IJCSN - International Journal of Computer Science and Network, Volume 3, Issue 6, December 2014
ISSN (Online): 2277-5420 **www.IJCSN.org**
Impact Factor: 0.274

On The Origin of the Fast Dynamics of the Internet Topology

[1] Tounwendyam F. Ouédraogo, [2] Oumarou Sié , [3] Boureima Zerbo

[1] Ecole Normale Suérieur, Université de Koudougou
Koudougou, BP 376, Burkina Faso

[2] Department d'informatique, Université de Ouagadougou
Ouagadougou, BP , Burkina Faso

[3] Departement d'informatique Université de Ouagadougou
Ouagadougou BP, Burkina Faso

Abstract - Internet evolves over time and its large scale topology contributes to make challenging the dynamics measurement and characterization. A restrictive approach that consists to focusing on the dynamics observed by a single monitor has shown a fast dynamics of Internet topology. The topology evolves at a pace much higher than expected. We observe new IP addresses continuously and already seen IP addresses disappear with the same pace, over a long period (ten months) of measurement. This phenomenon is not localized but concerns the whole Internet topology. We show that the routing is at the origin of the fast dynamics and not a topology change.

Keywords - *Internet, Fast dynamics, Topology, Routing.*

1. Introduction

Most Internet mapping at IP-level topology consists to perform many end-to-end measurements from different hosts using TRACEROUTE-like methods [18, 17, 10, 12, 7]. The maps obtained in this way are incomplete and biased [2, 8, 1, 16, 20]. The challenge to obtain an accurate map of the Internet is related to its topology dynamics. The dynamics due to the load balancing has been identified as leading to a bias in the measurements by the authors of these works [2, 1]. Recent contribution introduces an approach to study the Internet dynamics [10]. Instead of trying to map the whole Internet, their approach consists to focus on what a single monitor sees of the Internet dynamics. Preliminary results of data analysis obtained from several monitors around the World, has shown a particular feature of the Internet dynamics: the topology evolves at higher pace and more sustained than expected [11, 10]. We made the similar observations on a most long measurement. We show that the new IP addresses existed before their first observation and the disappeared IP addresses still alive on the Internet which proves that the routing is at the origin of the fast dynamics. We also evaluate the impact of destinations (target in the end-to-end measurement) on the fast dynamics. We find different contributions of destinations to the fast dynamics. A tens of destinations with the largest contributions provide only 11% of the 3 2795 IP addresses. Therefore, we cannot attribute the fast dynamics to a handful of destinations. The rest of the paper is organized as follows. In section 2, we present a description of the dataset. Section 3 gives an empirical proof of the cause of the fast dynamics. Section 4 studies the influence of the destinations (end-to-end measurement parameter) on the fast dynamics. Section 5 surveys the related work. We conclude and discuss the implications of our contribution in Section 6.

2. Dataset

For our work, we use the data described in [6]. Several measurements were conducted from more than a hundred monitors scattered around the world. Each monitor has used a set of destinations that stayed the same during the period of the measurements. The measurement has consisted to run periodically the TRACETREE tool, one round measurement at every quarter of an hour. At each round measurement TRACETREE collects a routing tree from a given monitor to a set of destinations in a TRACEROUTE-like manner. The measurements were conducted from more than hundred monitors around the world for different periods of time depending on the monitor, from weeks to several months. For a more comprehensive description of the measurement framework and the obtained data, see [10]. In this paper, we have therefore chosen to illustrate our results mainly with measurement performed from 2011 December to 2012 September of a monitor located in France [6]. The destinations are 3 000 IP addresses chosen randomly that replied to ICMP echo request messages when they are been selected. These ten months of measurement correspond 25 311 rounds, about 96 rounds per days.

3. Origin of the Fast Dynamics

In previous works [11, 5, 14], the sustained renewal of IP addresses has been observed as a phenomena of the Internet dynamics. The authors of [11] have shown that this phenomena is not due to the volatif IP addresses neither to dynamic addresses that are dynamically allocated to different hosts over time and they have been ruled out of the probable causes. We present in this section what is the main cause of this fast dynamics observed at IP-level topology.

3.1 Appearance of New IP Addresses

The evolution of the number of IP addresses observed since the beginning of the measurements shows clearly a growth, see Figure 1. We can distinguish two main phases of growth: the first phase (shorter) corresponds to an exponential growth where there is a high renewal of IP addresses; the second phase is a linear growth with less renewal of IP addresses.

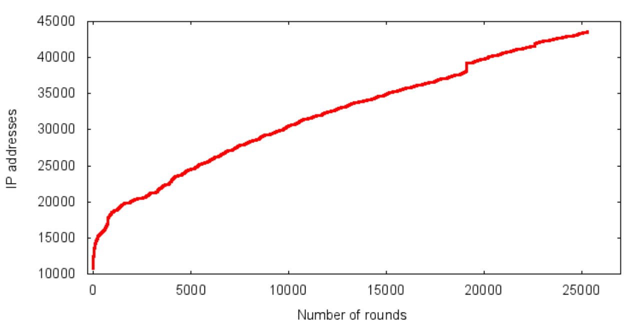

Figure 1: Evolution of the number of distinct IP addresses observed.

The pace of appearance of IP addresses is too fast to be considered as news addresses that did not exist before on the internet because most of these addresses are supposed to be servers or routers. To study this question, we consider data of ten others monitors that performed measurement in the same period. Among these ten monitors, we choose nine monitors as control monitors to observe the last one that is the main monitor. Considering the dates of observation of IP addresses, we can know what IP addresses observed by the main monitor have already been observed by the control monitors.

Figure 2 shows the evolution of IP addresses observed by the main monitor and those observed before by the control monitors. Most of the new IP addresses observed by the main monitor have been already observed. This means that these addresses existed before in the Internet, and they are just revealed by the routing dynamics. Even though the difference between the two curves becomes more and more important, at the end of the measurement around 80% of the new IP addresses that have already been observed by the control monitors.

It is important to consider the time elapsed between observation of the main monitor and observation of the control monitor of an IP address. A short time which means the address has been observed by the monitors at the same period increases the probability that the address to be really new in the Internet but a long time means the IP address is not new, but existed long time before. Among 22 008 IP addresses that already observed by the control monitors, 13 660 (about 62%) IP addresses are observed 7 days at least before the observation of the main monitor, and 3 790 IP addresses (about 17%) are observed by the main monitor in the same day (less than 24 hours later).

Notice that among the new IP addresses observed in the same moment by both controls and main monitors, some of these addresses could exist before the beginning of the measurement. Indeed, the measurements of control and main monitors began in the same time, so they may observe an address (that existed before or not) at the same moment.

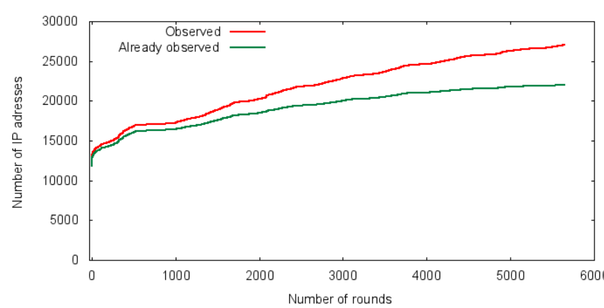

Figure 2: Number of distinct IP addresses observed by the main monitor and the control monitors.

In addition, the number of IP addresses already observed by the control monitors is underestimated. Figure 3 shows how the number of IP addresses already observed grows when the number of control monitors increases. It clearly appears that more control monitors allow identifying more IP addresses that existed before their first observation in the measurement.

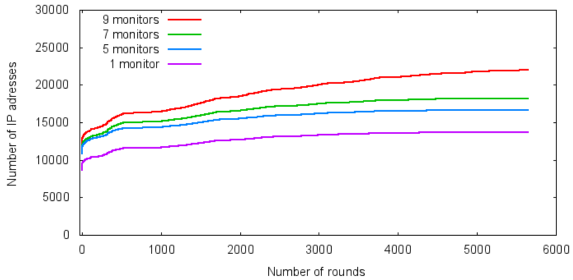

Figure 3: Number of IP addresses already observed depending of the number of control monitors.

In summary, with 80% of IP addresses and probably more that existed before on the Internet, we can conclude that the fast dynamics observed is mainly due to the routing dynamics that reveals existed IP addresses.

3.2 Disappeared IP Addresses

The number of IP addresses seen at each round measurement is stable and it seems not exceed a certain threshold[1], except some downward peaks[2], see Figure 4. This threshold is due to the tree structure that limits the number of IP addresses observable in a round measurement.

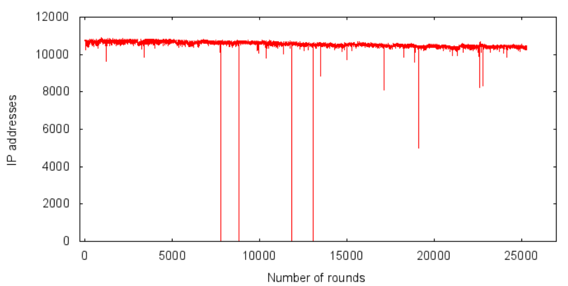

Figure 4: Number of IP addresses observed at each round measurement.

As the number of IP addresses seems to be the same at each round and that there is a renewal of addresses, then there are IP addresses already observed that disappear. Indeed, we have approximately the same number of addresses which appear and disappear what gives roughly an invariant number of addresses at each round. The curve of the evolution of the number of IP addresses which disappear looks like the inverted curve of the evolution of the number of new IP addresses, as shown in Figure 5.

[1] Notice that the value of this threshold may be different with others measurement parameters like the number of destinations

[2] These peaks could indicate a loss of connectivity at or near the monitor, or an event such as a major routing change or failure. Studying this is however out of the scope of this paper.

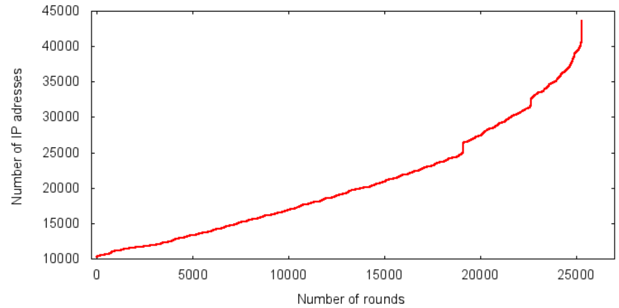

Figure 5: The number IP addresses that are observed before the round r and never after the round r as function of the number of round.

We conducted an experimental measurement that consists to know if the disappeared IP addresses of our measurement have also disappeared from the Internet.The measurement is done with the same parameters as those described in the section 2, except we use less destinations but this has not effect on the phenomena. We obtain 5 085 IP addresses in the six week-measurement. Each round measurement sees about 1 900 IP addresses. After the first round, 3 196 IP addresses are appeared and roughly the same number are disappeared, about 3 200 IP addresses. Simultaneously to the measurement, the IP addresses that disappear are regularly submitted to ICMP echo request twice a day, i.e. every 12 hours.

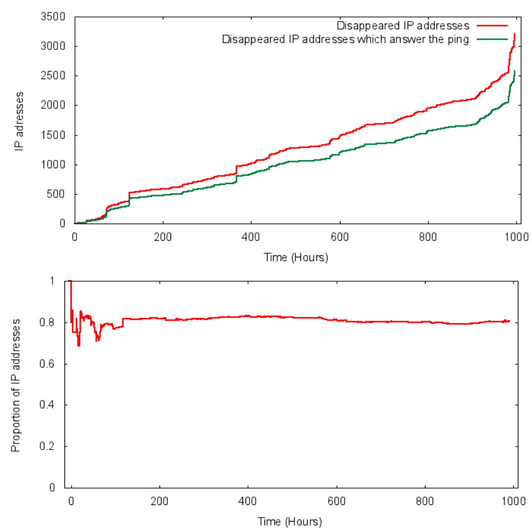

Figure 6: Top: Evolution of the number of IP addresses disappeared (solid line curve) and those among them that answered to the ICMP echo request (dashed curve). Bottom: Evolution of the percentage of IP addresses that still respond to the ICMP echo request among those that disappear.

Figure 6 shows that almost all the disappeared IP addresses continue to respond to the ICMP echo request. This rate is highest at the beginning of the measurement and stabilizes at about 80% until at the end of the measurement. The IP addresses that disappear from our measurement are just hidden by the routing dynamics because they still exist on the internet. The observed fast dynamics of the Internet is a routing dynamics, and not a topology change.

4. Influence of the Destinations

The method of measurement consists of periodically measuring the paths between monitor (source) and destinations (targets). The dynamics of path toward a destination will depend to its localization in the topology. However the destinations are chosen randomly among all IP addresses on the Internet. In this section we aim to study the role of the destinations in the observed dynamics, particularly if some destinations have significant influence in the observed fast dynamics. The common parts of routes toward destinations are merged to give a tree that is the outcome of the TRACETREE measurement. In order to obtain the new IP addresses observed by each destination, from the tree we must back to a set of separate paths of the destinations. We assume that a new IP addresses that appears in the path between the monitor (root) to a destination (leaf) belongs to this destination. If a new IP address appears in the common part of paths of several destinations, it will count once for each of them. Figure 7 shows an illustration of the separation of three routes, where two new addresses are observed in a round. The first IP address appears in the common part of the three routes and the second appears in the route of one destination. The new IP address in the common part is supposed to be seen by the three destinations taken separately. This assumption allows assigning realistically the new IP addresses observed to the destinations in given round measurement.

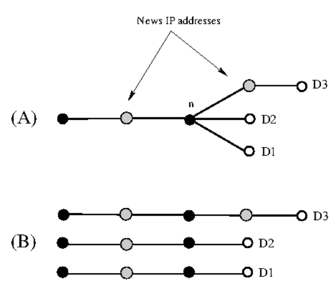

Figure 7: A: Tree of TRACETREE measurement. Destinations d1, d2 and d3 are leaves. The gray discs represent new IP addresses observed in the current round. B: Separation of the three paths from the tree.

The distribution of the number of IP addresses observed by each destination has a bell shaped with outliers which clearly disconnect on the right side, see Figure 8. These destinations have over than 400 IP addresses. Even these destinations get more IP addresses than others, we cannot attribute to them the renewal of IP addresses because they observe less than 3% of 3 2795 new IP addresses observed after the first round measurement. However the distribution shows that the destinations do not have the same contribution to the fast dynamics and this shows how an appropriate choice of the destinations may improve the dynamics measurement.

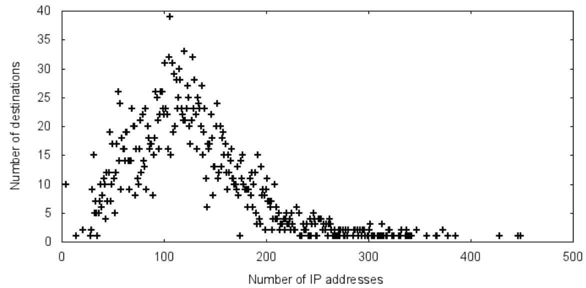

Figure 8: Distribution of the number of IP addresses observed by the destinations.

If we consider the new IP addresses that appear over time by destination, we observe that the renewal of addresses is less sustained. The destinations have a long period of time where they have not observed new IP addresses. These periods vary between two weeks to seven months depending on destination. This means that the most regular destination may spend two weeks without observing a new IP address and the less regular destination, until seven months without observing a new IP address. However in almost every round of measurement new IP addresses are observed, for instance in the last day of measurement 58 addresses have been observed.

The renewal of IP addresses is sustained by the majority of destinations. Then, the fast dynamics observed due to the routing is not due to handful of destinations which are located somewhere, meaning that the phenomena is general and characterizes the Internet dynamics at IP-level topology.

5. Related Work

Many research projects have been focused on the mapping challenge of the Internet topology. For instance some of them have concerned the methods of measurement that running probes from a limited number of hosts leads to miss some links and creates a bias on the observed degrees of the nodes, see [2, 8, 4, 19]. Others works have shown that the dynamics of Internet topology like load balancing makes measurement tools like TRACEROUTE less effective and leads to a view of internet topology with partial and false information, see [21, 2, 5, 15]. TRACETREE, Fast mapping and Interface Set Cover are recent methods of measuring the topology that try to take into account the evolution of the Internet [10, 5, 3]. The last paper uses a strategy similar to our observations on the destinations to improve the mapping. The destinations are selected dynamically depending to their impact on the mapping. The authors of work [9] studied the measurement process of different complex networks with the skitter data [18]. They observed a similar dynamics to us, Notice that the skitter data is collected at a lower frequency and for larger time scales than our data. A similar approach to ours is broached in this paper [13] but at AS-level of the topology and with a time-scale much longer than we do.

6. Conclusion

Our contribution has been to bring to light the cause of the fast dynamics of observed at ip-level topology. The phenomenon has been observed in a long measurement of several months. We have shown that the IP addresses which appear and disappear are both due to the routing dynamics. We have evaluated the impact of the destinations on the fast dynamics and shown that is not a dynamics of an area of the topology but concerns the whole Internet. From this paper some issues may lead to future work. Most of the Internet mapping methods is based upon the end-to-end measurement where destinations used as targets play a role important in the measurement. These methods can be improved by relevant choice of the destinations. Another direction for future work concerns the modelisation of the Internet dynamics. Our work is a contribution to the characterization of the dynamics which is essential to obtain more appropriate models.

References

[1] D. Achlioptas, A. Clauset, D. Kempe, and C. Moore. On the bias of traceroute sampling. In ACM STOC, 2005.
[2] D. Achlioptas, A. Clauset, D. kempe, and C. Moore. On the bias of traceroute sampling,or: Why almost every network looks like it has a power law. In ACM Symposium on Theory of Computing (STOC2005), 2005.
[3] R. Beverly, A. Berger, and G. Xie. Primitives for active internet topology mapping: Toward high-frequency characterization. In IMC, 2010.
[4] Q. Chen, H. Chang, R. Govindan, S. Jamin, S. Shenker, and W. Willinger. The origin of power laws in internet topologies revisited. In Proc. IEEE INFOCOM, Jun. 2002.
[5] I. Cunha, R. Teixeira, and C. Diot. Measuring and characterizing end-to-end route dynamics in the presence of load balancing. In PAM'11 Proceedings of the 12th international conference on Passive and active measurement, pages 235–244, 2011.
[6] Data. http://data.complexnetworks.fr/Radar/.
[7] B. Huffaker, D. Plummer, D. Moore, and k claffy. Topology discovery by active probing. In Proc. Symposium on Applications and the Internet, January 2002.
[8] A. Lakhina, J. Byers, M. Crovella, and P. Xie. Sampling biases in IP topology measurements. In IEEE INFOCOM, 2003.
[9] M. Latapy and C. Magnien. Complex network measurements: Estimating the relevance of observed properties. In IEEE infocom, 2008.
[10] M. Latapy, C. Magnien, and F. Ouédraogo. A radar for the internet. In Proc. first International Workshop on Analysis of Dynamic Networks (ADN), in conjunction with IEEE ICDM 2008, 2008.
[11] C. Magnien, F. Ouedraogo, G. Valadon, and M. Latapy. Fast dynamics in internet topology: preliminary observations and explanations. In ICIMP'09 Fourth International Conference on Internet Monitoring and Protection, 2009.
[12] Internet maps from mercator. http://www.isi.edu/div7/scan/mercator/maps.html.
[13] R. Oliveira, B. Zhang, and L. Zhang. Observing the evolution of internet AS topology. In ACM SIGCOMM, volume 37(4), 2007.
[14] J.-J. Pansiot. Local and dynamic analysis of internet multicast router topology. In Annales des telecommunications, vol. 62, pages 408–425, 2007.
[15] V. Paxson. End-to-end internet packet dynamics. IEEE/ACM Trans. Networking, 7(3):277–292, June 1999.
[16] P. De Los Rios. Exploration bias of complex networks. In Proceedings of the 7th conference on Statistical and Computational Physics Granada, 2002.
[17] Y. Shavitt and E. Shir. DIMES: Let the internet measure itself. ACM SIGCOMM Computer Communication Review, 35(5), 2005. See http://www.netdimes.org.
[18] Caida – skitter project. http://www.caida.org/tools/measurement/skitter/.
[19] N. Spring, R. Mahajan, and D. Wetherall. Measuring isp topologies with rocketfuel. In ACM SIGCOMM, 2002.
[20] F. Viger, B. Augustin, X. Cuvellier, B. Orgogozo, T. Friedman, M. Latapy, C. Magnien, and R. Teixeira. Detection and prevention in internet graphs. Computer Networks, 52:998–1018, 2008.
[21] Feng Wang, Nick Feamster, and Lixin Gao. Quantifying the effects of routing dynamics on end-to-end internet path failures. Technical report (TR-05-CSE-03) in Department of Electrical and Computer Engineering in the University of Massachusetts at Amherst, 2006.

Tounwendyam Frédéric OUEDRAOGO
-Master's degree in Mathematics and Computer at University of Ouagadougou Burkina Faso - PhD in Computer Science at the Complex Networks Research Team of LIP6 (UPMC) Paris- FRANCE. He is Assistant Professor at the University of Koudougou and Director Department of ICT promotion. Member of Research Teams LTIC (University of Ouagadougou) and Complex Networks of LIP6 (UPMC).

Oumarou SIE
PhD at the University of Rennes I France.
He is Leader of the Research Team LTIC Department d'informatique, Professor at the Department of Science and Technical of University of Ouagadougou . Advisor at the Regulatory Authority for Electronic Communications of Burkina-Faso.

Boureima Zerbo
PhD at the laboratoire des Sciences et Techniques Information Connaissance. He is Assistant Professor at the University of Ouagadougou. Advisor at the Regulatory Authority for Electronic Communications of Burkina-Faso.