BGP churn evolution: A perspective from the core

Ahmed Elmokashfi
Simula Reserach Laboratory

Amund Kvalbein
Simula Reserach Laboratory

Constantine Dovrolis
Georgia Institute of Technology

***Abstract*—The scalability limitations of BGP have been a major concern in the networking community lately. An important issue in this respect is the rate of routing updates (churn) that BGP routers must process. This paper presents an analysis of the evolution of churn in four networks in the backbone of the Internet over the last six years, using update traces from the Routeviews project. The churn rate varies widely over time and between networks, and cannot be understood through "black-box" statistical analysis. Instead we take a different approach with a focus on investigating the underlying reasons for BGP churn evolution. Through our analysis we are able to identify and isolate the main reasons behind many of the anomalies in the churn time series. We find that duplicate announcements is a major churn contributor, and responsible for most large spikes in the churn time series. Other intense periods of churn are caused by misconfigurations or other special events in or close to the monitored AS, and hence limiting these is an important mean to limit churn. We then analyze the remaining "baseline" churn, and find that it is increasing with a rate much slower than the increase in the routing table size.**

I. INTRODUCTION

The deployment of the BGP routing protocol has sustained tremendous growth over the last couple of decades and it is arguably one of the main technological reasons behind the Internet's success. Lately, however, there are significant concerns about the *scalability of BGP interdomain routing*. These concerns focus either on the growing routing table size (number of routable prefixes) or on BGP dynamics and instability (also known as "churn") [13]. Both factors are important, especially for routers at the core of the Internet. The growing size of the routing table requires increasingly larger fast memory, but it does not necessarily slow down packet forwarding as long as address lookups are performed using TCAMs or constant-time longest-prefix matching algorithms [17].

Churn, on the other hand, is a more serious concern because processing BGP updates can be computationally intensive (updating routing state, generating more updates, checking import/export filters), and it can trigger a wide-scale instability. If the current best route to a destination is modified, the global RIB and the FIBs on the line cards need to be updated. To make things worse, routing updates are known to be very bursty, with peak rates several orders of magnitude higher than daily averages. When the rate of updates becomes too high, the fear is that there will be (or there are already) periods when routers will be unable to maintain a consistent routing table.

An earlier study by Huston and Armitage reported an alarming growth in churn [7]. During 2005, the daily rate of BGP updates observed by a router in AS1221 (Telstra) almost doubled, while the number of prefixes grew by only 18%. Based on these measurements, the authors projected future churn levels and concluded that current router hardware will need significant upgrades in order to cope with churn in a 3-5 years perspective. It was this study that largely motivated our work.

Specifically, in this paper we present a longitudinal study of BGP churn spanning a longer time frame (6 years) and more monitors (routers in 4 tier-1 ISPs) than previous studies. Generally, the churn time series is very noisy, dominated by frequent large spikes, and "level shifts" that last for several weeks or even months. There are periods in which churn is slowly increasing, others in which it is decreasing, and major differences between monitors. One option could be to characterize the evolution of churn using "black-box" statistical or time series analysis methods. That approach would answer questions about the correlation structure and the marginal distribution of the underlying time series, attempting to fit the data in a standard time series model (e.g., ARIMA). That descriptive method, however, would not be able to *explain* what causes spikes, level shifts or trends in BGP churn.

We prefer, instead, to take Tukey's exploratory data analysis approach that focuses on the *causes behind the observed phenomena.* This direction allowed us to identify the underlying cause of most large spikes and of major level-shifts. These events are largely caused by anomalies or mistakes in BGP implementations or configuration, and they are mostly local effects at the measured network. This also explains why different networks and monitors experience very different churn. We also observe that the use of the rate-limiting MRAI timer can play an important role in reducing certain pathological sources of churn, but it is not able to protect a router from large spikes and level shifts. After removing updates attributed to such local effects, we find the remaining time series to be much smoother and show more consistency across monitors. Furthermore, we analyzed this remaining "baseline" time series for the presence of long-term growth. That analysis shows that this filtered version of churn has increased during the last six years by about 20-80%, depending on the monitor, which is significantly less than the growth rate of routable prefixes during the same time period (about 100%). We also find that the number of updates per prefix per day is decreasing over time, indicating that, on the average, the stability of a prefix has improved. We also examine the daily peak churn rate, measured as the busiest one-minute period in that day, and find that it would be reduced by more than an order of magnitude if updates that are redundant or caused by certain anomalies were filtered out.

The rest of this paper is organized as follows. Sec. II gives an overview of different underlying factors that create BGP churn. Sec. III describes our dataset. Sec. IV shows the major trends in BGP routing during our study period. Sec. V focuses on the role of the MRAI timer and of duplicate updates on the churn time series. Sec. VI looks at the role of large routing events that affect many prefixes. Sec. VII investigates major level shifts in the churn time series. Sec. VIII analyzes the growth of the baseline churn that remains after we remove pathologies, large spikes and level shifts. Sec. IX gives an overview of related work, and Sec. X concludes this work.

II. CHURN GROWTH FACTORS

Several different factors can influence BGP churn. First, it is expected that the rate of BGP updates a router receives will increase with the number of routable destination prefixes. Roughly speaking, each prefix corresponds to a destination network. If these destination networks fail and recover independently and with the same probability, we would expect a linear relation between the size of the routing table and churn. Of course this is a very simplistic model, but still we can expect that there is a positive correlation between the number of routable prefixes and churn.

The observed churn will also depend on the routing activity of individual prefixes at their origin AS. Over the past few years, it has become increasingly common for stub ASes to be multihomed to several providers [3]. Multihoming enables load-balancing by selectively announcing different prefixes to different providers. As this practice gradually becomes more common, we can expect that it contributes to increasing churn when a network destination becomes unreachable.

Another source of churn is routing events taking place in or between transit ASes. Such events include link failures (physical failures, router reboots, etc), policy changes that result in new preferred routes, or changes in the IGP or iBGP configuration of a transit AS. Importantly, these operations often affect a large number of prefixes at the same time. The amount of churn observed at a router after such events will also depend on the the topology and policies.

Topological properties of the AS-level Internet graph will also affect the churn rate [4]. Increased multihoming in the Internet increases the churn generated when a destination prefix is announced or withdrawn from the source AS. On the other hand, increased connectivity can reduce the impact of failures, if a local alternative is available.

Finally, there are BGP mechanisms and parameter settings that can reduce the observed churn. Two important mechanisms are the MinRouteAdvertiseInterval (MRAI) timer and Route Flap Damping (RFD). Additionally, the use of ingress/egress filtering and route reflectors in iBGP can limit churn. The interactions between different protocol implementations and configurations, or their impact on BGP churn, is far from well understood.

III. DATASET

Our analysis is based on BGP update traces collected by the Routeviews project [1]. Routeviews collectors run BGP sessions with several routers, referred to as *monitors*, in many networks. A 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. In addition, Routeviews dumps a snapshot of the routing table of each monitor every two hours; we use those snapshots to observe the growth of the routing table size over the last few years.

We focus on update traces from monitors at large transit networks in the core of the Internet. Specifically, we analyze the churn time series from four monitors at AT&T, Sprint, Level-3 and France Telecom (FT)[1]. The corresponding monitors belong to the *Default Free Zone* (DFZ), meaning that they do not have a default route to another provider, and so they know a route to practically all destination networks in the Internet.

Routeviews provides historical update traces spanning more than six years for these four monitors. In some cases, the IP address of the monitor had 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. Our time series cover the period from January 1 2003 to December 31 2008, giving us six years worth of routing updates from four backbone monitors. However, the AT&T monitor was unavailable during two and a half month in late 2003, while the Sprint monitor was unavailable in the last four months of 2008.

If the multi-hop BGP session between a monitor and the collector is broken and re-established, the monitor will re-announce all its known paths, giving large bursts of updates. This is a local artifact of the Routeviews measurement infrastructure, and it does not represent genuine routing dynamics. Hence, we use the method described in [20] to identify and remove updates caused by such “session resets”. After such filtering, our dataset consists of more than 1.2 billion updates. Note that the updates received from a monitor is not a good estimate for the total number of updates a backbone router must process. A router typically has several active BGP sessions, and so the total load on the router is the sum of the churn from all BGP sessions.

To confirm that the method of [20] is able to identify all session resets and the subsequent routing table transfers, we applied it on BGP updates collected from three multi-hop BGP sessions of the same nature as the Routeviews sessions. We compared table transfers detected by the used method against BGP session logs. The method was able to detect table transfers and their exact starting time in 11 out of 12 cases. The 12th case was also reported but with 7 seconds difference in the starting time. This indicates the effectiveness of the used method in detecting table transfers.

---

[1] Because of space limitations we limit ourselves to present results mainly from AT&T, Sprint, and Level-3, while discussing results from FT as needed. An extended version of the results that includes all monitors can be found in [5].

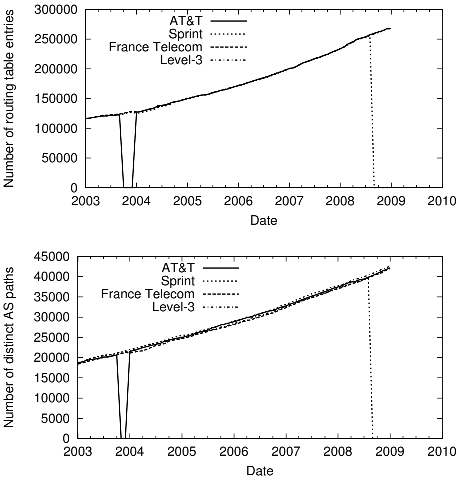

Fig. 1: Routing table size (top) and number of observed AS paths (bottom) during our study period.

Due to complex iBGP configurations using confederations or route reflectors, different edge routers of the same AS do not necessarily see the same set of paths to different destinations. We compared the churn time series of two monitors located in AS 2914 (NTT-A) and observed that their cross-correlation coefficient is high (0.93)[2]. Even though we cannot claim that this observation is true in general, it is reasonable to expect that two routers of the same AS would produce similar (but not identical) churn.

IV. THE "RAW" CHURN TIME SERIES

Before we focus on the churn time series, we first show two important aspects of growth in the BGP routing system. The top panel in Fig. 1 shows the number of routing table entries in the four monitors, sampled on a monthly basis. The number of entries in the different monitors is very similar, which is expected since these monitors are all DFZ routers. The *number of routable prefixes increased by 125% over the last six years*, from about 120K to 270K entries. The increase in the table size fits well with a quadratic function, with a coefficient of determination of 99.9%. The bottom panel in Fig. 1 shows the number of distinct AS paths (routing paths) in the routing tables (after removing the effects of AS path-prepending) again on a monthly basis. This metric has also increased dramatically (120%) during the last six years, from 19K to 42K paths.

One may expect that since the size of the routing table and the number of routing paths have more than doubled during the last six years, BGP churn should also show a similar consistent and significant increase. This is not the case however. The left column in Fig. 2 shows the "raw" BGP churn time series, measured as the number of BGP updates received daily from each monitor. Some high-level observations are necessary before we proceed with the analysis.

The raw time series is dominated by frequent and large spikes. At all monitors, there are days with dramatically higher churn than usual. We have truncated the y-axis of these plots to make the graphs more readable; in some days, the number of updates reached several millions. Large spikes are particularly frequent in the Level-3 monitor. Such spikes cannot be ignored as "statistical outliers"; instead, we need to understand what causes them.

There are several "level shifts". In addition to spikes, we see several periods of sustained increased activity that last for weeks or months. For example, we see a period that lasted about 6 months in mid-2006 at the Level-3 monitor. Again, level shifts cannot be viewed just as incidents of statistical non-stationarity; we need to understand what causes them.

There is little correlation between monitors of different ASes. The spikes and level shifts in the four monitors do not follow the same pattern. The cross-correlation coefficient between the different monitors is very low – between 0.04 and 0.13. This indicates that churn is highly dependent on the location and configuration of the corresponding router. We cannot understand the evolution of BGP churn by just looking at one monitor.

Churn is highly bursty even in large time scales. As seen in the left column of Fig. 2, churn is highly bursty even in the relatively large time scale of a day. We also examined the churn time series in shorter time scales (5 minutes and one hour) and observed that in some cases the majority of the daily churn is produced during short periods that last for few minutes.

It can be misleading to infer long-term trends from the raw churn time series. Because of the previous issues, it is clear that the blind application of statistical trend estimation methods can fail to detect a trend or it can produce misleading results. The approach we take in this paper is to first analyze what causes some major characteristics of the raw time series (spikes, level shifts, etc) and then, after we remove pathological or local effects, to apply statistical trend estimation on the remaining "baseline" churn.

V. IMPACT OF MRAI TIMER AND OF DUPLICATE UPDATES

We first analyze the deployment and impact of the MRAI timer during the last six years. Then, we examine the frequency of duplicate BGP updates in the churn time series.

A. MRAI timer deployment and impact

The BGP standard recommends using a jittered MRAI timeout that falls between 22.5 and 30 seconds in eBGP sessions. When MRAI is used in the BGP session between a monitor and the collector, we expect that consecutive updates from the monitor will arrive in bursts separated by at least 22.5 seconds. So, to detect whether the MRAI timer was deployed in the four monitors during the study period, we used the following two-steps approach.

1. For each monitor we selected one day from each week of the study period, which resulted in a sample of 288 days per

---

[2] This is the only AS in which Routeviews has more than one monitor throughout the study period.

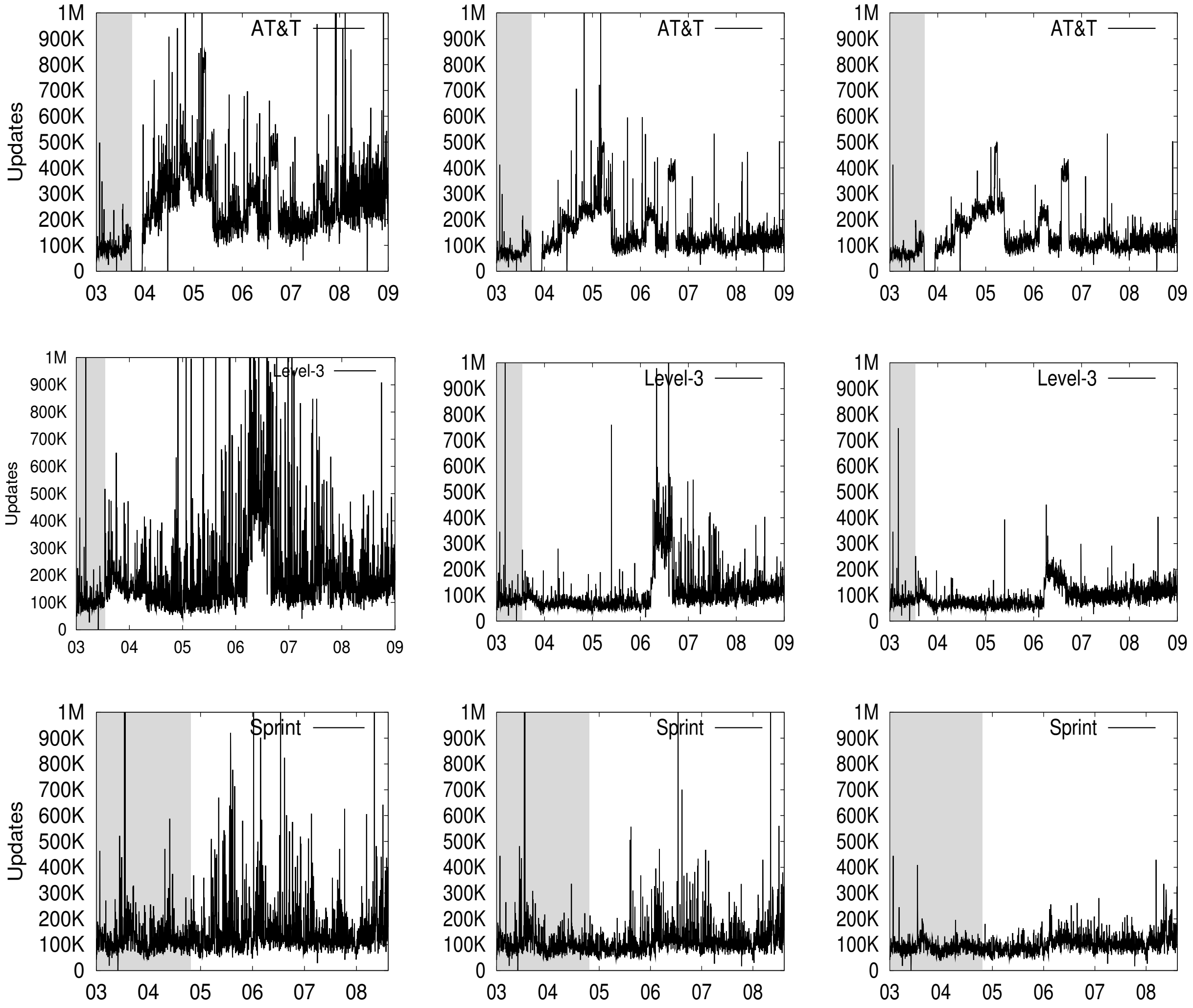

Fig. 2: Daily BGP churn: raw time series (left), after removing duplicates (middle), after removing duplicates and large events (right) at AT&T (top), Level-3, and Sprint (bottom).

monitor. We then calculated the distribution of update inter-arrival times for each day in the sample.

2. We calculated the fraction of inter-arrivals between consecutive updates that is less than 22.5 seconds. If that fraction is significant, MRAI was probably *not* deployed on the corresponding day. We find that setting the threshold for the fraction of inter-arrivals anywhere between 0.15 and 0.3 results in detecting the same periods.

Using this approach we inferred the MRAI deployment periods for our 4 monitors. We use grey shaded areas in the time series to show periods in which the MRAI timer was deployed. Note that the MRAI timer was disabled in all monitors after a certain point during our study period. It should also be noted that the MRAI timer is turned-off by default in the products of a major router vendor.

To investigate the impact of the MRAI timer, we measured the median daily churn rate in a three month period before and after each identified deployment transition. For each transition, we calculated the ratio (median churn without MRAI)/(median churn with MRAI). We found that the churn level increases when the MRAI timer is turned off, while it decreases when it is turned on. For example in Level-3, churn increased by a factor of 1.8 when the MRAI was turned off.

B. Duplicate updates

The conventional wisdom is that BGP implementations generate a large number of duplicate updates, which imposes an unnecessary processing load on routers. It has been pointed out that one reason for the large number of redundant updates is stateless BGP implementations that do not keep track of the last update sent to a peer [9].

We identified all duplicate updates (announcements and withdrawals) in our dataset. By “duplicate announcement”, we mean an announcement that is identical to the last seen announcement for the same prefix, i.e., no change in either

Monitor	AT&T	Level-3	FT	Sprint
2003	23.7%	40.7%	33.0%	7.2%
2004	47.6%	53.8%	45.2%	23.5%
2005	34.8%	61.7%	52.0%	41.1%
2006	31.8%	46.1%	43.5%	17.7%
2007	52.6%	42.3%	50.0%	14.3%
2008	59.6%	32.4%	43.2%	12.9%

TABLE I: Fraction of duplicate updates per monitor.

the AS-path or in any of the transitive route attributes. These announcements are redundant and can be viewed as a pathology of the BGP implementation at the corresponding monitor. A recent measurement study [15] attributed duplicates in BGP churn to interactions between iBGP and eBGP.

To our great surprise, we measured that, across all four monitors, duplicate announcements are responsible for about 40% of the churn during the study period! On the other hand, duplicate withdrawals are close to zero (except Level-3, where they account for about 1% of the updates). It is disappointing that almost half of the observed churn is not really necessary. This number is higher than the 16% of the duplicate announcements “AADupType1” reported earlier [11]; that study looked at monitors located in ASes of different sizes during a 6-month period in 2006. Our findings indicate that there is still much to be gained by deploying improved BGP implementations that keep additional state to avoid sending redundant updates to the global routing system.

The number of duplicate updates per day is highly variable, and shows no correlation across monitors. It is also difficult to identify a consistent long-term trend in the number of duplicates. Table I shows the fraction of duplicate announcements per year in each monitor. It is worth noting that the Sprint monitor shows a much lower fraction of duplicates than the other three monitors. These results indicate that the specific implementation of BGP and local configuration details can greatly influence the amount of redundant updates.

The second column in Fig. 2 shows the three time series after filtering out duplicate updates[3]. Note that removing duplicate updates has the additional benefit that most of the spikes are also removed. This indicates that redundant updates are not only responsible for a large fraction of churn, but they are also responsible for generating large bursts of churn that put the highest burden on router CPUs.

VI. LARGE EVENTS

After removing duplicates, we focus on “large routing events”, or simply *large events*, defined as events that affect a large number of prefixes at about the same time. The intuition is that incidents at the core of the Internet, such as link failures between two transit networks, have the potential to introduce instability to a large number of prefixes simultaneously, causing major churn spikes. At first we group prefix updates into single prefix events using an approach similar to what is used in [19]. Then, we group single prefix events that occur at about the same time into *events*. Next, we define a *large event* as an event that consists of more than a certain number of prefix events.

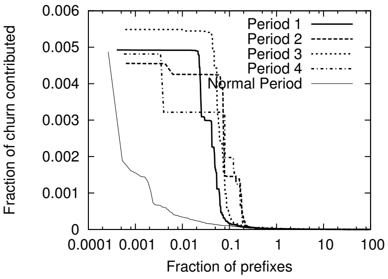

Fig. 3: Churn contribution from the most active prefixes during four level shifts at the AT&T monitor.

By examining the distribution of the number of affected prefixes per event in our four monitors, we decided to use a threshold of 2000 prefixes. With this definition, (at most) the top 0.2% of all events are considered large events. Different monitors experience large events with a different frequency. Also the time series of large events shows little or no correlation between different monitors. This implies that large routing events have a major impact only locally, and they are probably due to incidents close or within the monitor AS. Further details about the identification of large events, and the related thresholds, can be found in [5].

The third column in Fig. 2 shows the churn, after removing updates due to large events. Comparing this time series with the churn after removing duplicates, we see that most remaining large spikes in the duplicate-free churn are related to large events. Even though the remaining time series, after excluding the impact of duplicate updates and large events, is much smoother, it still shows several significant level shifts; they are the subject of the next section.

VII. ANALYZING LEVEL SHIFTS

The time series (for the AT&T and Level-3 monitors in particular) are still dominated by level shifts where the magnitude of churn differs substantially from the periods before and after. The presence of these level shifts makes it difficult to reliably detect any long-term trend in churn. It has proved hard to find an automated method for identifying these level shifts and finding their root cause. Instead, we make an in-depth analysis of the most dominant level shifts. We focus our analysis on the AT&T and Level-3 monitors.

AT&T: The AT&T time series involves several clear level shifts, in addition to a long period of increased activity spanning one and a half year from January 2004 to June 2005. In our detailed analysis, we manually identified four distinct level shifts.

The first activity period is the long period of increasing trend from December-11-2003 to March-01-2005. The second level shift started immediately after the first activity period and lasted for one month. The third and fourth level shifts took place from February-15-2006 to March-31-2006 and from July-31-2006 to September-25-2006 respectively.

Fig. 3 shows the fraction of total churn contributed by each prefix during our four activity periods, sorted by the activity

---

[3] Raw and filtered datasets are available at http://vefur.simula.no/bgp-churn/.

level of each prefix. From this plot we observe that there is a very small set of prefixes that contributed the majority of churn during each period. For comparison, we also include a curve for the churn in 2008, which does not contain any level-shifts. This clearly shows the abnormality of the level shift periods.

During period 1, a small set of 148 prefixes (i.e 0.1% of the total number of prefixes) contributed 49.8% of the total churn. We investigated the activity patterns of these prefixes by examining the inter-arrival times of their updates. The prefixes can be classified into three groups based on their median updates inter-arrival times.

The first group consists of prefixes with median update inter-arrival time at 58 seconds. When investigating their update patterns we find that these prefixes belong to AS 21617 and are reached using the path {7018,701,21617} where 7018 is the monitor AS number. During this period this group of prefixes flapped up and down almost every minute. It is reasonable to believe that this long-lasting and high-frequency flapping pattern is caused by a flapping link or misconfiguration.

The prefixes that fall into the second group have their median update inter-arrival time at 65 seconds. By examining them closely we find that they are originated by either the monitor AS (i.e. 7018) or its direct customers. During this period this group of prefixes exhibited a change which is either a withdrawal or a re-announcement approximately every minute. Moreover, the 90th percentile of the update inter-arrival times is approximately equal to the median, which confirms a strict periodicity in these updates. This makes us believe that these updates are caused by an anomaly that changes the path selection at regular intervals, rather than a flaky link or some adaptive load balancing method that would give a more irregular pattern.

The last group includes prefixes that have their median update inter-arrival time at 196 seconds. We find that these prefixes belong to AS 1938, and were changing their AS path between {7018 10888 24 11537 20965 2200 1938} and {7018 10888 11537 20965 2200 1938}. It is difficult to spot the root cause in this case. However, in the FT and Level-3 datasets we observe similar flapping patterns that involve changing some prefixes' next hop from AS 24 (NASA) to other ASes. Therefore, this activity might be caused by some instability in/near AS 24 that lasted for a long time.

Period 2 started immediately after the end of period 1, and lasted for one month. There is a small set of 170 prefixes that generates 71.7% of the total churn during this period. The main driving change behind this level shift is a small set of prefixes belonging to General Electrics AS that continuously changed their next hop AS between AS 80 (General Electrics) and AS 1239 (Sprint). Note that AS 80 is a stub AS and does not announce many prefixes. Still, the frequency of the route changes is high enough to create this radical increase in churn.

In period 3 and 4, we find that the level shifts are caused by leaking of private AS numbers into the global routing. Private AS numbers (ranging from 64512 to 65535) are used to divide large ASes into multiple smaller domains connected by eBGP, or they can be assigned to stub ASes that connect only to a single provider in order to conserve public AS numbers. Private AS numbers should always be removed from routing updates that are sent to the global BGP mesh. During these two level shifts, updates containing private AS numbers are responsible for 54.3% and 70.5% of total churn respectively.

Level-3: The data shows a clear level shift in the Level-3 time series from March-01-2006 to August-31-2006. By doing similar analysis as in the case of AT&T, we find that the increased activity can be attributed to a set of flapping prefixes, which changed their AS-PATH continuously from {3356 3561 4134 X} to {3356 1239 4134 X} or vice versa, where X represents the rest of the AS path. Here we see how AS 3356 (Level-3) alternates between the two different neighboring ASes 3561 (Savvis) and 1239 (Sprint) to reach AS 4134 (China-Backbone). Note here that Savvis is owned by Level-3 and hence the route through Savvis is preferred. When this route is lost, Level-3 will select the backup route through Sprint. The frequency of this flapping for each prefix is between once every 10 minutes and once every 20 minutes. However, China-Backbone is a major transit provider, and Level-3 selects it as the preferred path for more than 2000 destination prefixes. Hence, a single change will trigger a large number of updates.

The above analysis shows that level shifts are usually caused by specific failures or misconfigurations in or near the monitored AS. The left column in Fig. 4 shows the churn time series after filtering out all updates attributed to the level shift generators described above. In the following section we discuss and analyze different statistical properties and trends in the churn time series after removing the effect of level shifts.

VIII. THE GROWTH OF BASELINE CHURN

In this section, we analyze the growth of the churn time series after removing duplicate updates, large events, and the level shifts of the previous section. We refer to this time series as the churn "baseline". We also analyze the time series of peak churn, measured from the busiest 1-minute period of each day.

A. Baseline churn

After removing churn caused by anomalies and local effects, we are left with the baseline churn, which is a much smoother time series and shows more correlation across monitors (see Fig. 4). The cross-correlation coefficient between the AT&T, Level-3, and Sprint baseline time series is around 0.55, which is much greater than the highest value observed in the raw time series (0.14). This increase suggests that our approach has managed to filter most of the effects that are quite local to each monitor. The cross-correlation between the three North Americans monitors and France Telecoms remains low at (0.11). Geographical locality and presence could be a plausible explanation for this.

The application of linear regression on the baseline time series results in a low Pearson's correlation coefficient (0.03 to 0.42, depending on the monitor), since even the baseline

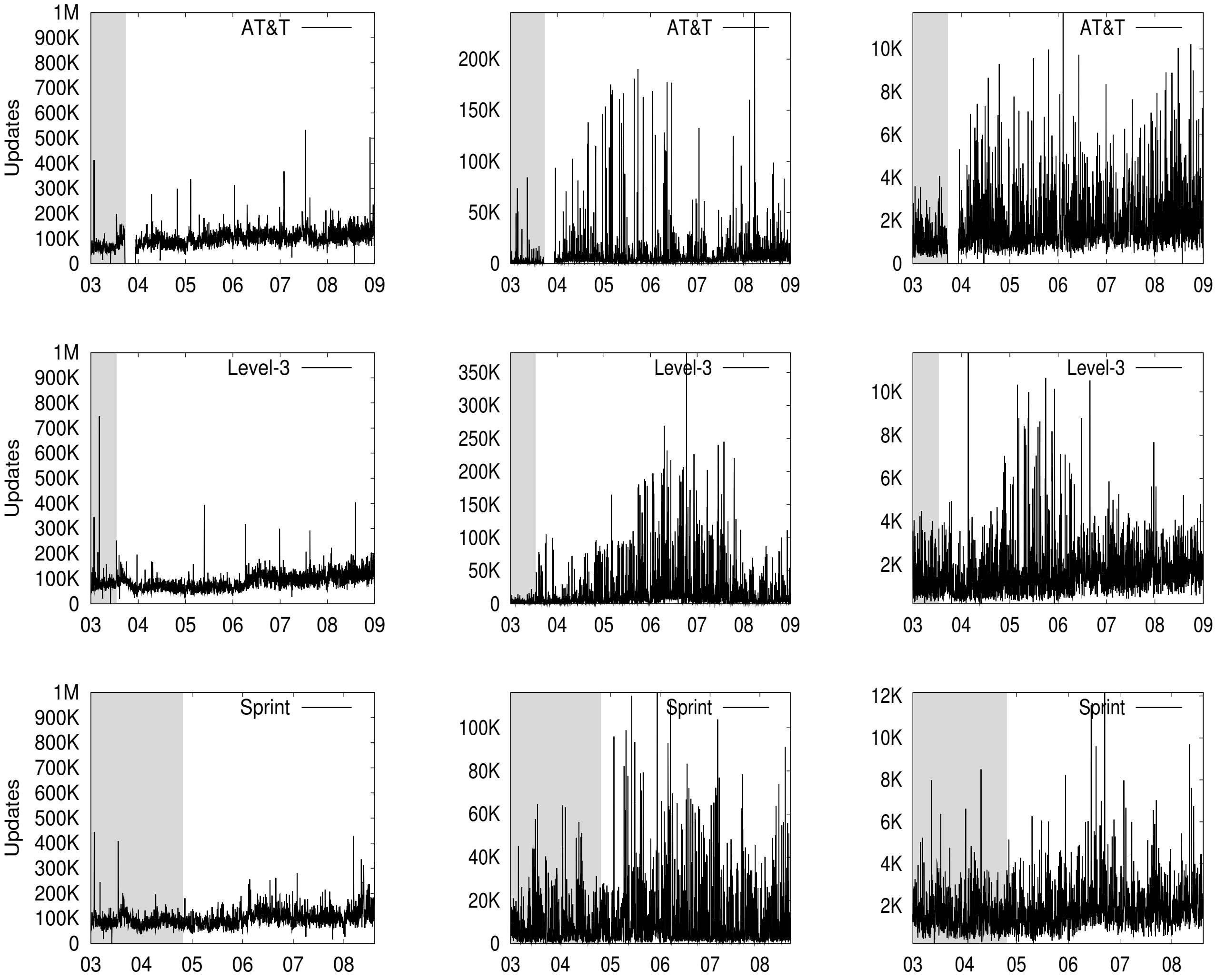

Fig. 4: Baseline daily total churn (left), 1-minute peak churn per day in the raw time series (middle), and 1-minute peak churn per day in the baseline time series (right) in AT&T (top), Level-3 (middle), and Sprint (bottom).

churn contains some spikes and small level shifts. Therefore, we rely on non-parametric statistics, which is more robust to outliers, and in particular on the Mann-Kendall statistical test for trend detection. The Mann-Kendall test reports that there is a statistically significant increasing trend in the baseline time series in all four monitors at a 90% significance level. Actually, both the non-parametric and parametric (linear regression) tests give similar estimates for the slope of the increasing trend. Table II presents the estimated slopes in additional updates per day.

The same table also shows the estimated relative churn increase during the six-year study period. This figure is calculated from the estimated slope using the median daily churn rate for the first 3 months as starting point. The two estimation techniques are in reasonable agreement with each other. There are, however, significant differences between the monitors. This is not surprising, since different monitors have different sets of customers and peers and different internal configuration.

A main observation from this analysis is that the increase in the baseline churn is relatively slow compared to the growth of the routing table size. The data presented in Fig. 1 shows that the number of routable prefixes has more than doubled over our study period, while the baseline churn has increased by 20-80%, depending on the monitor. This implies that the *daily number of updates per prefix in the baseline time series has decreased over the last six years*. Determining the reason for this improvement is beyond the scope of this paper. One potential explanation is better network management practices at most stub networks. The “densification” of the Internet [3] may have also helped, as it provides additional routing paths when the preferred route is lost.

B. Daily peak activity

The churn rates presented so far are daily averages. The peak churn rates in shorter timescales may be more important

TABLE II: Baseline churn growth: Mann-Kendall slope estimate in updates per day, and estimated relative churn increase during our study period. The parametric estimates are also shown.

Monitor	AT&T	Level-3	FT	Sprint
M-K slope	22.04	24.46	7.79	14.38
Est. increase	49.6%	69.3%	20.0%	29.4%
Lin. regr. slope	22.02	24.12	6.50	16.42
Est. increase	47.2%	81.5%	15.5%	33.3%

TABLE III: Daily 1-minute peak churn growth

Monitor	AT&T	Level-3	FT	Sprint
Raw peak churn				
M-K slope	1.03	1.57	0.27	-
Est increase	58.8%	258.9%	50.5%	-
Baseline peak churn				
M-K slope	0.32	0.43	0.10	0.29
Est increase	47.9%	90.9%	20.9%	39.3%

in terms of the processing load imposed on routers. Here, we examine the growth of the peak daily churn rate, measured as the *maximum 1-minute churn on each day*. We refer to this time series as the "daily 1-minute peak churn".

The plots in the second and third columns of Fig. 4 show the daily 1-minute peak churn in the raw time series and in the baseline time series, respectively. A first observation is that the daily peak activity in the raw time series is much higher than in the baseline time series: on average, there is an order of magnitude difference between the two time series across all monitors, and on some days the difference can reach up to two orders of magnitude! This confirms that local effects and protocol misconfigurations are responsible for most of the peak churn that routers have to process.

The Mann-Kendall test reports an increasing trend in the raw and baseline daily 1-minute peak churn across all four monitors. The exception is the raw time series at the Sprint monitor, where no trend could be detected. Table III presents the slope and the relative estimated increase at each monitor. The modest growth in the FT monitor is probably due to the use of the MRAI timer, which was switched on in late 2006 and contributed to a lower churn level (see below). The noisy nature of the raw time series makes it difficult to get accurate growth trends, and so these numbers should be viewed only as rough estimates.

We observe that the estimated relative growth in daily 1-minute peak churn rate is somewhat higher for the raw time series than for the baseline. This indicates that the impact, in terms of peak churn, of duplicates and other local effects increases with time. For the baseline time series, the increase in the daily 1-minute peak level is comparable to the increase in the total daily churn.

Finally, we investigate to what extent the daily 1-minute peak churn is influenced by the use of MRAI timers. We compare the median daily 1-minute peak churn calculated in a three month window immediately before and after each change in the MRAI configuration at the FT (two transitions, see [5] for details), Level-3, and Sprint monitors. Fig. 5 shows the churn in the 3-month period before and after the MRAI timer was turned on in late 2006 at the FT monitor, for the raw and baseline time series (the horizontal lines in the figures show the median level of churn). We find that the MRAI timer has no clear effect on the daily 1-minute peak churn in the baseline time series. However, in the raw time series, there is a clear increase in the peak churn when the MRAI timer is off. The peak churn increases by a factor 1.1 in Sprint, 2.0 in Level-3, and 3.7 and 2.8 in the first and second transitions in FT.

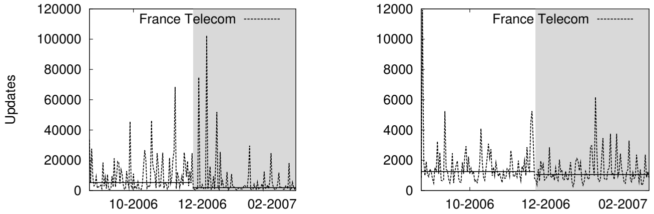

Fig. 5: Churn before and after MRAI timer is turned on at FT monitor (left: raw peak churn, right: baseline peak churn).

These findings show that the effect of the MRAI timer is much stronger on the raw time series than on the baseline. This implies that the MRAI timer is mostly effective at filtering out some duplicate updates and local effects.

IX. RELATED WORK

Interdomain routing dynamics and scalability have been active topics of research during the last decade or so.

Labovitz [9] was the first to show that BGP suffers from excessive churn caused by pathological protocol behavior and suggested practical ways to fix broken BGP implementations. In a follow-up work [10], they found that better router implementations had reduced churn by an order of magnitude, but that duplicate announcements still contributed much unnecessary churn. Our findings confirm that this is still the case, and that this type of updates is responsible for most large spikes in churn. Mahajan et al. reported [12] that BGP misconfigurations are pervasive and cause an increase in the processing load of routers. A recent measurement study [11] concluded that the state of BGP routing is now "healthier" than it was a decade ago.

The phenomenon of *path exploration* was discussed in [8]. The effectiveness of the MRAI timer to limit path exploration was discussed in [6].

The network topology plays a role in the observed churn as well. In a recent measurement study, it was shown that path exploration is less severe in the core of the Internet than at its periphery [14]. It has also been shown that events at the edge of the network affect a larger number of ASes than those at the core [21]. In a recent study [4], we investigated the role of various topological factors, including multihoming, hierarchy and peering links, and of the use of update rate limiting on BGP churn growth.

Another set of studies analyzed the contribution of different ASes and prefixes to the observed churn. Broido et al. [2] showed that a small fraction of ASes is responsible for most of the churn seen in the Internet. Similarly, several other papers [16], [18] reported that a small subset of prefixes are responsible for a large percentage of churn.

A recent study [7] reported that BGP churn increases at a much faster pace than the routing table size. During 2005, the daily rate of update messages almost doubled, while the size of the routing table grew by only 18%. Our study, based on a much longer study period and a larger number of monitors, gives a much more conservative estimate of the churn growth rate.

X. CONCLUSIONS

This study has investigated the evolution of churn at four monitors located in the core of the Internet over the last six years. The corresponding time series are very bursty, with large churn spikes and level-shifts. We have performed an in-depth analysis of the time series in order to identify and explain the main sources of churn.

We found that *the most severe update bursts are caused by local effects in the monitor AS*. In particular, we found that 40% of route announcements are redundant and they are not needed for correct protocol behavior. We also identified the underlying reasons for the most severe level-shifts in churn. These are normally caused by configuration mistakes or other anomalies in or close to the monitored AS. Our findings suggest that the most effective short-term solutions for limiting churn will be protocol improvements that filter out redundant updates, and methods that can detect (long-lasting) configuration mistakes and other anomalies that result in sustained high churn levels.

This study has also shown that *there is a long-term increasing trend in the identified baseline churn*, but at the same time, *the growth rate is relatively low*. We find that the *churn rate increases more slowly than the number of prefixes* in the routing table. While the routing table grew about 120% during our study period, the baseline churn rate grew only by 20 - 80%, depending on the monitor.

There can be several reasons why we only see a slow increase in the baseline churn compared to the growth of the routing table size. On one hand, configuration management systems and experience are improving. Also, the increasing connectivity in the Internet [3] can play a positive role, since more failures can be handled locally if an alternate route is known.

We also investigated the daily 1-minute peak churn rate, and found that this is an order of magnitude higher in the raw time series compared to the baseline. These time series are very noisy, but they appear to be slowly growing with time.

We plan to take this work forward in several directions in future work.

We want to analyze churn in a setting where we have more control of the measurement infrastructure. One extension is to set up monitoring sessions to several routers in the same AS, in order to assess the effect of intra-AS configurations. We also want to set up several monitoring sessions to the same router, in order to isolate the effect of configuration settings like the use of MRAI timers.

ACKNOWLEDGMENT

We would like to thank Samantha (Sau Man) Lo for her inputs and help during the early stages of this work. We also thank Tarik Cicic for his help with measurement data. We are very grateful to the RouteViews project, because this study would not be possible without their long-term efforts.

REFERENCES

[1] Routeviews project page. http://www.routeviews.org.

[2] Andre Broido, Evi Nemeth, and kc claffy. Internet expansion, refinement, and churn. In *European Transactions on Telecommunications*, January 2002.

[3] Amogh Dhamdhere and Constantine Dovrolis. Ten years in the evolution of the Internet ecosystem. In *In the Proceedings of ACM SIGCOMM/USENIX Internet Measurement Conference (IMC)*, 2008.

[4] Ahmed Elmokashfi, Amund Kvalbein, and Constantine Dovrolis. On the scalability of BGP: the roles of topology growth and update rate-limiting. In *Proceedings ACM CoNEXT*, 2008.

[5] Ahmed Elmokashfi, Amund Kvalbein, and Constantine Dovrolis. On the evolution of BGP churn in core networks. Technical Report 02-2009, Simula Research Laboratory, Available at http://folk.uio.no/amundk/churnevolution.pdf, June 2009.

[6] Tim Griffin and Brian Premore. An experimental analysis of BGP convergence time. In *Proceedings ICNP*, 2001.

[7] Geoff Huston and Grenville Armitage. Projecting future IPv4 router requirements from trends in dynamic BGP behaviour. In *Proceedings ATNAC*, Australia, dec 2006.

[8] Craig Labovitz, Abha Ahuja, Abhijit Bose, and Farnam Jahanian. Delayed Internet routing convergence. In *Proceedings ACM SIGCOMM*, pages 175–187, 2000.

[9] Craig Labovitz, G Robert Malan, and Farnam Jahanian. Internet routing instability. In *Proceedings ACM SIGCOMM*, Cannes, France, 1997.

[10] Craig Labovitz, G. Robert Malan, and Farnam Jahanian. Origins of Internet routing instability. In *Proceedings IEEE INFOCOM 1999*, New York, NY, March 1999.

[11] Jun Li, Michael Guidero, Zhen Wu, Eric Purpus, and Toby Ehrenkranz. BGP routing dynamics revisited. *Computer Communications Review*, April 2007.

[12] Ratul Mahajan, David Wetherall, and Tom Anderson. Understanding BGP misconfiguration. In *Proceedings ACM SIGCOMM*, pages 3–16, New York, NY, USA, 2002. ACM.

[13] David Meyer, Lixia Zhang, and Kevin Fall. Report from the IAB workshop on routing and addressing. http://tools.ietf.org/id/draft-iab-raws-report-02.txt, apr 2007.

[14] Ricardo Oliveira, Beichuan Zhang, Dan Pei, Rafit Izhak-Ratzin, and Lixia Zhang. Quantifying path exploration in the Internet. In *Proceedings IMC*, Rio de Janeiro, Brazil, oct 2006.

[15] Jong Han Park, Dan Jen, Mohit Lad, Shane Amante, Danny McPherson, and Lixia Zhang. Investigating occurrence of duplicate updates in BGP announcements. IETF GROW, jul 2009.

[16] Jennifer Rexford, Jia Wang, Zhen Xiao, and Yin Zhang. BGP routing stability of popular destinations. In *IMW '02: Proceedings of the 2nd ACM SIGCOMM Workshop on Internet measurement*, pages 197–202, New York, NY, USA, 2002. ACM.

[17] George Varghese. *Network Algorithmics*. Morgan Kaufmann, San Fransisco., 2004.

[18] Ricardo V.Oliveira, Rafit Izhak-Ratzin, Beichuan Zhang, and Lixia Zhang. Measurement of highly active prefixes in BGP. In *Proceedings IEEE Globecom*, 2005.

[19] Jian Wu, Zhuoqing Morley Mao, Jennifer Rexford, and Jia Wang. Finding a needle in a haystack: pinpointing significant BGP routing changes in an IP network. In *NSDI'05: Proceedings of the 2nd conference on Symposium on Networked Systems Design & Implementation*, pages 1–14, Berkeley, CA, USA, 2005. USENIX Association.

[20] Beichuan Zhang, Vamsi Kambhampati, Mohit Lad, Daniel Massey, and Lixia Zhang. Identifying BGP routing table transfers. In *MineNet '05: Proceedings of the 2005 ACM SIGCOMM workshop on Mining network data*, pages 213–218, New York, NY, USA, 2005. ACM.

[21] Xiaoliang Zhao, Beichuan Zhang, Andreas Terzis, Daniel Massey, and Lixia Zhang. The impact of link failure location on routing dynamics: A formal analysis. In *Proceedings ACM SIGCOMM Asia Workshop*, apr 2005.