Computer - Communication Networks; C.2.1 [**Computer-Communication Networks**]: Network Architecture and Design

IMC'04, October 25–27, 2004, Taormina, Sicily, Italy.

General Terms

Algorithms, Measurement, Performance

Keywords

Multi-Homing, Route Selection, AS Hierarchy

1. INTRODUCTION

The Internet is a collection of separately administered *autonomous systems* (ASes). The Border Gateway Protocol (BGP) is the de facto standard inter-domain routing protocol used among ASes to exchange routing information for global reachability. Many customer (stub) networks, especially large ones, are often connected to multiple ISPs. This practice is referred to as *multi-homing.* The primary objective of multi-homing is to provide enhanced reliability during the failures of a provider network. However, it is also desirable to exploit multi-homing for performance optimization such as load balancing and minimizing network latency. As studied in [1], there are significant performance benefits in *dynamic* provider selection based on previous delay measurements because different providers show different latencies for the same destinations. To enable "intelligent" provider selection, a multi-homed customer network can apply certain routing policies in the BGP path selection, e.g., by appropriately setting the $LOCAL\ PREFERENCE$ attribute (see section 2 and [2]).

In this paper, we address the problem of how to perform *scalable* route selection in a multi-homed *stub* network to optimize network latencies to various destinations as measured by RTT. A straightforward method is to simply perform RTT measurements (e.g., using ping) to each destination via each provider and select the one with the minimum RTT as the "best" route/next-hop provider to the destination. Such a method clearly does not scale, as the number of destinations on the Internet is very large. So the question is whether there is a more scalable method. To answer this question, we set up a small measurement infrastructure in a multi-homed stub network with two commercial providers and conducted a series of experiments to collect RTT and traceroute data to a large number of destination networks (network prefixes) via the two providers. By analyzing the data sets we collected and the BGP routing information, we investigate the potential factors that contribute to the performance difference through the two providers.

Our major findings and contributions are summarized as follows. First, our analysis reveals that the AS hierarchy often causes the AS paths via the different providers to merge at the core of the Internet – the so-called *Dense* or *Transit Core* [3], resulting in shared common segments to destination ASes. Due to these shared AS path segments, the difference in the RTT performance via different providers are determined to a large extent by the RTTs from the stub network to the Dense Core or Transit Core of the Internet AS hierarchy. Furthermore, the AS paths for many different destinations follow the same path to the core of the Internet and then diverge for their own destinations. Thus, those destinations have similar RTT differences. Second, based on the above finding that the AS paths have shared segments, we propose a route selection algorithm that can reduce the amount of measurement. Applying the algorithm on our dataset, we can reduce the amount of measurement to only 4% of that of the straightforward approach. We believe that our findings are not unique to the stub network used in our measurement study, but also hold in general for other stub networks. In particular, the prevalence of the shared AS path segment in the Internet is corroborated by our analysis of BGP data from the Routeview project [4].

The remainder of this paper is organized as follows. Section 2 provides some background on BGP and AS hierarchy, and describes our measurement set-up and experiment methodology. In section 3, we show that the shared segments of paths are prevalent in the Internet. We propose a scalable route selection algorithm and evaluate the performance in section 4. Section 5 concludes the paper.

1.1 Related Work

As multi-homing is increasingly adopted by customer networks for enhanced reliability and other performance benefits, companies such as [5, 6] are offering commercial products to exploit dynamic provider selection for improved performance. However, no scientific studies have been published demonstrating their effectiveness.

The first academic research that provides a systematic analysis of the potential performance benefits of multi-homing is the study in [1], where extensive measurement data collected from web hosting and data center facilities are used. Using the measurement data, the authors compare the performance (e.g., response time) of web servers located *in the same city* but connected to different ISPs, and treat them *as if they were multi-homed.* Their study shows that there are significant performance benefits in *dynamic* provider selection based on previous delay measurements. Our work differs in that our focus is on understanding the factors such as the AS hierarchy that contribute to the difference in the performance of upstream providers and devising a scalable route selection algorithm. Hence our study complements and furthers the study of multi-homing pioneered by [1].

2. BACKGROUND, MEASUREMENT SETUP AND EXPERIMENTS

In this section, we first provide some necessary background on Internet AS hierarchy and BGP. We then describe our measurement set-up, the experiments we conducted and data we collected for our study.

2.1 BGP and Internet AS Hierarchy

The Internet consists of more than 14,000 network domains, or ASes. BGP is the de facto inter-domain routing

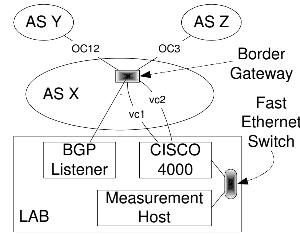

Figure 1: Multihomed Measurement Environment

protocol used among ASes to exchange routing information for global Internet connectivity [2]. BGP uses a basic path vector protocol to announce route updates – a list of ASes traversed, called AS path is included (among many other attributes) in the route updates as they propagate among ASes. The primary objective of BGP is to enable ASes to apply *policy control* in route selection, filtering and propagation. Because of the different business relations (e.g., customer-provider, peering [7]) formed among ASes and the resulting policies they use to filter, select and propagate BGP route updates, it is now well known that the (logical) Internet AS topology reveals a hierarchical structure, with a few ASes (so-called tier-1 ISPs) constituting the core of the Internet, where most of the Internet traffic traverses. In this study we adopt the classification of the Internet AS hierarchy proposed in [3], which categorizes ASes into five levels: *Dense Core* (level 1), *Transit Core* (level 2), *Outer Core*, *Regional ISPs*, and *Customer Networks*. The majority of ASes are customer (or stub) networks.

When an AS is connected to multiple ASes, it will likely receive multiple route announcements to a given destination network prefix, of which the AS selects the "best" route based on policy and other considerations. The standard BGP best route selection uses several attributes such as *Local Preference*, *AS Path* length and *MED* (Multi-Exit Discriminator) [2]. An AS typically influences the route selection process by manipulating the *Local Preference* attribute. To effectively take performance such as network latency into account when making route selection decisions, we need to understand the major factors that affect the network performance.

2.2 Measurement Setup

For our study, we set up a small measurement infrastructure in a multi-homed campus network (AS57, see Fig. 1). AS57 was connected to two commercial ISPs, AS1 and AS 3908, before Nov. 2003. After Nov. 2003, AS3908 has been replaced by AS7911. For brevity, AS57, AS1, AS3908, and AS7911 are referred to as AS X, AS Y, AS Z, and AS W, respectively. Both AS Y and Z belong to the Transit Core and AS W belongs to the Dense Core. The measurement environments are the same before and after the replacement. However, this replacement gives the effect of experimenting at two different stub networks. The measurement host in the lab has two IP addresses of a prefix in AS X. The border gateway router and CISCO 4000 are configured to select the outgoing ISP based on the source IP address of the packet. This way, the measurement host can choose the outgoing ISP by using different source IP addresses.

ping was used to measure the round trip time (RTT) from the measurement host to a given destination address via each of the two providers (AS Y - AS Z pair or AS Y - AS W pair depending on the measurement date). For each destination, two *ping* processes were launched concurrently through each of the two providers and measured the RTTs. The *traceroute* measurement was used to collect the router level path information. From the router-level paths, we also obtain the AS-level path information by mapping the IP addresses to their ASes using an AS mapping tool [8]. The BGP listener in the lab collects the BGP information.

2.3 Experiments and Data Processing

The measurement with the two providers, AS Y and AS Z was done from Aug. 15, 2003 to Oct. 29, 2003. Using BGP routing data collected on Aug. 15, 2003, we choose the set of all the network prefixes, and from each prefix, we randomly select two IP addresses (as in [8]). The resulting target IP address set contains a total of 246,932 IP addresses. By using *ping* messages, we collected RTTs of 65,631 "live" IP addresses. After discarding some destinations with inconsistent RTT responses and choosing only one destination for a prefix, we have obtained the final dataset consisting of the RTT value pairs (each through each provider) of *36,219* network prefixes. We, hereafter, use destinations and prefixes interchangeably. The traceroute measurement was done from Oct. 8, 2003 to Oct. 15, 2003 for the "live" IP addresses. We gathered 9,031 destinations with complete router level paths through the two providers. Since we removed the traceroute result with non-responding intermediate routers, the amount of *traceroute* data is less than that of the *ping* data.

The *ping* measurement with the two providers, AS Y and AS W was done from Feb. 17, 2004 to Mar. 21, 2004 for 259,918 IP addresses collected from the BGP data of Feb. 17, 2004. We collected the RTTs of 72,349 IP addresses. The traceroute measurement was done from Apr. 28, 2004 to May. 9, 2004 for the "live" 72,349 IP addresses. We gathered 14,934 destinations with complete router level paths through the two providers. Even though the AS Z has been replaced by AS W, the measurement environment and the experiments were almost the same. For more information about the measurement environment and the experiments, please refer to [9].

3. IMPACT OF THE SHARED SEGMENT ON THE DIFFERENCE OF RTTS

In this section, we analyze the collected data set and the BGP information to find out the potential factors that contribute to the difference of RTTs through different providers. We show that AS hierarchy causes the two paths to merge at the core of the Internet and thus the difference of RTTs to the destination can be estimated by the difference of RTTs to the router at which the two paths merge.

3.1 RTTs and Difference of RTTs

In this paper, we focus on the RTTs and the difference of RTTs through different providers. Before we show the measurement result, we define some notations. For each network prefix p in the final dataset,

- let $rtt_u(p)$ denote the (representative) RTT from our measurement host to p via the provider u, where u can be either Y, Z, or W.

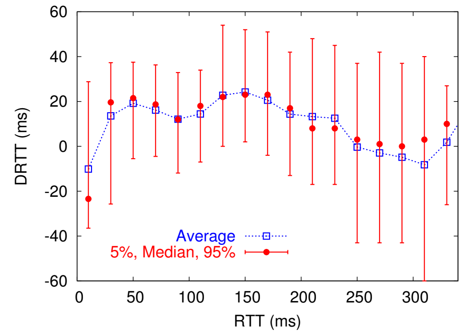

Figure 2: Distribution of DRTTs over RTTs to Prefixes : data with AS Y-Z pair

- let $drtt_{u,v}(p) = rtt_u(p) - rtt_v(p)$ be the difference of RTTs through two different providers, u and v. When the providers pair (Y-Z pair or Y-W pair) does not matter, we just use $drtt(p)$. Furthermore, we use DRTT as the abbreviation of $drtt(p)$.

We investigate whether the DRTTs are correlated to the network distance of a destination network from the stub network. We use $\min\{rtt_Y(p), rtt_Z(p)\}$, i.e., the minimum of the two RTTs as the network distance, and then group the prefixes into 20ms-bins accordingly: [0ms, 20ms), [20ms, 40ms), etc. Fig.2 shows the average, median, 5th, and 95th percentile of the DRTTs to the prefixes in each 20ms-bin. We do not see any clear correlation between the distance and the DRTTs. Especially in the range of [20ms, 200ms), the DRTT distributions do not differ very much: the averages and medians oscillate around 20ms. We conjecture that this phenomenon occurs because the two paths merge at some router so that the DRTT can be determined by the distance from the source to the router at which the two paths merge. To verify our conjecture, we investigate the prevalence of the shared segments among different destinations and the correlation between the shared segments and DRTT in the following.

3.2 Prevalence of Shared Segments

Before we discuss the prevalence of the shared segments, we introduce some notations. Fig.3 shows two AS paths from a source X via providers Y and Z to a destination D that merge at an AS M. Let *merging AS* be the AS at which the two AS paths merge. Similarly *merging router* is defined as the router at which the two router level paths merge. Let an AS (router) level *non-shared segment pair* denote the two paths from the source AS (router) to the merging AS (router). Likewise, an AS (router) level *shared segment* denotes the path from the merging AS (router) to the destination AS (router). In Fig.3, the pair of the segments $XY...M$ and $XZ...M$ is the *non-shared segment pair* for the two AS paths from X to D. It should be noted that in Fig. 3, the ingress routers of the two paths to the merging AS can be different.

In the path analysis, we first compute the ratio of the AS path length from X to the merging AS over the entire AS path length based on the BGP information collected on Feb. 17, 2004. This provides some insight on how many AS path pairs merge along the path to the destination and where the merging ASes are. We choose the path length through AS Y

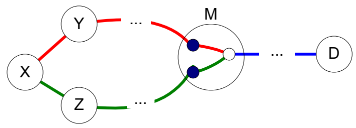

Figure 3: Two AS paths from X to D through the two providers Y and Z merge at AS M.

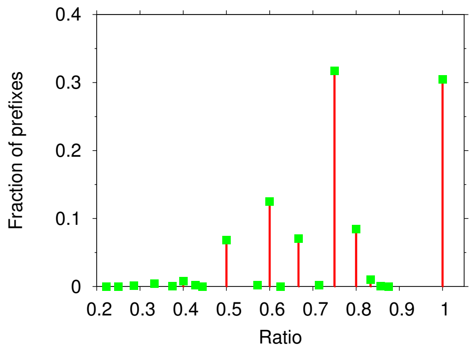

Figure 4: Ratio of the non-shared segment length over the whole path length in AS level (the provider Y-W pair) : Each bar with a dot represents the fraction of prefixes which have the corresponding ratio.

as the representative path length to the destination. Fig. 4 shows the distribution of this ratio. Around 30% of the prefixes have the ratio 1, i.e., their paths merge at the destination. For almost 70% of the prefixes, their paths merge along the way. This result shows that the shared paths are prevalent from the viewpoint of a stub network. The analysis on the AS hierarchy positions of the merging ASes in the BGP information shows that about 54% of the paths merge at the dense core and 21% of the paths merge at the transit core. This implies that the existence of AS hierarchy causes the paths to merge in the core of the Internet.

This result is not specific to the stub network AS X. Even though there are many real multihomed stub networks, we do not have access to the BGP information of the networks. Thus, to verify that the shared segments are prevalent in other networks, we use Routeviews data as the model of other stub networks. Routeviews data contains a number of BGP routing tables [4]. We use the table of RIBs in 'sh ip bgp' format collected on Oct 6th 2003. We select 2 ASes from a city and treat them as the two providers of an imaginary stub network called *virtual stub network*. We have 36 such AS pairs, which constitute 36 virtual stub networks. For each virtual stub network, we compute the fraction of AS path pairs merging at each hierarchy position. Our analysis shows that for most virtual stub networks, more than 50% of the AS path pairs merge at the dense core, which indicates that the shared segments are prevalent in the Internet.

We now look at the relationship between the merging AS and the merging router. Since there are multiple peering points between two ASes, the two router level paths may not merge at the merging AS. However, as can be seen in Fig. 5, for 83% to 92% of the destinations, the merging router is located in the merging AS. In Fig. 5, the x-axis represents the difference between the merging AS and the AS of the merging router. The difference 0 means that the merging AS

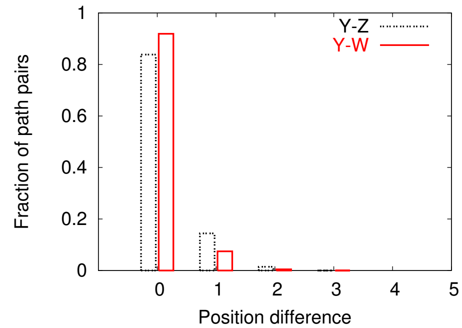

Figure 5: Merging AS vs AS of the Merging router : Y-Z represents the result with providers Y-Z pair and Y-W represents the result with providers Y-W pair.

is the same as the AS of the merging router. The difference 1 means that the AS of the merging router is the next hop AS of the merging AS towards the destination. This result shows that the shared segments are prevalent between two router level paths via the two providers and the merging routers are mostly located in the merging ASes. It should be noted that the the merging router may not be in the merging AS because of multiple peering points between ASes.

3.3 Dependency of DRTTs on the Non-Shared Segment Pair

We have three observations on the relationship between the existence of shared segments and DRTTs. First, the destinations with the same *router* level non-shared segment pairs have the same DRTT. Second, the DRTT to a destination can be accurately estimated by the DRTT to the merging router. Third, the DRTTs are independent from the RTTs to the destinations. as we see in 3.1, but they are related to the RTTs to the merging router.

The first observation is true because the DRTT is determined by the non-shared segment pair and if the destinations have the same non-shared segment pairs, their DRTTs should be the same. Regarding the second observation, the DRTT to a destination is the same as the DRTT to the merging router if the non-shared segment pair to the merging router is the same as the non-shared segment pair to the destination. However, since a router level path is determined by the destination IP address of the packet, if the destination IP address of the packet is the IP address of the merging router, the packet might follow a path different from the path taken when the destination IP address of the packet is the IP address of the destination. A traceroute-like probing tool can overcome this difficulty by using the IP address of the destination as the destination IP address of the packet.

The third observation is based on our intuition that many destinations with different distances have the same router level non-shared segments. To corroborate the above intuition, we group all the destinations according to their merging routers, i.e., if two destinations share the same merging router, they are included in the same group. Fig. 6 displays the range of absolute RTTs for destinations within each group. It is clear that destinations with wide range of RTTs may have the same merging router. These three

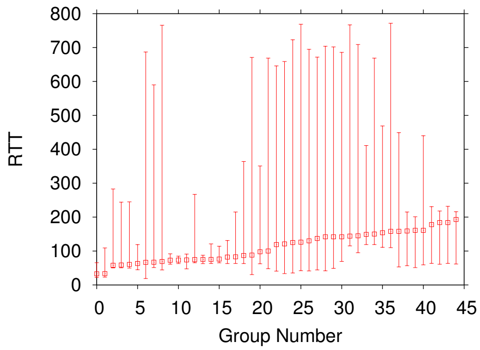

Figure 6: Min, Median, and Max of RTTs in each group categorized by merging routers. Groups are sorted by their Median. For the purpose of clarity, only the groups with more than 20 prefixes are shown.

SRS (input : prefix set $\mathcal{P}$, providers u and v)
1 Group the prefixes in $\mathcal{P}$ based on the non-shared segment pairs in the AS level.
2 For each group $\mathcal{G}$,
3 Randomly select a prefix p in $\mathcal{G}$
4 Run two traceroutes to p, one through each provider
5 Find out the two ingress routers, i_u and i_v in the merging AS of $\mathcal{G}$
6 Measure the RTT to i_u through provider u and the RTT to i_v through provider v, i.e. $rtt_u(i_u)$ and $rtt_v(i_v)$
7 Choose the provider with smaller RTT as the outgoing provider for all the prefixes in $\mathcal{G}$
8 Return the selected (prefix, provider) pairs

Figure 7: Scalable Route Selection Algorithm

observations on the shared segments provide the basis for devising a scalable route selection algorithm. Especially, the first observation allows the algorithm to reduce the measurement overhead. We look at the details of the algorithm in the next section.

4. ROUTE SELECTION ALGORITHMS

The route selection in a stub network based on the RTTs through different providers requires active probes to the destination prefixes. A straightforward method that simply performs RTT measurements to each destination via each provider clearly does not scale, as the number of destinations on the Internet is very large. In this section, we propose a *scalable* route selection algorithm with small accuracy degradation by exploiting the three observations discussed in 3.3. As discussed in 3.3, the DRTT to a destination can be estimated by the DRTT to the merging router. If many destinations have the same non-shared segment pairs, one measurement to the merging router will suffice. However, identifying the merging routers of the destinations requires traceroute to all the destinations, which is the same amount of measurement as that of the straightforward approach. We should have other probing targets that can be identified with much less measurement overhead.

We propose to use two *ingress* routers of the merging AS as the probing targets. This can dramatically reduce the DRTT measurement to the destinations. First, the ingress router pairs are likely to be the same for all the prefixes with the same AS level non-shared segment pair if all the ASes use the same policy such as *Early Exit* [1]. The ingress router pair of the merging AS of a destination is identified by two traceroute probes to the destination. Once the ingress router pair is identified, the DRTT to the destination is estimated by $rtt_u(i_u) - rtt_v(i_v)$, where i_u is the ingress router through u and i_v is the ingress router through v at the merging AS. We call $rtt_u(i_u) - rtt_v(i_v)$ the DRTT to the ingress router pair. This DRTT estimation is used as the DRTT estimations for all the prefixes with the same AS level non-shared segment pair. The algorithm based on this approach is described in Fig. 7.

The estimation based on the ingress routers of the merging AS may not be accurate when the two ingress routers are far away and the *merging router* is close to one ingress router. One way to overcome this problem is to subdivide the groups formed in Step 1 of SRS algorithm based on the next hop AS of the merging AS towards the destinations. Then, we use the ingress routers of the next hop AS as the probing targets. Since most of the merging routers are in the merging AS as can be seen in Fig. 5, the new ingress routers can be the same or at least close to each other. We call this version SRS-N (SRS with Next hop AS). One problem of SRS-N is that it unnecessarily subdivides the groups with close (even the same) ingress routers. Instead of subdividing the groups unanimously, we can subdivide only the groups with the ingress routers far away from each other. We call this version SRS-E (SRS with Error estimation). Although it is not easy to measure the network distance between two (remote) ingress routers from a stub network, one can roughly estimate the RTT between the two ingress routers by computing the differences $|rtt_u(i_u) - rtt_u(i_v)|$ and $|rtt_v(i_u) - rtt_v(i_v)|$ because if i_u and i_v are close, the two differences will be small.

4.1 Performance Evaluation

To properly evaluate the accuracy of the algorithms, we need to measure the RTTs to the destinations in a group and the ingress routers of a randomly selected prefix in the group simultaneously because route changes can cause some errors in RTT measurements. However, the data set has been collected over a long period, so it does not provide such information. Thus, in this evaluation, we just compute the absolute difference (error) between the DRTT to each destination and the DRTT to the ingress routers for the destination based on the traceroute data. Fig. 8 shows the cumulative distribution of the absolute errors. The keys with "SRS" show the results of using the ingress routers of the merging AS. The keys with "SRS-N" show the results of using the ingress routers of the next hop AS of the merging AS. The median absolute errors of "SRS" are about 10 ms and 15 ms for provider Y-W pair and Y-Z pair respectively.

[1]The assumption that the ingress router pairs (furthermore the router level non-shared segment pair) are the same for all the prefixes with the same AS level non-shared segment pair may not always be true, especially when the merging AS has multiple peering points and the neighboring ASes use different policies. Since the violation of the assumption can degrade the accuracy of the estimation, we plan to study the validity of this assumption in the future work.

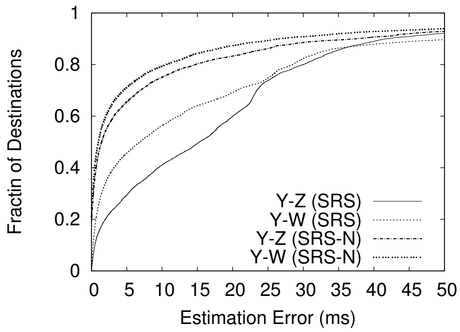

Figure 8: Cumulative distribution of the absolute errors of the DRTT estimation : The absolute error is the absolute difference of the DRTT to the ingress routers and the DRTT to the destinations.

However, the results of Y-Z and Y-W pairs with "SRS-N" show that 80% of the destinations have absolute errors of less than 10 ms. This evaluation can be incomplete due to several limitations. First, the RTT measurements using traceroute may not be accurate. Second, since we use different source IP addresses for the two packets through the two providers, the router level paths can be different from the merging router to the destination and from the destination back to the source because of the load balancing. Third, the result does not account for the fact that the ingress routers can be different for the prefixes in $\mathcal{G}$ in Fig. 7. Even with these limitations, the evaluation shows that the ingress routers can be used as the probing targets.

Next, we compute the measurement overhead reduction of SRS algorithm. The number of total prefixes is 129,959 in the BGP information for Y-W pair collected on Feb. 17, 2004. After applying SRS algorithm, the number of groups is 5,216. The required distance measurement is only 4% of that of the straightforward one. After applying SRS-N, the number of groups is 15,351, which is 11.8% of the measurements of the straightforward one. It should be noted that for each measurement for a prefix, additional traceroute measurement is required (line 4 in Fig. 7). The analysis of the BGP information of other dates shows the similar result. Regarding the performance of SRS-E, we believe that if we use SRS-E, the accuracy will be higher than SRS-N, but the measurement overhead will be smaller than SRS-N because SRS-E selectively subdivides the groups. The evaluation of the tradeoff between the number of groups and the accuracy is one of the future works.

Other approaches are possible to reduce the active probes. Instead of using active probing, the RTTs can be measured passively by observing TCP packets. However, it is not always possible to obtain the RTTs for a destination through both the providers. Another approach one might think of is to use active probes only to the small number of prefixes with large traffic [10]. Our approach can be used in conjunction with that approach to further reduce the active probes.

5. CONCLUSION AND FUTURE WORKS

In this paper, we studied the prevalence of the shared segment between the two paths through the two providers. Our study confirmed that majority of the destinations have shared segments and the merging ASes are mostly in the dense core. Based on the observations on the existence of the shared segments, we proposed scalable route selection algorithms. The preliminary evaluation showed that the SRS algorithms can reduce the measurement overhead and SRS-N can improve the accuracy of SRS.

Since this work is on going, we conclude the paper by emphasizing the future work. First, we plan to evaluate SRS algorithms with more consistent data sets. We plan to run more rigorous RTT measurements to ingress routers, merging routers, and destinations. Second, we plan to evaluate the tradeoff between the number of groups and the accuracy. The accuracy depends on the RTT between ingress routers. If we want to reduce the RTT between ingress routers, the number of groups will increase. We see a clear tradeoff between these two metrics. Finally, we plan to apply a similar method (using shared segments) on path quality monitoring. The overhead of monitoring the paths from a multi-homed stub network to other destinations can be reduced by exploiting the existence of the non-shared segments.

6. ACKNOWLEDGEMENTS

We thank our shepherd Constantine Dovrolis and the anonymous reviewers for their constructive feedback on this work. We also thank NTS people of University of Minnesota for providing the CISCO router and the necessary configuration for our testbed. This work was supported in part by the National Science Foundation under the grants ANI-0073819, ITR-0085824, and CAREER Award NCR-9734428.

7. REFERENCES

[1] Aditya Akella, Bruce Maggs, Srinivasan Seshan, Anees Shaikh, and Ramesh Sitaraman. A measurement-based analysis of multihoming. In *Proceedings of ACM SIGCOMM 2003*, karlsruhe, Germany, August 2003.

[2] Sam Halabi and Danny McPherson. *Internet Routing Architectures.* CISCO PRESS, second edition, 2001.

[3] Lakshminarayanan Subramanian, Sharad Agarwal, Jennifer Rexford, and Randy H. Katz. Characterizing the internet hierarchy from multiple vantage points. In *Proc. IEEE INFOCOM*, New York, NY, June 2002.

[4] University of oregon route views project. http://www.routeviews.org.

[5] Route science company. http://www.routescience.org.

[6] Proficient networks. http://www.proficient.net.

[7] Geoff Huston. Interconnection, peering and settlements–part i. In *Internet Protocol Journal*, Jun 1999.

[8] Zhuoqing Morley Mao, Jennifer Rexford, Jia Wang, and Randy H. Katz. Towards an accurate as-level traceroute tool. In *Proceedings of ACM SIGCOMM 2003*, karlsruhe, Germany, August 2003.

[9] Sanghwan Lee, Zhi-Li Zhang, and Srihari Nelakuditi. Impact of as hierarchy on multihoming performance : A stub network perspective. Technical report, Computer Science Department, Univ. of Minnesota, February 2004.

[10] N. Feamster, J. Borkenhagen, and J. Rexford. Guidelines for interdomain traffic engineering. *SIGCOMM Comput. Commun. Rev.*, 33(5):19–30, 2003.