Happy Packets to You!

Randy Bush, Timothy G. Griffin, Jun Li, Zhuoqing M. Mao, Eric Purpus, Daniel Stutzbach

Paper ID: 1568938735 Number of pages: 12

***Abstract*— Previous studies of the Internet control plane examined its quality from a number of different points of view, but very few considered the performance of the data plane as a measure of control plane performance. For example, a large number of routing updates during some period or slow routing convergence, are often cited as an indication of a poorly designed or behaving routing protocol. However, if the customers' packets are being delivered well, which we term "happy packets", while routers can handle routing updates without being overloaded, then how significant are these control plane measurements?**

We believe packet happiness should be the primary metric for measuring control plane performance, especially the quality of routing. Since handling routing changes is the fundamental function of any routing protocol, in this paper we measure packet happiness during controlled routing changes. We obtain results by using established data plane metrics – delay, drop, jitter and reordering (DDJ&R) – in order to evaluate BGP, the *de facto* standard inter-domain routing protocol. Moreover, we compare the results with routing updates observed by RouteViews, a commonly used resource for studying the control plane. We establish that generally there is *no clear correlation* between the number and duration of routing updates derived from partial views and the performance of the data plane, further cautioning the usage of metrics from the control plane alone in evaluating control plane quality.

***Index Terms*— control plane, data plane, BGP, happy packet, network measurement**

I. INTRODUCTION

We frequently hear comments about Internet control plane quality, such as

- Internet routing is fragile and collapsing,
- Yesterday was a bad routing day on the Internet,
- BGP is broken or is not working well,
- Changing protocol X to Y will improve routing, or
- Internet routing has been severely affected by event X (*e.g.*, power blackout, worm outbreak).

This research is funded by NSF award ANI-0221435. Randy Bush is with IIJ, Email: randy@psg.com; Timothy G. Griffin is with Intel Research, Email: tim.griffin@intel.com; Jun Li is with University of Oregon, Email: lijun@cs.uoregon.edu; Zhuoqing M. Mao is with University of Michigan, Email: zmao@eecs.umich.edu; Eric Purpus is with University of Oregon, Email: epurpus@cs.uoregon.edu; Daniel Stutzbach is with University of Oregon, Email: agthorr@cs.uoregon.edu.

To evaluate Internet control plane quality, a variety of measures have been used, including number and frequency of routing updates [1], [2], number of prefixes [3], size of routing tables [4], and the speed or completeness of routing convergence [5], [6]. For example, significant numbers of BGP updates are often equated with Internet instability [1], [2]. One such comment that is often expressed by people studying BGP is "There are too many BGP updates, so BGP must be broken."

But what is *good* routing? How can one decide that a particular measurement shows one routing scheme is superior to another? What metrics should be used?

While people often contend that one measurement is better or worse than another when measuring control plane quality, the measure which ultimately counts is whether the customer's packets reach their intended destination with good performance. If the customers' packets are *happy*, the routing system, and other components in the control plane, are doing their job well. So, while all those metrics above are indeed important, packet performance should be the primary metric for evaluating control plane performance.

It is indeed true that sometimes Internet routing may experience huge routing updates, but perhaps they are just like white blood cells; although their presence may signal a problem, they are often part of the cure, not necessarily part of the problem. As long as routers will not fall over due to a sharp load increase, it is not clear the control plane is in danger. Furthermore, BGP already has built-in timers such as the MinRouteAdvertisement timer [7] that can rate-limit the updates to alleviate router update processing overhead.

Also, it is not clear that high-quality packet delivery requires routing convergence as we speak of it today. Even though the routing system is in the middle of convergence, packets may still be able to reach their destinations smoothly. For example, if there is better routing information near the destination than at the source, it is still possible that there is sufficient information near the source to get the packet to the better informed space. Also, we know convergence can be sped up sometimes if announcement throttling (MinRouteAvertisementInterval Timer [7]) is reduced [5].

We set out to measure control plane quality by measuring the data plane. Fortunately, there are well-established metrics for determining the "happiness" of packets from the data plane: delay, drop, jitter and reordering (*DDJ&R*). Because handling routing changes is the fundamental function of routing protocols, we measure and evaluate these packet performance metrics (DDJ&R) during controlled routing changes. In this paper we focus on BGP since it is the *de facto* inter-domain routing protocol on today's Internet.

We study how data plane performance can be used to evaluate the quality of the Internet control plane. In particular, if during routing changes the packet delivery quality is only degraded to a limited degree (or even not degraded), we assert that the underpinning routing protocol is effective. Such a bounded degradation includes bounded lengthening of packet delay, acceptable packet loss rate and duration, low variation of jitter in receiving packets, and low percentage of out-of-order packets.

On the other hand, instead of directly measuring DDJ&R of packets in the data plane, one may be tempted to derive such information from continuously archived control plane data (such as those collected by RouteViews [8] or RIPE [9]) that have been commonly used for studying the control plane. Such data, albeit partial, is globally visible, and can be used conveniently to deduce information such as the duration that the path to a prefix is unavailable, or the number of BGP updates exchanged during this duration. The question is, however, can such metrics correlate with the DDJ&R of packets, thus potentially replacing or reducing the effort of DDJ&R measurement?

This paper is a first step toward successfully using the data plane to evaluate the control plane, and focuses on primary research issues. Some interesting topics, therefore, are out of the scope of this paper. For instance, when measuring the DDJ&R of packets, we do not consider possible effects of congestion, do not test all possible routing changes, and do not try to associate the performance of packets with *all* possible factors. Also, we will not fully address the implications of our findings on the design of future routing and application protocols. **Our contributions are mainly to establish data plane performance as the primary metric for control plane evaluation, design a method to achieve that, and analyze a common misconception that leverages a control plane data archive.**

The paper is organized as follows. Section II describes the background and related work. Section III illustrates how we collect data from both the data plane and the control plane, including the experiment methodology involving actively introducing routing changes and monitoring long-term continuous UDP data streams. We describe the measurement results in Section IV, and analyze the data plane performance during routing changes. Section V discusses other factors that may affect data plane performance and a potential alternative for predicting packet happiness. Open issues and future work are discussed in Section VI and we conclude the paper in Section VII.

II. Background and Related Work

Previous studies have focused on a number of different aspects of control plane performance. However, most have ignored the most important metric: the performance of the data plane.

Studies such as [1], [2] have examined BGP behavior under the stress during events such as the Code Red II and Nimda worms as well as the US east coast blackout in August of 2003. These studies base their findings on distant and partial views of BGP updates. We call this type of data *partial* because the routing updates are collected from a very small subset of all BGP speakers. This kind of data only contains views from those particular peers. The best-known examples of partial BGP data are those collected by the Oregon RouteViews project [8] and the raw data from RIPE's Routing Information Service [9].

Wang *et al.* [10] also studied BGP behavior under the stress of Code Red and Nimda and concluded that over 40% of observed updates can be attributed to monitoring artifacts, specifically monitoring BGP session resets, of these partial data sources. As a result, this calls into question the applicability of metrics such as the number of updates observed by these partial views. Observing significant spikes in the number of BGP updates in these cases was an observation of broken multi-hop peering sessions and not an observation of BGP's behavior as previously assumed.

Other approaches taken in studying control plane performance also analyzed updates from these partial views. Labovitz *et al.* discovered duplicate withdrawals accounting for over 60% of the total number of updates observed at major backbone peering points in their 1997 study [11]. Again, these studies depend on the applicability of partial data views to the overall performance of the control plane.

Studies such as [5], [12] have taken an analytical approach to assess BGP behavior. Griffin [5] has shown that BGP does not always converge and that checking for convergence criteria is an NP-hard problem given static knowledge of policy configurations among all peers. In addition to divergence in BGP, Griffin [12] describes scenarios in which dynamically setting multi-exit descriptors (MEDs) based on IGP protocol metrics can lead to MED oscillation. All these studies take a different approach from those using partial data view. However, none of these approaches examine what happens to packet performance.

Packet performance itself has been studied at length in end-to-end measurements conducted by Paxson *et al.* in [13]. While these measurements observed pathological behavior based on routing events, they are not used for understanding the control plane *per se*. In particular, it is not the foci of these works to study control plane quality via data plane performance. Labovitz *et al.* [14] also has done some preliminary study of how data plane behaves during routing changes. We extend their work by using a comprehensive set of data plane performance metrics and a more diverse set of probing locations.

Recent work by Feamster *et al.* has begun to address a possible correlation between the control plane and the data plane in [15] using timing and BGP update count information. This work also uses end-to-end active probing and *local* BGP data. Their findings show that cases exist where end-to-end path failure precedes BGP instability and where conversely, BGP instability *preceded* end-to-end path failure. This supports our hypothesis that control plane data alone is insufficient to predict data plane behavior.

III. Methodology

In this section, we describe how we collect data from both the data plane and the control plane of the Internet. From the data plane, we have designed a methodology to measure packet performance in terms of DDJ&R under controlled routing changes. From the control plane, we leverage the RouteViews archive of routing updates.

A. Data Collection from Data Plane

We designed a methodology to measure packet performance in terms of DDJ&R during *controlled* routing changes injected by the Beacon infrastructure [16]. In what follows we first describe how we control routing changes, then describe how we measure packet performance during the routing changes.

1) BGP Beacon: We inject controlled routing changes by using a *BGP Beacon* [16], an unused globally visible prefix with a published schedule for announcements and withdrawals. Our experimental setup consists of a multi-homed BGP Beacon 192.83.230.0/24 that has been active since September 2003. The Beacon router is housed in Seattle's major carrier hotel with 100+Mb connections to each of two global providers. Both providers are tier-1 ISPs providing international network connectivity. Periodically, the Beacon sends a BGP withdrawal message to one or the other provider, thus simulating the control plane changes of a multi-homed site losing a link to one of its providers. After the withdrawal message is sent, the Beacon sends an announcement to the failed provider to simulate the subsequent failure recovery when the link is repaired.

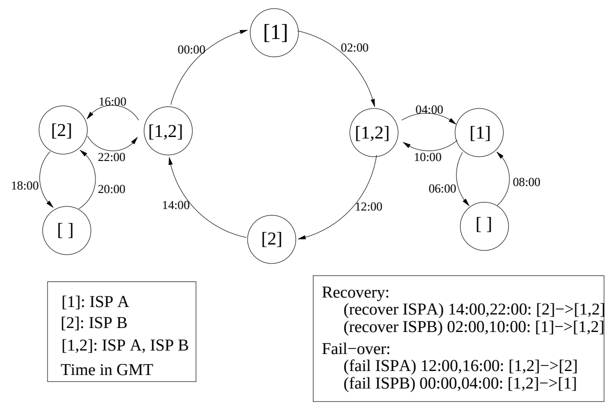


Fig. 1. Multi-homed Beacon transition schedule

The detailed transitions the Beacon goes through each day for this study are shown in Figure 1. Each circle indicates the state that the Beacon is in, expressed in terms of the upstream providers offering network connectivity to the Beacon prefix. Each arrow identifies a state transition as a result of a single BGP announcement or withdrawal message to one of the upstream providers. The label on the arrow indicates the time in GMT when the transition occurs. Throughout the paper we also 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.

For simplicity, we say that a site prefers ISP A when a probe site is given multiple paths toward the Beacon prefix and it chooses the one advertised by ISP A.

2) Data streams from PlanetLab probe nodes: To measure packet delivery performance during controlled routing changes, we select a set of geographically and topologically diverse probe sites from the PlanetLab infrastructure [17], a distributed wide-area platform for testing planetary-scale network services. During the period when a routing change is injected, every probe site sends streams of UDP packets at 50ms intervals toward the *test stream sink*, a host configured with a specific IP address from the Beacon prefix. To calibrate the performance, such streams are also sent during time periods when no routing change is injected. Every packet is stamped with a sequence number and a departure time. No other live hosts exist behind the Beacon prefix. The test stream sink records every packet it receives,

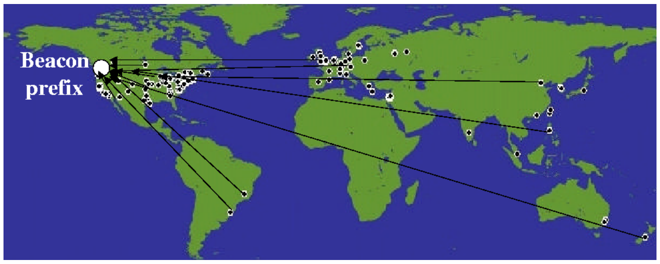


Fig. 2. PlanetLab experiment set up: UDP streams sent from many sites to the Beacon prefix

including the timestamp, the sequence number and the TTL value of the packet.

To calibrate the effect of routing changes, the streams of UDP data are also directed to another IP address that belongs to a prefix from the same AS as the Beacon prefix, but is not perturbed by our experiments. In addition to these UDP streams, at each PlanetLab node, both ping and traceroute measurements are also taken to record round-trip and IP-level routing information. Such measurements are taken at the same time as the UDP streaming. It is likely that the performance of the data plane depends on factors such as the network dynamics, the AS path, and the geographic location. To study this effect, we rotate across all 161 sites as probing sources.

3) Metrics: We define the DDJ&R metrics and describe how we measure them below:

- *Delay*: We can measure either one-way delay or roundtrip delay, but the former has to deal with clock skew problems on PlanetLab sites, and the latter has to consider asymmetric paths. We first find the mean one-way delay, and then adjust all the delays to be relative to the mean. We have 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 length of a time window with exceptionally high loss rate.
  A loss duration over a particular period is calculated as follows to filter out noise and statistically insignificant losses:
  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 to 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 that includes the maximal number of one-second time windows each of which are above the threshold, and that itself has 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 then 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 reordered packet. Reordering rate is defined as the percentage of packets per second.

B. Data Collection from Control Plane

One of our goals is to investigate the possible correlation between data plane performance (DDJ&R) and control plane dynamics (such as the number of routing updates and the duration of a routing change). Here, we describe how we collected control plane data and performed our control plane measurements. We discuss our investigation in detail in Section V-B.

Control plane data mainly are collected from the Oregon RouteViews [8] project, which consists of a few centralized monitors that receive routing data from a large number of diverse routing peers. The RouteViews monitors do not route any packets, but instead serve as a collected repository for archiving BGP updates. Updates are timestamped locally when they arrive, and are dumped to disk at 15-minute intervals.

We use RouteViews archives from route-views2.uoregon.edu to retrieve data related to a specific event at specific times. Other RouteViews monitors are not considered in our study due to inconsistent timestamps on the different monitors.

For each controlled routing change injected using our BGP Beacon scheme, we can observe its effect from the monitor's peers. We do this by filtering updates for the BGP Beacon prefix 192.83.230.0/24. Since every Beacon state transition happens at an exact hour, the updates which fall into the $[\text{-}10m, 10m]$ window, *i.e.*, $[\text{-}600s, 600s]$, are collected. We also count and record the number of updates received every second within the window. To reduce the impact of external routing changes not injected by our measurement setup, we use anchor prefixes as described in [16] to ignore beacon events when external routing changes occur.

Here, we define *BGP duration* as the time from the first update to the last update received during the event. Similarly, the *BGP update number* is counted as the total number of updates during the event.

We came across a similar difficulty with synchronized clocks between the BGP Beacon and RouteViews, just as we did with the PlanetLab nodes. As such, it is impossible for us to match exactly the time at which updates are received on the control plane with the time of the data plane measurements. However, since our correlation study involves the *duration* of data plane and control plane measurements, the slight clock skew does not affect our results.

IV. Results and Analysis

In this section we present our measurement results and analysis by answering the following question: Are packets happy during routing changes (especially when compared with the normal period), thus we can assert that a routing protocol such as BGP is performing well?

We have 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, then analyze how these results can be used to evaluate BGP performance. We also discuss whether these injected routing changes are sufficient to represent all possible routing changes.

In this section, we assume that in our experimental environment, the routing change is the only factor that affects packet happiness, thus DDJ&R information during routing changes can be used to judge the quality of a routing protocol underneath. We revisit this assumption in Section V.

A. Results

We use 128.95.219.192 to illustrate DDJ&R from an individual probe host, which represents the results from most other probe hosts (results for other hosts can be found at this Web site [18]). Figures 3(a) - 3(d) show DDJ&R measurements of UDP streams between 128.95.219.192 and the test stream sink over a 20-minute period under four different routing changes, AB-B, AB-A, A-AB, and B-AB, respectively. Most times the packet delay is acceptable, and no reordering was detected; thus it can be inferred that packet jitter is also acceptable during routing changes. As shown in Figure 3(b), however, a loss duration for about 30 seconds exists when the AB-A routing change happens. Also illustrated at the bottom of each figure is the number of BGP updates per second during the routing change as captured by RouteViews. In general, the packet streams from 128.95.219.192 are performing well, either during or outside the routing changes.

However, we also observe that unlike the majority of packet streams we observe, the DDJ&R of a packet

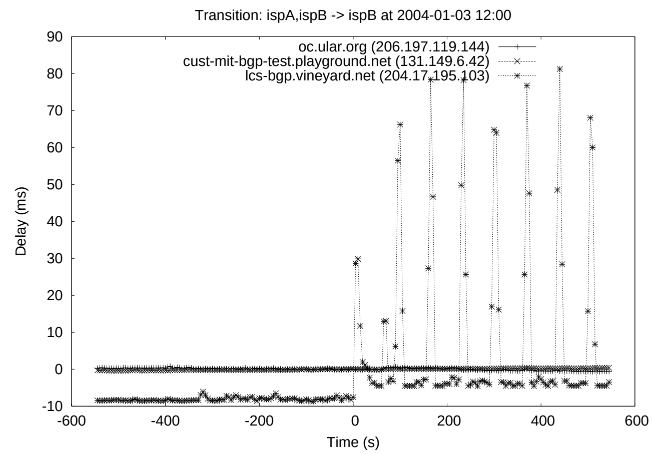


Fig. 4. Delay comparison of sampled data streams from different probe hosts

stream can also exhibit poor performance. For example, Figure 4 shows the worst case we have observed. We see that the delay of the UDP stream from the probe host lcs-bgp.vineyard.net at 2004-01-03 12:00 is much worse when compared to the delay of UDP streams from oc.ular.org or cust-mit-bgp-test.playground.net, where the latter is the common phenomenon. More details that include the drop and reordering information are shown in Figures 5(a) to 5(d), corresponding to different routing changes. The AB-B routing change also incurred the longest loss duration, $10s$, with 91 drops and 8 reorders. Notice that, however, the DDJ&R of packets during routing changes in this extreme case are not significantly worse than during the normal period, particularly since the DDJ&R during the normal period is already poor for this site.

The aggregated results in Figures 6(a) and 6(b) show the CDF distributions of delay and jitter respectively for the AB-A transition for all hosts. These graphs show CDF lines for three time windows during the routing changes: [-$5, 5$] minutes, [-$10, 10$] minutes, and [-10,-5]$, [5, 10]$ minutes. 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, and can be obtained from [18].

Aggregated loss rates for some of the probe sites preferring ISP B are shown in Figures 7(a) to 7(d), under four different routing changes, AB-A, A-AB, AB-B, and

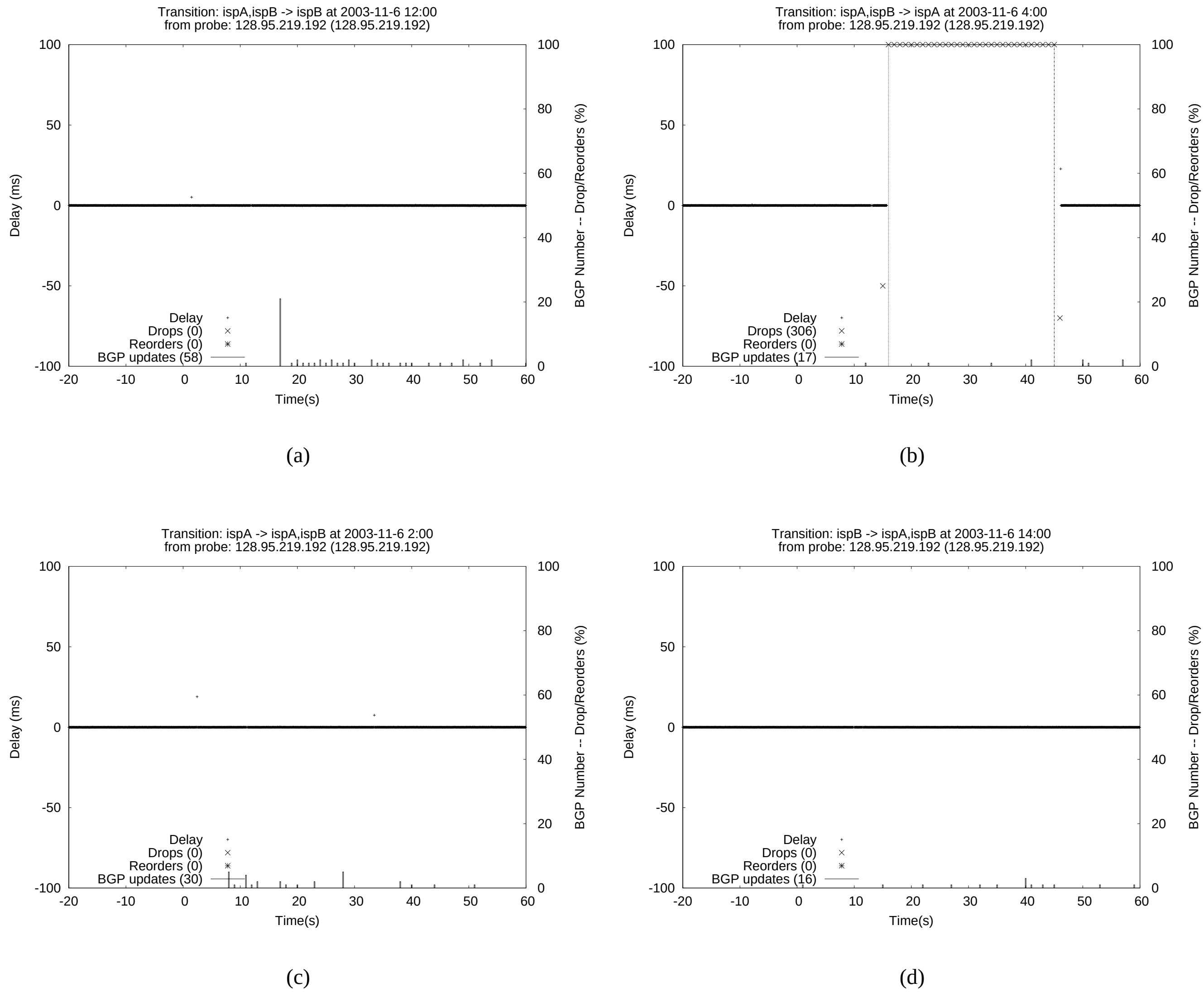


Fig. 3. DDJ&R of UDP streams from 128.95.219.192 toward the test stream sink

B-AB, respectively. We can see that Figures 7(a) and 7(b) concern losing and obtaining of the preferred path of a probe, respectively. Therefore, their loss rate during routing changes is more dramatic than that in Figures 7(c) and 7(d), where a probe site loses or obtains the path it does not prefer. It is interesting to notice that the packets from sites au and pl experienced lower loss rate during routing changes than during the normal period.

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

B. Analysis

From the results above, we can see that during those injected routing changes, BGP performance is acceptable in terms of DDJ&R in most cases. With a more than 90% chance the delay of UDP streams will be less than 10 ms. Also, with more than 90% probability, the jitter is less than $4ms$, and with 99% probability, less than $10ms$. 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%. In most cases, this is comparable to the loss rate at normal times without routing change. Even in the worst case as shown above, the DDJ&R during injected routing changes is still not significantly worse, especially when compared with the already poor DDJ&R during the normal period. Overall, although the DDJ&R performance of packets during routing changes is generally worse than during normal periods, it is acceptable, especially for non-realtime applications such as Web browsing. Therefore,

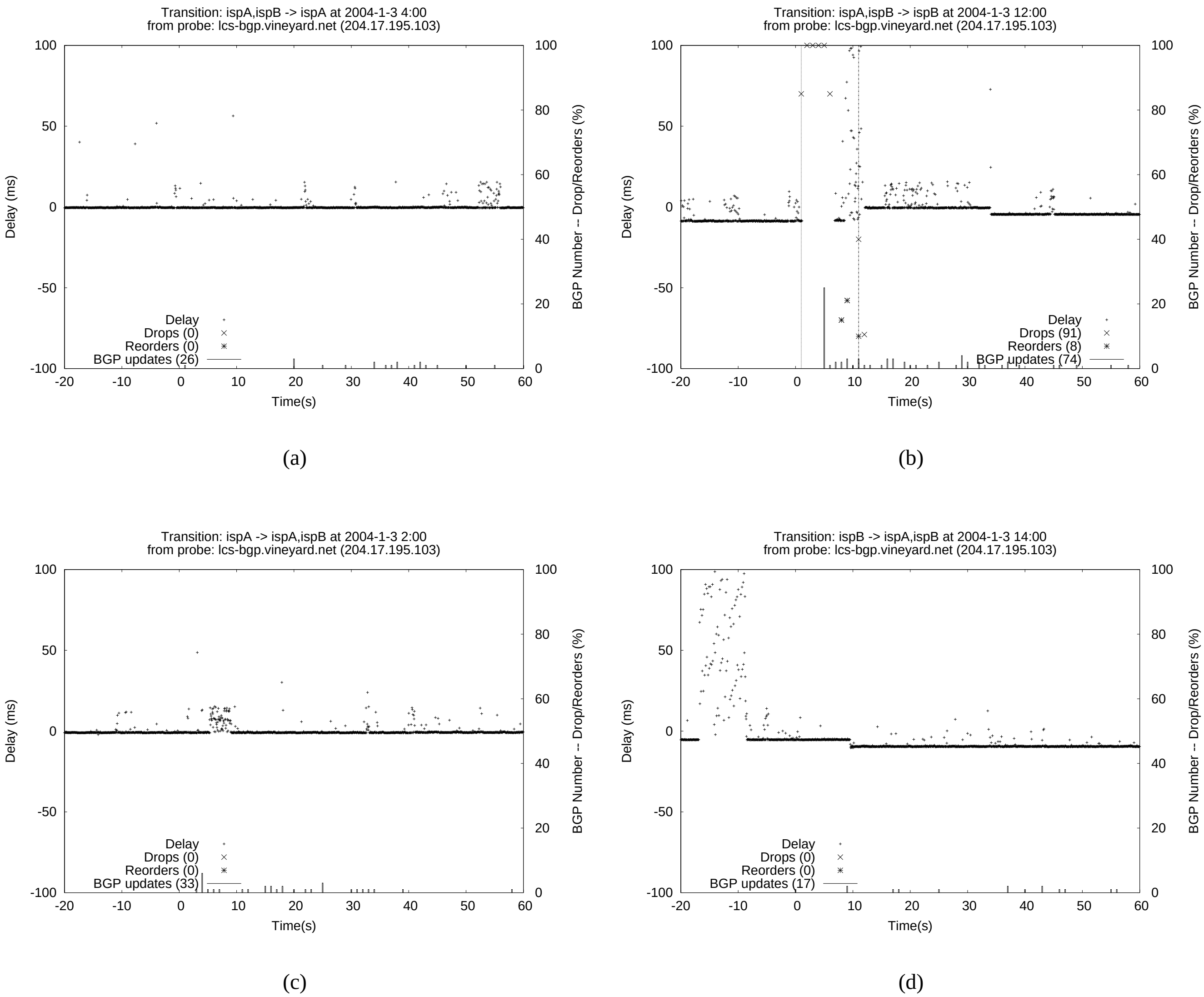


Fig. 5. DDJ&R of data streams from lcs-bgp.vineyard.net toward the test stream sink

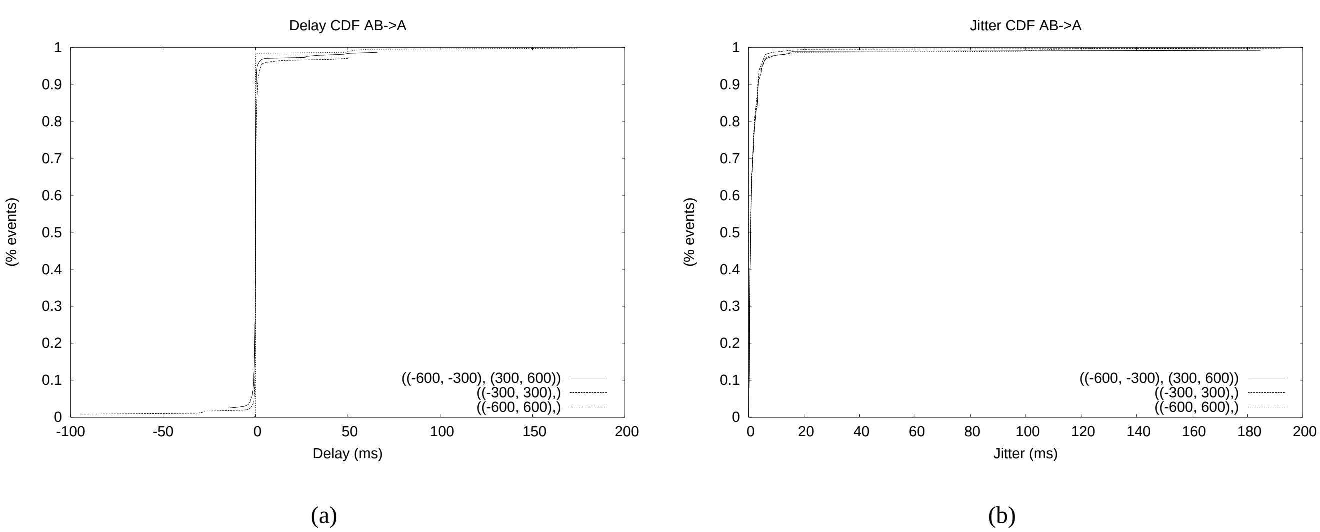


Fig. 6. CDF of delay and jitter for all hosts (AB-A)

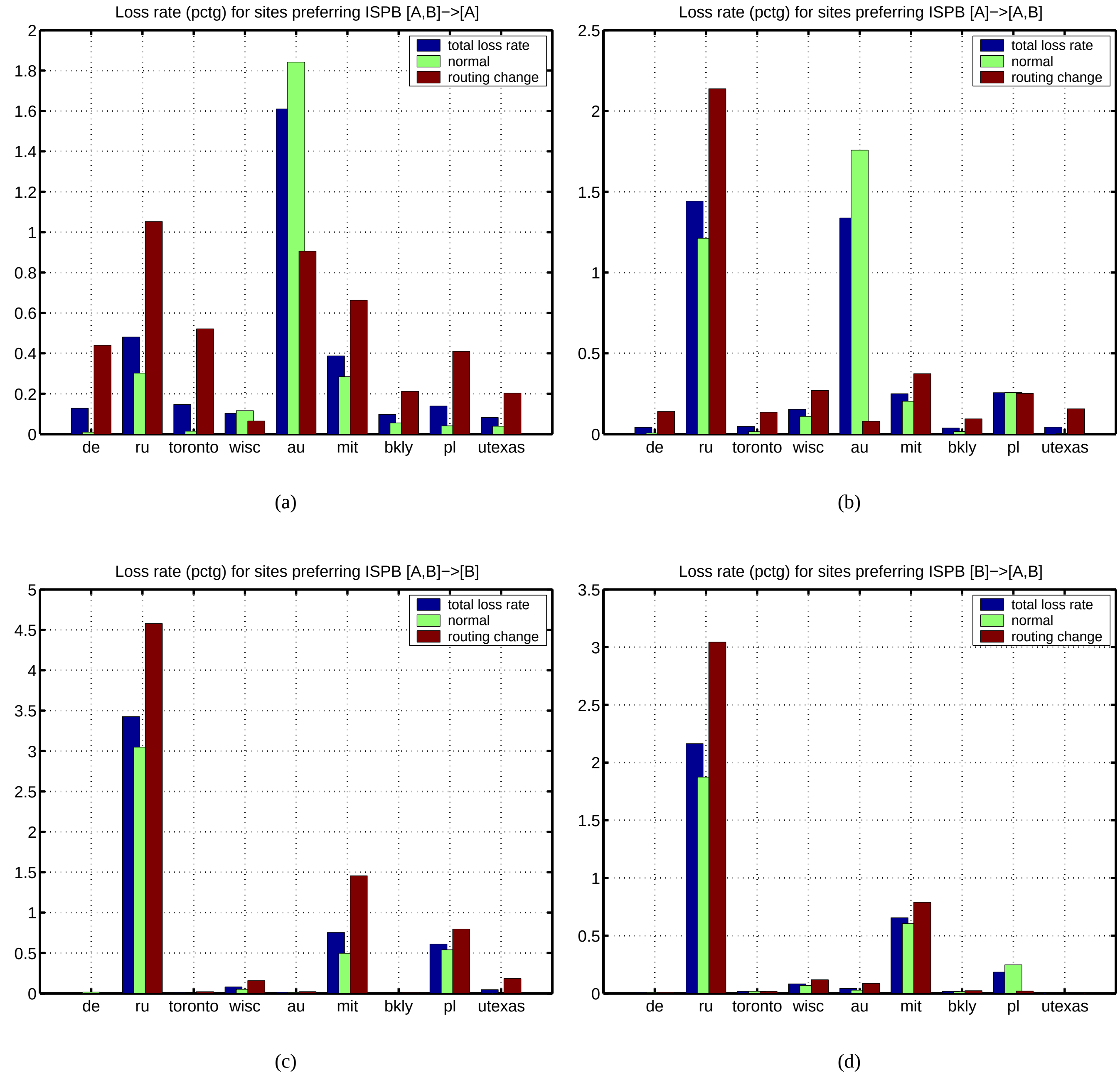


Fig. 7. Loss rate (%) for sites preferring B

we conclude that in most cases the BGP performance is satisfactory, especially in handling routing changes.

Our happiness measurement is also representative in reflecting most routing changes. The DDJ&R results between the test stream sink and a probe host is approximately the same as those between the Beacon router and the designated router of the probe host given the low latency between the host and its immediate local BGP router. Thus these results accurately reflect the happiness of packet stream between the two routers. Furthermore, the four different types of injected routing changes represent two basic cases of routing path fail-over and two basic cases of routing path recovery, and they can be regarded as primitives that all other routing changes can be derived, typically in the form of a sequence of these primitives. If packets are happy under these four types of routing changes, they will presumably be happy under other routing changes as well. Therefore, our happiness analysis for all UDP streams can approximate closely the packet delivery performance between two routers under a variety of routing changes.

V. Discussion

While we have shown the effectiveness of using data plane performance for control plane evaluation in Section IV, two major questions remain to be addressed:

- Is routing change the major reason that affects data plane performance, and therefore the DDJ&R of packets can be used to evaluate the routing protocol underneath?
- Are other metrics equivalent to packet happiness metrics for evaluating the control plane, especially those metrics from the control plane that have been commonly used?

We address the first question in Section V-A and the second in Section V-B.

A. Justifying Packet Happiness for Control Plane Evaluation

In Section IV, we assumed that routing changes are the only factor affecting data plane performance, therefore the DDJ&R of packets can be used to evaluate the routing protocol underneath. However, data plane performance is also affected by other factors, such as network congestion, high link error rate, *etc.* These are not directly related to routing changes. Given that the function of a routing protocol is to ensure good performance of the data plane (e.g., in case of congestion, identifying a better alternate route), we argue that DDJ&R is precisely the right measure for the control plane. Therefore, our analysis in Section IV still holds.

We show in this section that in our measurement environment, the packet delivery performance actually has little to do with simple static path characteristics such as hop count, and thus it is not easy to predict data plane performance directly using topology information. This emphasizes the need for using DDJ&R which captures the packet dynamics to evaluate the control plane.

We illustrate our discovery using the loss rate as a representative packet happiness indicator. Our measurements over all UDP streams have found that the loss rate has no direct correlation with either AS or router hop count. Figure 8(a) shows that there is little correlation between the packet loss rate and the AS hop count of a routing path. Similarly, the router hop count of a routing path also appears to be not correlated with packet loss rate, as shown in Figure 8(b).

B. Misconception on Inferring Packet Happiness

We believe packet happiness must be measured directly using DDJ&R measurements from the data plane. However, given the difficulty to accurately capture end-to-end DDJ&R, one may be tempted to use an easier but still seemingly effective approach: using control plane data, such as those collected by RouteViews or RIPE, to predict packet performance and hoping that this will be equivalent to measuring DDJ&R. For example, the duration from the time a path to a prefix is withdrawn to the time an alternate path is reestablished, or the number of BGP updates exchanged during this duration, *perhaps* can decide or indicate packet performance: the longer the duration is, or the more BGP updates are, the more unhappy the packets will be. Since those RouteViews or RIPE data are centrally managed and easily available, this approach indeed looks attractive.

In this section, we demonstrate that this seemingly good approach is actually misleading. We find little correlation between aggregate control plane information and DDJ&R. Using RouteViews data as the source of control plane information (see Section III-B), particularly the BGP duration and BGP update count, and using loss duration as the representative of direct measurement of packet delivery performance, we illustrate that such a correlation does not really exist.

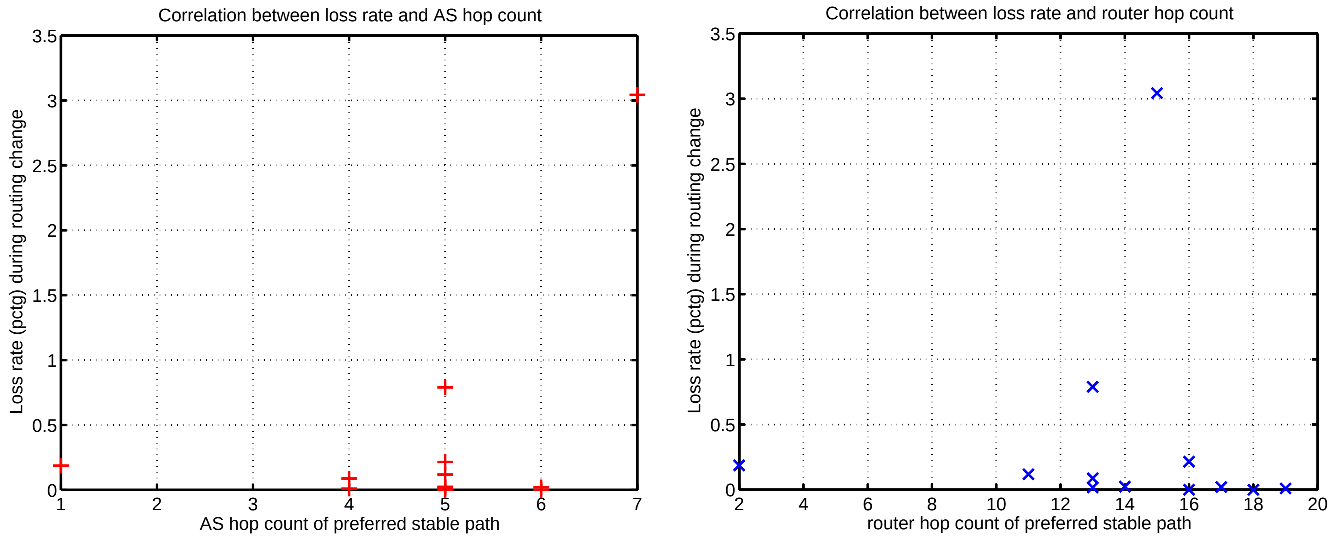
Figure 9(a) shows BGP duration versus loss duration when ISP B is no longer a provider. Were there a correlation between BGP duration and loss duration, there should be a curve or trend line matching most points. This kind of trend does not exist here. While the loss duration varies from almost 0 seconds to approximately 220 seconds, most BGP durations are around 100 seconds, with several being roughly 440 seconds.

Figure 9(b) is similar to Figure 9(a), except the route to ISP A is withdrawn this time. We observe a similar grouping of BGP durations across a range of loss durations. Figures 9(c) and 9(d) depict similar patterns for the recovery events A-AB and B-AB, both again showing no essential correlation between loss duration and BGP duration.

Figure 10 shows the aggregated BGP duration and loss duration data over all beacon events. This graph shows similar data to Figure 9 except with more pronounced grouping. BGP 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 duration falls within the same range while the loss duration is from 50 to 220 seconds. In other outlying points, BGP is above 300 seconds while loss duration remains between 0 and 50 seconds.

Overall, the graphs in Figures 9 and 10 show the difficulty of making claims about packet performance based on the duration of BGP chatter.

We now examine whether BGP update count can be used to predict packet performance, a commonly used metric by others [1], [2]. Figures 3 and 5 show the BGP



(a) (b)

Fig. 8. Packet loss rate vs. hop counts

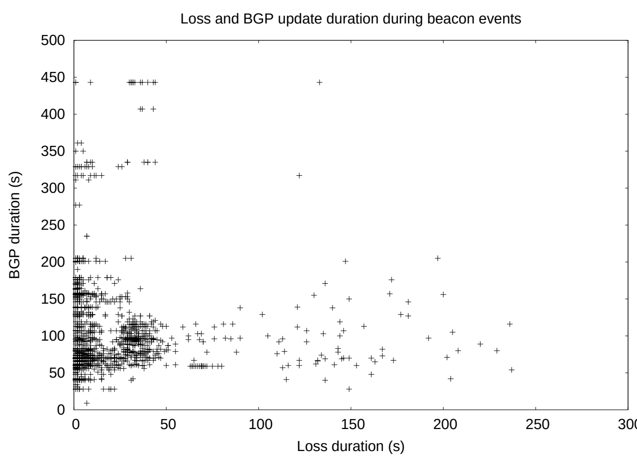


Fig. 10. Packet loss duration versus BGP update duration - All BGP events

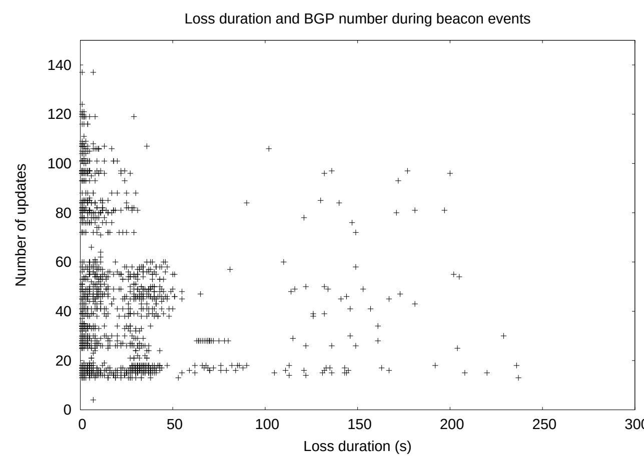


Fig. 11. Packet loss duration versus BGP update number - All BGP events

update rate during different routing changes. For the data streams shown in these figures, one can hardly discover a correlation between the BGP update rate and any metrics of delay, drop, jitter, or reordering. By showing an aggregated results, Figure 11 also demonstrates the little likelihood of a relationship between the number of BGP updates observed by our RouteViews-based partial view 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 number of BGP updates ranges from roughly 15 to 120 with some outliers. Unlike Figures 9 and 10, there are no clear groupings of control plane values. As a result, this makes it even harder to make claims about loss duration based on control plane measurements from partial BGP data.

The lack of correlation between loss duration and BGP duration, or between loss duration and number of BGP updates, strongly affirms our earlier hypothesis that partial knowledge from the control plane will be far from sufficient to predict packet performance. Neither guaranteed comprehensive nor representative, those data from RouteViews only represent partial knowledge related to the DDJ&R of packet delivery.

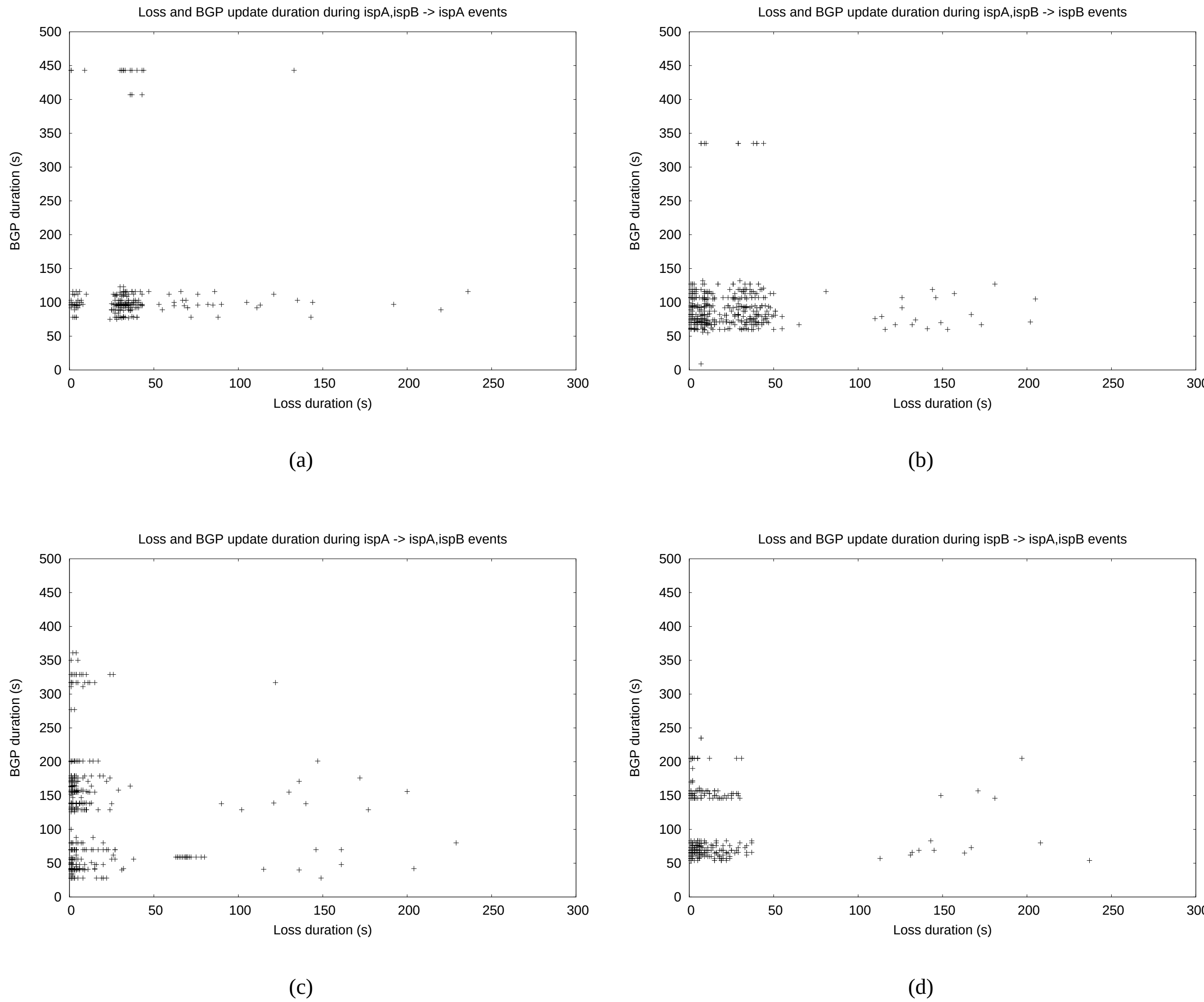


Fig. 9. Packet loss duration versus BGP update duration

VI. Open Issues and Future Work

Of course we believe there is more work to be done in this area.

This paper examines routing changes originating from a single site. We have yet to see a global Internet event, such as Code Red. It is likely unwise to make inferences from one site's behavior to global behavior during a large scale event. We have some measurement infrastructure to gather equivalent data should a large scale event occur.

When a routing change causes the data path to change, we suspect that some choices could traverse congested links. It might be useful to characterize performance of the paths chosen during routing changes.

As some of the graphs show, such as Figure 5(b), some data plane events, especially those over complex topologies, have complex structure. It would be useful to understand the causes of this complexity, and determine if and how it is related to control plane topology and/or performance.

Due to the way BGP obscures global knowledge, much of this and related work looks at how control plane information distant from the packet source and destination can predict or diagnose data plane performance. This leaves open questions of how much control plane information would allow better prediction and/or diagnosis. From what percentage of the routers in the Internet would RouteViews need feeds to be useful in such analyses? How does the topological distribution of control plane measurement influence the predictability of data plane performance?

VII. Conclusions

We believe that data plane performance is the best measure of control plane effectiveness. After all, the goal of the Internet is delivering the packets from the source to the 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.

We have found little to support assertions that BGP is not resilient or will not scale considerably beyond its current use. This is not to say that BGP is perfect; see T. Griffin's work on opacity of BGP policy [19], [20], [12] for an example.

Researchers and operators have the desire to predict and/or diagnose Internet performance by measuring parameters of the control plane, especially BGP. But, as BGP is a path vector protocol, one cannot have global knowledge of routing state without knowing the state of all routers; and this is considered infeasible as the Internet scales. This lack of global knowledge is why researchers and operators have tried to infer local behavior from distant data, the only data they/we can really ever get. So we look at remote and partial measurements such as RouteViews, or BGP feeds from some subset of the Internet, and try to infer from those data what local behavior will be between source and destination of a data plane/path.

But, as the Internet has scaled, the current management plane has not scaled with it. The portion of the global knowledge we can have is less and less. So we should be more and more critical of any inferences we might make from a weaker and weaker sample set.

We have shown in this paper that the data plane performance between a packet source and a packet destination, while routing changes are occurring at the destination, cannot be predicted by general distantly measured counts of BGP announcements. This is interesting because other researchers and operators have made such measurements and assumed a relationship which we have been unable to show.

Perhaps the best analogy for changes in the number of BGP updates is the white blood cell count (WBC) in humans. An elevated WBC is an indication of the body fighting some infection; i.e., the WBC is measure of the solution, not of the problem. We suggest that increases in the number of global BGP announcements are indications of the network healing itself, and indicate the proper operation of routing, not its failure.

References

[1] J. Cowie, A. T. Ogielski, B. Premore, and Y. Yuan, "Internet worms and global routing instabilities," in *Proc. of SPIE*, 2002.
[2] J. Cowie, A. T. Ogielski, B. Premore, E. Smith, and T. Underwood, "Impact of the 2003 blackouts on Internet communications," available at `http://www.renesys.com/news/2003-11-21/Renesys_BlackoutReport.pdf`, 2003.
[3] S. Bellovin, R. Bush, T. G. Griffin, and J. Rexford, "Slowing routing table growth by filtering based on address allocation policies," `http://www.research.att.com/~jrex/papers/filter.ps`, June 2001.
[4] G. Huston, "Analyzing the Internet BGP routing table," in *Internet Protocol Journal*, 2001.
[5] T. G. Griffin and B. J. Premore, "An experimental analysis of BGP convergence time," in *Proc. ICNP*, 2001.
[6] T. G. Griffin and G. Wilfong, "An analysis of BGP convergence properties," in *Proceedings of ACM SIGCOMM 1999*, 1999.
[7] Y. Rekhter and T. Li, "A Border Gateway Protocol 4," 1995, RFC 1771.
[8] University of Oregon, "Oregon Route Views project," .
[9] RIPE, "Routing information service raw data page," http://data.ris.ripe.net/.
[10] L. Wang, X. Zhao, D. Pei, R. Bush, D. Massey, A. Mankin, S. F. Wu, and L. Zhang, "Observation and analysis of BGP behavior under stress," in *Proc. ACM Internet Measurement Workshop*, 2002.
[11] C. Labovitz, R. Malan, and F. Jahanian, "Internet routing instability," *IEEE/ACM Trans. Networking*, 1998.
[12] T. Griffin and G. Wilfong, "Analysis of the MED oscillation problem in BGP," in *Proceedings of the 10th IEEE International Conference on Network Protocols*. 2002, pp. 90–99, IEEE Computer Society.
[13] V. Paxson, "End-to-end routing behavior in the Internet," *IEEE/ACM Trans. Networking*, 1997.
[14] C. Labovitz, A. Ahuja, A. Abose, and F. Jahanian, "An experimental study of delayed Internet routing convergence," in *Proc. ACM SIGCOMM*, 2000.
[15] N. Feamster, D. G. Andersen, H. Balakrishnan, and M. F. Kaashoek, "Measuring the effects of Internet path faults on reactive routing," in *Proc. ACM SIGMETRICS*, 2003.
[16] Z. Mao, R. Bush, T. Griffin, and M. Roughan, "BGP Beacons," in *Proc. ACM Internet Measurement Workshop*, October 2003.
[17] "PlanetLab," `http://www.planet-lab.org`.
[18] "Happy Packets Data," `http://bbgp-slave.uoregon.edu/~agthorr/graphs/index.html`.
[19] T. G. Griffin, F. B. Shepherd, and G. Wilfong, "Policy disputes in path vector protocols," in *Proc. ICNP*, 1999.
[20] T. G. Griffin and G. Wilfong, "An analysis of BGP convergence properties," in *Proc. ACM SIGCOMM*, 1999.