Toward Understanding the Behavior of BGP During Large-Scale Power Outages

Jun Li, Zhen Wu, and Eric Purpus
{lijun, zwu, epurpus}@cs.uoregon.edu
Computer and Information Science Department
University of Oregon

Abstract— **While the Internet continues to thrive, the resiliency of its fundamental routing infrastructure is not fully understood. In this paper, we analyze the behavior of the *de facto* inter-domain routing protocol, BGP, during a large-scale power outage that affected the connectivity of $3,175$ networks in dozens of cities in the eastern USA and Canada. By observing proper metrics of BGP, we study BGP behavior from both the global level and the prefix level. At the global level, our results show that many global BGP metrics remained stable during the blackout event; importantly, we do not find an increase in the number of BGP announcements, a metric that has been primarily used to indicate significant changes in routing. However, we observe an apparent increase in the number of withdrawals. At the prefix level, we introduce per-prefix AS-path graphs and study their evolution for affected prefixes during the blackout; we have found that in such graphs there is a sharp decrease in the number of edges and nodes as well as changes in node degrees.**

I. Introduction

As the Border Gateway Protocol (BGP) [1] provides the inter-domain routing throughout the Internet—a fundamental functionality for delivering Internet packets across different domains, failure or degradation of BGP can lead to various severe problems, ranging from packet loss or delay to the collapse of the whole Internet. It is therefore critical to understand how BGP reacts to adversity and how resilient BGP is.

As there are many different kinds of adversity, such as large-scale power outages, misconfiguration, and worm attacks, in this paper we study BGP under specific events; in particular, we choose the East Coast blackout of August 2003. Our objective is to characterize and analyze BGP behavior surrounding this event, and try to deduce the resiliency and responsiveness of BGP during such a large-scale adversity. We are particularly interested in discovering any primary distinction of BGP behavior during this blackout.

The East Coast blackout is reportedly the largest of its scale and affected the connectivity of 3,175 networks [2]. Shortly after 4 p.m. EDT on August 14, 2003, this power outage hit a large region in the eastern USA and Canada within two minutes. It was not until 9 p.m. EDT on August 15 that power was completely restored. More details of the outage, such as how networks of various sizes were affected, can be found in [2].

In this paper, we analyze BGP behavior mainly from two complimentary angles: the global-level BGP behavior analysis and the prefix-level analysis. We believe combining the results at these two levels will strengthen our analysis. Global level analysis may reveal the existence of anomalies or differences in BGP behavior. Prefix level analysis may reveal details about the nature of an anomaly or difference. Furthermore, unusual behavior at the prefix level may be hidden at the global level by the noise of the global statistics.

Our paper makes the following contributions. First, we propose several global metrics to characterize macroscopic BGP behavior. We discover that although no significant difference is observed in the number of total BGP updates or other metrics, the number of explicit withdrawal messages makes the blackout stand out. Second, we introduce per-prefix AS-path graphs to study BGP behavior from a microscopic point of view, and our results demonstrate that this is a promising approach to observe prefix-level routing changes or diagnose routing problems.

II. Methodology

A. Overview

Using BGP update data collected before, during, and after the blackout event, we gather BGP statistics at both the global and prefix level. At the global level, we gather statistics in terms of several representative metrics on BGP updates: the number of BGP updates, the number of explicit BGP withdrawals, interval of BGP updates, and top active prefixes. At the prefix level, we create graphs describing the routing paths toward a destination prefix AS from other AS sources in the Internet, *i.e.* per-prefix AS-path graphs, and examine the changes in these paths throughout the blackout event.

For both global-level and prefix-level analysis, we use the data before and after the blackout event to serve as a baseline of assumed normal behavior. Our graphs will have three vertical time-marking lines: the first line marking the beginning of the power outage on August 14, 20:10 UTC, the second the beginning of the incremental power recovery process on August 15, 4:00 UTC, and the third the end of the process on August 16, 2:30 UTC.

B. BGP Updates

BGP routers exchange updates between each other to update the path information for particular prefixes. Updates come in the form of either announcements or withdrawals, where paths are added or removed respectively. Each update is relevant to a list of prefixes that are included in the update. In announcements,

1-4244-0357-X/06/$20.00 ©2006 IEEE
This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE GLOBECOM 2006 proceedings.

each prefix updates include an AS-path to describe the path taken to reach that prefix.

Our study uses the BGP updates from the University of Oregon RouteViews archive [3]. By peering with ASes from around the world, RouteViews monitors continuously collect BGP updates from their BGP peers (those collected during a table dumping process are filtered in our study), resulting in a huge archive of BGP updates regarding reaching prefixes throughout the whole Internet.

C. *Per-Prefix AS-path Graphs*

We construct a per-prefix AS-path graph as follows. Starting with an empty graph, we filter all the updates for a particular prefix during a given window of time. For every announcement, we add the announced path to the graph if it does not already exist, based on the AS-path field for this prefix. The router-id of the announcement is also associated with that path to differentiate between multiple routers from the same AS. For every withdrawal, we match the router-id of the withdrawal with the path in the graph associated with the same router-id, and remove this path from the graph. We use the router-id to do this because the AS-path being withdrawn is not explicitly given, but instead is implied in the withdrawal. At this point, any unconnected AS nodes are removed from the graph.

III. Global-Level Statistics

Since there is already a huge amount of BGP updates propagating throughout the Internet, and only a small amount of BGP updates would be produced for those networks affected by the blackout event, one would expect that hardly any statistical difference of BGP can be observed. This is true as we analyze the results on the number of BGP updates, the interval of BGP updates, and top active prefixes. However, we do observe an apparent increase in the number of explicit BGP withdrawals. Note that because the amount of explicit withdrawals is only about 5% of all BGP updates, changes regarding explicit withdrawals are easier to notice.

A. *Number of BGP Updates*

The number of BGP updates has been a primary metric in observing the dynamics of BGP. Changes in the number of updates over a certain period directly indicates whether BGP peers are relatively chatty or quiet, or even perhaps experiencing some difficulties in exchanging routing information. Previous major events, such as Internet worms, often caused obvious surges in the number of BGP updates ([4], [5]). Naturally, we want to answer the same question: Did the blackout cause a similar surge in the number of BGP updates?

We observe no significant changes in the number of BGP updates during or after the blackout. Figure 1 shows the number of updates received by our RouteViews monitor every hour. Although there is an increase immediately after the power outage, it does not stand out compared to the number of updates when the blackout event is not happening. This is not surprising considering that a power outage is an ON/OFF event.

B. *Interval of BGP updates*

The interval between BGP updates (also called inter-arrival time) can also indicate the level of BGP activity and the general stability of routing paths. When no changes occur to the routing paths, no updates will be exchanged. Conversely, frequent path changes can lead to frequent routing information exchange through BGP updates, thus leading to smaller update intervals. In our measurement, we define the update interval as the time between two updates for the *same* prefix from the *same* RouteViews peer.

We analyze the interval distribution during the blackout period and compare it to two randomly chosen periods of the same length. As shown in Figure 2, we found no significant difference. There are three spikes for all three periods, around interval values of 30, 60, and 90 seconds, and most intervals (about 70-80%) are less than 100 seconds. Both the blackout period distribution and the reference distributions have these spikes, and do not differ significantly from one another. We believe the spikes are the effect of a configuration parameter—the Minimal Route Advertisement Interval (MRAI) timer—that was introduced to control BGP traffic overhead on routers by defining the minimal amount of time between advertisements from a single router for a particular prefix.

C. *Top Active Prefixes*

Due to the ON/OFF nature of the power outage, we suspect that during the blackout period those affected prefixes were *not* significantly more active than those unaffected prefixes. To evaluate our hypothesis, we study the hourly rate of BGP updates for the top $1,000$ active prefixes. As shown in Figure 3, we do not find significant variations during the blackout period. Although higher values exist immediately after the power outage, this less than significant variation can hardly be regarded as correlated with the blackout event. Additionally, our analysis has shown that the top 1,000 prefixes, which constitute less than 1% of the total number of unique prefixes in a routing table, typically contribute almost 40% of the total number of prefix occurrences every hour. And this distribution pattern did not change during the blackout. Finally, our analysis on the top 500 and 100 active prefixes also show that no significant differences can be observed during the blackout.

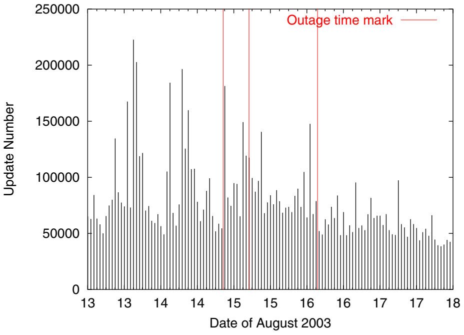

Fig. 1. Number of BGP updates every hour

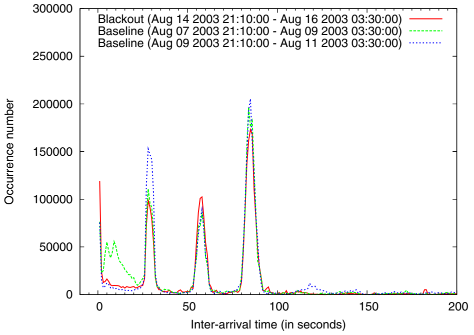

Fig. 2. Update interval distribution

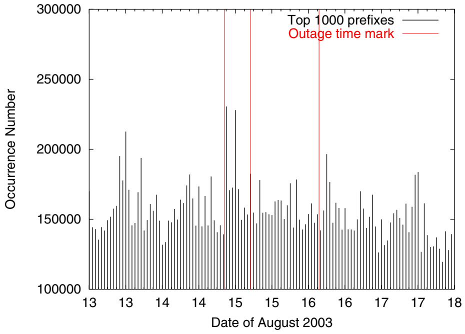

Fig. 3. Top active 1000 prefixes

D. Number of Explicit Withdrawals

We now study the metric that indeed experienced significant variation during the blackout—the number of explicit withdrawals. Withdrawals happen when a given path for reaching a particular prefix is no longer preferred or available. While a BGP peer can announce a different path to reach a prefix, also called an *implicit withdrawal*, the BGP peer can simply inform others through an *explicit withdrawal* that the path toward a given prefix is no longer valid.

Although power backup systems are common, the power outage caused some networks to become unreachable due to the sustained nature of the outage, extending beyond backup capabilities. BGP sessions that involve networks affected by the blackout can thus be broken. When BGP routers that peer with BGP routers from affected networks send out explicit withdrawals to notify other BGP routers, it will result in many withdrawal updates. The number of networks affected during the blackout directly affects the number of withdrawal updates. In the case of the East Coast blackout, as shown in Figure 4, a spike of almost 200% in the number of withdrawals occurred immediately after the power outage.

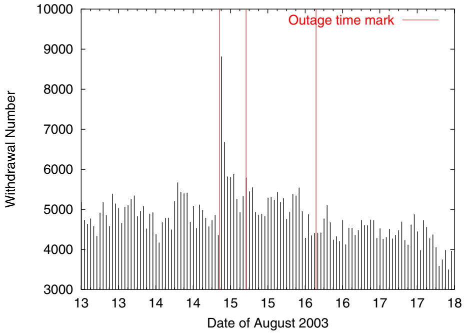

Fig. 4. Number of BGP withdrawals every hour

IV. Prefix-Level Statistics

Global-level BGP analysis can show certain changes during a large-scale failure such as the East coast blackout event, but such an analysis is aggregated over all networks. In order to pinpoint problematic prefixes or perform any prefix-level operations, it is important to zoom in and watch BGP behavior at the individual prefix level. We construct prefix graphs every five minutes before, during, and after the blackout, and record their characteristics at these points. In the following, we first present an example per-prefix AS-path graph, then describe what metrics and which prefixes to use for analysis, and finally present our results and analysis at the prefix level.

A. An Example Per-Prefix AS-Path Graph

Figure 5 shows an example of a per-prefix AS-path graph. The shaded nodes in the graph are ASes which contain a router that our RouteViews monitor directly peers with. The unshaded nodes represent ASes which we learn about from AS paths, but do not peer with. This graph has a number of interesting characteristics:

- AS 12021 is the origin AS, and contains routing loops to itself. These loops are formed when the AS path contains the AS number multiple times in a row. This is commonly known as prepending the origin AS to the AS path, and is often used as a method of traffic engineering for creating longer AS paths.
- AS 3130 announces two different paths—one through AS 2914, and one through AS 1239. Since each BGP router will only announce one path toward a prefix, we can conclude that there are actually two different routers from AS 3130 announcing two different paths. We can identify this behavior by tracking the **out-degree** in the per-prefix graphs. In this case, the node for AS 3130 has an out-degree of two.
- AS 12021 has four peers carrying its traffic toward this particular prefix, and two of those four peers have a larger number of peers routing traffic through them. Here we can characterize this kind of connectivity by **in-degree**. AS 701 has an in-degree of five, while AS 15290 has an in-degree of seven.

B. What Metrics, and Which Prefixes?

To analyze prefix-level BGP behavior during the blackout event, we examine the changes of per-prefix graphs throughout the blackout event. There are two issues to consider: what metrics to use to discover graph changes, and which particular prefixes to investigate.

Per-prefix graphs present us with a multitude of powerful methods for analyzing graph evolution. We start with the following graph characteristics: number of nodes, number of edges, node in-degree, and node out-degree. The number of

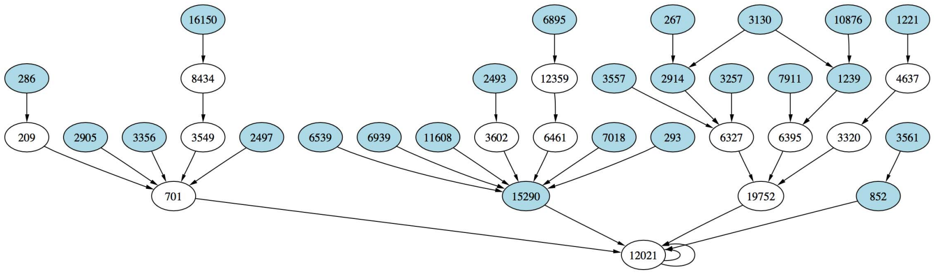

Fig. 5. An example per-prefix AS-path graph

nodes and edges can reflect the scale of the graph. Node degrees can indicate the connectivity of the graph. Variations in node degrees thus will reflect changes regarding how nodes are connected at different times. Note that here *in* and *out* are relative to the direction of the graph oriented towards the origin AS for a given prefix. In a graph where nodes have high in-degree, many ASes choose paths through a common AS. In a graph where nodes have out-degrees greater than one, multiple routers for a single AS choose different paths to reach a given prefix (particularly common for large ASes that span large regions).

Furthermore, we observe that compared to the total number of ASes in a graph, the number of ASes that advertise different routes from more than one router is small. This makes the out-degree of a majority of the nodes exactly one, so in the following we actually do not inspect node out-degree. Meanwhile, with the out-degrees mostly one, the shape of the graph is very tree-like and the number of edges and nodes in the graph are directly related to one another (if it is exactly a tree, the number of nodes is equal to the number of edges plus one). Adding or removing an edge from the graph often requires adding or removing a node as well. Therefore, the number of nodes and edges over time will be almost identical in shape, and we leave out the number of edges in the following.

We analyzed 63 prefixes with our per-prefix graph method. 17 of these prefixes were chosen randomly from ASes which are affected by the blackout according to the Renesys report [2]. The other 46 came from distinct ASes chosen at random. In the following, we use a prefix from an AS affected by the blackout as a representative to report our results and analysis.

C. Number of Nodes

Figure 6 shows the change in the number of nodes at five minute intervals. In particular, we see a drop in the number of nodes between the initial outage and the start of the power restoration. The reason of the drop is as follows. When a BGP router notices that a path to a particular prefix is no longer available, such as due to the unreachability of its next-hop BGP router, it will send out a withdrawal update to its BGP peers to which it has advertised the path. These peers will then further send out withdrawals of paths to this prefix. This phenomenon happens during the blackout when BGP speakers of affected networks become unavailable, sometimes after their power backup runs out. In managing the prefix graph for a particular prefix, every withdrawal of a path toward this prefix will cause a corresponding path to be removed from the graph. After receiving enough withdrawals, there may be no paths to the prefix in question which include a given node. That node will then also be removed from the graph. A decrease in the number of nodes thus indicates that a number of withdrawals for this prefix have occurred.

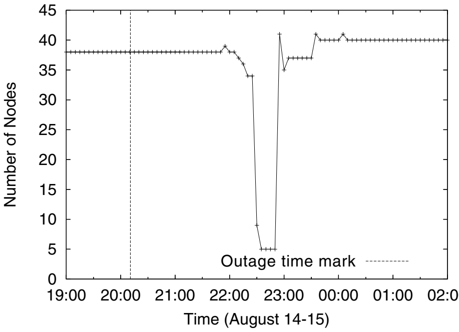

Fig. 6. Number of nodes during the time window

We can also see the recovery shortly after the drop. When a path towards a prefix becomes available again, the path is re-announced and propagates in a fashion similar to withdrawals. In managing the prefix graph, when announcements arrive for paths that do not already exist in the graph, the paths are then added to the graph. Figure 6 shows that the paths are quickly re-announced and the number of nodes remains fairly steady thereafter.

D. Node In-Degree

We look at the distribution of node in-degrees during the blackout period, including the average and maximum of their values. For a given graph of a particular prefix, the average is useful for observing overall graph connectivity, while the maximum shows the in-degree of the most *popular* AS, *i.e.*, the AS with the greatest number of upstream ASes for reaching the prefix.

Figure 7 shows the variation of the maximum in-degree during the blackout: a sudden *increase* followed by a sudden decrease and then a return to the original value. These two sudden changes reflect the changes in connectivity caused by the blackout. Suddenly, 15 ASes all switch to paths that have a common AS as the next hop, but then as the total number of possible paths drops, the maximum in-degree also drops.

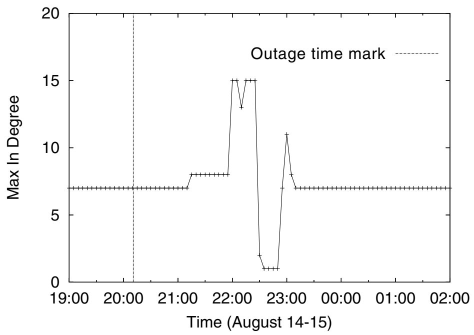

Fig. 7. Maximum in-degree during the East Coast blackout

The first spike in in-degree indicates that many BGP routers had that particular common AS as a secondary route. When the preferred routes became unavailable to these routers, many of them switched to the same route through this common AS. Figure 7 also has a recovery phase. When the blackout ends and the graph recovers to its original level of scale and connectivity shape, the maximum in-degree also returns to its original value.

V. Related Works

The literature contains a number of studies on generic behavior and dynamics of BGP (*e.g.* [6], [7], [8]). Regarding the impact of the East Coast blackout on the Internet, Cowie *et al.* [2] found that 3175 networks were affected by the blackout, with over 2000 networks suffering severe outage for more than 4 hours and over 1400 networks for longer than 12 hours.

BGP during a blackout shares commonalities with other studies on BGP under stress. These studies, including [4], [9], [10], [5], have focused on analyzing the global behavior of BGP, and the stresses are mainly from Internet worms.

Similar to our study using per-prefix AS-path graphs, some works have also tried to study BGP from a microscopic point of view. An earlier work [11] by Govindan *et al.* studied the routing stability in the context of prefix reachability: *prefix availability* and *prefix steadiness*. Researchers also studied the reliability of Internet paths toward a sample set of prefixes [12]. Also similar to our approach to visualizing the detailed routing toward a prefix, another independent work by [13] also proposed to construct AS-level graph representations for reaching a given prefix, which is mainly a visualization tool. Finally, anomalies in reaching particular prefixes have also been studied [14].

VI. Conclusions

While the Internet continues to thrive, BGP as the *de facto* inter-domain routing protocol over the Internet must be studied, especially when it is facing serious failures or attacks. In this paper, we analyzed BGP's behavior during a large-scale power outage that caused $3,175$ networks to lose their connectivity. We identified a number of metrics for analyzing BGP behavior, including both global level and prefix level, and applied these metrics to our analysis. Global metrics studied include number of updates, inter-arrival time, top active prefixes over a time window of two weeks surrounding the blackout period, and explicit withdrawals. We also introduced per-prefix AS-path graphs as another analysis method, and studied changes in graph characteristics such as the number of nodes and node in-degrees. Our results show that during the blackout event, there was an apparent increase in the number of explicit withdrawals, and at the prefix level, the per-prefix AS-path graphs of some prefixes also show a sharp decrease in the number of edges and nodes, as well as changes in node in-degrees. Nonetheless, our results also show that BGP can recover from the blackout event in a timely manner. Overall, the negative impact of the power outage was limited only to the affected areas, and BGP performed well during this large-scale power outage.

We also hope that our BGP behavior analysis techniques in this blackout event can be generalized and applied towards analyzing BGP's resiliency in other security, misconfiguration, or outage events.

References

[1] Y. Rekhter, T. Li, and S. Hares. A Border Gateway Protocol 4 (BGP-4). RFC 4271 (Draft Standard), January 2006.

[2] J. Cowie, A. Ogielski, B. Premore, E. Smith, and T. Underwood. Impact of the 2003 blackouts on Internet communications, November 2003.

[3] University of Oregon Route Views Project. RouteViews. `http://antc.uoregon.edu/route-views/`.

[4] J. Cowie, A. Ogielski, B. Premore, and Y. Yuan. Global routing instabilities during Code Red II and Nimda worm propagation. Technical report, Renesys, 2001.

[5] M. Lad, X. Zhao, B. Zhang, D. Massey, and L. Zhang. An analysis of BGP update surge during Slammer attack. In *Proceedings of the International Workshop on Distributed Computing (IWDC)*, 2003.

[6] C. Labovitz, G. Malan, and F. Jahanian. Internet routing instability. *IEEE/ACM Transactions on Networking*, 6(5):515–528, 1998.

[7] C. Labovitz, A. Ahuja, A. Bose, and F. Jahanian. Delayed Internet routing convergence. In *Proceedings of ACM SIGCOMM*, pages 175–187, 2000.

[8] T. Griffin and B. Premore. An experimental analysis of BGP convergence time. In *Proceedings of the Ninth International Conference on Network Protocols (ICNP'01)*, pages 53–61, November 2001.

[9] J. Cowie, A. Ogielski, B. Premore, and Y. Yuan. Internet worms and global routing instabilities. In *Proceedings of SPIE International symposium on Convergence of IT and Communication*, 2002.

[10] 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 Internet Measurement Workshop*, November 2002.

[11] R. Govindan and A. Reddy. An analysis of Internet inter-domain topology and route stability. In *Proceedings of IEEE INFOCOMM*, April 1997.

[12] X. Zhao, M. Lad, D. Pei, L. Wang, D. Massey, S. Wu, and L. Zhang. Understanding BGP behavior through a study of DoD prefixes. In *Proceedings of DISCEX III*, April 2003.

[13] L. Colitti, G. Battista, I. Marinis, F. Mariani, M. Pizzonia, and M. Patrignani. BGPlay. `http://www.ris.ripe.net/bgplay`.

[14] Ke Zhang, Amy Yen, X. Zhao, D. Massey, Felix S. Wu, and L. Zhang. On detection of anomalous routing dynamics in BGP. In *Proceedings of the International IFIP-TC6 Networking Conference*, pages 259–270, 2004.