THE COMMONS INITIATIVE: COOPERATIVE MEASUREMENT AND MODELING OF OPEN NETWORKED SYSTEMS*

Sascha D. Meinrath† and kc claffy‡

I. INTRODUCTION

Over the past several years, interest in municipal wireless and community networking has increased dramatically.[1] Thus far, these initiatives have generally focused on networking local communities. The next evolution in networking involves peering these networks together.[2] Research on broadband service provision is desperately needed to help forge new national telecommunications policies and inspire innovation in networking technologies.[3]

With this goal in mind, the Cooperative Association for Internet Data Analysis ("CAIDA") held a workshop to discuss—and ultimately propose—collaboration among researchers and networks to simultaneously solve three

---

* The COMMONS Strategy Workshop brought together leaders from across North America to discuss issues related to broadband service provision and to strategize about how to interconnect existing networks to one another utilizing national fiber assets. An earlier version of this article was published on the CAIDA Web site. The authors wish to thank all the participants in the COMMONS project.

† Sascha D. Meinrath is Research Director of the Wireless Future Program at the New America Foundation.

‡ kc claffy is the principal investigator for the Cooperative Association for Internet Data Analysis ("CAIDA") and Resident Research Scientist based at the San Diego Supercomputer Center.

1 *See* MUNIWIRELESS.COM, STATE OF THE MARKET REPORT (2006) (on file with author).

2 Peering is "a relationship established between two or more ISPs (Internet Service Providers) for the purpose of exchanging traffic directly, rather than doing so through a backbone Internet provider." NEWTON'S TELECOM DICTIONARY 705–06 (23d ed. 2007).

3 Press Release, Rep. Edward J. Markey, Markey Addresses Consumer Federation of America on Broadband and Internet Policy (Feb. 1, 2007), http://markey.house.gov/index.php?option=com_content&task=view&id=2577&Itemid=141.

acute and growing problems facing the Internet. First, there exists a self-reported financial crisis in the Internet infrastructure provider industry that poses a threat to broadband growth and American competitiveness. Second, a data acquisition crisis has stunted the field of network science. Finally, emerging community, municipal, regional, and state networks need additional broadband connectivity but face limited provider, service level, and usage options.

The Cooperative Measurement and Modeling of Open Networked Systems ("COMMONS") Initiative proposes to build or partner with a collaborative national backbone to connect participating community, municipal, regional, and state networks to one another and to the global Internet. The COMMONS provides a platform for Internet researchers to study this infrastructure. It also provides a low-cost medium for networks to peer with one another. This approach will provide vital research results for policymakers across the country and around the world.

Part II of this article describes the findings from the initial COMMONS Strategy Workshop held in December 2006. Part III outlines relevant open research problems identified by the participants. Part IV proposes a framework for the end-to-end interconnection of networks at all levels on a national scale. Finally, the article concludes with a discussion of the steps necessary to bring about such a networking arrangement. Further, this section highlights the potential benefits to the scientific community, network operators and developers, key decision makers, and the general public.

II. THE COMMONS STRATEGY WORKSHOP

A. Purpose

On December 12–13, 2006, CAIDA hosted the first COMMONS Strategy Workshop at the San Diego Supercomputer Center on the Campus of the University of California at San Diego.[4] Workshop participants included a collection of relevant individuals and entities, such as community and municipal wireless builders, measurement experts, wireless technologists and researchers, policymakers, legal and privacy experts, and industry participants.[5] Workshop attendees discussed the design, creation, and operation of an experimental infrastructure that could simultaneously address three core crises: (1) the gap between those who benefit from digital technology and those who do not, commonly known as the digital divide; (2) the scientific integrity of network re-

[4] *See* CAIDA, COMMONS Workshop Agenda, http://www.caida.org/workshops/commons/0612/agenda.xml (last visited Apr. 4, 2008).

[5] *See* CAIDA, COMMONS Workshop Participants List, www.caida.org/workshops/commons/0612/list.xml (last visited Apr. 4, 2008).

search; and (3) the inability to empirically inform policy decisions at a critical juncture in telecommunications history.

The United States is facing a worsening broadband crisis. Over the past half-decade, the United States has fallen behind a growing list of industrialized nations in delivery speeds, price per megabit, broadband penetration rates, and other facets of broadband service provision.[6] Rural and poor communities face additional obstacles, often receiving little or no broadband access or being forced to pay higher service rates when they do have access.[7] Due to regulatory, political, and market constraints of incumbent local exchange carriers and other broadband providers, Internet researchers have been unable to study mission-critical aspects of the Internet[8] and the state of its current robustness, capacity, usage, and problem areas. Therefore, potential solutions to these issues continue to be conjecture, rather than empirically-backed analysis.[9] Meanwhile, telecommunications regulators and policymakers have increasingly called for a methodologically sound study of Internet usage, analysis of potential failure points, and improvements to this vital infrastructure.[10]

The COMMONS provides an opportunity to address these shortcomings within a single national framework. By creating a national peering infrastructure that interested network operators may choose to join, the COMMONS will provide numerous opportunities and benefits for a range of different constituencies including: low-cost transport; the ability to buy bandwidth in bulk and share the cost-savings among the COMMONS partner organizations; and empirically grounded analyses of traffic flow, congestion points, and underutilized links.

---

[6] S. DEREK TURNER, FREE PRESS, BROADBAND REALITY CHECK II: THE TRUTH BEHIND AMERICA'S DIGITAL DECLINE 3–4 (2006), *available at* http://www.freepress.net/files/broadband_report.pdf.

[7] U.S. GOV'T ACCOUNTABILITY OFFICE, TELECOMMUNICATIONS, BROADBAND DEPLOYMENT IS EXTENSIVE THROUGHOUT THE UNITED STATES, BUT IT IS DIFFICULT TO ASSESS THE EXTENT OF DEPLOYMENT GAPS IN RURAL AREAS 4 (2006), http://gao.gov/new.items/d06426.pdf.

[8] COMM. ON NETWORK SCIENCE FOR FUTURE ARMY APPLICATIONS, NAT'L RESEARCH COUNCIL, NETWORK SCIENCE 7 (2005).

[9] *See* NAT'L SCI. & TECH. COUNCIL, FEDERAL PLAN FOR CYBER SECURITY AND INFORMATION ASSURANCE RESEARCH AND DEVELOPMENT (2006), *available at* http://www.nitrd.gov/pubs/csia/csia_federal_plan.pdf (exemplifying the interagency framework that is needed to provide an adequate level of cyber security and information assurance capabilities and technologies).

[10] *See* Bob Wallace, *NTCA: FCC Commissioner Calls for Sweeping Regulatory Reform: Candor from Copps Counts with Rural Telco Crowd*, TELECOMM. ONLINE, Feb. 7, 2006, http://www.telecommagazine.com/newsglobe/article.asp?HH_ID=AR_1706 (explaining that the United States lacks a national broadband strategy that effectively reaches rural America); *see also* Press Release, Rep. Edward J. Markey, *supra* note 3 (noting that the United States has low broadband subscribership within the global market).

Peering will be conditionally available to local, state, and federal government entities, academic institutions, community Internet initiatives, and commercial entities based upon three conditions. First, networks will make select operational data available to the COMMONS researchers under appropriate legal data sharing and privacy guards. Second, the attached networks must agree to develop and abide by the COMMONS's policies, which will be based upon research results of empirical data analyses of network usage. Third, participating networks must abide by the Acceptable Use Policies[11] created by the COMMONS Coordination Committee.[12] The COMMONS will impact both the Internet industry and policymaking broadly, by providing substantial real-world data on Internet traffic at the national level and informing analyses, regulatory discussion, and technological innovation. It also promises to raise the intellectual merit of the entire field of Internet science through increasing standards of data collection and sharing within the research community.

The telecommunications sector has not yet recovered from the privatization and commercialization of Internet infrastructure, which began in the early 1990s. After a decade of boom and bust, consolidation continues, though the number of Internet service providers left to consolidate has greatly diminished.[13] Furthermore, the largest of the remaining providers have publicly insisted that they will not be able to make the required investment to build-out broadband infrastructure without more flexible pricing strategies to recover costs.[14]

---

[11] CAIDA uses the following Acceptable Use Policies to govern its data collection activities: (1) passive monitors will run only strictly necessary services and will be kept up-to-date with necessary security patches and operating system upgrades to limit security risk; (2) only a minimal number of CAIDA personnel trained in protecting user privacy and secure handling of data will have accounts on passive data monitors; (3) no packet payloads will be permanently recorded without specific permission from the hosting site; (4) traces will not be released from CAIDA custody unless the IP addresses are anonymized using prefix-preserving anonymization (or other current state-of-the-art anonymization technology). CAIDA personnel and collaborators who are physically present in CAIDA offices may have access to non-anonymized packet headers for research purposes; (5) CAIDA will require registration from users who wish to download anonymized traces; and (6) traces will be distributed internationally to registered users, although we are bound by the Department of State's International Traffic in Arms Regulations. *See* COMMONS Site Acceptable Use Agreement, http://www.caida.org/projects/commons/aup/ (last visited Apr. 19, 2008).

[12] The COMMONS Coordinating Committee will consist of representatives from active partners on the COMMONS Project and will help oversee the COMMONS Project, provide feedback on the COMMONS documents, and help with outreach.

[13] *See, e.g.*, Tom Spring, *ISPs Share Urge to Merge*, CNN, Mar. 2, 1999, http://www.cnn.com/TECH/computing/9903/02/ispmerge.ldg/index.html (describing the rapid consolidation of Internet Service Providers during the late 1990s).

[14] Marguerite Reardon, *AT&T Chief, FCC Chair Clarify on Net Neutrality*, CNET-NEWS, Mar. 21, 2006, http://www.news.com/AT38T-chief%2C-FCC-chair-clarify-on-New-neutrality/2100-1034_3-6052239.html.

By removing from commercial providers the responsibility for supporting Internet service delivery to unprofitable areas, the COMMONS will measurably alleviate the economic pressure on these providers. Additionally, the COMMONS offers an unprecedented opportunity to establish standards of scientific integrity in the field of Internet research by providing rigorous empirical data against which to validate theory, modeling, and simulation activities. Furthermore, because the COMMONS will support public analysis of actual Internet traffic, it will inform debates on increasingly important technical, economic, policy, privacy, and social issues relating to the Internet. Finally, the COMMONS not only allows struggling community networks to cost-share a financially daunting component of their connectivity, but it also provides a forum for the cooperating networks and the research community to share lessons with one another.

B. Background

The first general purpose Internet backbone,[15] NSFNET, was funded and administered by the National Science Foundation to support the networking needs of the research and education community.[16] It was implemented to provide continuously collected data on the function and usage of the network.[17] When this backbone was decommissioned in the mid-1990s as a part of the government's strategy to privatize the network infrastructure, the attached academic networks transitioned to self-sustaining funding models.[18] Unfortunately, since that time, access to representative data on Internet traffic, topology, routing, and security has diminished as Internet researchers have struggled to conduct legitimate, reproducible scientific experiments under increasingly restrictive constraints. Furthermore, many of the measurements and lessons learned by commercial Internet providers over the last decade were not retained, or if retained, not shared for fear of providing an advantage to marketplace competitors.

Obstacles to the collection and analysis of traffic data on the commercial Internet pose not only formidable technical and engineering challenges, but also include more daunting legal, logistical, and proprietary considerations. Data acquisition is further complicated by the upgrades to new networking

---

[15] The backbone of a network is "the part of the communications network which carries the heaviest traffic. The backbone is also that part of a network which joins LANs together – either inside a building or across a city or the country." NEWTON'S TELECOM DICTIONARY, *supra* note 2, at 146.

[16] NAT'L. SCI. FOUND., THE INTERNET: CHANGING THE WAY WE COMMUNICATE 6 (2008).

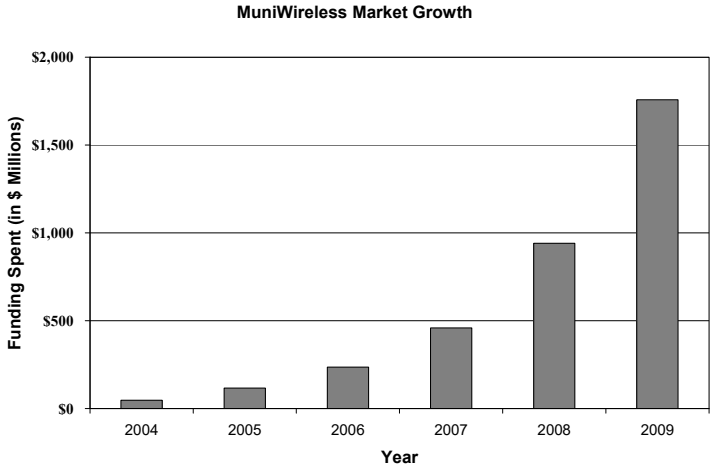
[17] *Id.* at 11.

[18] *Id.* at 12.

technologies that are generally prohibitively expensive or difficult for researchers to monitor. Diverting resources to statistics collection takes those resources directly away from forwarding of packets, which tends to drive commercial providers toward switching from vendors that sacrifice potential research functionality in exchange for performance. As a result, core backbone routers often do not have the functionality to gather the intricate data needed to support scientifically sound modeling, simulation, and analysis efforts.[19] In combination, these issues leave the Internet research community continually struggling to validate research theory. And yet, as the world becomes increasingly dependent on the Internet infrastructure, it becomes ever more critical to understand not only Internet structure, workload, and dynamics, but also the economic forces that constrain their evolution.

The Internet is at a unique turning point in its history—a time when wireless infrastructure that is "too inexpensive not to deploy" is starting to gain traction in community and public settings. Indeed, in the last several years, the growth of wireless access has increased dramatically. According to MuniWireless.com, in the municipal wireless market alone, the sector has grown from $47.4 million in 2004 to $235.5 million in 2006, with the market predicted to exceed $1 billion before the end of the decade.[20]

Figure 1. Growth of the Municipal Wireless Market (2004–2009).[21]



[19] kc claffy & Tracie Monk, *What's Next for Internet Data Analysis? Status and Challenges Facing the Community,* 85 PROCEEDINGS OF THE IEEE 1563–71 (1997).

[20] MUNIWIRELESS.COM, *supra* note 1, at 8.

[21] *Id.*

Looking at the United States, one can see both that the diversity of networks is substantial and that their numbers are rapidly increasing.

Table 1. Number and Type of Municipal Wireless Networks—July 2005 to December 2006[22]

Type of Network	July 2005	Feb. 2006	April 2006	June 2006	Sept. 2006
Region/Citywide	38	56	58	59	68
City Hot Zones	22	29	32	32	43
Municipal or Public Safety Use Only	28	32	35	35	35
Planned Deployments	34	59	69	121	135

Although the growth in municipal wireless networks may look promising, the United States is falling behind other countries with regard to Internet infrastructure penetration.[23] However, the extent to which it is lagging has been obfuscated by faulty and opaque measurement and analysis methodologies.[24] In his seminal analysis and report on the state of American broadband, S. Derek Turner, Free Press Research Director, states that policymakers must:

> [R]equire the FCC to improve its data collection on broadband markets. Policymakers cannot adequately assess the problems in the broadband market, nor identify the most appropriate solutions, if the FCC provides poor information. The starting point should be a more precise measure of which geographic areas have service (using a smaller unit than the ZIP code). Beyond that, carriers should be required to report the percentages of households where broadband service is available in every service area, the percentage of households that subscribe, and the average cost per megabit of throughput. This evidentiary record will allow an accurate analysis of the problems we face and foster solutions that will achieve results.[25]

As an example of policymakers' failures, until recently the Federal Communications Commission ("FCC") defined broadband as "data transmission speeds

---

[22] *Id.* at 29.

[23] *See* Rob Kelley, *Broadband Lag Could Hurt the U.S.*, CNNMONEY, June 17, 2005, http://money.cnn.com/2005/06/16/technology/broadband/index.htm.

[24] *See* TURNER, *supra* note 6, at 2, 4 ("The FCC uses misleading and meaningless measures of broadband coverage and competition."); *see also* Robert M. McDowell, Comm'r, FCC, Luncheon Address at the Broadband Policy Summit III (June 7, 2007), *available at* http://www.netcompetition.org/BB_Policy_Summit.pdf (explaining that the often-cited OECD broadband statistics are faulty in five respects, including measuring penetration per capita as opposed to per household, and geographic size and topology of the different countries in the study).

[25] *See* TURNER, *supra* note 6, at 36–37.

exceeding 200 kilobits per second (kbps), or 200,000 bits per second, in at least one direction . . . ."[26] On March 19, 2008, the FCC announced that they would raise the speed for basic broadband to 768 kbps.[27] Furthermore, the FCC has no national broadband plan. The closest thing the United States has to a national policy is a statement by George Bush calling for "universal, affordable" broadband access to all consumers by 2007.[28] In contrast, Japan has an active national initiative to bring 10 Gbps lines to all houses by 2010.[29] Note that current prices per megabit in Japan are around 30–50 cents, an order of magnitude lower than many places in the United States.[30] National broadband strategies have proven useful to the countries that have pulled ahead of the United States in terms of broadband penetration rates and price per megabit by helping structure and prioritize deployment of broadband infrastructure.

The COMMONS aims to improve the availability and reliability of Internet research by providing accurate empirical data on which Internet policy decisions can then be based. Peering networks such as the COMMONS envisions reduce the costs of research while benefiting researchers, policymakers, and the public at large.

III. QUESTIONS ADDRESSED BY THE COMMONS STRATEGY WORKSHOP

As a scientific and public-service hybrid, the COMMONS is a multi-faceted research and development endeavor that will allow measurable progress in those areas identified by participants and others. The COMMONS Strategy Workshop addressed seven key areas identified by workshop participants as critical to the initiative: (1) infrastructure issues; (2) regulatory harmonization and reform; (3) outreach and education; (4) research and technological development; (5) business model innovation; (6) expansion of broadband services; and (7) vision for the future. For participating networks, the incentive to join

---

[26] FCC CONSUMER FACTS: BROADBAND OR HIGH-SPEED INTERNET ACCESS 1 (2008), http://www.fcc.gov/cgb/consumerfacts/highspeedinternet.pdf.

[27] *See* Press Release, FCC, FCC Expands, Improves Broadband Data Collection (Mar. 19, 2008), http://hraunfoss.fcc.gov/edocs_public/attachmatch/DOC-280909A1.pdf

[28] *See* President George W. Bush, Remarks by the President in a Conversation on Homeownership (Mar. 26, 2004), *available at* http://www.whitehouse.gov/news/releases/2004/03/print/200040326-15.html.

[29] Corporation for Education Network Initiatives in California, *Technology*, FIRST MILE, Nov. 12, 2004, at 2.

[30] *See* ORG. FOR ECON. CO-OPERATION & DEV., OECD COMMUNICATIONS OUTLOOK 2007, at 222–23 (2007), *available at* http://213.253.134.43/oecd/pdfs/browseit/9307021E.pdf (noting prices as low as $.22 per Mbits in Japan versus $3.18 in the United States for October 2006).

the COMMONS is similar to that of participants of Internet2,[31] National LambdaRail,[32] Quilt,[33] and thirty-three state networks trying to execute similar agendas on a smaller scale.[34] Those incentives include collective buying power, ongoing access to extensive research data, affordable fiber infrastructure, and transparent and accountable collaboration.

A. Infrastructure Issues, Expansion of Broadband Services, and Business Model Innovation

As envisioned, the COMMONS will build neutral, open optical networks by peering on National LambdaRail, Internet2, or other allied networks. Such a goal necessitates understanding local and regional demands and identifying linkages among communities seeking unfettered connectivity. Creating an efficient buy-in process for regional networks to cooperate with one another avoids duplicating the efforts of individual networks.

Peering networks, as the COMMONS envisions, will expand service to low-income constituencies, as well as those with disabilities and those that are underserved. The project will reach out to community wireless networks as well as interconnecting with Canada and international networks.

With regard to innovative business models, the COMMONS will allow assessment of the financial sustainability and scalability[35] of aggregating broadband demand. It will foster lowered bandwidth pricing through transparent empirical analysis of backhaul[36] and backbone[37] structures, and will provide the ability to evaluate the economics and performance of "quality of service" against over-provisioning to meet user and application demands.

---

[31] *See* About Internet2, http://www.internet2.edu/about (last visited Apr. 4, 2008).

[32] *See* About National LambdaRail, http://www.nlr.net/about (last visited Apr. 19, 2008).

[33] *See* Quilt, Mission, http://www.thequilt.net (last visited Apr. 4, 2008).

[34] For example, the Illinois Century Network serves over 8,000 clients throughout the state, including libraries, schools, and municipalities. Ill. Dep't of Cent. Mgmt., Next Century Network, http://www.illinois.net/next (last visited Apr. 18, 2008).

[35] Scalability refers to the size to which something can grow relatively easily. NEWTON'S TELECOM DICTIONARY, *supra* note 2, at 812.

[36] A communications channel is backhauling when it takes traffic beyond its destination and back. In fiber networks, backhauling is a traffic management technique used to reduce the expense of multiplexing/demultiplexing. *Id.* at 145.

[37] *See supra* note 15.

B. Regulatory Harmonization and Reform, Outreach and Education, and Research and Technical Development

In addition to increasing connectivity and encouraging new business model innovation by improving accountability and research methodologies of both carriers and regulators, the COMMONS would provide empirical research to support national telecommunications policy. Such research would facilitate transparent negotiation among public and corporate interests for assets such as rights-of-way.[38]

In addition to educating policymakers and informing telecommunications policy, the COMMONS would aim to educate regulators and the public, correcting misinformation and myths using the best available empirical Internet data. These efforts would include answering concerns of users and organizations regarding online privacy. Furthermore, the data derived from the COMMONS would be provided to expert agencies for independent research and analysis.

Analysis of the COMMONS data will foster a better understanding of why networks become overloaded, including analyzing provisioning models in economic terms as well as technological terms. Furthermore, the COMMONS will provide data that facilitate the objective assessment of the social impact of proposed initiatives and will also pioneer acceptable use policies and research methodologies that balance privacy concerns against data-retention concerns.

C. Envisioning the Future

The potential impact of the COMMONS cannot be overestimated. It will directly benefit multiple constituencies across all levels of society, including private citizens, municipalities, businesses, corporations, network operators, content and service providers, public services, and the scientific community. It will also enable local and national politicians, regulators, and legislators to intelligently influence broadband policy. Future goals of the COMMONS involve developing a long-term vision for project assessment and follow-ups, building partnerships among and between academics and communities to support Internet science, and achieving universal affordable broadband access within the next ten years. Taken together, participants in the COMMONS envision a collaborative community research environment that truly will expand broadband research, access, and policy into the future.

---

[38] "Carriers and service providers . . . must obtain right-of-way to dig trenches or plant poles for cable systems, and to place wireless antennas." NEWTON'S TELECOM DICTIONARY, *supra* note 2, at 704.

IV. PROPOSED EXPERIMENT

As discussed at the COMMONS Strategy Workshop, this article proposes that Cisco,[39] CAIDA,[40] and a national backbone resource (to be determined) join together with community, local, municipal, regional, and state networks to support a large-scale, incentive-based experiment in end-to-end network workload, performance, economic, and behavioral measurement on an unprecedented national, inter-segment, inter-provider scale. Specifically, this article proposes to develop a requirements document and roadmap to support the use of a national OC-192[41] transit backbone[42] for community wireless networks and other public sector networks to reach each other. The project will include support for native multicasting[43] of public sector services (national, state, local), including classes for any schools interested in sharing them.

In exchange for free or low-cost transit, the attached networks would agree to collaborate with network researchers in specific ways. For example, the attached networks would allow researchers access to both historical and current operational data, in appropriately anonymous form to protect the users, to study the network. Networks would agree to permit or participate in occasional openly-reviewed experiments required to test new technologies, and would make customized end-user polling tools available to community network users who individually volunteer to participate in project-related behavioral research. A prerequisite to attachment would be a commitment to adherence to responsible general administrative guidelines as set by the Policy Board to ensure that project-funded resources are used in a secure and appropriate manner.

Participating network operators would be selected to provide a broad sample of access media, connection methods, and operational environments. Participation would be conditioned on willingness to facilitate and contribute to data collection under normal operating conditions and to collaborate in occasional network-based experiments.

---

[39] Cisco Systems, Inc. is a supplier of networking equipment and network management for the Internet. Cisco Home Page, http://www.cisco.com (last visited Apr. 19, 2008). The COMMONS Strategy Workshop included two representatives from Cisco. CAIDA, COMMONS Workshop Participants List, *supra* note 5.

[40] CAIDA, the Cooperative Association for Internet Data Analysis, provides tools and analyses promoting the engineering and maintenance of a robust, scalable global Internet infrastructure. CAIDA, Home Page, http://www.caida.org (last visited April 16, 2008). The COMMONS Strategy Workshop was hosted at CAIDA's offices at the San Diego Supercomputing Center.

[41] NEWTON'S TELECOM DICTIONARY, *supra* note 2, at 665–66.

[42] Transit backbones interconnect different networks, providing the "connective tissue" that facilitates peering and transport of data throughout the Internet.

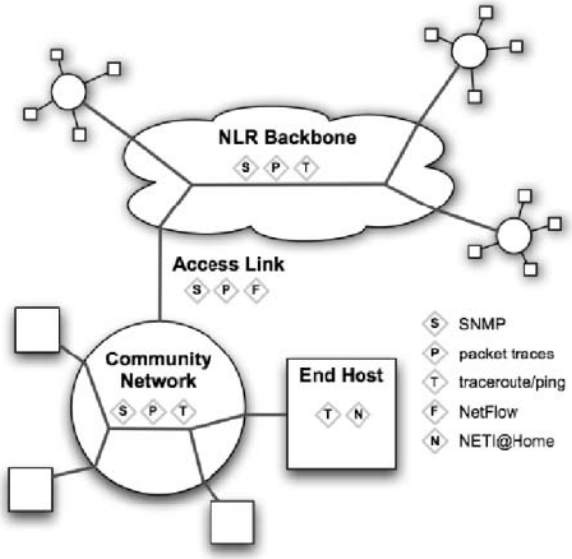
[43] Multicasting is a process whereby a data packet is sent to several different recipients (i.e., any device with a multicast address) instead of to a single machine. Unlike "broadcast" data packets, multicasting may go to many, but not necessarily all, devices within a network.

A. Possible COMMONS Measurement Architecture

Maintaining funding for Internet measurement infrastructure past the span of a given funded research project has thus far eluded the Internet research community. This failure has a substantial negative impact, not only on the goal of conducting scientifically sound Internet research, but also on all large-scale networking research that requires empirical validation. The COMMONS can directly address this crisis in a way no other existing network can. This article proposes the development and deployment of measurement infrastructure for the COMMONS that guarantees measurable progress toward restoring the intellectual strength of a wide range of Internet modeling, simulation, analysis, and theoretical research activities currently occurring without any validation. The measurement data gathered from this infrastructure will advance at least four areas: (1) support for validation of scientific research; (2) development of new measurement technology; (3) evaluation of proposed future Internet architectures; and (4) empirical answers to questions of critical national security and public policy importance.

Based on information collected during the COMMONS Strategy Workshop, this article proposes four distinct levels of measurement to support various policy and economic constraints as well as various research needs. This diversity of measurement types will facilitate correlational studies that were not previously possible. Figure 3 illustrates these four levels: (1) the backbone; (2) the backbone access link; (3) the attached community network; and (4) the end host. Figure 3 also lists examples of the types of measurements that the COMMONS should support at each specific level. Workshop participants recognize that measurements at each level will have various costs, precision, and utility. Indeed, an important early objective of the project will be to determine what data can be gathered, and at what granularity, to meet the needs of as much of the Internet research community as possible at the lowest overall cost.

Figure 2: Four Measurement Levels Incorporated into the COMMONS



1. Backbone Measurements

The COMMONS backbone will support measurements that provide an indication of overall backbone health and utilization (e.g., capacity, congestion, actual use, and throughput). As an optical fabric, the backbone could support direct measurement of wavelengths,[44] or provide SNMP[45] counters as supported in the attached equipment. The COMMONS initially proposed to collect a simple set of measurements. In the future, the COMMONS may want to invest additional development resources in optical network measurement hardware,[46] and will need to establish vetting procedures and clear disclosure procedures for adding any operational measurement functionality to the backbone.

---

[44] *See, e.g.,* Jörg B. Micheel, *LambdaMON – A Passive Monitoring Facility for DWDM Optical Networks, in* PASSIVE AND ACTIVE NETWORK MEASUREMENT 228 (Constantinos Dovrolis ed., 2005) (explaining that the lambdaMON system uses dense wavelength division multiplexing to passively monitor high performance optical networks).

[45] Simple Network Management Protocol ("SNMP") is the most common method by which network management applications can query a management agency using a supported Management Information Base. NEWTON'S TELECOM DICTIONARY, *supra* note 2, at 852.

[46] Optical network measurement hardware is specifically designed to collect data on

2. Access Link Measurements

Access links connect the peering router of the attached community network with the peering routers of the backbone. Examples of measurements on access links include: (1) passive measurement of packet traces[47] on attached links; (2) flow measurement; and (3) SNMP counters. A more holistic analysis of network functionality can only be garnered through a multi-level, multi-method approach to network data collection. By combining passive and active measurements, focusing on data collection on different links of a network, and ensuring that diverse data are gathered, the COMMONS will help researchers gain a far better understanding of the intricacies of real-world Internet conditions.

3. Community Network Measurements

Within an attached local network, a wider variety of measurement options are both possible and expected. All of the access link measurements can also apply to links within the community network. Additionally, networks would contribute to: (1) anonymous HTTP Web proxy logs; (2) anonymous DNS statistics;[48] (3) traceroute server logs; (4) large-scale traceroute probes;[49] (5) support for other active measurement experiments as approved by the COMMONS policy review board;[50] (6) BGP peering[51] with, for example, the COMMONS-routeviews;[52] (7) anonymous or aggregated Web cache logs for analy-

---

fiber optic (as opposed to copper-based) networks. Since light (instead of electricity) is used on fiber optic networks, both passive and active monitors are specifically built for this medium.

[47] Packet tracing refers to "the monitoring and reporting a particular packet addresses or types for diagnostic purposes." NEWTON'S TELECOM DICTIONARY, *supra* note 2, at 690.

[48] DNS statistics are data collected from domain name servers, which are machines that connect IP addresses (e.g., 208.77.188.166) to specific domains (www.example.com).

[49] Traceroute is a widely used tool for determining the path of a packet through a network. Traceroute logs and probes help identify each router that a packet passes through and can be useful in determining the topology of a network and locating areas of congestion.

[50] *See* CAIDA 2007–2010 Program Plan, http://www.caida.org/home/about/progplan/progplan2007 (last visited Apr. 19, 2008).

[51] BGP Peering is peering using Border Gateway Protocol. Border Gateway Protocol is a Gateway Protocol that routers (other non-router devices also may be involved as intermediaries) employ in order to exchange appropriate levels of routing information. When BGP peer routers (routers with a TCP connection for purposes of exchanging routing information) first establish contact, they exchange full routing tables, which are maintained in Routing Information Bases. Subsequent contacts involve the transmission of incremental changes only. NEWTON'S TELECOMM Dictionary, *supra* note 1, at 160.

[52] Accurate knowledge of autonomous system ("AS") relationships is a requirement of many research tasks. CAIDA collects and analyzes, on an ongoing basis, AS-level topology and AS relationships, and provides this data for use by the community. CAIDA's data are available from 2004 to present, with one file created per week in 2006 and one per month in

sis; and (8) making tools available to community network users who individually volunteer to participate in research.[53]

4. End Host Measurements

Finally, the COMMONS will actively support the investigation of a recent methodological advance in the field of Internet measurement—peer production. Several existing measurement projects have drawn on the inspiration of SETI@Home to develop client-based measurement software for use in a peer production model.[54] This model takes advantage of end users' volunteering their hosts to the measurement infrastructure by downloading and executing measurement software that sends gathered data or statistics back to the project's central processing site.[55] Receiving data from unknown, untrusted users presents particularly daunting, though not insurmountable challenges.[56]

For active measurements, the biggest challenges are: (1) deployment to insure low impact on the infrastructure;[57] (2) prevention of use of tools for Distributed Denial of Service attacks;[58] (3) accountability of measurement source in case of operational problems;[59] (4) analysis of bias due to self-selection of sources (by volunteers);[60] and (5) validation of the integrity of resulting data.[61]

---

prior years. Each file contains a full AS graph derived from RouteViews BGP table snapshots taken at eight-hour intervals over a five-day period. The AS relationships available are customer-provider (and provider-customer in the opposite direction), peer-to-peer, and sibling-to-sibling.

[53] *See* CAIDA 2007–2010 Program Plan, *supra* note 50.

[54] *See* David P. Anderson, Jeff Cobb, Eric Korpela, Matt Lebotfsky & Dan Werthimer, *SETI@home: An Experiment in Public Resource Computing* 45 COMM. OF THE ACM 56, 56 (2002) (explaining that Search for Extraterrestrial Intelligence at Home (SETI@Home) is a project in which "[m]illions of computer owners worldwide contribute computer time to the search for extraterrestrial intelligence, performing the largest computation ever.").

[55] *See, e.g.*, NETI@home, http://www.neti.gatech.edu (last visited Apr. 4, 2008) ("NETI@home is an open-source software package, named after the widely popular SETI@home, that collects network performance statistics from end-users.").

[56] *See* Yuval Shavitt & Eran Shir, *DIMES: Let the Internet Measure Itself,* 35 COMPUTER COMM. REV. 71, 73–74 (discussing how the DIMES project utilizes volunteers whose contributions must be quantified).

[57] One of the major tasks for network researchers is to conduct their studies without causing artifacts or disruptions to existing services. Because data is traveling along the Internet backbone at the speed of light, even outages lasting a second or two can cause enormous headaches.

[58] A distributed denial of service attack often uses seemingly innocuous tools en masse to lock up a targeted server. While the tools used in the COMMONS have enormously useful potential, it is also imperative that they be built and deployed in ways that prevent their use for malicious intent.

[59] Network monitoring systems need to protect network user privacy, ensure the integrity of the data collected, and be deployed in ways that allow network operators and system administrators to oversee them in cases where problems might arise.

[60] Methodologically, data are only as good as the analyses conducted. Sampling bias

Client-side passive measurement infrastructures have all the same problems as active measurement infrastructures, but in addition, they bear formidable privacy challenges. For certain measurement questions, such as provider or application prevalence on the Internet, there is substantial material incentive to manipulate a macroscopic Internet measurement system. Thus, it is essential for both participant and researcher to trust in the integrity of the measurement.

Researchers do not yet have an in-depth understanding of the methodological problems of scaling Internet measurement paradigms to incorporate peer production. This article proposes a strategic approach toward a model for peer production of Internet measurements that cross-validates client-produced data with more trusted measurements from controlled infrastructure measurement devices. COMMONS Strategy Workshop participants expect that the next few years will serve as a transition period during which researchers can determine if the same integrity can be achieved from client-side infrastructures as is now achieved from controlled infrastructures.

B. Meta-infrastructure to Support Protected Data Sharing

Because the COMMONS will initially be composed of cooperative Internet Protocol ("IP") networks, commercial counterincentives to rapid and broad sharing of information will not frame inter-network security and communication policies. If launched, the COMMONS will establish a security response team to aid participating networks with integrating security best-practices into their network operations. The cooperative structure of the COMMONS will also allow research into areas that have proved impossible in the unregulated commercial framework. Such research areas will include: testing secure naming and routing protocols; coordinating public and private sector response to and recovery from major Internet disruption; developing fundamentally new architectural components that might emerge from the National Science Foundation's Global Environment for Networking Innovations ("GENI") program;[62]

---

needs to be ruled out (or accounted for) in order to maximize the applicability of the conclusions stemming from the COMMONS research.

[61] Due to the complexities involved in real-world data collection, it will be important to systematically verify the reliability of the data collected to ensure that conclusions drawn from the COMMONS research are valid.

[62] Nat'l Sci. Found., Global Environment for Networking Innovations (GENI), http://www.nsf.gov/cise/cns/geni (last visited Apr. 4, 2008).

> GENI is designed to allow experiments on a wide variety of problems in communications, networking, distributed systems, cyber-security, and networked services and applications. The emphasis is on enabling researchers to experiment with radical network designs in a way that is far more realistic than they can today. Researchers will be able to build their own new versions of the "net" or to study the "net" in ways that are not possible today. Compatibility, with the Internet is NOT required. The purpose of GENI

implementing state-of-the-art wireless techniques and policies to promote efficient utilization of spectrum; and economic modeling of a wide variety of end-to-end paths with particular aim toward transparent techniques to analyze capital and operating cost accounting.

Respect for user privacy is essential to CAIDA's mission and the goals of the COMMONS. COMMONS participants are acutely aware of the sensitivities involved in sharing of Internet measurements. CAIDA already participates in projects such as the Protected Repository for the Defense of Infrastructure Against Cyber Threats ("PREDICT")[63] and DatCat,[64] both of which are framed by these sensitivities. The launching of both of these projects makes this a perfect time to establish the COMMONS to leverage these preexisting projects and build a laboratory supporting Internet researchers across the country.

CAIDA's leadership will ensure that the COMMONS measurement infrastructure satisfies the measurement needs of the larger Internet research community while protecting the best interests of participating networks and geographic communities. For active measurement infrastructure, the primary concerns are often how to coordinate measurement requests from a large and diverse group of researchers, and how to ensure integrity of the data when gathered by an unknown party. On the other hand, for passive measurement infrastructure, the primary concerns are often the cost of hardware for high-speed trace collection and protected access to trace data. In both cases, an incentive-based cooperative model will help ensure that as many needs as possible are cost-effectively met.

1. The PREDICT Project: A Legal Framework to Support Protected Data Sharing

In 2004, the United States Department of Homeland Security ("DHS") recognized the need to support the calibration of cyber security tools[65]—particularly those funded by government agencies—in real world environments.[66] After extensive consultation with privacy law experts, DHS is cur-

---

> is to give researchers the opportunity to experiment unfettered by assumptions or requirements and to support those experiments at a large scale with real user populations.

GENI FAQ, http://www.geni.net/faq.html (last visited Apr. 4, 2008).

[63] Protected Repository for the Defense of Infrastructure Against Cyber Threats (PREDICT), Operational Overview, http://www.predict.org [hereinafter PREDICT]. *See infra* Part IV.B.1.

[64] *See infra* Part IV.B.2.

[65] Examples of cyber security tools include samples of normal and malicious Internet traffic, malicious software samples, logs from machines compromised in targeted attacks, and other data to develop hardware and software that protects against and mitigates the effects of hacking attempts and malicious software.

[66] Dep't of Homeland Security, About the Cyber Security Research and Development

rently in the process of launching the PREDICT Project to allow researchers to request real-world datasets to assist their research.[67] The goal of the PREDICT Project is to facilitate the development of Internet defense technologies, products, models, and strategies.[68] DHS has made noteworthy progress in the legal and privacy aspects of infrastructure data access, specifically addressing the concerns of Internet Service Providers that want to support the research community, but are constrained by privacy concerns, laws, and policies.[69] The presence of such a framework bodes well for the viability of the COMMONS and demonstrates a flexible and participatory path to achieve both protection of privacy and support for empirical network science.

2. DatCat: Internet Measurement Data Catalog

For the past several years, CAIDA has been developing a data cataloging system for the Internet research community. Despite its necessity to scientific endeavor, data available to researchers are limited by legal, social, and technical constraints on its collection and distribution. Thus, the distribution of available data is a valuable service to the general research community. To this end, CAIDA has developed the Internet Measurement Data Catalog ("DatCat") to provide a searchable index of available data, enhance documentation of datasets via a public annotation system, and advance network science by promoting reproducible research.[70] Like the PREDICT Project, DatCat is a critical supporting infrastructure to the COMMONS.

C. Broader Impact: Economics, Regulations, and Policy

Measuring the Internet for a decade provides a reliable way to learn how economic issues impact a field's ability to make scientific progress. When the companies that own the infrastructure under study are declaring bankruptcy, measurements are scarce. Likewise, when the companies that own the infrastructure are competing against one another, whatever measurements do exist often are considered extremely sensitive or completely proprietary. Yet changing technologies, commercial strategies, and regulatory policies have brought dramatic restructuring of Internet service delivery at local, national, and global levels. Accompanying these changes are a variety of strong but conflicting

---

Center, http://www.cyber.st.dhs.gov/about.html (last visited Apr. 19, 2008).

[67] *See* PREDICT Home Page, http://www.predict.org (last visited Apr. 4, 2008).

[68] *Id.*

[69] *Id.*

[70] DatCat, Internet Measurement Data Catalog, http://www.datcat.org (last visited Mar. 25, 2008).

(and generally unverified) assertions concerning the relative feasibility, necessity, and superiority of different possible outcomes of this restructuring.

These assertions pose a grave dilemma for both researchers and policymakers. Given the increasingly critical role of information and communications technologies for national productivity, economic competitiveness, and even security, the costs of error could be grievous. Yet decision makers are often forced to operate in an information vacuum. They are often placed in the position of only having access to the information that the companies which would be affected by policy and regulatory changes are willing to share.

In the United States, both telecommunications companies and user advocates are unhappy with the current state of communications policy. Prices for services are higher than in many other industrialized nations, and broadband penetration is lower.[71] Telecommunications companies have increasing trouble attracting investment, and claim their broadband services need exemption from common carrier regulation in order to thrive.[72] Solutions tend to focus on industry-centric approaches to policy reform, such as how much price-control leverage should the government have over telecommunications carriers; how much freedom should telecommunications carriers have to price-discriminate; how much subsidization of telecommunications companies is necessary; and which entities should be forced to pay for universal service for rural areas. The dearth of Internet research makes it impossible to come up with reliable empirical answers to many of the questions to which regulators and politicians need answers. Thus, national telecommunications policy is forced to advance blindly at a time when the United States is losing its competitive broadband edge as compared to a growing number of industrialized countries. The COMMONS provides a collaborative environment for policymakers to help shape the research questions under study and offers a vital resource for regulators seeking to make decisions based upon empirical scientific research.

V. CONCLUSION

Building on the momentum from the COMMONS Strategy Workshop, outreach to community and municipal networks, particularly wireless initiatives,

---

[71] Robert D. Atkinson, *Framing a National Broadband Policy,* 16 COMMLAW CONSPECTUS 145, 146 (2007).

[72] This dynamic is probably best exemplified by the filings made by telecommunications companies during the Supreme Court *Brand X* case and proceedings at the FCC regarding Title I and Title II reclassification. *See* Nat'l Cable & Telecomm. Assoc. v. Brand X Internet Svc., 545 U.S. 967 (2005); *see also*, Eli M. Noam, *Beyond Liberalization II: The Impending Doom of Common Carriage*, 18 TELECOMM. POL'Y 435 (1994); WILLIAM JONES, THE COMMON CARRIER CONCEPT AS APPLIED TO TELECOMMUNICATIONS: A HISTORICAL PERSPECTIVE (1980), *available at* http://www.cybertelecom.org/notes/jones.htm.

has increased dramatically. In addition, CAIDA staff are identifying community networks to be involved in the first phase of the COMMONS, their measurement/data-collection capabilities, and the support resources needed to help maximize the utility of these measurements.

The COMMONS is one of the most important experiments the Internet research community has ever considered, and is an exciting and innovative partnership among industry, researchers, community organizers, network operators, fiber owners, municipalities, and policymakers. Through collaborative peering and rigorous data collection and analysis, the COMMONS facilitates both basic research and innovative improvements to the Internet. The COMMONS presents an unprecedented opportunity for establishing standards for scientific integrity for Internet research using rigorous empirical data to validate theories, models and simulations. In particular, the measurement data gathered through the COMMONS will lend unique strategic significance to National Science Foundation's GENI program. The potential outcome of this project promises researchers a clearer picture of the nature and characteristics of the current Internet than ever before possible, while informing discussions of future architectures and related design issues. The COMMONS also provides an opportunity for opening up the economics, ownership, and trust layers of Internetworking in much the same way the transport, network, and application layers of the Internet are open to innovation. At this critical juncture in telecommunications history, the COMMONS creates a much needed resource to help chart the future of the Internet.