ON/OFF Model: A New Tool to Understand BGP Update Burst

Xiaoliang Zhao, Daniel Massey
University of Southern California
Information Sciences Institute
Email: {xzhao, masseyd}@isi.edu

Mohit Lad, Lixia Zhang
Computer Science Department
U. of California, Los Angeles
Email: {mohit, lixia}@cs.ucla.edu

***Abstract*— BGP, the inter-domain routing protocol, can exhibit complex behaviors under various conditions. Although BGP log data have been made available in the recent years, the sheer size of the log data makes it difficult to interpret BGP behavior using only the raw BGP update messages and understanding the global routing dynamics in today's Internet remains a great challenge.**

In this paper we focus on the analysis of BGP update bursts, a commonly observed event that occurs at varying frequency. We define a BGP *update burst* as an occurrence of a large number of BGP updates that are separated by very short time intervals. To investigate the causes of such bursts we developed an ON/OFF model which can be used to classify the BGP bursts into two classes: stable routing changes and transient route flapping. A stable routing change means an existing route is replaced by a new route that lasts for a long time period, while transient route flapping means a series of routing updates occur for the prefix over a short time period but at the end of the burst the route is the same as the original route. By applying our ON/OFF model to BGP routing updates over the last two years, we found that the ON/OFF model is an effective way to identify stable routing changes, such as those caused by physical failures in the network, and that about half of the update bursts are caused by transient route flapping. Further investigation reveals the specific causes for a number of the transient flappings. Overall, the development of the ON/OFF model helps us make a significant step towards a complete understanding of the global routing dynamics.

I. Introduction

The Internet consists of large number of Autonomous Systems (AS) that exchange routing information with each other to learn the best path to the destinations. Presently, BGP (Border Gateway Protocol) is the *de facto* inter-AS routing protocol and is designed to adapt to link failures, AS topology changes and routing policy changes. BGP is a path vector based routing protocol and each BGP router advertises to neighbors (peers), entire AS path information to destinations. To exchange routing information, the BGP peers establish peering sessions. Whenever a new BGP session is set up between two peers, the complete routing tables are exchanged between them. After this initial exchange, routers only send update messages for routes that change or new routes that are added. Information exchanged by BGP is used for global routing. Therefore, faults or attacks in the BGP infrastructure can lead to problems such as denial of service and misdirected traffic.

Ideally, as a protocol, there would be a solid understanding of BGP's behavior, its response to faults, and its vulnerabilities to attacks. But in practice, the BGP infrastructure constitutes a large scale system and could exhibit complex behaviors under various conditions. BGP log data have been available in the recent years, provided by Oregon Route-Views [1] and RIPE [2]. In their services, there are one or more monitoring points, which are BGP routers that peer with routers within ISPs. A monitoring point archives its BGP routing table snapshots and the BGP updates received from its peers. These update messages that either signal route change or some route attribute change, are caused by events such as a physical link failure, the emergence of a better route, or simply a policy change. Due to the large scale deployment of BGP, and policies, events are hidden from the observers at the monitoring points. Instead, what we see at these monitoring points, is the results of the events. For instance, a physical link failure is an event that would cause the ends of the link to send update messages to their neighboring routers. Depending on how many of these routers use this link, we would have updates being propagated further. At a remote monitoring point, all we see is update messages, without any idea about what kind of event caused this update. This problem, as well as the sheer size of the log data, make it difficult to interpret BGP behavior using only the raw BGP updates messages. Therefore, understanding BGP dynamic behavior continues to be an open challenge.

In this paper, we propose a model that would be a significant step toward a complete understanding of the global routing dynamics. This paper is an attempt to demystify the events behind these updates as observed from monitoring points and to gain some high level insight into what these updates can tell us about the type of changes in BGP routes. In particular we study the event of BGP update message bursts. *BGP burst* refers to a series of updates triggered by routing changes. We show that with our model we can gain considerable insight into the events causing these bursts. We classify BGP bursts into two classes: *transient routing changes* and *non-transient routing changes*. A transient routing

change refers to a change in which a route, after a series of routing updates, is eventually restored back, while a non-transient change is one in which a route is replaced by another route for a significantly long time. Transient changes, if better understood, could be potentially beneficial for operational practices, such as optimizing some BGP parameters to better handle such changes.

By applying our ON/OFF model to BGP routing updates over the last two years, we found that the ON/OFF model is an effective way to identify stable routing changes, such as those caused by physical failures in the network, and that about half of the update bursts are caused by transient route flapping. Further investigation reveals the specific causes for a number of the transient flappings. Overall, the development of the ON/OFF model helps us make a significant step towards a complete understanding of the global routing dynamics.

The paper is organized as follows. Section II talks about our methodology used for the data processing. Section III presents the ON/OFF model. Section IV shows that, given a ON timer as five minutes, there are 50% of total BGP bursts are transient changes, as well as some statistics for duration distribution of BGP bursts are presented. Section V studies some cases of BGP bursts and found some of them are caused by worm activities, faults, which may suggest us to look back at protocol design more carefully to better response to those changes.

II. Data Source

We analyzed BGP routing updates collected by RIPE NCC[2] during several months in 2001 and 2002. RIPE NCC has eight data monitoring points (rrc00 - rrc07). We selected the rrc00 monitoring point and gathered data from the BGP routers listed in Table I. Some of these routers are located in global ISPs and others are located in regional ISPs. Geographically, routers are located in different countries including the United States, Japan and three European countries.

We chose the rrc00 monitoring point because it receives full routing tables from ISPs. If an ISP only provides partial routing tables and then withdraws its route to a prefix, this may indicate that ISP has lost its route to this prefix or may indicate the ISP has simply changed routes and the new route does not match the partial export policy.

It should also be noted that BGP updates are sent to the monitoring point via multi-hop BGP connections. In the operational Internet, nearly all ISP peerings are through BGP routers sharing a common physical link, where BGP updates are sent via TCP connection over single link/hop. However, the BGP monitoring point RRC00 peers with ISP routers via TCP connections that cross multiple route hops and links. When the multi-hop session fails, the monitoring point reports a session state change. Note that if a peering session is reset, all routes are implicitly withdrawn and, when the new peering session is started again, it involves a complete table exchange. In nearly all session reset we observed during the studied periods, the same routes are re-advertised when the session to the ISP router resumes. We attribute this behavior to lower stability of the multi-hop BGP sessions. We pre-process the update files to remove the updates that are generated due to session reset, resulting in a cleaned data set of BGP updates, for our analysis.

Location	ASes that rrc00's peers belong to
US	AS7018 (AT&T), AS2914 (Verio)
Netherlands	AS3333 (RIPE NCC) AS1103 (SURFnet) AS3257 (Tiscali Global)
Switzerland	AS513 (CERN), AS9177 (Nextra)
Britain	AS3549 (Global Crossing)
Japan	AS4777 (NSPIXP2)

TABLE I
RRC00's peering ASes that we examined

A routing change can be broadly of two types, one that changes AS path for a given prefix, including withdrawal and announcement of a newly reachable prefix, while the other type that does not. AS path changes may be due to many reasons, such as hardware failures, operational BGP session resets, or policy changes. A BGP update which does not convey new path information may change other BGP attributes, such as Multiple Exit Discriminator (MED), Community attributes. Such kind of updates may be caused by policy changes or bad software implementation choices. In this paper, we are only concerned about the AS path changes and do not look into details of other attribute changes. In the rest of the paper, all the route changes or updates are referred to those with AS Path Changes, unless specified otherwise.

The data was collected over the months of July 2001, September 2001, November 2001, Feburary 2002, July 2002, and August 2002. All of the data was examined by using the methods described in the following sections. But due to the paper size limitation, we only present results for some particular months and from some particular peers' point of view, but if not mentioned, the results for other months and peers are in general similar to the sample results.

III. ON/OFF Model

In this section, we develop the ON/OFF model and show its usefullness by analyzing BGP burst.

A. Bursty Nature of BGP updates

Figure 1 shows the number of updates on a hourly basis. As can be seen, the total number of updates per hour are normally below 100,000, but there are some spikes. Those spikes are indications that a large volume

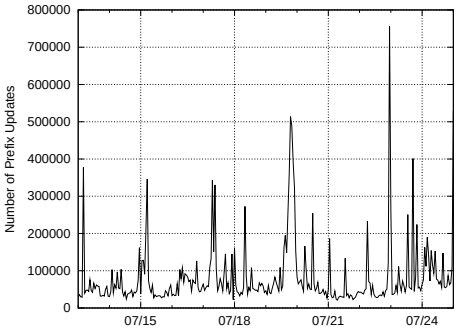


Fig. 1. Number of BGP Prefix Updates in Hourly Bins From July 13th, 2001 to July 24th, 2001

of updates coming as a burst. BGP burst, also noted in [3], is a phenomenon of interest because it could be an indication of routing instability, which might be resulted from routing device failure, configuration error, or even malicious attack. During BGP burst, routing paths may be altered, forwarding performace may be affected, and applications may experience delay and package loss.

We need a measure of BGP burst to estimate, what the level of instability is at any given instant of time. Given a trace of BGP updates, one straightforward way to analyze it would be to simply count the total number of updates for a given period. Such simple count could give us some clues on the occurence of some event. As a matter of fact, the spike on 07/18/2001 is corresponding a known topology event [1] and the spike on 07/19/2001 corresponds to a known worm attack [2]. However, those spikes do not tell whether bulk of them come from a small set of prefixes, or whether the updates come from a very large set of prefixes and are evenly distributed.

What we are really interested about is *the number of prefixes that are in the process of change at any given instant of time*. If a very large set of prefixes experience routing change simultaneously, it is a strong indication of occurrence of routing event warranting further investigation. Simple count of update messages is inadequate for this purpose. We will present the ON/OFF model in the rest of this section, which exactly captures the the number of prefixes that are in the process of change at any given instant of time.

B. Definition of ON/OFF Model

We build the ON/OFF model with a hypothetical case of updates for a single prefix.

Fig 2 shows an example on how we can build on the ON/OFF model. At a higher level, the state ON corresponds to an active state, where a prefix may be expected to have routing changes, while OFF corresponds to a steady state, where the prefix's route is

[1]Baltimore tunnel fire occurred at about 15:00 EST on July 18, 2001.

[2]Code Red worm spread out on July 19, 2001.

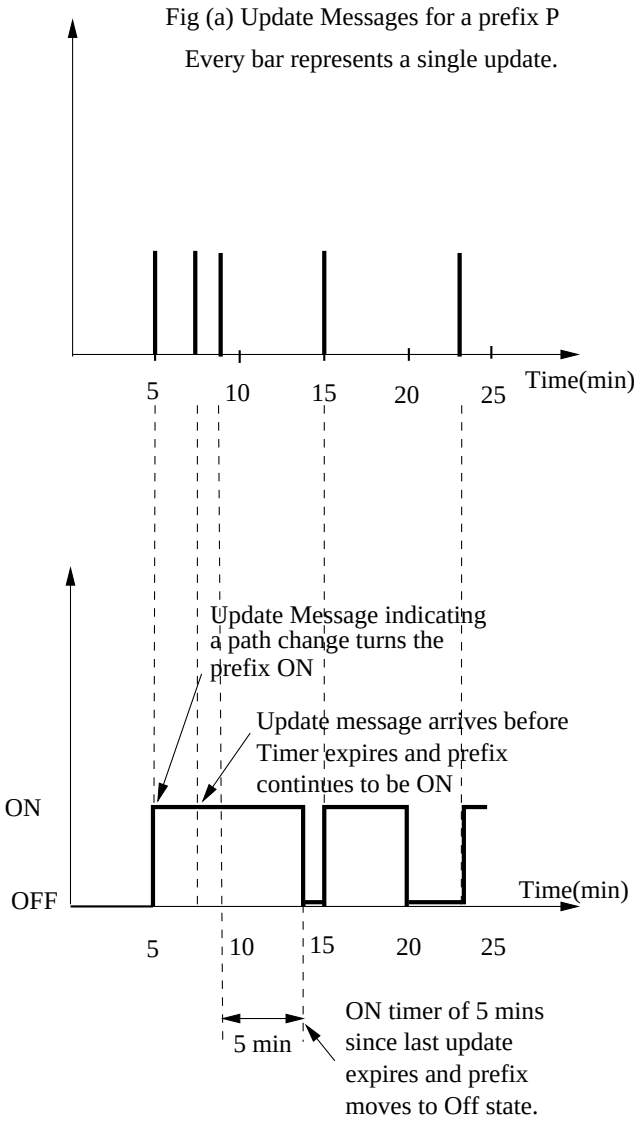


Fig. 2. ON/OFF state transitions for a prefix P with 5 minutes timer corresponding to updates received for that prefix

expected to stay for some time. Part a of the figure shows a sequence of update messages spaced in time. At time minute 5, on the arrival of the first update message, the prefix is turned ON, and the a timer, called *ON timer*, is started. This timer is used to account for temporary changes while alternate routes are being explored as well as to account for convergence problems [4]. In this hypothetical case, we consider this timer to be 5 minutes and thus the timer will expire at minute 10. As we can see, the second update arrives at a time t=minute 7, before the timer expires. This update keeps that the prefix stays at ON state. At this moment, the timer is restarted to wait for another 5 minutes to accommodate further updates. Similar

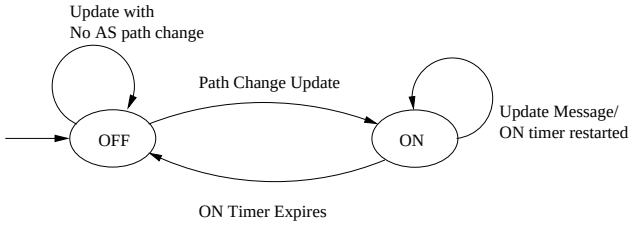


Fig. 3. On/Off State transition diagram

action is taken at t=minute 8, when the 3rd update message arrives. However from t=minute 9 to t=minute 14, there is no further update, and the timer expires thus pushing the prefix to OFF state as shown in part b of fig 2. The prefix continues to be OFF till the next update announcing a route change arrives for the prefix, which is at t=minute 15. Again, the prefix is turned ON and turns OFF when there is no further update for the next 5 minutes. Thus, the prefix, as observed from the monitoring point is moving between ON and OFF states, depending on the updates being generated and the time spacing between successive updates.

Defintion We define a prefix to be **ON**, if it has recently received an update indicating an AS path change and an **timer** with time t, that would turn it off, has not expired since its last update.

From the simple example discussed above, we can construct a state transition diagram for the ON and OFF states as in fig 3. Every prefix by default is in the OFF state. If we observe an update message for a prefix, but the new path announced is the same as the old one, the prefix will continue to be in the OFF state. However, if the new path announced is different from the OLD path, then the prefix moves on to ON state. In this state, the prefix waits for one of the two events to happen. Either, the ON timer expires, in which case the prefix moves back to OFF state, or there is an update message for the prefix before the ON timer expires. In the latter case, the prefix stays at ON state and the ON timer will be restarted.

Thus, a prefix that receives updates, all of which are within t of each other will keep the prefix *ON* for the entire period of updates. However, if even one update arrives more than t of the last update for the same prefix, then the prefix would have turned off as soon as the ON timer expired. The choice of this timer t is very critical, which will be discussed in detail later.

C. Implications of ON/OFF periods

The total amount of time, a particular prefix remains ON is called as *ON period*. Similarly, the duration of time, a prefix remains OFF is called as *OFF period*. ON period for a prefix indicates that the prefix has very recently undergone a change of path. Thus, if a prefix was to have a long ON period, it would imply that the prefix path is changing more frequently than others with shorter ON periods. Gaining an estimate of how many prefixes are ON at any instant, would give us an idea of how stable the routes were at that instant. If an external event like a topological change would affect BGP routes, then the instability should be reflected, by a noticeable increase in the ON ratings.

Initially, we set all prefixes at OFF state until path changes turn some prefixes to ON state. For a given prefix, the path used before the prefix entering ON state is recorded and compared with the path used after the prefix returning back to OFF state. If two paths are equal, we define it as a *transient change*, otherwise we call them a *non-transient change*.

Transient change is of interest because it may reflect some unexpected network events, such as a transient failure, which normally will be repaired very quickly, routing slow convergence, and other events.

D. Choice of ON Timer

The choice of ON timer plays an important role in our model. If we choose the timer value to be very small, we might divide an ON period into very small interval succesive ON periods, which means the routing change is still undergoing, but one single ON period cannot cover them. While if we choose the value to be very large, we would extend ON periods too long to cover uncorrelated routing changes. We ran experiments with different values like 5 mins, 10 mins, 15 mins, 20 mins and 1 hour to obtain different results. In this paper, we mainly present the results for 5 mins and 20 minues, which is based on the ditribution of inter-arrival time of path change updates. As shown in Figure 4, at least 50% of updates indicating path changes arrive within 300 seconds from the previous path change update. Therefore if we choose ON timer as 5 minutes, those updates will be covered by one ON period. If we choose ON timer as 20 minutes, more than 65% path change updates will be coalesced with other updates. The curve levels out roughly after 1200 seconds, so the longer timer may not make much differences.

IV. Results

This section shows the results obtained by applying ON/OFF model to historical BGP updates. As we stated eariler that the count of ON prefixes at a particular time could give us a hint about how stable the routes were at that instant. The distribution of ON period reveals how long it will take for a prefix to converge to a new path after a series of updates. Most importantly, we are wondering how many changes are the tranisent changes. The following sections will attempt to answer these questions.

A. Count of ON Prefixes

Initially, it is assumed that every prefix is at OFF state, *i.e.,* the total count of *on* prefixes is 0 at the start point. Whenever a prefix was turned ON, the total count will be increased by one at that time, and

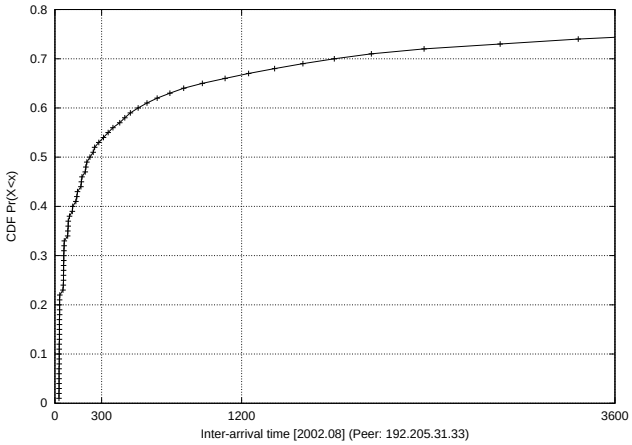


Fig. 4. Distribution of Inter-arrival time of Path Change Updates

the pair $< time, count >$ will be recorded. Similarily, whenever a prefix was turned OFF, the total count will be decreased by one and a new pair will be recorded as well. Figure 5 shows those pairs for August 2002 and September 2001 from ATT point of view, where X-axis represents the time, and Y-axis shows the total count of ON prefixes at that time.

Given an ON timer of 5 minutes, Figure 5(a) shows that on average, the total ON prefixes are less than 500 at any given time with some exceptions. In fact, we totally obtained 193,186 $< time, count >$ pairs in August 2002, and 96% of them have a count which is less than 500. Figure 5(b) shows that when the ON timer is increased to 20 minutes, the count is increased as well. Totally, there are 167,587 pairs, and 70% of them have a count which is less than 500. We have fewer pairs because a longer ON timer may collapse multiple ON periods into one, and consequently reduce both the number of ON and OFF periods. In addition, if a prefix remains at ON period longer, the population of ON prefixes at a particular time will be increased. We also observe a few spikes, some of those will be explained in the section V.

Figure 5(c) shows that count of ON prefixes for September 2001, with a ON timer as 20 minutes. One may quickly notice the sharp increase around September 18, which will be explained later in this paper as well.

B. ON Period Distribution

The duration of a prefix staying at an ON state is counted as an ON period. Figure 6 shows the distribution of ON periods for August 2002 and September 2001. Given the ON timer as 5 minutes, we totally observed 1,067,730 ON periods. Out of those ON periods, as shown in Figure 6(a), 38% are equal to 300 seconds [3], 75% are less than 409 seconds, and 95% are less than 665 seconds. Considering we already artifically add 5 minutes to the ON periods, the actual duration may be even shorter. These numbers may provide a general idea about how long a BGP burst would last.

Given a longer timer as 20 minutes, we have the similar observation: totally, we obtained 758,248 ON periods, and 27% are equal to 1200 seconds, 75% are less than 1519 seconds, 95% are less than 2783 seconds. Note that ON periods tend to be longer in this case, it is because two or more consecutive and closely-spaced ON periods may be combined into one by a long timer.

Although most of ON periods are relatively short-lived, some ON periods last extremely long, even with a short ON timer. Such long-lived ON periods most likely indicate the involved prefixes or networks suffered network problems. For example, the longest ON period obtained from the data is 53348 seconds. The further investigation revealed that one particular prefix has flapped between two paths almost at every minute, sometime being withdrawn, from the August 11th early morning until late night, and the prefix finally ended with being withdrawn. The same thing happened on August 12th again, but the prefix ended with a third path. Based on its flapping pattern and timing information, we conjecture it was caused by a configuraion error for the prefix during those two days. This example shows that ON/OFF model could be used to narrow down to a small set of prefixes which are more worthy to investigate than others.

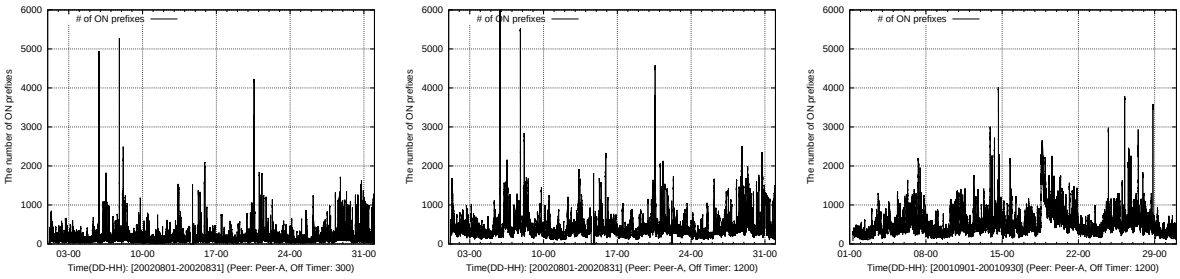
Furthermore, we count how many updates been sent during one ON period. The results show that for 5 minutes ON timer, 75% ON periods contain at most two updates, and 99% contain at most 8 updates. For 20 minutes timer, 75% contain at most 3 updates and up to 15 updates for 99% periods.

C. Transisent Changes

As we described earlier, an ON period implies a prefix was in the course of the convergence of routing changes. We are more concerned about which path will be used after the changes converge. If the new path is different from previously used path, it clearly indicates that either the old path is expericing some failures, or the policy has been changed to prefer a new path. If the new path is the same as the previous one, *i.e.,* a transient change occurred, it means that the routing path has been changed at least once but evetually old path was restored back to the routing table. The causes for such changes are not completely understood, some explanations will be provided in the next section. But first, one may wonder if the transient changes occur frequently, or rarely.

If we count one transition from an OFF period to the next OFF period as one *change*, the transient changes can be indentified if the path used in the an OFF period is the same as the path used in the next OFF period. Figure 7 show the daily percentage of transient changes over total changes in August 2002 and September 2001.

[3]It means that those ON periods only contain one update.

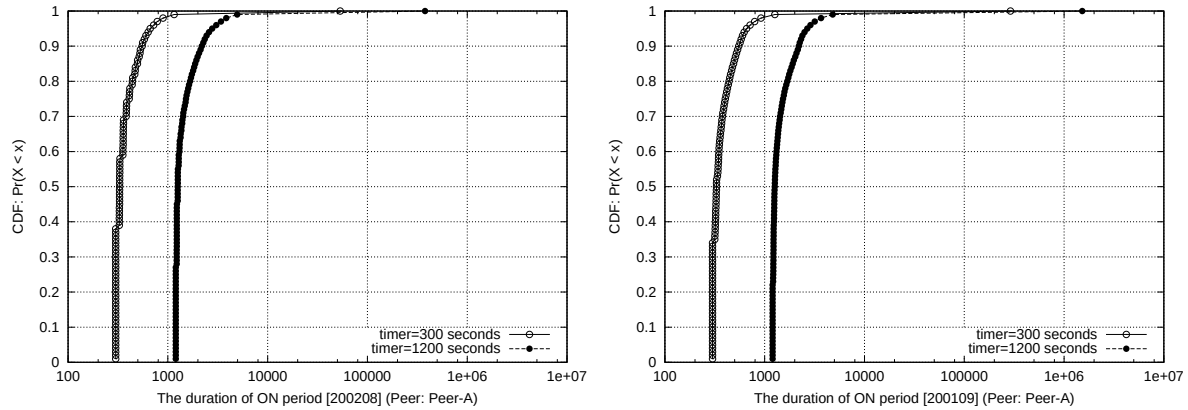


(a) August 2002, Timer=300 seconds

(b) August 2002, Timer=1200 seconds

(c) September 2001, Timer=1200 seconds

Fig. 5. Count of ON Prefixes.



(a) August 2002

(b) September 2001

Fig. 6. Distribution of ON Periods

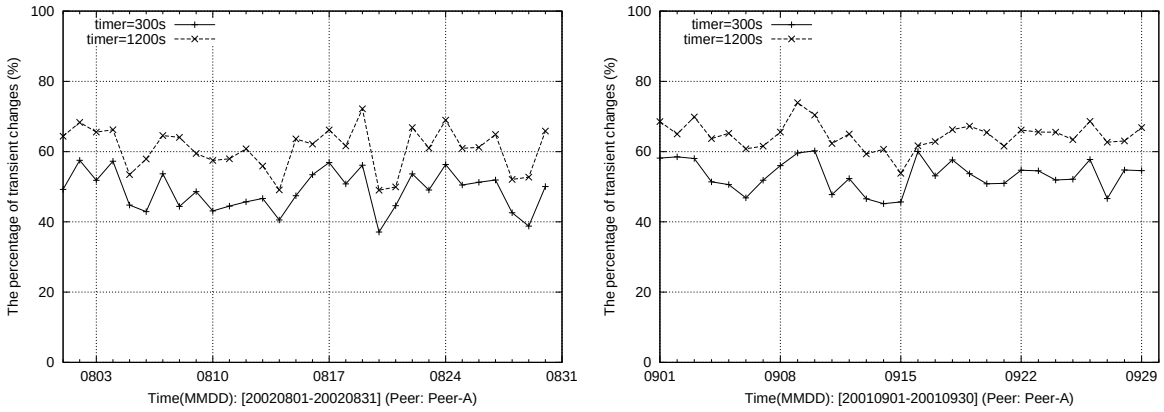
From this figure, we can see that transient changes count for around 50% of total changes every day. It is clear that transient changes happen quite often, which is consistent with observations from other studies [3]. If it were better understood, the causes for such changes may be better controlled, and the protocol may be finer tuned to react to such changes.

Note that a transient change involves at least two updates; the first one turns the prefix ON, and the last one restores the old path. Combining the previous results that most of ON periods only contain 2 or 3 updates, it implies that the type of transient change, which firstly fails over to a new path, then quickly returns back to the old path, counts for a decent number of transient changes.

One may also note that when the timer becomes longer, by comparing two curves in Figure 7(a), the transient changes tend to take more proportion of total changes. It is similar for other months and other peers, as shown in Figure 8(a). We already know that the longer timer will decrease the number of changes, hence the number of transient changes. However, it seems longer timer reduces the total number of changes more than the transient changes. For example, when the ON timer is changed from 5 minutes to 1 hour, the total number of changes is decreased by 44.27%, while the number of transient changes is only decreased by 16.8% in August 2002.

One reason might be a longer timer combines two non-transient changes into one transient changes. For example, a routing change like $Path_1 \rightarrow Path_2 \rightarrow Path_1$ is combined into $Path_1 \rightarrow Path_1$, thus the total number of changes is decreased by one, but the transient changes is actually increased by one. The increase of the proportion of the transient changes by longer timer seems support such explanation.

It is an interesting observation since it reveals that the routing to a prefix seems *stick* to a particular path; whatever the routing changes are and however long it takes, the path tends to be reused evetually. We could test this conjecture from another perspective by counting how many paths being used during OFF periods for each prefixes. A path used during OFF period will be



(a) August 2002

(b) Septermber 2001

Fig. 7. Percentage of Transient Changes

termed as a *non-transient path*[4]. A non-transient path normally will be used to forward traffic for a while, at least longer than ON timer. Therefore, counting the number of non-transient paths will reveal the stability of reaching a prefix. If a prefix is reachable via fewer non-transient paths, espeically if only via one path, it implies that the routing to the prefix is quite stable. When we increase the ON timer, if the conjecture is true, we should see more prefixes only have one non-transient paths. Figure 8(b) shows the results. The X-axis shows the different value for the ON timer, while the Y-axis shows the percentage of prefixes which are with one non-transient path, as well as the percentage which are with more than one non-transient paths. The figure shows that when the ON timer increases, more prefixes tend to be reachable via only one path, which supports our conjecture. Such *stickness* property of the routing system were also observed by other studies, such as [5] observed the path to reach top-level DNS servers are quite stable, and [6] also made the similar observation to popular sites.

V. An Explanation for Transient Changes

This section presents some cases we investigated based on some anomalies captured by our ON/OFF model, as well as the studies of the routing impacts caused by some known network events.

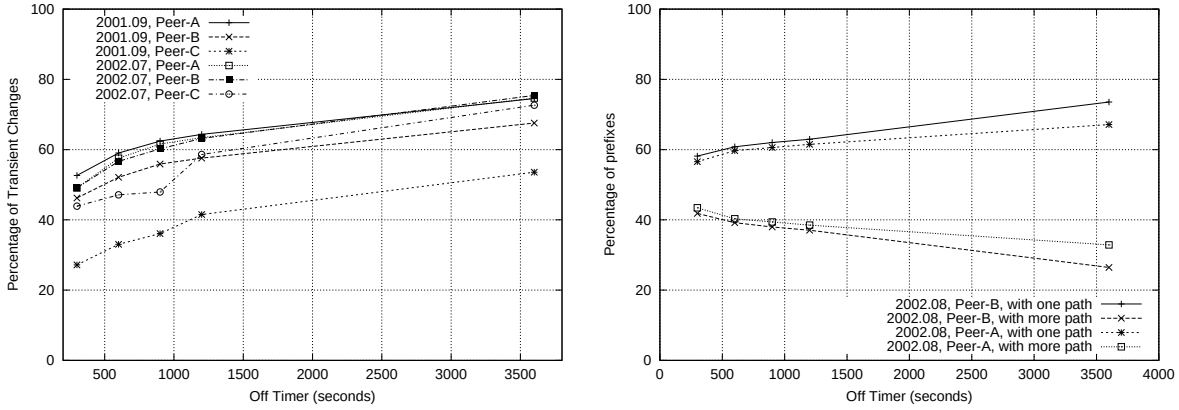
A. Code Red/Nimda Worm Attack

One event around September 18, 2001 that attracted lots of attentions was the Nimda worm. According to the SANS Institute, the scanning activity of the Nimda worm dramatically increased at approximately 1pm GMT on September 18, and abated in the following hours[7]. Effect of Nimda worm on BGP is examined in [8].

Figure 5(c) shows a sharp increase of ON prefixes around September 18, 2001, which an indication of a large number of prefixes experienced routing changes simultaneously on that day. In fact, comparing with the median number of ON prefixes every day in September, the number of changes on the day 18th was increased by 87.6%, from 41242 to 77373. The number of transient changes was increased by 124.64%, from 19701 to 44586. The increase of transient changes indicates that the excessive traffic caused by worm activities affected the routing stability. For example, if AS A multi-homed with two providers, saying B and C. Normally, the incoming traffic followed the path (BA). The worm traffic could congest this path, then A managed to swith to another path (CA) by withdrawing the prefix from B or changing community attributes to inform B or C the new route preference. Quickly, the new path was congested again, A had to switch back, and so on. From the outside point of view, such routing changes are a transient changes, given a proper ON timer.

On July 2001, the Code Red worm attacked the Internet. Our data also show a similar pattern of increased number of transient changes. Note that although the proportion of transient changes over the total changes did increase with a longer ON timer as described in Section IV, the proportion remained almost same as other days, as shown in Figure 7(b). It means that the increase of transient changes are proportional to the total changes, which implies that the number of non-transient changes was also increased by worm traffic. Rethink the above example, now the cases are after A switched to new path (CA), A will stay with the new path longer than the ON timer, which counts for a non-transient change. Figure 7(b) suggests that both cases happened during the worm attack.

[4]A non-transient path is contrast to the path appeared in ON periods, which is considered as a transient path.



(a) Percentage of transient changes with different ON timer

(b) Percentage of prefixes with one or more non-transient paths

Fig. 8. Stickness Property of the Routing System

B. Spikes on Aug. 2002

Figure 5(a) and (b) show there are few spikes of the total number of ON prefixes in August 2002, which may indicate some anomalies. However, if prefixes involved in a spike ended with non-transient changes, it may be caused by legitimate routing changes such as a new path may cause many prefixes switched to it simuteneously. Thus prefixes involved in a spike ended with transient changes are more interesting for an investigation.

An ON prefix could end with a transient change or non-transient change. Figure 9 shows the number of ON prefixes which ended with transient changes. Only one spike was singled out, which occurred around August 7, 19:18 to 19:23 GMT. The further investigation revealed that at that time, there were 2352 prefixes switched to use AS 1239 as a transit AS, but very quickly [5], they switched back to their original paths. Because the transit AS was not used very long, we believe this was caused by some kind of errors, such as misconfigurations.

C. Baltimore Tunnel Fire

On July 18th 2001 at about 18:10 GMT, a 60-car freight train carrying paper, wood and hazardous materials derailed and caught fire in the Baltimore tunnel. The Baltimore tunnel carried communication fibers constituting part of the backbone network, and the fire resulted in damage to these fibers [9]. However, no data is available on exact time of fiber damage, but based on the delay and packet loss observed by operators on nanog mailing list [9], it should be within several hours after the fire.

The recovery and restoration of the fibers in the tunnel, following the fire, took much longer due to the dangerous materious and extremely high temperatures inside(1500 degrees Fahrenheit). It was reported that operators and engineers worked overnight to reroute traffic to other cables and restored service to most customers by through afternoon July 19th, 2001 [9]. Additional fibers were also laid outside the tunnel in order to restore some of the links.

However, based on the communications with network operators, this event is rather regarded as an internal link failure than a link failure between different ASes. Consequently, as an inter-domain routing protocol, BGP would not affected too much. Our model also shows that there are no obvious anomalies observed on that day, which is consistent with the people's common belief.

VI. Related Work

In [3], Olaf and Anja proposed a method to generate realistic BGP traffic in test labs. First, they defined two concepts, *instability originator* and *instability burst*. *Instability originator*, referred to as any routing events affect prefixes, causes *instability burst*, referred to as a series of BGP update due to propagation of changes and updating routing tables. This paper used the similar method as our ON timer to determine the end of the burst, but with a much longer time window as 4000 seconds. Both of their work and our work share some similar results, for example, they also report that most of burst are short-lived, and transient changes are pervasive. But we examined various values of ON timer and its implications, thoroughly studied the path change patterns, as well as we examined much longer time period of BGP data to ensure the statistical significance of the results.

Rexford et. al. [6] studied the routing behavior for popular destinations. To compare the instability for different prefixes, they also merge multiple updates into

[5]The ON periods for those changes ranged from 326 to 384 seconds.

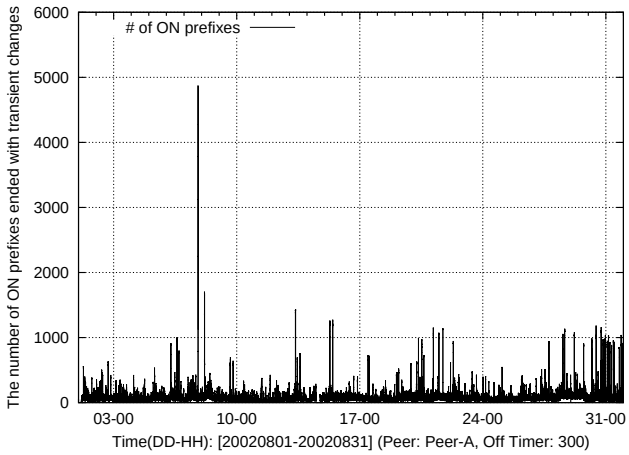


Fig. 9. Count for ON prefixes which ended with a transient change (ON timer = 5 minutes

one "event" based on the space between two consecutive updates. However, a much shorter time window of 45 and 75 seconds was used in their study. They also found that the routes to the popular are quite stable. Along the same line, Lan et. al. [5] studied the reachability and routing changes for top-level DNS servers in order to protect those servers from route spoofing attack. They also found that the routes to top-level DNS servers exhibit quite high stability. Comparing with their studies, we are more interested in BGP behavior for general prefixes because our ultimate goal is to gain a complete understanding of BGP behavior, which should not be limited by a particular set of prefixes.

VII. Conclusion

The sheer size of the BGP log makes it difficult to interpret BGP behavior using only simple analysis tools. In this paper we developed an ON/OFF model to study BGP behavior, which is an initial step toward a sold understanding BGP's performance under both normal and stressful conditions, its response to faults, and its vulnerabilities to attacks. By applying our ON/OFF model, we found that the ON/OFF model is an effective way to identify two type of inputs to BGP system, stable path changes and transient path changes. Transient changes more likely are caused by configuration erros, transient failures, and other unexpected events. And we found such type of changes are pervasive: about half of the update bursts can be classified as transient changes. Overall, the development of the ON/OFF model is an useful tool to help us make a significant step towards a complete understanding of the global routing dynamics.

References

[1] "The Route Views Project," http://www.routeviews.org/.
[2] RIPE, "Routing Information Service," http://www.ripe.net/ris/index.html.
[3] Olaf Maennel and Anja Feldmann, "Realistic bgp traffic for test labs," in *Proceedings of the ACM SIGCOMM*, 2002.
[4] C. Labovitz, A. Ahuja, A. Bose, and F. Jahanian, "Delayed Internet routing convergence," in *Proceedings of the ACM SIGCOMM*, August/September 2000.
[5] L. Wang, X. Zhao, D. Pei, R. Bush, D. Massey, A. Mankin, S. Wu, and L. Zhang, "Protecting bgp routes to top level dns servers," in *ICDCS03*, 2003.
[6] J. Rexford, Jia Wang, Zhen Xiao, and Yin Zhang, "Bgp routing stability of popular destinations," in *Proceedings of the ACM IMW 2002*, Oct. 2002.
[7] Networking System Adminisration and Security Institute (SANS), "Nimda worm/virus report," http://www.incidents.org/react/nimda.pdf.
[8] L. Wang, X. Zhao, D. Pei, R. Bush, D. Massey, A. Mankin, S. Wu, and L. Zhang, "Observation and analysis of bgp behavior under stress," in *Proceedings of the ACM IMW 2002*, 2002.
[9] NANOG, "The North American Network Operators' Group," http://www.nanog.org/.