

Summary of the Works-in-Progress Session at IMC 2016

David Choffnes
Northeastern University
choffnes@ccs.neu.edu

This article is an editorial note submitted to CCR. It has NOT been peer reviewed. The author takes full responsibility for this article's technical content. Comments can be posted through CCR Online.

Abstract

The Internet Measurement Conference provided a formal works-in-progress (WiP) session in November, 2016. This report describes the process for organizing and selecting participants, the format of the event, and summaries of the work presented.

CCS Concepts

•Networks → Network properties;

Keywords

Internet measurement, work in progress

1. INTRODUCTION

As conferences and even workshops become increasingly competitive and “high-stakes” venues for presenting early-stage or on-going work, there is a need for informal venues for researchers to present their unpolished research ideas or advertise their recently released research products such as datasets and tools. In November, 2016, the general chairs of the Internet Measurement Conference (IMC), Phillipa Gill and John Heidemann, invited me to run an event to fill this gap for Internet measurement research—a task which I happily accepted. The end result was a formal “Works-in-Progress” (WiP) session.

The WiP session was created as an opportunity for researchers to give a very short presentation at IMC about Works-in-Progress as well as tools or datasets that are available to the community. It was important to emphasize that the WiP is *not a publication venue*, but a forum for the exchange of ideas with peers from the community. In the interest of balancing the formality of a session with the informality of the work being presented, we opted for short “lightning” talks with time for just one or two questions afterward. The event ran for one hour after the end of the technical session on the first day of the conference.

The primary purpose of this editorial is to disseminate the tools and research that WiP participants opted to share publicly. To help guide those who want to organize a similar event in the future, the following section briefly describes how the event came to be and what worked well throughout the process of organizing and running the event.

2. LOGISTICS AND STATISTICS

We issued a call for WiPs on July 27th, 2016 and asked authors to submit titles and 200-word abstracts via HotCRP. While it may seem like overkill, HotCRP was useful for a variety of reasons.

First, there were 17 submissions. While it is a tiny fraction of submissions to the main conference, it was more than would be

easy to manage via e-mail alone. Second, given the first-come-first-serve policy for accepting WiPs into the program, it made it easy to tell who submitted when. Third, it allowed authors to upload their slides (and re-upload revised versions) as PDF “submissions” without the organizer needing to manage versions or files directly. This made the process of creating the master slide deck trivially easy on the day of the event. Fourth, it made mass communication with authors easy. Last, it facilitated gathering responses to questions about whether authors wanted to include abstracts or PDFs on the WiP website.

Out of the 17 submissions, only one was removed for being out of scope and one of the submissions had no presenter. One of the 15 remaining submissions was logically two separate projects, so in total there were 16 presentations for 60 minutes. Surprisingly, the session ended a few minutes early with time for everyone to present. The following steps helped make this happen.

Authors were allowed to use up to two slides and given up to four minutes to present their work and take questions. This was enforced via a tablet countdown clock ominously displayed to the presenters from the front row. The slides were gathered in advance and projected from a single laptop to avoid switching time. The presenters were also told the order of talks in advance and asked to come to the front of the room to take the stage immediately following the previous WiP. Thanks in large part to the presenters' preparation and compliance, the hour went off without a hitch.

The event was fairly well attended with at least 50 people in the room. Two factors that seemed to contribute to high attendance are that 1) the session occurred on the first day, when attendance at the conference was already at its peak, and 2) it was timed to fill the gap in the schedule between the technical session and the reception, thus providing a somewhat “captive” audience.

If there was one drawback, it was that the rapid-fire nature of the event did not give the audience sufficient time to chew on ideas to come up with questions, so there was limited Q&A during the session. However, multiple attendees indicated that the pace and short talk durations were perfectly appropriate given the nature of the WiP.

3. PRESENTATION SUMMARIES

Of the 16 presentations at the workshop, 12 submitted summaries for inclusion in this report. For a full list of presentations (and slides from those who published them), see <http://conferences.sigcomm.org/imc/2016/wip.html>. The summaries are listed in the order of presentation.

OpenINTEL: Large-Scale, Long-Term Active DNS Measurement Data for Network and Security Research

Roland van Rijswijk-Deij presented work on OpenINTEL, a large-scale active measurement platform for the DNS. It currently stores the results for a set of common DNS queries for all second-level domains in about 60% of the global DNS name space (covering most gTLDs, such as .com, .net, ..., and a number of ccTLDs). The data set already covers more than 21 months and is accessible to network measurement and security researchers on request, for more information please refer to the corresponding paper [1].

AMON: An Extensible Open Source framework for Online Monitoring, Statistical Analysis and Forensics of Multi-Gigabit Streams
Abhishek Balaji Radhakrishnan presented AMON (All-packet Monitor), an open-source tool that can process packets efficiently at multi-gigabit speeds. It leverages the high-performance packet monitor PF_RING and is readily deployable using off-the-shelf commodity hardware. AMON's data products (the so-called "databricks") can be employed for detection of high-impact events (such as denial-of-service attacks), real-time traffic visualizations, and identification of heavy-hitters. Some of the future directions include adding new modules for programming distributed instances of AMON, monitoring new applications such as DNS, and enabling privacy preserving data sharing techniques, etc.

A Continuing Study of the Active IPv6 Address Space

In this WiP talk, David Plonka gave an update on active IPv6 World-Wide Web (WWW) client activity, through the past 3 years. Using the methods that he and Arthur Berger introduced in their IMC 2015 paper, "Temporal and Spatial Classification of Active IPv6 Addresses," Dave showed the continuing growth and increased ephemeral nature of both IPv6 addresses and their subnet identifiers. In collaboration with Vasco Asturiano, who developed a first Hilbert curve visualization of IPv6 address space, he also showed compared Hilbert curve heat map and hierarchical Treemap visualizations of the active IPv6 address space, which comprises about 1 billion active WWW client /64 prefixes and 11 billion active WWW client addresses per month. (For more information on the IPv6 Hilbert curve visualization, visit <http://bl.ocks.org/vasturiano>.)

A Lightweight Looking Glass for Flexible DNS Diagnostics and Measurement

Casey Deccio presented work on DNS diagnosis and measurement. An understanding of Domain Name System (DNS) health through qualitative and quantitative measurement techniques can help inform deployment practices to maintain the robustness required for such a critical system. Carefully designed DNS queries and their responses (or associated network errors) form the basic data with which DNS health is assessed. However, because of the great network path diversity between DNS clients and servers, query/response visibility from diverse vantage points is necessary for greater fidelity of quantitative and qualitative measurement. In this presentation we discussed the challenges associated with measuring a large, distributed, network system such as the DNS, proposed a DNS looking glass solution for increased measurement fidelity, and demonstrated its functionality using open source tools.

A New MAP: Measurement and Analysis for Protocols Research Group

In this WiP talk, David Plonka introduced the new Internet Research Task Force (IRTF) Measurement and Analysis for Protocols Research Group (MAPRG) and welcomed IMC attendees to consider it as a venue to share their work. Along with Mirja Kühlewind, Dave chairs the MAPRG, which, although it is only recently chartered, has met previously, concurrent with IETF meetings. Thus, they expect to meet three times in 2017, giving In-

ternet measurement researchers the opportunity to present and discuss measurements and analyses of existing Internet protocols with standards practitioners and network operators. For more info on MAPRG participation, its mailing list, and video recordings of past presentations, visit <https://irtf.org/maprg>.

fling: A Middlebox Measurement Platform

Runa Barik presented work on *fling*. Middleboxes on Internet paths interfere with various protocols and protocol options. The transport layer in the network stack is ossified due to ubiquitous deployment of these boxes in the Internet. With the intention of detecting and measuring what these boxes do with arbitrary protocols and protocol options, we have developed a tool and a platform called *fling*. The *fling* web portal is currently hosted at <http://fling-frontend.nntb.no>.

PEERING: An AS for Us: Giving researchers control over real BGP routing

Ethan Katz-Bassett gave an update on the PEERING testbed, a BGP testbed that allows researchers to connect to real ISPs around the world and conduct experiments that exchange routes and traffic on the Internet. The testbed has been used by Katz-Bassett and by others in a number of research papers. Recently, it has been expanded with richer IXP connectivity and easier deployment of new sites and experiments.

Towards A Longitudinal Study of Adoption and Effectiveness of RPKI-Based Route Filtering

Ethan Katz-Bassett presented work he and collaborators are doing to measure adoption of the BGP security standard for Route Origin Validation (ROV), which can filter announcements with invalid origins. BGP security is only useful if adopted, yet only anecdotal evidence exists of the state of adoption. To fill this knowledge gap, they are making BGP announcements—some violating their ROAs—to conduct a long-running longitudinal study of adoption of route filtering, yielding a richer understanding of adoption and configurations than is possible via passive observation of existing (uncontrolled) announcements and potentially uncovering issues before they would otherwise manifest.

Mapping the Internet Backbone

M. Abdullah Canbaz presented work on mapping the Internet backbone. As the Internet lacks a common measurement mechanism, researchers develop ad-hoc mechanisms to uncover underlying connectivity. Several research groups have also developed public platforms where their extensive measurement data is shared with other researchers and network practitioner. In this work, we build on the data from such resources to extensively map the Internet backbone. To achieve such mapping, we first obtain IP addresses along with BGP announcements from different platforms, and then perform traces toward every observed IP address from each ingress of the destination AS. Finally, we perform unresponsive router, subnet and IP alias resolutions to reveal the underlying backbone of the router level Internet.

Crowdsourcing Physical Infrastructure Mapping with net.tagger

Robert Beverly presented work on *net.tagger*, a new project to crowd-source physical network infrastructure mapping. Using a mobile phone app (publicly available in the Google playstore), users "tag" by recording the location, picture, and meta-data of visible infrastructure markers such as manholes, dig warnings, pedestals, etc. This meta-data is copied to our server where it is integrated into a global infrastructure map and stored for later analysis. A significant future research opportunity and challenge is aggregating tags and deriving a higher-level understanding of the Infrastructure.

Browser based network troubleshooting using machine learning techniques

Max Bachl presented work on network troubleshooting. When an Internet problem occurs, it is hard or impossible for the average non tech-savvy user to determine the exact cause of the bad Internet connection. We are investigating novel pure web-based approaches to find the root causes of Internet problems using delay measurements and machine learning techniques. Specifically we use one binary classifier for each possible cause of bad Internet connections (such as bad wireless connection to the home router or cross traffic) and train them using delay measurements collected in lab experiments. When a user performs measurements using our web site the delay measurement results and we use our trained binary classifiers to tell the user if a problem is present and if there is a problem we can also name the specific problem.

A Look at the Optical Layer of Data Center Networks

Monia Ghobadi presented work on the optical layer of data center networks. Despite a ubiquitous presence of optical links in today's data center networks, little is known about even their basic characteristics in the wild. In this work, we conduct one of the first large-scale studies of the optical layer by monitoring 600K fiber optic links at Microsoft. Our analysis uncovers several findings that present opportunities to reduce cost and improve data center availability. First, we learn that optical links in data centers are vastly

over-engineered. Second, we find that number of packets dropped due to corruption dominates congestion. Based on these two observations we are proposing new schemes to improve the efficiency of data center networks. A paper using this dataset was recently accepted to NSDI [2].

4. CONCLUSION

In summary, the first formal WiP session at IMC was successful in terms of number of submissions, presentations, presentation quality, and audience attendance. The event ran smoothly in large part due to proper planning and organization, and it is my hope that this encourages and facilitates similar sessions in future at IMC and for other conference venues.

5. REFERENCES

- [1] R. van Rijswijk-Deij, M. Jonker, A. Sperotto, and A. Pras. A High-Performance, Scalable Infrastructure for Large-Scale Active DNS Measurements. *IEEE Journal on Selected Areas in Communications*, 34(6), 2016.
- [2] D. Zhuo, M. Ghobadi, R. Mahajan, A. Phanishayee, X. K. Zou, H. Guan, A. Krishnamurthy, and T. Anderson. RAIL: A Case for Redundant Arrays of Inexpensive Links in Data Center Networks. In *Proc. of NSDI*, 2017.