Global Routing Instabilities Triggered by Code Red II and Nimda Worm Attacks

James Cowie, Andy T. Ogielski, BJ Premore, and Yougu Yuan

Renesys Corporation
Hanover, NH 03750

December 2001

Abstract

We analyze the large, long-lasting, widespread instabilities of the global BGP routing system observed during the Code Red II and Nimda worm attacks in July and September 2001, respectively. The identification and characterization of global routing instabilities employs heuristic spatio-temporal correlation analysis of multiple BGP message streams collected from over 150 autonomous systems' border routers in the RIPE RIS project, and their correlation with the worm traffic is exposed by the analysis of TCP packet traces collected in several /16 networks during the worm attacks. We analyze router failure modes that can be triggered by such abnormal traffic and lead to destabilization of the BGP routing system. To further illustrate the occurrence of cascading routing failures we also present data on another type of global routing instabilities associated with common router misconfigurations generating malformed BGP update messages. Our results show previously unrecognized global routing failure modes, and suggest new research directions.

1 Introduction

The Internet is highly dynamic in two major respects: first, the user traffic exhibits enormous variability that has been extensively measured and studied [31, 36, 35], second, the IP routing topology changes at a relatively high rate. The global routing dynamics is considerably less understood, and to date there have been few measurement-based analyses [30, 28]. Recent availability of dedicated routers collecting raw BGP message traces from many default-free peers at several well-connected locations [39] offers the opportunity to accurately track the dynamic changes of the global routing topology (union of valid best routes from any location to every reachable prefix), delayed only by seconds to minutes. Such a detailed dynamical view of routing changes cannot be easily obtained from the analysis of periodically collected BGP routing tables, which have been successfully exploited to analyze the gross Internet topology [20, 16, 7] and its growth trends [22, 8, 42] at the autonomous system (AS) level.

The focus of this paper is on the emergence of long-lived routing instability events observed in the collective dynamic behavior of the streams of BGP route update messages between June and October 2001. The timestamped BGP message streams collected from 162 border routers from 115 autonomous systems and stored in the RIPE RIS database [39] have been analyzed and correlated at multiple resolutions. In July and September 2001 we detected hours-long periods of exponential growth and decay in the route change rates, across all default-free peering sessions and most prefixes, indicating significant widespread degradation in the end-to-end functioning of the global Internet. Contrary to popular lore, these events did not correlate with localized failures in the Internet infrastructure, such as fiber cuts or power outages. Instead, we have documented a compelling connection between global routing instabilities and the propagation phase of Microsoft worms such as Code Red and Nimda: What were believed to be attacks targeting solely Microsoft servers in fact also turned out to generate widespread end-to-end routing instabilities.

To explain these phenomena, we rely on case reports we received from many network operators and adminis-
Considering that certain properties, such as diameter, remain unchallenged despite growth [20]. Faloutsos, Faloutsos and Faloutsos described power-laws related to AS-level topology metrics such as the vertex-degree distribution [16]. Chen, Chang, Govindan, Jamin, Shenker, and Willinger subsequently showed that using additional data, a more complete picture of Internet topology could be obtained which did not match the power-laws as well [8]. All three of these studies relied heavily on BGP routing tables as a source of data. Broido and Claffy report specifically on the process of analyzing these tables, including using multiple merged tables, and further suggest a new method for reducing redundant data [1, 2]. While we also use BGP data extensively, our focus is not on analyzing the general features of the AS-level Internet topology, but on identification and characterization of instances of global routing instability.

The value of BGP table statistics has prompted a number of free public resources which periodically publish such information. Notable examples are those of Huston [21], NLANR [17], and APNIC [5]. Typical statistics are forwarding table size (number of prefixes), AS and prefix path length distribution, and prefix aggregation information. While these are valuable statistics, our analysis required a more detailed spatio-temporal look at BGP message arrivals.

The first research to look specifically at BGP messages was conducted by Labovitz, Malan and Jahanian [30]. Their data came from core border routers, and they observed that an extremely large fraction of BGP updates were pathological. Their continued study identified several origins of the misbehavior and eventually resulted in major reductions in BGP update volume [28]. While we too have observed continuously large volumes of BGP traffic, we take particular interest in anomalies in the collective dynamic behavior of BGP message traffic. In this report, we restrict ourselves to a search for the causes of these anomalies, which stand out in stark contrast to the typical pulse of inter-domain routing traffic.

3 Methodological Background

To set the context, consider that at the end of 2001 the global Internet has been a coalition of over 12,000 autonomous systems (ASes). An AS – a corporation, network provider, or another entity – is an administratively closed collection of IP networks communicating using a unified internal routing policy, that is globally identified by an assigned autonomous system number (ASN). Worldwide IP communication among the ASes is established in a multitude of business agreements, exclusively using the Border Gateway Protocol (BGP) [38] for distributing the routing information among BGP routers in...
all ASes. Therefore, BGP failures and instabilities immediately affect the connectivity of the global Internet.

When a BGP router’s preferred route to a given network prefix has changed, it sends out a BGP update message to each connected peer router. Therefore, by establishing BGP peering connections with a large number of BGP routers from well-connected organizations, analysis of traffic gathered at a single BGP monitoring point can provide a great deal of information about the way those organizations view the Internet, and about the dynamics of how paths change over a wide range of time scales. The RIPE RIS project [39] maintains several such monitoring points across Europe; they peer with many of the so-called global tier-1 providers, plus very many smaller regional European networks. Access to multiple BGP monitoring points provides opportunities to filter the localized infrastructure failures that are close to individual collection points, clearing the way to unambiguously identify and study routing instability features that affect large portions of the Internet simultaneously.

The results shown in this article are primarily based on analysis of time-stamped BGP messages collected at the RIPE NCC site (rrc00) and the AMS-IX site (rrc03), both in Amsterdam, the Netherlands. We also analyzed BGP traffic from other Internet exchanges that host RIPE RIS collection sites, including LINX (London), SFINX (Paris), CIXP (Geneva), and VIX (Vienna). Details of the extended analysis are omitted for brevity. The raw BGP message archives at RIPE contain various errors and anomalies that require certain amount of care in their analysis, in particular: occasional timestamp clock shifts, missing data, corrupted MRT headers, truncated BGP messages, and the opening and closing of collecting routers’ BGP sessions.

The RIPE NCC facility is particularly interesting, as it collects BGP routing updates from several large Internet providers via multi-hop links, and thus provides a good and fairly complete dynamic view – second by second – of the evolving state of global routing. The autonomous systems that had routers peering with the RIPE NCC collecting router in the period covered in this paper are listed in Table 1.

Table 1: Autonomous systems peering with RIPE NCC router.

<table>
<thead>
<tr>
<th>AS</th>
<th>Organization</th>
</tr>
</thead>
<tbody>
<tr>
<td>AS286</td>
<td>KPNQwest Backbone</td>
</tr>
<tr>
<td>AS513</td>
<td>CERN</td>
</tr>
<tr>
<td>AS1103</td>
<td>SURFnet</td>
</tr>
<tr>
<td>AS2914</td>
<td>Verio</td>
</tr>
<tr>
<td>AS3257</td>
<td>Tiscali Global</td>
</tr>
<tr>
<td>AS3549</td>
<td>Global Crossing UK</td>
</tr>
<tr>
<td>AS3549</td>
<td>Global Crossing USA</td>
</tr>
<tr>
<td>AS4608</td>
<td>Telstra Internet</td>
</tr>
<tr>
<td>AS4777</td>
<td>APNIC Tokyo</td>
</tr>
<tr>
<td>AS7018</td>
<td>AT&amp;T Internet4</td>
</tr>
<tr>
<td>AS9177</td>
<td>Nextra (Schweiz)</td>
</tr>
<tr>
<td>AS13129</td>
<td>Global Access Telecommunications</td>
</tr>
</tbody>
</table>

The BGP protocol contains route flap dampening mechanism that prevents a BGP router from sending too many messages about an unstable route [43]; and a timer (the Minimum Route Advertisement Interval Timer) that maintains a minimum separation between consecutive announcements to a given peer, with default value of 30 seconds. Therefore, if we see large increases in the number of BGP update messages, it’s an unambiguous sign that the diversity of network prefixes is rising. Furthermore, the duration of these BGP message traffic surges, and the rate of their growth are what distinguish global instabilities from the pervasive background noise, daily rhythms and localized failures.

We operationally define a global routing instability in terms of its rate, duration and diversity as follows: Exponential or similarly fast growth of the rate of prefix updates, high update rates lasting hours to days, with almost all prefixes churning, in BGP updates from almost all default-free peers.

Very short, high spikes in advertisement rates are very common – whenever a peer BGP session undergoes a hard reset, for example, a full table dump will follow. More surprisingly, our examination of the data indicates that localized failures in the core Internet infrastructure (fiber cuts, flooding, power failures, building collapses) tend to generate only short-term increases in the external BGP prefix advertisement rate, which decline in a matter of minutes as the highly redundant peering in the core Internet topology routes around the damage, although specific networks may remain unreachable until the damage is repaired.

Of far greater concern are the appearance of sustained rises in aggregate BGP update rates that last for hours. To expose their mechanisms, in this work we look for route rate change correlations among peers, origin ASes, prefixes, prefix lengths, and route lifetimes.
4 Global Routing Stability: Summer of 2001

Figure 1 shows the trends in the aggregated rate of BGP prefix advertisements received from all peers of the RIPE NCC collection router from 1 June through 24 September 2001. This time series represents a measure of the coarse-resolution route change activity for the entire Internet. The aggregation exposes several strong patterns and features, for example:

- One can observe strong weekly and daily trends in the median, an effect which may be due to interactions with either the diurnal patterns of traffic, or the diurnal patterns of activity by network operators performing routine maintenance on BGP routers. The weekly trend ramps up from Monday through Wednesday, and then steadily declines towards the weekend.

- Very common high and narrow peaks in update rates also tend to follow a weekly trend, thinning over the weekends, and are assumed to be dominated by router reconfigurations, session resets and other network maintenance, although other time-localized causes cannot be excluded at this time.

- Two strange non-periodic features jump out when the baseline is examined:
  - A nearly tenfold magnitude rise in the baseline on 19 July (see Figure 2)
  - A more rapidly rising and longer-lasting rise in the baseline on 18-22 September (see Figure 5).

Note that the above aggregated time series does not serve as a measurement of reachability over time—this would be achieved for instance by plotting the fraction of prefixes with valid routes, and watching for dips.

At this aggregation level, and in the absence of internal BGP updates in the aggregate, there are few interesting features when network infrastructure is broken at localized geographic points. Neither the Baltimore tunnel train wreck that severed multiple fiber links on 18 July nor the attacks of 11 September appear as features in this plot. These events did not destabilize the global Internet. In general, the high levels of routing activity following fiber cuts between tier-1 and other major providers remain localized within the immediately affected autonomous systems, and do not create external BGP message storms that are highly visible worldwide.
Figure 3: Left: BGP prefix advertisements. Right: BGP prefix withdrawals. Each row represents the time series from one BGP peer in RIPE NCC aggregated in 60-minute bins. Note the high correlation of the waves of withdrawals and advertisements across all peers on 19 July.

Figure 4: Code Red II port 80 probes per hour recorded in two unrelated /16 networks.

5 Code Red II and Nimda Events

In particular, we are concerned with the two non-periodic features visible on 19 – 20 July and 18 – 22 September. These two “storms” in BGP update rates correlate with the propagation phases of the Microsoft worms known as Code Red II (in July) and Nimda (in September).

On 19 July, we observed an exponentially growing eight-fold increase in the advertisement rate, over a period of about eight hours (all times are in GMT; subtract 4 hours for EDT). This surge faded over the same time scale as it arrived. When one considers the current estimates of BGP convergence times (several minutes [29]), it is more than a little disturbing to see a fundamental quantity like BGP advertisement rate exhibiting exponential growth for eight hours. One initial guess was a delayed effect from the 18 July Baltimore tunnel fiber cuts, and whose impact was highly visible in the discussions on network operators’ mailing lists such as NANOG [32], as network engineers tweaked routing for the next day or two. But this does not appear to have caused the BGP storm.

In order to gain a better understanding of the mechanism driving this BGP storm, we conduct a finer analysis. We began by separating the contributions of individual BGP peering sessions to the total BGP update message traffic, and by separately following the time courses of BGP route advertisement and withdrawal messages.

In Figure 3, BGP prefix advertisements and withdrawals are graphed along the z-axis as impulses. The x-axis is time, from 1 July through 31 July. The y-axis
going into the page) separates the contributions of the 13 individual peer ASes at the rrc00 collecting router. On other days, it is common for individual peers to contribute spikes of high prefix count, reflecting BGP sessions closing and opening close to the collection point. On 19 July, however, all peers experience a wave of updates, sustained for many hours.

Further analysis of the BGP message traffic has indicated that no specific autonomous system or set of autonomous systems seems to be generating the traffic surge, and that no specific IP prefix or set of prefixes was flapping significantly more than before the onset of the surge (see Section 5). Instead, the net effect was that routes to most of the approximately 110,000 prefixes in the Internet were changing more than normal: The data reflect a broad-based BGP update storm with no apparent single-point cause.

**Correlation with Code Red II Attack**

The time course of the 19 July BGP storm suggests that it has been triggered by the sudden spread of a variant of the Microsoft worm known as Code Red II [13]. Our analysis has been materially aided by quantitative data and qualitative reports on the network security mailing lists [23, 26, 33] concerning worm scans rates, sudden connectivity losses, ARP storms, and other worm effects. Ideal data on the worm-generated traffic storm would show the time series of worm activity for a good statistical sample of networks of known size, so that the global activity levels could be inferred by extrapolation. However, only a limited number of such datasets are available.

Figure 4 shows the Code Red II propagation data collected independently on two /16 networks (each nominally containing 64k IP addresses) during the entire day of 19 July. The original data were obtained from [19, 15], also summarized in [24, 25]. The two time series show the number of TCP SYN packets received in the two distinct /16 networks hour after hour. Note the virtually identical time course of these attacks as seen from different networks. These plots give a measure of the intensity of worm scanning traffic at all affected networks. We also analyzed the tcpdump trace for the network shown on the left [14], which further revealed very high diversity of source and destination IP addresses in the worm scan packets: during most of the worm attack on this /16 network there were over 3,000 unique attacking hosts in any one-minute interval, at rates exceeding 8,000 probes per minute. On a longer time scale, the corresponding aggregate rates were over 100,000 unique attacking hosts per hour, and about 500,000 probes per hour, that were probing over 90% of the address space in this network. We will return to this point later.

Further information on the Code Red II virus is available from [18].

**Nimda Worm Attack**

On Tuesday, 18 September, simultaneous with the onset of the propagation phase of the Nimda worm, we observed another BGP storm. This one came on faster, rode the trend higher, and then turned itself off, though much more slowly. Over a period of roughly two hours, starting at about 13:00 GMT, the rrc00 aggregate BGP advertisement rates exponentially ramped up by a factor of 25, from 400 per minute to 10,000 per minute, with sustained “gusts” to more than 200,000 per minute. The advertisement rate then decayed gradually over several days, reaching pre-Nimda levels by 24 September.

Figure 6: BGP prefix withdrawals in 15-minute periods. Each row is one BGP peer in RIPE NCC. Note the steep onset of a wave of withdrawals on 18 September. Plotted in the back is the Nimda worm scan rate (not to scale) observed in two /16 networks.
The analysis of the BGP storm triggered by the Nimda worm follows a similar course as the analysis of Code Red presented above. By separating the contributions of individual BGP peering sessions to the total BGP update message traffic at RIPE NCC collecting router, and by separately following the time courses of BGP route advertisement and withdrawal messages we can demonstrate correlation across all peers. Similar analyses were performed also for the other data collecting routers.

The steep exponential growth of the 18 September BGP storm is aligned with the exponential spread of Nimda, the most virulent Microsoft worm seen to that date. The Nimda worm exhibits extremely high scan rates, multiple attack modes generating very heavy traffic, and has been much more damaging than the July Code Red worm [37, 41].

We have also analyzed the tcpdump traces of the Nimda worm attack on a /16 network [14]. The time series in Figure 9 illustrates the trend in the worm scanning rate. Notice a faster onset than with the 19 July Code Red II attack. It is interesting to speculate that narrow peaks in the worm scan rate on the three or four occasions preceding the actual attack may represent trials of the worm.

6 Network Reachability Failures During the Code Red and Nimda Attacks

The temporal correlation of the worm attacks and routing instabilities shown above does not tell much about the statistics of affected prefixes and routes, and in particular does not characterize the populations of high churn networks. In this Section we argue that Nimda and Code Red triggered long-term BGP instabilities unlike any localized network failure, in particular that a detailed analysis shows no suspect prefixes—most prefixes churn, and no suspect routes—most routes churn. Since the analysis is qualitatively similar for both worm attacks, we will not duplicate each example.

Figure 7 compares the populations of prefixes churning during a quiet hour (weekend night of July 7, with very low rates of BGP updates) and during the peak hour of the Code Red II attack (July 19). The data represent all prefixes seen in the BGP updates (announcements and withdrawals) at the AMS-IX collecting router (rrc03). It is seen that the primary difference is in the sizes of prefix populations, with no other obvious patterns that might localize the sources of routing instability. We have also analyzed the count of withdrawals and announcements for each prefix shown in Figure 7, with the result that the primary difference is in their number, and again no obvious pattern can be found.

Figure 8 shows that all prefix lengths were similarly affected by the effects of the Nimda worm attack on global routing stability.

A preliminary comparative analysis of the BGP routing tables collected a few hours before and a few hours after the onset of the Nimda storm shows that the increase in prefix withdrawals during the worm period, while significant, did not result in long-lasting losses of reachability to the edge networks. These were, by and large, transient failures.

7 Mechanisms of BGP Router Failures During the Worm Attacks

In general, long–lasting, high–diversity, high–rate route churn can be produced by any mechanism that is causing a large number of BGP sessions to close and reopen repeatedly.

At this time we are not aware of any direct measurements showing step-by-step how the the worm propagation traffic causes BGP router failures, moreover, it is expected that different router models and software releases will show different failure modes, and different failure recovery. However, there is strong evidence that BGP sessions do fail under several stress conditions (such as excessive memory demands [6]), and due to software bugs. The router vendor advisories [11, 9, 10] list the primary causes of BGP protocol and router failures:

- router CPU overload
- out of memory
- overflows
- router software bugs

Case reports obtained by us from network engineers and administrators who observed and described the effects of the Code Red and Nimda worm attacks, as well as the router vendor advisories, indicate that while these particular worms do not infect the routers themselves, they do impact the routers indirectly via CPU overload (for example, protocol and slow path processing, high number of flows, high interrupt rates), excessive memory demands (for example, large numbers of routes or flows), and cache overflows (for example, ARP storms). Each of these may

\[1\] Another mechanism would require a very large number of non-BGP originated, globally announced networks to keep failing and recovering repeatedly
cause either a BGP session failure, or the router failure. The following list shows the aspects of the worm traffic, and worm-induced BGP traffic, that may lead to such failure conditions:

- traffic diversity (number of flows),
- traffic intensity and congestion losses,
- high BGP message load,
- vulnerability of HTTP servers in routers (management interfaces),
- failures in network gear (such as DSL routers and other components),
- IGP (Intra-AS) flapping and routing failures,
- proactive reconfiguration or disconnection of routers by operators.

Among the case reports we obtained detailed measurements of the effect of the worm traffic on router processing load. Figure 9 shows a striking correlation between the Nimda worm scan rate in one /16 corporate network in the Midwest, and a router CPU utilization in a tier-1 provider’s network on the West Coast. This router's CPU utilization increases four–fold from 10% to over 40% at the peak of Nimda attack.

So far, data analysis indicates that the most likely common cause of worm-induced router failures is the worm traffic diversity, that is abnormally high number of IP packet source-destination address pairs seen in a short time under conditions of heavy scanning traffic. This peculiar traffic pattern may not be included in standard router test scenarios, leaving the routers highly vulnerable to such traffic conditions. While failure details may vary, one failure mode may involve generation of extremely many short flow records in a very short time, stressing both the CPU and router memory. Another failure mode may be triggered by ARP storms on directly attached networks targeted by the worm probes.

Once BGP traffic increases significantly due to multiple router failures, another BGP router failure scenario becomes possible: Extremely high BGP update rates, impacting the router CPU and memory, can either lead to BGP message losses or to malloc failures. Such a scenario is particularly troubling, as it has the potential for triggering cascading failures.

Finally, it has to be noted that exponential spread of the worm may cause a large number of network operators from corporations and smaller ISPs at the Internet’s edge to independently shut down, or reboot, or attempt to
reconfigure their border routers; and the total amount of BGP message traffic grows exponentially with the number of edge domains feeling the effects of the worm.

However, according to simple bandwidth utilization estimates based on measured worm scan rates it appears that congestion losses of BGP messages are not a likely effect with these particular worm spread scenarios. Although one would expect BGP messages to be very high priority traffic, and thus not subject to congestion-related loss until the situation were dire indeed (such prioritization has been shown to be a good thing [40]), it’s not clear that network operators routinely enable this kind of prioritization.

In summary, we tentatively conclude that due to large diversity of network hardware and software the worm-induced routing instability results from the combination of effects listed above. This does not make it easy to propose a simple conceptual model amenable to theoretical or simulation-based analyses. However, we have also observed a common occurrence of relatively simpler global routing instabilities that may be much easier to analyze. They are briefly described in the following section.

8 Misconfiguration Instabilities

According to BGP standards [38], upon receiving a malformed BGP message the BGP speaker must send a notification to the offending peer, and subsequently close this BGP session. The standards do not specify how quickly this session may be re-established. However, not all deployed BGP routers conform to the RFC, creating the possibility of the following scenario and its variants:

1. A misconfigured router starts announcing a private AS number in a (confederation) ASPATH.
2. Certain routers ignore but propagate the malformed route
3. Other, RFC-compliant routers close and reopen the BGP sessions.
4. The combination may propagate wildly depending on the routers in the neighborhood.
5. Instability ends only when the original leak is plugged.

We have recorded numerous cases of BGP instabilities that were correlated with malformed ASPATH announcements, for instance in Figures 3 and 6 they can be seen as smaller “instability waves” correlated across all peers. In Figure 7 we show a recent instance recorded on October
8 that has developed into a substantial global instability event. We are currently investigating this in greater detail, to be reported separately. We have already begun building simulation models to understand the spread of such instabilities in greater detail.

9 Conclusions and Future Research

Analysis of BGP message streams collected from multiple routers is only beginning, and it is fair to say that we have barely scratched the surface. At this point, several broad observations are in order:

First, there is a strong interdependence between the routing topology and the end-user traffic: Route fluctuations and failures necessarily lead to data traffic failures along affected paths, and, vice versa, certain abnormal data traffic conditions may trigger routing failures as we show in this paper. These interactions are addressed to some extent in this article, and should be further investigated in future research.

Second, this work demonstrates the importance of correlating network measurements concerning both routing and user traffic. Not doing so may well result in incorrect analyses, as recently demonstrated by a commercial traffic monitoring service [27, 34] who denied any impact of the worm attacks on the Internet performance by relying solely on monitoring test traffic latencies and losses among agents placed in the largest Internet core backbones, and thus entirely missing the action.

In the context of large network instabilities, it is of interest to notice that according to recent research on the dynamics of complex engineered systems [4, 3] the Internet is likely to exhibit unexpected catastrophic failure modes.

The empirical analysis of BGP routing failures demonstrated in this paper directly leads to a number of new research directions. Among those are: the need for standardizing the router behavior during recovery from failures to prevent same failure recurrency, graceful degradation of routing protocols running on routers faced with excessive demands on their CPU and memory resources, and detection and containment of attacks on the routing infrastructure.

10 Acknowledgments

Thanks to Henk Uijterwaal and his group for the the RIPE RIS project which has been the source of the raw BGP message traffic archives explored in this paper. Vicki Irwin, Ken Eichman and Vern Paxson kindly offered us worm traffic traces from several /16 networks. Special thanks are due to very many network engineers and administrators for sending us detailed case histories and observations on router misbehaviors under worm traffic stress conditions. Finally, we acknowledge interesting discussions with Dave Donoho, Tim Griffin, and others at the 2001 Leiden Workshop on Multiresolution Analysis of Global Internet Measurements.

References


