Characterization of BGP Recovery Time under Large-Scale Failures

Amit Sahoo
Dept. of Computer Science
Univ. of Califonia, Davis
Davis, CA 95616
Email:asahoo@ucdavis.edu

Krishna Kant
Intel Corporation
Hillsboro,OR 97124
Email: krishna.kant@intel.com

Prasant Mohapatra
Dept. of Computer Science
Univ. of Califonia, Davis
Davis, CA 95616
Email:pmohapatra@ucdavis.edu

***Abstract*— Border gateway protocol (BGP) is the standard routing protocol between various autonomous systems (AS) in the Internet. In the event of a failure, BGP may repeatedly withdraw some routes and advertise new ones until a stable state is reached. It is known that the corresponding *recovery time* could stretch into hundreds of seconds or more for isolated Internet outages and lead to high packet drop rates. In this paper we characterize BGP recovery time under large-scale failure scenarios, perhaps those caused by disastrous natural or man-made events. We show that the recovery time depends on a variety of topological parameters and can be substantial for massive failures. The study provides guidelines on reducing the impact of BGP convergence delay on the Internet.**

Keywords: Autonomous system (AS), Border gateway protocol (BGP), Massive failures, Recovery time.

I. Introduction

BGP (Border Gateway Protocol) is the predominant protocol used for inter-domain routing in the Internet. BGP belongs to the class of *path vector* routing protocols wherein each node maintains multiple ordered paths to reach each destination. One of these paths is chosen at any time according to some given policy. When this primary path fails, BGP withdraws this path and selects the next best backup route. The new route is advertised to its neighbors. However there is no guarantee that the backup route is still valid. In case the backup route has failed, it will be withdrawn only after a withdrawal is sent by the neighbor which advertised it, and another backup route is chosen. This absence of information about the validity of a route can cause BGP to go through a number of backup routes before selecting a valid one. The cycle of withdraws/advertisements can continue for a considerable amount of time and this delay is known as the *recovery time* (or *convergence delay*).

Internet routing sports other classes of routing protocols as well, such as the *link state* and *distance vector* protocols. Link state protocols flood the entire network with information about the cost to reach their immediate neighbors whereas distance vector protocols advertise the cost of the best path for each destination to their neighbors. The flooding approach of link state protocols makes them generally inappropriate for inter-AS use. Distance vector algorithms, on the other hand, suffer from the *count-to-infinity* problem, in which nodes may continuously increase their cost to reach an inaccessible node. Both distance vector and link state protocols are generally used within an autonomous system (AS). Inter-AS routing primarily uses BGP because of its better scalability, flexibility and configurability. In particular, the scalability of BGP has been a critical factor in the explosive growth of the Internet over the last decade.

Numerous studies [7], [8], [4], [5], [12], [15], [19] have been carried out to study the fault tolerance and recovery characteristics of BGP. In particular, it is was shown by Labovitz et al. [8] that the convergence delay for isolated route withdrawals can take > 3 min in 30% of the cases. Zhao et al. [19] showed that packet loss rate can increase by 30x and packet delay by 4x during recovery. There have been some efforts to create analytical models for BGP convergence delay. These studies have identified factors that affect the recovery time and also computed lower and upper bounds. However it is still difficult to estimate the convergence delays for a fault in an arbitrary network. The problem is complicated further if we consider multiple failures and there has not been any work in that area.

The primary reason why large scale failures in the Internet have not been studied is their low probability of occurrence. But it is easy to see that large scale failures can cause a significant disruption to the Internet routing infrastructure, not only in the affected ASes but also in the rest of the Internet. Large scale failures can occur because of a number of reasons such as malicious attacks on the Internet infrastructure, earthquakes, major power outages, massive hurricanes, etc. Recent events have shown that communication networks are needed the most during times of crisis, and that increases the importance of a short convergence delay. Thus it is vital that we have a good understanding of BGP convergence behavior after large scale failures.

A large scale failure would typically take out numerous routers belonging to multiple autonomous systems (ASes).For large scale failures, it is more likely that there is a contiguous area of complete failure. However, scenarios where the affected routers are sparsely distributed over a large area can also be envisioned. In either case, scenarios where only the links (but not the routers) fail are not very likely, and are not considered. Our objective is to study the recovery

characteristics of BGP networks after multiple BGP router failures and to identify the factors that affect the convergence process.

A. Related Work

There has been a fair amount of work on the analysis of BGP convergence properties. However, most publications have examined simple networks or a specific set of sources and destinations only. Although many parameters affecting the convergence time have been identified, it is still not possible to estimate the convergence time for a set of simultaneous failures in an arbitrary network. In this section we talk about the important papers published in this area and the conclusions therein.

Craig Labovitz et al [8]injected faults in the Internet and measured the convergence times from 5 different ASes. The authors observed different convergence times for the same event from different ASes. The authors also computed the lower and upper bounds for the convergence time for a completely connected graph. In a follow-up paper Labovitz et al.[9] concluded that the convergence time for a route is proportional to the length of the longest possible backup path from the source to the destination. Obradovic [12] also arrived at a similar conclusion.

Griffin and Premore [5] studied the effect of BGP's MRAI (*minimum route advertisement interval*) [13] timer on the convergence time after a fault in simple BGP networks. They found that as the value of the MRAI timer is increased, the convergence time first goes down and then increases. The number of update messages however, stabilizes after decreasing initially. The authors concluded that there is an ideal value for the MRAI timer for each source AS, and a fixed default MRAI value will not be optimal in terms of the convergence time.

In this section we have briefly discussed the related work in this area. The rest of the paper is organized as follows. Section II describes the methodology of our study and the assumptions that we made about the network characteristics. We present and discuss the results of our experiments in Section III. Finally we have the conclusion and the references.

II. Study Approach

We used a number of synthesized topologies for our studies and varied their parameters to analyze the effect of these parameters on the recovery times. A modified version of BRITE [11] was used for topology generation and BGP simulations were carried out using SSFNet [14].

A. Topology Generation

BRITE can generate topologies with a configurable number of ASes and with multiple routers in each AS. BRITE supports a number of AS topology generation schemes such as Waxman [17], Albert-Barabasi [1], and GLP [2]. In the Waxman scheme, the probability of two ASes being connected is proportional to the negative exponential function of distance between the two ASes. The Albert-Barabasi and GLP models use *preferential connectivity* and *incremental growth* for edge creation. In these schemes the probability of connecting to a node is proportional to the degree of that node. Both these schemes try to generate a power-law degree distribution, however the results are generally not satisfactory if the number of nodes (ASes) is less than a thousand. We modified BRITE to allow more flexible degree distributions so that we do not have the aforementioned problem and it is possible to assess the impact of degree on recovery time in more controlled settings (e.g., uniform degree, mixture of high and low degree, etc.). We also modified the code to generate variable number of routers for the ASes. The number of routers for each AS was selected from a heavy tailed distribution.

Geographical placement is essential for studying large scale failures since such failures are mostly expected to be geographically contiguous (e.g. an earthquake zone). However, directly using the geography of actual Internet is not only difficult (precise identification & location of routers is a hard problem) but also considerably limits the scenarios that can be studied. Instead, we placed all ASes and their routers on 1000x1000 grid. Studies of real internet have found that the geographical extent of an AS is strongly correlated to the AS size (i.e., number of routers in the AS) [10]. Here we assume a perfect correlation and make the geographical area (the region over which the routers of an AS are placed) of an AS proportional to its size (number of routers). In particular, the routers of the largest AS are distributed over the entire grid. For smaller ASes, the area is reduced proportionately. The routers of an AS are distributed randomly over the geographical area assigned to it. In most cases, we used a uniform distribution; however, we also studied the impact of clustered router placement on recovery time.

Internet studies also show that larger ASes are better connected [16]. This is handled as follows: We first create a sequence of AS degree values according to the selected AS degree distribution and sort them. Similarly, the AS list is also sorted according to the number of routers in the ASes . The inter-AS degree of an AS in the AS list is then set to the value at the corresponding location in the inter-AS degree list. This creates a perfect correlation between AS sizes and degree. Again, although a perfect correlation is unlikely in practice, it is a reasonable approximation for our purposes.

Normally we did not take geographical location into account when creating inter-AS edges, but we did run a few cases where we used a Waxman (distance-based) connectivity function. The ASes are linked together using a pseudo-preferential connectivity in the sense that one of the ends of an edge is selected randomly but the other end is selected according to the degree of the node. Once an inter-AS edge has been created, we randomly select a router from one of the ASes and connected it to the closest router in the other AS. We used the default Waxman scheme for creating the intra-AS edges. However we observed that distance based connections inside the ASes did not have any significant impact on the convergence delays.

With the above changes, we could generate networks with

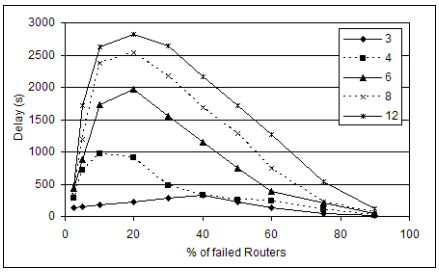

Fig. 1. Recovery time for constant degree networks

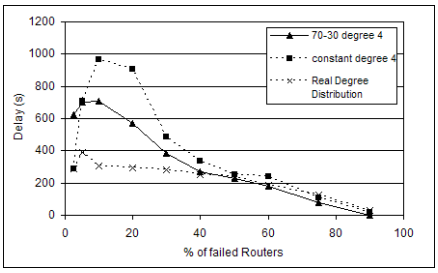

Fig. 2. Recovery time for different degree distributions

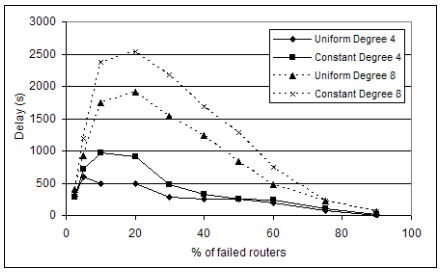

Fig. 3. Recovery time for Constant vs. Uniform degree

arbitrary inter-AS degree distributions. This allowed us to experiment with degree distributions with different decay characteristics, distributions extracted from real networks besides the uniform and constant degree distributions.

B. BGP Simulation

We used SSFNet for our experiments because it has been used extensively in the research community for large scale BGP simulations and BRITE can export topologies in the format used by SSFNet. We used OSPFv2 as the intradomain routing protocol. We used the standard values for the BGP MRAI timers, 30 seconds for eBGP and 0 seconds for iBGP. All the timers were jittered as specified in RFC 1771 [13]and are reduced by up to 25%. We used a mesh of iBGP peering instead of route reflection [6] inside the ASes as the number of routers in the ASes is not very large. We did not model processing and link delays as we wanted to study scenarios where the BGP routers are not overloaded. For the experiments, we simultaneously failed a group of routers and then analyzed the update messages generated by BGP. In most of the cases we failed the routers in a circular region around the center of the map. A failure will lead to both route withdrawals as well as route replacements. We observed the recovery time and messages sent out per unit time, for failures of different magnitudes.

III. Experimental Results

In studying the impact of massive BGP router failures on recovery time, the following parameters are the most relevant:

1) Magnitude of failure, in terms of the number of routers.
2) Inter-AS degree distribution (average degree & its variability).
3) Failure area – contiguous ("area failure") vs. scattered.
4) MRAI value.
5) Extent of the impact of distance on inter-AS connections.
6) Geographic distribution of routers.

We study the effects of these factors through our experiments. We used 200 ASes for our experiments, unless stated otherwise. Each AS had between 1-100 routers drawn from a heavy tailed distribution with an average of about 5. Our initial experiments indicated a considerable variability and complexity in BGP recovery time behavior. Consequently, for the experiments discussed below, we varied only one parameter at a time and also considered several simple topologies in addition to those modeled after the real topologies. We carried out multiple runs for each case, and we plotted the graphs using the average values. We discuss the estimated errors for our simulations in Section III-E.

A. Degree Distribution

For studying the impact of degree distribution, we first examined how the recovery time with a "realistic" degree distribution would compare against one with a constant or uniform degree with the same average value. The average measured inter-AS degree from the Internet AS-level topology is about 8.0 [18]. However, the Internet has over 22000 ASes and the maximum inter-AS degree is in the thousands. For our 200 AS network we decided to restrict the maximum degree to 50 and used the degree distribution in the range 1-50. This gave us a degree distribution which decays as a power law with an exponent of about -1.9. The average degree was very close to 4. We found that, a topology with a constant inter-AS degree equal to 4 yielded a recovery time 5-6 times as high as the realistic case! This prompted us to examine the recovery time as a function of the degree distribution. We started off with topologies in which all the ASes have a constant inter-AS degree. To avoid contamination of results due to other factors, routers were located uniformly in this case and distance wasn't considered while creating the inter-AS edges. Fig. 1 shows the recovery time (in seconds) as a function of the failure magnitude (in terms of fraction of routers failed).

In all cases, the recovery time increases initially with the size of the failure to some maximum value and then slowly rolls off. It is seen that a higher degree consistently increases the recovery time. The sharpness of initial increase also increases with the degree. This happens because, the number of possible backup paths goes up as the degree is increased. The recovery time rises initially because, a larger failure translates into more failed routes and more failed backup routes. However as the number of failed nodes continues to grow, the residual network gets smaller and hence the length of the backup routes explored during the convergence process is shortened. This causes the decline in the convergence delay. It must be noted that the connectivity in the network must keep decreasing with the size of the failure. However, we are only

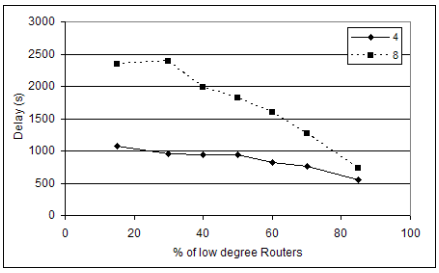

Fig. 4. Recovery time vs. percentage of low degree nodes

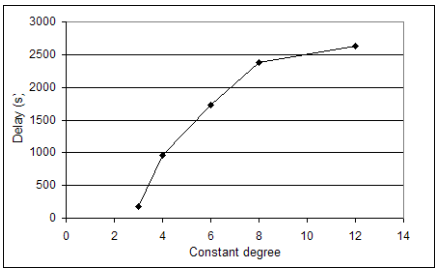

Fig. 5. Recovery time vs. constant degree for 10% failure

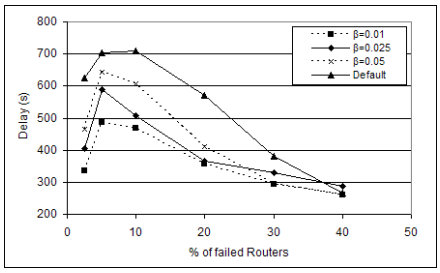

Fig. 6. Recovery time vs. distance based connectivity

looking at the BGP convergence delay here. Common sense dictates that larger failures are less probable than smaller ones. Thus for a network like the Internet, only the left end of the graph might be of any realistic interest.

Two, rather surprising observations can be made from the results. First, for a reasonable connectivity (recall that "realistic" average connectivity is 4), the recovery time shoots up close its maximum value at a much smaller failure percentage (10%) than one would have suspected. The practical implication of this that we don't need truly large failures to experience a very high recovery time. The second surprising observation is that the average degree is not at all a reliable indication of the recovery time. This is illustrated in Fig 2, which compares the convergence delays for the "realistic" topology mentioned earlier, a topology constant inter-AS degree of 4.0 and a third topology (referred to as the 70-30 case henceforth) where 70% of the nodes have low connectivity (1-3) and the other 30% have a high connectivity value (8 or 9) such that the average degree is 4.0. It is seen that variable connectivity helps bring down the maximum recovery time considerably.

The effect of variability in the degree can also be seen in Fig 3. Here we compare the convergence delays for topologies with constant inter-AS degree, with topologies that have a uniform degree distribution but with the same average degree as the constant case. For the uniform case, the inter-AS degree is uniformly distributed in the range [1..2x-1], where x is the desired average degree. Again we see that the convergence delays for the uniform case are significantly lower than constant case.

We have seen that the 70-30 and the uniform distributions yield lower recovery times than the constant connectivity case, and the convergence delays for the power law degree distribution are lower still. The reason for this behavior is that the overall recovery time is a result of two opposing factors with respect to degree:

A Number of routes: Higher degree translates into more routes, which means that during a failure, the number of withdrawn routes as well backup routes is higher.

B Route Lengths: Higher degree nodes on the other hand reduce the distance between other nodes. This helps in a quicker propagation of updates known to one node to other nodes. In other words, high degree nodes can act as "short circuits" and actually help lower the recovery time.

We find that, the effect (A) is generally much stronger than (B). As a result, a uniform increase in the degree of most nodes results in higher recovery times as shown in Fig 1. The effect (B) can be seen in Fig 2 where the convergence delay for the 70-30 case is less than the topology with constant inter-AS degree. Thus, the presence of a small percentage of high degree nodes can provide the beneficial short circuit effect and lower the recovery time. This can be seen more clearly in Fig 4 which shows the maximum recovery time as a function of the fraction of nodes that have a low degree. Recall that in Fig 2 we showed a situation where 70% of nodes have degree in the range 1-3 and others have a higher degree (8 or 9). In Fig 4, we use a similar idea except that percentage of nodes with low degree is varied while maintaining the same average degree. This means that as the fraction of high degree nodes decreases, their degree goes up. Fig. 4 shows the curve for average degrees of both 4.0 and 8.0. It is seen that the curves show a definte decreasing trend. This reinforces the idea that a small number of well connected nodes among a large number of poorly connected nodes forms the ideal situation for low recovery time.

The arguments above still fail to explain why a distribution (e.g., power law) should yield lower recovery time than the fixed low-high mixture of degrees. This result follows by applying the above arguments recursively. We can lower the recovery time by again splitting the high degree fraction into parts: a large subset with lower than average degree, and a smaller subset with a much higher degree. Note that a recursive high-low degree partitioning is akin to cascade multifractal construction and in the limit yields the log-normal distribution.

One issue not addressed above is the behavior of the convergence delay as a function of average degree (with the degree distribution kept the same). This is shown more clearly in Fig. 5 where we show the convergence delay of 10% failure for topologies with constant inter-AS degree. It is seen that the curve shows a diminishing return behavior, which may appear counter to the explanation of effect (A) above. The explanation lies in the fact that the convergence delay depends on the lengths of the longest backup routes explored during the convergence process. If the degree is already high, increasing it further doesn't lead to a proportional increase in the lengths of the longest routes.

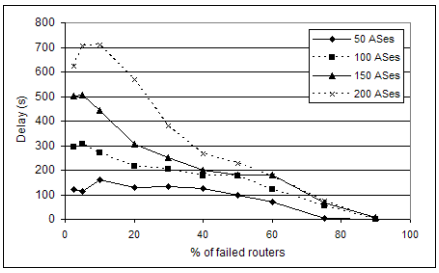

Fig. 7. Recovery time vs. AS size

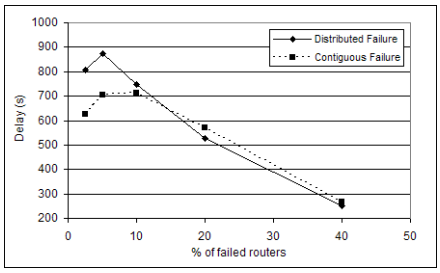

Fig. 8. Effect of distributed failure

B. Distance-based Connections

As stated earlier, in reality, routers connect preferentially to other routers that are nearby [10]. For small ASes, a similar property should hold with respect to AS-AS connectivity. For large, ASes, the concept of "nearby AS" may not be very meaningful; however, for uniformity, we conducted experiments with distance based inter-AS connectivity where the inter-AS distance was defined to be the distance between the "center" of the respective ASes. In our model the largest ASes cover almost the entire area of the map and hence their "location" will always be close to the center of the map. However, the heavy tailed distribution, which is used to generate the number of routers for each AS, ensures that the number of large ASes is small and hence the location is much more meaningful for the rest of the ASes. We used the Waxman connectivity scheme for creating the inter-AS edges. The probability that two ASes are connected was proportional to $e^{-d/\beta M}$ where d is the distance between the "locations" of the two ASes, M is the maximum possible distance and β is a dimensionless parameter. For our experiments we varied the values of β and observed the variation in the convergence delay.

For all the cases we used the same degree distribution: 70% low degree (1-3) nodes and 30% high degree nodes. The overall inter-AS degree for the topologies was equal to 4. Fig 6 shows the results. It is clear that the convergence delay goes down as the decay rate β is decreased, i.e. as the probability of connecting to closer ASes is increased. The reason for the behavior is simple. A decrease in β leads to more links between geographically proximate ASes, and this means that these ASes now have less links connecting them to the rest of the network. The failure of a bunch of ASes in a contiguous area has less effect on the rest of the network, and hence the convergence delays go down.

C. Network Size

In Fig 7 we show the effect of the size of the network on the convergence delay. As expected, we see that the convergence delay increases with the number of ASes in the network. That is because the number and the length of the routes go up with the size. The interesting thing to note here is that the convergence delays go up even if the number of failed routers stays the same as the number of ASes grows. Thus for large networks, even moderate sized area failures could result in long recovery times. Given the continued growth of the Internet, we expect that BGP recovery times will continue to increase. This clearly points to the need for stop-gap mechanisms that can avoid substantial packet losses or route resolution errors during the recovery process.

D. Other Observations

In all the results that we have discussed till now, we considered a contiguous area of failure. However there can be scenarios in which the failed routers are sparsely distributed over a large area. Possible reasons could be a worm attack on the world wide web, an attack on routers sharing the same vulnerable software, etc. So we experimented with a few topologies in which the failed nodes were randomly distributed over the map. The results are shown in Fig. 8. We see that the maximum convergence delays for the distributed failure are greater than that for the contiguous failure case. That is because in a contiguous failure, a number of the failed edges are between the failed routers (intra-AS edges are distance dependent) and hence do not have any effect on the convergence process. That is no the case with a distributed failure and hence the the overall effect is greater.

By default, the router placement in our experiments was uniform over the entire grid. We examined a few cases in which the distribution of the routers was non- uniform. For this, the entire grid was divided up into 5x5 blocks, and within each block a consistent non-uniform placement pattern was used. This pattern made the routers most likely to be located near the center and with decreasing probability towards the edges. No distance based connections were used in this case. It was found that non-uniform placement did not change the convergence time in any significant way.

E. Simulation Errors

In this section we discuss the error estimates for our simulations. Fig. 9 shows the standard deviations for the convergence delays for topologies with the "realistic" degree distribution. For most of the results that we have shown for topologies with 200 ASes and an average degree of 4, the coefficient of variation (V) was within 10%. V was more than 10% for a few cases where the magnitude of failure was low or very high. In general we observed that V was higher for extreme values of the magnitude of failure and lower for medial ones. For small failures, there can be a lot of variability (between the multiple

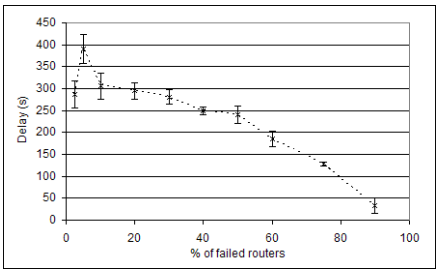

Fig. 9. Real degree distribution (Standard Deviations)

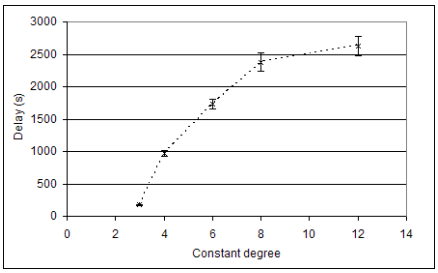

Fig. 10. Recovery time for 10% failure (Standard Deviations)

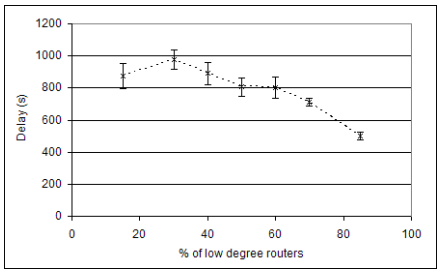

Fig. 11. Effect of degree distribution (Standard Deviations)

runs) in the connectivity of the nodes that we fail and that is the reason the *V* for the recovery time is high. The same holds true for the residual network when we have very large failures. The *V* for these cases can be improved by having more runs for them. In Fig. 10 we show the deviations for "10% failure" convergence delays for topologies with constant degrees. V is less than 10% for all the data points in the figure; however *V* does increase with the degree, and it is high (20-50%) when we have an average degree of 8 or 12 and a low ($<$10%) or a high ($\geq$75%) failure magnitude. Finally we show the "10% failure" convergence delays for topologies with variable degree distributions in Fig. 11. The topologies have the same average degree(4) but the fraction of low degree nodes is different. Again *V* for all the data points is less than 10%.

IV. Conclusions

In this paper we studied the recovery time of BGP for large-scale failure scenarios. The study sheds light on how inter-domain routing in the Internet will behave under natural or man-made disaster scenarios. It was found that the recovery time increases sharply as the magnitude of failure grows to about 10% (of routers)and then rolls off. This means that multiple failures can lead to much longer periods of instability as compared to single failures. Furthermore, even with a fixed number of failed routers, the recovery time increases as the number of ASes increases. Therefore, the recovery time for large scale failures in the Internet can be expected to keep increasing in the future. The paper also points to a number of other interesting aspects about BGP recovery time. In particular, degree distribution has a stronger influence on the recovery time than distance based connectivity or clustered location of routers. Also, a heavy tailed distribution for connectivity (which is present in the Internet today) and distance based connectivity (which is highly likely to exist in the Internet) actually help in bringing down the recovery time.

The future work includes a more thorough study of BGP recovery mechanisms with an aim of devising new schemes to a) reduce the recovery time for large scale failures, and b) to reduce the impact of the recovery process on packet loss and delays.

References

[1] A.L. Barabasi and R. Albert, "Emergence of Scaling in Random Networks," *Science*, pp. 509–512, Oct. 1999.

[2] T. Bu and D. Towsley, "On Distinguishing between Internet Power Law Topology Generators," in *Proc. IEEE INFOCOM 2002*, vol. 2, New York, Jun. 23–27, 2002, pp. 638–647.

[3] "The Border Gateway Protocol". [Online]. Available: http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/bgp.htm

[4] Dan Pei, B. Zhang, et al., "An analysis of convergence delay in path vector routing protocols," *Computer Networks*, vol. 30, no. 3, Feb. 2006, pp. 398–421.

[5] T.G. Griffin and B.J. Premore, "An experimental analysis of BGP convergence time," in *Proc. ICNP 2001*, Riverside, California, Nov. 11–14, 2001, pp. 53–61.

[6] Bassam Halabi, *Internet Routing Architectures*. Cisco Press, 1997.

[7] C. Labovitz, G. R. Malan, and F. Jahanian, "Internet Routing Instability," *IEEE/ACM Transactions on Networking*, vol. 6, no. 5, pp. 515–528, Oct. 1998.

[8] Labovitz, C., Ahuja, et al., "Delayed internet routing convergence," in *Proc. ACM SIGCOMM 2000*, Stockholm, Sweden, Aug. 28–Sep. 1, 2000, pp. 175–187.

[9] C. Labovitz, A. Ahuja, et al., "The Impact of Internet Policy and Topology on Delayed Routing Convergence," in *Proc. IEEE INFOCOM 2001*, vol. 1, Anchorage, Alaska, Apr. 22—26, 2001, pp. 537–546.

[10] A. Lakhina, J.W. Byers, et al., "On the Geographic Location of Internet Resources," *IEEE Journal on Selected Areas in Communications*, vol. 21 , no. 6, pp. 934–948, Aug. 2003.

[11] A. Medina, A. Lakhina, et al., "Brite: Universal topology generation from a user's perspective," in *Proc. MASCOTS 2001*, Cincinnati, Ohio, August 15–18, 2001, pp. 346-353.

[12] D. Obradovic, "Real-time Model and Convergence Time of BGP," in *Proc. IEEE INFOCOM 2002*, vol. 2, New York, Jun. 23–27, 2002, pp. 893–901.

[13] Y. Rekhter and T. Li, "Border Gateway Protocol 4," RFC 1771, Mar. 1995.

[14] "SSFNet: Scalable Simulation Framework". [Online]. Available: http://www.ssfnet.org/

[15] G. Siganos and M. Faloutsos, "Analyzing BGP Policies: Methodology and Tool," in *Proc. IEEE INFOCOM 2004*, vol. 3, Hong Kong, Mar. 7–11, 2004, pp. 1640-1651.

[16] H. Tangmunarunkit, J. Doyle, et al, "Does Size Determine Degree in AS Topology?," *ACM SIGCOMM Computer Communication Review*, vol. 31, issue 5, pp. 7–10, Oct. 2001.

[17] B. Waxman, "Routing of Multipoint Connections," *IEEE Journal on Selected Areas in Communications*, vol. 6, no. 9, pp. 1617–1622, Dec. 1988.

[18] B. Zhang, R. Liu, et al., "Measuring the internet's vital statistics: Collecting the internet AS-level topology ," *ACM SIGCOMM Computer Communication Review*, vol. 35, issue 1, pp. 53–61, Jan. 2005.

[19] X. Zhao, D. Pei, D. Massey, and L. Zhang, "A study on the routing convergence of Latin American networks," in *Proc. LANC 2003*, La Paz, Bolivia, Oct. 4–5, 2003, pp. 35–43.