Measuring Internet Performance when Broadband is the New PSTN

Bill Lehr[1]
Steve Bauer[2]
David Clark[3]

Massachusetts Institute of Technology[4]

** PRELIMINARY DRAFT – PLS DO NOT CITE W/O AUTHOR APPROVAL **

Abstract

As we consider the transition from the telephony PSTN to broadband as the "new PSTN," an important aspect of the transition will be the ways in which we measure and assess service performance. This presents a number of important challenges because (a) broadband is inherently more complex (in basic technology, market structure, and the range of services and use/user needs it addresses); (b) there is a desire to rely increasingly on markets instead of direct government control (public utility regulation) to manage the ecosystem (investment, pricing, and service choice – which includes technologies, business models, and market-based regulatory frameworks). For both reasons, the measurement metrics play a more critical and important role. Internet measurements and performance data are relevant for communications policy (identifying market failures, crafting and implementing appropriate remedies, and assessing the efficacy of such remedies), for traffic management (operational control of networks), and as critical "market intel" (efficient markets require information about the quality of the goods and services that are being transacted to operate). In this paper, we address some of the challenges and opportunities posed by the need to develop new platforms for Internet performance measurement.

Specifically, we believe broadband measurement must go beyond simple measurements like speed toward a richer conceptualization of the quality of experience. This will have important implications for the design of metrics and measurement infrastructures since the user's experience is impacted by both component (link) level and system (end-to-end) level factors that will engage cross-layer, cross-value chain dialog. Another area we believe will be worthy of special focus will be how we address the challenge of assessing and providing reliability for the broadband Internet.

[1] Corresponding author: wlehr@mit.edu, 617-258-0630

[2] bauer@mit.edu

[3] ddc@csail.mit.edu

[4] The authors would like to acknowledge support from NSF Awards 1040020, 1040023, the MIT Communications Futures Program (http://cfp.mit.edu), and the MITAS Project (http://mitas.csail.mit.edu). All opinions expressed herein are those of the authors alone.

1. Introduction

The Internet used to be an application overlay on the wires of the Public Switched Telecommunications Network (PSTN). From its origins as a government-funded research network in the 1960s to its emergence as the first mass-market platform for data communications in the 1990s to the conversion to broadband access after 2000, the Internet has evolved into *the* preferred infrastructure platform for all electronic communications worldwide. Today, telephony is just another of the many services that are supported on the Internet. The largest providers of broadband Internet access services are the incumbent telephone and cable television providers that own the wired and wireless last-mile networks that were originally built to deliver legacy fixed and mobile telephony and cable television services. Just as in the past when Plain Old Telephone Service (POTS) was viewed as essential infrastructure for society and the economy – analogous to other basic infrastructures such as electric power, water, and roads -- today, broadband Internet access is now regarded as essential infrastructure. As such, there is an enduring public policy interest in ensuring universal access to affordable broadband service. The broadband Internet is in the process of becoming the "new PSTN."

To understand what it means for the broadband Internet to become the new PSTN, it is important to understand how both the Internet and the PSTN have evolved. Our focus will be on service performance measurement, which constitutes an integral component of any policy framework for PSTN management. At the highest level, we will explain how the increased complexity, heterogeneity, and dynamism of the broadband Internet induces increased reliance on market-based metrics, which are themselves more complex, heterogeneous, and dynamic as a result. Compared to measuring and managing the performance of POTS, measuring and managing the performance of the broadband Internet is inherently more difficult and will engage a wider range of stakeholders and market participants.

In subsequent sections, we trace the impact of three distinct evolutionary processes that underlie the transition to the broadband Internet as the new PSTN and what this implies at the 50-thousand-foot level for the ecosystem for Internet performance metrics. We then focus more specifically on the performance metric-relevant issues that impact FCC policy in its efforts to promote universal service, ensure a healthy market for interconnection, and the reliability of our critical PSTN infrastructure to highlight some of the challenges. Next, we review the current status of Internet performance measurement, and conclude with our thoughts on the policy challenges that loom and how these might be addressed in the near and longer term.

2. The broadband Internet as the new PSTN

The Public Switched Telephony Network (PSTN) is both a network and a regulatory construct. There are important performance metric issues associated with both. As a network, the technology, industry, and market structures have changed substantively as we have evolved from telephony to the broadband Internet. As a regulatory construct, we recognize that the PSTN is essential infrastructure for society and the economy, and as such there is an enduring public interest in ensuring universal and affordable access. Ensuring the provision of essential PSTN services is a key focus of communications regulatory policy, and over time, has given rise to a huge bulwark of legacy rules, standards, and regulatory apparatus. In transitioning to broadband as the new PSTN, we need to confront what of this legacy apparatus needs to be mapped into the new world of a broadband PSTN and how best to accomplish that goal.

To best understand some of what is involved, it is worthwhile considering the three distinct and simultaneous evolutionary processes that underlie the transition to the broadband Internet as the "new PSTN":

- (1) the transition from a telephone network to a broadband Internet platform
- (2) the evolution of the Internet toward becoming a cloud computing utility
- (3) the transition from public utility regulation to market-based regulation

2.1. From telephone network to broadband Internet

The legacy PSTN was a telephone network purpose-built to provide universal telephony access, and to serve as a wired platform for end-to-end leased lines for enterprise data networking. It was essentially designed to support a single-service: end-to-end circuit-switched voice-grade telephony (with a 4KHz dynamic range that was encoded into a 64Kbps digital signal).[5] It was engineered end-to-end as a "silo-network" and originally owned and managed end-to-end as the AT&T Bell System.

In contrast, the Internet was a small set of protocols (the narrow waist of the hourglass) that supported best-effort packet-switched data communications between "peer" hosts, where the "peers" might be individual nodes or entire networks. It was originally provided as a networking overlay on top of the PSTN infrastructure, using the same leased line infrastructure and dial-up telephony access services as supported enterprise data networking. The Internet provided "best-effort" service that was unsuitable for carrying delay-sensitive telephony traffic until the capacity and capabilities of the Internet improved, a limitation that is no longer applicable.

The Internet has been augmented with lots of new functionality and greatly expanded capacity so that the Internet is now used widely as *the* platform for all electronic communication services from multimedia video (over-the-top television, interactive gaming) to telephony (VoIP and videoconferencing) to data (which, ultimately, includes all kinds of traffic). Data traffic is inherently more heterogeneous in its quality-of-service or performance requirements (i.e., intrinsic data rate, delay tolerance, bit error rate requirements, etc.).

[5] Over time, with the addition of enhanced calling features, voice-mail, and numerous other advancements, the range of telephony services expanded. Similarly, with the introduction of virtual private networking, frame-relay, and ATM services, the range of business data services expanded beyond leased lines.

From a PSTN perspective, figuring out where to draw the line between what constitutes the "Internet" is complex and contentious (is it just the narrow waist? Or, what of the complementary network intelligence and functionality that exists in today's broadband Internet should be included?). We do not even have a clear definition of what constitutes "broadband service" so have no equivalent to the "voice-grade circuit" definitions that were accessible in legacy telephone regulation.

2.2. Evolution of Internet into a Cloud Computing Utility

The Internet is evolving into a cloud computing utility, expanding beyond its original focus on providing end-to-end packet transport, to include general support for on-demand access to distributed computing resources. A growing amount of functionality and intelligence (state) is moving into the Internet to complement, and at times, substitute for intelligence and functionality in the end-hosts.

The Internet is evolving into a platform for cloud services (software-as-a-service, data centers, computing-on-demand, content-delivery networks, etcetera). Figuring out what of this "cloud" of resources should be deemed to be part of the Internet, and even more importantly, to be part of the "new broadband PSTN" versus what are "overlays" or uses of the PSTN is far from clear.[6] Moreover, ISPs do lots of active network (traffic) management using techniques like DPI[7], MPLS[8], variable bitrate encoding[9], and a host of other techniques that impact the user experience. This makes it harder to specify

[6] Companies like Microsoft, Google, and Amazon provide key infrastructure components for Internet-based cloud services, but are not typically regarded as ISPs. There are also a number of content delivery services like Akamai that offer functionality that is also provide, at times, by ISPs. What portion of these businesses operations (if any) should be included as part of the "new PSTN"? See Clark et al. (2006) for a discussion of the role of overlays in the Internet.

[7] Deep packet inspection (DPI) is used to identify what applications are in use, and such information may be used for real-time resource allocation to ensure that certain applications (e.g., peer-to-peer file sharing) do not crowd-out access for other applications like VoIP. Of course, DPI might also be used to implement price discrimination strategies or for (potentially) privacy-invading market research. As with other techniques, we note that the welfare implications of DPI and other network management techniques are topics worthy of discussion, without offering any opinions pro or con here.

[8] MultiProtocol Label Switching (MPLS) is widely used by ISPs and enterprise VPNs to support traffic prioritization and manage traffic flows.

[9] Variable bitrate encoding can allow application performance to be adaptive to variable network conditions. A mobile video service provider might down-code a video stream (e.g., from HD to NTSC) without noticeable degradation in the video experience for a viewer looking at the video on a small mobile handset screen, but might have to increase the data channel allocation or buffer the stream if the user is actually viewing the video on a large HD screen (e.g., using AppleTV to stream from a user's smartphone to their widescreen TV at home). Alternatively, YouTube takes advantage of the higher initial data rate offered by many cable modem services (i.e., "Powerboost") to pre-load local buffers to support a better user viewing experience in the event of subsequent congestion.

performance metrics since there are so many more things that one is concerned about; and, in particular, makes it less desirable to rely on input-based regulations that overly constrain network management decisions. We would like to preserve the freedom over implementation options to take advantage of context-specific technical advances.[10]

2.3. From Public Utility Regulation to Markets

At the same time that we were transitioning from circuit-switched telephony to the broadband Internet platform, we have been moving from legacy public utility regulation toward market-based regulation. This is natural in the sense that public utility regulation is a telephony legacy, and an absence of regulatory oversight is the Internet legacy.[11] Thus as we move from the world of a telephone PSTN, it is easy to move toward an unregulated PSTN based on the Internet. However, this misses the point of what it means for the Internet to become the new PSTN, and by so doing, assume the public interest obligations associated with its new role as basic infrastructure. To understand better how to reconcile the need for greater regulatory oversight[12] with the changing regulatory dynamics of regulating the PSTN in today's world, it is useful to summarize some of the high-level differences between the legacy PSTN and a new broadband PSTN.

From a metrics/measurement perspective, the challenge changes from one of the regulator collecting the information and data needed to implement public utility regulation (e.g., through rate-base and other regulatory proceedings) to ensuring the existence of a rich market ecosystem for market-relevant/performance-relevant information. In this new environment, the regulators position is more specialized and

---

[10] Input-based regulation includes rules that specifically identify the technology to be employed, as for example, in the definition of what comprises a voice-grade circuit in POTS telephony. Analogously, one might define the Internet service solely by reference to the IP protocols. That sort of minimalist view has been espoused by some Internet purists, who regard everything else (above and below the narrow-waist of the Internet as not "Internet." Unfortunately this approach does not really address the challenge of how to regulate the new broadband PSTN since it leaves out much of the functionality that is critical to supporting today's Internet. However, attempting to specify how one should implement such other technology (e.g., network management, content-delivery services, policy-based routing, etcetera) would overly constrain innovation and the ability to adopt context-dependent solutions. An alternative is to move toward output-based regulation (e.g., "broadband Internet service should be capable of supporting voice telephony").

[11] The Telecommunications Act of 1996 does not mention the Internet, leading some critics to argue that it was outdated on the date of its passage. Others have argued that leaving the Internet out, helped protect the Internet from premature regulation.

[12] An important point of this paper is that regulatory oversight does not imply any particular form of regulatory intervention. In the absence of any market failures or an enduring public interest, there really is no need for regulatory oversight. With an enduring public interest, there is a need for regulatory oversight, but there may still not be a need for active intervention (e.g., there may not be a market failure). We do not want to argue here about whether there are market failures (e.g., a potential last-mile bottleneck issue), but believe that the potential that such a conclusion is justified either today or as a valid risk for the future is sufficiently clear that policy needs to stand ready to address this issue.

limited: the regulator has a unique ability to compel disclosure of information that might otherwise not be available to the market and can have a strong impact on metrics by its decisions of how metrics may be instantiated in regulations (e.g., minimum performance standards) or reflected in government reports.[13]

At the highest level, the "new PSTN" will be significantly more heterogeneous, dynamic, and complex than the legacy PSTN, and control will be more decentralized and distributed (see Table 1).

Table 1: A new broadband PSTN is very different from legacy

New broadband PSTN is more:	Because:
Heterogeneous	More classes of users (sophisticated and mass market), uses (with divergent QoS requirements), and contexts (wider range of technologies supported than in legacy PSTN).
Dynamic	Resource sharing is even more essential and needs to be flexible to respond to market conditions over varying time scales (real-time, provisioning, investment). This is because capacity usage of the platform is inherently less predictable than in days of silo-based telephony (more bursty traffic); because of need to customize functionality to address heterogeneous needs (data traffic is inherently more heterogeneous than telephony); and to accommodate the more rapid pace of technological/market change and inherent uncertainty.
Complex	Because broadband Internet is a platform for many services, composable out of a more diverse array of components across the industry value chain, protocol stack, and general economy. Because of all of the reasons above.
Decentralized	Internet is inherently less centralized and hierarchical than traditional telephony networks; boundary between edge and core is more ambiguous (with implications for assignment of control over functionality); and transition to competition in components has broken the end-to-end, silo-structure of legacy telephone networks and the resulting industry structure it spawned.
Distributed	Ownership of key assets is distributed across a wider range of stakeholders from edge-network operators (end-user equipment), applications, content providers, content delivery networks, Web and cloud services providers, and ISPs. The range of asset-owners with

[13] Government reporting may be taken as authoritative by consumers, and as such, the choice of metrics and how they are reported and explained may influence market equilibria. Consumers may naturally look to the government to tell them which data is important for help in interpreting the data. Government data collection and reporting behavior also may be taken as a signal of future potential regulatory action. For example, the credibility of the regulator's threat to take enforcement action will depend, in part, on whether the regulator is believed to have the information needed for effective enforcement.

	heterogeneous regulatory classifications and market-positions is greater.

For the past 50 years, we have been moving from heavy-handed public utility regulation of telecommunication markets toward increased reliance on market-based competition to ensure socially desirable outcomes. The viability of efficient competition presumes the absence of market failures (a tautology); and when competition is viable, a principal justification for regulation disappears. In communication policy circles, the drive from public utility regulation to market-based competition has been motivated by the belief that technical innovations and market growth have increased the viability of competition,[14] while at the same time, the dynamism and increased complexity of telecommunications markets have increased the costs of command-and-control style regulatory oversight.

A key challenge of Public Utility Regulation is asymmetric information: the regulated firm knows more about the technology and market conditions than the regulator does. Moreover the firm is typically better able to respond quickly to changing circumstances. The inherent information asymmetry and time lags in regulatory processes pose a risk of inefficiency that is magnified as the environment becomes more dynamic, complex, and uncertain. One way to address this challenge is to allow firms more discretion in how they run their businesses. As telecommunications markets have become more complex and dynamic, we have seen a succession of communications policy regulatory reforms that have had the effect of increasing the scope for firm discretion and toward increased reliance on market-based competition to ensure socially desirable outcomes. Examples include the shift from rate of return to price cap regulation, the progressive deregulation

[14] Innovation that drives market growth and falling costs can turn what was a natural monopoly into a market capable of sustaining viable competition. It is also possible that a new technology might have the opposite effect. In markets for last-mile competition, it was long believed that wired networks were a natural monopoly, but then the rise of the Internet, wireless, and a host of other technologies (e.g., broadband-over-powerline) made it look as if intermodal competition between multiple last-mile facilities-based providers might be feasible. Now, with the consolidation among wireless and wired last-mile providers, and the move to much higher data rate and capital intensive (per subscriber) broadband services, there is a growing concern that we are heading back to (maybe never left) a world in which last-mile access remains a bottleneck. Of course, and at least since the 1990s, analysts have continuously disagreed on the viability of current and future competition in last-mile services. We are not interested in engaging that debate here, but note that recognizing broadband as "basic infrastructure," *ceteris paribus*, increases the need for broadband regulation (i.e., there is an enduring public interest in ensuring efficient provision of broadband services). But, recognition of the increased need for broadband regulation does not imply that the choice of regulation is obvious. First, if one concludes either that current (or prospective) competition in broadband services is sufficiently effective or that all of the regulatory options would be worse than any competitive problems, one might reasonably conclude that the best regulation is to do nothing. However, a watchful but non-intrusive regulator is still valuable since it is very possible that current (and prospective) conditions are influenced by the potential threat of regulatory intervention or that conditions might change so as to render regulatory intervention desirable.

of segments of the end-to-end telephone network,[15] and the drive toward "technical neutrality" in regulation.[16]

3. The need for new performance metrics

As the regulatory role shifts from the regulator to the market, the need for publicly available performance metrics increases.[17] Markets need information to function efficiently and in a distributed/decentralized market, there are a larger number of stakeholders who need a greater range of performance-relevant information for efficient collective decision-making.[18] Moreover, when the markets are changing rapidly, the

---

[15] Telephone equipment markets were opened to competition in the 1960s, long-distance telephone services in the 1980s, and local telephone services in the 1990s. In a succession of decisions from 2002 to 2005, the FCC classified broadband services as "information services," thereby exempting broadband from Title II regulation of basic telecommunication services and the attendant common carriage regulatory burdens.

[16] Technically neutral regulation aspires to be unbiased with respect to the choice of technology by the regulated firm, allowing firms discretion in their choices of technology, with the competitive market serving as the arbiter of which technologies succeed. Cable/telephone convergence and the rise of intermodal competition helped drive reforms to eliminate asymmetric regulation of telephone services, although much remains to be done. In spectrum policy, PCS licensing in the U.S. was "technologically neutral" in allowing 2G mobile service licensees to choose the technology, whereas in Europe and many other markets, regulators mandated adoption of GSM. Although a commonly espoused aspiration, true technical neutrality is rarely, if ever, achieved in practice.

[17] Contrast today's environment with the legacy world of regulating the AT&T Bell System of black-phone to black-phone telephony. In that earlier world, it was sufficient for the regulator to know virtually everything about AT&T but only enough of that data needed to be shared more widely to allow third-party auditing of regulatory efficiency. Consumers and businesses did not need to know about the relative performance of alternative telephone services or customer premises equipment in the absence of competition. The new broadband PSTN is a composite system of multiple components, and policymakers need to be (implicitly) concerned about the potential for competition not just in the end-to-end systems but also in the component markets (where the components include end-user equipment, access network technologies, core networks, cloud services, etcetera).

[18] The narrow-waist architecture of the Internet decentralizes and distributes decision-making, which implies a greater number of interfaces across which (potentially) market-based transactions may occur with (potentially) different "buyers" and "sellers" transacting different "goods." Each of these markets may require different information or have a different perspective on how to interpret information. Note that while more public information is needed and of more types, that does not mean that everyone needs to know the same things. There just needs to be "enough" public information among enough decision-makers to allow efficient decision-making. Figuring out precisely what that means in practice will be context dependent and difficult. As the PSTN becomes more mission-critical for the economy and more capable of collecting granular data (e.g., about per-user behaviors as a consequence of the expansion of location-aware services and embedded sensors), there will be valid reasons to limit access to data to protect privacy and security (e.g., avoid making critical network assets easier targets for terrorist attack).

metrics need to evolve and adapt so their development will also need to shift more toward the market.[19]

In the legacy PSTN, we naturally looked to the regulator as the principal source of public information about performance. In a world of legacy public utility regulation the principal justifications for public disclosure of performance data are as documentary support of regulatory decision-making and to allow third-party analysts to audit the regulator's control.[20] Reliance on market forces to induce efficient decisions implies the existence of choice and informed choice about what services or devices to purchase or network architecture to invest in depends on the existence of trustworthy performance data.[21] In markets, the sources of performance data will necessarily multiply and embracing the multiplicity of sources is one of the ways that markets ensure efficiency.[22] The regulatory authorities should not and cannot reasonably hope to be the sole authoritative source of performance data. A wider array of market participants will and should be expected to contribute to the ecosystem for performance metrics and data.[23]

In this richer environment, regulators like the FCC will continue to have a very important role in the metrics and performance measurement space because of their special regulatory power which can influence market outcomes both directly (through regulations and subsidies) and indirectly (by addressing information gaps and asymmetries and by the threat of future interventions).

In the balance of this section, we will focus on three examples of performance-relevant measurement issues that the FCC will need to concern itself with as we transition to the new broadband PSTN:

---

[19] The same bureaucracy that renders regulatory decision-making slow, will impact the regulator's ability to evolve metrics and reporting practices in response to changing market circumstances.

[20] The watcher needs watching. For example, to protect against capture of the regulator on the one hand, or inefficiency or exploitation by the regulator on the other.

[21] There are obviously degrees of "trust" one may have in performance data which depends on many things, including the provenance of the data (who is providing it and what are their incentives), the knowledge of the consumer (how capable are they of interpreting the data), and the purpose for which the data is to be used (the degree of reliability needed depends on the decision to be made with the data).

[22] Fundamentally, markets reflect the decentralization of decision-making authority which allows markets to be better at solving complex information problems (e.g., contrast Soviet-style central planning with western-style capitalism).

[23] The market participants will include end-users (running monitoring agents on their home and enterprise networks who may blog or otherwise publicize their data), third-party market analysts (from the investment and market research community, from consumer and industry advocacy groups, and from academia administering surveys and engaging in primary data collection efforts for non-profit, paid and advertising-supported dissemination), industry providers (ISPs, content/ application providers, and technology vendors releasing information as required by regulation or voluntarily to strategic partners or for marketing purposes), *and* policymakers (in the FCC, PUCs, FTC, and Homeland Security seeking to inform the public and support their regulatory efforts).

- Achieving Universal Service Goals
- Regulating Internet Interconnection
- Ensuring a reliable PSTN

Each of these challenges represents a proto-typical policy-based rationale for why the FCC needs to develop new performance metrics for the new broadband PSTN, but we have ordered them sequentially in order of increasing difficulty. Whereas establishing performance metrics to define quality standards for basic broadband access services (a requisite for implementing universal service goals) seems relatively straightforward even if remaining contentious, identifying the right metrics and data reporting strategies for monitoring interconnection markets and for ensuring system reliability are more complicated and speculative at this point.

3.1. Achieving Universal Service Goals

A key role of government is to ensure universal access to essential infrastructure (like roads, water, electric power, and now, the broadband Internet). Whether necessary or not, a substantial component of legacy PSTN regulation is associated with the collection and distribution of universal service subsidies, most of which has been focused on ensuring affordable access to fixed-line telephony services.

The U.S. National Broadband Plan appropriately recognized the necessity to migrate and reform existing universal service programs from telephony to broadband support mechanisms. The FCC is currently in the midst of transitioning programs like the high cost funds (in excess of $4.5 Billion per year) into the broadband-focused Connect America Fund.[24]

To effectively target collection and distribution of such funding, the FCC has to have a definition of broadband and what consumers are under-served by broadband, or equivalently, do not meet some minimum standard of broadband performance. Of course, there may be multiple purpose-specific "definitions" that may be applicable, just as in the early days of broadband, the FCC defined broadband (*for data collection purposes*) as offering a data rate in excess of 200Kbps.[25] We might need to have separate definitions

---

[24] See Report and Order and Further Notice of Proposed Rulemaking, In the matter of Connect America Fund *et al.* (WC Docket No. 10-90 *et al.*), Before the Federal Communications Commission, Released November 18, 2011, available at: http://transition.fcc.gov/Daily_Releases/Daily_Business/2012/db0206/FCC-11-161A1.pdf).

[25] This definition was long criticized as reflecting too low a goal for broadband speed, but such criticism was often misguided. The FCC specified the 200Kbps standard for its data collection efforts, it did not designate that as the goal for what broadband service should be nor required it as a minimum performance standard. The 200Kbps standard had the virtue (in the late 1990s) of excluding from broadband data collection ISDN, satellite-based broadband, and mobile broadband services, but still captured most DSL and cable modem offerings that comprised the bulk of offered data services at the time. Of course, to analysts, more granular data is always more desirable to enable the tracking of quality differences across regions and time; but the need for granularity has to be balanced against data collection costs and confidentiality considerations.

for fixed broadband and mobile broadband performance since the two services offer distinct value propositions. For example, mobile broadband is typically personalized (associated with a single user) and provides anywhere access but usually with lower speed and availability[26] than fixed broadband service, which is provided to a household (and therefore supports multiple users).[27]

The focus to date has been on characterizing broadband data services based on the peak data rate, which although positively correlated with other metrics of interest, offers a very incomplete picture of service performance. As we have discussed elsewhere and further below, even figuring out what the right speed metric should be and how to report it is far from trivial.

While the need to have a definition for broadband to implement universal service programs provides a sufficient justification for such a definition, adoption of a definition could have much broader market implications. The FCC might choose to instantiate such a definition in minimum performance standards, if it is successful in establishing its authority to regulate broadband.[28] However, it remains an open policy question as to whether it would be a good idea for the FCC to adopt an explicit definition of what constitutes broadband service, or leave it to the market to define what constitutes broadband. The latter approach amounts to an implicit definition since presumably the FCC will be compelled to intervene if the market fails to support an adequate level of service. Moreover, what constitutes an adequate level of service will involve more than minimum performance standards, but will concern the distribution of service offerings (choice among providers, service tiers, pricing, and variation across communities). Furthermore, increased performance variation may not be a bad thing: for example, serving everyone with at least a minimum level of service and most people with much higher quality service (e.g., fiber-to-the-home) may be much better than simply ensuring everyone has a minimum level of service. Indeed, the ability to achieve peak data rates that far exceed average rates is a key design feature of the Internet.

Regardless of whether the FCC adopts an explicit or implicit approach to defining what constitutes appropriate performance for broadband service, the FCC will need to monitor

---

By analogy, one might imagine the FCC adopting a universal-service-program-specific definition for broadband to enable funding decisions (figuring out which broadband proposals are eligible for funding in particular programs).

[26] Mobile service is available in more locations than fixed service, but it likely to be less available at any particular location than where fixed service is provided. Which offers higher "availability" depends on one's definition of availability.

[27] With upgrading of mobile broadband to 4G LTE and the expansion of tethering services, the customer experience of performance may be narrowing.

[28] In a series of decisions from 2002 to 2005, the FCC reclassified broadband as an information service, thereby insulating it from the application of the Title II regulatory authority applied to basic telecommunications services. The FCC relied on its authority under Title I to regulate broadband, but the Court has rejected that claim which the FCC's legal authority to regulate broadband in question.

that that level of performance is being delivered over time. Again, a sufficient rationale for that need is to monitor that universal service funds are being used appropriately, while an additional rationale might include enforcement of performance standard regulations (if implemented).

The FCC will need to craft policies specifying how the data will be collected, managed, and retained. Presumptively, a key source of such data will be ISPs, as well, as external measurement sources the FCC may avail itself of. Consideration will need to be given to the costs of data collection, management and reporting, and how those costs are recovered.

3.2. Regulating Internet Interconnection

Ensuring end-to-end connectivity in the telephony PSTN over the history of deregulation and the emergence of new types of service providers (mobile, long-distance, VoIP) gave rise to a complex and often conflicting array of interconnection regulations. These imposed obligations to interconnect, and often specified the prices for terminating traffic. Collectively, these are referred to as carrier interconnection and intercarrier compensation regulations.

In contrast, interconnection in the Internet has been unregulated. ISPs negotiate bilateral interconnection agreements that historically were bifurcated into revenue-neutral peering agreements and transit agreements. Under the latter, the transit service providing ISP was compensated for assuming the obligation to deliver traffic to its destination.

As with universal service, the National Broadband Plan recognized the need to reform PSTN interconnection regulations in order to eliminate the accumulated inefficiencies of the legacy regime and to transition to the broadband future.

Precisely what the FCC should do to regulate Internet interconnection is highly contentious with many advocating that the FCC do nothing, keeping Internet interconnection unregulated. We tentatively agree that it would be ill-advised for the FCC to actively intervene in Internet interconnection markets today, but recognize that in keeping with the Internet becoming the new PSTN, that the FCC has an enduring interest in ensuring an open and efficient market for Internet interconnection. We have explained elsewhere how FCC attempts to regulate network management practices as part of its framework for ensuring an open Internet (sometimes referred to as Network Neutrality) necessarily engages concerns over interconnection policy.[29]

With respect to performance measurement, the FCC needs to monitor the health and activity of interconnection markets to enable it to identify if problems arise and to provide an empirical basis for crafting remedies if market failures appear evident. Today, there is little publicly available information other than anecdotal about the scope and

[29] See Clark, Lehr, Bauer (2009) and Clark, Lehr, Bauer (2011) on interconnection policy.

nature of ISP interconnection agreements. To adequately assess the health of interconnection markets, policymakers will need data about actual traffic flows as well as about the terms and conditions in Interconnection agreements. We have noted elsewhere how markets for Internet interconnection have become more complex and dynamic with a wider range of interconnection agreements than prevailed in the earlier world of simple peering and transit agreements.[30] As the recent flap in Europe over the French regulator's request for disclosure of interconnection agreement information demonstrates, figuring out how best for regulators to empirically monitor Internet interconnection markets remains a difficult question.[31]

3.3. Ensuring a reliable PSTN

When the Internet was a nice-to-have application offering best-effort service, users readily tolerated outages and variable service quality. The Internet was allowing users to do things they could either not do otherwise or accomplish only at greater expense (e.g., VoIP as opposed to traditional switched telephony). As the Internet becomes essential, mission-critical infrastructure, the need to ensure the reliability of the broadband PSTN becomes much more important.

The need for performance metrics to assess reliability arises both in the context of monitoring the reliability of mass market broadband services, as well as ensuring the reliability of the overall PSTN. We might think of the first challenge as akin to monitoring the performance of retail banking and credit markets, while the latter is akin to monitoring the performance of the banking system (e.g., avoiding events like the financial sector meltdown of 2008). Implementing a framework for monitoring the reliability performance of the Internet cloud will be complex and will involve new types of metrics and processes to ensure systemic reliability.[32]

4. Current Performance Metrics Practices

In this section, we consider the current status of broadband performance measurement, with a focus on evaluating the performance of mass-market broadband access services. While this is important, it is worth noting that for end-users, what matters ultimately is end-to-end performance. While the access connection is a key element in determining the quality of end-to-end service and a worthy focus for regulatory interest in its own right, there are multiple other elements (both on the home and server side) that are not directly controllable by the broadband access provider that can and do have a significant impact

---

[30] See Faratin *et al.* (2007) and Clark, Lehr, Bauer (2011).

[31] See, for example, http://www.afnic.fr/en/about-afnic/news/general-news/5711/show/afnic-publishes-its-reply-to-the-public-consultation-by-arcep.html.

[32] See, Lehr (2012, forthcoming) on Internet cloud reliability.

on the quality of the user experience. How these components interact and complement each other varies by type of application.[33]

There currently exist two main classes of broadband performance data. These two classes are distinguished by the type of data gathered, the methods in which the data is gathered, and the entities involved in the process. For the purposes of this paper, we designate each class by the main entities involved: 1) Data collected by broadband network operators (e.g. Verizon, Comcast, Time Warner Cable) and 2) Data collected by all other entities (e.g. end-users, application/content providers, academics, and third-party consultants[34]). Many of these non-ISP sources of data rely on a mix of edge-based (often web-based) and in-network server-based measurements. To the extent the third-party providers are able to co-locate their measurement infrastructure in provider networks and ISPs may collaborate with third-parties in collecting measurements, the distinction between service provider and third-party data sources will blur over time.

In recognition of the need for better metrics and coordination among a larger set of stakeholders, the FCC recently launched the Next-Generation Measurement Architecture Standardization and Outreach Group (NMASOG) to "facilitate the standardization and broad deployment of a comprehensive Internet broadband performance measurement architecture, protocols, metrics, and data structures."[35] The goal of this group is to provide "requirements and input to standardization bodies, and describe how the output of these external organizations can be best used to create a reliable, accurate and sustainable broad performance measurement infrastructure." See Figure 1 for a diagram that envisions some of the main components and measurement and control protocols of this architecture.

---

[33] For example, with cacheable data like streaming video that is not real-time (e.g., YouTube), the server and user-side application can cooperate to take advantage of variable network bandwidth to deliver a better user experience (smoother delivery of video) by buffering content locally. Technologies like ECN/re-ECN that would enable explicit congestion notification could allow applications and network components to inform each other before congestion results in dropped packets, and with the additional information, cooperate on delivering better end-to-end performance. While technically feasible, there are a number of business (incentive) challenges that need to be overcome before such capability is widely available. For a discussion of ECN/re-ECN, see http://www.bobbriscoe.net/projects/refb/ or http://www.caida.org/workshops/isma/1102/slides/aims1102_sbauer.pdf.

[34] The third-party consultants include market researchers, investment analysts, and traffic analysis technology providers like Sandvine (a provider of deep packet inspection technology).

[35] A mission statement was drafted in May 2012, see http://tinyurl.com/6qbxq4t.

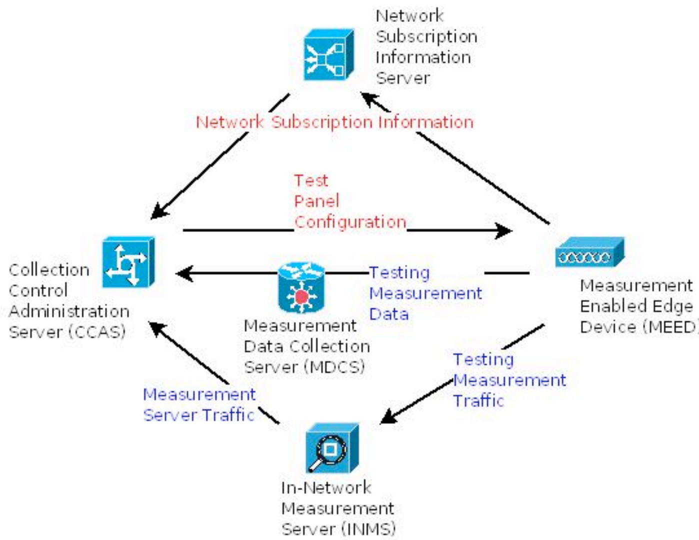

Figure 1: Next-Generation Measurement Architecture

The key take away from this diagram is that multiple entities and devices are involved in the testing/measurement and data collection process. They imagine a future in which a "large number, if not all, consumer edge devices, such as cable or DSL modems, will be equipped with measurement capabilities 'out of the box'. At any given time, only a small fraction of such *measurement-enabled edge devices* (MEED) will actively participate in performance measurements, or may only participate in some portion of such measurements." The network providers are a key component as well as they supply a service that authoritatively determines "the connection parameters of the broadband connection, such as the service provider (ISP), geographic location, network technology, service class, rate caps, burst capability and provisioned connection speed."

Until such an integrated architectural approach comes to fruition, however, the classification of relevant broadband data as originating from provider or edge-based measurements will remain informative. We summarize some of the main measurement and data collection activities in each of these two categories in the following two subsections.

4.1. Data collected by broadband network operators

This operational community has the most direct (and at times exclusive) access/visibility into some data that has relevance to the regulators. In particular we focus on two types of data that providers are in exclusive possession of today: 1) reliability/outage data; and 2) usage data.

4.1.1. Reliability/Outage Data

The FCC currently requires PSTN providers to electronically report information about significant disruptions or outages that meet specified thresholds set forth in Part 4 of the

FCC's rules (47 C.F.R. Part 4).[36] In the context of telephone calling (an easily defined service), it is relatively easy to identify and unambiguously specify metrics such as call-completion percentages (i.e., a simple binary characterization of the state of the system can record whether a call attempt succeeded in establishing a voice-grade connection). This stands in marked contrast to what outage and reliability reporting might mean for broadband networks.

We discussed the challenge of measuring broadband reliability in an earlier paper that we submitted to the FCC's *Notice of Inquiry* that was launched in April 2011 to gather data and stakeholder perceptions on Internet reliability.[37] We recap some of the relevant points here. There is wide variation in what one might reasonably mean by reliability. A common metric for reliability is "availability" which is the amount of time that a system is expected to be in-service (or free of "failures").[38] It is often expressed as a statistical time measure (e.g., Mean Time To Failure, MTTF) or the percent of time over some period that the system is available for service. In addition to MTTF, reliability engineers are also interested in the Mean Time To Restore (MTTR), which when coupled with information about the relative costs of outages, recovery, and failure avoidance options provides the basic elements for a cost-benefit analysis of good reliability management.

Applying these concepts, in our own previous work, we proposed three distinct types of reliability metrics that we expect to be useful for assessing broadband services. These metrics were:

Reliability of performance: By "reliability of performance," we mean that the quality-of-service provided by broadband meets or exceeds some target level of performance over some specified timeframe. There are multiple measures that one might be concerned with, including the speed (data rate), latency, jitter, or bit error rate. One might be interested in some composite (weighted index) of multiple of such measures. Defining meaningful target performance levels and time frames is an ongoing challenge. One strategy might be to develop profiles that are appropriate for different classes of application or usage (e.g., "reliability of streaming video" or "reliability of VoIP," or more generically "reliability of broadband suitable to support usage of typical broadband subscriber" – where defining what constitutes a typical subscriber is, itself, a difficult challenge).

---

[36] See Network Outage and Report System (NORS) http://transition.fcc.gov/pshs/services/cip/nors/nors.html

[37] See Lehr *et al.* (2011) for a discussion of broadband reliability.

[38] What constitutes a "failure" is a key component of metric design. Does the event result in a disruption of service that is noticeable to the subscriber? Does the disruption in service cause more than negligible harm (cost)? Is the event measurable? What are probabilities of Type I *and* II errors? And so on.

Reliability of connectivity: An even more basic notion of broadband service reliability focuses on Internet connectivity.[39] Any particular server may be down for either a brief or long time. The failure of a single server or even subset of servers need not constitute an outage of broadband service. There is some arbitrariness to the definition of what constitutes a sufficient loss of connectivity to be classified as an outage. Exploring different metrics (and implementations) is part of the challenge.

Reliability of core services: A third notion of reliability is the availability of certain "core" Internet services. The potential list of target services or functionalities, and the metrics for assessing whether they are operating properly, is long. Several of the obvious candidates include services like VoIP, DNS, email, and the Web. Each of these services may fail in multiple ways (e.g., no connectivity, diminished performance, localized or system-wide loss of availability, etcetera).

The recent announcement[40] by the FCC requiring outage reporting for interconnected VoIP services is an example of compelling reporting in the "core services" category above. Over time we expect that operators' obligations to report data across all of the categories will grow.

For providers, gathering and reporting this information will be a non-trivial task both from an operational and technical perspective. The main reason for this is that neither the deployed systems nor the operational procedures have been designed with these reporting tasks in mind. A network operator we spoke with said, "I want to fix problems not document them." At a fundamental level the Internet has been designed with the same approach; failures are routed around not documented. This is not to imply that outage reporting is impossible, rather the intent is to provide additional intuition why extracting this information from a broadband infrastructure is challenging.

4.1.2. Usage Data

The Internet is often compared to the network of highways, streets, and roads that make up the transportation system. For transportation networks, it is generally recognized that traffic data (i.e. the volume of traffic, congestion information, incident reports, etc.) is as important to understanding the state of the network as is information about where the roads or links actually are. The same is true for the Internet.

In transportation networks, traffic data is valuable over both short time scales (e.g., allowing real time traffic management to re-route commuters around a rush hour

---

[39] Sundaresan et al. (2011) propose defining “availability as the faction of the time that home users have IP connectivity to their access ISP.” This definition may suggest that connectivity is limited to packets traversing the access link, but that would be a rather limited notion. For many, connectivity would imply being able to get packets to/from off-net locations, but determining what fraction of which targets need to be unavailable and for how long to qualify as an outage is complex.

[40] See http://www.fcc.gov/document/fcc-requires-outage-reporting-interconnected-voip-improve-911

accident) and over longer time scales (e.g., for planning maintenance cycles and capacity expansion investments). During periods of congestion,[41] traffic data and real-time traffic management via lights, tolls, and special commuter lanes has proved especially important in enabling more efficient utilization of the existing transportation infrastructure. Improving the efficiency of the existing infrastructure delivers benefits in the form of reduced commute times (contributing directly to labor productivity), improved safety, and reduced pollutant emissions through intelligent traffic management policies.

On the Internet, traffic data is similarly important to network operations. Over short time scales ranging from less than a second to hours or days, traffic data is an input into systems (both automated and human-centered) that make routing decisions (e.g., balancing loads across different network links), identify suspected or actual security or transmission failures, and implement traffic management policies.[42] Over longer time scales measuring months or years, traffic data is vital to capacity planning and provisioning, allowing capacity to be efficiently installed in advance of demand, thereby better accommodating future traffic growth without congestion-related disruptions. Thus traffic data is essential to almost all the practical dimensions of network management and to the political, regulatory, and theoretical questions of what constitutes good, acceptable, or socially desirable network management.

While traffic conditions on the highway and roadways can be observed externally (via both technical sensors and human observations), information about Internet traffic and the level of congestion of the different autonomous networks which collectively compose the Internet is limited. While individual network operators generally have a good idea about the state of their own networks, outside stakeholders have little visibility into the state of traffic on networks. Networks can be probed and tested by outside observers to derive some measurements, but the scope and confidence of such measurements is limited compared to the accuracy and breadth of information available to network operators.

The majority of users have very little visibility or understanding of what is happening to their traffic once it enters a network. Without better visibility, it is not surprising that there are widely diverging opinions about the true state of networks. For example, what are the congestion and utilization levels now and in the predictable future? Or, what are the effects of different traffic management policies?[43]

The problem is that this limited visibility by outside stakeholders into the traffic and congestion state of networks makes it hard to have confidence in the regulatory and investment decisions that affect such networks. On the one hand, traffic management policies that are efficient and 'fair' could be disrupted or private investments in

---

[41] For a discussion of congestion in the Internet, see our companion paper, Bauer, Clark, Lehr (2009).

[42] See Subramanian (1999) for an introduction to network management.

[43] Over 10,000 comments were filed in the FCC's proceeding (07-52) regarding Comcast's management of peer-to-peer traffic (see http://fjallfoss.fcc.gov/prod/ecfs/comsrch_v2.cgi).

expanding capacity could be deterred. On the other hand, network operators could be exploiting their control to thwart or discourage disruptive new innovations and competitors (either intentionally or accidentally).

The amount of information that *could* be collected by network operators from their networks is enormous. Individual network elements, which include routers, switches, servers, caches, and subscriber modems, can report hundreds of different statistics. With hundreds or thousands of elements in a network, and millions of subscriber lines, the volume of potential data is enormous. For example, one network operator we spoke with indicated that the total volume of data records could exceed 300 terabytes of data a year. Collecting and transporting the raw data in real-time to the network operations center where it can be processed, analyzed, and managed presents a difficult challenge that incurs significant operational costs. Determining what data to archive and how to compress/summarize the data and manage access presents complex statistical, logistical, and policy challenges.

In spite of the costs, network operators do systematically collect real-time traffic data because it is essential for successful network operation. The data is an input into strategic and operational decision-making across virtually all ISP functions. The data informs decisions about the capacity of internal links, routing policies, security policies, and interconnection contracting. It is used for high availability and disaster recovery planning, for financial projections, employee evaluations, technical strategy discussions, and sales and marketing. In larger network operations, there are specialized departments focused on managing the collection and analysis of network traffic data, and the sharing of relevant portions and views of the data across the organization.

4.2. Data collecting by all other entities

We now turn to broadband performance data collected by the wide variety of other entities in the Internet ecosystem. These include end-users, content providers, academics, and others. To date, the primary focus of broadband measurement by all of these other entities has been testing and reporting the "speeds" achieved by broadband services. The one we summarize below are a subset of those available but are some of the most widely cited and discussed.

Taken as a whole, we expect these sorts of sources to play a more important role in the future measurement ecosystem. No single source is adequate. There needs to be an active market in performance measurement that allows the multiple sources of data to compete and complement each other.

4.2.1. Samknows

Both Ofcom in the UK and the FCC in the United States run large-scale measurements of broadband networks utilizing dedicated measurement boxes and infrastructure provided by SamKnows.[44] In the FCC version of the Samknows study, over 7,000 broadband subscribers run specially built access points which conduct numerous tests every hour of every day.[45] The current list of tests is provided in the following table. The data from this set of tests is publically released by the FCC and has been utilized in academic research projects, the advertising campaigns of competing broadband providers, and as input to other regulatory proceedings.

Metric	Primary measure(s)
Web browsing	The total time taken to fetch a page and all of its resources from a popular website
Video streaming	The initial time to buffer, the number of buffer under-runs and the total time for buffer delays
Voice over IP	Upstream packet loss, downstream packet loss, upstream jitter, downstream jitter, round trip latency
Download speed	Throughput in Megabits per second
Upload Speed	Throughput in Megabits per second
UDP latency	Average round trip time of a series of randomly transmitted UDP packets
UDP packet loss	Average round trip time of a series of randomly transmitted UDP packets
Consumption	Volume of data downloaded and uploaded by the panelist
Availability	The total time the connection was deemed unavailable
DNS resolution	The time taken for the ISP's recursive DNS resolver to return an A record for a popular website domain name
ICMP latency	The round trip time of five regularly spaced and schedule ICMP packets
ICMP packet loss	The percentage of packets lost in the ICMP latency test
FTP throughput	Throughput in Megabits per second at which a file can be downloaded from or uploaded to an FTP server
Peer-to-peer	Throughput in Megabits per second at which a file can be downloaded

Table 1: Tests running on Samknows test boxes

In the U.S., the data collected by the SamKnows infrastructure has already resulted in several published reports. These reports helped inform market stakeholders that the broadband speeds delivered by the biggest ISPs were higher than some had thought, suggesting that earlier claims that these ISPs were under-delivering on their speed promises were overblown.[46]

In contrast with web-based or other testing methods, the SamKnows approach provides valuable insight by allowing measurements that isolate the broadband access connection and provide a platform for implementing a flexible array of tests. However, as noted already, while a useful addition to the performance measurement toolbox, it is not the only tool needed to assess performance by the market or policymakers. For example, the

---

[44] See http://www.samknows.com/broadband/index.php.

[45] See http://www.samknows.com/broadband/uploads/Methodology_White_Paper_20111206.pdf.

[46] For example, on the basis of data from March 2011, the FCC concluded that "for most participating broadband providers, actual download speeds are substantially closer to advertised speeds than was found in data from early 2009" (see http://transition.fcc.gov/cgb/measuringbroadbandreport/Measuring_U.S._-_Main_Report_Full.pdf).

SamKnows infrastructure has yet to be applied to the networks of broadband resellers (e.g., companies like Speakeasy.net that offer broadband services using facilities leased from facilities-based providers) or mobile broadband provider networks. And, even with 1000s of boxes deployed, this testing infrastructure only represents a tiny share of total subscriber lines. While adequate for addressing certain questions, it is unlikely to be suitable for documenting types of performance variations that end-users might care about (e.g., which ISP offers the best performance in a specific location?). Nevertheless, we believe that efforts like the SamKnows project adopted by Ofcom in the UK, the FCC in the US, and currently underway in the EC,[47] reflect praiseworthy attempts by policymakers to enhance the collective knowledgebase of performance metrics.

4.2.2. Measurement Lab:

Measurement Lab (M-Lab) is a distributed measurement platform run by consortium of academic and industry players that includes Google, the New America Foundation, the PlanetLab Consortium, and others. The goal of M-Lab is to provide an open platform to enable researchers to easily deploy and share Internet measurement tools. They seek to "advance network research and empower the public with useful information about their broadband connections."[48] They currently host nine different broadband measurement tools running on a global infrastructure of measurement nodes. The tools and their measurement purpose are listed in the following table.

Measurement Tool	Purpose
NDT (Network Diagnostic Tool)	Test your connection speed and receive sophisticated diagnosis of problems limiting speed.
Glasnost	Test whether certain applications or traffic are being blocked or throttled on your broadband connection.
NPAD (Network Path & Application Diagnostics)	Diagnose common problems that impact last-mile broadband networks.
Pathload2	See how much bandwidth your connection provides
ShaperProbe	Determine whether an ISP is performing traffic shaping.
BISmark	Apply to host a router device to test Internet connectivity over time.
WindRider	Detect whether your mobile broadband provider is performing application or service specific differentiation.
SideStream	Collects statistics about the TCP connections used by the measurement tools running on the M-Lab platform.
Neubot	Performs periodic tests to measure network performance and application-specific traffic throttling.

Table 2: Measurement Lab Testing Tools

The M-Lab initiative is noteworthy in several respects. First, M-Lab is committed to being an open platform to help ensure transparency and support the free exchange of ideas about new and existing metrics. Given the complexity and dynamism of the

[47] In October 2011, the EU announced plans to deploy a 10,000 node SamKnows testing infrastructure for Europe (see http://www.samknows.com/broadband/news/commission-launches-eu-wide-broadband-performance-study-11053.html).

[48] See http://www.measurementlab.net/.

broadband Internet and the difficulty (and undesirability) of prematurely attempting to standardize on a minimal subset of metrics, it is very important that methods be fully and transparently documented so analysts can be clear about what the data *is* as a necessary prerequisite for discussing what any particular metric might *mean*. Second, M-Lab is committed to providing an evolvable platform on which researchers from academia and industry can collaborate to develop new metrics and reporting tools. A key challenge of performance measurement is how to summarize the huge amounts of potential data to gain useful (intelligible, discriminating) insights into phenomena of interest. Third, M-Lab is committed to making the data and tools widely available so that over time there will be a large compendium of test data which can be mined by researchers seeking to understand emerging trends as a well as transient phenomena. The M-Lab sponsors are looking to expand the platform globally.

4.2.3. Akamai and Sandvine Reports

Akamai is a company that provides content and application delivery services over the Internet. They operate a content delivery network (CDN) that consists of tens of thousands of servers in many of the largest 1000 networks on the Internet. Their delivery network serves content for a wide selection of web sites and thus observe web traffic from a wide selection of end users.[49] For both billing and operational reasons they keep fairly extensive logs of clients that connect to their servers to download content. They have analyzed these logs and produced a "State of the Internet" report quarterly since the first quarter of 2008.[50] In addition to reporting on connection speeds, they also present data on attack traffic, Internet penetration and broadband adoption, and general trends seen in the data.

Sandvine is a company that provides traffic analysis technology, including deep packet inspection technology. Like Akamai, in the course of their business, they gain unique insight into Internet traffic patterns. In recent years, they have been publishing a series of reports on Internet traffic trends and emerging insights that are widely reviewed by analysts.[51]

There are also a host of market research firms and investment analysts that collect and publish data, either in for-fee research reports or as part of their marketing efforts.

Taken together, the SamKnows, M-Lab, and Akamai/Sandvine/Other performance metrics efforts illustrate important trends in the space. All of three offer distinct and complementary benefits that contribute all-important competition for ideas in the emerging market for broadband performance metrics. We expect policymakers will need to take advantage of all such sources to appropriately address their future needs for

---

[49] Selected customer list is available at: http://www.akamai.com/html/customers/customer_list.html

[50] For example, see http://www.akamai.com/stateoftheinternet/.

[51] See http://www.sandvine.com/news/global_broadband_trends.asp.

empirical research in support of their policy mandate to regulate the new broadband PSTN.

5. Policy Recommendations

In approaching the transition to the new broadband PSTN, the FCC will confront numerous challenges. In the following sub-sections we provide some high-level guidance on how the FCC should address its performance measurement responsibilities.

5.1. Recognize that metrics serve multiple purposes

Performance measurements serve multiple purposes, and understanding the purpose for a metric will help in its interpretation, and may also help in managing the total costs of performance measurement. First, the ISPs and other firms engaged in offering Internet functionality [52] and end-customers operating enterprise networks will engage in significant performance measurements to facilitate real-time network management (e.g., load balancing, fault monitoring, and SLA verification tasks), network provisioning, and longer-term, strategic planning and investment. There will be significant duplication and partial overlapping in such measurement efforts, and their will be strategic considerations that will have a strong impact on impeding incentives to share, but if shared appropriately, such information would prove very useful to all market stakeholders and policymakers. Such sharing could help in filling gaps (no one sees everything), cross-validating data,[53] and in sharing the costs of performance management.[54]

What constitutes "appropriate sharing" will depend on the context. For example, ISPs might be willing to share more detailed information if the disclosure of such information is protected, if the sharing is voluntary, or if they are compensated (at least partially) for the costs of data collection. The desire to preserve confidentiality may be reasonably motivated by the desire to protect the privacy of subscribers, to protect competitors (or regulators) from having too detailed a picture of the ISPs actual operations, or for security reasons; but too limited public disclosure limits the value of the information to the market. For example, in earlier efforts to document the progress of broadband availability, the FCC obscured the number of ISPs offering service in a zip code if the number of providers was less than three – severely diminishing the value of the reported

[52] Hereafter, we will use "ISPs" broadly to refer to ISPs, content delivery providers, other cloud resource providers, and technology providers whose business is in the provision and operation of Internet-related services.

[53] The opportunity to cross independent data sets of similar phenomena makes it easier to identify anomalies, which might include network-specific effects, data errors, or mis-representations. The latter may be helpful for performance verification in contract enforcement. To enforce a contract it is not sufficient that the contracting parties know a mistake has occurred, it may be necessary to verify the mistake by a third-party enforcer.

[54] Having more than one entity measure the same phenomena may be desirable, but some sharing of measurement responsibilities could economize on total measurement costs.

data for third-party analysis of broadband deployment progress. When it comes to data on traffic, we do not need public disclosure of per-user behavior (application usage, MB traffic over time), but we do need reliable information to characterize the statistical distribution of per-subscriber behavior over time.[55]

Other important motivations for performance measurement include the need for users to be able to make informed market-based decisions. Buyers need to be able to evaluate service offerings to choose options that offer the right mix of quality and price for their needs, and for market-discipline to work through the collective actions of buyers and sellers. For such decision-making, highly granular data is usually not necessary. It might be sufficient for buyers of broadband to understand that a particular class of service is suitable for on-line gaming whereas another class of service is more appropriate to light-users, and to be able to understand which providers' have options in which markets to meet those needs. In addition, however, broadband consumers may want more real-time performance monitoring of their access connections to diagnose home networking problems. Sampling of such data, if publicly shared, could help support third-party monitoring of the reliability of service provider offerings. Designing such a sampling regime is complex both from a statistical and business/policy perspective. An appropriate sampling methodology needs to anticipate the uses of the information (e.g., if the data is to be used to document specific reliability failures then it is more sensitive than if it is to be used as one of many inputs into some broad measure of performance "quality"). To be representative and to keep data overhead costs manageable,[56] relatively coarse sampling is likely to be desired (not all users and not continuously for any user).

A final important rationale for collecting performance measurements will be to facilitate policymaking, which includes implementing policies already undertaken (e.g., administering universal service mechanisms), diagnosing potential market failures in anticipation of future actions (e.g., monitoring the Interconnection markets), and helping protect against market failures (e.g., threats to systemic reliability by auditing overall system performance).

In addressing this last need, which is the principal focus of this paper, the FCC should leverage its knowledge of the evolving ecosystem of measurement in crafting its metric-specific interventions. The market-based need by ISPs and the needs of end-users (both mass market and enterprise customers) to engage in significant performance

[55] Our MITAS project (http://mitas.csail.mit.edu) used a unique sample of per-subscriber data provided voluntarily by ISPs to help characterize the distribution in per-subscriber behavior, while providing a mechanism for respecting valid data confidentiality concerns. A reason why we need to know the distribution in per-subscriber data is to help tease apart the contributions to growth of average versus peak users, which is desirable to better craft solutions (technical, pricing, and policy) to optimally address changing traffic patterns.

[56] It will be important to manage costs both in aggregate (the potential for BigData-explosion is large) and individually will be important. Minimizing the individual costs of performance measurement is important to ensure end-users have an incentive to participate. The success of efforts like SETI@home demonstrate the potential of generating significant levels of participation with no direct compensation being required.

measurement means there will be an expanding source of market-based data available. Some of this data will be provided by ISPs and end-users to the market voluntarily (through industry consortia and marketing materials) and some without the providers express consent (third party analysts and from overlay measurement platforms). The FCC's need for such data are likely to be less detailed and granular than the market participants, who in many cases, will be in a better position to collect, retain, and manage such data.

The availability of such data sources will reduce the need for the FCC to compel reporting but it will not eliminate the need for the FCC to collect, monitor, and interpret data (for the general public). The FCC's role may become more curatorial (selecting among public sources of data to complement whatever data the FCC collects directly, and offering commentary on the appropriate interpretation of the data). And, the FCC should stand ready to compel data collection from ISPs and other industry participants if adequate data is not either provided voluntarily or otherwise obtainable by the market.

The FCC may encourage and support incentives by end-users and ISPs to collect and manage the requisite data by targeted actions. These may include blessing industry collaborations like BITAG and standardization efforts, engaging in dialogs with initiatives like M-Lab, and by promoting access to open-source measurement platforms and clients.

5.2. No 'one size fits all' solution is desirable

In an earlier paper, we discussed why there is no single broadband speed metric that is right for all circumstances,[57] and in any richer characterization of broadband performance, speed measurement ranks as a relatively-easy problem. While we do not think it would be appropriate to designate a single metric, we also recognize the need to enable comparability and avoid confusion. Too many metrics that are mutually inconsistent would be as problematic as no metrics. Furthermore, to make interpretation more tractable, there will need to be greater recourse to index metrics (which weight and consolidate component metrics into a summary statistic). The problem with index metrics is that the choice of weights is very important and is likely to vary by context.

The FCC should continue to resist attempts to over-simplify measurement and data reporting to make it easier for end-users to interpret or consume. The FCC should err on the side of complexity in the interests of preserving information. If the measurement components are reported, than analysts can compose their own indices based on their own choice of weights. The only valid rationales for summarizing data are to control costs or protect confidentiality, but such justifications need to be carefully considered to avoid abuse.

[57] See Bauer et al. (2011).

In its role as curator, the FCC should provide interpretations of particular metrics it may choose to embrace (from third parties) or originate in-house. The FCC can point to and promote efforts by industry standards bodies to define best practices for performance measurement and reporting.[58]

5.3. Support transparency and disclosure of methods

While it might not be always desirable to transparently disclose all data publicly, it is desirable and generally feasible to disclose the methods used to collect data and implement reported metrics. Performance measurements which do not provide detailed discussions of the methods have, at best, only limited value.

Ideally, the code used to generate the measurements would be open-source. However, this may be too tough a requirement in the case of privately-provided code. Among the sources of information that we expect to be of growing importance in the future will be for-profit market-research enterprises which compete, in part, on the quality of their proprietary code for collecting, analyzing, and presenting data. It would be unreasonable to expect such providers to voluntarily publicly disclose all of their code, and such disclosure may not be necessary to verify their measurement methods. Confidential disclosure or other techniques may provide a sufficient basis for verifying the trustworthiness of certain measurement metrics.

Because of its special role and authority with respect to performance monitoring and reporting, the FCC should be specially careful to rely, wherever possible, on open-source code and transparent methods. Furthermore, the processes used by the FCC to select among methods and metrics should be open and transparent. Because the metrics will need to evolve with the markets, and because metrics from many sources will need to be considered, ensuring an open and flexible process for evolving metrics will be increasingly important. This is another reason why FCC should be more willing to acknowledge that data errors will occur and need to corrected. It will not have the sole responsibility for managing performance monitoring and should not assume excessive responsibility. This will place a greater burden on the market, but that is the price of giving market forces greater scope for discretion.

5.4. Shift to ex post enforcement

Finally, we believe that the increased complexity and dynamism and shift of responsibility from command and control management to reliance on market forces will

[58] Standards bodies define standards for specific metrics as well as recommendations for best-practices and processes for developing and predicting standards. For example, the IEEE 1413 standard provides standards for reliability predictions that are appropriate to the different contexts in which reliability may need to be assessed (e.g., before a design is formalized, when only rough estimates are possible, as well as later when more detailed and empirically-justified estimates are possible). See Elerath and Pecht (2012).

make it harder for the FCC to craft or rely on ex ante rules to manage market outcomes. An example of this is the shift in Network Neutrality regulation to rely on case-by-case analyses to identify market problems and craft explicit remedies, rather than to attempt to specify ex ante (specific) limits on behavior. From a performance metrics perspective, this might be construed to imply that the FCC should not be able to collect data on performance metrics unless it is first able to demonstrate a reasonable expectation that there is a problem it needs to remedy. However, we would disagree with such a position.

As noted earlier, the FCC needs to monitor a range of activities not just to implement positive interventions that it is already engaged in (such as universal service), but also to identify prospective problems that may require intervention in the future. This gives the FCC a broad mandate for monitoring market performance, with wide scope for potentially compelling targeted information, but with the proviso that it use that authority sparingly. Wherever possible, the FCC should leverage market-based (third-party or voluntary) data for assessing performance. To do that effectively, however, the FCC may need to deploy some in-house infrastructure like the SamKnows infrastructure or conduct FCC-sponsored primary research to facilitate its verification and interpretation of data from other sources, and to fill such gaps as exist.

6. Conclusions

In this paper, we focus on the performance metrics challenge that will confront communications policymakers in the FCC and similar regulatory authorities around the globe as we transition to broadband as the new PSTN. In becoming essential infrastructure, the broadband Internet becomes unavoidably the focus for government regulation. It is no longer merely an application that relies on the PSTN, but is part of the core fabric of the PSTN. As such, policymakers have a public interest mandate to ensure universal access to this basic infrastructure, like they do with respect to access to clean water, safe roads, and reliable electric power.

The question is not whether we need to regulate the Internet, but how we should regulate it. The transition from telephony to broadband does not eliminate the policy goals that motivated legacy regulation of the telephone networks, but it does beg the question as to what portions of that legacy regulation need to be re-mapped into the broadband Internet world. To the extent market-based competition is sufficiently viable or market-failures that existed with respect to legacy telephone services no longer apply, we may be able to simply eliminate classes of regulation (e.g., do we need "must-carry" broadcast rules in a world of over-the-top video on the Internet?). However, the continuation of programs like universal service and legitimate concerns over last-mile bottlenecks or imperfections in interconnection markets imply that there will be continuing role for communications policy regulatory enforcement. Of course, in keeping with the secular trend of the last fifty years, the expectation is that market-forces will continue to play a key (if not dominant) role in enforcing the desired policy outcomes (allocative, productive and dynamic efficiency).

For all of these reasons, we will need performance metrics to formulate, target and enforce effective communication policies. The challenge for meeting this needs is complicated by the growing complexity of broadband (as compared to legacy telephony) and the more dynamic market and regulatory environment in which the PSTN exists. In this new environment, we need a robust market for metrics. Policymakers will need to rely on a diverse and evolving set of metrics from multiple sources – from efforts initiated and controlled by the regulators directly, from data provided by network operators and others with a vested interest in collecting granular data (not just for policy decision-making, but also for real-time operational management), and increasingly, from other third-party sources.

7. References

Bauer, Steve, David Clark, and William Lehr (2010), "Understanding Broadband Speed Measurements," MITAS Working Paper, June 2010, available at: http://mitas.csail.mit.edu/papers/Bauer_Clark_Lehr_Broadband_Speed_Measurements.pdf

Clark, D., W. Lehr, and S. Bauer (2011) "Interconnection in the Internet: the policy challenge," 39th Research Conference on Communications, Information and Internet Policy (www.tprcweb.com), Alexandria, VA, September 2011. (pdf=http://people.csail.mit.edu/wlehr/Lehr-Papers_files/clark%20lehr%20bauer%20TRPC2011%20%20Interconnection%20RC2.pdf)

Elerath, J. G. and M. Pecht. 2012. "Ieee 1413: A Standard for Reliability Predictions." Reliability, IEEE Transactions on, 61(1), 125-29.

Faratin, P., D. Clark, S. Bauer, P. Gilmore, A. Berger and W. Lehr (2008), "The Growing Complexity of Internet Interconnection," Communications & Strategies, No. 72, 4th Quarter 2008, pages 1-21.

Lehr, W., S. Bauer, M. Heikkinen, and D. Clark (2011) "Assessing broadband reliability: Measurement and policy challenges," 39th Research Conference on Communications, Information and Internet Policy (www.tprcweb.com), Alexandria, VA, September 201