Watching Data Streams Toward a Multi-Homed Sink Under Routing Changes Introduced by a BGP Beacon

Jun Li, Randy Bush, Zhuoqing Mao, Timothy Griffin, Matthew Roughan, Daniel Stutzbach, Eric Purpus*

Abstract. Certain popular metrics have been used to reflect the instability of the Internet control plane, such as the volume and duration of BGP updates. However, it is unclear whether there is a direct relationship between those metrics and the data plane performance, especially as the Internet is becoming more densely connected and many networks become multi-homed. To clarify this, we measured data streams from a number of PlanetLab nodes toward a sink behind a multi-homed BGP Beacon, which can introduce scheduled BGP routing changes to potentially affect the performance of those data streams. In particular, as an important first step, we measured the delay, drop, jitter, and reordering of these data streams and compare them against the volume and duration of BGP updates. We found these data streams were only slightly affected and there is little correlations between these selected metrics from the two planes. Further work includes the correlation of the data plane performance with other control plane metrics while considering more types of routing changes.

Keywords: control plane, data plane, routing, BGP, Internet performance, network measurement

1 Introduction

Internet health should be judged by the performance of its data plane, as this is what affects its users. Meanwhile, the Internet control plane constantly experiences certain instabilities due to routing changes. The question is, while the ultimate goal of routing changes is to discover the best data delivery paths—thus improving data plane performance, is it possible that routing changes may actually cause a certain level of disruptions at the control plane and further degrade the data plane performance?

Certain studies have been conducted to address some aspects of the relationship between the two planes. For example, Labovitz et al. found that during

* This research is funded by NSF award ANI-0221435. Jun Li is with University of Oregon; Randy Bush is with IIJ; Zhuoqing Mao is with University of Michigan; Timothy Griffin is with Cambridge University; Matthew Roughan is with University of Adelaide; Daniel Stutzbach is with University of Oregon; Eric Purpus is with University of Oregon. Point of contact: lijun@cs.uoregon.edu.

delayed BGP convergence, end-to-end Internet paths will experience increased packet loss and latency [1]. Feamster et al. found that BGP instability correlates well with path faults that appear in the network core. Research also found that most BGP updates are about prefixes that do not receive much traffic [2, 3].

Nonetheless, excluding studies above, much research has mostly focused on each of the two planes separately. Among the many questions concerning the two-plane relationship, the following must still be answered: (1) If a site has a richer connectivity, such as multi-homed, will the data delivery performance toward this site vary or degrade when routes toward this site change? This question is becoming more important as the Internet becomes more densely connected and more networks become multi-homed. (2) Is there a correlation between the data plane and the control plane in terms of certain metrics from the two planes?

In this paper, we conduct a measurement study to answer these questions. In particular, at the control plane, we focus on BGP, the *de facto* standard inter-domain routing protocol. We set up a BGP Beacon [4] that can introduce specific routing changes according to a configured schedule. The Beacon is multi-homed and it is connected to two large global providers. At the data plane, we generated and measured data streams from PlanetLab[5] nodes toward a sink behind the BGP Beacon. Note this sink is thus also multi-homed. To address Question (1) above, we then measured the performance of received packet streams in the presence of the routing changes introduced by the Beacon.

There are well-established metrics for performance of the data plane—*delay, drop, jitter,* and *reordering* (DDJ&R)—of packet delivery toward a destination. These are also metrics that users really care about! On the other hand, to measure the instabilities of the control plane, the volume and duration of routing updates have been widely used. (There are also other control plane metrics, but we study these two first in this paper.) Related to Question (2) above, we will investigate whether there is a correlation between the data plane and the control plane in the above context in terms of these metrics about the two planes.

Our study has found that during those introduced routing changes, packet delivery is only degraded to a limited degree (or sometimes even not degraded). Such a bounded degradation includes bounded lengthening of packet delay, acceptable packet loss rate and duration, low variation of jitter in receiving packets, and a low percentage of out-of-order packets. This indicates that the impact of those routing changes on reaching the multi-homed site is not significant. It is also not obvious that the DDJ&R of packet streams during the routing changes correlates with the control plane dynamics at the same period in terms of those selected control plane metrics.

The paper is organized as follows. Section 2 presents our methodology in generating and measuring routing changes at the control plane and our methodology in setting up and measuring data plane performance. Section 3 reports the data stream delivery performance when routing changes are introduced. Section 4 studies whether there is a correlation between DDJ&R and popular control plane metrics. Section 5 describes related work. Finally, we conclude the paper and point out open issues in Section 6.

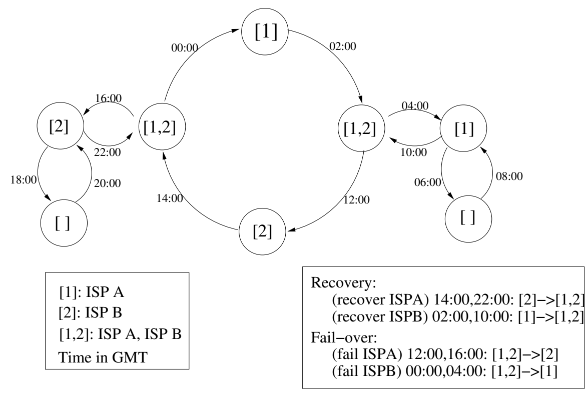


Fig. 1. Multi-homed Beacon schedule

2 Methodology

2.1 Measurement Setup

When a site incurs a BGP routing change toward a destination, it will lose and/or gain a path toward that destination. We injected routing changes by using a *BGP Beacon* [4], an otherwise unused globally visible prefix with a known schedule for announcements and withdrawals. We used a multi-homed BGP Beacon 192.83.230.0/24 that has been active since September 2003. The Beacon routers are housed in one of Seattle's major carrier hotels with 100+Mbps connections to two global providers. Both providers are transit-free ISPs providing international network connectivity.

Every two hours, the Beacon sends a BGP withdrawal or announcement message to one or both providers, thus simulating the control plane changes of a multi-homed site losing and restoring a link to one or more of its providers. Note that switching paths between two global providers will very likely lead to significant routing changes. The specific transitions the Beacon goes through each day for this study are shown in Figure 1.

Throughout the paper we use *AB-A, AB-B, A-AB, B-AB* to represent "fail ISP B," "fail ISP A," "recover ISP B," and "recover ISP A," respectively. Also for simplicity, we say that a site prefers ISP x when a probe site, if given multiple paths toward the Beacon prefix, chooses the one advertised by ISP x.

2.2 Data Collection from Data Plane

To measure packet delivery performance during controlled routing changes, we selected 161 geographically and topologically diverse probe sites from the PlanetLab [5] (a few RON [6] nodes were also used), and measured data streams from them toward a *test stream sink* over an extended period of four months.

With this setup, if all data streams display similar data plane characteristics when specific routing changes occur, we can regard the routing changes to be the primary reason. The topology diversity setup (which can be further improved based on the discovery from [7]) can help us leave out the topology as a major factor in affecting the characteristics, and measuring data streams over an extended period can help us not worry about the temporal conditions—such as congestion—as major reasons in affecting the characteristics.

During the period in which a routing change was injected, every probe site sends streams of UDP packets at 50ms intervals toward the *test stream sink*, which is a host configured with the specific IP address within the Beacon prefix. UDP was used instead of TCP, because the latter's performance depends on a vast number of factors, so UDP is a better measure of simple network performance. To calibrate the performance, such streams were also sent during time periods when no routing change was injected. Every packet was stamped with a sequence number and a departure time. No other live hosts existed behind the Beacon prefix. The test sink recorded every packet it received, including the timestamp, the sequence number, and the TTL value of the packet.

The DDJ&R of these data streams are measured as follows:

- *Delay*: We can measure either one-way delay or round-trip delay, but the former is subject to clock skew on PlanetLab sites, and the latter to asymmetric paths. Therefore, we first find the mean one-way delay, and then adjust all the delays to be relative to the mean. Measuring delays relative to the means makes the steady-state mean show up as 0 ms. We found this relative one-way delay works well for evaluating delay dynamics from the same probe host.
- *Drop* (or *Loss*): Drops are detected as gaps in sequence numbers which are never filled. Two common metrics are used in our work: *loss rate* and *loss duration*. Loss rate is the percentage of dropped packets per second. Loss duration is the time span with exceptionally high loss rate. To filter out noise and statistically insignificant losses, the loss duration is calculated as follows:
  1. Compute the loss rate in every one-second time window. We choose a one-second window to reduce smaller windows' sensitivity to a small number of losses.
  2. Set a threshold of the average loss rate plus two standard deviations. Thus, if a particular host is experiencing regular high loss, we still look for *exceptionally* high drop percentages.
  3. Find the interval which includes the maximum number of one-second time windows each of which are above the threshold, and that themselves have a loss rate above the threshold.
- *Jitter*: Jitter is computed as the discrete first derivative of the delay. For each received packet, if the previous sequence number is received the jitter is the delta between their delays.
- *Reordering*: Based on the sequence numbers in packet streams, when a packet arrives out of the expected sequential order, it is counted as a reordered packet. Reordering rate is defined as the percentage of reordered packets per second.

2.3 Data Collection from Control Plane

Control plane data were collected from the Oregon RouteViews [8] project, which consists of distributed monitors that receive routing data from a large number of diverse routing peers. The RouteViews monitors do not forward packets, but instead generate a repository of archived BGP updates. Updates are timestamped locally when they arrive, and are dumped to disk at 15-minute intervals. We used RouteViews archives from route-views2.uoregon.edu to retrieve data related to a specific event at specific times. Other RouteViews monitors were not considered in our study due to inconsistent timestamps on the different monitors.

For each routing change injected by our BGP Beacon, we observed its effect from the monitors' peers, filtering updates for the BGP Beacon prefix 192.83.230.0/24. Since every Beacon state transition happened at an exact hour, the updates which fell into a surrounding [-10m, 10m] window were collected. To reduce the impact of external routing changes not injected by our measurement setup, we used anchor prefixes [4] to ignore beacon events when real external global routing changes occur.

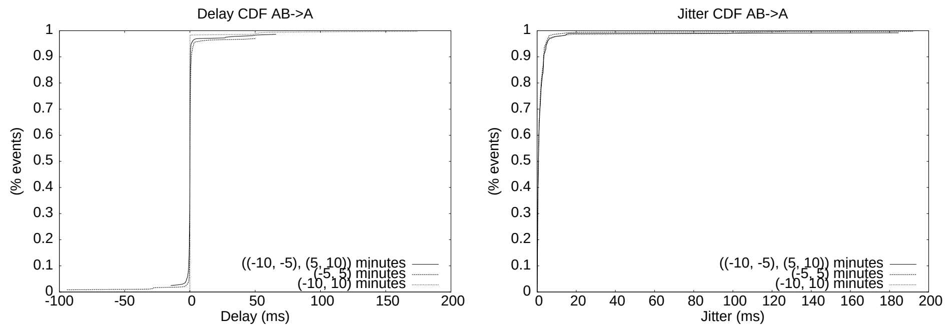
We focus on two popular control plane metrics in this paper. We define *BGP duration* (or *BGP update duration*) as the time from the first update to the last update received during an event. Similarly, the *BGP update volume* is counted as the total number of updates during the event.

3 DDJ&R of Data Streams During Routing Changes

We measured delay, drop, jitter, and reordering from every probe host toward the test stream sink over a period of four months. We report the measurement results in this section.

Figures 2(a) and 2(b) show the CDF distributions of delay and jitter respectively for the AB-A transition for all probing hosts. These graphs show CDF lines for three time windows: [-5, 5] minutes, [-10,-5], [5, 10] minutes, and [-10, 10] minutes. As the routing change disturbance did not last longer than 10 minutes as we found, these three windows capture the (potential) differences between delay and jitter measurements during the routing changes, each side of the routing changes, and during the whole event. We can see that the distributions for both delay and jitter during each of these windows is almost identical. This suggests that the packets were experiencing no significant overall changes in performance during the AB-A events. Similar results exist for the AB-B, A-AB, and B-AB events.

The loss rates for probe sites that preferred ISP B are similar to the loss rates for those that preferred A. Figures 3(a) to 3(d) show aggregated loss rates for some probe sites that preferred ISP B, under four different routing changes, AB-A, A-AB, AB-B, and B-AB, respectively. We can see that Figures 3(a) and 3(b) concern losing and obtaining the preferred path of a probe, respectively. Therefore, their loss rate during routing changes is more dramatic than that in Figures 3(c) and 3(d), where a probe site loses or obtains the path it does



(a) (b)

Fig. 2. CDF of delay and jitter for all hosts (AB-A).

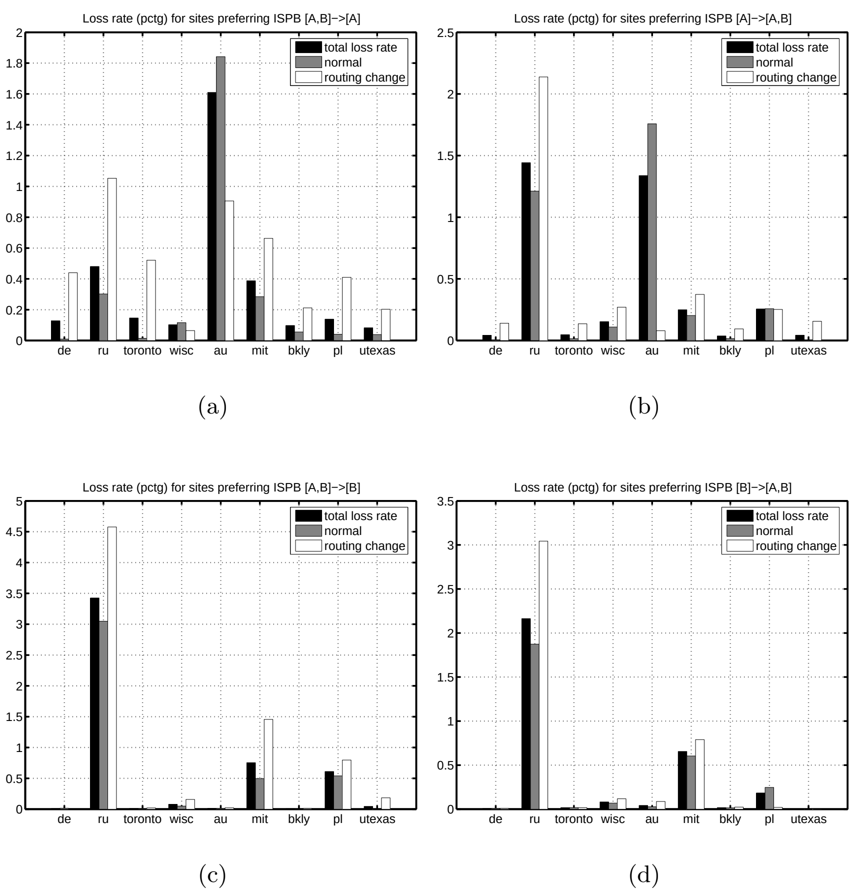


Fig. 3. Loss rate (%) for sites preferring B.

not prefer. It is interesting that packets from a couple sites (*au* and *pl*) experienced lower loss rate during routing changes than during the normal period, which could be due to the high congestion of the new path or due to the almost instantaneous path switch that did not cause much packet loss.

We do not show aggregated results of reordering since the number of reordered packets is close to zero during injected routing changes.

From the results above, we can see that, in most cases, during the artificial routing changes, data plane performance is acceptable in terms of DDJ&R; hence BGP is doing its job well. The increase in the delay of UDP streams will be less than 10 ms, with more than 90% probability. Also with more than 90% probability, the jitter is less than 4 ms, and less than 10ms with 99% probability. Loss and reordering rates are also generally low. With the exception of two international probe sites, *ru* and *au*, the loss rate is generally lower than 1% during routing changes, and often lower than 0.4%. Overall, in most cases, the BGP performance is satisfactory.

4 Correlating DDJ&R with BGP Updates Metrics

We found little correlation between DDJ&R and those BGP update metrics we studied.

First, it is difficult to make claims about packet performance based on the duration of BGP chatter. Figure 4(a) shows BGP update duration versus loss duration when ISP B is no longer a provider. Were there a correlation between BGP update duration and loss duration, there should be a curve or trend line matching most points. This sort of trend does not exist here. While the loss duration varies from almost 0 seconds to approximately 220 seconds, most BGP update durations are around 100 seconds, with several being roughly 440 seconds.

Figure 4(b) is similar to Figure 4(a), except the route to ISP A is withdrawn. We observe a similar grouping of BGP update durations across a range of loss durations. Figures 4(c) and 4(d) depict similar patterns for the recovery events A-AB and B-AB, both again showing essentially no correlation between loss duration and BGP update duration.

Figure 5 shows the aggregated BGP update duration and loss duration over all beacon events. This graph shows similar data to Figure 4 except with more pronounced grouping. BGP update durations range from 10 to 200 seconds and loss durations range from near 0 to 50 seconds. We also see a large number of outlying points. In some cases, BGP update duration falls within the same range while the loss duration is from 50 to 220 seconds. In other outlying points, BGP update duration is above 300 seconds while loss duration remains between 0 and 50 seconds.

Note that a majority of loss durations are shorter than BGP durations. This is because a packet from a source could still reach its destination before the source converges on a new path to the destination. Some intermediate routers along the way, which are closer to the destination, could have already converged on the new path. The packet can thus follow the converged path from those

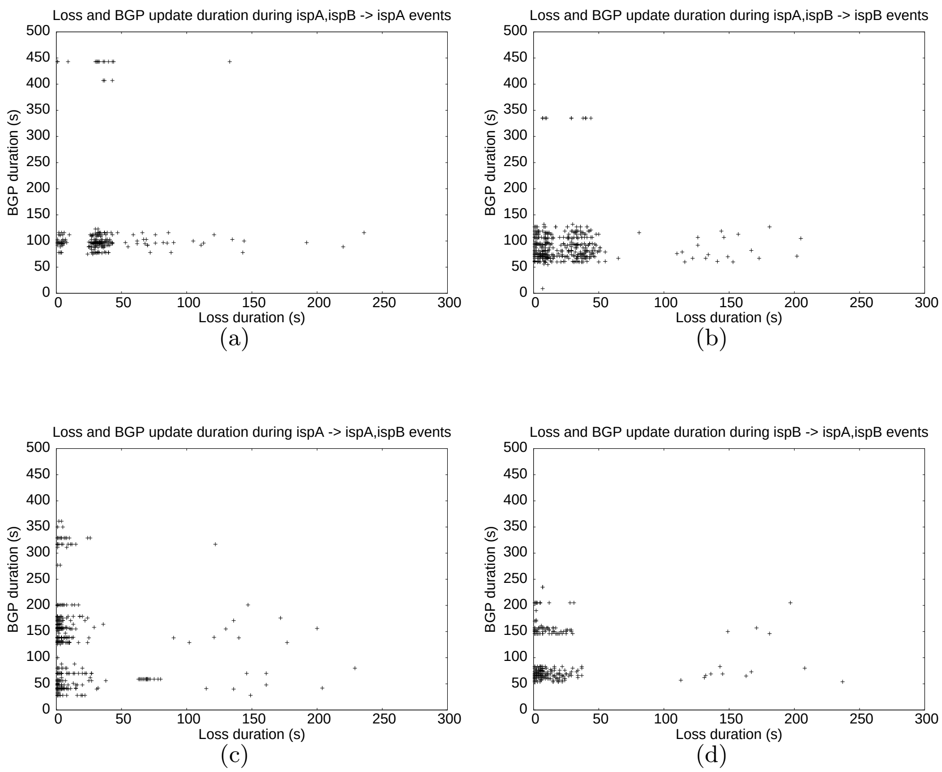


Fig. 4. Packet loss duration versus BGP update duration.

intermediate routers. In another words, the whole path that the packet takes may not be optimal, but still leads the packet to its destination.

Moreover, as demonstrated by Figure 6, there is little likelihood of a relationship between the volume of BGP updates observed by RouteViews and the loss duration. Here we see an even more pronounced range of values from the control plane measurements with the same loss durations. The volume of BGP updates ranges from roughly 15 to 120 with some outliers. Unlike Figures 4 and 5, there are no clear groupings of control plane values, making it even harder to correlate a data plane metric (loss duration) with a control plane metric (BGP update volume).

The lack of correlation between loss duration and BGP update duration, or between loss duration and BGP update count, strongly affirms that these BGP metrics cannot be used to evaluate the data plane performance in this context where data streams are delivered to a multi-homed site. There might be a different metric or set of metrics from the control plane that could be correlated to DDJ&R of data streams, for which further research is needed.

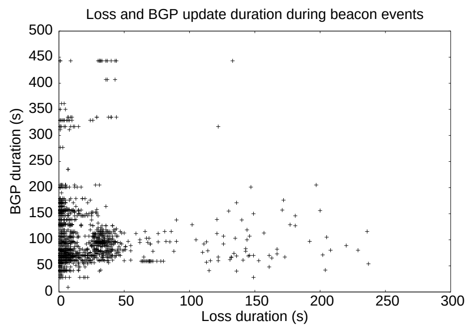


Fig. 5. Packet loss duration versus BGP update duration - All BGP events

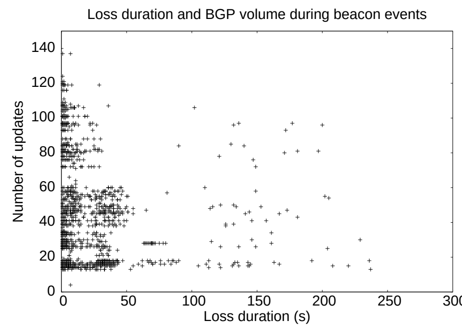


Fig. 6. Packet loss duration versus BGP update volume- All BGP events

We suggest that increases in the volume of global BGP announcements or BGP update duration could indicate the network healing itself, thus a proper operation of routing, not necessarily its failure that will lead to severe data plane degradation.

5 Related Work

Packet performance itself has been studied at length in end-to-end measurements conducted by Paxson et al. [9]. While these measurements observed pathological behavior based on routing events, they are not used for understanding the relationship between the data plane and the control plane *per se*.

For studies that are concerned with the relationship of the two planes, certain aspects have been studied. An earlier work by Craig et al. found that the BGP convergence can be fairly long, and during which the data delivery performance toward the affected prefixes would degrade [1]. End-to-end Internet paths will experience intermittent loss of connectivity, as well as increased packet loss and latency. However, it is unclear that as the Internet becomes more densely connected how data plane performance would be affected by routing changes. We extend their work by measuring the performance from a diverse set of probe locations toward a multi-homed sink, and studying the correlation between metrics from the two planes.

The relationship between BGP updates and traffic volume has also been studied. Rexford et al. showed that only a small percentage of BGP updates are related to "popular" prefixes [2]. Agarwal et al. further found that even a smaller percentage of them would have an impact on the majority of traffic [3].

In measuring the effects of Internet path faults on reactive routing, Feamster et al. found that end hosts with multiple connections to the Internet are more resilient to path faults (outage) [10]. They also studied the possible correlation between the control plane and the data plane in terms of BGP updates and path failures, and found that end-to-end path failures could precede BGP instability and conversely, BGP instability could also precede end-to-end path failures.

6 Conclusions

We believe that data plane performance is the best measure of control plane effectiveness. After all, the goal of the Internet is to deliver packets from a source to a destination. Though we only studied BGP, we believe that data plane performance should be a significant metric in judging the efficacy of other routing protocols, *e.g.*, intra-domain routing protocols such as IS-IS and OSPF.

Through watching data streams toward a multi-homed sink during routing changes, we have shown in this paper that the data plane performance while routing changes are occurring is generally good and not significantly affected. We further found that there is little correlation between the widely used performance metrics for the data plane, *i.e.* delay, drop, jitter, and reordering that users care about, and those for the control plane, which are BGP duration and BGP update volume in our study here.

We believe that this paper paves the way for future work. Insofar as demonstrating that data plane performance for a multi-homed site is *not* necessarily affected during a routing change and that a correlation does *not* exist between selected data plane and control plane metrics, this study is sufficient. However, more research can be done under different experimental configurations. For example, the routing changes we introduced are next to the destination of data streams, and one may need to investigate more general routing changes. Ensuring no data bias may be also important when using data from PlanetLab and RouteViews or any other experimental environment. Additionally, we studied BGP update duration and volume, but there might be other control plane metrics that may correlate well with data plane metrics.

References

1. Labovitz, C., Ahuja, A., Bose, A., Jahanian, F.: Delayed internet routing convergence. IEEE/ACM Trans. Networking **9** (2001) 293–306
2. Rexford, J., Wang, J., Xiao, Z., Zhang, Y.: Bgp routing stability of popular destinations. In: Proc. ACM Internet Measurement Workshop. (2002)
3. Agarwal, S., Chuah, C.N., Bhattacharyya, S., Diot, C.: The impact of bgp dynamics on intra-domain traffic. ACM SIGMETRICS **32** (2004) 319–330
4. Mao, Z., Bush, R., Griffin, T., Roughan, M.: BGP Beacons. In: Proc. ACM Internet Measurement Workshop. (2003)
5. PlanetLab: An open platform for developing, deploying, and accessing planetary-scale services. (`http://www.planet-lab.org`)
6. Andersen, D.G., Balakrishnan, H., Kaashoek, M.F., Morris, R.: Resilient overlay networks. In: Proc. 18th ACM SOSP, Banff, Canada (2001)
7. Banerjee, S., Griffin, T.G., Pias, M.: The interdomain connectivity of planetlab nodes. In: PAM. (2004)
8. University of Oregon: (Oregon Route Views project) http://www.routeviews.org/.
9. Paxson, V.: End-to-end routing behavior in the Internet. IEEE/ACM Trans. Networking (1997)
10. Feamster, N., Andersen, D.G., Balakrishnan, H., Kaashoek, M.F.: Measuring the effects of Internet path faults on reactive routing. In: ACM SIGMETRICS. (2003)