The 9th Workshop on Active Internet Measurements (AIMS-9) Report

kc claffy
UCSD/CAIDA
kc@caida.org

David Clark
MIT/CSAIL
ddc@csail.mit.edu

ABSTRACT

For almost a decade, CAIDA has hosted its Workshop on Active Internet Measurements (AIMS-9). This workshop series provides a forum for stakeholders in Internet active measurement projects to communicate their interests and concerns, and explore cooperative approaches to maximizing the collective benefit of deployed infrastructure and gathered measurements. On 1-3 March 2017, CAIDA hosted the ninth Workshop on Active Internet Measurements (AIMS-9). Materials related to the workshop are at http://www.caida.org/workshops/aims/1703/.

CCS Concepts

•**Networks** → **Network measurement; Public Internet;** *Network dynamics;*

Keywords

active Internet measurement, validation

1. INTRODUCTION

On 1-3 March 2017, CAIDA hosted the ninth Workshop on Active Internet Measurements (AIMS-9) as part of our series of Internet Statistics and Metrics Analysis (ISMA) workshops. This workshop series provides a forum for stakeholders in Internet active measurement projects to communicate their interests and concerns, and explore cooperative approaches to maximizing the collective benefit of deployed infrastructure and gathered measurements. Discussion topics this year included: existing and needed measurements to inform Internet policy; how to make measurement results more accessible to policy analysts and policymakers; existing and proposed active measurement platforms, architectures, methods, and tools; observation of path transparency (and lack thereof) and its policy implications; QoS and QoE measurement; integration of diverse measurement data to support innovative analysis; and classroom use of network measurement data. kc claffy (CAIDA) and David Clark from MIT's CSAIL and Internet Policy Research Institute co-hosted this year's workshop. We provide our informal thoughts on some of the more interesting insights that we took from the workshop. This report does not cover each topic discussed; materials presented at the workshop are linked from http://www.caida.org/workshops/aims/1703/.

ACM ISBN 978-1-4503-2138-9.
DOI: 10.1145/1235

2. MEASUREMENTS TO INFORM POLICY

While the policy community may continue to focus on network neutrality, it never became a central focus of Internet measurement research, for several reasons. Most importantly, the area is not amenable to rigorous scientific inquiry, since terminology is itself not rigorous, e.g., "reasonable network management." More practically, current discrimination is more likely to have an economic rather than a technical form, e.g., zero rating or discriminatory pricing of interconnection; evaluating these types of discrimination requires different sorts of data and analysis. The only discussion at this year's workshop relevant to network neutrality was a discussion of traffic throttling by mobile providers, reporting that several of the major U.S. providers have been slowing video streams. This behavior seemed to stop with the release of the FCC's 2015 Open Internet order, which classified Internet access as a telecommunications service covered by Title II of the Telecommunications Act. It will be interesting to see if this practice recurs since the current FCC suggests it will weaken or undo this action.

As a counterpoint to this discussion, Scott Jordan (UC Irvine) presented a list of the network measurements that are of interest to the FCC, including: more detailed information about the performance of access links (measuring both variation among users and variation over time for one user); behavior of links interconnecting ISPs; network management practices (such as throttling); Quality of Experience (QoE); usage; mobile performance metrics, etc. Most of these metrics are technically, financially, and politically challenging to effectively measure. Agencies such as the FCC, which would benefit from these sorts of measurements, have no demonstrated ability or appetite to fund the necessary work. Network operators, in turn, would have to provide ground truth for validation of measurement, modeling, and inference methods. Both sides contribute to the gulf between policy and technology.

We also spent a later session discussing the role of measurement in informing public policy. The Internet measurement community is highly technical, largely based in the

field of computer science, and tends to publish in technical conferences such as the Internet Measurement Conference (IMC), CoNEXT, SIGCOMM, NSDI, and security conferences. David Clark undertook an exercise to catalog all policy-relevant research published in the most recent Internet Measurement Conference. First he gave an overview of the literature represented there. He found a wide range of methods, often creative, and often opportunistic, meaning they are taking advantage of data that is readily available or easy to obtain. The distinction between active and passive measurement does not capture the diversity of work. Some of the more interesting papers combined different measurement methods and/or datasets. David also observed that the archive of historical data in the IMC papers (not always shared) was as important as the deployed infrastructure to support measurement.

Most of the published papers were deeply technical, providing information to network and service providers about how to better design their systems. But several studies were highly relevant to policy makers, most obviously network and application performance, and the disruption of end-to-end path security or integrity. Yet a wide gulf between the technical and policy communities continues, with the policy community largely unaware of Internet measurement research results or methods. Papers written for technical network research conference such as IMC do not often make results accessible to policy makers. There is a relevant policy conference – the Telecommunications Policy Research Conference (TPRC) – but aside from the two organizers of AIMS, perhaps only two other people attending this AIMS workshop had ever attended TPRC or were even aware that it existed. There was broad interest in understanding how to better bridge this gulf, but AIMS participants acknowledged the low incentive to invest their time in helping to do so: activities to bridge the gulf are hard to fund and do not generally contribute to academic advancement.

3. MEASUREMENTS OF OUTAGES

One current focus of the measurement research community is how to measure and report Internet *outages*, a term that still lacks a standardized definition. Different types of measurements reveal different sorts of events and impairments in the current Internet. One form of outage is where a region of the Internet becomes unreachable from one or a set of active measurement vantage points. Users inside cannot get out and users outside cannot get in. Some of these outages are due to natural disasters (e.g., hurricanes, earthquakes), others are due to operator error, and others are deliberate, such as the well-publicized disconnection of Egypt from the Internet during the Arab spring.

Researchers detect outages across the global Internet in several ways. One approach is to send active probe packets that comprehensively cover the IP address space, and look for changes in responses. We heard about different approaches to active probing for outages, with a discussion of the issues that arise, including inconsistent responses and irritation at a few destinations about the constant probing. A second approach is to monitor the routing protocols of the Internet (the Border Gateway Protocol, or BGP) and look for changes in routing information that indicate a region is inaccessible. Yet another method is to use unsolicited background traffic, observable via a network telescope (announced but mostly unassigned address space), which arises from pervasive sources of malicious or inadvertent traffic from all parts of the Internet; if a large enough telescope stops observing traffic from a specific region, that region has probably lost connectivity to the global Internet. Researchers have developed sophisticated analysis methods that allow correlation of these various signals of possible outages, in order to gain more confidence in inferences and a more accurate view of overall connectivity.

Another type of outage is a failure in an exchange point or co-location facility where lots of ISPs interconnect. These outages might be caused by a power failure or equipment failure. Because of the highly distributed pattern of interconnections in the Internet, a single failure at an exchange point may not disconnect regions of the network for more time than the routing protocols require to reconfigure the forwarding table to use new paths. However, users may see degraded performance due to loss of capacity or longer paths. Researchers at CAIDA, TU Berlin, and MIT have devised methods that can detect and localize these sorts of outages to a physical location within a city.

4. OBSERVING PATH TRANSPARENCY

Another focus for measurement is the extent to which encrypted connections are being compromised by a so-called "man in the middle" (MITM) attacks. One might think that with encryption, the end-points of a connection are assured of having protected communication. However, not so. As an old intelligence guy once said: "Amateurs think you have to break the crypto, professionals just steal the encryption keys." Or, in this case, just fool each end of the connection into thinking that the encryption key they have was shared with the other end, while in fact it was shared with the device along the path (Man-in-the-middle (MITM) attack). This vulnerability exists because of weaknesses in a part of the Internet called the Certificate Authority hierarchy, used to confirm that a web site is who it claims to be. If an actor can induce creation of a false certificate, he or she can completely invalidate the expected assurance provided by end-to-end encryption.

This exploit can take several forms, but the interesting measurement question is how prevalent false certificates are. Daniel Zappala (BYU) surveyed recent work in this area and reported that, depending on how the measurement is carried out, the number of flows subject to a MITM intervention varied between 1 in 1500 to as high as 1 in 10 or 25 in more. Most of these compromises are not malicious. Many companies that issue computers to their employees pre-configure them to allow key falsification so that the employer can monitor and safeguard employee communication. Other MITM interceptions were carried out by advanced firewalls to monitor content of communications, presumably in the interest of safety of the users. However, the researchers did find evidence of malicious MITM attacks – perhaps 1% of the observed incidents. One might expect this fraction to grow, given the pressures (both malicious and relatively benign) to intercept and observe Internet traffic. Both policy and technical factors influence the extent to which users should (or must) accept this practice.

Brian Trammel (ETH) presented related recent work on measurement of *path transparency*: determining the extent of impairment by accidental and purposeful manipulation at the transport layer. He described a tool for one-sided measurement of many targets from a single source, with simul-

taneous passive observation of generated packets to observe interference. The Path Transparency Observatory aims to provide comparability, reduction, and visibility of path data from different sources.

5. EVOLUTION OF TRAFFIC

Network operators, who have struggled to add capacity to deal with the flood of ever-higher quality video, often ask "what comes next?" Is there an application beyond video that will continue to drive the need for more bandwidth? Mike Wittie measured traffic in the emerging application area called *augmented reality*, a generalization of virtual reality in which imagery presented to the user integrates real world with virtual information. The technical requirements for augmented reality are highly variable, depending on the exact application. But real-time augmented reality, in which the user sees a virtual overlay on the physical world that changes with the head motion of the user, has an interesting latency requirement, which has a greater effect on user satisfaction than available bandwidth. If an application delivers imagery of the virtual overlay with latency less than 20 ms, then the latency must actually be less than 5ms, because between 20 ms and 5 ms, a user does not perceive the latency difference, but experiences motion sickness. This observation suggests that while Internet designers can strive to remove latency from the system, even the baseline latency of the speed of light will require application designers to modularize their systems so that some of the most latency-sensitive elements are co-located with the user. There are limits to what can originate in the cloud.

6. DOMAIN NAME SYSTEM

Another topic of interest was the Domain Name System. Mattijs Jonker (U. Twente) presented an update on the OpenINTEL proejct, which captures and archives results to comprehensive DNS querying of many TLDs, enabling a view into the evolution of the Internet over time. This data supports modeling behavior of the DNS under attack, and how services subject to DDoS attack try to evade the attacker. Neil Spring (U Maryland) presented on the extent to which the anycast addressing mechanisms used to diversify the DNS were actually working as expected to route queries to topologically nearby servers. John Heidemann (USC) presented on how to test a new variant of the DNS by feeding in a captured set of DNS queries.

7. MEASUREMENT FROM THE EDGE

Many measurement projects involve active probing of the network from the edge, reflected in the name of this workshop series (Active Network Measurement). These efforts include topology mapping, outage detection, access link performance evaluation, and scanning the address space for security vulnerabilities. This year's workshop included status reports on several such efforts, supporting software infrastructure (RADclock) to improve timestamp precision, as well as a new effort that crowd-sources active measurement to assess which networks properly support specific security best practices.

Alexander Marder (U Penn) addressed the challenges of mapping AS borders with a new algorithm that combines two previously separate methods (bdrmap and MAP-IT, two methodologies published separately at IMC2016) to improve AS annotations for routers, more accurately identify inter-AS links, and work with existing data at Internet scale.

Ricky Mok (CAIDA) described an early effort to design and build a platform to leverage crowd-sourced video upload and traceroute measurements triggered by observed congestion events on the network. His project seeks to include subjective QoE assessments and to measure the impact of network events on the client's performance and QoE.

Srikanth Sundaresan (Princeton) offered a new approach to detect whether congestion occurs on the access link or the wireless link (or somewhere else), by probing the TCP path using send TTL-limited packets within a TCP flow.

Eric Gaston (NPS) presented a new tool, *Yarrp6*, for IPv6 topology discovery. IPv6 topology discovery introduces new challenges given the enormity of the IPv6 address range and rate limiting on the network. Yarrp6 offers a new high-speed stateless traceroute technique that enables comparison of Transport Protocol (ICMPv6, UDP, TCP SYN, and TCP ACK Paris traceroute) on forward IP path inference.

Matthew Zekauskas (Internet2) gave an update on a new perfSONAR software module for scheduling, supervising and archiving measurements. This module includes a complete replacement for the Bandwidth Test Controller (BWCTL) component of perfSONAR. The new software has a simplified code base, RESTful API, and standardized, documented JSON data format.

Ramakrishna Padmanabhan (U. Maryland) presented his efforts to classify IP addresses (e.g., into University, Enterprise, and Cellular gateway) by analyzing traffic patterns from a large CDN's download manager logs.

All Internet measurement research projects struggle with the difficulty of sustaining operation of their research infrastructure. Deploying hundreds or thousands of measurement probes and supporting them in the field is unglamorous, costly work for which it is hard to get funding. This problem is compounded in cellular networks. We discussed one proposal – UCSD's PacketLab – for greatly reducing the functionality in a probe. Today, an experimenter downloads a program or script onto a probe that carries out an experiment. An alternative approach is that the experimenter's program runs on a server, and sends much lower-level instructions to the probe to send individual packets. By reducing the complexity of the task done on the probe, and standardizing the interface between the probe and the experimental server, it might be possible to use a rich variety of devices as probes, and share probes among multiple experiments. Practical implementation of such an approach would require resolution of many issues, including avoiding disruptive interactions among experiments and malicious use of probes, or more generally preventing experiments from disrupting the Internet. Despite the obstacles, this idea generated a great deal of interest among participants.

8. INTEGRATION OF DIVERSE DATA

It is perhaps a measure of the maturation and increasing sophistication of the field that some of the most interesting work is not reporting on what was learned from a single experiment, but what can be learned from integration of results from a range of experiments, where the data is potentially gathered by different teams. CAIDA's Internet Outage Detection and Analysis (IODA) project illustrates this sort of integration. Such work is complex, since data is not always in a consistent format, of consistent quality, or easily

retrieved. CAIDA reported on a new proposed effort – called Platform for Advanced Network Data Analysis, (PANDA) – which will integrate many of their diverse datasets from different measurement efforts into an integrated interface or *science gateway*. This interface is intended to allow the more sophisticated researcher to use these different datasets in a unified manner, and as well enable casual browsing of the data by a broader set of users, including (potentially) policymakers. Such platforms could also be used to make data from other research teams available.

9. MEASUREMENT DATA IN EDUCATION

Olivier Fourmaux (UMPC) described a promising new platform for the e-learning community as part of the EU-funded FORGE project. The platform will provide an educational layer of e-learning technologies, facilities and tools for the EU Future Internet Research and Experimentation (FIRE) initiative that includes 27 facilities offering wired, wireless, OpenFlow, Cloud, or other network resources. The layer helps automate resource discovery, selection, reservation, provision, experiment execution, control, monitoring results collection and resource release.

Tanja Zseby described her use of recent measurement data in two of her network security classes. The courses use CAIDA's Network Telescope Data to familiarize students with network data analysis methods, give students in-depth understanding of TCP/IP flow behavior, deepen network security knowledge, enable general scientific work skills, increase exploratory and forensic analysis skills, and awaken the scientist in each student.

ACKNOWLEDGMENTS. The workshop was funded by the Department of Homeland Security (DHS) Science and Technology Directorate, Cyber Security Division (DHS S&T/CSD) contracts HHSP 233201600012C and FA8750-12-2-0326 and the National Science Foundation award CNS-1513283. The work represents the position of the authors and not necessarily that of DHS or NSF.

10. WORKSHOP PARTICIPANTS

The main reason we continue this workshop is the enthusiastic participation it attracts from some of the brightest and most productive people in the community. We are grateful for their engagement and insights, many of which are reflected in this report.

- Robert Beverly (NPS)
- David Choffnes (Northeastern University)
- Taejoong Chung (Northeastern University)
- Danilo Cicalese (Telecom Paristech)
- David Clark (MIT)
- Ann Cox (DHS S&T Cyber Security Division)
- Casey Deccio (Brigham Young University)
- Ahmed Elmokashfi (Simula Research Laboratory)
- Olivier Fourmaux (UPMC Sorbonne UniversitÃl's)
- Eric Gaston (Naval Postgraduate School)
- Yossi Gilad (Boston University/MIT)
- John Heidemann (University of Southern California / Information Sciences Institute)
- Amir Herzberg (Bar Ilan University)
- Mattijs Jonker (University of Twente)
- Scott Jordan (UC Irvine)
- Erin Kenneally (DHS)
- Yonggon Kim (KAIST CSRC)
- Scott Kirkpatrick (Hebrew University)
- Robert Kisteleki (RIPE NCC)
- Kirill Levchenko (UC San Diego CSE)
- Alexander Marder (University of Pennsylvania)
- James Martin (Clemson University)
- Ramakrishna Padmanabhan (University of Maryland)
- Erik Rye (US Naval Academy)
- Aaron Schulman (UC San Diego CSE)
- Neil Spring (University of Maryland)
- Srikanth Sundaresan (Princeton)
- Brian Trammell (ETH Zurich)
- Darryl Veitch (University of Technology Sydney)
- Mike Wittie (Montana State University)
- Daniel Zappala (Brigham Young University)
- Matthew Zekauskas (Internet2)
- Tanja Zseby (TU Wien)
- kc claffy (UC San Diego/CAIDA)
- Alberto Dainotti (UC San Diego/CAIDA)
- Amogh Dhamdhere (UC San Diego/CAIDA)
- Marina Fomenkov (UC San Diego/CAIDA)
- Vasileios Giotsas (UC San Diego/CAIDA)
- Bradley Huffaker (UC San Diego/CAIDA)
- Young Hyun (UC San Diego/CAIDA)
- Alistair King (UC San Diego/CAIDA)
- Ricky Mok (UC San Diego/CAIDA)
- Joshua Polterock (UC San Diego/CAIDA)