A practical study of the problems of current Internet routing tables

Arnav Ghosh, Bruce Hartpence, Daryl Johnson
School of Informatics,
Rochester Institute of Technology
Rochester, NY USA

Abstract - The phenomenal growth of the Internet amazes even the creators of this worldwide network. Apart from the constantly changing data protocols and services that the Internet has to adapt to, the sheer volume of users has been one of the biggest challenges the Internet is coping with. This paper is directed towards the study of Internet scale routing tables in a lab environment to understand the dynamics of route processing by routers, and the effect of increasing the number of routing table entries on the overall performance of the router in terms of packet forwarding. This study is an effort to simulate an Internet scale network in the lab to shed light on some of the practical problems of the Internet routing table size and its performance and the security implications.

1. Introduction

In the 1970s, an experimental network known as ARPANET was started to connect a few research universities and some government bodies. It was an experiment which never ended and has evolved to become the world’s largest network; the Internet. The incredible achievement of the Internet is that it enables us to access information and data from all around the globe within seconds, in spite of traffic variations and departure from the original scope. The Internet as we know is a huge mesh, interconnecting networks all around the world. The meshed Internet architecture makes it inherently resilient to most network failures and is the strength of the Internet, making it tolerant to link failures, node failures and sometimes even network segment failures.

The Internet is based on Internet Protocol (IP) addressing which was first developed in 1980. IP addresses are logical addresses that make identification and communication possible across multiple networks. IPv4 addresses were initially divided into classful groups of A, B, C, D and E. Classful addressing and routing protocols meant that addresses in a particular class would be advertised to different networks across classful boundaries with the class specific subnet mask. This was inefficient and resulted in wastage of IPv4 addresses. In 1993, researchers came up with Classless Inter-Domain Routing (CIDR), which allowed networks belonging to a specific class being advertised with subnet mask of varying lengths. This was a great step and allowed efficient use of the IPv4 address space. CIDR was being widely used all around the world but as more and more IPv4 addresses were allocated, a hidden problem of routing table explosion emerged even though the routing table entries could be aggregated. Since CIDR allowed the propagation of IPv4 networks with any possible valid subnet masks, ISPs, enterprises, universities, etc started advertising their network block in smaller chunks rather than one huge network. For example, Apple Computer Inc. owns the entire 17/8 subnet and would advertise a single 17.0.0.0/255.0.0.0 block before the existence of CIDR. At present there are more than a hundred route entries in the 17/8 network which is advertised on the Internet which can be verified by using a publicly accessible router server at http://www.routeviews.org [13].

The sheer volume of users and nodes that are now connected to the Internet is the other important factor that has led to the huge increase in the current Internet routing table size. There are 1,966,514,816 Internet users as of June 30th, 2010 [15]. This means almost 1/3rd of the world’s population are now connected to the Internet. The Internet has penetrated almost each and every country in the world. This wide spread geography of the Internet meant that large chunks of IP addresses would have to be broken down to cater the needs of individuals around the globe. This has led to the wide spread use of CIDR to break down classful addresses into smaller subnets to be distributed among various nations, ISPs, cities, businesses and homes. This exponential increase in the users and nodes connected to the Internet has eventually led to large number of advertised networks; meaning that the routers at the core of the Internet would have to keep all these advertised networks in their routing table and then make appropriate routing decisions. For each packet that a router has to route, it has to perform a route lookup and if the routing table is large, the route lookup can cause significant delay.

Border Gateway Protocol or BGP is the routing protocol of the Internet. BGP is a path vector routing protocol making complex routing decision based on various negotiated policies. If there are multiple routes to a single network, BGP selects the best route and places that particular route on the routing table. Since the Internet is a huge mesh there could be many possible ways to reach a particular network. The BGP Routing Information Base or RIB, is a table containing all possible routes to all the networks in use. From this table, the best possible route is selected. As of July 2010, the number of RIB
entries in core routers of the Internet has reached more than 325000 [1]. The huge size of the RIB table is another significant problem that vendors are coping with. Processing such a large number of entries causes significant load on the routers and directly affects routing latency.

2. Problem Description

The unprecedented growth of the Internet has led to some fundamental routing problems. The routing table and the BGP RIB sizes have become enormous and continue to increase in size with each passing day. The global routing table size increases almost 16 percent annually compared to the RAM speed, which increases only 10 percent annually [7]. The routers at the core of the Internet make routing decisions after going through these large tables, which affects their efficiency. If a router reaches a point where it does not have the memory to store any more routes, it might simply crash or start behaving abnormally [2]. The presence of these large tables could also lead to new kind of Denial of Service or DoS attack.

The other issue with large routing and BGP tables is the time that it takes for these routers to converge. If a router crashes or there is a change in routing information for a large network segment, these routers have to process the new routing information, which takes significant amount of time even with enough memory and CPU cycles. Routers can only process a limited amount of information in a given interval of time. CPU cycles are allocated in a way that the new routing information does not overwhelm the router. This can cause a significant amount of down time. Vendors and ISPs have started to look beyond traditional routing architectures and many of them have started migrating to Multi Protocol Label Switching or MPLS which is a relatively new switching technology. In order to run a network with BGP, all core routers as well as the edge routers need to run BGP whereas when using BGP in conjunction with MPLS, only the edge routers need to run BGP, thus reducing routing load.

The vendors are trying hard to keep up with this growth but since the rate at which the Internet is growing is much higher than the rate at which hardware technology is progressing, they have started to realize that they could be eventually heading towards a dead end.

3. Hypothesis

Large routing tables affect the performance of routers and studying route lookup process on routers with large routing tables could enable us to generate a new exploitation technique to slow down the routing process on routers by forcing lookup of infrequently used routes.

This research will help in validating the hypothesis by thorough and in-depth study of routers with large routing information both in static and dynamic routing scenarios. Analysis of empirical data collected during the experiment using Cisco 2811 series routers will be used in this validation.

4. Related Work

This section presents a few related research ideas and results that are important to this study of current Internet routing tables. The Literature review has been broadly classified into two sections. The first section has previous research information about Internet routing instability and their causes. This section is important because it is directly related to this paper, which focuses on overloading routers to test their performance parameters. The second section deals with routers handling the route information and how they process routing tables. This section is important as this could have security implications.

The authors of papers [3] and [8] were among the first to research the various reasons of Internet routing instability. They pointed out that routing instability in the Internet could originate due to router configuration errors, link problems, lack of router resources, etc. Any one of these could cause routes to appear or disappear (route flaps), which could eventually lead to degradation of end to end performance of the Internet.

The authors of the paper [4] performed lab based analysis of a few commonly used routers from different vendors. The main goal of this research was to study the effects on a router when it is overloaded with excessive routing information, like large routing updates, large routing tables, etc. An incorrectly configured router could sometimes introduce a large amount of routing information into the Internet routing system. This could result other routers in the chain to crash or operate incorrectly and could eventually lead to cascading failures.

The authors tested a few routers from Cisco and Juniper Networks. They experimentally loaded the routers beyond the specified limit. Different routers running different firmware versions responded differently from each other, but the overall result proved that the routers would either stop responding to that particular peer which caused the problem, reset all BGP peers, or the particular interface would simply stop responding. Any of these would result in network disruptions and at some point would require manual intervention. The researchers were able to conclude that it is possible for routers to fail if they were inadequately equipped to handle large amount of routing information injected into the network.

This paper goes a step further in order to determine the effects an overloaded router has on routing traffic.

The authors of research paper [5] performed experiments in the AT&T backbone to understand the stability of BGP in the Internet. They wanted to establish a relationship between the popularity of prefixes and its BGP stability. Previous research had shown that BGP updates were very frequent and could cause traffic delivery problems. This research was the first one to understand the dynamics of the BGP updates. With this research, the authors were able to establish the fact that a few popular destinations were responsible for carrying majority of the traffic and had relatively stable routes. The majority of the BGP updates and instability are caused by unpopular destinations, which do not carry a lot of traffic.

The presentation [6] was a study to understand the Autonomous System or AS level dynamics of the internet traffic flow. This study also reinstated the fact that a small
percentage of the AS level topology is used to send and receive majority of the traffic on the Internet.

The author of the paper [9] primarily focuses on the importance of service guarantee in networks. The author then explores the various DoS attacks, which are difficult to fight and could easily disrupt genuine network traffic. With real time traffic in the picture, Quality of Service or QoS is very important and the author concludes with the fact the DoS can impact and reduce the QoS offered by networks.

Following the lines of the papers [5] and [6], my research would try and determine whether the popular destinations that carry the majority of the traffic are routed faster than the other traffic and also determine whether a router arranges its routing table lookup in the order of most used prefixes.

My research would also look at the security implications of the way routers handle the routing table. As mentioned by paper [9], DoS attacks can degrade network performance; my research would try to establish whether it is possible degrade network performance by generating legitimate network traffic (similar to some DoS attacks) directed towards not so frequently used prefixes of the routing table.

5. Methodology
This research is Quantitative in nature. The research results and conclusions are based on data collected from the lab-based setup and often compared to real world data.

This research was completed in two phases. The first phase was directed towards studying the effects of loading the routers with large amounts of routing information. The focus of the second phase of this experiment was to understand how routers processes large amounts of routing information and whether it is possible to exploit this behavior of routers. The second phase of this research also focuses on determining whether frequently used routes are processed faster than others.

Each of the phases in turn has two distinct case studies based on the differences in setup. The first case study deals with experiments using only static routes to load the routers. The second case study deals with experiments using BGP as a routing protocol to load the routers.

Since the routers used in the experiments had only 256Kbytes of non volatile configuration memory, the generated configurations file, which had a size starting at 3.78 Mbytes, had to be copied to the running memory which was 256 Mbytes. TFTP was used to copy the configuration file to the running configuration memory using the following command: 

`copy tftp running-config`

Once the basic network connectivity was complete as shown in Figure 1, the respective devices were configured with appropriate IP addresses and connectivity from PC1 to PC2 was verified. Router R1 and R3 had default routes pointing towards R2 to simplify route processing on them. This configuration is called the base configuration for the devices where all the experiments would start. At this point a Perl script was used to generate a configuration file for router R2, which at the first instance of the experiment contained 60000 static routes pointing towards router R3, as shown below.

```plaintext
ip route 193.196.189.0 255.255.255.0 FastEthernet0/0 192.168.1.10
ip route 193.196.190.0 255.255.255.0 FastEthernet0/0 192.168.1.10
ip route 193.196.191.0 255.255.255.0 FastEthernet0/0 192.168.1.10
ip route 193.196.192.0 255.255.255.0 FastEthernet0/0 192.168.1.10
ip route 193.196.193.0 255.255.255.0 FastEthernet0/0 192.168.1.10
```

Once the file is copied to the running memory, the routers take significant amount of time to process this configuration file and actually use this. In this experiment I loaded the routers with static routes starting from 60,000 through 40,000 in 20,000 route increments. The time to process the configuration file at 60,000 routes was about 50 minutes. This value increased as the number of routes increased and while loading the router with 400,000 routes, the router required a total of 18 and 1/2 hours to process the file, which was copied into the running memory. During the period that the router slowly churns through the configuration file, it is almost non responsive and even the console connection to the router becomes non interactive.

Once the router R2 has processed the entire configuration and reaches a stable state to pass traffic, router R3 is configured to have 5 loopback interfaces with IP addresses randomly selected from the static routes configured on router R2. These 5 IP addresses become the destinations for which data is collected. The computer, PC2, is also used to collect data. Once the data is collected, the router R2 is reset to base configuration and the previous steps are repeated with 20,000 additional static routes maxing out at 400,000 routes.
Once the basic network connectivity is complete as shown in Figure 2, the respective devices are configured with appropriate IP addresses and basic connectivity from PC1 to PC2 is verified. Routers R1 and R3 have default routes pointing toward R2 to simplify route processing on them. This configuration is called the base configuration for the devices where all the experiments would start.

At this point a Perl script is used to generate a configuration file for router R3, which at the first instance of the experiment contained 60000 static routes. Once the router R3 has processed the entire configuration and reaches a stable state, eBGP is enabled on R2 and R3. This populates the routing table on R2 with large amount of BGP routes. Once router R2 has processed the entire configuration and reaches a stable state, router R3 is configured to have 5 loopback interfaces with IP addresses randomly selected from the BGP routing table on router R2. These 5 IP addresses become the destinations for which data is collected. Computer PC2 is also used to collect data.

Once data is collected, routers R2 and R3 are reset to base configuration and the prior steps are repeated with 20,000 additional static routes. The upper limit of this experiment is 200,000 BGP routes due to lack of memory on the routers.

Once the basic network connectivity is complete as shown in Figure 3, the respective devices are configured with appropriate IP addresses and basic connectivity from PC1 to PC2 is verified. Router R3 has default routes pointing toward R2 to simplify route processing on it. This configuration is called the base configuration for the devices where all the experiments would start.

At this point the configuration file generated during Phase 1 containing 400,000 static routes is copied to router R2’s running configuration memory.

Once router R2 has processed the entire configuration and reaches a stable state, the complete routing table from router R2 is captured on a text file as shown below.

```
Gateway of last resort is not set

S      200.102.174.0/24 [1/0] via 192.168.1.10, FastEthernet0/0
S      200.85.157.0/24 [1/0] via 192.168.1.10, FastEthernet0/0
S      200.68.140.0/24 [1/0] via 192.168.1.10, FastEthernet0/0
S      200.51.251.0/24 [1/0] via 192.168.1.10, FastEthernet0/0
S      200.34.234.0/24 [1/0] via 192.168.1.10, FastEthernet0/0
```

With the entire routing table (400,000 routes) on a text file, I selected 50 IP addresses; every 8,000 (400000/50) routes from the routing table. These equally spaced IP addresses would form the test samples to determine the processing time of different routes positioned at different locations of the routing table.

Router R3 is configured to have 50 loopback interfaces with IP addresses previously selected from the routing table of R2. Data for all 50 IP addresses is collected from PC1, which directly connects to R2.
Once the basic network connectivity is complete as shown in Figure 4, the respective devices are configured with appropriate IP addresses, and basic connectivity from PC1 is verified. Router R3 has default routes pointing toward R2 to simplify route processing. This configuration is called the base configuration for the devices where all the experiments would start.

At this point, the configuration file generated during Phase 1 containing 200,000 static routes is copied to the router R3’s running configuration memory.

Once router R3 has processed the entire configuration and reaches a stable state, eBGP is enabled on R2 and R3. This populates the routing table on R2 with a large amount of BGP. This routing table is captured to a text file using the command shown below,

```
R2#sh ip route
Codes: C - connected, S - static, R - RIP, M - mobile, B - BGP
 D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
 N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
 E1 - OSPF external type 1, E2 - OSPF external type 2
 i - IS-IS, s - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
 ia - IS-IS inter area, * - candidate default, U - per-user static route
 o - ODR, P - periodic downloaded static route

Gateway of last resort is not set
B  193.187.122.0/24 [20/0] via 192.168.1.10, 1d00h
B  193.170.107.0/24 [20/0] via 192.168.1.10, 1d00h
B  193.153.88.0/24 [20/0] via 192.168.1.10, 1d00h
```

With the entire routing table (200000 routes) on a text file, I selected 50 IP addresses, every 4000 (200000/50) routes from routing table. These equally spaced IP addresses would form the test samples to determine the processing time of different routes positioned at different locations of the routing table. Router R3 is configured to have 50 loopback interfaces with IP addresses previously selected from the routing table of R2. Data for the 50 IP addresses is collected from PC1, which is directly connected to R2.

6. Results

Data collected during the Phase 1 of this experiment was router memory, packet processing time and jitter. Router memory was a measure to check the load on the router with large routing tables and the `show ip route summary` command was used on the router to get this information. Traceroute was used to collect packet processing times and jitter value was collected using the tool iperf [15]. The packet processing time and jitter were used to determine network performance.

![Figure 5. Phase 1 Case Study 1 Route count Vs Routing table memory](image)

Figure 5 is a clear representation of the increasing memory usage with the increase in routing table size. This parameter displayed behavior as expected at the start of the experiment.

![Figure 6. Phase 1 Case Study 1 Route count Vs Processing time](image)

Figure 6 depicts the increasing trend of packet processing time with the increase in routing table size. At the start of the experiment, I had expected a more linear increase in the packet processing times. As seen in the graph, though the overall trend of packet processing time increases with the increase in routing table size, there are a few unexplained spikes in the packet processing time.
Figure 7 compares jitter to the increasing routing table size. The graph maintains a steady level until the routing table size reaches 340,000. At this point the router runs out of memory to incorporate the entire routing table into the Forwarding Information Base (FIB) table. The FIB table is used by the Cisco Express Forwarding (CEF) [11] mechanism used by Cisco routers to speed up packet forwarding without having to perform traditional lookups. With the failure of CEF, the jitter value skyrockets from about 2ms to around 35ms. This in turn meant that packet delay variation had increased significantly now that the router had to go through the actual routing table before forwarding packets.

Figure 8 represents the increasing memory utilization with the increase in BGP routes. This increase is almost linear and indicates that the memory utilization of the router is directly proportional to the number of routes.

Figure 9 represents packet processing time against BGP route count. During the start of this experiment, I had expected the packet processing time to be fairly constant around the 1 – 2 ms second mark since this was the behavior shown during the static route testing until the 200,000 route mark. There were a few unexplained spikes in the packet processing time, but the overall trend is similar to the static route test at similar route counts. The packet processing time at 220,000 should have been theoretically infinite as the router was never able to process as many routes.

Figure 10 represents the jitter against the BGP router count. The jitter value can be seen to be fairly constant around 2 – 3 ms until about 180,000 – 200,000. Loading the router with 200,000 BGP routes is only possible by turning CEF (Cisco Express Forwarding) off. Similar to the static route case, once CEF is turned off the jitter value skyrockets to about 30 – 35ms.

Packet processing time was the only data collected during the phase 2 of this experiment. The data was collected at each of the 50 sample IP addresses.
7. Observations

The important observation from phase 1 of this experiment was the fact that loading routers with enough routing information affects its performance and effectively the performance of the network. Figures 6, 7, 9 and 10 clearly depict this behavior.

The most important observation in the phase 2 of this experiment is the positioning of the samples in the routing table. As seen in the figures 11 and 12, the high packet processing times are not located just at the end of the routing table but also at the beginning of the routing table. At the start of this experiment, my assumption was that the router would take more time to process packets which had destinations at the end of the routing table since it would have to go through more entries to reach there. The result of this experiment proved that the theory is wrong, and one possible reason could be that the router does not use the routing table in the sequence that is displayed when using the `show ip route` command. The other significant observation from the phase 2 of this experiment was the fact that the routers do not process frequently used subnets or destinations any faster than the ones which are not used as often. Successive lookups for the same destination in a quick succession of time yielded results which proved the hypothesis wrong.

8. Conclusions

The findings of this research conclude that large routing tables have a significant impact on the router performance. The other important conclusion is the uncertainty in predicting the behavior of routing table processing on a router.
Phase 1 of this research clearly indicated that increasing the routing table size on a router increases the packet processing times on the router. Post CEF failure, the jitter values also become fairly high indicating the performance degradation of the network. Based on this part of the research it can be concluded that it is very necessary to provision router correctly if we need good network performance. We should always keep in mind that Internet routing tables could be unstable at times and inject unusually large amounts of routing information to routers. In case the routers do not have enough memory and processor cycles to process these spikes, they would crash, causing large downtimes as these routers need a lot of time to stabilize and start processing packets after a crash or a reboot.

Phase 2 of the research indicated that predicting the route lookup behavior is difficult. The routers do not process routes at the top of the table any faster than the ones at the bottom. Routes at the bottom of the routing table have processing time similar to ones at the top of the table. The other important conclusion from this research was that routers do not process frequently used routes faster than the routes which are not. This behavior of routers made it almost impossible to formulate an exploitation script to make use of the route processing procedure.

9. Future Work

This research can be further extended by using different routers models as well as routers from different vendors like Juniper Networks etc. This research was performed on a router which was not processing packets other than the experimental packets. A good next step would be to study routers with significant traffic other than the experimental traffic to take the study even closer to actual Internet routers.

Traffic generators and advanced visualization tools would enhance this study and could be used when this study is taken to the next step.

10. Acknowledgments

I would like to thank Professors Bruce Hartpence and Daryl Johnson for their ideas that helped me complete this study.

11. References