LIP6 Project proposal:
Measuring event impact and propagation in the internet

General Information

Supervision

Timur Friedman, Olivier Fourmaux (NPA team, LIP6, CNRS/SU)
Lionel Tabourier (Complex Networks team, LIP6, CNRS/SU)
Kevin Vermeulen (LAAS, CNRS)

Location

LIP6, Sorbonne Université, 4 Place Jussieu, 75005 Paris.

Duration and remuneration

Up to 6 months internship - the remuneration of the internship is based on the standard rate of academic internship in France (4.05 euros/h, 35 h/week) with an additional compensation for food and transport.

Profile

This internship is directed at Master students (preferably Master 2 students) with a background in computer science. Good coding skills are requested for the internship, knowledge of a widely-used language in learning, such as python, is preferable but not mandatory. Background in computer networking, system building, and graph theory are at the heart of the internship, so a background in those areas is an asset, but not mandatory.

Contact

`lionel.tabourier@lip6.fr` , `kevinmylinhvermeulen@gmail.com`

Scientific description

Context

Understanding the impact of internet anomalous events at internet scale, such as performance degradation, outages, or attacks, is a challenging problem. If techniques and systems have been designed to detect outages at some particular internet facilities [7], or detect congestion between interdomain links [5], there exists no internet scale system to monitor events across all autonomous systems (ASes), and thus, we have no clear understanding on the impact of an event on the internet.

The BGP protocol allows ASes to interconnect, so that each AS can reach the prefixes containing the IP addresses of another AS via the routes received with BGP. As most internet events rarely last more than tens of minutes [2], to capture them, we need to run traceroutes towards each BGP prefix announced by all the ASes very frequently. And in addition to these background measurements, we need to be able to run even more targeted measurements during an event, in order to have a precise understanding of the behavior of the internet paths before and after this event. Unfortunately, public measurement systems, such as RIPE Atlas [9] and CAIDA Ark [1] do not offer such measurements or the possiblity to run them. They either perform meshed traceroutes between hundreds of sources and destinations at short intervals (15 minutes) [9], or perform traceroutes to one destination per BGP prefix from a hundred of vantage points every day [1]. This is neither sufficient to have an internet scale coverage nor to cover most internet events.

Methodology

We propose to design this missing measurement system, that will run background traceroutes at high speed [4, 12] every 15 minutes from a few vantage points to one destination in each BGP prefix announced by any AS. When an event is detected, we will run targeted measurements using propagation algorithms to understand how this event spreads on the internet. The next sections describe our ideas to detect the events and to monitor their propagation.

Event detection

To detect the events, we propose to use systems such as the internet health report [2], which monitors events on the internet, such as AS disconnections or paths latency inflation, using anomaly detection techniques in public traceroute measurements [6] and RIPE Atlas probes disconnections [10]. Examples of other data sources are BGP communities [7] or even spike in latencies using data from gaming streaming platforms [3]. Finding other data sources of internet events is also part of the project to improve coverage of our event detection.

Propagation models

Once an event has been detected, our system will select targeted traceroutes to understand how the event propagates. Namely, we would like to know which AS were impacted by an event, and how the network reacted. There is a variety of models describing propagation on graphs extracted from real-world data, including maps of the internet at the AS-level. They range from random walks [8] to cascade failures [11], typically used to model spreading phenomena. We would like to investigate to what extent these models can be fitted to actual observations and evaluate if the results obtained can be extrapolated to predict the consequence of future events. Note that internet routing differs significantly from other diffusion processes on graphs, as it involves routing choices which are not based strictly on the direct neighborhood of a node. That might lead us to design new propagation models adapted to this specific problem.

Expected outcome

This internship is part of a multiple years research project, and although the main lines of the methodology are well defined, a publication after the internship would be a bonus. We expect the intern to write a report and present a poster of some preliminary results that would serve as a basis for future work on the topic of a doctoral student. If the internship runs smoothly, the intern could be considered to apply for a PhD position.

References

[1] Ark. `https://www.caida.org/projects/ark`, last accessed on September 12, 2018.

[2] Internet Health Report website. `https://ihr.iijlab.net/ihr/en-us/`.

[3] Catalina Alvarez and Katerina Argyraki. Using gaming footage as a source of internet latency information. In *Proceedings of the 2023 ACM on Internet Measurement Conference*, pages 606–626, 2023.

[4] Robert Beverly. Yarrp'ing the Internet: Randomized high-speed active topology discovery. In *Proc. ACM IMC*, 2016.

[5] Amogh Dhamdhere, David D Clark, Alexander Gamero-Garrido, Matthew Luckie, Ricky KP Mok, Gautam Akiwate, Kabir Gogia, Vaibhav Bajpai, Alex C Snoeren, and Kc Claffy. Inferring persistent interdomain congestion. In *Proc. ACM SIGCOMM*, 2018.

[6] Romain Fontugne, Cristel Pelsser, Emile Aben, and Randy Bush. Pinpointing delay and forwarding anomalies using large-scale traceroute measurements. In *Proceedings of the 2017 Internet Measurement Conference*, pages 15–28, 2017.

[7] Vasileios Giotsas, Christoph Dietzel, Georgios Smaragdakis, Anja Feldmann, Arthur Berger, and Emile Aben. Detecting peering infrastructure outages in the wild. In *Proceedings of the conference of the ACM special interest group on data communication*, pages 446–459, 2017.

[8] Naoki Masuda, Mason A Porter, and Renaud Lambiotte. Random walks and diffusion on networks. *Physics reports*, 716:1–58, 2017.

[9] RIPE NCC. RIPE Atlas, 2019. `https://atlas.ripe.net/`.

[10] Anant Shah, Romain Fontugne, Emile Aben, Cristel Pelsser, and Randy Bush. Disco: Fast, good, and cheap outage detection. In *2017 Network Traffic Measurement and Analysis Conference (TMA)*, pages 1–9. IEEE, 2017.

[11] Lucas D Valdez, Louis Shekhtman, Cristian E La Rocca, Xin Zhang, Sergey V Buldyrev, Paul A Trunfio, Lidia A Braunstein, and Shlomo Havlin. Cascading failures in complex networks. *Journal of Complex Networks*, 8(2):cnaa013, 2020.

[12] Kevin Vermeulen, Justin P Rohrer, Robert Beverly, Olivier Fourmaux, and Timur Friedman. Diamond-miner: Comprehensive discovery of the internet's topology diamonds. In *Proc. USENIX NSDI*, 2020.