Modeling as a necessary step for understanding Internet-wide route propagation

Steve Uhlig
TU Delft
suh@info.ucl.ac.be

Olaf Maennel
University of Adelaide
olaf.maennel@adelaide.edu.au

Wolfgang Muehlbauer
TU Munich
wolfgang@net.in.tum.de

Abstract— **Most of the insight we have today about the Internet-wide route propagation comes from studies performed more than 5 years ago [1], [2], [3]. Since then, insight about the routing system has largely focused on the behavior of individual ASes [4], [5], [6], [7] or particular prefixes [8]. The way BGP routes propagate across the Internet, how fast, and whether BGP does it reasonably well are largely open questions. To argue in favor or against potential alternatives to BGP, the research community must be able to provide answers to such questions. We argue that since measurements cannot reveal the whole picture of the Internet route propagation, building models of the Internet is one of the most reasonable ways to improve our understanding of this aspect of the interdomain routing system. [9] constitutes a first step in this direction.**

I. Measuring BGP routes propagation

So far, measurements have served the purpose of assessing the quality of the BGP route propagation across the Internet [1], [2], [3]. The time BGP routes take to propagate seems to be highly variable. Aspects likes CPU load, BGP timers and route flap damping [10], [11], [12] have been shown to contribute to this sometimes poor propagation time. Even though some prefixes suffer from slow convergence time, it is unclear what factors are most important to understand how a ”typical” prefix propagates across the Internet topology and how it would propagate if changes are brought to the Internet.

Measurements are inherently limited, as shown by the poor visibility of BGP events that trigger routing changes [13]. Relying on BGP informa tion, even gathered from a large number of vantage points scattered all over the Internet [14], [13], [9], provides an unclear picture of the BGP convergence.

Given the limitations of Internet-wide measurements of the routing system, recent works have focused on the behavior of single ASes. Teixeira et al. [5], [6], [15] studied the impact of hot-potato routing on the dynamics of the selection of BGP routes inside a large tier-1 AS, and measured its impact on the traffic matrix. Agrawal et al. [7] measured the impact of external BGP advertisements received by another tier-1 on its traffic. While Teixeira et al. concluded that the sensitivity of a single AS can be pretty important, Agrawal et al. concluded the opposite. This superficial contradiction was settled by putting together the context of the measurements and of the two studied ASes [16]. Concluding anything from the current studies about the behavior of ASes in the Internet is very tedious. Most studies of individual ASes focused on tier-1 ASes, whose routing behavior is carefully monitored and tuned. Other types of ASes might exhibit very different routing behaviors than tier-1 providers. Concluding anything about the evolution of interdomain routing from measurements only seems elusive.

II. Why do we need AS-topology models?

Since measurements alone cannot provide a clear picture of the BGP route propagation, we need to complement measurements with simulations. The purpose of simulations is to explore different possible ”Internets” that are not only consistent with what data tells us [9], but also to go beyond the information that lies within the BGP data. By trying to simulate different Internets, we hope to be able to find out what is the granularity at which the Internet must be modeled to correctly capture how routes might actually propagate.

We believe that there are two essential concepts that need to be defined to make a simulation model render the reality of the Internet: *routing domains* and *policies*.

A. Routing domains

Typically, subsets of routers inside an AS choose to reach most prefixes using the same exit point [5]. While a desirable and probably unattainable goal for modeling the routing in the Internet is to achieve the same topological granularity as the real one, for practical purposes this might not be necessary at all. If the goal of a model is to capture the choices that are visible from a set of vantage points, then accurately modeling the details that do not impact this visibility are useless.

We define the notion of a *routing domain* as a set of routers (in the real Internet) that choose the same path towards a set of destination prefixes. The choice of a path by a router inside an AS depends on the diversity available to the router and to the BGP decision process. Still, we might not need the full diversity present in the routing tables of all routers to have a pretty good idea of what the routing domains should be in a given AS. [9] showed that the number of routers required per AS to account for the information lying in the BGP data varies much between ASes. From the viewpoint of BGP data, most ASes need only to be made of a single BGP router [9]. On the other hand, a few ASes located in the core of the Internet need up to tens of routers. This hints at the difficulty of correctly modeling an AS since different ASes will require different topological granularities in a simulation model.

In [9], we implemented the notion of a routing domain through entities we called *quasi-routers*. A *quasi-router* in our simulations represents a consistent choice of the BGP routes by several real routers. There are two main issues when having to implement a quasi-router in a simulation model. First, the number of quasi-routers in a given AS depends on the available data, i.e. several quasi-routers can be claimed to exist only if we see in data that several distinct routes towards a prefix have been propagated. Since there can be many routes that are known to an AS but will never be propagated outside this AS [17], it is difficult to argue that a model based on BGP data may capture the actual the routing diversity available to some AS in the Internet.

B. Policies

Several works have tried to infer the likely policies between ASes [18], [14], [19]. Whether or not the inferred policies are correct and complete, those works have relied on a one router per AS granularity, i.e. two ASes could only have one peering relationship between each other. This is a big but reasonable simplification of reality. For building a faithful model of the Internet, sticking to simple peering relationships is not acceptable. Not only are peering relationships more complex than customer-provider or peer-peer, but multiple peerings between large ASes occur quite often. [9] showed that to account for the diversity available in the BGP data, it is necessary to go beyond both the atomic structure of ASes and the simplistic peering relationships between ASes.

III. Towards better models of the Internet

As explained above, the main difficulty in modeling the AS-level Internet is that the topological details of an AS and the definition of its policies are interdependent. The topological details of an AS will impact how coarse the peering relationships between ASes will be in a model (see previous section). For correctly predicting the AS paths chosen inside an AS [9], a fine enough granularity of peerings between ASes must be defined. Multiple routes need to be propagated across some ASes to propagate the diversity visible from a set of vantage points. We believe that this routing diversity can be used as a lever to find the right granularity for intra-AS topology and inter-AS policies inside a model of the Internet.

For an Internet model to be of practical interest, it must have enough routing diversity as to be close enough from reality. By definition, a model must go beyond observations, hence have more diversity than the observed one. The topological granularity of the model must hence be finer than the coarsest one that needs to explain BGP data. On the other hand, the additional information introduced by the model to capture more of reality than the data should be as consistent as possible with (unobserved) reality.

So far, only [20], [21], [9] tried to go beyond the simplistic one router per AS model of the Internet. Still, [20], [21], [9] suffer from important limitations like not taking into account the internals of an AS or relying on assumptions about peering relationships. The research community lacks experience in modeling the Internet routing system. We believe that before changing the interdomain routing system, the research community should first get experience into the most relevant aspects of its behavior. Building models that capture the propagation of the routes across the whole Internet is part of this effort. By playing with such models in what-if scenarios, we believe that much insight can be gained about what is right and wrong in the current interdomain routing system.

Acknowledgments

Part of this work was supported by the Belgian National Fund for Scientific Research (Steve Uhlig) and by the Australian Research Council (ARC) grant DP0557066 (Olaf Maennel).

References

[1] C. Labovitz, G. Malan, and F. Jahanian, “Internet Routing Instability,” *IEEE/ACM Transactions on Networking*, vol. 6, pp. 515–528, October 1998.
[2] C. Labovitz, A. Ahuja, and F. Jahanian, “Experimental Study of Internet Stability and Backbone Failures,” *In Proceedings of the Twenty-Ninth Annual International Symposium on Fault-Tolerant Computing*, June 1999.
[3] C. Labovitz, A. Ahuja, and A. Bose, “Delayed Internet Routing Convergence,” *In Proceedings of SIGCOMM*, pp. 175–177, August 2000.
[4] R. Teixeira and J. Rexford, “A measurement framework for pin-pointing routing changes,” in *Proc. ACM SIGCOMM Network Troubleshooting Workshop*, 2004.
[5] R. Teixeira, A. Shaikh, T. G. Griffin, and J. Rexford, “Dynamics of Hot-Potato Routing in IP Networks,” in *Proc. ACM SIGMETRICS*, 2004.
[6] R. Teixeira, A. Shaikh, T. G. Griffin, and G. M. Voelker, “Network sensitivity to hot-potato disruptions,” in *Proc. ACM SIGCOMM*, 2004.
[7] S. Agarwal, C.-N. Chuah, S. Bhattacharyya, and C. Diot, “The Impact of BGP Dynamics on Intra-Domain Traffic,” in *Proc. ACM SIGMETRICS*, 2004.
[8] Z. M. Mao, R. Bush, T. G. Griffin, and M. Roughan, “BGP Beacons,” in *Proc. ACM IMC*, 2003.
[9] W. Muehlbauer, A. Feldmann, O. Maennel, M. Roughan, and S. Uhlig, “Building an AS-topology model that captures route diversity,” in *Proc. ACM SIGCOMM*, 2006.
[10] S. Agarwal, C.-N. Chuah, S. Bhattacharyya, and C. Diot, “Impact of BGP dynamics on router CPU utilization,” in *Proc. Passive and Active Measurement Workshop (PAM)*, 2004.
[11] A. Feldmann, H. Kong, O. Maennel, and A. Tudor, “Measuring bgp pass-through times,” in *Proc. Passive and Active Measurement Workshop (PAM)*, 2004.
[12] Z. M. Mao, R. Govindan, G. Varghese, and R. Katz, “Route flap damping exacerbates Internet routing convergence,” in *Proc. ACM SIGCOMM*, 2002.
[13] A. Feldmann, O. Maennel, M. Mao, A. Berger, and B. Maggs, “Locating Internet Routing Instabilities,” in *Proc. ACM SIGCOMM*, 2004.
[14] L. Subramanian, S. Agarwal, J. Rexford, and R. H. Katz, “Characterizing the Internet hierarchy from multiple vantage points,” in *Proc. IEEE INFOCOM*, 2002.
[15] R. Teixeira, N. G. Duffield, J. Rexford, and M. Roughan, “Traffic matrix reloaded: Impact of routing changes,” in *Proc. Passive and Active Measurement Workshop (PAM)*, 2005.
[16] R. Teixeira, S. Agarwal, and J. Rexford, “Bgp routing changes: merging views from two isps,” *SIGCOMM Comput. Commun. Rev.*, vol. 35, no. 5, 2005.
[17] S. Uhlig and S. Tandel, “Quantifying the bgp routes diversity inside a tier-1 network,” in *Networking*, May 2006.
[18] L. Gao, “On Inferring Autonomous System Relationships in the Internet,” in *Proc. IEEE Global Internet*, 2000.
[19] G. Battista, M. Patrignani, and M. Pizzonia, “Computing the Types of the Relationships Between Autonomous Systems,” in *Proc. IEEE INFOCOM*, 2003.
[20] Z. M. Mao, J. Rexford, J. Wang, and R. Katz, “Towards an accurate as-level traceroute tool,” *A presentation to ACM SIGCOMM*, 2003.
[21] Z. M. Mao, L. Qiu, J. Wang, and Y. Zhang, “On AS Level Path Inference,” in *Proc. ACM SIGMETRICS*, 2005.