Computer Networks Project 2
Analyzing BGP behavior during large-scale routing events

Release date: May 6, 2024
Due date: May 31, 2024, 12 PM

1 Introduction

In this project you will be analyzing the border gateway protocol (BGP)'s behavior during normal times and during large-scale (global) routing events.

As an example of a large-scale routing event, we consider the massive route leak initiated by Telekom Malaysia (AS4788) that caused significant network problems for the global routing system. Starting at 08:43 UTC June 12th, 2015, AS4788 Telekom Malaysia started to announce about 179,000 of prefixes to Level3 (AS3549, the Global crossing AS), which in turn accepted these and propagated them to their peers and customers. Since Telekom Malaysia had inserted itself between these thousands of prefixes and Level3 it was now responsible for delivering these packets to the intended destinations. This event resulted in significant packet loss and Internet slow down in all parts of the world. Read further about the event here.

Furthermore, the announcement of such large number of prefixes caused BGP routers to receive overwhelming number of update messages, and increased the convergence time of the global routing system.

In this project you will analyze the BGP update messages received by a set of BGP routers distributed across the globe using the update message archives collected by the RouteViews project. RouteViews has multiple *collectors* where each collector has multiple BGP sessions with BGP routers (i.e., *monitors*) in a variety of ISPs across the globe, and logs the update messages sent on each of these sessions.

2 Project

2.1 Your task

We provide you with three traces of update messages (i.e., `updates.20150611.0845-0945.txt`, `updates.20150612.0845-0945.txt`, and `updates.20150613.0845-0945.txt`) received by the same RouteViews collector at the same time of the day (8:45 AM - 9:45 AM) for three consecutive days (11th - 13th June, 2015). Your task is to analyze each of these files separately, and answer the following questions for each file separately. The answer for each file should be written in a `.yaml` file with the same name as the input file (i.e., `updates.20150611.0845-0945.yaml`, `updates.20150612.0845-0945.yaml`, and `updates.20150613.0845-0945.yaml`). Each line of each yaml file corresponds to one of the following questions, where the key should specify the question number and the value should be the answer (i.e., `<question number>: answer`). In answering the following questions, you need to ignore the effect of prefix aggregation. If the *final* answer to any question was not an integer, round it up to the smallest integer larger than the answer.

2.2 Input file format

Each input file we provide you with is the result of running `bgpdump` on a RouteViews archive file. Each line in the provided files specifies a single BGP update message received by the collector. Each line consists of multiple fields associated with different fields and metadata of a BGP update message. Different fields in each line are separated by `|`, each of which specifying the following fields in their order of appearance:

- BGP Protocol
- timestamp (in epoch format)
- W/A/B (withdrawal/announcement/routing table)
- Peer IP (address of the monitor)
- Peer ASN (ASN of the monitor)
- Prefix
- ASPath (as a list of AS numbers separated by space)
- Origin Protocol (typically always IGP)
- Next Hop
- LocalPref
- MED
- Community strings
- Atomic Aggregator
- Aggregator

Please note that some of these fields could have no values. In that case, if they are in the middle of a line, you would observe more than `|` characters in a row and if multiple fields at the end of a line have no value, they are just removed from the line.

Furthermore, please note that to complete this project you might not necessarily need all the aforementioned fields.

2.3 Questions

Questions you should answer regarding each input file are:

1. What is the number of all update messages?
2. What is the number of announcements?
3. What is the number of withdrawals?
4. What is the number of prefixes for which at least one BGP update is received?
5. What is the number of prefixes for which at least one announcement is received?

6. What is the number of prefixes for which at least one withdrawals is received?

We define an update burst as the sequence of more than one BGP update messages received for the same prefix, where the time interval between the timestamps of each two consecutive update messages in the sequence is shorter than 4 minutes. Each burst is an indicator of an event at the destination network (prefix originator). Given this definition, answer the remaning questions:

7. What is the total number of bursts?
8. What is the maximum number of bursts per prefix?
9. How long is the longest burst (in seconds)?
10. What is the average of the longest burst of all prefixes (in seconds)?
11. How many prefixes experience no bursts at all?
12. How many prefixes experience an average burst longer than 10 minutes?
13. How many prefixes experience an average burst longer than 20 minutes?
14. How many prefixes experience an average burst longer than 30 minutes?

3 Grading

Your grading is only based on the `.yaml` output file you provide us with, and the correct answer to each question for each input will receive 1 point, and any wrong answer will receive 0 points. Thus, correct answer to all questions would receive 42 points.

4 Submission

For each student we have created an empty repository at `https://gitlab.inf.ethz.ch/PRV-PERRIG/networks-course/project-routeviews/cn-2024-routeviews/<netid>-routeviews-project`. Please push the three requested `.yaml` files to your own repository before the deadline.

5 Acknowledgment

This project has been adapted from one of Princeton's computer networks course (COS-461)'s assignments.