Towards Impactful Routing Research: Running Your Own (Emulated) AS on the (Real) Internet

Brandon Schlinker[1], Kyriakos Zarifis[1], Italo Cunha[2], Nick Feamster[3], Ethan Katz-Bassett[1], and Minlan Yu[1]

[1]Department of Computer Science, University of Southern California

[2]Department of Computer Science, Universidade Federal de Minas Gerais

[3]School of Computer Science, Georgia Institute of Technology

ABSTRACT

Interdomain routing has long been primed for an overhaul. Designed for a previous era and since dragged into today's networks through complex configurations, little progress has been made on aligning interdomain routing capabilities with network operator needs. Yet substantial growth has occurred in tangential areas such as datacenter networking. We believe that the barriers to performing interdomain routing research must be lowered to enable researchers and operators to better understand existing problems and evaluate potential improvements. To this end, we propose a testbed that enables the emulation of autonomous systems and their internal routing domains. The testbed allows these emulated networks to advertise routes and exchange traffic with the real Internet, exposing them to the Internet's complex policies and inconsistencies. Our testbed enables experiments to have complete control over the network topology and glue-logic between routing domains, including choice of IGP protocol, routing engine, and route redistribution. In addition, experiments can exchange routes and traffic with ISPs in multiple peering locations across the Internet, enabling experiments which emulate complex, geographically distributed networks.

Categories and Subject Descriptors

C.2.2 [**Computer-Communication Networks**]: Network Protocols—*routing protocols*

Keywords

Internet routing; interdomain routing; routing protocols; network architecture; testbeds

CoNEXT Student Workshop'13, December 9, 2013, Santa Barbara, CA, USA.
Copyright 2013 ACM 978-1-4503-2575-2/13/12 ...$15.00.
http://dx.doi.org/10.1145/2537148.2537153.

1. BACKGROUND

The Internet is composed of thousands of private networks, including end-user and tier-1 ISPs, enterprise networks, and content providers. The Border Gateway Protocol (BGP) exchanges logical route advertisements in parallel with the physical connectivity interconnecting private networks. These advertisements are combined with frequently complex routing policies to enable border routers to forward traffic towards its destination. An Interior Gateway Protocol (IGP) is the internal equivalent of BGP and responsible for directing traffic within an autonomous system (AS), channeling traffic amongst internal and border routers [3].

Under ideal circumstances, both of these protocols achieve the goal of exchanging traffic between networks. However, as the sophistication and scale of networks has increased, the stability of these two protocols has been undermined. BGP is susceptible to configuration errors due to the complexity of its configuration [4]. Route reflectors and route redistribution, both of which further increase complexity and may result in less than ideal routing, are often necessary as the size of a network grows [3, 1]. Explicitly defining the route traffic should take remains difficult and often proves impossible due to the complexity of the routing decision process [3]. In addition, there has been little focus on new and upcoming problems, such as the need to explicitly define routes to achieve quality-of-service or dynamically adjust routes given the state of a neighboring network.

We believe that the overhead required to accurately emulate the interdomain routing complexities of large autonomous systems has suppressed experiments in academia and industry, thus hampering innovation. Although simulation engines minimize the overhead of such experiments, they are no replacement for evaluation in a real world environment [5]. To this end, we propose a testbed that emulates complex AS topologies and enables the emulated ASes to advertise routes and exchange traffic with the Internet. Our testbed enables experiments to have complete control over the network topology and glue-logic between routing domains, including choice of IGP protocol, routing engine(s), and route redistribution. In addition, experiments can exchange routes and traffic with ISPs in multiple peering locations across the Internet, enabling experiments which emulate complex, geographically distributed networks.

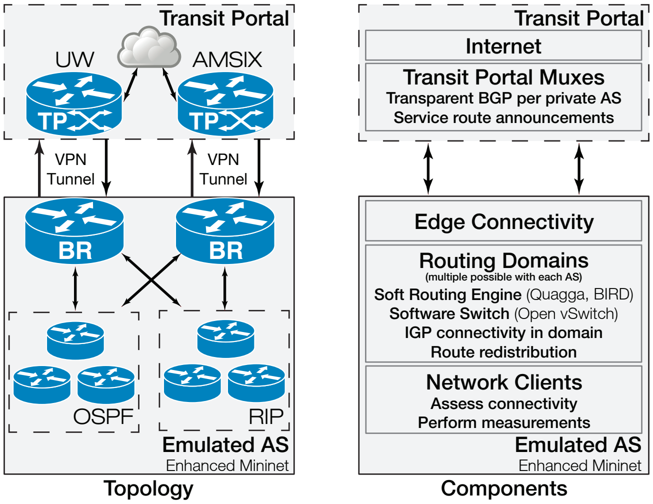

Figure 1: Example network topology hosted from a single enhanced Mininet instance alongside components required at each layer

2. TESTBED FUNDAMENTALS

2.1 Benefit to Community

The testbed provides operators and researchers with the ability to easily emulate complex autonomous system topologies while facilitating the exchange of routes and traffic with peers in Internet eXchange Points (IXPs) around the world. Thus, the testbed enables even a large ISP to emulate all aspects of their network infrastructure, including the complexities of their peering policies. For example, a large ISP could emulate their network infrastructure and experiment with different routing configurations and route advertisements. Information gained from the experiment could then be applied to the ISP's existing network. Overall, a researcher or operator faces significantly reduced overhead when setting up an experiment using the testbed and gains immediate access to often restricted AS-level Internet connectivity.

2.2 Enabling Interdomain Research

The testbed provides automated configuration of Transit Portal, linked with an enhanced version of Mininet. These two components are focused on very separate tasks.

Transit Portal is a service with BGP multiplexers spread throughout the world [6]. Each multiplexer provides isolated BGP connectivity for multiple experiments concurrently under a single global AS number. Designed primarily to provide BGP connectivity to downstream nodes, it lacks any intradomain capabilities and primarily enables the announcement and withdrawal of routes from a single endpoint. Mininet is an abstraction layer on top of Linux kernel namespaces which enables the rapid generation of virtual networks matching a user-specified topology [2]. Mininet was designed for SDN network research and thus lacks support for legacy IGP protocols, interdomain routing and interaction with the general Internet through AS-level connectivity. Our modifications to Mininet enable the emulation of virtual routing domains which are interconnected via traditional routing protocols (OSPF, BGP) and software routing engines (such as Quagga).

Acting independently, neither tool is capable of emulating complex interdomain routing scenarios, such as those of a large ISP. In contrast, our testbed will enable the emulation of such an environment, including the glue-logic which exists between different routing domains in a large network, exchange of route advertisements with upstream peers at multiple entry points of presence, and complex internal topologies, as shown within Figure 1.

3. CONDUCTING EXPERIMENTS

An experiment begins with the installation of the customized Mininet software and it's dependencies on a system provided by the experimenter. This varies from other testbed frameworks, such as Emulab, which colocate the entire experiment on a remote system to provide the requisite resources. Since Mininet uses Linux kernel namespaces, resource requirements are minimized and a network topology such as the one shown within Figure 1 can easily be executed on a standard laptop computer. Support for scaling topologies across multiple physical machines (when necessary) will be provided in future work.

Following installation, the characteristics of an autonomous system are defined using an extended version of Mininet's conventional topology syntax. For each routing domain, the routing engine and IGP protocol will be defined. Interconnections will also be defined amongst border routers within these routing domains and the upstream Transit Portal BGP muxes. In addition to supporting these new semantics, all traditional Mininet topology functionality is preserved, providing end-host and switch support along with the ability to emulate hybrid SDN/legacy networks.

When Mininet instantiates an instance of the topology, it will automatically create all required links, perform configuration of the routing engines, and connect to the Transit Portal service. Connections to the configured Transit Portal muxes will be established via VPN. The credentials for these VPN connections will be provided through a registration process and serve as input to Mininet during topology configuration. Experiments can then be conducted by connecting to individual containers within Mininet and running measurement applications and/or modifying network routes.

3.1 Future Experiments

We have held brief discussions with researchers interested in using this testbed, including the developers of the *Software Defined Internet Exchange* (SDX) at Georgia Tech. For SDX, potential applications of the testbed include the emulation of SDX peers and remote Internet exchange points.

4. CONCLUSIONS

We proposed a platform for use by operators and researchers wishing to examine the characteristics of interdomain routing schemes. Our platform leverages existing technologies to maximize the level of control available to the experimenter, including control over the network topology, routing engine(s), and routes advertised to the public Internet. We hope that the introduction of this platform will enable researchers to conduct experiments which bring a greater understanding of existing interdomain protocols and novel new techniques to improve interdomain routing.

5. REFERENCES

[1] N. Gvozdiev, B. Karp, and M. Handley. Loup: the principles and practice of intra-domain route dissemination. In *Proceedings of the 10th USENIX conference on Networked Systems Design and Implementation*, nsdi'13, pages 413–426, Berkeley, CA, USA, 2013. USENIX Association.

[2] B. Lantz, B. Heller, and N. McKeown. A network in a laptop: rapid prototyping for software-defined networks. In *Proceedings of the 9th ACM SIGCOMM Workshop on Hot Topics in Networks*, page 19. ACM, 2010.

[3] F. Le, G. G. Xie, D. Pei, J. Wang, and H. Zhang. Shedding light on the glue logic of the internet routing architecture. In *Proceedings of the ACM SIGCOMM 2008 conference on Data communication*, SIGCOMM '08, pages 39–50, New York, NY, USA, 2008. ACM.

[4] R. Mahajan, D. Wetherall, and T. Anderson. Understanding bgp misconfiguration. In *ACM SIGCOMM Computer Communication Review*, volume 32, pages 3–16. ACM, 2002.

[5] R. V. Oliveira, D. Pei, W. Willinger, B. Zhang, and L. Zhang. In search of the elusive ground truth: the internet's as-level connectivity structure. In *ACM SIGMETRICS Performance Evaluation Review*, volume 36, pages 217–228. ACM, 2008.

[6] V. Valancius, H. Kim, and N. Feamster. Transit portal: Bgp connectivity as a service. *ACM SIGCOMM Computer Communication Review*, 40(4):463–464, 2010.