Discovering the optimization of routing paths in Overlay networks

Nicolas BLAYE, Jianfei PAN
Christophe LOHR, Isabel AMIGO, Sandrine VATON

Abstract

This paper reviews the project Overlay networks routing optimization. As an introduction, the paper introduces a prototype using RIPE Atlas for measuring overlay and using shortest paths algorithm for finding the optimal routing paths in overlay.

Keyword

Overlay, shortest paths algorithm, RIPE Atlas

Introduction

With the evolution of internet services, comes an ever greater need for storage and data processing. Companies like Facebook, Orange or Netflix have many data centers and servers scattered around the world to manage these data. In order to ensure the best possible quality of service, it is necessary that these hotspots are connected in the best way possible, however, because of the different strategies chosen by the routers on the Internet, routing is not always on an optimal level of quality of service (delay, bandwidth, availability ...). In the following, we will consider these hotspots as an overlay, ignoring the architecture of the roads between these points (which can be extremely complicated).

The objective of this project is to propose a solution for the measurement of an overlay network based on a metric and find the optimal routing between nodes (data centers, servers ...).

In this paper we will present our prototype for finding the optimal routing paths. The first part will introduce the main definition overlay with which we could do our hypothesis and abstraction. The next section presents the tool (RIPE Atlas) we will use to measure the overlay, presenting the problem of representativeness of the measure. Another part will introduce the prototype and the results.

Overlay networks

introduction-> what, why there are paths not optimals

An overlay network is a computer network that is built on top of another network. Nodes in the overlay network can be thought of as being connected by virtual or logical links, each of which corresponds to a path, perhaps through many physical links, in the underlying network.[1]

Overlay network can be considered as a full graph: every node is connected. The default routing path should be the direct connection between the two nodes. But as the different routing rules exist on the lower level of the network, it’s possible that the default routing path is not the optimal one. In the following, we consider the RTT(Round-trip delay time) as the principal metric for the measurement of the connections in overlay.

Project Atlas

RIPE Atlas is a global network of probes that measure Internet connectivity and reachability, providing an unprecedented understanding of the state of the Internet in real time.[2]

There are thousands of active probes in the RIPE Atlas network all over the world. RIPE Atlas users who host a probe can also use the entire RIPE Atlas network to conduct customised measurements that provide valuable data about their own network.

For our prototype, we use some probes as hotspots for our custom overlay, and we do ping measurement between one another to measure the complete graph.

The downsides of RIPEAtlas lies in the disponibility of those probes (they are managed by users), so we need to carefully select our probes by looking at their disponibility statistics, and the representativeness of the network: probes are installed by people that know more than the average person about Internet, and they tend to have a better connection than normal people. In our case, we went to simulate a datacenter overlay, so this is not a problem, but it would be if we were interested in end user.

The Model

Before the conception of the prototype, we did several hypothesis for the model:

- The connections are symmetric: the one-way trip time is the same in the two directions. This Hypothesis allows use us the RTT as the metric for connections and use an undirected graph.
- The overlay can be modelized as an undirected complete graph: every node is connected with the others. This hypothesis comes from the principal abstraction of the overlay.
- The influence of RIPE Atlas measurement on the RTT can be ignored.
- The RIPE Atlas measurement (ping) is allowed by the chosen probs.

With all these hypothesis, we have the model for the prototype: using shortest paths algorithm to find all optimal paths in an undirected complete graph in which the weight of connections between nodes is represented by the RTT measured with RIPE Atlas.

Prototype

Using RIPE Atlas probes, we built a prototype in python to simulate an overlay. Each node of the overlay is a probe, and each link is a ping measurement between two probes. Our prototype is divided into 3 parts:

- Measure : running the measurement using RIPE Atlas API (with Cousteau)

- Algo: calculating the shortest paths using Djikstra or Floyd Warshall
- Visu: generating an interactive graph in a GUI using Graph-tool

Difficulties and solutions

Algorithm choice

We consider two algorithms for the shortest paths in our project: Dijkstra algorithm and Floyd-Warshall algorithm. Dijkstra algorithm is designed for finding all the shortest paths from one point, however Floyd-Warshall calculate all the shortest paths between each pair of points.

In theory, they have the same complexity: $O(V^3)$ where V is the number of points in the graph. [reference: rapport biblio] So we did a benchmark to test the two algorithms and we found Floyd-Warshall is usually faster than the other one so we decided to use Floyd-Warshall as a default algorithm for our prototype, even though the calculating time of the algorithm is negligible compared to the time needed for the measurement.

Probe choice

For the probe choice, we looked at maps with the repartition of google data center around the world [3], and try to find a corresponding probe for each data center, with a map picturing the probe repartition [4].

This give us a first set of 35 probes to test our prototype. For some probes, we had to find a compromise between having a high disponibility and being close to the datacenter. For example in the US, the low amount of probes made the choice harder than in Europe. The detail of how we choose the probes can be found within the directory of the project.

Results

We did some measurement with our prototype, and we found that in every case, there is always more than one path that is shorter than the direct one. We even had a routing path that was passing through two intermediate nodes (the direct link was down). Sometimes, the shorter path is only a few milliseconds shorter, but we had some cases where it was more than divided by two (excluding the case where the direct link was down).

With our prototype, we can conclude that there is research to be done in that area to optimize the routing by using an overlay.

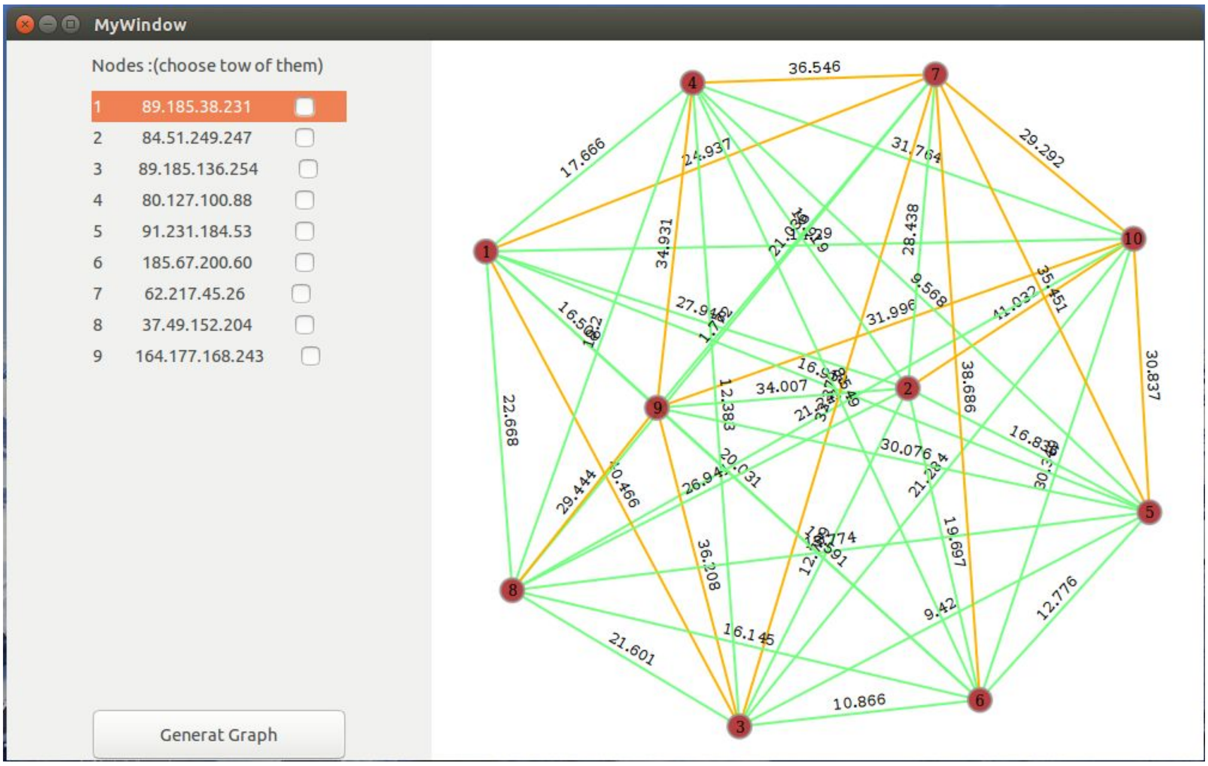

Diagram1 Interface of 9 nodes

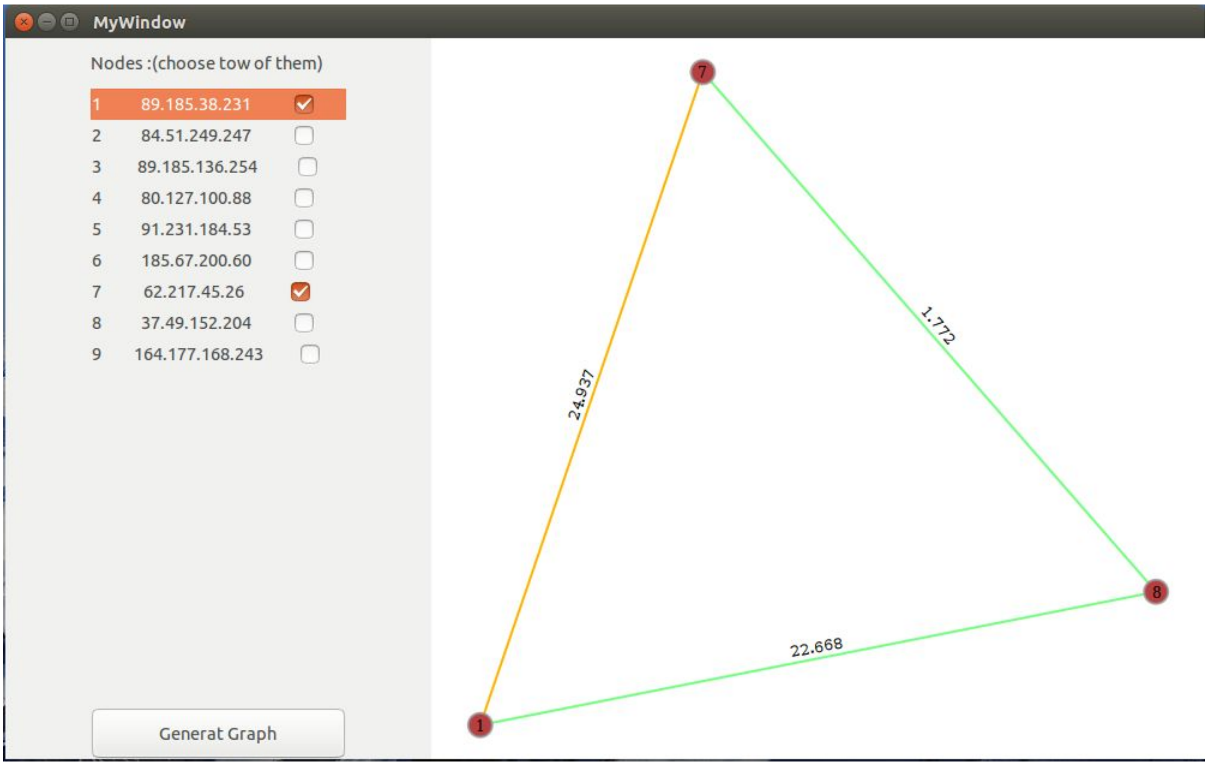

Diagram2 Routing path no-optimal between nodes number 1 and number 7

Conclusion

The results of our prototype shows that the default routing path may be not optimal. The situation simulated by our prototype exists in the reality and we think this method can be used to optimise the routing in overlays. For example, the big technical companies have many data centers all over the world, and our prototype can be used for measuring the overlay of theses data centers and help to find optimal routing paths. With these optimizations, the data can be transferred more efficiently.

Because of the limit of the time for our project, we have some perspectives for the project that we didn't have time to do :

Statistic Datas:

The situation of the overlays changes all the time. The results of measurement is not stable. So in this point of view, the results is not persuasive. However the prototype collects the results of every measurement.The idea is using these static datas to automate the measurement and the analysis of the routing situation of the Overlay and give a more stable solution by using IA algorithm for example.

Optimization of UI:

For the visualization part, we use a python library for generating the graph. For the moment, the graph is just a set of nodes and connections. An optimization is to draw the nodes in a map with which we can see clearly the geographical distribution of the nodes. With this optimization, the optimization of routing paths will be more obvious and we could easily find shorter routes that are not obvious. An obvious route is for example a route from Washington to Munich will often be quicker passing by London's node. In our measurement, we faced the case where a route from Zurich to Paris was twice faster by going through Munich.

References :

[1] https://en.wikipedia.org/wiki/Overlay_network

[2] https://atlas.ripe.net/

[3] http://nauges.typepad.com/.a/6a00d8345167aa69e200e551e51c228834-popup

[4] https://atlas.ripe.net/results/maps/network-coverage/?filter=