Revisiting BGP Churn Growth

Ahmed Elmokashfi
Simula Research Laboratory
ahmed@simula.no

Amogh Dhamdhere
CAIDA / UC San Diego
amogh@caida.org

ABSTRACT

In the mid 2000s there was some concern in the research and operational communities over the scalability of BGP, the Internet's interdomain routing protocol. The focus was on *update churn* (the number of routing protocol messages that are exchanged when the network undergoes routing changes) and whether churn was growing too fast for routers to handle. Recent work somewhat allayed those fears, showing that update churn grows slowly in IPv4, but the question of routing scalability has re-emerged with IPv6. In this work, we develop a model that expresses BGP churn in terms of four measurable properties of the routing system. We show why the number of updates normalized by the size of the topology is constant, and why routing dynamics are qualitatively similar in IPv4 and IPv6. We also show that the exponential growth of IPv6 churn is entirely expected, as the underlying IPv6 topology is also growing exponentially.

Categories and Subject Descriptors

C.2.2 [**COMPUTER-COMMUNICATION NETWORKS**]: Network Protocols – *Routing Protocols*

General Terms

Measurement

Keywords

BGP; routing dynamics; scalability

1. INTRODUCTION

The Internet has sustained tremendous growth over the last couple of decades, but not without concerns about the scalability limitations of underlying protocols [17]. During 2005, the daily rate of BGP updates seen by a measurement point inside the Telstra network (AS1221) almost doubled. The number of IPv4 prefixes, meanwhile, grew by only 18% [13], suggesting that router hardware would need significant upgrades within 3-5 years in order to cope with projected churn. On the contrary, our later study using more monitors and spanning a longer time frame (seven years) revealed that after filtering out temporary artifacts, BGP churn in the IPv4 topology grows more slowly than the number of IPv4 prefixes [7]. This slow growth was also later confirmed by others [11].

Studying churn in the IPv6 topology has not received much attention, mostly due to the fact that IPv6 deployment is still in its nascent stages. Geoff Huston presented the first study comparing update churn in IPv4 and IPv6 in late 2011 [12]. He found that while churn in IPv4 does indeed appear "flat", it increases exponentially in IPv6. That article further speculated on the reasons why churn in IPv4 and IPv6 shows such fundamentally different characteristics, and raised concerns about the future scalability of the developing IPv6 ecosystem. Our recent work [5] showed that BGP dynamics in IPv4 and IPv6 are qualitatively similar when seen in relation to the size of the underlying topology – the magnitude of update churn normalized by the size of the underlying topology is *constant* for both IPv4 and IPv6.

Despite this prior empirical work on measuring and characterizing BGP churn (see Section 2 for more related work), we lack a model that can express the aggregate amount of BGP churn in terms of a few easily measurable properties of the topology and routing system. Such an explanatory model is useful to predict how BGP churn will evolve in the future, given that the interdomain topology, routing policies, and traffic characteristics are continuously changing. For example, with the impending exhaustion of IPv4 addresses and the advent of IPv4 transfer markets, routing tables may become more fragmented, and routing policies more granular. We would like to be able to predict the impact of such changes on BGP churn, but currently do not have the tools to do so.

In this paper, we develop a model that expresses BGP update churn as a function of certain measurable factors such as the average AS path length of the underlying topology, the fraction of prefixes announced by an AS that experience routing updates simultaneously, and the likelihood of observing routing updates to a given AS. Our model illustrates why routing dynamics are qualitatively similar in both IPv4 and IPv6. We believe that this model can help to track and explain the evolution of update churn in both IPv4 and IPv6 in the future. As the adoption of IPv6 is likely to be an ongoing process for years to come, a continuing look at the properties of BGP churn in IPv6, along with explanatory models, would help reveal insights into how the transition is progressing.

2. RELATED WORK

The characteristics of BGP churn and its impact on routing stability and scalability have been the subject of much research over the last decade and a half. Seminal papers by Labovitz et al. [14, 15] studied pathologies of updates and BGP convergence. Subsequent work investigated BGP path exploration and the activity patterns of prefixes [18, 19]. Geoff Huston has been monitoring and reporting on BGP churn and instabilities via his website and blog [10]. In 2006, Huston and Armitage warned of an alarming growth in BGP update churn [13]. Later, using data from more vantage points, Huston revised the growth trend to a slow (linear) growth [11]. Concerns about the inability of BGP to cope with increasing churn [17] and the impact of routing updates on router CPU utilization [1] spurred some of our own previous research [6, 7] and other efforts [3, 12] on characterizing the evolution of update churn. Our findings indicate that the sustained level of churn grows at a pace slower than the routing table size, and shows the same trend as the growth in the number of ASes. In 2007, Li et al. [16] revisited BGP update dynamics, comparing with a study conducted a decade earlier. They found that BGP update dynamics showed fewer pathologies as compared to the previous study. These measurement studies have revealed insights into BGP update churn and its evolution, but have not attempted to model the underlying factors that affect the amount of churn in the routing system.

In the area of modeling BGP dynamics, there is much work on modeling BGP convergence (see [8] and related references). More recently, Valler et. al [20] modeled the propagation of BGP instability, exploring the interaction between BGP dynamics and topological amplification. Zhao et. al [23] proposed an ON/OFF model which they used to classify bursts of BGP updates as "stable routing changes" or route flapping. Zhao et al. [24] developed a model that captures the impact of the location of failures on the resulting routing dynamics. The goal of our paper is not to capture BGP dynamics and their propagation on short timescales; instead, we attempt to model the sustained level of update churn, and how this depends on various properties of the routing system.

3. DATASETS

Our analysis of routing dynamics and network sizes of the IPv4 and IPv6 infrastructures is based on BGP updates and routing table dumps collected by the Routeviews project. Routeviews collectors run BGP sessions with routers (or *monitors*) in many networks. Each monitor sends a BGP update to the collector every time there is a change in the preferred path from the monitor to a destination prefix. We use update traces from two Routeviews collectors: Routeviews6 for IPv6 data and Oregon-IX for IPv4 data. The IPv4 updates span the period from Jan 1, 2003 to Feb 16, 2012; the IPv6 updates span the period from May 7, 2003 through Feb 16, 2012.

We analyze data from monitors in five networks that contributed both IPv4 and IPv6 routing data throughout the study period: AT&T (AS7018), Hurricane Electric (AS6939), NTT-America (AS2914), Tinet (AS3257), and IIJ (AS2497). All these networks are large transit providers, except IIJ which is a small transit provider. We have also looked into networks of different types (e.g., APAN which is a Content/Access/Hosting provider according to our classification) and they showed similar overall trends as our five monitors. We cannot include graphs from all monitors due to space constraints. AT&T's IPv4 monitor was unavailable for three months in 2003, and its IPv6 monitor was unavailable between May 2005 and May 2007, while Tinet's IPv6 monitor was unavailable between June 2008 and June 2010, causing the observed gaps in the time series for these monitors. In some cases, the IP address of a monitor changed during our study period. We identified the corresponding IP addresses and concatenated the update time series after confirming that they correspond to the same actual monitor.

If the multi-hop BGP session between a monitor and the collector is broken and re-established (session reset), the monitor re-announces all its known paths, producing large bursts of updates. This is a local artifact of the Routeviews measurement infrastructure, and does not represent genuine routing dynamics. We use the method developed by Zhang et al. [22] to identify and remove updates caused by session resets.

We use topology and AS classification data to evaluate our model in Section 5. This data comes from our previous work [4], which used a machine learning decision tree classifier to classify ASes into four classes in terms of their function and business goals – Enterprise Customer (EC), Small Transit Provider (STP), Large Transit Provider (LTP), and Content/Access/Hosting Provider (CAHP).

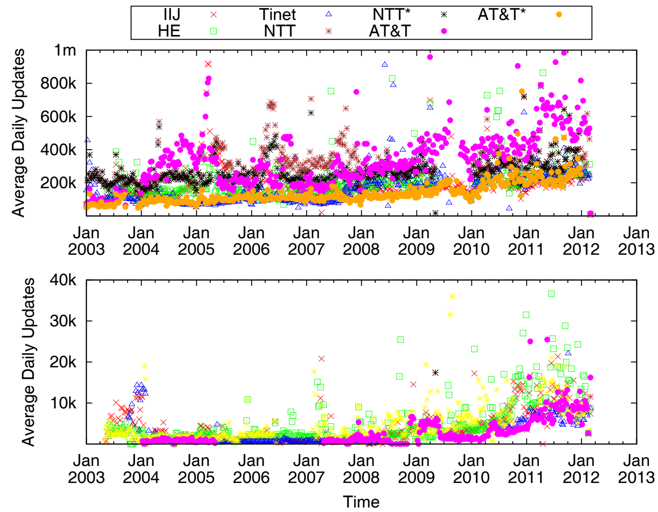

Figure 1: Churn over time in IPv4 (top) and IPv6 (bottom), measured as the average number of daily updates received from each monitor over a week. The AT&T* and NTT* time series are obtained by removing updates related to non-stationary periods from the AT&T and NTT raw time series respectively, as described in [7].

4. COMPARING CHURN EVOLUTION IN IPV4 AND IPV6

Figure 1 shows the raw BGP churn time series for IPv4 and IPv6 respectively. Each point on the plot is the average number of daily updates received from that monitor over the course of a week. We also examined hourly and daily aggregates of updates, and found the same growth pattern persisted at these different granularities, albeit noisier [1]. Our goal is to study the evolution of update churn, but long-term trends in IPv4 churn can be difficult to spot due to the presence of non-stationary periods, causing jumps in daily churn that lasted for weeks or months, such as for AT&T from Jan'04 to Jun'05 and for NTT from Sep'07 to Dec'07. These periods are usually caused by misconfigurations or other monitor-specific events (e.g., flapping), and are not related to the long-term evolution of churn. We reuse results from our previous work to analyze the AT&T and NTT time series after removing updates related to these periods. The filtered time series are shown in Fig 1 (top panel) as AT&T* and NTT*. We refer the reader to [7] for a detailed discussion about the non-stationary periods and their removal from the AT&T time series.

We observe two distinct differences between IPv4 and IPv6 BGP dynamics. First, the growth trends are qualitatively different in IPv4 and IPv6. Second, IPv6 churn exhibits more similarity across monitors than IPv4 churn. We quantify these differences in the remainder of this section.

Churn growth trends

To examine the difference between the growth of churn in IPv4 and IPv6, we divide each time series into two segments. The first starts on January 1, 2004 and ends on June 30, 2008, while the second starts on July 1, 2008 and ends on December 31, 2011. We then compute the ratio of average daily churn in the last six months of a segment to that in the first six months of the same segment. We denote this ratio as G. The IPv4 time series shows similar G values for both segments. Averaging across monitors, G is 1.45 in the first segment and 1.47 in the second, indicating that the trend

[1]We show the weekly average of daily churn to avoid visual clutter.

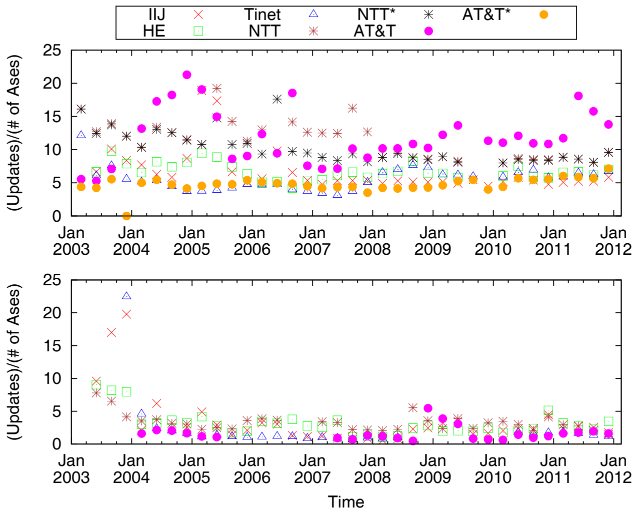

Figure 2: Churn growth in relation to topology size in IPv4 (top) and IPv6 (bottom). BGP churn, in both IPv4 and IPv6, grows linearly with the number of ASes.

of churn growth in IPv4 has not changed significantly between the first and second half of the measurement duration. In IPv6, on the other hand, G is ≈ 1 in the first segment (denoting no growth), while it is 9.67 in the second segment, almost an order of magnitude higher. This confirms that the trends in IPv6 churn evolution are qualitatively different from those in IPv4.

Similarity across monitors

We use Kendall's τ rank correlation coefficient [9] to estimate the extent of correlation in daily churn between all pairs of vantage points in IPv4 and IPv6 respectively. The IPv4 pairs exhibit little correlation; $\tau \leq 0.4$ for all pairs. The IPv6 pairs, however, demonstrate strong positive correlation with $\tau > 0.5$ for seven out of the ten pairs. This is likely because the IPv6 topology is much smaller and thus provides less isolation (i.e., routing changes will have a larger scope of impact).

Churn as a function of topology size

Churn in IPv4 and IPv6 grows at different rates, but so do the underlying topologies. The IPv4 topology now grows linearly, while the IPv6 topology grows exponentially, as shown in [5]. To understand how churn has evolved with respect to network size, we track growth in the number of updates, normalized by the size of the underlying AS topology. To calculate this metric, we bin the total number of updates per day into three-month windows, find the median daily churn for each window, and divide it by the average number of ASes in the graph during that time window. Figure 2 plots this metric for IPv4 (top) and IPv6 (bottom). In IPv6 this number has remained mostly stable since Jan 2004 at $\approx$ 2 to 3 updates per AS.

In IPv4, except for the AT&T and NTT monitors, this metric stabilized in 2006 at ≈ 5 to 8 updates per AS. Other monitors that peer with the Oregon-IX collector show similar behavior. We confirmed that the discrepancy in AT&T and NTT is due to the presence of non-stationary periods as mentioned previously. The AT&T* time-series, that is obtained after filtering out these periods, shows a ratio of 5 updates per AS. Similarly, the filtered NTT* time series exhibits a stable ratio between 7 and 8 starting from 2006. We emphasize that this stability does not imply that different vantage points experience the same level of churn. It means that churn seen by different vantage points grows linearly as a function of topology size. In fact, different vantage points experience varying levels of churn depending on their topological location (i.e., different networks have different sets of neighbors) and configurations. The ratios in Figure 2 illustrate this fact. For instance, given that there were about $38K$ ASes at the end of 2011, a variation in the number of updates per AS between 5 and 8 means that our vantage points experienced daily churn levels between $190K$ updates and $304K$ updates.

Our analysis reveals that *BGP dynamics measured in relation to the topology size in IPv4 and IPv6 are qualitatively similar, and the number of updates grows at the same rate as the number of ASes.* However, we still lack an explanation of why this is so, and of why the multiplicative constants turn out to be between 5 and 8 for IPv4 and between 2 and 3 for IPv6. Hence, a model that breaks down the observed churn into easily measurable parameters can be instrumental for answering these questions. Such a model is important in two respects. First, it acts as a tool for estimating the expected sustained level of churn given the values of a few topological and routing parameters. Hence, it can help predicting churn levels when these parameters vary, as well as detecting anomalous jumps in the sustained level of churn. Second, it can be used as an educational tool for explaining the elementary factors behind BGP churn and their contribution to the overall picture. We develop such a simple model for BGP churn in the next section.

5. DISSECTING BGP CHURN

BGP churn results from a complex interaction that involves BGP configuration, the interdomain topology, and routing incidents. To understand the dependence of churn on these properties, we develop a simple approach that dissects churn into a set of constituents.

At a high level, BGP routing is a collection of prefixes and ASes. In terms of dynamics, however, we can view it as a collection of activity units. An activity unit is a single prefix or a group of prefixes originated by the same AS and share a common routing fate, i.e., whenever a BGP update is received for one prefix in the activity unit, updates for all other prefixes in that activity unit are received as well. An activity unit is different from a BGP policy atom [2], which is a group of prefixes that share the same AS path. In fact, all activity units are atoms but not vice versa; prefixes may share the same AS path but differ in the router level path and consequently have different fates. To estimate the average number of updates $\overline{U(T)}$ in a time window of length T, we need to quantify the following:

1. The average number of activity units A that undergo a change in T.
2. The average number of times n an activity unit experiences a change in T.
3. The average number of updates U_A observed following a routing change to an activity unit.

$\overline{U(T)}$ can be expressed formally as

$$\overline{U(T)} = AnU_A \tag{1}$$

In the remainder of this section, we investigate and quantify each term that contributes to $\overline{U(T)}$ in Equation 1 above.

5.1 The activity unit

The average granularity of an activity unit, denoted as μ, represents the average number of prefixes of the same origin AS that share a common fate. We believe that μ changes over relatively large time spans (months or years), because it is related to routing practices followed by network operators. Therefore, we measure μ using four snapshots from each year in our study period. In every year, we use all BGP updates observed in February, May, August,

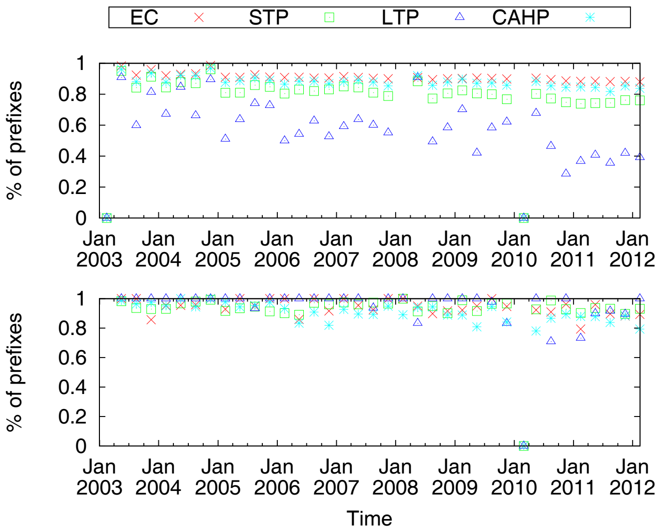

Figure 3: The average size of the activity unit from the perspective of the IIJ monitor. IPv4 (top) and IPv6 (bottom). In IPv4, except for LTPs, ASes of different types mostly act as a single unit. In IPv6, ASes of all types act a single unit.

and November. To measure μ, we start by grouping together updates for the same prefix that are temporally close to each other, and refer to these as *events*. The rationale is to cluster updates that are likely triggered by the same underlying routing change. We employ the prefix event definition suggested by Wu et. al [21]: Two updates for the same prefix are part of the same event if the time interval between them is not more than 70 seconds.

After identifying all single prefix events, we investigate the temporal correlation between events affecting prefixes that are originated by the same AS. Assume that we record prefix events that affect prefixes $(P_1, P_2, P_3, ..., P_n)$, which belong to the same AS, at times $(t_1, t_2, t_3, ..., t_n)$ in an ascending order. We traverse the event list and assign every two consecutive prefixes (P_{i+1}, P_i) to the same activity set if $(t_{i+1} - t_i) \leq Th_G$ seconds. We set Th_G to 60 seconds (i.e., about two MRAI timer intervals). Smaller values may lead to false negatives, since the MRAI timer can increase the spacing of outgoing updates that are initially spaced by a few seconds to a full MRAI timer interval. Setting Th_G to a bigger value in the order of minutes can result in false positives. This process leaves us with multiple activity sets. As a first step, we remove all redundant sets. The remaining sets, however, may not be disjoint. For instance, an IGP change inside the monitor AS can trigger a change in preferred egress points. This leads to the re-announcement of all prefixes belonging to the same origin AS at about the same time. Consequently, our approach will place all these prefixes in one activity set. We therefore, follow a conservative approach to prune superfluous correlations. Our approach compares each activity set to all other sets to identify common prefixes. We then remove these common prefixes from the larger activity set if their fraction is less than Th_U of its size. More formally, when comparing two sets S_i and S_j where $|S_i| > |S_j|$, we remove all prefixes that belong to $P = \{p | p \in S_i \text{ and } p \in S_j\}$ from S_i if $\frac{|P|}{|S_i|} \leq Th_U$, otherwise we remove them from S_j; we set Th_U to 0.8. In other words, the common prefixes are removed from the bigger activity set, unless they constitute 80% of its membership. Hence, our approach is very conservative in grouping prefixes into the same activity set. This way we obtain disjoint activity sets and minimize the membership of each set by limiting it to the smallest number of prefixes that appear to change simultaneously.

We experiment with different values of Th_G and Th_U to check whether our estimated activity set sizes are sensitive to the values chosen above (i.e. $Th_G = 60$ and $Th_U = 0.8$). While fixing Th_U to 0.8, we set Th_G to a range of values between 5 and 80 seconds. Across monitors and snapshots and for the largest class of ASes (i.e., ECs) the average number of prefixes in the largest activity set does not change by more than 1.8% and 0.2% when setting Th_G to 5 seconds and 80 seconds respectively. We also fix Th_G to 60 seconds and experiment with two values of Th_U, 0.7 and 0.9. The largest fraction of prefixes that are active together for ECs does not change by more than 0.023% and 0.025% when setting Th_U to 0.7 and 0.9 respectively. We, however, expect that the choice of thresholds would have a larger effect on ASes known to have sophisticated routing policies such as LTPs. This is indeed the case, although to a small extent. For LTPs, The average number of prefixes in the largest activity set varies by 3.9% and 0.9% when setting Th_G to 5 seconds and 80 seconds respectively. The robustness of our grouping to different threshold values suggests that prefixes of the same AS, at least for ECs, usually share the same routing fate.

After grouping prefixes into different activity sets, we calculate for each origin AS i the largest fraction of prefixes that frequently change together, denoted as f_i. Next, we compute the average of f_i across all ASes of the same type. The top panel in Figure 3 illustrates the average maximum fraction of prefixes that change together for different AS types in the IPv4 topology, while the lower panel shows the same for IPv6. A larger fraction indicates a coarser granularity (i.e., more prefixes originated by the same AS change together). Both plots are from the perspective of the IIJ monitor; other monitors show similar results, which we omit due to space constraints. We observe that for IPv4, μ is becoming *finer* (decreasing fraction in Figure 3) over time across all AS types. Also, μ becomes coarser (increasing fraction in Figure 3) as we move from Large Transit Providers to Enterprise Customers. The finer granularity for LTPs is expected since these networks serve a large number of non-BGP speaking customers, operate in larger geographical areas, and may advertise different prefixes in different regions. It is thus natural that a smaller fraction of their prefixes share the same fate. Most (92%) ASes in the Internet according to our AS classification are ECs, and hence we can safely approximate μ using the f_i for ECs. The upper panel in Figure. 3 demonstrates that f_i for ECs is ≈ 0.95 – meaning that virtually all prefixes originated by an EC change together. We also investigated the distribution of f_i for ECs. The median is one throughout our study period, and the first quartile is over 0.9 in two thirds of the snapshots and always larger than 0.81. Furthermore, the lower panel in Figure. 3 shows for IPv6 and across AS types that prefixes originated from the same AS act, on average, as a single unit. This is plausible since IPv6 adoption is still in early stages with little traffic, which nullifies the need for sophisticated routing strategies (e.g., traffic engineering). Based on these observations we conclude that virtually all prefixes originated from the same AS, in both routing systems, act on average as a single unit. Hence, *the total number of activity units is equal to the number of ASes.*

5.2 How many activity units?

Having established that prefixes originated from the same AS usually act as a single activity unit, the next question is how many ASes are seen active in a time window of length T. In other words, what is the average number of unique origin ASes we receive routing updates to in T. To answer this question, we analyze BGP updates between February 1st and May 1st in each year of our study period [2]. We identify the origin AS of each updated prefix and

[2] For 2012 we consider the period from mid-November 2011 to the

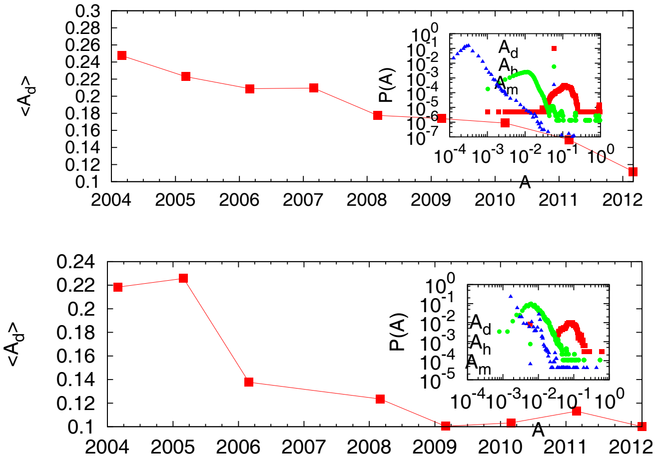

Figure 4: The average fraction of origin ASes with updated prefixes over time in IPv4(top) and IPv6 (bottom), from the perspective of the HE monitor. The inset plots show the distribution of this fraction measured per day A_d, hour A_h, and minute A_m in a typical month.

then count the unique number of origin ASes seen every minute, hour, and day. We hypothesize that this fraction will decrease over time due to topological densification [4] which makes a route less central, limiting the impact scope of routing changes.

Figure 4 shows the evolution of the average daily fraction of uniquely updated origin ASes for IPv4 (top panel) and IPv6 (lower panel). The mean shows a clear decreasing trend, confirming our hypothesis. Furthermore, the IPv4 and IPv6 values are comparable, suggesting that there are no fundamental differences in the likelihood that a destination experiences a routing change. Each monitor shows a decreasing trend, albeit with different rates of decrease.

The inset plots show the distribution of uniquely updated origin ASes measured per minute (A_m), hour(A_h), and day (A_d), respectively. A_m demonstrates greater variability over several orders of magnitude. The variability decreases as we increase T. A_d for instance, exhibits clear symmetry around its mean with minimal variability. These distributions remain stable over the years and only shift towards the left as a result of a decreasing mean.

5.3 How many times is an AS active in T?

The second constituent of our model is the average number of times n that an AS is active in T. To estimate n, we count the number of prefix events per day for prefixes originated by the same AS for every day between February 1st and May 1st in each year in our study period [2]. We then divide by the average number of prefixes per AS to get an estimate of how many times an origin AS is seen per day, and next average across all origin ASes seen in that day. We limit this analysis to ECs because of their coarse activity granularity as shown in Section 5.1. Finally, we calculate the median of average daily values. The median is more suitable here as opposed to the mean since it is more robust to outliers caused by highly active prefixes [18].

In IPv4, the number of times an origin AS is seen per day is between 1.1 and 1.4. In IPv6, however, this number fluctuates in a wider range between 2 and 7.5. This implies that IPv6 routing is ≈ 6 times less stable than IPv4. Interestingly, this suggests that the IPv6 routing system, though qualitatively similar to that of IPv4, *is less stable*. A plausible explanation is that the IPv6 deployment for some organizations may be experimental ("non-production"), causing a larger likelihood of repeated failures for those ASes. We plan to investigate this point more closely in future work.

end of our study period in mid-February 2012.

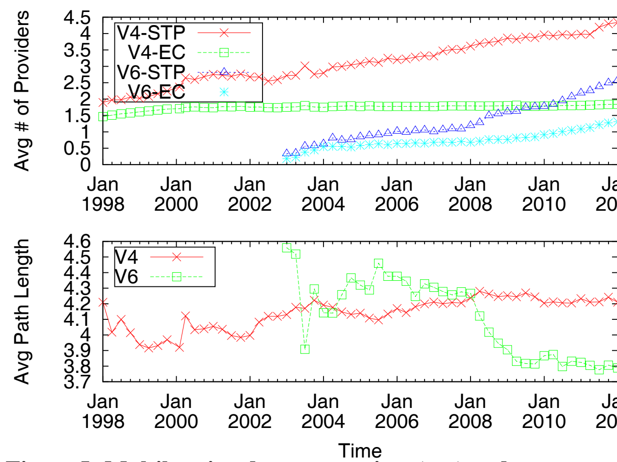

Figure 5: Multihoming degree over time (top) and average path length over time (bottom).

5.4 Activity magnitude

The last variable in our model is the average number of updates U_A observed following a routing change to an activity unit, i.e., how many updates do we expect to see following a routing change towards a destination AS? This question can be split into two parts. First, how many prefixes does an AS announce on average? Second, how many updates do we observe on average following a single prefix event (U_E)? Multiplying these two quantities answers our initial question.

Prefixes per AS

We find that the average number of prefixes per AS grows slowly in both routing systems. In IPv4, this number grew by 24% from 7.8 in 2004 to 9.7 in early 2012. In the same period, the routing table size and the number of ASes grew by 204% and 145% respectively. Interestingly, the average number of prefixes per AS shows similar growth in IPv6, where it grew by 23% from 1.3 in 2004 to 1.6 in 2012. The routing table size and the number of ASes grew during the same period by 1969% and 1639% respectively. The number of prefixes and the number of ASes grow linearly in IPv4, and both these quantities grow super-linearly in IPv6. This causes the observed slow growth in the average number of prefixes per AS for both IPv4 and IPv6.

Average number of updates per prefix event (U_E)

The average number of updates per prefix event U_E is directly related to the path vector nature of BGP, unlike the granularity of the activity unit and the average number of prefixes per AS, which depend on how networks are engineered and how IP addresses are allocated. Before measuring U_E, we develop a simple expression for the magnitude of U_E in terms of the properties of the underlying topology.

Assume that we have a routing monitor at the top of the AS-level hierarchy observing the activity of a prefix originated by a stub AS [3]. The monitor and the origin AS are, on average, l hops apart, where l is the average AS path length seen from the monitor. Assuming for simplicity a strict hierarchical structure, there are $l-1$ transit ASes between the origin and the monitor ASes. The average number of paths available at the monitor can be approxi-

[3]Most of the route monitors we have used to measure update churn in Section 4 are indeed high-degree ASes, and are thus near the top of the AS-level hierarchy

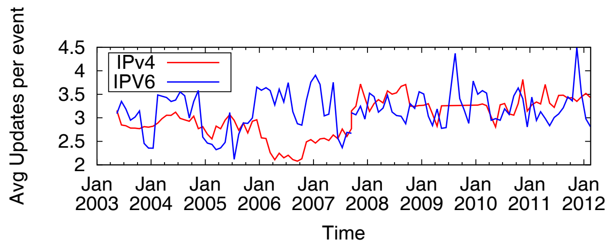

Figure 6: Average updates in a prefix event from the perspective of the HE monitor

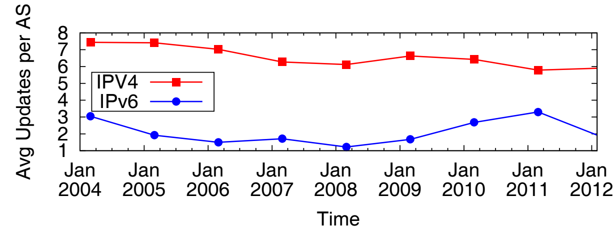

Figure 7: Validating churn dissection

mated as

$$Pa = d_C {d_M}^{l-1} \quad (2)$$

where d_C is the average provider degree of stubs and d_M is the average provider degree of transit providers.

When a route to a prefix changes, BGP starts exploring available alternate paths until it converges to a new route, or completely withdraws the affected prefix. The alternate paths are ranked based on their preference and explored accordingly. A low ranked path may never be explored if the decision process picks a higher ranked alternative or if it is withdrawn earlier during the convergence process. Thus, due to the path vector nature of BGP, the likelihood of exploring a path decreases as its rank increases. To capture this effect, we assume that the likelihood to explore the i^{th} path (L_E) is inversely related to its rank. More precisely, L_E is equal to C_i/i, where C_i is a multiplicative factor that is function in the rank of the i^{th} path. Note that the expected number of explored path is directly related to U_E. Equation 3 gives an estimate of U_E seen by our monitor above.

$$U_E = \sum_{i=1}^{Pa} \frac{C_i}{i} \quad (3)$$

The contribution of the numerator C_i varies depending on the type of routing change. In transient changes it diminishes for low ranked paths, because BGP quickly converges to an alternative path that is highly ranked. It is, however, more important in routing changes that end with a withdrawal of the affected prefix, since many alternative paths are likely to be explored. Given that most routing changes do not lead to the complete withdrawal of a prefix [19], we assume that the likelihood to explore a path is mainly determined by its rank and thus ignore the contribution of C_i. The summation in Equation 3 is a partial sum of harmonic series that converges to $\approx ln(Pa) + \gamma$ for a large i, $\gamma = 0.5772$ is called the Euler's constant. Substituting Pa from Equation 2 gives:

$$U_E \approx ln(d_C) + (l-1)ln(d_M) + \gamma \quad (4)$$

The top panel in Figure 5 shows the evolution of provider degree for ECs (d_C) and STPs (d_M) in IPv4 and IPv6. The bottom panel illustrates the evolution of the average path length. The small measured value of d_C in both IPv4 and IPv6 suggests that it has limited influence on U_E. U_E grows as the logarithm of d_M, and d_M grew from 3 in 2003 to 4 in 2011. Consequently, the influence of d_M on U_E increased only marginally, from 1.1 to 1.3. Hence, U_E is mainly influenced by the average path length; $U_E \approx l - 1$.

Figure 6 shows the evolution of the measured average number of updates in a single prefix event from the perspective of the Hurricane Electric (HE) monitor. We omit the graphs for other monitors (which show similar values) due to space constraints. To focus on routing changes that involve path exploration, we only consider prefix events that include at least one AS-PATH change. In IPv4, this number mostly oscillates between 3 and 3.5 updates per prefix event. This matches our approximation, which predicts that the average number of updates should be one less than the average path length; figure 5 shows that the average IPv4 AS path length has been between 4 and 4.3 since 1998. The average number of updates in IPv6 has remained stable at around 3, while the average IPv6 AS path length was between 4 and 4.5 before 2008 and has dropped to around 3.8 since. The approximation $U_E \approx l - 1$ thus holds for both IPv4 and IPv6, indicating that the simplifications made above have little effect on the approximated U_E. Hence, when averaging over a large number of events, the likelihood to explore a path is mainly determined by its rank and that the contribution of the numerator C_i evens out.

5.5 Putting things together

In the previous subsections, we have measured the average granularity of the routing activity unit, then estimated and verified our model's constituents empirically. Next, we proceed to check whether our model captures the observed stability in the sustained level of churn with respect to the topology size. We start by setting the length of the time window T in Equation 1 to one day, then divide both sides of the equation by the number of ASes present in the routing system per day. The equation can be re-written as:

$$U_a(d) = \overline{A_d} n U_A \quad (5)$$

Where $U_a(d)$ is the number of daily updates per AS, $\overline{A_d}$ is the average daily fraction of uniquely updated origin ASes, n is the number of times an origin AS is seen active per day, and U_A is the average number of updates we expect to see following a routing change towards a destination AS. Note that we have empirically quantified $\overline{A_d}$ and n in Section 5.2 and Section 5.3 respectively. Furthermore, we have shown analytically in Section 5.4 that U_A can be expressed as $\overline{P_{as}}(l-1)$, where $\overline{P_{as}}$ is the average number of prefixes per AS and l is the average AS-PATH length. Based on this we can re-write Equation 5 above as follows.

$$U_a(d) = \overline{A_d} n \overline{P_{as}} (l-1) \quad (6)$$

Finally, we estimate $U_a(d)$ by substituting the measured values in the previous subsections back into Equation 6. Figure 7 shows the estimated U_a for IPv4 and IPv6. U_a is between 5.8 and 7.5 for IPv4, while it is between 1.5 and 3.2 for IPv6, closely matching the actual values measured in Section 4.

6. DISCUSSION AND CONCLUSIONS

The Internet's interdomain routing system has grown and evolved tremendously over the last two decades, increasing concerns about possible scalability limitations. Although worries about the scalability with respect to dynamics were shown to be overblown [7, 11], the reasons behind the slow growth of update churn remained unexplained. Also, BGP dynamics in the IPv6 Internet were largely not studied. We presented a model that dissects BGP update churn into its principal components. Our model is the first to express BGP churn in terms of four feasibly measurable properties of the routing system. Hence, it can be used both as an educational and measurement tool to explain and monitor the factors behind BGP churn and their interplay. Our findings confirm that IPv4 and IPv6 BGP dynamics are *characteristically similar*, in that the growth trends match those of the underlying topologies. But, the IPv6 Internet is less stable than IPv4; we see up to 6 times more routing events per origin AS per day in IPv6. Our model identifies several factors that affect the observed routing dynamics. It will be interesting to measure how these factors evolve, and then use our model to estimate the impact of that evolution on BGP churn.

Prefixes of the same AS mostly share the same fate. We find, empirically, that all prefixes for a majority of ASes share the same routing fate, thus acting as a single atomic unit. In the near future, however, IPv4 run-out and IP address transfer markets could make the routing system more fragmented, meaning that a smaller fraction of prefixes from the same AS may share routing fate.

Prefixes and ASes grow in a qualitatively similar fashion. The number of ASes and the number of prefixes grow linearly in the IPv4 routing system. Both numbers, however, grow super-linearly in the IPv6 routing system. This makes the average number of prefixes per AS increase at a pace order of magnitude slower than the number of prefixes and ASes. Consequently, the average magnitude of activity that an AS contributes remains small compared to the overall growth in the routing system. Prefix deaggregation and IP transfers may cause an increase in this metric.

The stability of the average AS path length, multihoming, and densification. The extent of BGP path exploration is mainly determined by the depth of the hierarchy. A stable average path length results in convergence sequences with a stable length. In other words, the prefix activity footprint remains, on average, invariant over time. The average AS path length (as measured from Routeviews/RIPE monitors) has largely been stable over time, presumably due to a *densification* process that increases the average degree in the interdomain topology [4]. Increasing multihoming – mainly in the core – increases the number of available alternate routes, and also limits the impact scope of routing changes. This is clearly captured by the decrease in the fraction of unique active ASes seen from our vantage points. The average path length, multihoming trends, and densification are the result of complex interconnection incentives of ASes which could change over time, thus affecting routing scalability.

Going forward, this work can be extended in several directions. Setting up a system to continuously monitoring the constituents of our model can help tracking and predicting changes in the sustained level of churn. Such changes might happen in the foreseeable future given the unanticipated effects of the impending exhaustion of the IPv4 address space. It is also interesting to investigate how the observed stability in churn with respect to the topology size changes if the properties of the routing system change. We also plan to study the origin of the measured difference in stability between IPv6 and IPv4.

Acknowledgments

We would like to thank Kc Claffy and the anonymous reviewers for their constructive comments. Ahmed Elmokashfi is supported by the Norwegian Research Council, grant number 209954/S10. Amogh Dhamdhere is supported by the National Science Foundation (NSF) under grant CNS-1111449.

7. REFERENCES

[1] S. Agarwal, C.-N. Chuah, S. Bhattacharyya, and C. Diot. Impact of BGP Dynamics on Router CPU Utilization. In *Proc. Passive and Active Measurement Conference*, 2004.

[2] A. Broido and k. claffy. Analysis of RouteViews BGP data: policy atoms. In *Network Resource Data Management Workshop*, Santa Barbara, CA, May 2001.

[3] L. Cittadini, W. Muhlbauer, S. Uhlig, R. Bush, P. Francois, and O. Maennel. Evolution of Internet address space deaggregation: Myths and reality. *IEEE Journal on Selected Areas in Communications*, 28(8), 2010.

[4] A. Dhamdhere and C. Dovrolis. Twelve years in the evolution of the Internet ecosystem. *IEEE/ACM Transactions on Networking*, 19(5), Oct. 2011.

[5] A. Dhamdhere, M. Luckie, B. Huffaker, A. Elmokashfi, E. Aben, and K. Claffy. Measuring the Deployment of IPv6: Topology, Routing and Performance. In *Proc. of ACM SIGCOMM IMC*, Nov 2012.

[6] A. Elmokashfi, A. Kvalbein, and C. Dovrolis. On the Scalability of BGP: the roles of topology growth and update rate-limiting. In *Proc. ACM CoNEXT*, Dec 2008.

[7] A. Elmokashfi, A. Kvalbein, and C. Dovrolis. BGP churn evolution: A perspective from the core. *IEEE/ACM Transactions on Networking*, 20(2), April 2012.

[8] T. Griffin and G. Wilfong. An Analysis of BGP Convergence Properties. In *Proc. ACM SIGCOMM*, 1999.

[9] M. Hollander and D. A. Wolfe. *Nonparametric statistical methods*. Wiley, second edition, 1999.

[10] G. Huston. The BGP instability report. http://bgpupdates.potaroo.net/.

[11] G. Huston. BGP in 2009 (and a bit of 2010). Presentation at ARIN XXV meeting, 2010.

[12] G. Huston. The BGP World is Flat, Nov. 2011. http://www.potaroo.net/ispcol/2011-12/flat.html.

[13] G. Huston and G. Armitage. Projecting future IPv4 router requirements from trends in dynamic BGP behaviour. In *Proc. ATNAC*, Australia, Dec 2006.

[14] C. Labovitz, A. Ahuja, A. Bose, and F. Jahanian. Delayed Internet Routing Convergence. *IEEE/ACM Transactions on Networking*, 9(3), June 2001.

[15] C. Labovitz, G. R. Malan, and F. Jahanian. Origins of Internet routing instability. In *Proc. IEEE INFOCOM*, 1999.

[16] J. Li, M. Guidero, Z. Wu, E. Purpus, and T. Ehrenkranz. BGP routing dynamics revisited. *Computer Communications Review*, apr 2007.

[17] D. Meyer, L. Zhang, and K. Fall. Report from the IAB workshop on routing and addressing. RFC 4984, 2007.

[18] R. Oliveira, R. Izhak-Ratzin, B. Zhang, and L. Zhang. Measurement of highly active prefixes in BGP. In *Proc. IEEE GLOBECOMM*, 2005.

[19] R. Oliveira, B. Zhang, D. Pei, R. Izhak-Ratzin, and L. Zhang. Quantifying Path Exploration in the Internet. In *Proc. ACM SIGCOMM IMC*, 2006.

[20] N. Valler, M. Butkiewicz, B. Prakash, and M. Faloutsos. Non-binary information propagation: Modeling BGP routing churn. In *Proc. IEEE INFOCOM Compter Communications Workshop*, 2011.

[21] J. Wu, Z. M. Mao, J. Rexford, and J. Wang. Finding a needle in a haystack: pinpointing significant BGP routing changes in an IP network. In *Proc. USENIX NSDI*, 2005.

[22] B. Zhang, V. Kambhampati, M. Lad, D. Massey, and L. Zhang. Identifying BGP routing table transfers. In *Proc. ACM SIGCOMM workshop on Mining network data*, 2005.

[23] X. Zhao, D. Massey, M. Lad, and L. Zhang. On/Off Model: A New Tool to Understand BGP Update Burst. *University of California, Los Angeles technical report*, 2004.

[24] X. Zhao, B. Zhang, D. Massey, A. Terzis, and L. Zhang. The impact of link failure location on routing dynamics: A formal analysis. In *Proc. ACM SIGCOMM Asia Workshop*, 2005.