

The COMMONS Initiative: Cooperative Measurement and Modeling of Open Networked Systems

Sascha Meinrath, CAIDA, sascha@caida.org
kc claffy, CAIDA, kc@caida.org

Executive Summary

On December 12-13, 2006, the Cooperative Association for Internet Data Analysis (CAIDA) held a workshop to discuss and ultimately propose a collaboration among researchers and networks to simultaneously solve three acute and growing problems facing the Internet: a self-reported financial crisis in the Internet infrastructure provider industry that poses a severe threat to broadband growth and U.S. competitiveness; a data acquisition crisis which has deeply stunted the field of network science; and a dilemma within emerging community, municipal, regional, and state networks, who need (additional) broadband connectivity but face severely 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. COMMONS Peering will be conditionally available to city, county, state, and federal government entities, academic institutions, community Internet initiatives (e.g., community wireless networks), and commercial entities based upon the following three conditions: (1) networks will make select operational data available to COMMONS researchers (under appropriate legal data sharing and privacy guards); (2) the attached networks must agree to develop and abide by COMMONS policies which will be based upon research results of empirical data analyses of network usage; and, (3) participating networks must abide by the Acceptable Use Policies set by the COMMONS project coordination committee.

CAIDA hosted 40 networking visionaries from across North America at the UCSD campus to discuss technical, policy, and operational issues related to the COMMONS initiative. This “COMMONS Strategy Workshop” brought together representatives from industry, community and municipal networks, regional and state networks, as well as Internet researchers, community organizers, and developers building next-generation data communications technologies. Workshop participants also included heads of research, infrastructure, media, and policy organizations, as well as telecommunications lawyers. The diversity of workshop participants helped ensure that all major project stakeholders were represented in the proceedings and provided invaluable insights into the potential pitfalls and opportunities the COMMONS Project faces in coming years.

At a time when research on broadband service provision is desperately needed to help forge new national telecommunications policies and inspire innovation in networking technologies [1], the COMMONS Project is an innovative platform that will provide vital research results for policy makers across the country and around the world. This report describes the findings from the COMMONS Strategy Workshop, outlines a set of relevant open research problems identified by participants, and concludes with recommendations that will benefit the scientific community, network operators and developers, key decision-makers, and the general populace.

1. COMMONS workshop

On December 12-13, 2006, CAIDA hosted the first COMMONS Strategy Workshop (Workshop) at the University of California's San Diego Supercomputer Center. Workshop participants included an unprecedented collection of relevant players: community and municipal wireless builders, measurement experts, wireless technologists and researchers, policymakers, legal and privacy experts, and industry participants. Workshop attendees discussed the design, creation, and operation of an experimental infrastructure that could simultaneously address three core crises: the digital divide, scientific integrity of network research, and inability to empirically inform policy decisions at a critical juncture in telecommunications history. **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. COMMONS Peering will be conditionally available to city, county, state, and federal government entities, academic institutions, community Internet initiatives (e.g., community wireless networks), and commercial entities based upon the following three conditions: (1) networks will make select operational data available to COMMONS researchers (under appropriate legal data sharing and privacy guards); (2) the attached networks must agree to develop and abide by COMMONS policies which will be based upon research results of empirical data analyses of network usage; and, (3) participating networks must abide by the Acceptable Use Policies set by the COMMONS project coordination committee.** While the COMMONS Project will broadly impact the Internet industry, both commercial and non-profit, as well as policymaking, 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 increased standards of data collection, curating, and sharing in the research community.

2. Motivation

The United States is facing a worsening broadband crisis -- over the past half-decade, the US 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 [2]. Rural and poor communities are being doubly discriminated against -- often receiving little or no broadband access and being forced to pay higher service rates when they do have access [3]. Due to regulatory, political, and market constraints of Incumbent Local Exchange Carriers (ILECs) and other broadband providers, Internet researchers have been unable to study mission-critical aspects of the Internet [4] and the state of its current robustness, capacity, usage, problem areas, which means that potential solutions to these issues remains more of an area of conjecture than one with empirically-backed analysis [5]. Meanwhile, telecommunications regulators and policy-makers have increasingly called for methodologically sound study of Internet usage, analysis of potential failure points, and improvements to this vital infrastructure [6, 1]. COMMONS provides an opportunity to address all of these shortcomings within a single national initiative. By creating a national peering infrastructure that interested network operators can choose to join, COMMONS will provide numerous opportunities and benefits for a range of different constituencies. However, there are also myriad potential obstacles and difficulties, which the purpose of this workshop was to articulate and analyze.

The telecommunications sector has not yet recovered from the privatization and commercialization of Internet infrastructure which began in 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. 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, i.e., differential pricing by type of traffic [7]. Legal scholars have long argued that this development is arguably a constitutional threat -- to the First Amendment -- since providers would thus have a lever to control how users of their infrastructure communicate [8]. From a technical and historical perspective, First Amendment issues on the Internet are not new, they surfaced most loudly in the fight against spam [9]. However the scope and depth of control being promulgated by network owners in the face of economic pressure is potentially far more invasive than previously seen.

By offloading the responsibility for supporting Internet service delivery to “unprofitable areas” from commercial providers, the COMMONS initiative will measurably alleviate the economic pressure on these providers. Additionally, 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 testbed will support public analysis of actual Internet traffic, it will inform debates on increasingly important technical, economic, policy, privacy, and social issues related to the Internet. Finally, the COMMONS project 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 learned with one-another.

3. Background

The first general purpose Internet backbone, NSFNET, was funded and administered by the National Science Foundation to support the networking needs of the research and education community and was instrumented to provide continuously collected data on the function and usage of the network. When this backbone was decommissioned in the mid-1990s as a part of the US government’s strategy to privatize the network infrastructure, the attached academic networks transitioned to self-sustaining funding models. Unfortunately, ever since that time, access to representative data on Internet traffic, topology, routing, and security has diminished as Internet researchers have struggled to do legitimate, reproducible science 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 often include more daunting legal (privacy), logistical, and proprietary considerations. Data acquisition is further complicated by the upgrades to new networking technologies that are generally prohibitively expensive or impossible for researchers to monitor. Statistics collection functionality takes resources directly away from forwarding of packets/frames, which tends to drive commercial providers toward switches from vendors that sacrifice potential research functionality in exchange for forwarding performance/throughput. As a result, core backbone routers often do not have the functionality to gather the fine-grained data needed to support scientifically sound modeling, simulation, and analysis efforts. In combination, these issues leave the Internet research community continually struggling with how to validate research theory. And yet, as the world becomes increasingly dependent on the Internet as critical infrastructure, it also becomes ever more critical to understand not only Internet structure, workload, and dynamics, but also the economic forces that constrain its evolution.

We are also at a unique inflection point in Internet 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 astounded. According the MuniWireless.com 2006 State of the Market Report [10], in just the municipal wireless market alone, this sector has grown from \$47.5 million in 2004 to \$235.5 million in 2006, with the market predicted to well exceed \$1 billion before the end of the decade:

- 2004 -- \$47.5 million
- 2005 -- \$116.9 million (147% annual increase)
- 2006 -- \$235.5 million (102% annual increase)
- 2007 -- \$459.6 million (95% annual increase, projected)
- 2008 -- \$941.0 million (105% annual increase, projected)
- 2009 -- \$1,757.5 million (87% annual increase, projected)

Graphically, one can easily see the exponential growth in these data communications services:

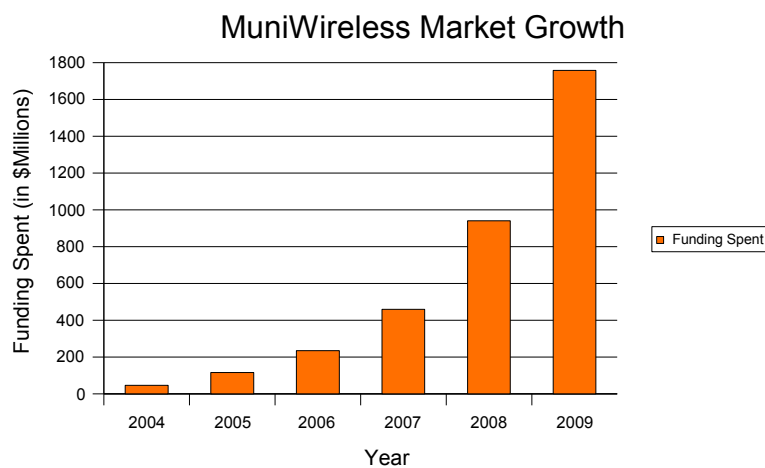


Figure 1. Growth of the municipal wireless market (2004-2009).

When one takes a look at what is happening within the United States, one can see both that the diversity of networks is substantial and that their numbers are rapidly increasing:

Type of network	Jul 2005	Feb 2006	Apr 2006	Jun 2006	Sep 2006	Dec 2006
Region/citywide	38	56	58	59	68	79
City hotzones	22	29	32	32	43	49
Muni or public safety use only	28	32	35	35	35	36
Planned deployments	34	59	69	121	135	149

Table 1. Number and type of municipal wireless networks – July 2005 to December 2006.

Once we map these networks geographically, one can see that the COMMONS Project is tapping into a national phenomenon with critical mass to support an incredible research endeavor:

Wireless Cites and Counties

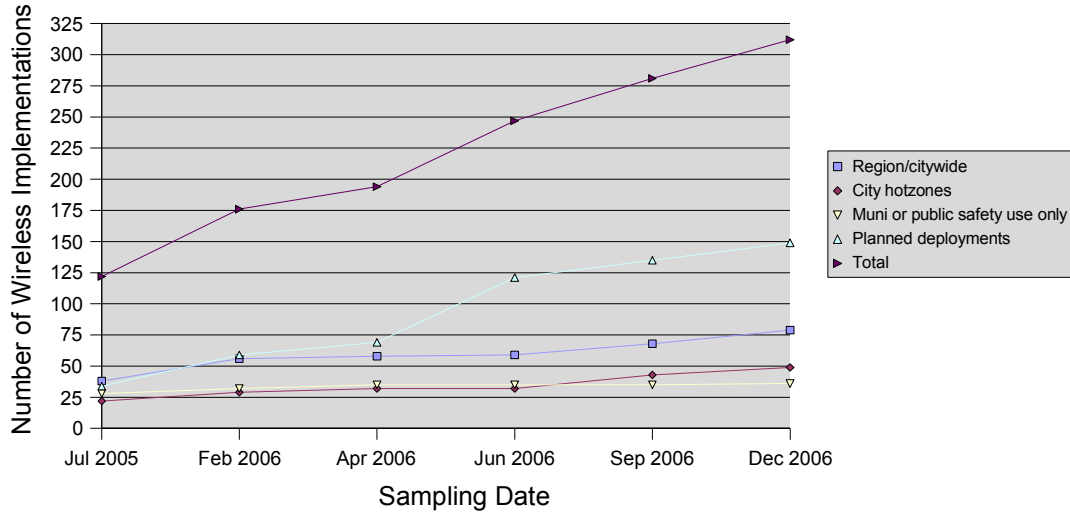


Figure 2. Number and type of municipal wireless networks – July 2005 to December 2006.

Community Internet Across America

Organized By Network Status

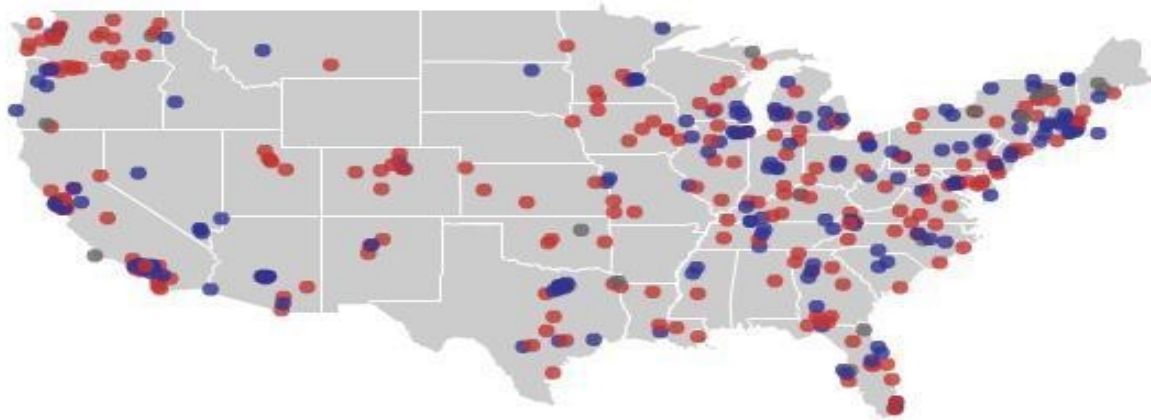


Figure 3. Potential community network partners for the COMMONS Project. Representatives from a sampling of these networks attended the December COMMONS Strategy Workshop.

Though wireless broadband numbers may look promising, the United States is falling dangerously behind with regard to Internet infrastructure penetration, although even the extent to which we're falling behind has been obfuscated by faulty and opaque measurement and analysis methodologies. In his seminal analysis and report on the state of US broadband, Derek Turner states that policymakers must, “require 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.” As just one example, the US defines broadband as 200kbps in a single direction and has no national broadband plan (the closest thing the US has had is a statement by George Bush calling for "universal, affordable" broadband access to all US consumers by 2007) – meanwhile, Japan has an active national initiative to bring 10Gbps lines to all houses by 2010 [11]. Note that current prices per megabit in Japan are around 30-50 cents per megabit, several orders of magnitude lower than anywhere in the U.S.

4. Open Problems

The workshop addressed seven key areas identified by workshop participants as critical to the initiative: infrastructure issues, regulatory harmonization and reform, outreach and education, research and technological development, business model innovation, expansion of broadband services, and vision for the future. For participating networks, the incentive to join COMMONS is similar to that of participants of Internet2, NLR, Quilt, and 33 state networks trying to execute similar agenda at smaller scale: collective buying power, ongoing access to extensive research data, affordable fiber infrastructure, and transparent and accountable collaboration. As a scientific and public-service hybrid, COMMONS is a multi-faceted R&D endeavor that will allow measurable progress on many issues of the Internet as emerging critical cyberinfrastructure:

INFRASTRUCTURE ISSUES

- building neutral, open, optical networks
- understanding local/regional demands, identifying linkages among communities seeking unfettered connectivity
- peering on NLR, Internet2, CENIC, and/or other allied networks, rather than duplicating efforts
- creating efficient buy-in processes for regional/area networks to cooperate

REGULATORY HARMONIZATION AND REFORM

- providing empirical research to support national telecommunications policy
- facilitating transparent negotiation among public and corporate interests for e.g., right-of-way
- improving accountability and research methodologies of carriers and regulators

OUTREACH AND EDUCATION

- educating the public, policymakers and regulators about infrastructure issues
- answering concerns of users and organizations regarding online privacy
- correcting misinformation and myths using the best available empirical Internet data
- providing network data to expert agencies for independent research and analysis

RESEARCH AND TECHNOLOGICAL DEVELOPMENT

- pioneering acceptable use policies and research methodologies that balance privacy and data-retention concerns
- analyzing provisioning models in economic as well as technological terms
- filtering spam and gaining a better understanding of why networks become overloaded
- preserving the end-to-end principle including support for end-to-end encryption of common applications
- assessing social impact of the proposed initiative

BUSINESS MODEL INNOVATION

- fostering lowered bandwidth pricing through transparent empirical analysis of backhaul as well as backbone cost structures
- deploying useful/P2P applications (e.g., telephony)
- assessing the financial sustainability and scalability of aggregating broadband demand
- providing peering locations as well as transit
- evaluating economics and performance of QoS vs overprovisioning to meet user and application demands

EXPANSION OF BROADBAND SERVICES

- expanding service to low income, disabled, non-profit and other un(der)served constituencies
- reaching out to community wireless networks
- interconnecting with Canada and international networks

VISIONING FOR THE FUTURE

- developing a long-term vision for project assessment and follow-ups
- building partnerships among and between academics and communities to support Internet science
- achieving universal affordable broadband access within the next 10 years

The potential impact of the proposed COMMONS initiative is impossible to overestimate. It will directly benefit multiple constituencies across all levels of our society:

- citizens/consumers, small businesses, and corporations
- fiber owners, network operators, and content/service providers
- municipalities, schools, and hospitals,
- local and national politicians, regulators, and legislators
- the scientific community

Taken together, participants in the COMMONS Project envision a collaborative community research environment that will truly expand our horizons into the 21st century and beyond.

5. Community-generated research topics

One of the main goals of the COMMONS Strategy Workshop was to generate research questions whose answers would materially and conceptually alter our view of cyberinfrastructure. The questions generated were wide-ranging, sometimes concrete and sometimes rhetorical, but provide valuable insight into the most salient research domains for COMMONS partners and the general public:

- How do we build universally accessible, open and scalable networks?
- Whose equipment works best? Which open-source wireless software/network works best?
- How do different frequencies/antennas affect network performance and end-user capacity?
- How to create a “Consumer Reports” for equipment (e.g., with standardized testing environments and benchmarks, rigorous and replicable methodology, and real-world applicability)?
- How do we build a network that we can actually test equipment on, swap out routing protocols, etc.?
- How do we create a systematic research and testing paradigm on a national scale?
- How can we build a clearing house for community wireless, best practices, economic models, and other useful information for small and medium enterprises;

how do we disseminate this information?

- How do we map the effects of local and state policies on broadband service outcomes?
- What are the inefficiencies in the current backbone? How can we quantify and expose these problems?
- What are the cost benefits of building more peer-to-peer architectures?
- How can network operators determine whether they should peer or pay for transit?
- How do community networks get access to POPs and high-speed connectivity?
- How are communities making use of a peered infrastructure?
- How do we explore different network ownership models?
- How do we support routing metrics using autonomous system numbers (ASNs)?
- How can ASNs help with better user control and traffic flow, yet not have the upstream vendor control what traffic you get?
- How do we more efficiently provide cheap connectivity?
- How do we know what different people and communities are paying for their broadband services?
- How has vertical integration affected broadband networks and Internet service provision?
- How do we improve transparency of market data?
- How do we measure the quality of usage?
- How do we measure the numbers of computing devices on a network?
- How does the network traffic behave?
- How does latency and jitter on the network change over time?
- Where does traffic on the network go? How can we geographically map network traffic?
- How much network traffic is (a)symmetrical and how much is file-sharing?
- How is (multi)media being used on the network?
- How is QoS affected by different traffic patterns?
- How can we monitor alerts and attacks in real-time to prepare for and defend against cyber attacks?
- How can we use protocols like Unicast Reverse Path Forwarding (URPF) to help prevent cyber attacks?
- How well is the DNS working? What improvements can we make to DNS?
- What is the true cost of spam to the major ISPs?
- How can we help solve the spam problem? What are better mechanisms for firewalling networks?
- What are the demographics and profiles of network users?
- How is local information used by network participants?
- Where are users on the network compared with where their data is stored?
- How do we build innovative and mutually beneficial collaborations among and between community groups and the research community?
- How do we anonymize data while still preserving the maximum utility possible for research?
- How do we negotiate dissemination of data and protect individuals' right to privacy?

Given the scope of interests already identified by current COMMONS participants, an equally bold initiative capable of addressing many of these research questions should be undertaken.

6. Proposed experiment

We propose that Cisco, CAIDA, 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, we propose to develop a requirements document and roadmap to support the use of a national OC-192 transit backbone for

community wireless networks and other public sector networks to reach each other. The project will include support for native multicasting of public sector services (national, state, local), including K-12 and university 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, for example:

1. Allowing researchers access to historical and current operational data, where necessary in appropriately anonymized form to protect the users, to study the network;
2. Agreeing to permit and/or participate in occasional openly reviewed experiments required to test new technologies;
3. Making customized end-user polling tools available to community network users who individually volunteer to participate in project-related behavioral research;
4. Committing to adherence to responsible general administrative guidelines as set by the Policy Board to insure that project-funded resources are used in a secure and appropriate manner, e.g., ingress filtering to prevent spoofing, sufficient logging to support DOS traceback, 24x7 accessible networking support contact, network neutrality requirements.

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

7. Possible COMMONS measurement architecture

Maintaining funding for Internet measurement infrastructure past the span of a given funded research project has eluded the Internet research community. This failing has a substantial negative impact, not only on the goal of conducting scientifically sound Internet research, but on all large-scale networking research that requires empirical validation. COMMONS can directly address this crisis in a way no other existing network can. We propose the development and deployment of measurement infrastructure for 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 our capabilities in at least four areas: support for validation of scientific research; development of new measurement technology; evaluation of proposed future Internet architectures; and, empirical answers to questions of critical national security and public policy importance.

Based on information collected during the COMMONS Strategy Workshop, we propose 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 not previously possible. Figure 4 illustrates these four levels: the backbone; the backbone access link; the attached community network, and the end host. The legend in the figure lists examples of the types of measurement we propose COMMONS support at each specific level. We recognize that measurements at each level will have various costs, precision, and utility. Indeed, a key early deliverable 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 lowest overall cost. The next sections describe these four levels of measurement in more detail.

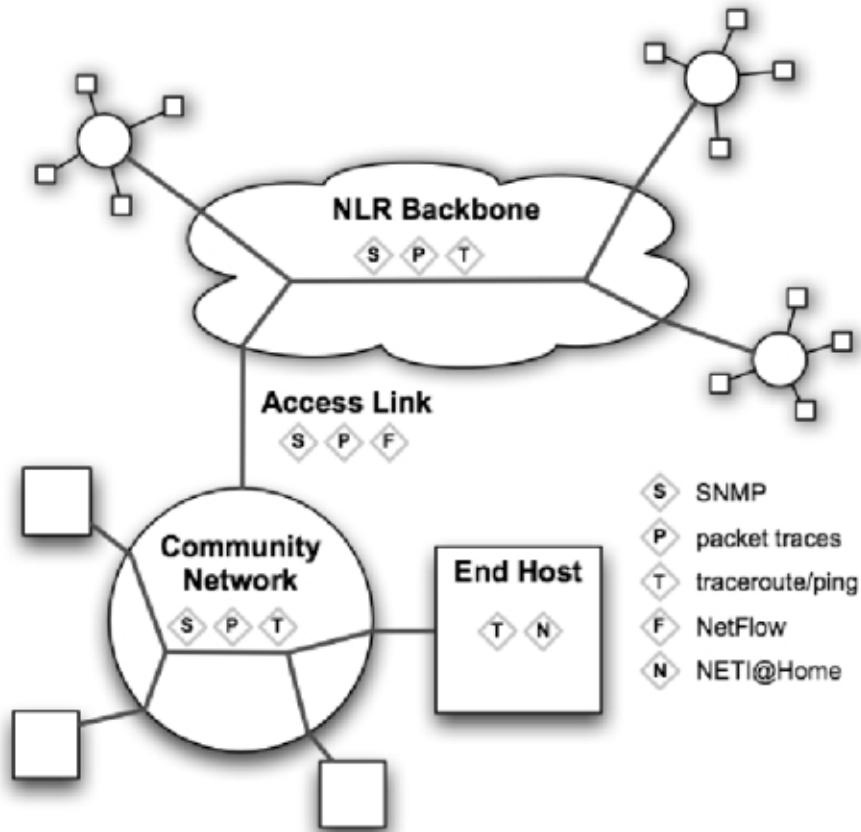


Figure 4: four measurement levels incorporated into the COMMONS Project.

7.1 Backbone measurements

The COMMONS backbone will support measurements that provide an indication of overall backbone health and utilization. As an optical fabric, the backbone could support direct measurement of wavelengths [12], or provide SNMP counters as supported in the attached equipment. It could also support a matrix of pairwise latency measurements between POPs. We initially propose to collect a simple set of measurements: the closest we can get to a traffic ingress/egress matrix using existing hardware, an interPOP delay matrix. In the future, the COMMONS Project may want to invest additional development resources in optical network measurement hardware, and will need to establish vetting procedures and clear disclosure procedures for adding any operational measurement functionality to the backbone.

7.2 Access link measurements

Access links connect the peering router of the attached community network with peering routers of the backbone. Example measurements on access links include: (1) passive measurement: packet traces on attached links; (2) flow measurement, e.g., netflow/ipfix output; (3) SNMP counters, e.g. utilization, errors; (4) traceroute/ping measurements to a probeset; and, (5) inter-POP delay.

7.3 Community/local network measurements

Within an attached community/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 some subset of: (1) anonymized HTTP web proxy logs; (2) anonymized DNS statistics, aggregated via e.g., dsc; (3) traceroute server logs; (4) large-scale traceroute probes; (5) support for other active measurement experiments as approved by COMMONS policy review board; (6) BGP peering with, for example, COMMONS-routeviews; (7) anonymized or aggregated web cache logs for analysis; and, (8) would make tools available to community network users who individually volunteer to participate in research.

7.4 End host measurements

Finally, the COMMONS Project 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 [13] model, taking 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. Receiving data from unknown, untrusted users presents particularly daunting, though not insurmountable, challenges thus far ignored by many existing projects [14, 15].

For active measurements, the biggest challenges are: deployment to insure low impact on the infrastructure; prevention of use of tools for DDOS attacks; accountability of measurement source in case of operational problems; analysis of bias due to self-selection of sources (by volunteers); and validation of the integrity of resulting data. Client-side passive measurement infrastructures have all the same problems as active ones but, in addition, they bear formidable privacy challenges. For certain measurement questions, such as provider or application prevalence on the macroscopic Internet, there is substantial material incentive to manipulate a macroscopic Internet measurement system. Thus, both participant and researcher trust in the integrity of the measurement is essential.

Researchers do not yet have an in-depth understanding of the methodological problems of scaling Internet measurement paradigms to incorporate peer production. We propose 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 participants expect that the next few years will serve as a transition period whereby researchers can determine if we can achieve the same integrity from client-side infrastructures that we now can from controlled infrastructures.

8 Meta-infrastructure to support protected data sharing

Because COMMONS will initially be composed of cooperative IP networks, commercial counter-incentives to rapid and wide sharing of information will not frame inter-network security and communication policies. If launched, COMMONS will establish a security response team (whose operation will be reviewed by the policy board) to aid participating networks with integrating security best-practices into their network operations. The cooperative structure of COMMONS will also allow research into areas that have proved impossible in the unregulated commercial framework, e.g., testing of secure naming and routing protocols, coordinating public and private sector response to and recovery

from major Internet disruption; development of fundamentally new architectural components that might emerge from NSF GENI program; implementation of state-of-the-art wireless techniques and policies to promote efficient utilization of spectrum; economic modeling of a wide variety of end-to-end paths, with particular aim toward transparent techniques for analysis of capital and operating cost accounting.

Respect for user privacy is essential to CAIDA's mission and the goals of the COMMONS Project. COMMONS participants are acutely aware of the sensitivities involved in sharing of Internet measurements, and CAIDA already participates in projects such as PREDICT [16] and leads DatCat [17], both framed by these sensitivities. The launching of both of these projects makes this a perfect time to establish the COMMONS initiative to leverage these pre-existing projects and build an in-vivo 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. CAIDA has extensive collaborative experience with many facets of both 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; while, 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.

8.1 PREDICT: A legal framework to support protected data sharing

In 2004, the US Department of Homeland Security (DHS) recognized the need to support the calibration of cyber security tools (particularly those funded by government agencies) in real world environments. After extensive consultation with privacy law experts, DHS is now in the process of launching the PREDICT Project [16] to allow researchers to request real-world datasets to assist their research. The goal of PREDICT is to facilitate the development of Internet defense technologies, products, models and strategies. DHS has made noteworthy progress in the legal and privacy aspects of infrastructure data access, specifically addressing the concerns of ISPs who want to support the research community but are constrained by privacy concerns, laws, and/or policies. The presence of such a framework bodes well for the viability of COMMONS and demonstrates a flexible and participatory path to achieve both protection of privacy and support for empirical network science.

8.2 DatCat: Internet measurement data catalog

For the past several years, CAIDA has been developing a data cataloging system for use by the Internet research community. Despite its necessity to scientific endeavor, data available to researchers is severely 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 [17] to: (1) provide a searchable index of available data; (2) enhance documentation of datasets via a public annotation system; and, (3) advance network science by promoting reproducible research. Like PREDICT, DatCat is a critical supporting infrastructure to the COMMONS Project and the synergies between the projects are numerous.

9 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 who own the infrastructure under study are declaring bankruptcy, measurements are scarce. And when the companies who own the infrastructure are competing against each other, whatever measurements do exist are often considered extremely sensitive or completely proprietary. And yet, changing technologies, commercial strategies, and regulatory policies have brought dramatic restructuring of Internet service delivery at the local, regional, national, and global levels. Accompanying these changes are a variety of strong but conflicting (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 policy makers. 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 – often being placed in the position of only having access to the information that the companies who would be affected by policy and regulatory changes are willing to share.

In the US, both telecommunications companies and user advocates are unhappy with the current state of communications policy. Prices for services are higher than many other industrialized nations, and broadband penetration is lower. Telecommunications companies are having increasing trouble attracting investment and claim their broadband services need exemption from common carrier regulation in order to thrive. Solutions tend to focus on industry-centric approaches to policy reform (e.g., 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, which entities should be forced to pay for universal service for rural areas, among many other questions). The problem is that the dearth of Internet research makes it impossible to come up with reliable empirical answers to many of the questions regulators and politicians need answers to. Thus, national telecommunications policy is forced to advance blindly at a time where the US is losing its competitive (broadband) edge when compared with a growing number of industrialized countries. The COMMONS Project provides a collaborative environment for policy-makers to help shape the research questions under study and offers a vital resource for regulators seeking to make decisions based upon empirical scientific research.

10 Next steps

As the workshop ended, participants formed working groups to continue discussions of most immediate tactical value: (1) possible organizational structures, and (2) economics and business models. Future working groups will likely include: infrastructure; operational security; measurement and research; and privacy.

Building on the momentum from the workshop, outreach to community and municipal networks (particularly wireless initiatives) has increased dramatically. In addition, CAIDA staff are developing a draft Roadmap (currently scheduled to be completed in June 2007), identifying community networks that will be involved in the first phase of the COMMONS Project, their measurement/data-collection capabilities, and the support resources needed to help maximize the utility of these measurements.

11 Conclusions

The COMMONS Project 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 Project facilitates both basic research and innovative improvements to the Internet. COMMONS presents an unprecedented opportunity for establishing standards for scientific integrity for in-vivo Internet research using rigorous empirical data to validate theories, models and simulations. Acutely relevant to NSF's current strategic interests, the measurement data gathered via the COMMONS project will lend unique strategic significance to NSF'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 Project also provides an opportunity for opening up the "economics, ownership, and trust layers" of Inter-networking 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, COMMONS provides a much needed resource to help us chart the future of communications.

ACKNOWLEDGMENTS. The COMMONS Strategy Workshop was supported by NSF grant proposal #0714690, Cisco Systems, CAIDA members, and the San Diego Supercomputer Center. We especially thank all the workshop participants for their insights and support. More information on the COMMONS Project is available online at: www.caida.org/projects/commons

12. References

- [1] Ed Markey (2007). Speech to the Consumer Federation of America on Broadband and Internet Policy. February 1, 2007. Accessed February 3, 2007 and available online at: http://markey.house.gov/index.php?option=com_content&task=view&id=2577&Itemid=141
- [2] S. Derek Turner. Broadband Reality Check II. Free Press, September 2006.
- [3] "Broadband Deployment is Extensive throughout the United States, but it is Difficult to Assess the Extent of Deployment Gaps in Rural Areas," Government Accountability Office, Report to Congressional Committees, GAO-06-426, May 2006.
- [4] National Research Council. "Network Science." 2006. National Academies Press.
- [5] "Federal Plan for Cyber Security and Information Assurance Research and Development." Report by the Interagency Working Group on Cyber Security and Information Assurance. April, 2006. National Science and Technology Council.
- [6] Bob Wallace (2006). "NTCA: FCC Commissioner Calls for Sweeping Regulatory Reform -- Candor from Copps Counts with Rural Telco Crowd." Telecommunications Online. February 7, 2006. Accessed February 3, 2007 and available online at: http://www.telecommagazine.com/newsglobe/article.asp?HH_ID=AR_1706
- [7] Sascha Meinrath (2006). "The Ed Whitacre (AT&T CEO) Network Neutrality Flip-Flop." March 23, 2006. Accessed February 3, 2007 and available online at: http://www.saschameinrath.com/2006mar23the_ed_whitacre_at_t_ceo_network_neutrality_flip_flip_0
- [8] Tim Wu and Laurence Lessig (2003). "Regarding Ex Parte Submission in CS Docket No. 02-52." August 23, 2003. Accessed February 3, 2007 and available online at: http://www.freepress.net/docs/wu_lessig_fcc.pdf
- [9] R. Jonas Geissler, Whether 'Anti-Spam' Laws Violate The First Amendment, 2001 Journal of Online Law art 8.
- [10] Muniwireless LLC. The 2006 municipal wireless state of the market report.
- [11] "Internet connections will hit 10Gbit by 2010 in Japan," 2004/8/29 <http://techcrazy.theppn.org/archives/2004/08/29/internet-connections-will-hit-10gbit-by-2010-in-japan/>
- [12] Jorg Micheel -- Lambda_mon - a passive monitoring facility for dwdm optical networks. In Constantinos Dovrolis, editor, PAM, volume 3431 of Lecture Notes in Computer Science, pages 228–235. Springer, 2005.
- [13] Yochai Benkler. Coase's penguin, or Linux and the nature of the firm.
- [14] Yuval Shavitt and Eran Shir. DIMES: Let the Internet Measure Itself. Technical report, Tel-Aviv University EE Department, 2005.
- [15] [Neti@home: http://www.neti.gatech.edu/](http://www.neti.gatech.edu/)
- [16] Department of Homeland Security. PREDICT project: Protected Repository for Defense of Infrastructure against Cyber Threats. <http://www.predict.org/>.
- [17] Datcat: Internet measurement data catalog. <http://www.datcat.org/>.

Appendix: COMMONS Strategy Workshop Mission Statement, 13 December 2006

Through discussion, debate, and compromise, COMMONS Strategy Workshop participants consensed unanimously on the following shared Statement of Purpose:

COMMONS is...

... a community of interest comprised of local networks, broadly defined, who seek to:

- utilize a AUP (Acceptable Use Policy) neutral communications infrastructure,
- mitigate geographic disparities in the cost of Internet access and wide-area network connectivity,
- enable competitive environment in the connectivity market,
- benefit from economies of scale, and
- realize a vision for ownership in an open, universal and scalable backbone infrastructure.

... a community of network operators who:

- recognize that the future usefulness and security of the Internet depends upon the availability of empirical network data,
- support the availability of that empirical data to the academic research community, and
- insist that the data collected and utilized be handled in a manner respectful of personal privacy.

... a community of interest of people who seek to exchange ideas on, and devote resources to, a wide variety of common interests in furtherance of the objectives stated above, including:

- approaches to social problems that leverage innovative technology and business strategies,
- political, legal, and regulatory strategies, and
- community networking.

By peering interested broadband networks via a national fiber infrastructure, COMMONS will build a cooperatively-run national backbone that interconnects participating community and municipal networks both to one-another and to the global Internet. Peering would be conditionally available to city, county, state, and federal government entities, academic institutions, community wireless initiatives, commercial entities, and allied organizations. In exchange for transit, attached networks will make measurement data available to Internet researchers, policy analysts, and technology developers following appropriate legal data-sharing frameworks and privacy safeguards. COMMONS provides an unprecedented opportunity for establishing standards for scientific integrity for in-vivo Internet research and provides rigorous empirical data to validate theories, models and simulations. Acutely relevant to NSF's current strategic interests, the measurement data gathered via the COMMONS project will provide unique strategic significance to NSF's GENI program and 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.