Controlling Real Cloud Experiments from BGP to the Server (and Back)

Todd Arnold,[‡] Brandon Schlinker,[†] Ítalo Cunha,[§] Ethan Katz-Bassett[‡]
[‡]Columbia University, [†]University of Southern California, [§]Universidade Federal De Minas Gerais

ABSTRACT

One of the impediments in performing Internet routing research is the lack of infrastructure capable of supporting experiments with both *control* and *realism*. Measurement and experimentation platforms provide minimal control over routing, limited connectivity, and restrict what operations are available. Meanwhile, the proprietary nature of routing policies stymies simulations from providing realism.

To address these deficiencies, we present the Peering Testbed, which is capable of safely providing both realism and control. Peering has a dozen PoPs on three continents, each participating in BGP sessions on the Internet. The system's novel approach to support experiments allows each experiment to quickly and easily customize its interaction with the Internet. We will demonstrate new features of the testbed: federating with other testbeds to enable experiments that span the network edge, WAN, and data center network in a unified manner without the need for traversing the public Internet. This allows researchers to create their own global network, with data centers, similar to that of cloud providers, with control over how traffic is delivered to and from services on the *real* Internet.

CCS CONCEPTS

• **Networks** → **Network experimentation**; *Routing protocols*;

KEYWORDS

Internet Routing, Border Gateway Protocol, Traffic Engineering, Testbeds

1 INTRODUCTION

Efforts by cloud/content providers to rearchitect their networks for improved network performance is motivating renewed research in the Internet routing ecosystem [16, 22]. These providers are investing billions in infrastructure expansion [17], transforming the Internet's architecture from a hierarchical model where most networks have connectivity with transit providers and a handful of peers to a flattened model in which cloud/content providers build out private intercontinental backbones and interconnect widely [8, 16, 22].

This "flattening" of the Internet is spurring innovation in new designs for traffic engineering of huge volumes of content, which requires sophisticated controllers to make the best use of cloud providers' expansive connectivity [16, 22], especially at Internet eXchange Points (IXPs). However, researchers face barriers to work in this increasingly important research area, and the available tools cannot support both control and realism in parallel, forcing compromise.

Platforms such as RIPE Atlas and PlanetLab provide visibility into the Internet's state at a given point in time. However, they are unable to provide a method to control routing traffic. Simulations and emulations provide experimenters with complete control to build complex topologies and networks. However, the fidelity of simulated networks in replicating the Internet's behavior is restricted by limited visibility into networks' interconnectivity and routing, limiting realism.

Peering is a testbed capable of safely providing routing experiments with both realism and control, and has 12 Points of Presence (PoPs) spread across three continents, connecting to ~1000 networks. The testbed's new functionality includes a high-capacity backbone interconnecting PoPs (via provisioned VLANs across Internet2's backbone) and direct VLAN connectivity between PoPs and configurable data centers (via CloudLab [2], which allows configuration down to bare metal). Combined, these resources provide researchers control over interdomain PoPs, data centers, and an interconnecting WAN, qualitatively equivalent to a cloud provider.

2 PEERING BACKGROUND

Infrastructure Overview: Peering's established infrastructure consists of commodity servers running the Bird routing software, interconnected with other networks via BGP. These servers, or PoPs, are in 12 locations around the world, at both IXPs and academic institutions. The rich connectivity at IXPs provides Peering with nearly 1000 peer Autonomous Systems (ASes), including 18 transit providers for IPv4 and

SIGCOMM Posters and Demos '18, August 20–25, 2018, Budapest, Hungary

ACM ISBN 978-1-4503-5915-3/18/08…$15.00
https://doi.org/10.1145/3234200.3234247

six for IPv6, meaning we have at least 18 distinct routes to every Internet IPv4 prefix and at least six for IPv6.

Researchers execute experiments on systems they control, which enables the platform to support a variety of configurations. Researchers can locally run applications such as a web server, a Tor relay, or combine PEERING connectivity with emulated intradomain topologies. If more computational resources are required, a researcher can use CloudLab or commercial cloud providers.

Experiments use a VPN connection to authenticate with a PoP and then to establish a BGP session, so both data and control plane traffic passes over the VPN connection. Each experiment is allocated dedicated, globally routable address space to interact with the Internet. Since each experiment has access to a local router and BGP, they have full control over both ingress and egress traffic for their experiment.

Research Using Peering: PEERING is open to the community via a proposal process [3]. Since 2016, the authors granted over two dozen requests for experimental access to PEERING.

To date, experiments performed on PEERING were part of 16 publications, including 6 at SIGCOMM [4–7, 9–16, 18–21]. The majority of experiments performed on PEERING were conducted by researchers not affiliated with PEERING, and many are in domains outside of PEERING operators' expertise. In particular, PEERING supported experiments in the security domain such to demonstrate security flaws in applications such as Tor and Bitcoin and to evaluate RPKI adoption.

3 NEW FUNCTIONALITY

We expanded the functionalities of the PEERING testbed to support a new range of experiments. The functionalities are:

Backbone Connectivity: We worked with research and education networks, such as Internet2 [1], to establish circuits between several PEERING PoPs. These circuits provide an experiment connected to one interconnected PoP with visibility into all routes at every other interconnected PoP via the backbone. The added visibility and connectivity allows clients to direct announcements and traffic to external peers at any of the PoPs connected to the backbone. The dedicated circuits between PoPs enables experiments to mimic control of the backbone networks of cloud providers and to use connectivity at multiple PoPs.

PEERING PoPs at CloudLab Sites: CloudLab [2] provides researchers with access to bare-metal systems to conduct cloud computing and data center experiments. PEERING has PoPs with backbone connectivity at all CloudLab locations, providing sub-millisecond delay connectivity from CloudLab to PoPs. Interdomain and backbone interconnectivity is an important component of modern cloud for any user facing service. The federation with CloudLab provides that functionality to researchers.

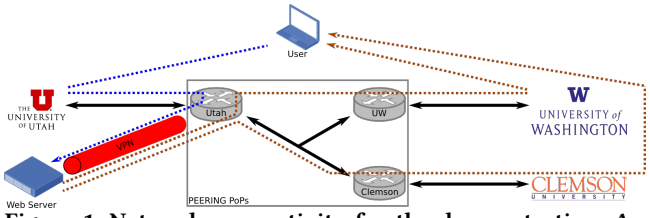

Figure 1: Network connectivity for the demonstration. An experiment is hosting a Web server at Utah CloudLab. Traffic for the server enters via the Utah PoP. The experiment can choose to return traffic via any PoP across the backbone. The addition of backbone connectivity and PoPs co-located with CloudLab resources provides experiments with a network qualitatively similar to that of cloud providers.

Types of Experiments: Using the new functionalities, either individually or combined, enables a broad range of experiments. Researchers can holistically examine the interactions between the various components of cloud provider networks (e.g. routing policies, traffic engineering, load balancing). Experiments have the ability to manipulate all aspects of how components interact with each other and external networks on the Internet, and can analyze performance, the tradeoffs between different implementations, or control system design.

4 DEMONSTRATION

Our demonstration illustrates the utility of the new backbone connectivity and federation with CloudLab, and is intended to raise awareness of this functionality and potential applications in others' research. Primarily, it shows how an experiment can use the backbone connectivity, fine-grain control of ingress and egress, and data centers to mimic a cloud provider and its traffic engineering capabilities.

We created a VM in the Utah CloudLab location which runs the PEERING client software, hosts a web service, and is labeled as "Web Server" in the lower left corner of Figure 1. The VM connects to the Utah PoP and announces its network via the Utah PoP's upstream. This directs all inbound traffic (the blue dashed line in Figure 1) to ingress at the Utah PoP. We demonstrate how a PEERING client can control the ingress location to another PoP, as well as how to dynamically configure the web server to send return traffic across the backbone for egress at either UW or Clemson (the orange dashed lines in Figure 1). Ingress and egress traffic can be at the same or different PoPs, and traffic to/from CloudLab and the Internet traverses a dedicated virtual WAN.

Our demonstration highlights functionality that was not previously available to researchers from existing cloud providers, or on PEERING or CloudLab individually, prior to this integration. By extending control to experiments run on the Internet, researcher are able to create and realistically assess traffic engineering control systems like those employed by cloud providers [16, 22].

REFERENCES

[1] Internet2. https://www.internet2.edu/.

[2] CloudLab. https://www.cloudlab.us.

[3] Peering. https://peering.usc.edu.

[4] Ruwaifa Anwar, Haseeb Niaz, David Choffnes, Ìtalo Cunha, Phillipa Gill, and Ethan Katz-Bassett. Investigating Interdomain Routing Policies in the Wild. In *IMC*, 2015.

[5] Maria Apostolaki, Aviv Zohar, and Laurent Vanbever. Hijacking Bitcoin: Routing Attacks on Cryptocurrencies. In *IEEE Symposium on Security and Privacy (SP)*, 2017.

[6] Henry Birge-Lee, Yixin Sun, Annie Edmundson, Jennifer Rexford, and Prateek Mittal. Using BGP to Acquire Bogus TLS Certificates. In *HotPETs*, 2017.

[7] Henry Birge-Lee, Yixin Sun, Annie Edmundson, Jennifer Rexford, and Prateek Mittal. Bamboozling Certificate Authorities with BGP. In *USENIX Security*, 2018.

[8] Yi-Ching Chiu, Brandon Schlinker, Abhishek Balaji Radhakrishnan, Ethan Katz-Bassett, and Ramesh Govindan. Are We One Hop Away from a Better Internet? In *IMC*, 2015.

[9] Arpit Gupta, Laurent Vanbever, Muhammad Shahbaz, Sean P. Donovan, Brandon Schlinker, Nick Feamster, Jennifer Rexford, Scott Shenker, Russ Clark, and Ethan Katz-Bassett. SDX: A Software Defined Internet Exchange. In *SIGCOMM*, 2014.

[10] Umar Javed, Ìtalo Cunha, David R. Choffnes, Ethan Katz-Bassett, Thomas E. Anderson, and Arvind Krishnamurthy. PoiRoot: Investigating the Root Cause of Interdomain Path Changes. In *SIGCOMM*, 2013.

[11] Ethan Katz-Bassett, David R Choffnes, Ítalo Cunha, Colin Scott, Thomas Anderson, and Arvind Krishnamurthy. Machiavellian Routing: Improving Internet Availability with BGP Poisoning. In *HotNets*, 2011.

[12] Ethan Katz-Bassett, Colin Scott, David R. Choffnes, Ítalo Cunha, Vytautas Valancius, Nick Feamster, Harsha V. Madhyastha, Thomas Anderson, and Arvind Krishnamurthy. LIFEGUARD: Practical Repair of Persistent Route Failures. In *SIGCOMM*, 2015.

[13] Simon Peter, Umar Javed, Qiao Zhang, Doug Woos, Arvind Krishnamurthy, and Thomas Anderson. One Tunnel is (Often) Enough. In *SIGCOMM*, 2014.

[14] Andreas Reuter, Randy Bush, Ìtalo Cunha, Ethan Katz-Bassett, Thomas C. Schmidt, and Matthias Wählisch. Towards a Rigorous Methodology for Measuring Adoption of RPKI Route Validation and Filtering. *SIGCOMM Comput. Commun. Rev.*, April 2018.

[15] Raja R Sambasivan, David Tran-Lam, Aditya Akella, and Peter Steenkiste. Bootstrapping Evolvability for Inter-Domain Routing with D-BGP. In *SIGCOMM*, 2017.

[16] Brandon Schlinker, Hyojeong Kim, Timothy Cui, Ethan Katz-Bassett, Harsha V Madhyastha, Ìtalo Cunha, James Quinn, Saif Hasan, Petr Lapukhov, and Hongyi Zeng. Engineering Egress with Edge Fabric. In *SIGCOMM*, 2017.

[17] Ben Treynor Sloss. Expanding our global infrastructure with new regions and subsea cables. https://blog.google/topics/google-cloud/expanding-our-global-infrastructure-new-regions-and-subsea-cables/.

[18] Peng Sun, Laurent Vanbever, and Jennifer Rexford. Scalable Programmable Inbound Traffic Engineering. In *SIGCOMM SOSR*, 2015.

[19] Yixin Sun, Anne Edmundson, Laurent Vanbever, Oscar Li, Jennifer Rexford, Mung Chiang, and Prateek Mittal. RAPTOR: Routing Attacks on Privacy in Tor. In *USENIX Security*, 2015.

[20] Yixin Sun, Anne Edmundson, Nick Feamster, Mung Chiang, and Prateek Mittal. Counter-RAPTOR: Safeguarding Tor Against Active Routing Attacks. In *IEEE Symposium on Security and Privacy (SP)*, 2017.

[21] Vytautas Valancius, Bharath Ravi, Nick Feamster, and Alex C. Snoeren. Quantifying the Benefits of Joint Content and Network Routing. In *SIGMETRICS*, 2013.

[22] Kok-Kiong Yap, Murtaza Motiwala, Jeremy Rahe, Steve Padgett, Matthew Holliman, Gary Baldus, Marcus Hines, Taeeun Kim, Ashok Narayanan, Ankur Jain, et al. Taking the Edge off with Espresso: Scale, Reliability and Programmability for Global Internet Peering. In *SIGCOMM*, 2017.