Internet Routing Dynamics and NSIS Related Considerations

Charles Shen and Henning Schulzrinne

Department of Computer Science
Columbia University
{charles, hgs}@cs.columbia.edu

Sung-Hyuck Lee and Jong Ho Bang

SAMSUNG Advanced Institute of Technology
{starsu.lee, jh0278.bang@samsung.com}

February 24, 2005

Abstract

Design of Next Step In Signaling (NSIS) protocol and IP routing interaction requires a good understanding of today's Internet routing behavior. In this report we present a routing measurement experiment to characterize current Internet dynamics, including routing pathology, routing prevalence and routing persistence. The focus of our study is route change. We look at the types, duration and likely causes of different route changes and discuss their impact to the design of NSIS. We also review common route change detection methods and investigate rules to determine whether a route change happened in a node's forward-looking or backward-looking direction is detectable. We introduce typical NSIS deployment models and discuss specific categories of route changes that should be considered in each of these models. With the NSIS deployment models in mind, we further give experimental evaluation of two route change detection methods - the packet TTL monitoring method and a new delay variation monitoring method.

Contents

1 Introduction

Some Internet applications require special network treatment in order to function properly. For example, real-time multimedia applications need Quality of Service (QoS) support for their delay or loss sensitive packets; applications running in Network Address Translation (NAT) and firewall environment need to have these NAT and firewall configured appropriately. Request for special network resources or services may be accomplished by invoking Internet signaling protocols that install and manipulate control information along the path of the application flow. A well known example is the Resource ReSerVation Protocol (RSVP) [1]. Originally designed for QoS signaling alone, RSVP has been extended in many ways for other purposes, such as NAT and firewall traversal [2], traffic-engineering [3, 4], mobility transparent QoS [5]. Generally the trend is to develop a single signaling protocol suitable for various applications. But the "RSVP plus extensions" approach did not turn out to be an optimal one [6]. Therefore the Internet Engineering Task Force (IETF) has chartered the Next Step In Signaling (NSIS) working group to develop a new generic protocol for signaling [7].

Effective operation of the NSIS protocol relies on the consistency of the path where control state information is installed and the path where data flow is forwarded. If this consistency is broken, the portion of data path that is not covered by related control mechanisms may severely affect the application performance. In the real-time QoS application example, lack of QoS in some segments of the end-to-end path may lead to an unacceptable overall delivery quality.

Internet route change is a major cause to the divergence of the signaling state information path and data forwarding path. Therefore, interaction between NSIS protocol and route change is an important aspect that needs to be addressed. The first step toward this problem is to detect the route change. The NSIS community has proposed three categories of route change detection methods [8]: routing monitoring (local trigger, extended trigger), packet monitoring (GIMPS C-mode monitoring, data plane monitoring) and GIMPS D-mode probing. The applicability of each method is yet to be studied.

The goal of this report is to characterize various route changes in today's Internet, and discuss their impact to NSIS design. Meanwhile, we try to evaluate the applicability of certain route change detection mechanisms. The first half of this report presents an Internet routing measurement carried out from April to August, 2004. We describe the measurement methodology and details of our data sets. Then we look at various statistics of routing pathologies, route changes, AS path changes as well as route prevalence and route persistence. We compared our results with some of the related work (Section 2) and summarized our findings from an NSIS perspective. The latter half of this report deals specifically with route change detection. After a brief overview of existing route change detection methods, we grouped route changes into two categories from a node's point of view: those happen in the node's forward-looking direction and those in its backward-looking direction. We use an example to illustrate how to determine whether route changes in either direction can be detected in any given network. We then look at route changes from the network's point of view and classify route changes according to their scales, specifically, Inter-Autonomous System (Inter-AS) and various kinds of intra-AS route changes. With that in mind, we introduce different NSIS deployment models and which types of route changes should be considered in each model. We further extend our discussion to the evaluation of two packet monitoring based route change detection methods. One is the simple, low cost TTL monitoring approach. The other is packet delay variation monitoring, a method that has not yet been discussed in the NSIS working group documents. Finally we summarize our results and present future work.

2 Related Work on Internet Routing Measurement

Recent years have seen many research work addressing different aspects of Internet routing measurement. Some propose ways to improve the deployment of measurement infrastructure and to perform a sound Internet measurement [9, 10]; some study improvement of measuring a particular path property such as path Round Trip Time (RTT) [11]; Some based directly on routing protocol messages to study interdomain Border Gateway Protocol (BGP) routing behavior [12–15] and how intradomain protocols may influence the path selection process in BGP [16].

We believe that the work most related to ours is Paxson's end-to-end routing dynamics study [17–20]. Paxson devised a measurement framework where a number of sites around the Internet host a specialized measurement service called Network Probe Daemon (NPD) and conducted a comprehensive study on end-to-end routing and end-to-end packet dynamics. He and others further developed the NPD framework to an infrastructure called National Internet Measurement Infrastructure (NIMI) [21], which attempts to address the limitation and biases of NPD. Using NIMI and a pool of public servers, Zhang et. al. performed another measurement and studied the constancy of Internet path properties [22–24]. They explored various notions of constancy in the context of routes, loss, delay and throughput and found that overall, one can count on constancy on at least the time scale of minutes.

Paxson's measurement was carried out during year 1994 and 1995 and Zhang's during winter 1999-2000. Both measurements used Poisson intervals to achieve an unbiased time average computed from measurements; both characterized routing pathology, routing prevalence and routing persistence. Zhang compared their measurements with the previous one. The conclusion was that the routing during the second measurement had neither improved nor degraded significantly since first measurement. Routing prevalence in the second measurement showed a considerable increase, but because of the difference in the methodology of the two measurements, the authors did not conclude whether this increase of prevalence values indeed reflects change in Internet routing behavior.

3 Routing Measurement Experiments and Data Sets

3.1 Traceroute Utility

To characterize routing behavior we used the `traceroute` utility [25] to measure Internet paths. `traceroute` finds out intermediate hops between two sites by launching UDP probe packets to an unlikely destination port value with an incrementing Time To Live (TTL) starting from 1. Each probe expects an ICMP "time exceeded" reply from an intermediate gateway or an ICMP "port unreachable" reply if the destination host is reached. By default three probes are sent at each TTL setting and a line is printed showing the TTL, address of the gateway and round trip time of each probe. If the probe answers come from different gateways, the address of each responding system will be printed. If there is no response within a 5-second timeout interval, a `*` is printed for that probe. It also indicates possible causes of timeout with `!H` (host unreachable), `!N` (network unreachable), `!P` (protocol unreachable). By default a maximum of 30 hops are elicited by `traceroute` to avoid being trapped in possible routing loops.

Although `traceroute` is a handy tool for route recording, one should bear in mind several limitations and measurement difficulties when using it.

First, it is hard to know whether probes of different hops actually take the same route as pervious hops since routes may change between two successive probe packets. The general assumption is that if a route appears self-consistent and shows no signs of multiple routing for any of its hops, it is treated as a valid measurement.

Second, routers and some hosts have multiple IP addresses. Some of them set the source address of the ICMP reply message as the interface on which this message is sent, rather than the address on which the probe is received. When an intermediate route has this behavior, it is difficult to know which router interface is actually in the path.

Other issues associated with using `traceroute` include the presence of rate limiting routers and firewall configurations. More details can be found in [20, 26].

3.2 Measurement Methodology

Our measurement used 24 public `traceroute` servers located in US, Iceland, Netherlands, Australia, Germany, Switzerland, Bulgaria, Sweden, and Thailand. The data collection script is written in Tcl. At a scheduled time point, our program spawns two threads for a pair of selected `traceroute` servers. Each thread performs a remote login to one server and invokes a `traceroute` to the other server in the pair. The results are returned and stored in a local file server. Since records obtained from different servers carry different formats, they are shaped to a uniform output format, supplemented with AS mapping information (Section 3.3), and fed into MySQL (an open source relational database) for further processing.

The majority of our experiments used independent and exponentially distributed time intervals between consecutive measurements, similar to that in [20, 22]. This gives unbiased, additive random sampling so all instantaneous signal values are sampled with equal probability. The Poisson nature of such measurements ensures that the proportion of times our measurement observes a given state is asymptotically equal to the amount of time that the Internet spends in that state [27]. We use this property to characterize route persistence and prevalence (Sections 4.7 and 4.8). Our exponential interval measurements have an average inter-measurement time of either 2.75 hours or 11.5 hours, limited by the number of participating servers and their allowed `traceroute` frequency.

To illustrate a more accurate picture of short-scale routing dynamics we also conducted experiments with a 10-minute fixed measurement interval on a few designated `traceroute` servers. These measurements also give us some understanding about possible impact of infrastructure restriction (i.e., relatively long measurement interval) on the measurement results.

3.3 IP - AS Mapping

In preparation for AS-level analysis, we need to map IP addresses in all `traceroute` records to their corresponding AS numbers. To do this, we first look up the Routing Assets Database (RADB) and its mirrored databases [28]. If a valid AS mapping entry is not found, we use `Level 3`'s looking glass [29] and AT&T's telnet server [30] which can display AS information via a `traceroute`. If all these fail, we check the records where the specific IP address appears. If this address has a hostname that identifies itself as part of a site with a known AS, then we assign that AS number to it. If no hostname can be derived, we look at the IP address range and neighboring hops in the same record and make a best guess.

3.4 Collected Data Sets

The four data sets (DS I through DS IV) we obtained through our experiments will be frequently referred to in later sections of this document. The selection of `traceroute` sites and measurement rates in each data set are determined by agreements from the respective server administrators and the geographic locations of the servers.

3.4.1 Data Set I

Table 1 lists the 12 sites that participated in DS I. The data collecting time is from 05:14:02 AM EDT, April 9, 2004 to 00:09:34 AM EDT, April 24, 2004. The sampling is Poisson with a rate averaged at one `traceroute` every 15 minutes for each site. This corresponds to an average of 2.75 hours sampling interval for each individual path.

3.4.2 Data Sets II and III

Data sets II and III contain the above 12 sites and additional 12 sites listed in Table 2. Each site is exponentially sampled on average every 30 minutes. This corresponds to an average of 11.5 hours sampling interval for each individual path. DS II is taken from

Site	Location
www.valkaryn.net	Los Angeles, CA
www.slac.stanford.edu	Stanford, CA
lava.net	Honolulu, HI
www.fh-friedberg.de	Frieberg, Germany
www.lf.net	Germany
swice2.switch.ch	Geneva, Switzerland
backbone.acad.bg	Sofia, Bulgaria
stockholm1.sunet.se	Stockholm, Sweden
traceroute.teragen.com.au	Melbourne, Australia
www.hafey.org	Sydney, Australia
www.megamirror.com	Sydney, Australia
www.debug.net	Frankfurt, Germany

Table 1: 12 sites participated in data set I

Site	Location
www.net.berkeley.edu	Berkeley, CA
www.opus1.com	Tucson, AZ
www.telcom.arizona.edu	Tucson, AZ
www.fmp.com	Leander, TX
www.washington.edu	Seattle, WA
www.cynexx.net	Los Angeles, CA
www.io.com	Austin, TX
www.rhnet.is	Iceland
www.cistron.net	Netherlands
www.globedom.com	Germany
proxy1.syd.connect.com.au	Sydney, Australia
corporate.pacific.net.th	Thailand

Table 2: Additional 12 sites participated in data sets II and III

12:20:11 AM EDT, May 22 to 12:24:19 PM EDT, Jun 13, 2004 and DS III is taken from 03:05:23 PM EDT, June 14 to 10:34:14 AM, July 06, 2004.

3.4.3 Data Set IV

In DS IV, we conducted a 10-minute fixed interval and a 2-hour exponential interval `traceroute` measurement simultaneously on four paths: `lava.net` (Honolulu, HI) to `www.globedom.com` (Germany) and its reverse path; `www.lf.net` (Germany) to `traceroute.teragen.com.au` (Melbourne, Australia) and its reverse path. This data set is collected from July 14 to August 3, 2004.

4 Analysis of Routing Dynamics

4.1 General Statistics and Routing Pathologies

Table 3 lists the general statistics of the `traceroute` records for our two main data sets, DS I and DS II. The first two rows show that 99.6% requests in DS I and 99.1% requests in DS II managed to get a result back from the `traceroute` server.

Item description	Data set I	Data set II
Number of requests issued	17226	25690
Number of records received	17156	25456
At least one hop completely missing	2988	4092
Infrastructure failure	32	11
Same IP address in multiple hops	26	614
Same hop with multiple IP addresses	749	2916
Other failures	198	873
Total number of "clean" records	13323	17743

Table 3: General statistics of data sets I and II

The third row shows that 16-17% of the records contain at least one hop that cannot not be identified at all. This situation occurs when there is any router along the path that is unable to respond to `traceroute` probes properly or when the correctly generated probe response is unable to reach the `traceroute` server. This situation does not necessarily reflect a connectivity problem and should more appropriately be seen as measurement limitations. As far as record processing is concerned, these measurements are inherently ambiguous because we could not tell if the route is the same as that observed at other instances. So they will be excluded for the main route change analysis.

Rows 3 through 5 in Table 3 list number of records showing various pathologies and will be further discussed below. There are other unsuccessful records that do not belong to any of these causes. For example, although `traceroute` normally probes a maximum of 30 hops toward the destination before it stops trying, some records simply ended at a hop in the middle of the path before all probes are attempted, without indicating any infrastructure failure. This situation is likely caused by communication problems with the `traceroute` server, destination host or the local files server. All such records went to the "Other failures" category. The remaining records are called "clean" records on which our main route change analysis will be based.

4.2 Infrastructure Availability Estimation

The first group of records we looked at for routing pathologies are records showing infrastructure failures, i.e., records where `traceroute` printed a host, network or protocol unreachable mark. The number of these records gives us a rough estimate of the network availability. As Table 3 shows, the corresponding value is 99.81% in DS I and 99.96% in DS II. Compared with data in 1994 [20] and 1999 [22], our data have a smaller proportion of failed measurement attempts and represent an equal or slightly higher infrastructure availability.

4.3 Same IP Address in Multiple Hops

The second type of records we looked at are those with at least one same IP address appearing in multiple hops of a `traceroue` record. The number of such records are relatively small (0.15% in DS I and 2.4% in DS II) as listed in the fifth row of Table 3. A detailed examination of these records revealed three causes to generate such records.

4.3.1 Capturing the Middle of Route Change

Records in this group captured the middle of a route change. An example is shown in the following three consecutive records between `backbone.acad.bg` and `www.hafey.org`.

```
Location: backbone.acad.bg --to-- www.hafey.org         Status: ok
Time: Sun Apr 18 06:29:59 EDT 2004 --to-- Sun Apr 18 06:30:22 EDT 2004
```

```
  1 62.75.33.246 20 msec 20 msec 20 msec
  2 athens-3-athens-2.grnet.gr (194.177.209.21) [AS 5408] 32 msec 20 msec 20 msec
  3 grnet.gr1.gr.geant.net (62.40.103.57) [AS 20965] 20 msec 20 msec 20 msec
  4 gr.uk1.uk.geant.net (62.40.96.98) [AS 20965] 100 msec 100 msec 104 msec
  5 uk.nl1.nl.geant.net (62.40.96.181) [AS 20965] 100 msec 100 msec 100 msec
  6 nl.de1.de.geant.net (62.40.96.101) [AS 20965] 96 msec 96 msec 100 msec
  7 ffm-b2-pos2-3.telia.net (213.248.77.89) [AS 1299] 208 msec 196 msec 208 msec
  8 ffm-bb2-pos2-3-0.telia.net (213.248.64.177) [AS 1299] 116 msec 116 msec 116 msec
  9 prs-bb2-pos7-0-0.telia.net (213.248.65.117) [AS 1299] 120 msec 116 msec 116 msec
 10 ldn-bb2-pos7-0-0.telia.net (213.248.65.113) [AS 1299] 120 msec 116 msec 116 msec
 11 nyk-bb2-pos6-0-0.telia.net (213.248.65.94) [AS 1299] 192 msec 192 msec 188 msec
 12 nyk-bb1-pos0-0-0.telia.net (213.248.80.133) [AS 1299] 192 msec 188 msec 192 msec
 13 nyk-i2-pos1-0.telia.net (213.248.82.10) [AS 1299] 192 msec 192 msec 188 msec
 14 POS3-1.IG4.NYC4.ALTER.NET (208.192.177.29) [AS 701] 192 msec 192 msec 192 msec
 15 904.at-1-0-0.XR3.NYC4.ALTER.NET (152.63.19.238) [AS 701] 188 msec 188 msec 188 msec
 16 0.so-2-0-0.XL1.NYC4.ALTER.NET (152.63.17.29) [AS 701] 192 msec 192 msec 192 msec
 17 0.so-2-0-0.TL1.NYC8.ALTER.NET (152.63.0.153) [AS 701] 192 msec 188 msec 192 msec
 18 0.so-0-1-0.TL1.LAX9.ALTER.NET (152.63.9.230) [AS 701] 272 msec 272 msec 272 msec
 19 0.POS6-0.IR1.LAX9.ALTER.NET (152.63.0.145) [AS 701] 272 msec 268 msec 272 msec
 20 so-7-0-0.IR1.LAX12.ALTER.NET (137.39.31.221) [AS 701] 272 msec 272 msec 272 msec
 21 so-3-0-0.TR1.SYD2.Alter.Net (210.80.48.133) [AS 703] 416 msec 420 msec 416 msec
 22 so-3-2-0.XR1.SYD2.Alter.Net (210.80.48.134) [AS 703] 424 msec 424 msec 424 msec
 23 411.AT-6-0-0.GW5.SYD2.ALTER.NET (210.80.32.250) [AS 703] 428 msec 424 msec 428 msec
 24 sisgroup-gw.aspac.customer.alter.net (203.166.42.254) [AS 703] 416 msec 420 msec 420 msec
 25 203.98.85.253 [AS 10145] 440 msec 436 msec 436 msec
 26 mail.hafey.org (203.98.94.7) [AS 10145] 424 msec 416 msec 416 msec


Location: backbone.acad.bg --to-- www.hafey.org         Status: ok
Time: Sun Apr 18 11:46:24 EDT 2004 --to-- Sun Apr 18 11:46:51 EDT 2004

  1 62.75.33.246 [AS 6453] 16 msec 20 msec 20 msec
  2 athens-3-athens-2.grnet.gr (194.177.209.21) [AS 5408] 20 msec 20 msec 16 msec
  3 grnet.gr1.gr.geant.net (62.40.103.57) [AS 20965] 16 msec 20 msec 20 msec
  4 gr.de2.de.geant.net (62.40.96.94) [AS 20965] 112 msec 100 msec 100 msec
  5 de2-2.de1.de.geant.net (62.40.96.54) [AS 20965] 100 msec 100 msec 96 msec
  6 ffm-b2-pos2-3.telia.net (213.248.77.89) [AS 1299] 104 msec 100 msec 100 msec
  7 ffm-bb2-pos2-3-0.telia.net (213.248.64.177) [AS 1299] 100 msec 100 msec 100 msec
  8 prs-bb2-pos7-0-0.telia.net (213.248.65.117) [AS 1299] 104 msec 100 msec 100 msec
  9 ldn-bb2-pos7-0-0.telia.net (213.248.65.113) [AS 1299] 100 msec 100 msec 100 msec
 10 nyk-bb2-pos6-0-0.telia.net (213.248.65.94) [AS 1299] 176 msec 176 msec 176 msec
 11 nyk-bb1-pos0-0-0.telia.net (213.248.80.133) [AS 1299] 172 msec 184 msec 176 msec
 12 nyk-i2-pos1-0.telia.net (213.248.82.10) [AS 1299] 176 msec 172 msec 176 msec
 13 POS3-1.IG4.NYC4.ALTER.NET (208.192.177.29) [AS 701] 192 msec 192 msec 192 msec
 14 904.at-1-0-0.XR3.NYC4.ALTER.NET (152.63.19.238) [AS 701] 192 msec 192 msec 192 msec
 15 0.so-2-0-0.XL1.NYC4.ALTER.NET (152.63.17.29) [AS 701] 196 msec 192 msec 196 msec
 16 0.so-2-0-0.TL1.NYC8.ALTER.NET (152.63.0.153) [AS 701] 192 msec 192 msec 192 msec
 17 0.so-0-1-0.TL1.LAX9.ALTER.NET (152.63.9.230) [AS 701] 268 msec 272 msec 272 msec
 18 0.POS6-0.IR1.LAX9.ALTER.NET (152.63.0.145) [AS 701] 272 msec 272 msec 272 msec
 19 so-7-0-0.IR1.LAX12.ALTER.NET (137.39.31.221) [AS 701] 272 msec 272 msec 272 msec
 20 so-3-0-0.TR1.SYD2.Alter.Net (210.80.48.133) [AS 703] 484 msec
    so-7-0-0.IR1.LAX12.ALTER.NET (137.39.31.221) [AS 701] 408 msec 404 msec
 21 so-3-0-0.TR1.SYD2.Alter.Net (210.80.48.133) [AS 703] 428 msec 536 msec 452 msec
 22 so-3-2-0.XR1.SYD2.Alter.Net (210.80.48.134) [AS 703] 468 msec 424 msec 428 msec
 23 411.AT-6-0-0.GW5.SYD2.ALTER.NET (210.80.32.250) [AS 703] 428 msec 428 msec 424 msec
 24 sisgroup-gw.aspac.customer.alter.net (203.166.42.254) [AS 703] 420 msec 420 msec 420 msec
 25 203.98.85.253 [AS 10145] 436 msec 436 msec 436 msec
 26 mail.hafey.org (203.98.94.7) [AS 10145] 420 msec 420 msec 420 msec


Location: backbone.acad.bg --to-- www.hafey.org         Status: ok
Time: Sun Apr 18 11:47:26 EDT 2004 --to-- Sun Apr 18 11:47:48 EDT 2004

  1 62.75.33.246 [AS 8717] 20 msec 20 msec 20 msec
  2 athens-3-athens-2.grnet.gr (194.177.209.21) [AS 5408] 16 msec 24 msec 20 msec
  3 grnet.gr1.gr.geant.net (62.40.103.57) [AS 20965] 20 msec 20 msec 20 msec
  4 gr.uk1.uk.geant.net (62.40.96.98) [AS 20965] 120 msec 84 msec 88 msec
  5 uk.nl1.nl.geant.net (62.40.96.181) [AS 20965] 96 msec 100 msec 104 msec
  6 nl.de1.de.geant.net (62.40.96.101) [AS 20965] 100 msec 96 msec 100 msec
  7 ffm-b2-pos2-3.telia.net (213.248.77.89) [AS 1299] 100 msec 104 msec 100 msec
  8 ffm-bb2-pos2-3-0.telia.net (213.248.64.177) [AS 1299] 100 msec 100 msec 100 msec
```

```
 9 prs-bb2-pos7-0-0.telia.net (213.248.65.117) [AS 1299] 100 msec 100 msec 100 msec
10 ldn-bb2-pos7-0-0.telia.net (213.248.65.113) [AS 1299] 100 msec 100 msec 100 msec
11 nyk-bb2-pos6-0-0.telia.net (213.248.65.94) [AS 1299] 172 msec 188 msec 172 msec
12 nyk-bb1-pos0-0-0.telia.net (213.248.80.133) [AS 1299] 176 msec 180 msec 188 msec
13 nyk-i2-pos1-0.telia.net (213.248.82.10) [AS 1299] 172 msec 172 msec 172 msec
14 POS3-1.IG4.NYC4.ALTER.NET (208.192.177.29) [AS 701] 192 msec 188 msec 192 msec
15 904.at-1-0-0.XR3.NYC4.ALTER.NET (152.63.19.238) [AS 701] 192 msec 192 msec 192 msec
16 0.so-2-0-0.XL1.NYC4.ALTER.NET (152.63.17.29) [AS 701] 188 msec 192 msec 192 msec
17 0.so-2-0-0.TL1.NYC8.ALTER.NET (152.63.0.153) [AS 701] 192 msec 192 msec 188 msec
18 0.so-0-1-0.TL1.LAX9.ALTER.NET (152.63.9.230) [AS 701] 268 msec 272 msec 272 msec
19 0.POS6-0.IR1.LAX9.ALTER.NET (152.63.0.145) [AS 701] 268 msec 276 msec 268 msec
20 so-7-0-0.IR1.LAX12.ALTER.NET (137.39.31.221) [AS 701] 272 msec 272 msec 268 msec
21 so-3-0-0.TR1.SYD2.Alter.Net (210.80.48.133) [AS 703] 420 msec 416 msec 420 msec
22 so-3-2-0.XR1.SYD2.Alter.Net (210.80.48.134) [AS 703] 424 msec 428 msec 424 msec
23 411.AT-6-0-0.GW5.SYD2.ALTER.NET (210.80.32.250) [AS 703] 444 msec 428 msec 424 msec
24 sisgroup-gw.aspac.customer.alter.net (203.166.42.254) [AS 703] 420 msec 484 msec 432 msec
25 203.98.85.253 [AS 10145] 440 msec 436 msec 636 msec
26 mail.hafey.org (203.98.94.7) [AS 10145] 544 msec 444 msec 416 msec
```

The first record was taken at 06:29:59 EDT, Apr 18, 2004 and showed that the original path went through UK, NL, DE as follows:

```
 3 grnet.gr1.gr.geant.net (62.40.103.57) [AS 20965] 20 msec 20 msec 20 msec
 4 gr.uk1.uk.geant.net (62.40.96.98) [AS 20965] 100 msec 100 msec 104 msec
 5 uk.nl1.nl.geant.net (62.40.96.181) [AS 20965] 100 msec 100 msec 100 msec
 6 nl.de1.de.geant.net (62.40.96.101) [AS 20965] 96 msec 96 msec 100 msec
 7 ffm-b2-pos2-3.telia.net (213.248.77.89) [AS 1299] 208 msec 196 msec 208 msec
```

In the second record taken at 11:46:24 EDT on the same day, the path skipped the UK and NL routers and appeared to go through two DE routers directly as follows:

```
 3 grnet.gr1.gr.geant.net (62.40.103.57) [AS 20965] 16 msec 20 msec 20 msec
 4 gr.de2.de.geant.net (62.40.96.94) [AS 20965] 112 msec 100 msec 100 msec
 5 de2-2.de1.de.geant.net (62.40.96.54) [AS 20965] 100 msec 100 msec 96 msec
 6 ffm-b2-pos2-3.telia.net (213.248.77.89) [AS 1299] 104 msec 100 msec 100 msec
```

After another 62 seconds, the third record, taken at 11:47:26 EDT, showed that the path had come back to the original one as of 06:29:59 EDT. During the short interval of route change, we observed the IP addresses 137.39.31.221 and 210.80.48.133 appeared in more than one hop in the second record as shown below, possibly indicating there are still packets going in both paths.

```
19 0.POS6-0.IR1.LAX9.ALTER.NET (152.63.0.145) [AS 701] 272 msec 268 msec 272 msec
20 so-7-0-0.IR1.LAX12.ALTER.NET (137.39.31.221) [AS 701] 272 msec 272 msec 272 msec
21 so-3-0-0.TR1.SYD2.Alter.Net (210.80.48.133) [AS 703] 416 msec 420 msec 416 msec
22 so-3-2-0.XR1.SYD2.Alter.Net (210.80.48.134) [AS 703] 424 msec 424 msec 424 msec
23 411.AT-6-0-0.GW5.SYD2.ALTER.NET (210.80.32.250) [AS 703] 428 msec 424 msec 428 msec

18 0.POS6-0.IR1.LAX9.ALTER.NET (152.63.0.145) [AS 701] 272 msec 272 msec 272 msec
19 so-7-0-0.IR1.LAX12.ALTER.NET (137.39.31.221) [AS 701] 272 msec 272 msec 272 msec
20 so-3-0-0.TR1.SYD2.Alter.Net (210.80.48.133) [AS 703] 484 msec
   so-7-0-0.IR1.LAX12.ALTER.NET (137.39.31.221) [AS 701] 408 msec 404 msec
21 so-3-0-0.TR1.SYD2.Alter.Net (210.80.48.133) [AS 703] 428 msec 536 msec 452 msec
22 so-3-2-0.XR1.SYD2.Alter.Net (210.80.48.134) [AS 703] 468 msec 424 msec 428 msec
23 411.AT-6-0-0.GW5.SYD2.ALTER.NET (210.80.32.250) [AS 703] 428 msec 428 msec 424 msec
```

4.3.2 Capturing Temporary Loss of Connectivity

Records in this group captured a temporary loss of connectivity somewhere in the network. When the connectivity returns, the same router appeared again. These cases usually indicate infrastructure failures. An example is shown below by three consecutive `traceroute` records between `lava.net` and `backbone.acad.bg`.

```
Location: lava.net --to-- backbone.acad.bg   Status: ok
```

```
Time: Mon Apr 12 15:28:18 EDT 2004 --to-- Mon Apr 12 15:29:40 EDT 2004

 1  iiwi-fe-0-0-0.lava.net (64.65.64.30)  0.615 ms  0.514 ms  0.498 ms
 2  s2-6.hsa1.hon2.bbnplanet.net (4.24.134.17)  0.585 ms  0.428 ms  2.469 ms
 3  so-0-1-0.mpls1.Honolulu2.Level3.net (4.68.112.89)  1.919 ms  0.993 ms  0.675 ms
 4  so-0-2-0.bbr1.SanJose1.Level3.net (209.247.9.114)  59.613 ms  61.544 ms  60.102 ms
 5  so-6-0-0.edge1.SanJose1.Level3.net (209.244.3.138)  59.997 ms  59.928 ms  59.881 ms
 6  globalcrossing-level3-oc48.SanJose1.Level3.net (4.68.127.54)  59.929 ms  60.003 ms  60.644 ms
 7  pos1-0-2488M.cr2.FRA2.gblx.net (67.17.65.65)  215.692 ms  216.287 ms  216.119 ms
 8  so5-0-0-2488M.ar2.FRA2.gblx.net (67.17.65.54)  215.521 ms  215.391 ms  215.342 ms
 9  geant-de1-de.so-7-0-0.ar2.FRA2.gblx.net (208.48.23.146)  217.927 ms  218.589 ms  227.638 ms
10  de1-1.de2.de.geant.net (62.40.96.130)  217.904 ms  217.946 ms  218.055 ms
11  de2.gr1.gr.geant.net (62.40.96.93)  297.503 ms  297.298 ms  297.666 ms
12  grnet-gw.gr1.gr.geant.net (62.40.103.58)  297.299 ms  297.268 ms  299.221 ms
13  athens-2-athens-3.grnet.gr (194.177.209.22)  299.368 ms  298.070 ms  298.830 ms
14  * * *
15  * * *
16  * * *
17  backbone-oteglobe.lines.acad.bg (194.141.252.1) 323.397 ms *  317.218 ms


Location: lava.net --to-- backbone.acad.bg   Status: timeout
Time: Mon Apr 12 17:18:26 EDT 2004 --to-- Mon Apr 12 17:28:27 EDT 2004

 1  iiwi-fe-0-0-0.lava.net (64.65.64.30)  0.519 ms  0.465 ms  0.424 ms
 2  * * *
 3  * * *
 4  * * *
 5  * * *
 6  * * *
 7  * * *
 8  * * *
 9  * * *
10  * * *
11  * * *
12  * s2-6.hsa1.hon2.bbnplanet.net (4.24.134.17)  1.114 ms !H *
13  * * *
14  * * *
15  * s2-6.hsa1.hon2.bbnplanet.net (4.24.134.17)  0.415 ms !H *
16  * * *
17  * * *
18  * * *
19  * * *
20  * * *
21  * * *
22  s2-6.hsa1.hon2.bbnplanet.net (4.24.134.17)  0.561 ms !H * *


 #CSTL#   Location: lava.net --to-- backbone.acad.bg   Status: timeout
 #CSTT#   Time: Mon Apr 12 18:34:24 EDT 2004 --to-- Mon Apr 12 18:44:24 EDT 2004
 1  iiwi-fe-0-0-0.lava.net (64.65.64.30)  0.407 ms  0.434 ms  0.414 ms
 2  * * *
 3  * * *
 4  * * *
 5  * * *
 6  * s2-6.hsa1.hon2.bbnplanet.net (4.24.134.17)  1.012 ms !H *
 7  * * *
 8  * * s2-6.hsa1.hon2.bbnplanet.net (4.24.134.17)  0.703 ms !H
 9  * * *
10  * * *
11  * * *
12  * * *
13  * * *
14  * * *
15  * * *
16  * * *
17  * * *
18  * * *
19  * * *
20  * * *
21  * * *
```

```
22  * * *
23  s2-6.hsa1.hon2.bbnplanet.net (4.24.134.17)  0.793 ms !H * *
```

The second and third records above, taken at 17:18:26 and 18:34:24 EDT, Apr 12, 2004, displayed the host `s2-6.hsa1.hon2.bbnplanet.net` (4.24.134.17) multiple times. Comparing to the measurement immediately before these two at 15:28:18 EDT on the same day, we see connectivity problems associated with the host `s2-6.hsa1.hon2.bbnplanet.net`. It is sometimes unreachable (indicated by `*`), sometimes reporting the destination `backbone-oteglobe.lines.acad.bg` unreachable, and sometimes fine (with intervals up to 210 seconds between its successful responses to `traceroute` probes).

4.3.3 Skipping Behavior

A few records originated from `swice2.switch.ch` (Swiss Education and Research Network) displayed a strange skipping behavior. An example is shown in the following three consecutive records.

```
Location: swiCE2.switch.ch --to-- lava.net   Status: ok
Time: Thu Apr 15 03:27:53 EDT 2004 --to-- Thu Apr 15 03:28:06 EDT 2004

  1 swiCE3-G4-3.switch.ch (130.59.36.6) 0 msec 4 msec 0 msec
  2 so-0-1-0.ar2.CDG2.gblx.net (64.212.70.61)  12 msec 8 msec 12 msec
  3 pos5-0-2488M.cr1.CDG2.gblx.net (67.17.65.113)  8 msec 12 msec 8 msec
  4 so0-0-0-2488M.cr2.LON3.gblx.net (67.17.64.38)  20 msec 16 msec 16 msec
  5 so7-0-0-2488M.ar2.LON3.gblx.net (67.17.66.30)  20 msec 16 msec 20 msec
  6 Level-3public-peering.ge-5-0-0.ar2.LON3.gblx.net (208.51.239.162) 16 msec 16 msec 20 msec
  7 ae-0-53.mp1.London1.Level3.net (212.187.131.129)  16 msec 20 msec 20 msec
  8 so-1-0-0.mp1.London2.Level3.net (212.187.128.49)  16 msec 20 msec 16 msec
  9  * so-1-0-0.bbr1.Washington1.Level3.net (212.187.128.138)  88 msec 92 msec
 10 ge-0-0-0.mpls1.Honolulu2.Level3.net (4.68.128.13)  212 msec 212 msec 212 msec
 11 so-7-0.hsa1.Honolulu2.Level3.net (4.68.112.90)  212 msec 212 msec 212 msec
 12 s1.lavanet.bbnplanet.net (4.24.134.18)  212 msec 212 msec 212 msec
 13 lava.net (64.65.64.17)  216 msec 212 msec 212 msec


Location: swiCE2.switch.ch --to-- lava.net   Status: ok
Time: Thu Apr 15 05:57:25 EDT 2004 --to-- Thu Apr 15 05:57:36 EDT 2004

  1 so-0-1-0.ar2.CDG2.gblx.net (64.212.70.61)  16 msec
    swiCE3-G4-3.switch.ch (130.59.36.6) 0 msec 0 msec
  2 so-0-1-0.ar2.CDG2.gblx.net (64.212.70.61)  16 msec 8 msec 12 msec
  3 pos5-0-2488M.cr1.CDG2.gblx.net (67.17.65.113)  8 msec 16 msec 8 msec
  4 so0-0-0-2488M.cr2.LON3.gblx.net (67.17.64.38)  24 msec 16 msec 20 msec
  5 so7-0-0-2488M.ar2.LON3.gblx.net (67.17.66.30)  16 msec 16 msec 16 msec
  6 Level-3public-peering.ge-5-0-0.ar2.LON3.gblx.net (208.51.239.162)  16 msec 16 msec 16 msec
  7 ae-0-53.mp1.London1.Level3.net (212.187.131.129)  20 msec 32 msec 20 msec
  8 so-1-0-0.mp1.London2.Level3.net (212.187.128.49)  16 msec 20 msec 20 msec
  9 so-1-0-0.bbr1.Washington1.Level3.net (212.187.128.138)  88 msec 92 msec 92 msec
 10 ge-0-0-0.mpls1.Honolulu2.Level3.net (4.68.128.13)  212 msec 212 msec 212 msec
 11 so-7-0.hsa1.Honolulu2.Level3.net (4.68.112.90)  212 msec 212 msec 212 msec
 12 s1.lavanet.bbnplanet.net (4.24.134.18)  212 msec 212 msec 212 msec
 13 lava.net (64.65.64.17)  216 msec 212 msec 216 msec


Location: swiCE2.switch.ch --to-- lava.net   Status: ok
Time: Thu Apr 15 06:59:15 EDT 2004 --to-- Thu Apr 15 06:59:27 EDT 2004

  1 swiCE3-G4-3.switch.ch (130.59.36.6) 4 msec 0 msec 0 msec
  2 so-0-1-0.ar2.CDG2.gblx.net (64.212.70.61) 12 msec 8 msec 12 msec
  3 pos5-0-2488M.cr1.CDG2.gblx.net (67.17.65.113)  8 msec 12 msec 8 msec
  4 so0-0-0-2488M.cr2.LON3.gblx.net (67.17.64.38)  16 msec 16 msec 16 msec
  5 so7-0-0-2488M.ar2.LON3.gblx.net (67.17.66.30)  16 msec 16 msec 20 msec
  6 Level-3public-peering.ge-5-0-0.ar2.LON3.gblx.net (208.51.239.162)  16 msec 16 msec 20 msec
  7 ae-0-53.mp1.London1.Level3.net (212.187.131.129)  16 msec 20 msec 16 msec
  8 so-1-0-0.mp1.London2.Level3.net (212.187.128.49)  20 msec 20 msec 16 msec
  9 so-1-0-0.bbr1.Washington1.Level3.net (212.187.128.138)  92 msec 96 msec 92 msec
 10 ge-0-0-0.mpls1.Honolulu2.Level3.net (4.68.128.13)  212 msec 216 msec 212 msec
```

```
11 so-7-0.hsa1.Honolulu2.Level3.net (4.68.112.90)  212 msec 212 msec 212 msec
12 s1.lavanet.bbnplanet.net (4.24.134.18)  212 msec 212 msec 212 msec
13 lava.net (64.65.64.17)  212 msec 212 msec 216 msec
```

From the first record at 03:27:53 EDT on April 15, 2004, we see `swiCE3-G4-3.switch.ch` (130.59.36.6) the first router and `so-0-1-0.ar2.CDG2.gblx.net` (64.212.70.61) the second on the path. However, the following record at 05:57:25 EDT on the same day shows that both routers appear as the first hop, making router 130.59.36.6 occur twice in that record. The third record about one hour later at 06:59:15 EDT returned to the normal path. According to the route administrator, it is not clear whether this is a measurement problem or issue with the router. It could be that the first-hop router `swiCE3` erroneously forwards the first of the three TTL=1 packets from `swiCE2` on to the next router `ar2.CDG2.gblx.net`, or `swiCE2` sends the first packet with TTL=2 instead of TTL=1. But the latter is less plausible.

Overall, analysis of records with at least one same IP address appearing in multiple hops of the record confirmed the two most common routing dynamics that we must deal with, namely route changes and route failure.

4.4 Same Hop with Multiple IP Addresses - Route Fluttering

Another group of records we looked at are those with a single hop containing multiple different IP addresses. In other words, the three consecutive `traceroute` probes for the same hop in a given path return different IP addresses. There are 4.4% and 11.5% of such records in DS I and DS II respectively, which do not seem to be negligible. Using DS I as the example, we defined an entry of one hop with multiple different IP addresses as one "instance" and identified totally 1194 such instances for all the 749 such records (one record may contain more than one instance). Among them, there are 20 distinct IP address pairs plus one distinct IP address triple (i.e., the same hop returns three different IP addresses). After a closer look at these IP addresses, we found that 7 of the 20 distinct IP address pairs (accounting for 7 instances) as well as the only triple instance, captured the middle of route changes; 2 of the 20 distinct IP address pairs (accounting for 2 instances) captured temporary network outages; 3 of the 20 distinct IP address pairs (accounting for 19 instances) resulted from the skipping behavior of site `swice2.switch.ch` as mentioned above. The remaining 8 of the 20 IP address pairs did not show any of the above causes, but they occurred repeatedly and shared all remaining instances. Table 4 lists details about these 8 router pairs.

The rapidly variable routing exhibited by these routers are called "fluttering" [20]. These are route changes occurring on the time scale of seconds to minutes between successive `traceroute` probes. One mechanism that could cause this behavior is when a router alternates between multiple next-hop routers in order to split load among the links to those routers. As pointed out in [20], such behavior is explicitly allowed in RFC 1812 "Requirements for IP Version 4 Routers" [31] as load splitting. The same RFC also cautions that there are situations for which this behavior is inappropriate. We infer that routers before these router pairs could have this option turned on. Also from the host names shown in the table, we infer that all the fluttering IP address pairs belong to the same domain and are physically close routers, at least in the same city. They could even be two interfaces of the same router.

In short, route fluttering tends to be local in scale and have less impact on NSIS. However, in cases where that is not true, it causes high rate of route changes and is hard to detect and deal with. Although the behavior was mentioned in the 1994 measurement, route fluttering is still commonly seen in today's Internet.

4.5 Analysis on Route Change

The following analysis is based on all "clean" records, i.e., we excluded records that meet at least one of the following conditions: have at least one hop completely missing; have at least one same IP address in multiple hops; have at least one hop with multiple IP

Domain	IP address	Hostname	Occurrence
Sprint	144.223.27.146 144.232.230.30	sl-telia1-1-0.sprintlink.net sl-telia1-4-0.sprintlink.net	83
Broadwing	216.140.0.66 216.140.0.70	s3-0-0.a1.hywr.broadwing.net s4-0-0.a1.hywr.broadwing.net	10
Global Crossing	67.17.65.53 67.17.65.77	pos5-0-2488M.cr2.FRA2.gblx.net pos6-0-2488M.cr1.FRA2.gblx.net	353
Global Crossing	67.17.65.54 67.17.65.78	so5-0-0-2488M.ar2.FRA2.gblx.net so4-0-0-2488M.ar2.FRA2.gblx.net	111
Global Crossing	67.17.65.57 67.17.65.81	pos11-0-2488M.cr2.FRA2.gblx.net pos11-0-2488M.cr1.FRA2.gblx.net	111
Global Crossing	67.17.65.58 67.17.65.82	so2-0-0-2488M.ar2.FRA3.gblx.net so3-0-0-2488M.ar2.FRA3.gblx.net	352
Global Crossing	67.17.64.66 67.17.74.157	pos6-0-2488M.cr2.SFO1.gblx.net pos8-0-2488M.cr1.SFO1.gblx.net	4
Verio	129.250.2.183 129.250.5.136	p16-3-0-0.r01.snjsca04.us.bb.verio.net p16-7-0-0.r00.snjsca04.us.bb.verio.net	141

Table 4: Rapidly fluttering router pairs in DS I

Item description	DS I	DS II
Total number of route changes	359	3775
TTL invisible route changes	222	2837
Proportion of TTL invisible RCs	62%	75%

Table 5: Route Change (RC) statistics in DS I and DS II

addresses; have infrastructure failure marks or other failures. This treatment is similar to that of [20]. As a result, 78% and 70% of the total records in DS I and DS II are identified as "clean" ones.

4.5.1 Number of Route Changes

Table 5 shows the number of route changes and how many of them do not affect total number of hops, i.e., are TTL-invisible, for data sets I and II. The numbers show that the majority of route changes are TTL-invisible. But TTL-invisible route changes are more likely to be trivial (local) changes that involve only a single or just a few hops. To confirm this, we took a closer look at the most frequently changed paths. Table 6 lists the paths that experienced more than 40 route changes in DS II.

For the most frequently changed path from `www.washington.edu` to `stockholm1.sunet.se`, we found that only 3 out of all 50 changes happened in the middle of the path. All remaining changes are caused by server balancing within the University of Washington network which affects the very first few hops only. Similar situation was also found for a number of other host sites, including `io.com`, `opus1.com`, `berkeley.edu`, and `connect.com.au`. It can be seen that all most frequently changed paths in the Table 6 contain at least one of these sites.

The route changes we are most interested are those non-trivial ones that take place in the middle of a path, possibly affecting routers at different geographic locations. So we perform a "router merging" to ignore those very frequent but only local route changes in order to focus on non-trivial route changes.

Occurrence	Source	Destination
50	www.washington.edu	stockholm1.sunet.se
50	www.washington.edu	www.net.berkeley.edu
47	www.washington.edu	proxy1.syd.connect.com.au
47	www.washington.edu	www.cistron.net
46	www.washington.edu	lava.net
44	www.washington.edu	www.debug.net
43	www.washington.edu	www.fh-friedberg.de
43	www.net.berkeley.edu	www.washington.edu
43	www.washington.edu	www.io.com
42	www.washington.edu	swice2.switch.ch
42	stockholm1.sunet.se	www.washington.edu
42	www.washington.edu	www.opus1.com
42	proxy1.syd.connect.com.au	www.washington.edu
42	lava.net	www.washington.edu
42	www.megamirror.com	www.io.com
41	www.io.com	www.washington.edu
41	www.rhnet.is	www.io.com

Table 6: Paths with more than 40 route change occurrence in DS II

4.5.2 Router Merging for TTL-Invisible Route Changes

The candidate routers to be merged are obtained through a list of router pairs that are responsible for all TTL-invisible route changes and appeared in the same positions in adjacent `tracroute` records. Table 7 shows all these router pairs in DS II that occur more than 40 times.

The number in the first column of Table 7 indicates the exact number of TTL-invisible route changes this pair caused. The other columns provide the IP addresses and hostnames of the router pairs.

In summary we merged routers that are in either of the following groups.

1. load balancing routers in host networks where the `traceroute` source or destination site is located. They include the routers in sites `washington.edu`, `io.com`, `opus1.com`, `berkeley.edu`, `connect.com.au` as identified above.
2. among all pairs of routers that were responsible for TTL-invisible route changes and appeared in the same positions in adjacent `tracroute` records, those pairs whose hostnames indicate that they belong to the same company or institute, and they either differ only in the last byte of their IP addresses, or occurred more than a threshold number of times in the data set being processed, which is 20 for DS I and 40 for DS II.

The above two conditions are very much overlap. Load balancing routers in host networks covered by the first rule almost always appear on top of the router pair list obtained from the second rule. Router pairs that are most likely responsible for TTL-invisible route changes usually belong to the same provider and have similar hostnames as well as IP addresses, which means the two sub-conditions in the second rule are both satisfied most of the time. But in some cases, it could also happen that a router pair satisfying the second condition has similar hostnames but their IP addresses differ by three bytes, as we will see later in an example in section 4.9.3.

The number of route changes after router merging are listed in Table 8. It shows that more than half of TTL-invisible changes were eliminated in both data sets.

Note that the merging process, including choosing the merging threshold, involves approximation and is unlikely to be absolutely accurate. Even after merging, we still

Occurrence	IP address A	IP address B	Hostname A	Hostname B
200	206.224.87.12	206.224.87.13	www.io.com	www-03.io.com
181	206.224.87.11	206.224.87.12	www-01.io.com	www.io.com
174	206.224.87.11	206.224.87.13	www-01.io.com	www-03.io.com
114	140.142.15.233	140.142.3.7	www4.cac.washington.edu	www.washington.edu
114	140.142.15.163	140.142.3.7	www.washington.edu	www.washington.edu
112	198.107.150.12	198.107.151.12	hnsp2-wes-ge-1-0-0-0.pnw-gigapop.net	hnsp2-wes-ge-1-0-1-0.pnw-gigapop.net
108	140.142.3.35	140.142.3.7	www.washington.edu	www.washington.edu
105	140.142.15.233	140.142.3.35	www4.cac.washington.edu	www.washington.edu
90	140.142.3.1	140.142.3.33	zinc-FE-4-7.cac.washington.edu	zinc-FE-3-11.cac.washington.edu
88	140.142.15.163	140.142.3.35	www.washington.edu	www.washington.edu
88	140.142.15.163	140.142.15.233	www.washington.edu	www4.cac.washington.edu
84	140.142.15.225	140.142.3.33	kirby-FE4-13.cac.washington.edu	zinc-FE-3-11.cac.washington.edu
79	140.142.15.161	140.142.3.33	astrovac-FE3-8.cac.washington.edu	zinc-FE-3-11.cac.washington.edu
73	140.142.15.161	140.142.3.1	astrovac-FE3-8.cac.washington.edu	zinc-FE-4-7.cac.washington.edu
67	140.142.15.161	140.142.15.225	astrovac-FE3-8.cac.washington.edu	kirby-FE4-13.cac.washington.edu
61	193.10.68.37	193.10.68.93	se-ov.nordu.net	se-ov.nordu.net
58	140.142.15.225	140.142.3.1	kirby-FE4-13.cac.washington.edu	zinc-FE-4-7.cac.washington.edu
56	193.10.68.38	193.10.68.94	dk-gw2.nordu.net	dk-gw2.nordu.net
51	198.107.150.5	198.107.151.5	cnsp1-wes-ge0-0-0.pnw-gigapop.net	cnsp1-wes-GE0-1-0.pnw-gigapop.net
46	140.142.150.24	140.142.153.23	uwbr2-GE1-2.cac.washington.edu	uwbr1-GE2-0.cac.washington.edu
46	64.245.55.36	64.245.55.4	64.245.55.36	64.245.55.4
45	144.232.1.106	144.232.1.42	sl-gw22-ana-10-0.sprintlink.net	sl-gw22-ana-9-0.sprintlink.net
45	198.107.150.30	198.107.151.30	prs1-wes-ge-0-0-0-0.pnw-gigapop.net	prs1-wes-ge-1-0-0-0.pnw-gigapop.net
43	140.142.150.24	140.142.155.23	uwbr2-GE1-2.cac.washington.edu	uwbr1-GE3-1.cac.washington.edu
42	140.142.153.23	140.142.155.23	uwbr1-GE2-0.cac.washington.edu	uwbr1-GE3-1.cac.washington.edu
41	140.142.150.23	140.142.150.24	uwbr1-GE0-1.cac.washington.edu	uwbr2-GE1-2.cac.washington.edu

Table 7: Router pairs responsible for more than 40 TTL-invisible route changes in DS II

TTL-invisible Route Changes	DS I	DS II
Number before merging	222	2837
Number after merging	107	862
Percentage reduction	52%	70%

Table 8: Route changes after router merging for DS I and DS II

Item description	DS I	DS II
AS path changes	66	296
AS path change over route change (before merging)	18%	7.8%
AS path change over route change (after merging)	27%	16%
TTL-invisible AS path change over AS path change	23%	17%
TTL-invisible AS path change over TTL-invisible route change (before merging)	6.8%	2.1%
TTL-invisible AS change over TTL-invisible route change (after merging)	14%	5.9%

Table 9: Number of AS path changes

see some router pairs with like hostnames causing most of the remaining TTL-invisible route changes, indicating more rooms for merging. On the other hand, there could also be a few merged router pairs that were actually not tightly coupled and should be treated separately. Solving these problems requires exact information of the specific router pair which is difficult to obtain.

4.5.3 Hop Count Difference for TTL-Visible Route Changes

Route changes that result in a different number of total hops are TTL-visible. Figure 1 shows the histogram of proportion of such changes vs. the magnitude of hop count differences. A negative number of hop count difference indicates that the number of hops becomes smaller after the route change. We found in both data sets, about 54% of such route changes involve only one hop count difference; 77% in DS I and 89% in DS II of such route changes involve fewer than 3 hops in hop count difference.

4.6 AS Path Change Analysis

4.6.1 Number of AS Path Changes

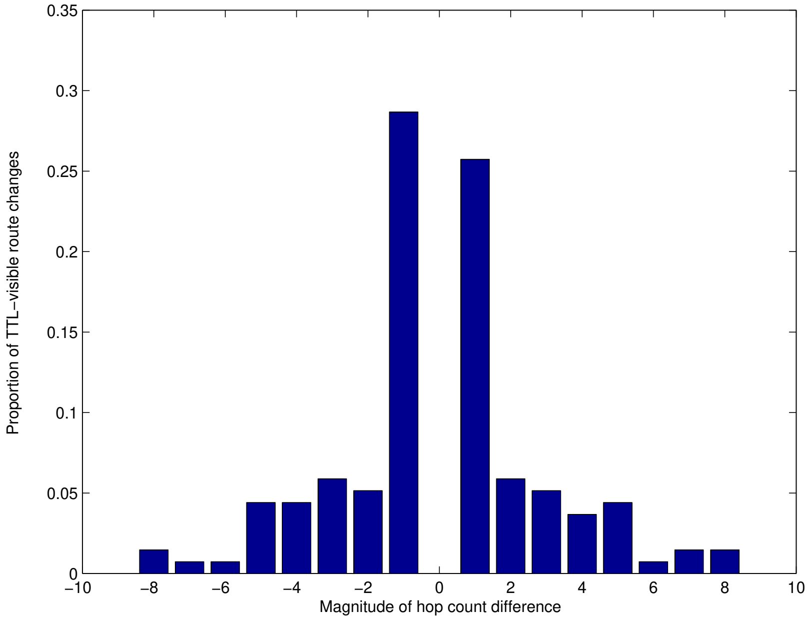
Some major route changes involve changes of AS paths. Table 9 shows the number of AS path changes in data sets I and II.

Not surprisingly, AS path changes only form a small portion of all route changes, from our data sets the number is at most about a quarter even after router merging. The majority of AS path changes appear to lead to changes in hop counts and are therefore TTL-visible.

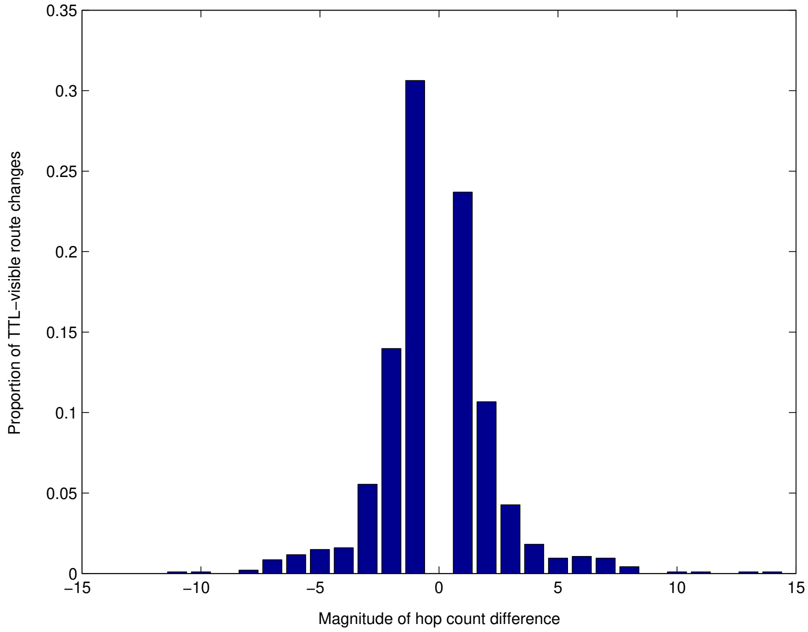
4.6.2 AS Count Difference Distribution

Figure 2 shows the proportion of AS path changes vs. the magnitude of AS count difference. It can be seen that over half AS path changes incur no difference in AS count. Almost all AS path changes involve no more than two ASes.

We also found in Figure 2 a rough symmetry centered at AS count difference of zero, a similar situation can be found in Table 1 for hop count difference. Both of these may partially be explained by the fact that most Internet paths are governed by a dominant route discussed in section 4.7. So the effect of route changes that change the hop counts in the path tend to be canceled by another route change returning to the dominant route.

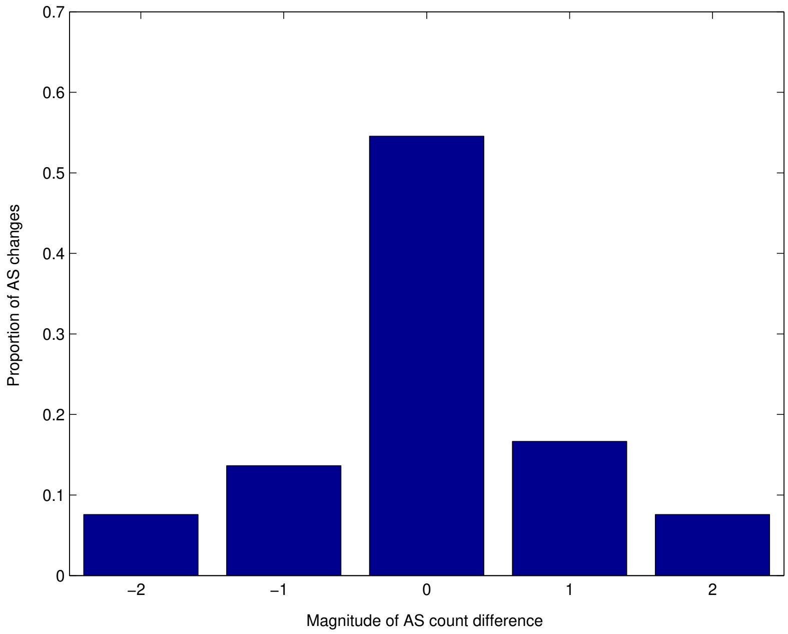


(a) DS I

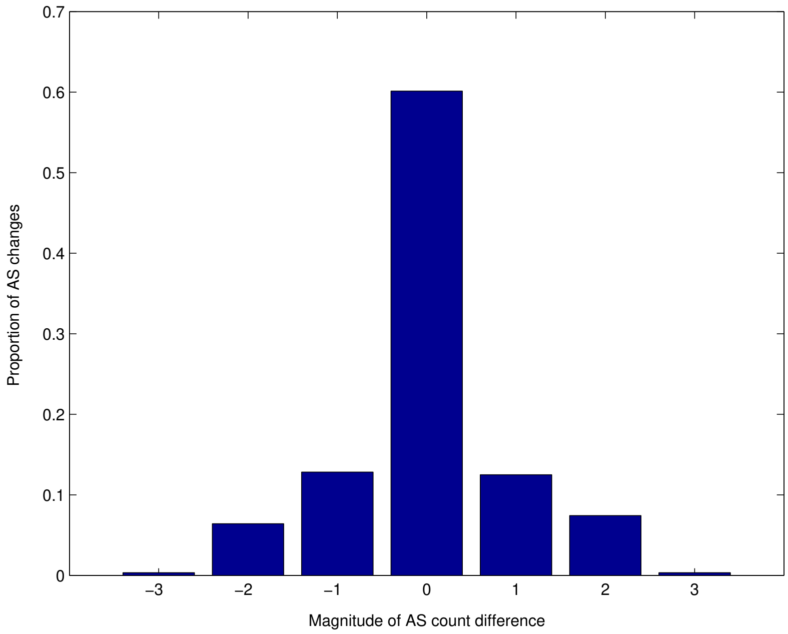


(b) DS II

Figure 1: Hop count difference distribution of TTL-visible route changes



(a) DS I



(b) DS II

Figure 2: Magnitude of AS difference distribution

Site	Connected ASes
www.valkaryn.net	AS911:Wiltel Communications Group AS2914:Verio
lava.net	AS1239:SprintLink Backbone AS3356:Level 3 Communications
www.slac.stanford.edu	AS293:ESnet AS2150:Cenic.net
www.lf.net	AS8472:BT Ignite in Germany AS20646:celox Telekommunikationsdienste GmbH
stockholm1.sunet.se	AS11908:Verestar AS8674:Netnod Internet Exchange Sverige AB

Table 10: Examples of multi-homed host sites

Second AS in the path	Third AS in the path
AS2914:Verio	AS1239:SprintLink Backbone AS3356:Level 3 Communications
AS3549:Global Crossing, Ltd.	AS7911:Wiltel Communications Group AS3356:Level 3 Communications AS22298:Ewan
AS11908:Verestar	AS1299:TeliaNet Global Network AS3356:Level 3 Communications
AS3257:Tiscali Intl Network	AS1200:Amsterdam Internet Exchange (AMS-IX) Peering AS AS1239:SprintLink Backbone AS3356:Level 3 Communications AS8674:Netnod Internet Exchange Sverige AB
AS293:ESnet	AS1239:SprintLink Backbone AS3356:Level 3 Communications
AS20965:The GEANT IP Service	AS3549:Global Crossing, Ltd. AS1299:TeliaNet Global Network

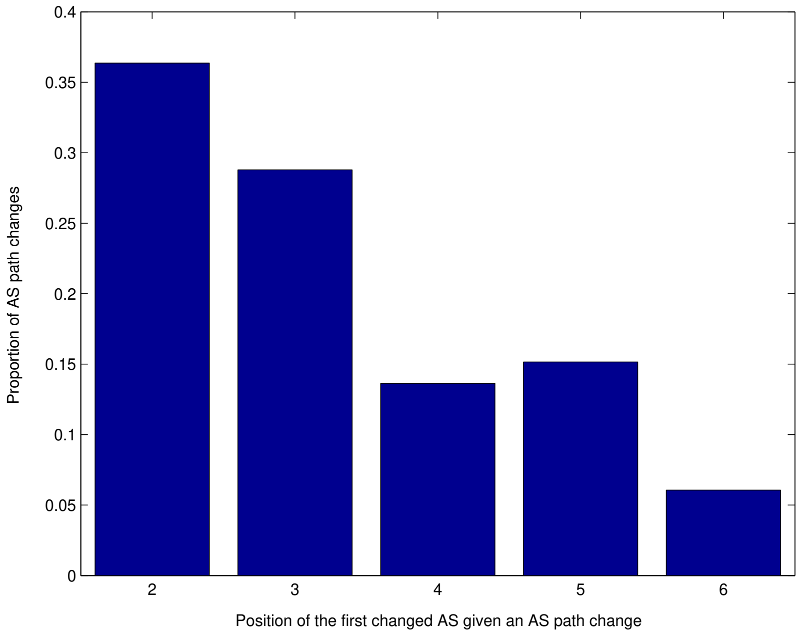
Table 11: Examples of multi-homed providers

4.6.3 AS Path Change Position and Multi-Homing

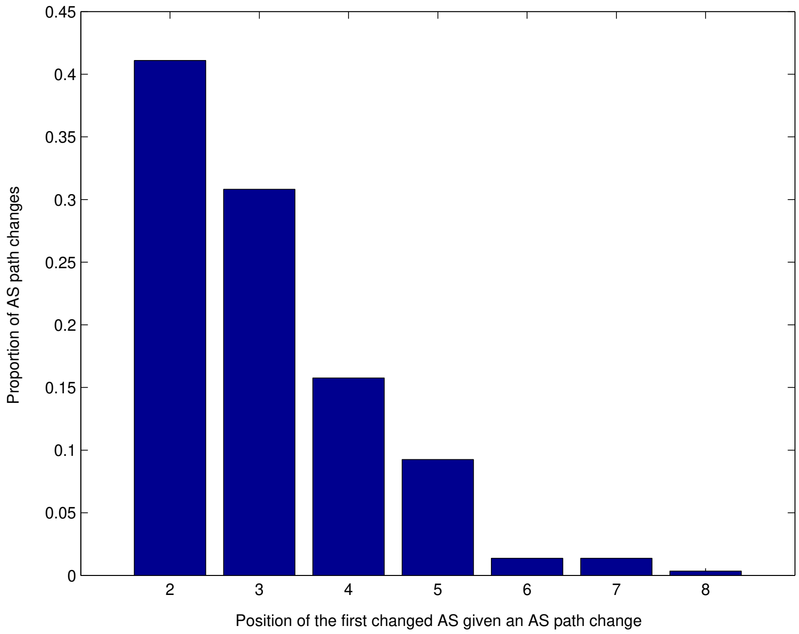
Figure 3 shows the position of the first changed AS given an AS path change. This information is helpful in understanding possible multi-homing behavior of certain sites. As we can see, over two thirds of the AS changes start from the second or the third AS. Change starting from the second AS indicates that the end system site may be connecting to multiple providers. Change starting from the third AS may indicate a smaller (regional) provider connecting to multiple larger (core) providers. Table 10 and 11 list some examples of these two categories respectively.

4.7 Dominant Route and AS Path Prevalence

Route prevalence denotes how likely it is that a path at an arbitrary point in time uses a particular route [20]. The independent and exponential interval measurement used in our experiments allows us to estimate the period of time that the path stays at one route by counting the number of times we observe that particular route during the experiment. The most frequently seen route for a path is then the most likely route or dominant route for that path. The dominant route prevalence value for a given path can then be computed by



(a) DS I



(b) DS II

Figure 3: AS change location distribution

the number of times the dominant route for the given path is sampled (n_{dom}) over the total number of times (n_{tot}) the given path is sampled during the experiment, i.e. n_{dom}/n_{tot}.

Dominant route prevalence can also be defined for a specific source site. In that case, the prevalence value will be the number of times the dominant route from this source to all destinations are sampled ($\sum_{j=1}^{d} n_{dom_j}$, d is the number of destinations, and n_{dom_j} the number of sampled dominant routes to destination j) over the total number of samplings from the source to all destinations ($\sum_{j=1}^{d} n_j$, n_j is the total number of samples from the source to destination j). Dominant route prevalence for a specific destination site may be very similarly defined.

Applying the dominant route and route prevalence concepts to AS paths, we can define dominant AS paths and AS path prevalence values in very much the same way.

4.7.1 Dominant Route and AS Path Prevalence for Individual Paths

Figure 4 shows the cumulative distribution of the dominant route prevalence for all paths in DS I and DS II. From the figure, we see that at host level, over half of the paths in DS I have a dominant route prevalence value of at least 0.77 before router merging and 0.95 after merging. The respective values for DS II are 0.55 and 0.7. At AS level, over 70% of paths have a dominant AS path prevalence value of 1 in both data sets. These findings agree with the results and conclusions made in [20] based on 1994 data that in general Internet paths are strongly dominated by a single route.

It is interesting to note that [22] actually obtained a considerably increased prevalence curve based on data of year 1999, but the authors did not conclude whether this increase indeed reflects change in Internet routing behavior because of possible effect of differing measurement methodologies. Our prevalence values obtained in 2004 are more close to the 1994 data, rather than the 1999's, indicating that measurement methodologies might indeed be playing a non-negligible effect.

4.7.2 Dominant Route Prevalence per Source or Destination

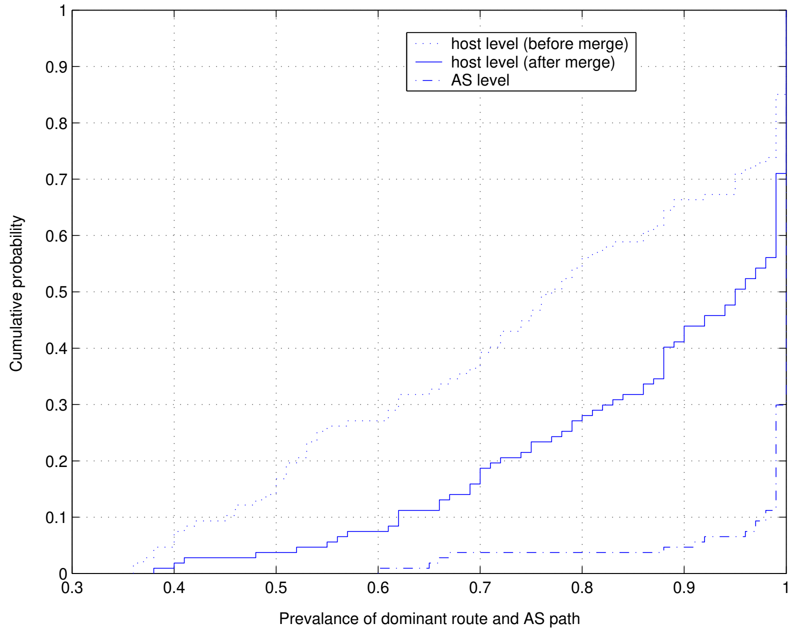
The dominant route prevalence per source or per destination are shown in Tables 12 and 13. The large spreads of prevalence values for different sources or destinations show clear site variations. Although in most cases the prevalence values for the same site as source or destination are similar, there are a few cases where these two values differ a lot. For example, the prevalence values of `www.fmp.com` as source or as destination differ by 0.23. This indicates possible asymmetric routing to and from this site.

As expected, router merging has the effect of increasing the prevalence values. In some cases, e.g., for `www.washington.edu`, the increase is substantial, obviously because most of the route changes are caused by the host site itself and they are eliminated after router merging. This again shows that router merging allows us to focus our attention on what actually happens inside the network.

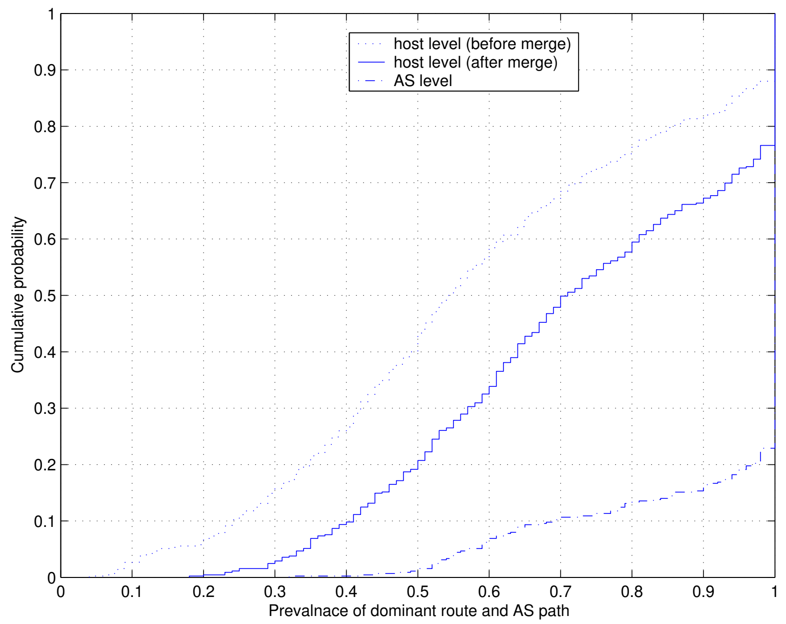
4.8 Route Persistence

Route persistence illustrates how long the path will stay in one route and it is usually difficult to evaluate because a series of measurements at particular points in time do not necessarily indicate a lack of change and then change back in between the measurement points. Following a similar approach as in [20], we looked at consecutive measurements made in different intervals at small, medium and large scales and computed the likelihood of seeing a route change during each particular interval scale. If that likelihood for a specific source or destination is too high, the site will be identified as an outlier and all related records corresponding to that site as source or destination will be removed. This way we make sure that we can trust the shorter scale measurements observing the same route as indicating that the route does indeed persist during the interval between the measurements. Then we can proceed to look at longer time scales.

Taking DS I as an example, first we looked at consecutive measurements made with interval of fewer than 60 seconds. There were 85 pairs of such measurements, none of



(a) DS I



(b) DS II

Figure 4: Cumulative distribution of route prevalence for all paths

Site	Before merge		After merge	
	Source	Destination	Source	Destination
backbone.acad.bg	0.68	N/A	0.75	N/A
lava.net	0.73	0.88	0.86	0.96
stockholm1.sunet.se	0.76	0.87	0.85	0.91
swiCE2.switch.ch	0.77	0.8	0.84	0.84
traceroute.teragen.com.au	0.8	0.74	0.9	0.89
www.debug.net	0.76	0.81	0.89	0.97
www.fh-friedberg.de	0.75	0.68	0.86	0.8
www.hafey.org	0.83	0.6	0.94	0.82
www.lf.net	0.75	0.72	0.92	0.84
www.megamirror.com	0.61	0.67	0.9	0.91
www.slac.stanford.edu	0.9	N/A	0.9	N/A
www.valkaryn.net	0.65	0.75	0.86	0.8

Table 12: Dominant route prevalence per source or destination for DS I

Site	Before merge		After merge	
	Source	Destination	Source	Destination
backbone.acad.bg	0.67	0.5	0.75	N/A
corporate.pacific.net.th	0.48	0.55	0.59	0.59
lava.net	0.55	0.56	0.67	0.75
proxy1.syd.connect.com.au	0.51	0.54	0.54	0.62
stockholm1.sunet.se	0.58	0.7	0.72	0.77
swiCE2.switch.ch	0.63	0.62	0.78	0.71
traceroute.teragen.com.au	0.65	0.62	0.87	0.73
www.cistron.net	0.65	0.6	0.76	0.72
www.cynexx.net	0.64	0.57	0.77	0.65
www.debug.net	0.74	0.65	0.86	0.73
www.fh-friedberg.de	0.62	0.62	0.71	0.67
www.fmp.com	0.57	0.35	0.65	0.42
www.globedom.com	0.53	0.62	0.58	0.68
www.hafey.org	0.52	0.66	0.64	0.73
www.io.com	0.53	0.28	0.62	0.68
www.lf.net	0.78	0.65	0.88	0.81
www.megamirror.com	0.6	0.57	0.79	0.73
www.net.berkeley.edu	0.71	0.78	0.76	0.85
www.opus1.com	0.52	0.55	0.8	0.74
www.rhnet.is	0.46	0.54	0.79	0.99
www.slac.stanford.edu	0.66	N/A	0.74	N/A
www.telcom.arizona.edu	0.46	0.45	0.53	0.52
www.valkaryn.net	0.54	0.7	0.6	0.79
www.washington.edu	0.09	0.21	0.72	0.69

Table 13: Dominant route prevalence per source or destination for DS II

them exhibited a route change. Next we looked at consecutive measurements made within interval of fewer than 10 minutes. There were total of 808 pairs of such measurements, only one route change was found among them. There were also 17 triple measurements made within 10 minutes. Triple measurements could provide additional indication of possible route changes if they were of patterns `A, B, A` or `A, B, C`. There were no such patterns found in the triple measurements within 10 minutes. With this information, the likelihood of seeing a route change during the 10 minutes interval for a source (destination) was computed by dividing the number of route changes within 10 minutes originated from (destined to) the source (destination) by total number of records originated from (destined to) the source (destination).

The above procedures are repeated for DS I with interval 1 hour and 6 hours. The whole procedure is also performed for DS II. The computed likelihood of seeing a route change in different time scales is listed in Table 14 and 15.

No outliers were identified for DS I. We plotted the distribution of the duration of long-lived routes before and after router merging in Figure 5. When plotting these figures, we used the same assumption as in [20] regarding duration estimation. First, there is no missing route change in between the measurement intervals; for a sequence of measurements all observing the same route, the route's duration is at least the span of the measurements; if at time a we observe route A, and then the next measurement at time b observes route B, our best guess is that route B began half way between these measurements. For routes observed at the beginning (end) of the measurement period, but not spanning the entire measurement period, the starting (ending) time is assigned as follows. If the next (previous) measurement also observed that route, then the route is assumed to persist for at least that much time into the past (future). If the next (previous) measurement did not observe that route, then the lone observation of the route is taken as its starting (ending) time.

Figure 6 shows the distribution of long-lived route durations for DS II. Figures 6(a) and 6(b) are results before and after router merging. Figure 6(c) is obtained after removing sites which showed a route change probability greater than 10% in any of the three time scales. They include `www.io.com`, `www.net.berkeley.edu`, `www.washington.com`, `proxy1.syd.connect.com.au`, `www.fh-friedberg.de`, `www.valkaryn.com` as sources and `traceroute.teragen.com.au`, `proxy1.syd.connect.com.au`, `www.io.com`, `www.washing-ton.com`, `www.fmp.com` as destinations.

Clearly, router merging or outlier removal tend to increase the proportion of paths with longer durations. From figures 5 and 6 we can see that the majority, or 69% and 83% of paths last more than one day; 58% and 55% of the routes persist under a week, these values generally agree with previous results in [22], where a total of about 1/3 of Internet routes are short-lived; as well as results in [20], where about half the routes lasting under a week and half over a week.

4.9 Impact of Choosing Measurement Interval

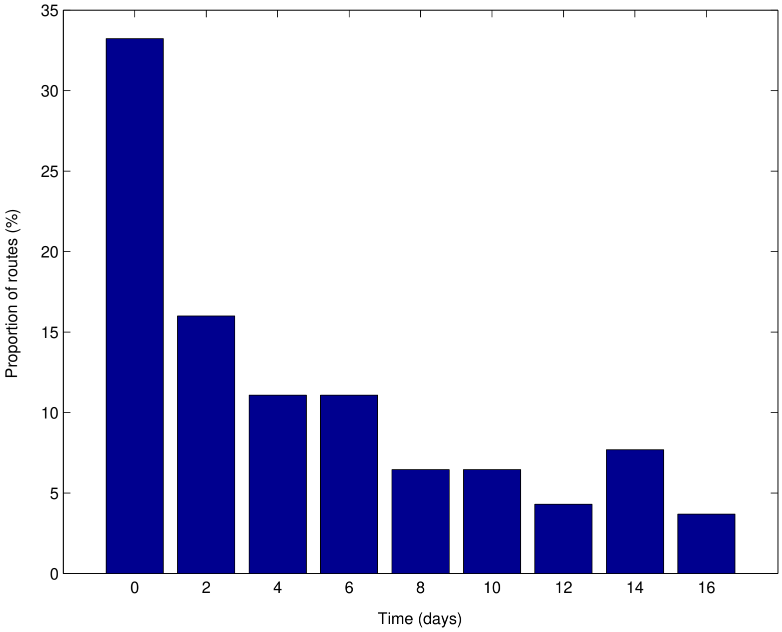
The average sampling rate of a particular `traceroute` server in our experiments is about 15 to 30 minutes. Taking into account the total number of paths between all servers, the average sampling rate for an individual path becomes 2.75 hours in DS I and 11.5 hours in DS II. To better understand whether some route dynamics is missed between the measurement intervals when the sampling rate is relatively low, we collected DS IV, where we chose four paths listed in table 16 to perform both short scale fixed interval (10 minutes) samplings and medium scale exponential interval (2 hours on average) samplings. To make the results more plausible, we checked the route change frequency of these four paths from DS II. According to Figure 7, the histogram of number of route changes for all paths in DS II (before router merging), 81% of total paths exhibited fewer than 10 route changes. The four paths selected in DS IV experienced 3 to 7 route changes in DS II, as shown in the last column (Count) of Table 16. So they are among the majority of paths in terms of route change frequency. Below we look at the measurement results for each of these four paths.

	10 minutes		1 hour		6 hours	
Site	Source	Destination	Source	Destination	Source	Destination
backbone.acad.bg	0	NA	0.30	NA	1.05	NA
lava.net	0	0	1.06	0.5	2.52	1.25
stockholm1.sunet.se	0	0	0.54	0.24	1.53	0.83
swice2.switch.ch	0	0	0	1.07	1.12	1.78
traceroute.teragen.com.au	0	0	0.66	0.99	1.84	1.62
www.debug.net	0	1.3	0.55	1.26	1.25	0.97
www.fh-friedberg.de	0	0	0	1.44	0.65	2.34
www.hafey.org	1.7	0	1.81	0.89	1.73	1.69
www.lf.net	0	0	0.8	0.45	1.46	1.63
www.megamirror.com	0	0	1.07	0	1.67	0.74
www.slac.stanford.edu	0	NA	0.28	NA	0.49	NA
www.valkaryn.net	0	0	1.25	0.24	2.06	1.67

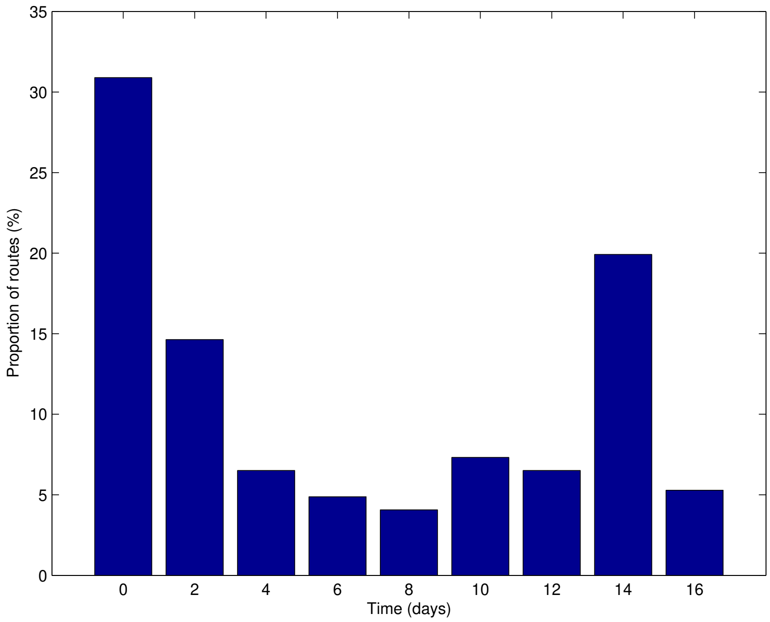
Table 14: Likelihood (%) of seeing route changes in different time scales for DS I

	10 minutes		1 hour		6 hours	
Site	Source	Destination	Source	Destination	Source	Destination
backbone.acad.bg	0	NA	3.63	NA	8.08	NA
corporate.pacific.net.th	0	0	2.78	0	2.56	4.21
lava.net	0	8.33	2.70	4.17	6.42	4.27
proxy1.syd.connect.com.au	0	10	14.49	14.63	13.06	13.06
stockholm1.sunet.se	0	10	3.33	1.54	2.97	2.40
swice2.switch.ch	6.67	6.25	4.05	3.95	4.17	3.97
traceroute.teragen.com.au	0	12.5	1.72	5.06	2.06	6.19
www.cistron.net	0	0	0	1.54	3.65	3.95
www.cynexx.net	0	7.69	3.45	5.33	2.68	7.2
www.debug.net	0	0	0	1.1	1.37	3.26
www.fh-friedberg.de	0	0	11.32	0	9.06	4.35
www.fmp.com	0	NA	0	0	1.81	26.32
www.globedom.com	0	0	6.78	6.1	7.48	5.77
www.hafey.org	0	0	0	5	4.4	5.41
www.io.com	30	7.14	22.37	25	21.92	23.17
www.lf.net	0	0	0	4.41	1.29	3.45
www.megamirror.com	0	8.33	0	7.69	1.92	9.32
www.net.berkeley.edu	22.22	0	10.96	1.47	10.4	1.48
www.opus1.com	0	0	3.51	1.14	7.29	1.81
www.rhnet.is	0	NA	0	0	1.82	0
www.slac.stanford.edu	0	NA	2.6	NA	3.26	NA
www.telcom.arizona.edu	0	0	4.65	0	2.80	2.71
www.valkaryn.net	0	0	12.31	0	11.24	1.37
www.washington.edu	25	4.76	20	15.73	20.20	16.43

Table 15: Likelihood (%) of seeing route changes in different time scales for DS II

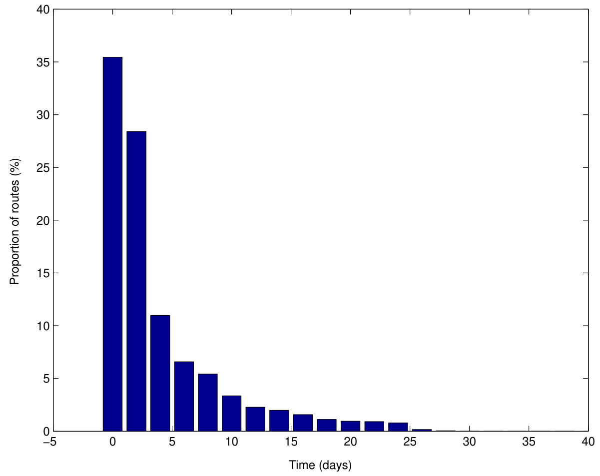


(a) Before router merging

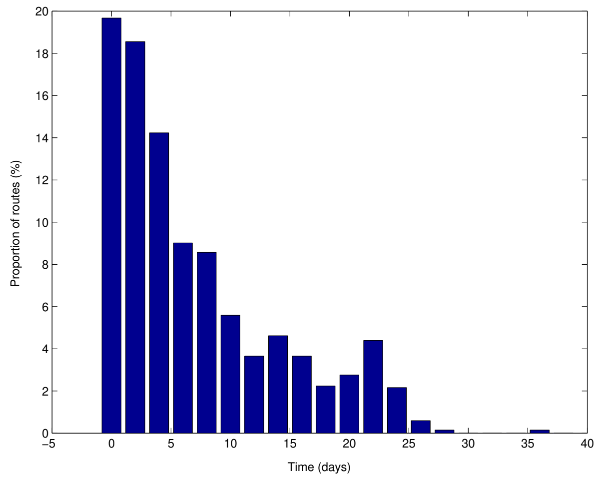


(b) After router merging

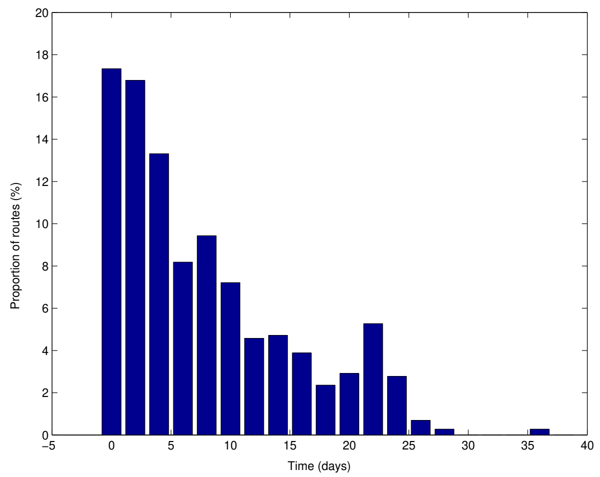
Figure 5: Distribution of route duration for DS I



(a) Before router merging



(b) After router merging



(c) After router merging and outlier removal

Figure 6: Distribution of route duration for DS II

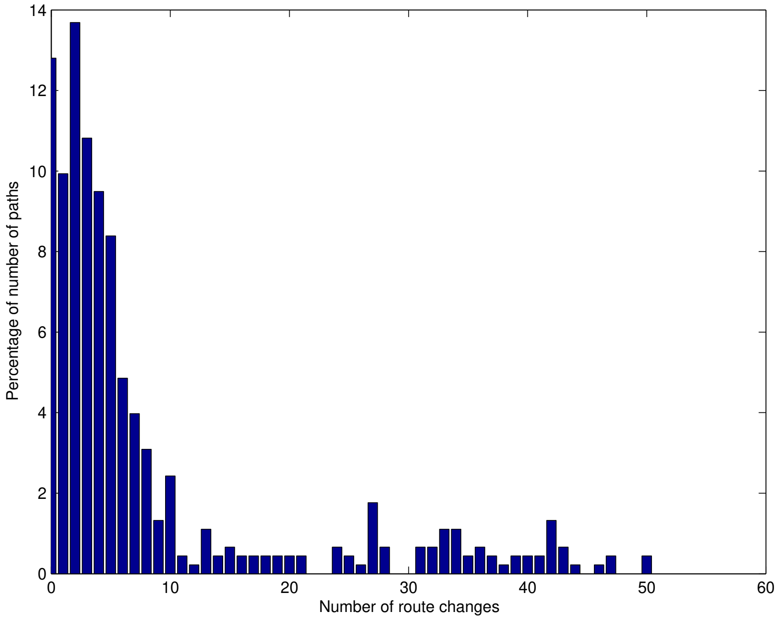


Figure 7: Number of paths vs. route change frequency in DS II

Source	Destination	Count
www.lf.net	traceroute.teragen.com.au	6
traceroute.teragen.com.au	www.lf.net	6
lava.net	www.globedom.com	3
www.globedom.com	lave.net	7

Table 16: Four paths used for short interval measurement

Measurement time	Route id	AS path id
Wed Jul 14 02:39:55 AM	1	1
Mon Jul 19 06:20:38 PM	4	1
Mon Jul 19 09:10:38 PM	1	1
Thu Jul 22 06:21:01 PM	2	1
Fri Jul 23 07:51:09 PM	1	1
Sat Jul 24 06:51:16 PM	3	1
Sat Jul 24 10:01:17 PM	1	1
Sun Jul 25 06:11:20 AM	3	1
Sun Jul 25 08:11:20 AM	1	1

Table 17: Route changes for the path from `lf` to `teragen` (10-minute interval measurement)

Measurement time	Route id	AS path id
Wed Jul 14 02:39:55 AM	1	1
Mon Jul 19 08:14:11 PM	4	1
Mon Jul 19 11:53:10 PM	1	1
Fri Jul 23 01:20:13 AM	2	1
Fri Jul 23 10:14:18 PM	1	1
Sat Jul 24 07:06:50 PM	3	1
Sat Jul 24 10:30:22 PM	1	1
Sun Jul 25 07:19:04 AM	3	1
Sun Jul 25 09:10:17 AM	1	1

Table 18: Route changes for the path from `lf` to `teragen` (2-hour interval measurement)

4.9.1 Path from `www.lf.net` to `traceroute.teragen.com.au`

We observed 8 route changes in both the 10-minute and 2-hour interval measurements for the path from `www.lf.net` to `traceroute.teragen.com.au`. Tables 17 and 18 show the time when route changes are observed in the two measurement schemes. Route ids are unique identifiers assigned to each different route of the path. With route ids we can track the exact type of route changes. AS path ids are defined similarly. Note that the first row in the tables shows the initial status when measurement is started and is not a route change.

It can be seen that the route with id=1 appears to be the dominant route for this path. If we assume that we do not miss any route changes in the 10-minute interval measurement, since there is no route change in any two consecutive measurement instances, all routes last at least 10 minutes. The shortest route holding time of this path during the measurement period will then be about 2 hours (minus or plus 10 minutes) which corresponding to route id=3 in its second instance starting at 06:11:20 AM, Jul 25, 2004. Therefore, this path appears to be fairly stable. There is also no AS change for this path since only one unique AS path is observed.

For this particular case, we see that the 2-hour exponential interval measurement did an equally good job in capturing all route changes seen by the 10-minute fixed interval measurement.

4.9.2 Path from `traceroute.teragen.com.au` to `www.lf.net`

Tables 19 and 20 show observed route changes for the path from `traceroute.teragen.com.au` to `www.lf.net`. Identifying the dominant route is less obvious in these two tables since

Measurement time	Route id	AS path id
Wed Jul 14 02:39:55 AM	1	1
Sat Jul 24 06:51:16 PM	4	1
Sat Jul 24 10:11:17 PM	1	1
Sun Jul 25 06:11:20 AM	4	1
Sun Jul 25 08:11:20 AM	1	1
Wed Jul 28 01:51:43 AM	3	1
Thu Jul 29 01:11:57 PM	2	1
Sat Jul 31 02:02:10 AM	5	1
Sat Jul 31 02:22:10 AM	2	1
Sat Jul 31 09:12:12 AM	5	1
Sat Jul 31 10:52:12 AM	2	1
Sat Jul 31 11:12:13 AM	5	1
Sat Jul 31 12:32:13 PM	2	1
Mon Aug 02 04:22:25 AM	5	1
Mon Aug 02 04:32:25 AM	2	1
Mon Aug 02 04:52:25 AM	5	1
Mon Aug 02 05:02:25 AM	2	1
Tue Aug 03 10:52:35 AM	5	1

Table 19: Route changes for the path from `teragen` to `lf` (10-minute interval measurement)

Measurement time	Route id	AS path id
Wed Jul 14 02:39:55 AM	1	1
Sat Jul 24 07:06:50 PM	4	1
Sat Jul 24 10:30:22 PM	1	1
Sun Jul 25 07:19:04 AM	4	1
Sun Jul 25 09:10:17 AM	1	1
Wed Jul 28 02:08:11 AM	3	1
Thu Jul 29 01:58:29 PM	2	1
Sat Jul 31 02:10:07 AM	5	1
Sat Jul 31 04:12:11 AM	2	1
Sat Jul 31 11:34:09 AM	5	1
Sat Jul 31 12:46:59 PM	2	1

Table 20: Route changes for the path from `teragen` to `lf` (2-hour interval measurement)

Measurement time	Route id	AS path id
Sat Jul 31 09:12:12 AM	5	1
Sat Jul 31 10:52:12 AM	2	1
Mon Aug 02 04:22:25 AM	5	1
Mon Aug 02 04:32:25 AM	2	1
Mon Aug 02 04:52:25 AM	5	1
Mon Aug 02 05:02:25 AM	2	1
Tue Aug 03 10:52:35 AM	5	1

Table 21: Missed route changes by the 2-hour interval measurement for path from `teragen` to `lf`

different routes appeared similar number of times. Although both measurement schemes captured 5 unique routes, the 10-minute measurement recorded 17 route changes but the 2-hour interval measurement recorded only 10. Table 21 lists the missed route changes separately. It shows that all missed route changes are changes between route id=2 and route id=5. The missed changes lasted variable periods of time, e.g., changes made to route id=5 last from 10 minutes to almost 2 hours. We show the route change happened at 04:22:25 AM EDT, Aug 02, 2004 below as an example.

```
Location: traceroute.teragen.com.au --to-- www.lf.net   Status: ok
Time: Mon Aug 02 04:12:25 AM EDT 2004 --to-- Mon Aug 02 04:12:44 AM EDT 2004

 1  core.mel.teragen.com.au (203.22.251.136)  0.643 ms  0.430 ms  0.395 ms
 2  fa0-1-5.que31-cs1.vic.uecomm.net.au (203.94.134.49)  0.937 ms  0.776 ms  0.781 ms
 3  atm4-0-0.10211.agg1.cit190.uecomm.net.au (218.185.29.229)  1.308 ms  1.312 ms  1.293 ms
 4  218.185.31.122 (218.185.31.122)  1.470 ms  1.770 ms  1.347 ms
 5  sl-gw1-mel-6-0-0.sprintlink.net (203.222.35.229)  2.452 ms  1.739 ms  1.865 ms
 6  sl-bb21-syd-1-0.sprintlink.net (203.222.33.18)  14.312 ms  14.255 ms  14.758 ms
 7  sl-bb21-syd-15-1.sprintlink.net (203.222.32.34)  14.253 ms  14.023 ms  14.231 ms
 8  sl-bb21-sj-3-2.sprintlink.net (144.232.8.130)  163.028 ms  163.768 ms  163.619 ms
 9  sl-bb25-sj-12-0.sprintlink.net (144.232.3.210)  163.558 ms  163.313 ms  163.440 ms
10  sl-bb23-sj-15-0.sprintlink.net (144.232.3.249)  193.231 ms  163.525 ms  163.324 ms
11  sl-bb25-rly-0-0.sprintlink.net (144.232.20.10)  240.284 ms  239.655 ms  240.112 ms
12  sl-bb21-rly-11-0.sprintlink.net (144.232.14.157)  239.633 ms  240.475 ms  239.536 ms
13  sl-bb20-tuk-2-0.sprintlink.net (144.232.20.114)  259.407 ms  259.997 ms  259.707 ms
14  sl-bb21-tuk-15-0.sprintlink.net (144.232.20.133)  247.616 ms  247.851 ms  248.803 ms
15  sl-bb21-lon-14-0.sprintlink.net (144.232.19.70)  317.441 ms  316.855 ms  322.521 ms
16  sl-gw10-lon-15-0.sprintlink.net (213.206.128.46)  317.183 ms  316.233 ms  317.488 ms
17  sle-btglo1-1-0.sprintlink.net (213.206.159.42)  316.490 ms  316.798 ms  316.380 ms
18  t2c2-ge7-0.uk-ilf.eu.bt.net (166.49.208.78)  318.087 ms  317.395 ms  317.448 ms
19  t2c2-p8-0.uk-lon2.eu.bt.net (166.49.195.126)  317.987 ms  317.910 ms  318.064 ms
20  t2c1-p1-0.de-fra.eu.bt.net (166.49.195.230)  344.164 ms  344.407 ms  344.694 ms
21  t2a4-ge6-0.de-fra.eu.bt.net (166.49.172.20)  344.372 ms  343.653 ms  344.087 ms
22  166-49-147-130.eu.bt.net (166.49.147.130)  344.087 ms  344.311 ms  344.220 ms
23  pos-2-0-c1.s.core.de.ignite.net (195.182.99.69)  346.922 ms  361.482 ms  346.580 ms
24  fa-1-0-0-a1.s.core.de.ignite.net (195.182.97.195)  346.886 ms  346.858 ms  347.105 ms
25  212.60.208.70 (212.60.208.70)  342.239 ms  341.944 ms  341.961 ms
26  ftp.LF.net (212.9.160.37)  342.451 ms  342.943 ms  343.104 ms


Location: traceroute.teragen.com.au --to-- www.lf.net   Status: ok
Time: Mon Aug 02 04:22:25 AM EDT 2004 --to-- Mon Aug 02 04:22:46 AM EDT 2004

 1  core.mel.teragen.com.au (203.22.251.136)  0.313 ms  0.143 ms  0.113 ms
 2  fa0-1-5.que31-cs1.vic.uecomm.net.au (203.94.134.49)  0.671 ms  1.215 ms  0.405 ms
 3  atm4-0-0.10211.agg1.cit190.uecomm.net.au (218.185.29.229)  1.387 ms  1.065 ms  1.321 ms
 4  218.185.31.122 (218.185.31.122)  1.352 ms  1.414 ms  1.343 ms
 5  sl-gw1-mel-6-0-0.sprintlink.net (203.222.35.229)  154.995 ms  28.438 ms  200.747 ms
 6  sl-bb20-syd-1-0.sprintlink.net (203.222.33.18)  13.907 ms  14.014 ms  13.979 ms
 7  sl-bb21-syd-15-1.sprintlink.net (203.222.32.34)  14.166 ms  14.047 ms  14.012 ms
 8  sl-bb22-sj-3-3.sprintlink.net (144.232.8.130)  163.314 ms  163.230 ms  162.751 ms
 9  sl-bb20-sj-15-0.sprintlink.net (144.232.3.166)  164.420 ms  163.864 ms  163.991 ms
```

```
10  sl-bb20-stk-12-0.sprintlink.net (144.232.20.98)  165.300 ms  165.956 ms  165.293 ms
11  sl-bb25-pen-12-0.sprintlink.net (144.232.20.166)  236.276 ms  234.757 ms  235.330 ms
12  sl-bb24-pen-14-0.sprintlink.net (144.232.16.77)  234.363 ms  234.673 ms  234.745 ms
13  sl-bb23-pen-8-0.sprintlink.net (144.232.16.65)  235.330 ms  234.463 ms  234.785 ms
14  sl-bb21-tuk-3-0.sprintlink.net (144.232.20.112)  237.090 ms  237.158 ms  237.079 ms
15  sl-bb21-lon-14-0.sprintlink.net (144.232.19.70)  305.071 ms  305.187 ms  305.503 ms
16  sl-gw10-lon-15-0.sprintlink.net (213.206.128.46)  304.769 ms  304.655 ms  306.098 ms
17  sle-btglo1-1-0.sprintlink.net (213.206.159.42)  305.179 ms  305.072 ms  304.945 ms
18  t2c2-ge7-0.uk-ilf.eu.bt.net (166.49.208.78)  304.553 ms  304.446 ms  304.845 ms
19  t2c2-p8-0.uk-lon2.eu.bt.net (166.49.195.126)  305.762 ms  304.672 ms  305.442 ms
20  t2c1-p1-0.de-fra.eu.bt.net (166.49.195.230)  331.004 ms  331.339 ms  331.082 ms
21  t2a4-ge6-0.de-fra.eu.bt.net (166.49.172.20)  331.072 ms  331.787 ms  331.585 ms
22  166-49-147-130.eu.bt.net (166.49.147.130)  331.278 ms  331.535 ms  330.963 ms
23  pos-2-0-c1.s.core.de.ignite.net (195.182.99.69)  336.185 ms  334.703 ms  362.433 ms
24  fa-1-0-0-a1.s.core.de.ignite.net (195.182.97.195)  335.013 ms  335.089 ms  334.950 ms
25  212.60.208.70 (212.60.208.70)  330.294 ms  330.064 ms  331.493 ms
26  ftp.LF.net (212.9.160.37)  330.724 ms  332.213 ms  331.165 ms
```

```
Location: traceroute.teragen.com.au --to-- www.lf.net   Status: ok
Time: Mon Aug 02 04:32:25 AM EDT 2004 --to-- Mon Aug 02 04:32:49 AM EDT 2004

 1  core.mel.teragen.com.au (203.22.251.136)  0.645 ms  0.426 ms  0.402 ms
 2  fa0-1-5.que31-cs1.vic.uecomm.net.au (203.94.134.49)  0.775 ms  0.775 ms  0.758 ms
 3  atm4-0-0.10211.agg1.cit190.uecomm.net.au (218.185.29.229)  1.299 ms  1.310 ms  1.294 ms
 4  218.185.31.122 (218.185.31.122)  1.486 ms  1.404 ms  1.340 ms
 5  sl-gw1-mel-6-0-0.sprintlink.net (203.222.35.229)  2.088 ms  2.089 ms  2.044 ms
 6  sl-bb20-syd-1-0.sprintlink.net (203.222.33.18)  14.130 ms  14.171 ms  14.897 ms
 7  sl-bb21-syd-15-1.sprintlink.net (203.222.32.34)  13.916 ms  13.937 ms  14.138 ms
 8  sl-bb21-sj-3-2.sprintlink.net (144.232.8.130)  163.287 ms  163.707 ms  163.797 ms
 9  sl-bb25-sj-12-0.sprintlink.net (144.232.3.210)  163.384 ms  163.530 ms  182.597 ms
10  sl-bb23-sj-15-0.sprintlink.net (144.232.3.249)  176.881 ms  177.369 ms  198.263 ms
11  sl-bb25-rly-0-0.sprintlink.net (144.232.20.10)  239.508 ms  247.188 ms  239.600 ms
12  sl-bb21-rly-11-0.sprintlink.net (144.232.14.157)  253.311 ms  242.509 ms  239.506 ms
13  sl-bb20-tuk-2-0.sprintlink.net (144.232.20.114)  260.084 ms  259.363 ms  277.311 ms
14  sl-bb21-tuk-15-0.sprintlink.net (144.232.20.133)  247.974 ms  249.217 ms  262.653 ms
15  sl-bb21-lon-14-0.sprintlink.net (144.232.19.70)  509.878 ms  356.110 ms  316.486 ms
16  sl-gw10-lon-15-0.sprintlink.net (213.206.128.46)  316.286 ms  317.462 ms  316.681 ms
17  sle-btglo1-1-0.sprintlink.net (213.206.159.42)  316.332 ms  316.882 ms  317.958 ms
18  t2c2-ge7-0.uk-ilf.eu.bt.net (166.49.208.78)  335.655 ms  317.311 ms  317.306 ms
19  t2c2-p8-0.uk-lon2.eu.bt.net (166.49.195.126)  318.165 ms  317.774 ms  317.917 ms
20  t2c1-p1-0.de-fra.eu.bt.net (166.49.195.230)  343.895 ms  344.890 ms  354.856 ms
21  t2a4-ge6-0.de-fra.eu.bt.net (166.49.172.20)  343.756 ms  364.800 ms  393.646 ms
22  166-49-147-130.eu.bt.net (166.49.147.130)  344.594 ms  344.460 ms  344.444 ms
23  pos-2-0-c1.s.core.de.ignite.net (195.182.99.69)  346.601 ms  346.817 ms  346.510 ms
24  fa-1-0-0-a1.s.core.de.ignite.net (195.182.97.195)  346.958 ms  347.127 ms  347.844 ms
25  212.60.208.70 (212.60.208.70)  341.891 ms  342.459 ms  341.384 ms
26  ftp.LF.net (212.9.160.37)  342.218 ms  343.784 ms  341.752 ms
```

The three records above show the original route, the first record after route change, and the first record after the route comes back to the original state. It turns out that this is a route change that affects the following five routers:

```
 9  sl-bb25-sj-12-0.sprintlink.net (144.232.3.210)  163.558 ms  163.313 ms  163.440 ms
10  sl-bb23-sj-15-0.sprintlink.net (144.232.3.249)  193.231 ms  163.525 ms  163.324 ms
11  sl-bb25-rly-0-0.sprintlink.net (144.232.20.10)  240.284 ms  239.655 ms  240.112 ms
12  sl-bb21-rly-11-0.sprintlink.net (144.232.14.157)  239.633 ms  240.475 ms  239.536 ms
13  sl-bb20-tuk-2-0.sprintlink.net (144.232.20.114)  259.407 ms  259.997 ms  259.707 ms
14  sl-bb21-tuk-15-0.sprintlink.net (144.232.20.133)  247.616 ms  247.851 ms  248.803 ms

 9  sl-bb20-sj-15-0.sprintlink.net (144.232.3.166)  164.420 ms  163.864 ms  163.991 ms
10  sl-bb20-stk-12-0.sprintlink.net (144.232.20.98)  165.300 ms  165.956 ms  165.293 ms
11  sl-bb25-pen-12-0.sprintlink.net (144.232.20.166)  236.276 ms  234.757 ms  235.330 ms
12  sl-bb24-pen-14-0.sprintlink.net (144.232.16.77)  234.363 ms  234.673 ms  234.745 ms
13  sl-bb23-pen-8-0.sprintlink.net (144.232.16.65)  235.330 ms  234.463 ms  234.785 ms
14  sl-bb21-tuk-3-0.sprintlink.net (144.232.20.112)  237.090 ms  237.158 ms  237.079 ms
```

Therefore, this path is apparently less stable and the 2-hour interval measurement is not sufficient to accurately capture the route dynamics. On the other hand, since the frequent route changes for this path is clearly restricted to the `Sprintlink` domain, and it incurs no TTL change, the best party to detect and deal with such changes will be the network service provider.

4.9.3 Path from `lava.net` to `www.globedom.com`

On this path we observed 382 route changes in the 10-minute interval measurement and 36 route changes in the 2-hour interval measurement. This is quite surprising because we have checked that this path is not a highly fluctuating path based on our analysis on DS II. The actual number of route changes shows that the path is experiencing constant changes. The short interval measurement contains roughly ten times the number of samples of the long interval measurement, and it recorded roughly ten times the number of route changes of the latter. One possibility that could have caused this is that the characteristics for this path had completely changed between the time we conducted DS II measurement and the time of DS IV measurement. This is confirmed by a quick check on records of both data sets. A typical route from `lava.net` to `www.globedom.com` in DS II contains 13 hops as shown below.

```
Location: lava.net --to-- www.globedom.com      Status: ok
Time: Tue May 25 12:26:34 PM EDT 2004 --to-- Tue May 25 12:26:44 PM EDT 2004

 1  iiwi-fe-0-0-0.lava.net (64.65.64.30)  0.526 ms  0.518 ms  0.461 ms
 2  s2-6.hsa1.hon2.bbnplanet.net (4.24.134.17)  0.455 ms  0.399 ms  0.870 ms
 3  so-0-1-0.mpls1.Honolulu2.Level3.net (4.68.112.89)  0.597 ms  0.594 ms  0.555 ms
 4  so-0-0-0.bbr2.Washington1.Level3.net (64.159.1.158)  114.564 ms  114.050 ms  114.119 ms
 5  as-0-0.mp1.London2.Level3.net (4.68.128.101)  192.255 ms  192.572 ms  192.249 ms
 6  so-3-0-0.mp2.Frankfurt1.Level3.net (212.187.128.29)  207.701 ms  207.785 ms  207.102 ms
 7  ge-11-2.ipcolo1.Frankfurt1.Level3.net (195.122.136.115)  207.435 ms  207.534 ms  207.578 ms
 8  212.162.1.131 (212.162.1.131)  229.697 ms  229.085 ms  229.524 ms
 9  internic-01361-win-ix-i1.c.telia.net (213.248.79.146)  230.609 ms  230.841 ms  230.399 ms
10  fwext-129-6400.internic.at (212.126.64.19)  231.655 ms  231.328 ms  231.576 ms
11  net-65-cisco3600.internic.at (212.126.65.2)  225.932 ms  225.820 ms  225.635 ms
12  route65-3.internic.at (212.126.65.3)  225.584 ms  225.543 ms  228.113 ms
13  www.globedom.com (212.126.66.8)  235.085 ms  234.095 ms  235.209 ms |
```

A typical route for the same path in DS IV, however, contains 21 hops.

```
Location: lava.net --to-- www.globedom.com      Status: ok
Time: Wed Jul 14 02:59:54 AM EDT 2004 --to-- Wed Jul 14 03:00:05 AM EDT 2004

 1  iiwi-fe-0-0-0.lava.net (64.65.64.30)  0.583 ms  0.448 ms  0.431 ms
 2  noio-fe-0-0-3-16.lava.net (64.65.127.97)  0.521 ms  0.450 ms  0.441 ms
 3  sl-gw1-prl-8-0-0.sprintlink.net (160.81.200.65)  1.591 ms  1.459 ms  1.410 ms
 4  sl-bb21-prl-2-0.sprintlink.net (144.232.30.2)  1.470 ms  1.354 ms  1.341 ms
 5  sl-bb22-stk-5-3.sprintlink.net (144.232.8.34)  55.814 ms  55.780 ms  55.781 ms
 6  sl-bb26-pen-12-0.sprintlink.net (144.232.20.162)  124.395 ms  124.460 ms  124.400 ms
 7  sl-bb27-pen-15-0.sprintlink.net (144.232.16.94)  124.446 ms  124.325 ms  124.516 ms
 8  sl-bb21-nyc-2-0.sprintlink.net (144.232.20.96)  126.528 ms  126.775 ms  126.591 ms
 9  sl-bb23-nyc-3-0.sprintlink.net (144.232.7.109)  126.597 ms  126.605 ms  126.825 ms
10  sl-gw27-nyc-14-0.sprintlink.net (144.232.7.22)  126.308 ms  126.299 ms  126.275 ms
11  sl-telia1-4-0.sprintlink.net (144.232.230.30)  122.655 ms  122.644 ms  122.706 ms
12  ldn-bb1-pos6-0-0.telia.net (213.248.65.89)  223.511 ms  223.504 ms  225.303 ms
13  prs-bb1-pos6-0-0.telia.net (213.248.64.9)  230.682 ms  230.812 ms  231.212 ms
14  ffm-bb1-pos6-0-0.telia.net (213.248.64.109)  240.366 ms  240.380 ms  240.270 ms
15  win-b1-pos5-0.telia.net (213.248.65.70)  247.136 ms  247.531 ms  248.075 ms
16  win-ix-i1-pos0-0-0.telia.net (213.248.77.150)  253.753 ms  253.643 ms  256.749 ms
17  internic-01361-win-ix-i1.c.telia.net (213.248.79.146)  258.264 ms  248.626 ms  248.056 ms
18  fwext-129-6400.internic.at (212.126.64.19)  249.454 ms  249.230 ms  249.028 ms
19  net-65-cisco3600.internic.at (212.126.65.2)  257.106 ms  260.760 ms  257.410 ms
20  route65-3.internic.at (212.126.65.3)  258.117 ms  258.019 ms  263.780 ms
21  www.globedom.com (212.126.66.8)  252.762 ms  253.796 ms  252.398 ms |
```

We infer from these records that `lava.net` is multi-homed with `Sprintlink` and `Level3`, as illustrated in figure 8. Apparently `lava.net` has switched its primary outgoing

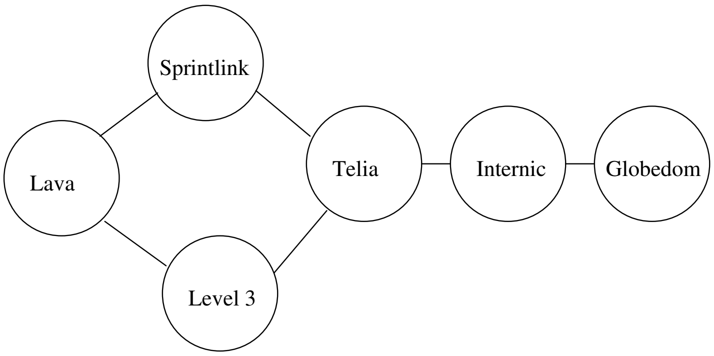


Figure 8: AS paths between `lava` and `globedom`

ISP from `Level3` in May to `Sprintlink` in mid-July. Therefore some of the `Sprintlink` routers are likely the cause of the suddenly increased number of route changes observed. Following a similar procedure in the router merging part of the DSI and DS II analysis, we listed the router pairs that were responsible for all TTL-invisible route changes and appeared in the same positions in adjacent `tracroute` records for this path in Tables 22 and 23.

Our guess is proved true. We see that the pair of `Sprint` routers in the first row of these tables, possibly doing load balancing, are responsible for virtually all route changes. Note that these two routers have similar hostnames but their IP addresses differ by three bytes. A query to the RADB [28] shows that both address ranges of 144.232.0.0/16 and 144.223.0.0/16 belong to `Sprintlink` AS 1239. We applied similar router merging techniques as in previous analysis to this path and listed the resulting route changes in Table 24 and 25.

Now the difference between the two measurement schemes is more plausible. The 2-hour measurement missed 8 out of 15 route changes. It also missed 2 unique routes and 1 unique AS paths. We looked at its undetected AS path change in particular, as shown below.

```
Location: lava.net --to-- www.globedom.com   Status: ok
Time: Mon Jul 26 05:41:31 PM EDT 2004 --to-- Mon Jul 26 05:41:43 PM EDT 2004

 1  iiwi-fe-0-0-0.lava.net (64.65.64.30)  0.530 ms  0.494 ms  0.724 ms
 2  noio-fe-0-0-3-16.lava.net (64.65.127.97)  0.422 ms  0.493 ms  0.486 ms
 3  sl-gw1-prl-8-0-0.sprintlink.net (160.81.200.65)  629.200 ms  2.733 ms  351.098 ms
 4  sl-bb21-prl-2-0.sprintlink.net (144.232.30.2)  1.797 ms  1.861 ms  1.422 ms
 5  sl-bb22-stk-5-3.sprintlink.net (144.232.8.34)  55.848 ms  56.112 ms  55.892 ms
 6  sl-bb26-pen-12-0.sprintlink.net (144.232.20.162)  202.193 ms  346.650 ms  178.191 ms
 7  sl-bb27-pen-15-0.sprintlink.net (144.232.16.94)  124.536 ms  124.831 ms  124.600 ms
 8  sl-bb21-nyc-2-0.sprintlink.net (144.232.20.96)  251.045 ms  135.503 ms  279.002 ms
 9  sl-bb23-nyc-3-0.sprintlink.net (144.232.7.109)  200.571 ms  209.132 ms  225.675 ms
10  sl-gw27-nyc-14-0.sprintlink.net (144.232.7.22)  127.004 ms  126.506 ms  126.746 ms
11  sl-telia1-4-0.sprintlink.net (144.232.230.30)  127.667 ms  122.916 ms
    sl-telia1-1-0.sprintlink.net (144.223.27.146)  122.858 ms
12  nyk-bb2-pos0-0-0.telia.net (213.248.80.134)  123.721 ms  122.914 ms  123.807 ms
13  ldn-bb2-pos6-0-0.telia.net (213.248.65.93)  197.548 ms  197.553 ms  196.812 ms
14  prs-bb2-pos6-0-0.telia.net (213.248.65.114)  205.973 ms  204.139 ms  203.873 ms
15  ffm-bb2-pos6-0-0.telia.net (213.248.65.118)  213.809 ms  214.309 ms  214.553 ms
16  win-b1-pos5-1.telia.net (213.248.65.74)  241.523 ms  242.805 ms  246.350 ms
17  win-ix-i1-pos0-0-0.telia.net (213.248.77.150)  391.680 ms  374.195 ms  243.089 ms
18  internic-01361-win-ix-i1.c.telia.net (213.248.79.146)  238.210 ms  234.474 ms  235.108 ms
19  fwext-129-6400.internic.at (212.126.64.19)  235.726 ms  236.642 ms  235.698 ms
20  net-65-cisco3600.internic.at (212.126.65.2)  246.089 ms  246.415 ms  260.978 ms
21  route65-3.internic.at (212.126.65.3)  258.089 ms  247.813 ms  254.910 ms
22  www.globedom.com (212.126.66.8)  244.034 ms  242.199 ms  238.544 ms


Location: lava.net --to-- www.globedom.com   Status: ok
```

Count	IP address A	IP address B	Hostname A	Hostname B
368	144.223.27.146	144.232.230.30	sl-telia1-1-0.sprintlink.net	sl-telia1-4-0.sprintlink.net
2	213.248.64.9	213.248.65.146	prs-bb1-pos6-0-0.telia.net	hbg-bb1-pos5-0-0.telia.net
2	213.248.64.21	213.248.65.89	kbn-bb1-pos2-1-0.telia.net	ldn-bb1-pos6-0-0.telia.net
2	213.248.64.109	213.248.64.42	ffm-bb1-pos6-0-0.telia.net	ffm-bb1-pos7-0-0.telia.net

Table 22: Router pairs responsible for all TTL-invisible route changes from `lava` to `globedom` (10-minute interval measurement)

Count	IP address A	IP address B	Hostname A	Hostname B
30	144.223.27.146	144.232.230.30	sl-telia1-1-0.sprintlink.net	sl-telia1-4-0.sprintlink.net
2	213.248.64.9	213.248.65.146	prs-bb1-pos6-0-0.telia.net	hbg-bb1-pos5-0-0.telia.net
2	213.248.64.21	213.248.65.89	kbn-bb1-pos2-1-0.telia.net	ldn-bb1-pos6-0-0.telia.net
2	213.248.64.109	213.248.64.42	ffm-bb1-pos6-0-0.telia.net	ffm-bb1-pos7-0-0.telia.net

Table 23: Router pairs responsible for all TTL-invisible route changes from `lava` to `globedom` (2-hour interval measurement)

Measurement time	Route id	AS path id
Wed Jul 14 02:49:54 AM	2	1
Tue Jul 20 07:20:45 PM	5	1
Wed Jul 21 03:00:47 AM	2	1
Thu Jul 22 06:41:00 PM	3	1
Fri Jul 23 11:31:06 AM	4	1
Fri Jul 23 08:31:09 PM	1	1
Sat Jul 24 06:51:16 PM	6	1
Sat Jul 24 10:21:17 PM	1	1
Sun Jul 25 06:11:19 AM	6	1
Sun Jul 25 08:41:22 AM	1	1
Mon Jul 26 05:51:31 PM	8	2
Mon Jul 26 06:01:31 PM	1	1
Mon Jul 26 07:21:31 PM	8	2
Mon Jul 26 07:51:31 PM	1	1
Tue Jul 27 08:11:40 PM	7	1
Tue Jul 27 10:51:41 PM	1	1

Table 24: Route changes for path from `lava` to `globedom` (10-min interval measurement)

Measurement time	Route id	AS path id
Wed Jul 14 11:12:20 AM	2	1
Tue Jul 20 07:29:19 PM	5	1
Wed Jul 21 06:35:38 AM	2	1
Fri Jul 23 01:48:34 AM	4	1
Fri Jul 23 05:21:01 PM	6	1
Sat Jul 24 05:29:10 AM	1	1
Sat Jul 24 09:00:28 PM	3	1
Sun Jul 25 01:08:55 PM	1	1

Table 25: Route changes for path from `lava` to `globedom` (2-hour interval measurement)

```
Time: Mon Jul 26 05:51:31 PM EDT 2004 --to-- Mon Jul 26 05:51:40 PM EDT 2004

 1  iiwi-fe-0-0-0.lava.net (64.65.64.30)  0.496 ms  0.578 ms  0.418 ms
 2  s2-6.hsa1.hon2.bbnplanet.net (4.24.134.17)  1.338 ms  0.457 ms  0.487 ms
 3  so-6-0.hsa2.Honolulu2.Level3.net (4.68.113.10)  0.580 ms  2.595 ms  2.338 ms
 4  so-0-1-0.mpls2.Honolulu2.Level3.net (4.68.112.93)  1.880 ms  1.313 ms  0.678 ms
 5  so-0-0-0.bbr2.Washington1.Level3.net (64.159.1.158)  114.331 ms  114.555 ms  114.237 ms
 6  as-0-0.mp1.London2.Level3.net (4.68.128.101)  188.196 ms  189.879 ms  187.447 ms
 7  so-3-0-0.mp2.Frankfurt1.Level3.net (212.187.128.29)  202.070 ms  201.913 ms  202.795 ms
 8  ge-11-0.ipcolo1.Frankfurt1.Level3.net (195.122.136.39)  202.545 ms  203.216 ms  204.200 ms
 9  212.162.1.131 (212.162.1.131)  232.028 ms  232.130 ms  231.653 ms
10  internic-01361-win-ix-i1.c.telia.net (213.248.79.146)  228.025 ms  227.711 ms  228.884 ms
11  fwext-129-6400.internic.at (212.126.64.19)  228.640 ms  229.858 ms  228.605 ms
12  net-65-cisco3600.internic.at (212.126.65.2)  247.478 ms  251.467 ms  238.717 ms
13  route65-3.internic.at (212.126.65.3)  240.108 ms  240.052 ms  241.507 ms
14  www.globedom.com (212.126.66.8)  231.671 ms  238.644 ms  247.040 ms


Location: lava.net --to-- www.globedom.com   Status: ok
Time: Mon Jul 26 06:01:31 PM EDT 2004 --to-- Mon Jul 26 06:01:42 PM EDT 2004

 1  iiwi-fe-0-0-0.lava.net (64.65.64.30)  0.694 ms  0.916 ms  0.454 ms
 2  noio-fe-0-0-3-16.lava.net (64.65.127.97)  0.565 ms  0.494 ms  0.478 ms
 3  sl-gw1-prl-8-0-0.sprintlink.net (160.81.200.65)  60.913 ms  35.030 ms  38.235 ms
 4  sl-bb21-prl-2-0.sprintlink.net (144.232.30.2)  1.910 ms  1.553 ms  1.374 ms
 5  sl-bb22-stk-5-3.sprintlink.net (144.232.8.34)  55.875 ms  56.093 ms  55.838 ms
 6  sl-bb26-pen-12-0.sprintlink.net (144.232.20.162)  125.022 ms  124.469 ms  245.824 ms
 7  sl-bb27-pen-15-0.sprintlink.net (144.232.16.94)  124.680 ms  124.963 ms  124.478 ms
 8  sl-bb21-nyc-2-0.sprintlink.net (144.232.20.96)  126.732 ms  127.005 ms  126.610 ms
 9  sl-bb23-nyc-3-0.sprintlink.net (144.232.7.109)  127.194 ms  126.998 ms  126.730 ms
10  sl-gw27-nyc-14-0.sprintlink.net (144.232.7.22)  126.487 ms  126.546 ms  126.453 ms
11  sl-telia1-4-0.sprintlink.net (144.232.230.30)  124.424 ms  122.629 ms  123.039 ms
12  nyk-bb2-pos0-0-0.telia.net (213.248.80.134)  124.076 ms  124.554 ms  124.304 ms
13  ldn-bb2-pos6-0-0.telia.net (213.248.65.93)  197.530 ms  196.797 ms  199.989 ms
14  prs-bb2-pos6-0-0.telia.net (213.248.65.114)  203.878 ms  203.899 ms  204.450 ms
15  ffm-bb2-pos6-0-0.telia.net (213.248.65.118)  213.609 ms  213.660 ms  213.957 ms
16  win-b1-pos5-1.telia.net (213.248.65.74)  242.758 ms  241.488 ms  242.834 ms
17  win-ix-i1-pos0-0-0.telia.net (213.248.77.150)  242.751 ms  249.337 ms  243.986 ms
18  internic-01361-win-ix-i1.c.telia.net (213.248.79.146)  235.028 ms  234.550 ms  235.197 ms
19  fwext-129-6400.internic.at (212.126.64.19)  235.954 ms  235.970 ms  236.201 ms
20  net-65-cisco3600.internic.at (212.126.65.2)  254.512 ms  246.773 ms  245.215 ms
21  route65-3.internic.at (212.126.65.3)  267.603 ms  252.695 ms  282.819 ms
22  www.globedom.com (212.126.66.8)  240.527 ms  244.280 ms  248.759 ms
```

From these three records, we see that although the site `lava.net` had shifted its primary outgoing ISP so most of the time in July this path went through `Sprintlink`, it still sometimes turned to the old `Level3` path for a short while (about 10 minutes to half an hour in the above measurement). This short diversion was unfortunately not caught by the 2-hour interval measurement.

4.9.4 Path from `www.globedom.com` to `lave.net`

The path from `www.globedom.com` to `lave.net` is the reverse path of the previous one. Since the previous one has completely changed its characteristics due to change of outgoing ISP, we first want to find out whether that is also true for the reverse path, i.e., whether `lave.net` has also switched its primary ISP for its incoming traffic. Following are two typical records taken from May and July for the path in consideration.

```
Location: www.globedom.com --to-- lava.net      Status: ok
Time: Tue May 25 12:28:13 AM EDT 2004 --to-- Tue May 25 12:28:20 AM EDT 2004

 1  fw0.globedom.com (212.126.66.1)  2.019 ms  0.169 ms  0.152 ms
 2  net-65-cisco3600.internic.at (212.126.65.2)  1.796 ms  1.414 ms  3.268 ms
 3  ciscointerxion.internic.at (212.126.65.1)  3.239 ms  2.544 ms  3.319 ms
 4  fwext-64-100M.internic.at (212.126.64.12)  7.316 ms  4.427 ms  5.064 ms
 5  win-ix-i1-feth0-1-0.telia.net (213.248.79.145)  3.43 ms  4.308 ms  17.697 ms
 6  win-b1-pos1-7.telia.net (213.248.77.149)  7.942 ms  11.904 ms  17.489 ms
```

```
 7  ffm-bb2-pos0-3-2.telia.net (213.248.65.73)  31.902 ms  44.592 ms  26.917 ms
 8  prs-bb2-pos7-0-0.telia.net (213.248.65.117)  32.846 ms  41.895 ms  42.756 ms
 9  ldn-bb2-pos7-0-0.telia.net (213.248.65.113)  49.837 ms  49.154 ms  50.112 ms
10  nyk-bb2-pos6-0-0.telia.net (213.248.65.94)  115.639 ms  123.788 ms  117.417 ms
11  nyk-bb1-pos0-0-0.telia.net (213.248.80.133)  124.43 ms  111.728 ms  124.481 ms
12  nyk-i1-pos3-0.telia.net (213.248.82.14)  124.534 ms  124.12 ms  124.507 ms
13  so-0-1-0.edge1.NewYork1.Level3.net (209.244.160.161)  120.943 ms  120.276 ms  120.958 ms
14  ge-2-1-0.bbr2.NewYork1.Level3.net (64.159.4.149)  121.106 ms  120.089 ms  120.941 ms
15  ge-0-0-0.mpls1.Honolulu2.Level3.net (4.68.128.13)  234.892 ms  197.915 ms  196.902 ms
16  so-7-0.hsa1.Honolulu2.Level3.net (4.68.112.90)  218.83 ms  234.192 ms  225.058 ms
17  s1.lavanet.bbnplanet.net (4.24.134.18)  219.423 ms  208.485 ms  234.995 ms
18  malasada.lava.net (64.65.64.17)  234.939 ms  222.995 ms  217.172 ms
```

```
Location: lava.net --to-- www.globedom.com      Status: ok
Time: Thu Jul 15 02:10:02 AM EDT 2004 --to-- Thu Jul 15 02:10:12 AM EDT 2004

 1  fw0.globedom.com (212.126.66.1)  0.167 ms  0.126 ms  0.122 ms
 2  net-65-cisco3600.internic.at (212.126.65.2)  3.496 ms  2.108 ms  3.39 ms
 3  ciscointerxion.internic.at (212.126.65.1)  3.41 ms  5.714 ms  3.397 ms
 4  fwext-64-100M.internic.at (212.126.64.12)  3.394 ms  5.708 ms  3.419 ms
 5  win-ix-i1-feth0-1-0.telia.net (213.248.79.145)  3.346 ms  5.624 ms  3.401 ms
 6  win-b1-pos1-7.telia.net (213.248.77.149)  3.471 ms  5.737 ms  3.398 ms
 7  ffm-bb2-pos0-3-2.telia.net (213.248.65.73)  31.805 ms  33.594 ms  44.344 ms
 8  prs-bb2-pos7-0-0.telia.net (213.248.65.117)  42.599 ms  42.029 ms  42.347 ms
 9  ldn-bb2-pos7-0-0.telia.net (213.248.65.113)  49.756 ms  49.208 ms  49.811 ms
10  nyk-bb2-pos6-0-0.telia.net (213.248.65.94)  124.403 ms  123.859 ms  124.572 ms
11  nyk-i1-pos3-0.telia.net (213.248.82.22)  124.505 ms  123.994 ms  124.372 ms
12  so-0-1-0.edge1.NewYork1.Level3.net (209.244.160.161)  120.844 ms  116.667 ms  117.268 ms
13  ge-2-1-0.bbr2.NewYork1.Level3.net (64.159.4.149)  117.294 ms  119.986 ms  126.248 ms
14  ge-0-0-0.mpls1.Honolulu2.Level3.net (4.68.128.13)  237.294 ms  237.166 ms  238.198 ms
15  so-7-0.hsa1.Honolulu2.Level3.net (4.68.112.90)  234.269 ms  237.44 ms  238.076 ms
16  s1.lavanet.bbnplanet.net (4.24.134.18)  252.383 ms  252.298 ms  252.298 ms
17  malasada.lava.net (64.65.64.17)  252.433 ms  252.091 ms  252.412 ms
```

These two records show that `lava.net` is still using `Level3` as its main incoming ISP, despite of its switching to `Sprintlink` for outgoing traffic in July. This testifies the asymmetric routing problem in the presence of multi-homing. The route changes observed for this path are displayed in Tables 26 and 27.

The 2-hour interval measurement missed 5 out of 13 route changes and 2 out of 6 routes compared to the 10-minute interval measurement. We show the missed route change at 11:11:06 AM EDT, Jul 23, 2004 in the following three consecutive records.

```
Location: www.globedom.com --to-- lava.net   Status: ok
Time: Fri Jul 23 11:01:05 AM EDT 2004 --to-- Fri Jul 23 11:01:11 AM EDT 2004

 1  fw0.globedom.com (212.126.66.1)  0.193 ms  0.127 ms  0.124 ms
 2  net-65-cisco3600.internic.at (212.126.65.2)  3.793 ms  2.101 ms  3.385 ms
 3  ciscointerxion.internic.at (212.126.65.1)  3.402 ms  5.733 ms  3.402 ms
 4  fwext-64-100M.internic.at (212.126.64.12)  3.399 ms  5.708 ms  3.402 ms
 5  win-ix-i1-feth0-1-0.telia.net (213.248.79.145)  96.071 ms  82.753 ms  244.348 ms
 6  win-b1-pos1-7.telia.net (213.248.77.149)  7.129 ms  6.152 ms  3.576 ms
 7  ffm-bb1-pos0-3-2.telia.net (213.248.65.69)  17.809 ms  17.144 ms  17.948 ms
 8  prs-bb1-pos7-0-0.telia.net (213.248.64.110)  31.693 ms  28.114 ms  28.553 ms
 9  ldn-bb1-pos7-2-0.telia.net (213.248.64.10)  35.578 ms  32.769 ms  35.512 ms
10  nyk-bb1-pos0-2-0.telia.net (213.248.65.90)  135.183 ms  133.74 ms  135.056 ms
11  nyk-i1-pos3-0.telia.net (213.248.82.14)  135.514 ms  134.439 ms  135.123 ms
12  so-0-1-0.edge1.NewYork1.Level3.net (209.244.160.161)  132.592 ms  122.952 ms  124.82 ms
13  ge-2-1-0.bbr2.NewYork1.Level3.net (64.159.4.149)  125.44 ms  123.864 ms  124.912 ms
14  ge-0-0-0.mpls1.Honolulu2.Level3.net (4.68.128.13)  245.226 ms  244.327 ms  245.183 ms
15  so-7-0.hsa1.Honolulu2.Level3.net (4.68.112.90)  193.342 ms  129.732 ms  241.996 ms
16  s1.lavanet.bbnplanet.net (4.24.134.18)  259.419 ms  260.009 ms  260.121 ms
17  malasada.lava.net (64.65.64.17)  259.426 ms  253.332 ms  256.897 ms
```

```
Location: www.globedom.com --to-- lava.net   Status: ok
Time: Fri Jul 23 11:11:06 AM EDT 2004 --to-- Fri Jul 23 11:11:13 AM EDT 2004
```

Measurement Time	Route id	AS path id
Wed Jul 14 02:39:55 AM	4	1
Thu Jul 15 06:10:07 PM	2	1
Tue Jul 20 06:10:45 PM	4	1
Tue Jul 20 06:40:45 PM	5	1
Tue Jul 20 07:00:45 PM	4	1
Wed Jul 21 10:40:54 PM	2	1
Fri Jul 23 11:11:06 AM	6	1
Fri Jul 23 11:21:06 AM	2	1
Fri Jul 23 11:31:06 AM	1	1
Mon Jul 26 06:11:31 PM	3	1
Tue Jul 27 07:31:40 PM	1	1
Thu Jul 29 06:11:58 PM	3	1
Sun Aug 01 09:02:18 AM	1	1
Tue Aug 03 08:12:33 AM	3	1

Table 26: Route changes for path from `globedom` to `lave` (10-minute interval measurement)

Measurement time	Route id	AS path id
Wed Jul 14 02:39:55 AM	4	1
Thu Jul 15 08:18:07 PM	2	1
Tue Jul 20 06:59:43 PM	4	1
Thu Jul 22 12:52:21 AM	2	1
Fri Jul 23 12:31:50 PM	1	1
Mon Jul 26 10:29:57 PM	3	1
Tue Jul 27 07:27:19 PM	1	1
Thu Jul 29 06:13:47 PM	3	1
Sun Aug 01 12:50:46 PM	1	1

Table 27: Route changes for path from `globedom` to `lave` (2-hour interval measurement)

```
 1  fw0.globedom.com (212.126.66.1)  0.39 ms  0.2 ms  0.126 ms
 2  net-65-cisco3600.internic.at (212.126.65.2)  4.565 ms  2.933 ms  3.561 ms
 3  ciscointerxion.internic.at (212.126.65.1)  3.57 ms  2.994 ms  3.559 ms
 4  fwext-64-100M.internic.at (212.126.64.12)  3.587 ms  2.992 ms  3.558 ms
 5  win-ix-i1-feth0-1-0.telia.net (213.248.79.145)  3.559 ms  2.99 ms  3.561 ms
 6  win-b1-pos1-7.telia.net (213.248.77.149)  3.572 ms  6.543 ms  3.557 ms
 7  ffm-bb1-pos0-3-2.telia.net (213.248.65.69)  17.785 ms  16.957 ms  17.768 ms
 8  prs-bb1-pos7-0-0.telia.net (213.248.64.110)  28.444 ms  28.326 ms  28.438 ms
 9  ldn-bb1-pos7-2-0.telia.net (213.248.64.10)  35.562 ms  34.981 ms  35.558 ms
10  nyk-bb1-pos0-2-0.telia.net (213.248.65.90)  135.532 ms  144.865 ms  135.028 ms
11  nyk-bb2-pos0-0-0.telia.net (213.248.80.134)  134.944 ms  134.407 ms  135.21 ms
12  nyk-i1-pos3-0.telia.net (213.248.82.22)  142.359 ms  141.491 ms  109.9 ms
13  so-0-1-0.edge1.NewYork1.Level3.net (209.244.160.161)  125.037 ms  123.558 ms  124.383 ms
14  ge-2-1-0.bbr2.NewYork1.Level3.net (64.159.4.149)  124.401 ms  123.629 ms  124.375 ms
15  ge-0-0-0.mpls1.Honolulu2.Level3.net (4.68.128.13)  245.573 ms  244.554 ms  245.217 ms
16  so-7-0.hsa1.Honolulu2.Level3.net (4.68.112.90)  245.212 ms  244.882 ms  245.117 ms
17  s1.lavanet.bbnplanet.net (4.24.134.18)  257.054 ms  258.663 ms  259.646 ms
18  malasada.lava.net (64.65.64.17)  259.392 ms  258.558 ms  258.608 ms


Location: www.globedom.com --to-- lava.net   Status: ok
Time: Fri Jul 23 11:21:06 AM EDT 2004 --to-- Fri Jul 23 11:21:11 AM EDT 2004

 1  fw0.globedom.com (212.126.66.1)  0.191 ms  0.128 ms  0.125 ms
 2  net-65-cisco3600.internic.at (212.126.65.2)  3.3 ms  2.886 ms  3.404 ms
 3  ciscointerxion.internic.at (212.126.65.1)  6.939 ms  3.834 ms  3.402 ms
 4  fwext-64-100M.internic.at (212.126.64.12)  3.411 ms  5.678 ms  3.522 ms
 5  win-ix-i1-feth0-1-0.telia.net (213.248.79.145)  3.499 ms  6.371 ms  3.403 ms
 6  win-b1-pos1-7.telia.net (213.248.77.149)  6.946 ms  5.722 ms  3.406 ms
 7  ffm-bb1-pos0-3-2.telia.net (213.248.65.69)  17.704 ms  30.98 ms  17.768 ms
 8  prs-bb1-pos7-0-0.telia.net (213.248.64.110)  28.42 ms  27.377 ms  28.432 ms
 9  ldn-bb1-pos7-2-0.telia.net (213.248.64.10)  35.068 ms  34.966 ms  35.556 ms
10  nyk-bb1-pos0-2-0.telia.net (213.248.65.90)  135.125 ms  134.265 ms  135.063 ms
11  nyk-i1-pos3-0.telia.net (213.248.82.14)  135.121 ms  134.49 ms  135.102 ms
12  so-0-1-0.edge1.NewYork1.Level3.net (209.244.160.161)  124.397 ms  123.854 ms  124.392 ms
13  ge-2-1-0.bbr2.NewYork1.Level3.net (64.159.4.149)  124.404 ms  123.802 ms  140.995 ms
14  ge-0-0-0.mpls1.Honolulu2.Level3.net (4.68.128.13)  244.526 ms  244.338 ms  245.346 ms
15  so-7-0.hsa1.Honolulu2.Level3.net (4.68.112.90)  151.299 ms  231.937 ms  151.797 ms
16  s1.lavanet.bbnplanet.net (4.24.134.18)  237.395 ms  258.485 ms  259.316 ms
17  malasada.lava.net (64.65.64.17)  259.433 ms  258.356 ms  255.409 ms
```

It can be seen that at 11:11:06 AM EDT, July 23, 2004, an extra router `nyk-bb2-pos0-0-0.telia.net` (213.248.80.134) (hop 11 in the second record above) was inserted into the path and caused the increment by 1 of total hop count in the path. This change appears to be local based on the hostnames of related routers. The duration of change lasts fewer than 20 minutes because the measurement taken 10 minutes later showed the extra router no longer existed in the path.

Overall, our analysis on data set IV confirms that choosing an appropriate measurement interval is very important to correctly characterize the dynamics of a path. While there are stable paths that allow a relatively longer measurement interval, many highly instable paths require a much higher measurement rate. Moreover, the presence of multi-homing may cause a path to shift from a very stable one to a very fluctuating one when different ISP is chosen for packet forwarding.

5 Route Change Detection Methods Overview

NSIS protocols install and manipulate state information in routers along the data path. In order to keep the consistency of the data path and the path on which signaling state information resides, it is important to detect a route change once it occurs and update state information in affected routers appropriately. Three categories of route change detection mechanisms are currently being discussed by the NSIS community: routing monitoring, packet monitoring and signaling probing [8].

Routing monitoring generates a trigger once a route change is determined from a routing table. The ability to deliver such a trigger depends on the routing protocol in use.

	Detection methods		Change observed	Detection module
Backward-looking	Routing monitoring	Local trigger	Routing	Routing or OS
		Extended trigger		
	Packet monitoring	C-mode signaling	TTL or interface	NTLP/NSLP
		Data	TTL or interface	NSLP
			Packet absence	
	Probing	D-mode signaling	Peer address	NTLP
Forward-looking	Routing monitoring	Local trigger	Routing	Routing or OS
		Extended trigger		
	Packet monitoring	C-mode signaling	TTL or interface	NTLP/NSLP
	Probing	D-mode signaling	Peer address	NTLP

Table 28: Summary of routing change detection methods

Most routers can provide local trigger to indicate route changes to the next hop. Some other routers maintain extended topology information of the network and are capable of providing an extended trigger for route changes happened in other parts of the network. Examples of these routers include those running link state routing protocols such as Open Shortest Path First (OSPF), and interdomain routing protocol BGP with AS_PATH information.

Packet monitoring approaches can be based on data or signaling packets. The parameters monitored can be change of TTL, interface or loss of packets all together. Generally speaking, signaling packet monitoring is easier to implement than data packet monitoring; data packet monitoring is more effective when a long signaling refresh interval has to be used due to scalability concerns; the reliability of monitoring TTL or interface change depends on the correlation between the route change and the resulting TTL or interface change, since only a subset of route changes affect the TTL or forwarding interfaces; the reliability of monitoring traffic absence is affected by the actual traffic pattern. If data packet absence is part of the normal characteristics of the traffic flow, false detection is likely to occur with the data absence approach.

Signaling probing requires the signaling transport entity to periodically re-discover its neighboring signaling peers. This approach is simple and effective, but it is always difficult to strike a balance between not generating excessive probing traffic and keeping track of all route changes with acceptable delay.

6 Backward-Looking vs. Forward-Looking Route Change Detection

From a single node's perspective, a route change could occur in its backward-looking (upstream) direction, or its forward-looking (downstream) direction. Table 28 lists the different detection methods grouped into these two categories.

Any route change generates a chain of two crossover nodes, or a divergence node and convergence node pair. In extreme cases, the sender could be the divergence node or the receiver could be the convergence node. In an example network shown in Figure 9, when the path changes from A-B-C-D-E-F to A-B-G-E-F, node B becomes the divergence node and node E the convergence node. In a network with NSIS entities (NEs) [7], we will have the concept of divergence (convergence) NEs. When the divergence (convergence) node itself is an NE, that node is also the divergence (convergence) NE. Otherwise the divergence (convergence) NE is the NE closest to the divergence (convergence) node in the upstream (downstream) direction. In Figure 9, if squares represent NEs and circles represent normal routers, we have node A as the divergence NE and node F as the convergence NE.

Excluding the signaling probing method whose usefulness is determined by the probe

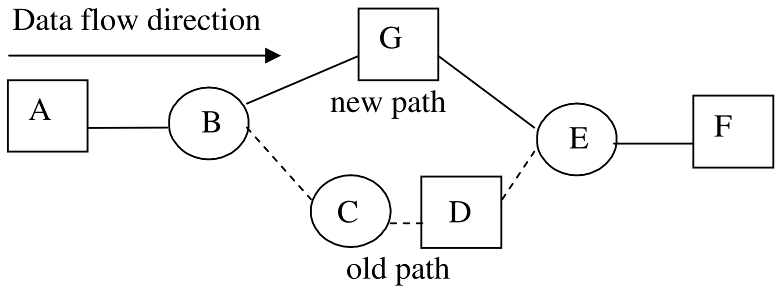


Figure 9: An intra-AS route change example

frequency, we can conclude that detecting a route change in backward-looking direction requires a convergence NE or at least one NE downstream to the convergence NE to be capable of routing or packet monitoring; detecting a route change in forward looking direction requires the divergence NE, or at least one NE upstream to the divergence NE to be capable of routing or packet monitoring.

Next we use the example topology in Figure 9 to illustrate details about when and which method would be effective in detecting a forward looking-looking or backward-looking route changes. We assume the domain is running OSPF with a common two-level hierarchy, where the whole OSPF AS is divided into several areas. Within a specific area, each router maintains an area link-state database. The detailed topology of the area is hidden from all other areas. The exchange of routing information between areas is done through area border routers, which together form an OSPF backbone area essentially running a distance vector algorithm. We also assume all nodes in Figure 9 belong to a single OSPF area and all NEs are capable of performing routing and packet monitoring where applicable.

6.1 Backward-Looking Route Change Detection

Various nodes in the network may detect the change between B and E as a backward-looking route change by the following methods.

Data packet TTL monitoring this method is effective when the two branches between the divergence and convergence nodes have different number of hops. In this example this condition holds and node F can use data packet TTL monitoring to detect the route change.

Signaling TTL monitoring in most cases, the signaling packets monitored for TTL change are refresh packets sent peer-to-peer. If the following two conditions are satisfied, then signaling TTL monitoring is applicable.

- The immediate downstream NE peer of the divergence NE remains in the flow path after a route change;
- The two branches between the divergence and convergence nodes have different number of hops;

In our example, the divergence NE, node A's immediate downstream NE peer in the old path is node D. Since D is no longer in the new path, it will not be able to detect the change based on signaling TTL monitoring. Otherwise if node D is not an NE, then node F would be the immediate downstream NE peer of node A and node F will succeed in detecting this route change by signaling TTL monitoring.

Data (signaling) packet interface monitoring this method is effective when the convergence NE is co-located with the convergence node, and the data (signaling) packets need to change incoming interfaces to the convergence node after the route change. In the example, the convergence node is not an NE although the flow's incoming interface does change. So this method is not applicable.

Data packet absence monitoring this method is effective when there are NEs in the old branch between the divergence and convergence node pair. In this example, node D is such an NE and can detect the route change based on data packet absence.

Routing monitoring this method is effective if the convergence NE has extended routing information till the divergence node and the affected area. In this example, the convergence NE F is supposed to detect the route change and deliver an extended trigger.

6.2 Forward-Looking Route Change Detection

Various nodes in the network may detect the change between B and E as a forward-looking route change by the following methods.

Data packet TTL monitoring this method is not applicable to forward-looking route change detection.

Signaling TTL monitoring in most cases, the signaling packets monitored for TTL change are refresh packets sent peer-to-peer. If the following two conditions are satisfied, then signaling TTL monitoring is effective.

- The immediate upstream NE peer of the convergence NE remains in the flow path after a route change;
- The two branches between the divergence and the convergence nodes have different number of hops.

In our example, the convergence NE, node F's immediate upstream NE peer in the old path is node D. Since node D is not in the new path, node F cannot detect the change based on signaling TTL monitoring. Otherwise if node D is not an NE, then node A would be the immediate upstream NE peer of node F and node A will succeed in detecting this route change by signaling TTL monitoring.

Signaling (data) packet interface monitoring data packet interface monitoring is not applicable. For the signaling packet interface monitoring to work, the divergence NE needs to be co-located with the divergence node, and the signaling packet need to change its outgoing interfaces from the divergence node after the route change. In the example, the divergence node is not an NE although the flow outgoing interface does change. So this method is not applicable.

Routing monitoring If the divergence node is co-located with the divergence NE, local trigger is applicable. Otherwise, extended trigger is effective when the divergence NE has routing information till the divergence node and the affected area. In our example, the divergence node B is not an NE and unable to provide local trigger. Since we assume all nodes are within an OSPF area, node A, which is the divergence NE, knows the route change occurred at node B and the affected links B-C and B-G. Node A is thus able to provide an extended trigger for the route change.

6.3 Recommendations to Enhance Route Change Detectability in Both Directions

The example above explained the detectability of an OSPF intra-area route change. We can similarly examine other types of route changes, such as OSPF inter-area, Routing Information Protocol (RIP) intra-AS, BGP inter-AS route changes. A common conclusion after analyzing all these possible route changes is that, to increase the chance of detecting route changes, it is recommended to have NEs equipped with route change detection mechanisms at least in all "special" routers, namely, AS border routers, OSPF area border routers, OSPF backbone area routers.

The reasoning to the above conclusion is not hard to understand. By doing so we are sort of separating the whole path into several independent (possibly hierarchical) NSIS network domains, each being taken care of by at least its ingress and egress nodes. In such a path, route changes within any of these domains stand a good chance to be detected by either a downstream NE or an upstream NE or both.

7 NSIS-Affecting Route Changes and NSIS Deployment Models

In all discussions so far we have not explicitly differentiated a generic route change from a route change that NSIS should indeed worry about. From a network perspective, all generic route changes can be classified into the different scales including inter-AS and intra-AS route changes. Intra-AS route changes may further be grouped into ingress-point, egress-point and mid-point route changes, depending on the position of the router in the AS where the route change occurred. We need to consider all inter-AS and intra-AS route changes only in a full NSIS network where all nodes are NEs. In a realistic scenario we are much more likely to have a mixed deployments of NEs along with normal routers. In that case, the actual route changes NSIS protocol needs to deal with are only a subset of all generic route changes. We call these route changes as NSIS-affecting route changes. The common property of all NSIS-affecting route changes is that all of them involve NE level changes. The route change shown in Figure 9 is an example of an NSIS-affecting route change.

To better understand NSIS-affecting route changes, we need to make assumption about NSIS deployment models. We list below several most likely deployment models in a mixed environment of NEs and normal routers, together with the respective NSIS concerned route changes in each model.

AS model in AS model each AS deploys a central NE that is responsible for NSIS related signaling for this AS. The NSIS-affecting route changes in this model include all route changes that involve AS changes, i.e., inter-AS route changes.

Entry model in entry model the ingress routers of each AS are NEs. NSIS-affecting route changes in this model include inter-AS and intra-AS ingress point route changes.

Border model in border model both ingress and egress routers of each AS are NEs. NSIS specific route changes in this model include inter-AS, intra-AS ingress point and intra-AS egress point changes.

Edge model in edge model the access routers of the source and destination sites are NEs. NSIS-affecting route changes in this model include inter-AS route change involving the first or last AS, intra-AS ingress point route change of the first AS as well as intra-AS egress point route change of the last AS.

Generic model the generic model is defined as any mixed models that cannot be clearly mapped to any of the above four categories. NSIS-affecting route changes in this model can not be broken down to the basic route change categories of different scales mentioned above. It is necessary to observe whether change of NEs is involved in a route change to identify whether the change is NSIS-affecting. Our example in Figure 9 used this model.

8 Evaluation of TTL Monitoring Based Route Change Detection

Having discussed general aspects of common route change detection methods, network route change detectability and introduced NSIS deployment models, we would like to provide some experimental evaluation of certain detection methods. Evaluation of routing monitoring requires a more complicated testbed, so our current focus is on packet monitoring mechanisms. In particular, we looked at the simple and low cost data TTL monitoring approach based on our routing measurement results.

TTL monitoring approach detects a change in route by the change in packet TTL, or the number of hops in the path. We start by assuming a full NSIS deployment model where all route changes count. Table 29 shows the percentage of TTL-visible route changes and AS path changes in our measurement data sets I, II and III.

The percentage of overall TTL-visible route changes does not seem very promising. However, even in an all NE network, the route changes that concern us most are always

Item description	DS I	DS II	DS III
TTL-visible route changes	38%	25%	23%
AS path change / route change	18%	8%	8%
TTL-visible AS path changes	77%	83%	88%

Table 29: TTL-visible route changes and AS path changes

Model type	DS I	DS II	DS III
AS model	77%	83%	88%
Entry model	51%	41%	40%
Border model	45%	39%	38%
Edge model	74%	90%	92%

Table 30: Effectiveness of TTL detection in four mixed NSIS deployment models

non-trivial ones. If we focus on AS level changes, we can see that TTL did a fairly good job in detecting roughly 4 out of 5 such changes.

On the other hand, we have seen using router merging (details in 4.5.2) we can easily reduce the number of TTL-invisible route changes by 52% and 70% in DS I and DS II respectively. This indicates that the majority of TTL-invisible route changes are trivial and less relevant.

In addition to the all NE model, we also examined the effectiveness of TTL detection mechanism in four mixed NSIS deployment models. The results are shown in Table 30.

The table shows that the TTL detection mechanism works better in these mixed models than in the all NE model. This is because most NSIS-affecting route changes in the mixed models are non-trivial changes, where TTL detection tends to be more effective. It can also be seen that the TTL detection ratio tends to be higher when the deployment of NEs is more sparse, in which case fewer route changes become NSIS-affecting. Therefore, the edge model and the AS model has higher detection ratio than the entry model and the border model. It is not possible to check the effectiveness of TTL detection mechanism in a generic mixed model when the detailed deployment information is unknown, but the results are expected to fall between the full NE model and the four mixed models we have studied.

Overall, the TTL monitoring detection method is still a reasonably good way to detect non-trivial route changes in all above NSIS deployment models, although TTL monitoring detection alone is clearly not enough.

9 Route Change Detection Based on Delay Variation

In this section we propose a new packet monitoring route change detection mechanism which is not yet mentioned in current documents of the IETF NSIS working group. Generally speaking, when packets for a given path takes a different route, they will experience different delays in the path. This variation in one-way-delay may be treated as an indication of possible route change.

We evaluated the delay variation route change detection method by data obtained from the RIPE NCC's Test Traffic Measurements Service (TTM) [32]. TTM measures key parameters of the network connectivity, in particular routing vectors, one-way-delay, packet-loss, and bandwidth, among about a hundred Internet sites around the world. The parameters of specific interest to us are the one-way-delay and routing measurements. One-way-delays are measured on a Poisson distributed interval at 30 seconds average, and routes are gathered randomly by `traceroute` about 6-10 times per hour. Our assessment used data between September 1 to September 7, 2004 for the eight source and destination pairs listed in Table 31.

Pair id	Site A (source/destination)	Site B (source/destination)
1	INFN, Trieste, IT	LINX, London, UK
2	RIPE NCC Amsterdam, NL	CERN at Starlight, Chicago, IL
3	Commercial ISP A, NJ	Commercial ISP B, CA
4	XO Comms. Inc., VA	Internet Initiative Inc., Yokohama, JP
5	Grnet, Athens, GR	University Waikato, Hamilton, NZ
6	CERN, Geneva, CH	Univ. of Tokyo, Tokyo, JP
7	Univ. of Melbourne, Melbourne, AU	FUNET, Espoo, FI
8	Univ. of Waikato, Hamilton, NZ	Internet2/UCAID, Ann Arbor, MI

Table 31: Source and destination pairs in evaluating the delay variation detection method

In our preliminary evaluation, we used a simple algorithm with fixed threshold to determine whether a route change has occurred based on delay variation. At the beginning, the path starts with an initial minimum delay, which is also the initial value of current delay. Then we look at delays in a sliding window containing the next six consecutive samples. The sliding window moves to the right as new samples become available. A route change is declared if six consecutive samples in the whole window all exhibit a delay smaller (if the delay is decreasing) or larger (if the delay is increasing) than the current delay by at least three milliseconds. When this happens, the current delay is also updated with the minimum delay of the current sliding window. The values of the threshold, namely six samples of window size and three millisecond difference are verified manually as being appropriate in our data sets to effectively find route changes while minimizing false positive detections. We plan to develop an adaptive mechanism to choose these values during the next step.

Compared with the TTL monitoring method, delay variation detection is also more likely to catch non-trivial route changes, where the magnitude of delay variation is expected to be larger. There will be overlaps in route changes detected by the two mechanisms. However, since the two methods are monitoring completely different parameters (hop count vs. delay), either of them stands a chance to detect some route changes that the other one would fail. Therefore they could be used as two complementary techniques to increase the overall chance of route change detection. Figure 10 illustrates some typical examples showing how delay variation detection works and how route changes detected by delay variation may or may not be visible to TTL monitoring. The dots in the figure represent samples of one-way-delays and the solid line represents the TTL of the path.

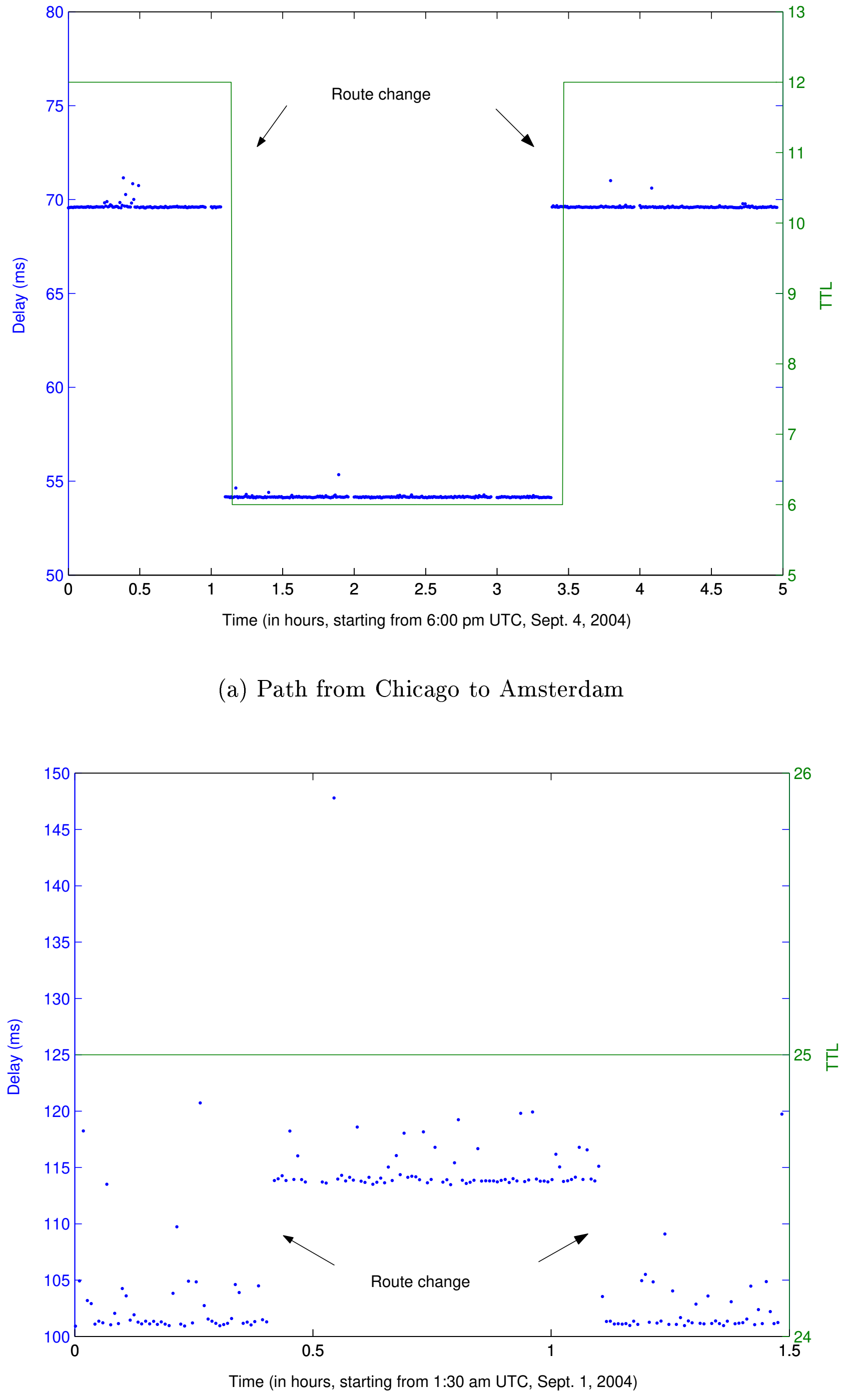
Figure 10(a) shows the path from CERN at Starlight, Chicago, IL to RIPE NCC Amsterdam, NL and the time from 6:00pm UTC to 11:00pm UTC on September 4, 2004. The corresponding routing records obtained by `traceroute` are as follows:

```
From  2004/09/03 19:58:01 To  2004/09/04 18:54:23

  1 192.91.246.129 ar5-chicago-vlan9.cern.ch
  2 192.65.184.54 cernh5-so-000.cern.ch
  3 192.65.192.7 cernh7-vlan2.cern.ch
  4 192.65.184.173 swice3-10ge-1-2.switch.ch
  5 130.59.36.5 swiCE2-G3-3.switch.ch
  6 62.40.103.17 switch.ch1.ch.geant.net
  7 62.40.96.30 ch.fr1.fr.geant.net
  8 62.40.96.90 fr.uk1.uk.geant.net
  9 62.40.96.181 uk.nl1.nl.geant.net
  10 62.40.103.98 surfnet-gw.nl1.nl.geant.net
  11 195.69.144.68 Hostname not found
  12 193.0.0.228 tt01.ripe.net

From  2004/09/04 19:10:30 To  2004/09/04 21:20:39 3944542 28

  1 192.91.246.129 ar5-chicago-vlan9.cern.ch
```



(a) Path from Chicago to Amsterdam

(b) Path from New Zealand to Michigan

Figure 10: Delay variation and TTL based route change detection

```
  2 192.91.246.125 abilene-cern-s1.cern.ch
  3 198.32.8.83 nycmng-chinng.abilene.ucaid.edu
  4 unknown  no response
  (Removed) 130.59.36.5 swiCE2-G3-3.switch.ch
  (Removed) 62.40.103.17 switch.ch1.ch.geant.net
  (Removed) 62.40.96.30 ch.fr1.fr.geant.net
  (Removed) 62.40.96.90 fr.uk1.uk.geant.net
  (Removed) 62.40.96.181 uk.nl1.nl.geant.net
  (Removed) 62.40.103.98 surfnet-gw.nl1.nl.geant.net
  5 195.69.144.68 Hostname not found
  6 193.0.0.228 tt01.ripe.net

From  2004/09/04 21:29:20 To  2004/09/09 16:10:46

  1 192.91.246.129 ar5-chicago-vlan9.cern.ch
  2 192.65.184.54 cernh5-so-000.cern.ch
  3 192.65.192.7 cernh7-vlan2.cern.ch
  4 192.65.184.173 swice3-10ge-1-2.switch.ch
  5 130.59.36.5 swiCE2-G3-3.switch.ch
  6 62.40.103.17 switch.ch1.ch.geant.net
  7 62.40.96.30 ch.fr1.fr.geant.net
  8 62.40.96.90 fr.uk1.uk.geant.net
  9 62.40.96.181 uk.nl1.nl.geant.net
  10 62.40.103.98 surfnet-gw.nl1.nl.geant.net
  11 195.69.144.68 Hostname not found
  12 193.0.0.228 tt01.ripe.net
```

The second record above showed a route change that caused the TTL to decrease from 12 to 6, and the new route lasted about 2 hours. The third record above showed a route change back to the original 12-hop one and remained unchanged afterwards. Both transitions are exactly captured in Figure 10(a) by the solid, TTL detection line. On the other hand, we can clearly see from the dotted data points that the one-way-delay of the path dropped by 15 ms when the route switched to a new one containing only half of the original number of hops. The delay went up by the same amount when the route came back to the original one. Note that a careful inspection of this figure also shows that the change in one-way-delay seemed to happen ahead of the change of TTL, by about 3 minutes. Similar situation was also found in a few other paths. There could be two possible explanations. First, there is only one route change. TTL and delay monitoring methods were indeed observing the same route change; second, there were actually two route changes and an additional TTL-invisible route change happened 3 minutes before the TTL-visible route change. If the former is true, then both methods are equally good in detecting route changes, but probably with different response time; if the latter is true, then either method will miss one of the two route changes. The combination of both methods is necessary to give the correct number of route changes. Unfortunately, due to the limited sampling rate in the measurement, we are not able to confirm which one is true at this time. Our current assumption is the former, i.e. treating the two methods as declaring the same route change.

Figure 10(b) shows the path from University of Waikato, Hamilton, NZ to Internet2/UCAID, Ann Arbor, Michigan, between 1:30 AM UTC and 3:00 AM UTC, September 1, 2004. The routing records around this period are shown below:

```
From  2004/09/01 00:21:17 To  2004/09/01 01:53:34

  1 192.107.171.142 gtw.wand.net.nz
  2 192.107.171.49 192-107-171-49.network.waikato.ac.nz
  3 130.217.2.6 pegasus.network.waikato.ac.nz
  4 203.167.234.85 Hostname not found
  5 218.101.61.14 jcore2-ge-0-2-0-927-acld.auckland.clix.net.nz
  6 203.98.50.251 ge-0-2-0-1.xcore1.acld.telstraclear.net
  7 203.98.42.67 jb1.wils.telstraclear.net
  8 157.130.245.21 GigabitEthernet5-2.GW1.LAX1.ALTER.NET
  9 152.63.112.214 0.so-0-0-0.CL2.LAX1.ALTER.NET
  10 152.63.115.150 0.so-1-0-0.TL2.LAX9.ALTER.NET
  11 152.63.15.250 0.so-1-0-0.TL2.SAC1.ALTER.NET
```

```
  12 152.63.54.9 0.so-7-0-0.XL2.SAC1.ALTER.NET
  13 152.63.52.229 POS7-0.BR5.SAC1.ALTER.NET
  14 204.255.174.210 Hostname not found
  15 154.54.2.221 p6-0.core01.sjc01.atlas.cogentco.com
  16 66.28.4.93 p4-0.core01.sfo01.atlas.cogentco.com
  17 66.28.4.70 p15-0.core02.sfo01.atlas.cogentco.com
  18 66.28.4.129 p15-0.core01.den01.atlas.cogentco.com
  19 66.28.4.30 p5-0.core01.mci01.atlas.cogentco.com
  20 66.28.4.34 p5-0.core02.ord01.atlas.cogentco.com
  21 66.28.64.194 g6.ba21.b002281-1.ord01.atlas.cogentco.com
  22 38.112.7.10 Merit.demarc.cogentco.com
  23 198.108.23.253 ge-0-1-0x24.aa1.mich.net
  24 198.108.90.118 so-0-0-0x0.ucaid2.mich.net
  25 207.75.164.254 ripe.internet2.edu

From  2004/09/01 01:55:37 To  2004/09/01 02:13:25

  1 192.107.171.142 gtw.wand.net.nz
  2 192.107.171.49 192-107-171-49.network.waikato.ac.nz
  3 130.217.2.6 pegasus.network.waikato.ac.nz
  4 203.167.234.85 Hostname not found
  5 218.101.61.14 jcore2-ge-0-2-0-927-acld.auckland.clix.net.nz
  6 203.98.50.251 ge-0-2-0-1.xcore1.acld.telstraclear.net
  7 203.98.42.67 jb1.wils.telstraclear.net
  8 157.130.245.21 GigabitEthernet5-2.GW1.LAX1.ALTER.NET
  9 152.63.112.214 0.so-0-0-0.CL2.LAX1.ALTER.NET
  10 152.63.115.150 0.so-1-0-0.TL2.LAX9.ALTER.NET
  11 152.63.15.250 0.so-1-0-0.TL2.SAC1.ALTER.NET
  12 152.63.54.9 0.so-7-0-0.XL2.SAC1.ALTER.NET
  13 152.63.52.229 POS7-0.BR5.SAC1.ALTER.NET
  14 204.255.174.210 Hostname not found
  15 154.54.2.221 p6-0.core01.sjc01.atlas.cogentco.com
  16 66.28.4.93 p4-0.core01.sfo01.atlas.cogentco.com
  17 66.28.4.70 p15-0.core02.sfo01.atlas.cogentco.com
  18 66.28.4.133 p14-0.core02.dfw01.atlas.cogentco.com
  19 66.28.4.38 p15-0.core01.mci01.atlas.cogentco.com
  20 66.28.4.34 p5-0.core02.ord01.atlas.cogentco.com
  21 66.28.64.194 g6.ba21.b002281-1.ord01.atlas.cogentco.com
  22 38.112.7.10 Merit.demarc.cogentco.com
  23 198.108.23.253 ge-0-1-0x24.aa1.mich.net
  24 198.108.90.118 so-0-0-0x0.ucaid2.mich.net
  25 207.75.164.254 ripe.internet2.edu

From  2004/09/01 02:42:27 To  2004/09/01 03:20:37

  1 192.107.171.142 gtw.wand.net.nz
  2 192.107.171.49 192-107-171-49.network.waikato.ac.nz
  3 130.217.2.6 pegasus.network.waikato.ac.nz
  4 203.167.234.85 Hostname not found
  5 218.101.61.14 jcore2-ge-0-2-0-927-acld.auckland.clix.net.nz
  6 203.98.50.251 ge-0-2-0-1.xcore1.acld.telstraclear.net
  7 203.98.42.67 jb1.wils.telstraclear.net
  8 157.130.245.21 GigabitEthernet5-2.GW1.LAX1.ALTER.NET
  9 152.63.112.214 0.so-0-0-0.CL2.LAX1.ALTER.NET
  10 152.63.115.150 0.so-1-0-0.TL2.LAX9.ALTER.NET
  11 152.63.15.250 0.so-1-0-0.TL2.SAC1.ALTER.NET
  12 152.63.54.9 0.so-7-0-0.XL2.SAC1.ALTER.NET
  13 152.63.52.229 POS7-0.BR5.SAC1.ALTER.NET
  14 204.255.174.210 Hostname not found
  15 154.54.2.221 p6-0.core01.sjc01.atlas.cogentco.com
  16 66.28.4.93 p4-0.core01.sfo01.atlas.cogentco.com
  17 66.28.4.70 p15-0.core02.sfo01.atlas.cogentco.com
  18 66.28.4.129 p15-0.core01.den01.atlas.cogentco.com
  19 66.28.4.30 p5-0.core01.mci01.atlas.cogentco.com
  20 66.28.4.34 p5-0.core02.ord01.atlas.cogentco.com
  21 66.28.64.194 g6.ba21.b002281-1.ord01.atlas.cogentco.com
  22 38.112.7.10 Merit.demarc.cogentco.com
  23 198.108.23.253 ge-0-1-0x24.aa1.mich.net
  24 198.108.90.118 so-0-0-0x0.ucaid2.mich.net
  25 207.75.164.254 ripe.internet2.edu
```

Pair id	TTL-visible	Delay variation detected	TTL-invisible	Delay variation detected
1	4	3	2	2
1R	2	1	2	2
2	2	2	1	0
2R	2	2	2	0
3	0	0	1	0
3R	2	2	3	0
4	13	3	3	3
4R	9	3	3	3
5	2	0	0	0
5R	3	0	9	7
6	0	0	1	1
6R	0	0	4	2
7	0	0	2	0
7R	0	0	0	0
8	1	0	4	3
8R	1	0	0	0
Total	41	16	37	23

Table 32: Delay variation detection of TTL-visible and TTL-invisible route changes

The second record above showed a route change at 1:55 AM UTC, which was exactly reflected in Figure 10(b) by a 12 ms delay increase. The delay statistics shows that the route changed again, probably to the original one, at around 2:13 AM UTC, resulting in a 12 ms delay decrease. The second record above confirmed that the new route ends at 2:13 AM UTC. The third record above confirmed that the route returned to the original one (as of the first record) after the second change. Note that the route from 2:13:25 AM UTC to 2:42:27 AM UTC is not known due to sampling frequency limitation. But the three records combined together give a coherent explanation of the boundary of the route changes detected by delay variation. Clearly the TTL value remained at 25 all the time so route changes in this figure cannot be detected by TTL monitoring.

Next we list in Table 32 the effectiveness of the delay variation mechanism in detecting TTL-visible and TTL-invisible route changes for all eight source and destination pairs. The "Pair ids" in the first column correspond to the same pair ids in Table 31. The letter "R" means reverse direction of the specific pair.

Table 32 showed that there are roughly half TTL-invisible and half TTL-visible route changes. Among the TTL-visible route changes, over one third of them appeared to be detectable by delay variation. This is not particularly interesting because TTL monitoring can detect all route change in this category anyway. More important here are route changes that are not visible to the TTL mechanism. From our data set we see that over 60% of TTL-invisible route changes manifest themselves as delay variation visible. This shows that delay variation monitoring may be a promising route change detection mechanism complementing the TTL method.

10 Conclusions and Future work

In order to understand the current Internet routing dynamics to support the design of NSIS and IP routing interaction, we conducted a routing measurement between April and August 2004. We compared our results with two previous measurements dated 1994 and 1999 [20, 22] and found that the overall routing has neither improved nor degraded significantly since the previous measurements. One important aspect that all measurements have shown is that the Internet paths are strongly dominated by a single route, but there exist significant site to site variations. Therefore the NSIS routing interaction mechanism is highly recommended to adopt an adaptive approach to cope with these variations.

Our route change statistics shows that route changes occur over a wide range of time scales, ranging from seconds to days, and over different network scales, ranging from intra-AS local changes to inter-AS level changes. The majority of route changes are found to be involving no change of total hop counts (i.e., TTL-invisible). However, a large proportion of these TTL-invisible route changes are caused only by a small set of (mostly local) router pairs belonging to the same service provider or host network, which are probably doing load splitting or load balancing. The router administrators have the best information on how these routers are configured and therefore are the best persons to take care of these route changes. Load splitting is a cause to extremely short route changes at the scale of seconds or minutes [20] and is still found to be common today. This behavior is controlled by a special option of an IPv4 router [31] and may be turned off. Load balancing usually lead to short or medium scale route changes and are seen in both host and service provider routers. Extra mechanisms need to be deployed in the networks to allow NSIS to function properly in the presence of load balancing. These mechanisms may include special route monitoring to quickly detect the route change and notify related NSIS modules about the change, or establishing redundant state information in balancing routers. The latter approach tend to cost more resources but renders a more prompt response time.

Our additional short interval routing measurement confirmed that the accuracy of measuring path characteristics using a longer measurement interval depends heavily on the stability of the path being studied. Comparing a 10-minute fixed interval and a 2-hour exponential interval measurement, we have seen examples of stable paths where both measurements capture the same number of routes, AS paths and their changes; we also have examples of less stable paths where the 2-hour measurement missed about half the number of routes, AS paths and their changes. Therefore, an ideal mechanism to estimate path characteristics might be to start with a relatively high rate of measurement, and reduce the rate accordingly if the path is found to be stable.

Multi-homing is often seen in our measurements and causes AS-level route changes. We saw examples of a site using one ISP as its primary outgoing ISP in May and June and then switch to another one as its primary outgoing ISP in July. This switching happened to cause a dramatic change in path characteristics, from a very stable path to a very fluctuating path. Moreover, the site still occasionally used the old ISP as outgoing ISP for a while, causing short term AS-level route changes that cannot be overlooked. The site is also using one ISP as the incoming ISP and another as the outgoing ISP, making the path apparently asymmetric. If the multi-homed site has control over these behaviors, it may deploy appropriate mechanisms to notify related NSIS entities about any multi-homing caused switching to allow a fast NSIS recovery.

After analyzing the measurement data, we discussed common route change detection methods and concluded that detecting a route change in backward-looking direction requires a convergence NE or at least one NE downstream to the convergence NE to be capable of routing or packet monitoring; detecting a route change in forward looking direction requires the divergence NE, or at least one NE upstream to the divergence NE to be capable of routing or packet monitoring.

We also introduced typical NSIS deployment models including AS model, entry model, border model and edge model and associated with each of them the NSIS-affecting route changes. Our subsequent evaluation of TTL monitoring approach found that, while simple and low cost, the TTL monitoring approach seemed to be reasonably good especially in detecting non-trivial route changes and in some mixed NSIS deployment models. But we also conclude that the TTL monitoring approach alone is not enough, and proposed a new route change detection mechanism based on delay variation monitoring. Our preliminary evaluation showed that it is a promising mechanism capable of capturing over half route changes missed by the TTL method.

Possible future work of this study include: effectively estimate path characteristics using an adaptive algorithm; given the path characteristics, find proper methods to configure the NSIS protocol parameters, such as refresh intervals; set up a routing monitoring testbed and evaluate the applicability of routing monitoring based route change detection;

develop a rigorous algorithm for the delay variation based route change detection; specify detailed NSIS recovery procedures after a route change is detected.

11 Acknowledgement

We would like to thank all `traceroute` sites that gave us permission to run our experiment. Some of their site administrators kindly shared their opinions on routing pathologies related to their sites. We would also like to thank the RIPE NCC Test Traffic Measurement Service for providing related routing data and Henk Uijterwaal for his comments.

References

[1] R. Braden, L. Zhang, S. Berson, S. Herzog, and S. Jamin, "Resource ReSerVation Protocol (RSVP) – Version 1 Functional Specification," *RFC 2205*, September 1997.

[2] M. Shore, "The TIST (Topology-Insensitive Service Traversal)," *Internet Draft, Internet Engineering Task Force*, May 2002, Work in progress.

[3] D. Awduche, L. Berger, D. Gan, and et. al., "RSVP-TE: Extensions to RSVP for LSP Tunnels," *RFC 3209*, December 2001.

[4] L. Berger, Editor, "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions," *RFC 3473*, January 2003.

[5] C. Shen, W. Seah, A. Lo, H. Zheng, and M. Greis, "Mobility Extensions to RSVP in an RSVP-Mobile IPv6 Framework," *Internet Draft, Internet Engineering Task Force*, 2003, Work in progress.

[6] H. Schulzrinne, X. Fu, C. Pampu, and C. Kappler, "Design of CASP - a Technology Independent Lightweight Signaling Protocol," in *Proc. of the 1st International Workshop on Inter-domain Performance and Simulation (IPS 2003)*, February 2003.

[7] R. Hancock, I. Freytsis, G. Karagiannis, J. Loughney, and S. Van den Bosch, "Next Steps in Signaling: Framework," *Internet Draft, Internet Engineering Task Force*, 2004, Work in progress.

[8] H. Schulzrinne and R. Hancock, "GIMPS: General Internet Messaging Protocol for Signaling ," *Internet Draft, Internet Engineering Task Force*, 2004, Work in progress.

[9] V. Paxson, "Strategies for Sound Internet Measurement," *Proc. of ACM Internet Measurement Conference*, October 2004.

[10] P. Barford, A. Bestavros, J. Byers, and M. Crovella, "An Architecture for Large-Scale Internet Measurement," in *Proc. of the ACM SIGCOMM Internet Measurement Workshop*, November 2001.

[11] Z. Wang, A. Zeitoun, and S. Jamin, "Challenges and Lessons Learned in Measuring Path RTT for Proximity-based Applications," in *Proc. of Passive and Active Measurement Workshop (PAM'03)*, San Diego, CA, April 2003.

[12] F. Wang and L. Gao, "Inferring and Characterizing Internet Routing Policies," in *Proc. of the ACM SIGCOMM Internet Measurement Workshop*, 2003.

[13] R. Mahajan, "Understanding BGP Misconfiguration," in *Proc. of ACM SIGCOMM*, August 2002.

[14] C. Labovitz, A. Ahuja, and A. Bose, "Delayed Internet Routing Convergence," in *Proc. of the conference on Applications, Technologies, Architectures, and Protocols for Computer Communications*, June 1999.

[15] Craig Labovitz, G. Robert Malan, and Farnam Jahanian, "Internet Routing Instability," *IEEE/ACM Transactions on Networking*, vol. 6, no. 5, pp. 515–528, 1998.

[16] R. Teixeira, A. Shaikh, T. Griffin, and J. Rexford, "Dynamics of Hot-Potato Routing in IP Networks," in *Proc. of ACM SIGMETRICS/Performance*, San Diego, CA, June 2004.

[17] V. Paxson, "End-to-end Internet Packet Dynamics," *IEEE/ACM Transactions on Networking*, vol. 7, no. 3, pp. 277–292, 1999.

[18] V. Paxson, "End-to-End Internet Packet Dynamics," in *Proc. of ACM SIGCOMM '97*, June 1997.

[19] V. Paxson, "End-to-end Routing Behavior in the Internet," *IEEE/ACM Transactions on Networking*, vol. 5, no. 5, pp. 601–615, 1997.

[20] V. Paxson, *Measurements and Analysis of End-to-End Internet Dynamics*, Ph.D. thesis, University of California, Berkeley, 1997.

[21] V. Paxson, A. Adams, and M. Mathis, "Experiences with NIMI," in *Proc. of Passive & Active Measurement*, 2000.

[22] Y. Zhang, *Characterizing End-to-End Internet Performance*, Ph.D. thesis, Cornell University, 2001.

[23] Y. Zhang, N. Duffield, V. Paxson, and S. Shenker, "On the Constancy of Internet Path Protperties," in *Proc. of ACM SIGCOMM Internet Measurement Workshop*, November 2001.

[24] Y. Zhang, V. Paxson, and S. Shenker, "The Stationary of Internet Path Properties," Tech. Rep., ACIRI, May 2000.

[25] Van Jacobson, "LBNL's Network Research Group," http://www-nrg.ee.lbl.gov/.

[26] L. Amini, *Models and algorithms for resource management in distributed computing cooperatives*, Ph.D. thesis, Columbia University, 2004.

[27] R. Wolff, "Poisson Arrivals See Time Averages," *Operations Research*, vol. 30, no. 2, pp. 223–231, 1982.

[28] Routing Assets Database, "http://www.radb.net," .

[29] Level 3 Looking Glass, "http://www.level3.com/LookingGlass," .

[30] AT&T Looking Glass, "http://route-server.ip.att.net," .

[31] F. Baker, Editor, "Requirements for IP Version 4 Routers," *RFC 1812*, June 1995.

[32] RIPE NCC Test Traffic Measurements, "http://www.ripe.net/ttm," .