MEASURING PERFORMANCE WHEN BROADBAND IS THE NEW PSTN
BY WILLIAM LEHR, STEVEN BAUER, AND DAVID D. CLARK

The transition of telecommunications services from the old Public Switched Telephone Network (PSTN) to the broadband Internet does not alleviate the needs of consumers for basic services. The Internet universe, however, is mostly unregulated. As broadband becomes the new PSTN, how do we ensure that services are delivered appropriately, and how can this be measured? Doctors Lehr, Bauer, and Clark propose an updated set of measurements and address the policy considerations they raise, in order to ensure that broadband services indeed serve the public interest.

INTRODUCTION

The Internet used to be an application overlay on the wires of the Public Switched Telephone 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 telephone 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 – 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.”
There are many considerations that will arise as the broadband Internet takes on the status of critical infrastructure. Ensuring the provision of essential PSTN services was 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 consider which components of this legacy apparatus need to be mapped into the new world of a broadband PSTN and how best to accomplish that goal. Our focus will be on service measurement, which constitutes an integral component of any policy framework for the PSTN, old or new. 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. 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.

The old PSTN was about one application – voice telephony. As we look at the broadband Internet and describe it as “the new PSTN,” we do not limit ourselves to simple voice communication. We use this term more generally to describe that set of services and supporting infrastructure that are now “critical infrastructure” for our nation. When we set out to measure our communications ecosystem, what parts are critical, what parts do we really care about, and what parts should be our focus as we discuss measurement? These questions are part of our exploration in this article.

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 below.

To understand what it means for the broadband Internet to become the new PSTN, it is important to understand the evolution of the Internet itself. In the next section, 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 for Internet performance metrics. We then focus more specifically on the performance metric-relevant issues that impact FCC policy in its efforts to pursue three important goals: promote universal service, ensure a healthy market for interconnection, and ensure the reliability of our critical PSTN infrastructure. These examples will allow us to highlight some of the challenges for measurement and monitoring. Next, we review the current status of Internet performance measurement, and follow that discussion with our thoughts on the looming policy challenges and how these might be addressed in the near- and longer-term.
Table 1: A New Broadband PSTN is Very Different from Legacy

<table>
<thead>
<tr>
<th>New broadband PSTN is more:</th>
<th>Because:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heterogeneous</td>
<td>More classes of users (sophisticated and mass market), uses (with divergent QoS requirements), and contexts (wider range of technologies supported than in legacy PSTN).</td>
</tr>
<tr>
<td>Dynamic</td>
<td>Resource sharing is even more essential and needs to be flexible to respond to market conditions over varying time scales (i.e., real-time, provisioning, and investment) in part because traffic and capacity utilization is inherently less predictable than in the days of silo-based telephony (i.e. traffic is more bursty); because of the need to customize functionality to address needs of more heterogeneous traffic; and to accommodate the more rapid pace of technological/market change and the uncertainty this brings.</td>
</tr>
<tr>
<td>Complex</td>
<td>Because the 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; and because of the reasons cited above.</td>
</tr>
<tr>
<td>Decentralized</td>
<td>The Internet is inherently less centralized and hierarchical than traditional telephony networks; the boundary between the edge and core is more ambiguous, which has implications for the assignment of control over functionality; and the transition to competition in components has broken the end-to-end, silo-network structure of legacy telephone networks and the resulting industry structure it spawned.</td>
</tr>
<tr>
<td>Distributed</td>
<td>Ownership of key assets is distributed across a wider range of stakeholders from edge-network operators (end-user equipment), applications and content providers, content delivery networks, Web and cloud services providers, and ISPs. The range of asset-owners with heterogeneous regulatory classifications and market-positions is greater.</td>
</tr>
</tbody>
</table>

THE BROADBAND INTERNET AS THE NEW PSTN

In this section, we discuss 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 general service platform; (2) the evolution of the Internet toward becoming a computing utility; (3) the transition from public utility regulation to market-based regulation.

From Telephone Network to Service Platform

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 designed to support a single service: end-to-end circuit-switched voice-grade telephony. It was engineered top-to-bottom as a “silos-network” and was originally owned and managed end-to-end as the AT&T Bell System.

In contrast, the Internet was designed as a general service platform defined by a small set of protocols that support packet-switched data communications between end-points. It was originally built as a networking overlay on top of the old PSTN infrastructure, using leased lines and dial-up
telephony access services. 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.

What was once an overlay on the circuit-switched PSTN technology is now a native platform in its own right. As the capacity and functionality has improved, 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).¹

Figuring out where to draw the line around what constitutes “the Internet” is complex and contentious. We used the term “the Internet” above without carefully defining it. The term is often used in a rather inclusive way to describe the totality of the Internet ecosystem – both the packet-level infrastructure that transports bytes and the higher application levels that build on that service to create and define the user experience. In common parlance and for most consumers with little understanding (or interest) in the industry structure of the telecommunications industry or its technologies, “the Internet” refers to that bundle of services that includes access to the World Wide Web, e-mail, and other data services that require the consumer to use broadband Internet access. We are comfortable with this usage when precision is not required, but any discussion of measurement must be clear as to exactly what is being measured, and which actors are in a position to carry out the measurements. Two definitional distinctions are central to any careful discussion of service performance and measurement.

First, the firms that build the packet-level infrastructure are usually different from the firms that build the higher-level applications. In the old telephone system, whether one was concerned with tracking physical outages in circuits or the voice quality of a call, the same firm was responsible for either issue. If we are concerned with the quality of the user experience in one or another Internet application, the firms that operate the packet-level infrastructure may not be in a position to measure this. Over the Internet, services build on top of lower level service platforms in a layered structure, the organization of the firms follows these service layers, and thus measurement itself will have to take on a layered quality.

Second, the firms that have built this packet-level infrastructure are using it for purposes other than building “the Internet.” The Internet Protocol (IP) is a specification that defines a technology. Many networks are built out of this technology but are not “the Internet,” or that portion of the global collection of networks using IP technology that are part of the public Internet. This distinction may seem obscure, but is in fact critical to any discussion of measurement, and more broadly to any discussion about regulation. Not all IP networks will be part of the future broadband PSTN and will be regarded as critical infrastructure. Specifically to the point of this article, when carriers (e.g., facilities owners) provide telephony using Internet technology (called Voice over Internet Protocol or VoIP), the packet transport they use is often distinct from the transport of packets for “the Internet.” VoIP is not necessarily Voice over the Internet (VoInternet), and measuring “the Internet” does not necessarily tell us how future voice telephony will perform.³

To clarify this perhaps subtle distinction, a useful way to distinguish different networks built out of the IP technology is whether packets can be directly exchanged between them. A corporate network may be designed to be totally disconnected from “the Internet,” or it may be connected to it to some degree.⁴ If a network used for VoIP is not connected at the packet level to the Internet, but achieves interoperation only at the higher service level of the specific application (e.g., telephony), then we should regard these as separate IP networks. When it is important in the discussion that follows, we will talk about a firm implementing an “IP platform,” out of which a higher-level platform such as the global Internet can be built.

**Evolution of the Internet into a Computing Utility**

At the same time that we see Internet technology becoming a native platform for essentially all services, we see “the Internet” becoming a much richer space of capabilities and functions. The Internet, broadly defined, is evolving into a 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.⁵ A growing amount of functionality and intelligence is moving into

---

² Distinguishing between private and public networks that may use the same technology is not always clear and may rely, in part, on the associated regulatory constructs. Similarly, distinguishing between what are private IP networks versus “the Internet” may also be ambiguous, and one needs to rely on regulatory definitions for clarification. A central theme of this article is to highlight what some of the issues that define the “broadband Internet as the future PSTN” entail for measurement.

³ For example, many international VoIP providers offered their telephone services over privately managed and provisioned transport networks, using IP technology but not the public Internet; similarly, Comcast offers telephony services over an IP network that is separate from the Internet.

⁴ Some analysts distinguish between Intranets (enterprise networks that use IP technology but are separate from the global Internet), Extranets (enterprise networks that allow connections to some outside users, e.g. supply chain partners or customers), and the Internet (which refers to the global public Internet). See for example Jack Schofield, “What are Intranets and Extranets?” BBC, Sept. 9, 2010, accessed Aug. 3, 2013, http://www.bbc.co.uk/webwise/guides/intranets-and-extranets.

the Internet to complement, and at times, substitute for intelligence and functionality in the end-hosts.

The term “cloud” has been used to describe this trend; that term as well is subject to many potential interpretations – it can mean software-as-a-service, data centers, computing-on-demand, content-delivery networks, and so on. Figuring out what components of this “cloud” of resources should be deemed to be part of the Internet is again an issue. More specifically, what parts of this larger ecosystem should be part of what we consider the “the new PSTN” – the critical infrastructure for the future? Should we have equal expectations for broadband access service to the home, content delivery overlays, and data centers? What about parts of the packet transport infrastructure provided by firms that are not traditional carriers or ISPs? The richness and complexity of the current services makes it harder to specify performance metrics since there are so many more things that one is concerned about.

**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. Thus, as we move from the world of a telephone PSTN, it would be easy to move toward an unregulated PSTN based on the Internet. However, such a trend would miss the point of what it means for the Internet to become the new PSTN, and by so doing, the Internet would inherit the public interest obligations associated with its new role as basic infrastructure. A broad-ranging debate is required to decide which of the regulatory concerns framed around the old PSTN should be carried forward and re-defined in the context of the new PSTN.

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. From a metrics/measurement perspective, this implies a change from the regulator collecting information and data needed to implement public utility regulation (e.g. through rate-base and other regulatory proceedings) to gathering data in order to ensure the existence of a vibrant market ecosystem. As the regulatory role shifts from the regulator to the

---

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 provided, at times, by ISPs. What portion of these businesses operations (if any) should be included as part of the “new PSTN?” For a discussion of the role of overlays in the Internet, see Peyman Faratin, David Clark, Steven Bauer, William Lehr, Patrick Gilmore, and Arthur Berger, “The Growing Complexity of Internet Interconnection,” *Communications & Strategies* 72 (2008): 51-71.

7 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.
market, the need for publicly-available performance metrics increases. Markets need information to function efficiently, and in a distributed/decentralized market there are more stakeholders who need a greater range of performance-relevant information for efficient collective decision making. Moreover, when the markets are changing rapidly, the metrics need to evolve and adapt, so their development will also need to shift more toward the market.

In this new market-centric measurement environment, the regulator’s position is more specialized and 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 used in regulations (e.g. minimum performance standards) or reflected in government reports. At the same time, the sources of performance data will necessarily multiply, and embracing the multiplicity of sources is one of the ways that markets ensure efficiency. 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.

8 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 some 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, etc.

9 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 may have a different perspective on how to interpret information. Note that while more public information is needed and of more types, this 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 this means in practice will be difficult and context-dependent. 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).

10 The same bureaucracy that makes regulatory decision making slow will impact the regulator’s ability to evolve metrics and reporting practices in response to changing market circumstances.

11 Consumers may give additional weight to government reporting, assuming it is authoritative. Hence, the government’s choice of metrics, how they are explained, and how they are used may influence market equilibria. Consumers may naturally look to the government to tell them what 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.

12 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 (the investment and market research community, consumer and industry advocacy groups, and academics 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 Federal Communications Commission, state and local public utility commissions, the Federal Trade Commission, and Homeland Security personnel seeking to inform the public and support their regulatory efforts).
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 of segments of the end-to-end telephone network, and the drive toward “technological neutrality” in regulation. At the same time, the dynamism and increased complexity of telecommunications markets have increased the costs of command-and-control style regulatory oversight. The increasing complexity and rapid rate of evolution of the Internet also implies a need for different regulatory approaches. In particular, the complexity makes it less desirable to rely on input-based regulations that excessively constrain network management decisions. Regulation should strive to preserve the Internet’s freedom of implementation options to take advantage of context-specific technological advances.

**THE ROLE OF MEASUREMENT: THREE EXAMPLES FROM “THE NEW PSTN”**

In 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: (1) achieving universal service goals; (2) regulating Internet interconnection; and (3) ensuring a reliable PSTN.

---


14 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.

15 Technologically 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 United States 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 technological neutrality is rarely, if ever, achieved in practice.

16 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. Attempting to specify how one should implement the range of technology and function (e.g. network management, content delivery services, policy-based routing, etc.) 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”).
Each of these challenges represents a prototypical policy-based rationale for why the FCC needs to develop new performance metrics for the new broadband PSTN. 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 currently 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.

**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 have been focused on ensuring affordable access to fixed-line telephony services.

The U.S. National Broadband Plan \(^{17}\) appropriately recognized the necessity of migrating and reforming 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. \(^{18}\)

To effectively target the collection and distribution of such funding, the FCC has to have both a definition of what constitutes an acceptable broadband service and a mechanism for identifying which consumers are under-served according to that definition. Of course, there may be multiple purpose-specific “definitions” that may be applicable. For example, in the early days of broadband, the FCC defined the term *broadband* as offering a data rate in excess of 200 Kbps for data collection purposes (but not to reflect their view of what ought to be an appropriate level of service). \(^{19}\) We might need to have separate definitions for fixed broadband and mobile broadband performance since the two services offer distinct value propositions. But any undertaking to expend public funds on universal service will require a definition of what offerings are adequate to fulfill that objective.

---


\(^{19}\) This definition was long criticized as reflecting too low a goal for broadband speed, but such criticism was often misguided. The FCC specified the 200 Kbps standard for its data collection efforts, and it did not designate that as the goal for what broadband service should be nor require it as a minimum performance standard. The 200 Kbps standard had the virtue (in the late 1990s) of excluding ISDN, satellite-based broadband, and mobile broadband services from broadband data collection, 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. 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).
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 use such a definition to establish minimum performance standards (assuming the FCC is successful in establishing its authority to regulate broadband). However, even if the FCC’s jurisdiction were unchallenged, it would not be obvious that the best approach would be for the FCC to adopt an explicit definition of broadband service, rather than leaving it to market forces to define what constitutes appropriate levels of that service. 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, and 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 that 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 whether 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 such 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.

---


21 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 in 2010 the Circuit Court of Appeals for the District of Columbia rejected that claim, which put the FCC's legal authority to regulate broadband into question. See Comcast v. Federal Communications Commission 2010 U.S. App. LEXIS 7039 (D.C. Cir. 2010).

22 Bauer, Clark, and Lehr, “Understanding Broadband Speed Measurements.”
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 over 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 ISP providing transit service 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.

The policy goal here is somewhat diffuse, since the range of harms and problems that might arise is hypothetical. In general, if the Internet is now critical infrastructure, there is now a public interest in ensuring that the degree of interconnection is such that the system functions as an integrated whole, with adequate performance and resilience, and so as to limit any abuses of market power. 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, the FCC has an enduring interest in ensuring an open and efficient market for interconnection. We have explained elsewhere how the FCC’s attempts to regulate network management practices as part of its framework for ensuring an open Internet (sometimes referred to as network neutrality) necessarily engage concerns over interconnection policy.24


With respect to measurement, the FCC needs to monitor the health and activity of interconnection markets to enable it to identify problems that arise and to provide an empirical basis for crafting remedies if market failures occur. Today, there is little publicly available information other than the anecdotal about the scope and nature of ISP interconnection agreements. To adequately assess the health of interconnection markets, policymakers will need technical data about actual traffic flows, points of congestion, and degrees of redundancy, as well as about the terms and conditions in interconnection agreements. Markets for Internet interconnection have become more complex and dynamic with a wider range of interconnection agreements than those that prevailed in the earlier world of simple peering and transit agreements.\(^{25}\) As the recent flap in Europe over the French regulator’s request for disclosure of interconnection agreement information demonstrates, figuring out how best regulators can empirically monitor Internet interconnection markets remains a difficult question.\(^{26}\)

Interconnection is one area for which it is important to remember the distinction between “the Internet” and various other IP-based transport platforms. Different IP networks, such as the one used for “the Internet” and the one used for carrier VoIP, may have different physical points of interconnection. The nature of redundancy, capacity engineering, business agreements, and so on may be different for these different networks, and measurements (and policy) for one may have nothing to do with another.

**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 a 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 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 in 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.\(^{27}\) It will be necessary to monitor the degree of redundancy in different parts of the network (e.g. points of interconnection), and a range of engineering decisions that relate to resilience, ranging from spare capacity to backup power.

---

25 Faratin, Clark, Bauer, Lehr, Gilmore, and Berger; Clark, Lehr, and Bauer.


As noted above, these issues will span a number of industries – not just the carriers with their history of regulatory obligations, but also the providers of content and applications, value-added services, and the equipment and devices that enable use of the Internet. Addressing these issues will necessitate a debate as to which aspects of the ecosystem should be within the scope of concerns about reliability. Once again, one cannot discuss reliability without carefully describing what the system of concern is. Is it the packet transport infrastructure that supports “the Internet?” Is it a different IP network that supports next generation telephony? Is it the data centers that support higher-level Internet services or the content delivery networks that will bring us future television?

**CURRENT PERFORMANCE METRICS PRACTICES**

In this section, we consider the current status of broadband performance measurement. Much attention today is focused on evaluating the performance of mass-market broadband access services. While access performance is important, what matters ultimately for end-users 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, and which can and do have a significant impact on the quality of the user experience. How these components interact with and complement each other varies by type of application.  

There currently exist two main classes of broadband performance data. These two classes are distinguished by the type of data gathered, the methods by which the data is gathered, and the entities involved in the process. For the purposes of this article, 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). 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 that 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.  

---

28 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 Bob Briscoe, “Congestion Exposure (Conex), Re-Feedback & Re-ECN,” Mar. 8, 2012, accessed Aug. 5, 2013, http://www.bobbriscoe.net/projects/refb/; or Steven Bauer and Robert Beverly, “Measuring the Current State of ECN Support in Servers, Clients, and Routers,” workshop presentation, The Cooperative Association for Internet Data Analysis, Feb. 3, 2011, accessed Aug. 5, 2013, http://www.caida.org/workshops/isma/1102/slides/aims1102_sbauer.pdf.  

29 The third-party consultants include market researchers, investment analysts, and traffic analysis technology providers like Sandvine (a provider of deep packet inspection technology).
In recognition of the need for better metrics and coordination among a larger set of stakeholders, the FCC 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.” The goal of this group was to provide “requirements and input to standardization bodies, and describe how the output of these external organizations can best be used to create a reliable, accurate, and sustainable broadband performance measurement infrastructure.” See Figure 1 below for a diagram that envisions some of the main components and measurement and control protocols of this architecture.

![Figure 1: Next-Generation Measurement Architecture](image)

The key take-away from this diagram is that multiple entities and devices are involved in the testing/measurement and data collection process. NMASOG imagines 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 because they supply a service that authoritatively determines “the connection

---

31 Ibid., 3.
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.”  

In response to this FCC outreach group initiative, in June of 2013 the Internet Engineering Task Force (IETF) chartered the Large-Scale Measurement of Broadband Performance (LMAP) working group to standardize a measurement system for performance measurements of broadband access devices such as home and enterprise edge routers, personal computers, mobile devices, and set top boxes (whether wired or wireless). The Broadband Forum (BBF) also recently formed the working group Broadband Access Service Attributes and Performance Metrics (WT-304) to enable standards-based testing and reporting of broadband performance. These standardization efforts may someday fundamentally change the quantity and quality of measurement data that is available.

Until such integrated architectural approaches are standardized and deployed, 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 subsections.

**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 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.

**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 Commission’s rules. In the context of telephone calls (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.

In April 2011, the FCC launched a Notice of Inquiry to gather data and stakeholder perceptions on Internet reliability. There is wide variation in what one might reasonably mean by reliability.  

---

32 Ibid., 1-2 (emphasis in original).
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”). This is often expressed as a statistical time measure (e.g. Mean Time to Failure, or MTTF) or the percentage 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 previous work, we proposed three distinct types of reliability metrics that we expect will be useful for assessing broadband services:

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 several 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” – in which defining what constitutes a typical subscriber is itself a difficult challenge).

Reliability of connectivity: An even more basic notion of broadband service reliability focuses on Internet connectivity. 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, e-mail, and the World Wide Web. Each of these services may fail in multiple ways (e.g. no connectivity, diminished performance, localized or system-wide loss of availability, etc.).


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 the probabilities of Type I and II errors? And so on.

39 Lehr, Bauer, Heikkinen, Clark.

40 Sundaresan et al. propose defining availability “as the fraction 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. See Srikanth Sundaresan, Nick Feamster, Renata Teixeira, Anthony Tang, W. Keith Edwards, Rebecca E. Grinter, Marshini Chetty, and Walter de Donato, “Helping Users Shop for ISPs with Internet Nutrition Labels,” Proceedings of the 2nd ACM SIGCOMM Workshop on Home Networks (2011): 13-18.
The February 2012 announcement by the FCC that it will require 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 a 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 that the intent is to provide additional intuition on why extracting this information from a broadband infrastructure is challenging.

**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 accident) and over longer time scales (e.g. for planning maintenance cycles and capacity expansion investments). During periods of congestion, 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. 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


43 For an introduction to network management, see Mani Subramanian, *Network Management: An Introduction to Principles and Practice* (Reading, MA: Addison-Wesley, 2000).
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 highways 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 that 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?  

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 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 per 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

44 Over 10,000 comments were filed in the FCC’s proceeding (07-52) regarding Comcast’s management of peer-to-peer traffic.
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.

**Data Collection 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, regulators, 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 ones we summarize below are simply 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.

**SamKnows:** Ofcom in the United Kingdom, the FCC in the United States, and the European Commission run large-scale measurements of broadband networks utilizing dedicated measurement boxes and infrastructure provided by SamKnows.\(^{45}\) 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.\(^ {46}\) The current list of tests is provided in Table 2 below. The data from this set of tests has been publically released by the FCC and utilized in academic research projects, the advertising campaigns of competing broadband providers, and as input to other regulatory proceedings.

In the United States, 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.\(^ {47}\)

---


Table 2: Tests Running on SamKnows Test Boxes

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

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 by policymakers. For example, the 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. Also, even with thousands of boxes deployed, this testing infrastructure only represents a tiny share of total subscriber lines. While adequate for addressing certain questions, it may not be suitable for documenting some 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, the FCC, and the EC reflect praiseworthy attempts by policymakers to enhance the collective knowledge base of performance metrics.

**Measurement Lab:** Measurement Lab (M-Lab) is a distributed measurement platform run by a 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.”

---

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 Table 3 below.

**Table 3: Measurement Lab Testing Tools**

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

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 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 that can be mined by researchers seeking to understand emerging trends as well as transient phenomena.

**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 websites and thus observes web traffic from a wide

---

49 There are also a host of other market research firms and investment analysts that collect and publish data, either in for-fee research reports or as part of their marketing efforts. Here we highlight some of the most significant and most often cited.
selection of end users. 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. 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 Sandvine’s business, the company gains 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.

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

CAIDA Topology Data: In contrast to the usage data from Sandvine and Akamai, CAIDA (the Cooperative Association for Internet Data Analysis at the University of California-San Diego) collects data about the connectivity of the Internet. The organization maps patterns of interconnection at various levels of detail, which allows it to document issues such as outages (whether due to natural or manmade causes) and possible business relationships between the actors that make up the Internet.

Measuring IP Infrastructure

This discussion is about third-party measurement centered on “the Internet.” But the distinction between the Internet and alternative IP-based platforms becomes critical here, in that other versions of the IP platform, such as the one now being set up by carriers to offer VoIP, are not designed to allow the attachment of general purpose, third-party edge devices. For the typical consumer of VoIP, the carrier provides the edge-device, which connects to the home telephone wiring of the customer. So there are few if any options to attach edge devices that can measure and monitor performance, reliability, and so on. For example, the topology data gathered by CAIDA, described above, cannot be gathered unless it is possible to attach third-party measurement devices to the network in question. What this implies is that if the regulator is concerned with the service

performance parameters of these alternative IP platforms, they are going to have to rely on (and perhaps compel) data that is gathered and reported by the provider, whether that data is gathered “in the net” (by looking at packet flows), or “at the edge” (in the edge device that implements the VoIP service).

The provider may choose to provide instrumented edge devices (the MEED concept in Figure 1 above) or may choose to allow third-party measurement devices to be attached to the network (e.g. the Sandvine devices described above). However, these decisions will be at the discretion of the service provider unless they are compelled to provide such instrumentation. There may be adequate motivation for these providers to gather (and perhaps release) this data, but there can be no expectation that such actions will be undertaken voluntarily.

**POLICY RECOMMENDATIONS**

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

**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 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 there will be strategic considerations that will have a strong impact on impeding incentives to share data. However, 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, and in sharing the costs of performance management.

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.

---

54 Hereafter, we will use “ISPs” broadly to refer to Internet service providers, content delivery providers, other cloud resource providers, and technology providers whose business is the provision and operation of Internet-related services.

55 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 misrepresentations. The latter may be helpful for performance verification in contract enforcement. To enforce a contract it is not sufficient that the contracting parties know that a mistake has occurred, and it may be necessary for a third-party enforcer.

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

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 online 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, relatively coarse sampling is likely to be desired—not for 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.,

58 The present authors’ 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.
59 It will be important to manage costs both in the aggregate (the potential for BigData explosion is large) and individually. 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 demonstrates the potential of generating significant levels of participation with no direct compensation being required.
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, the FCC should leverage its knowledge of the evolving ecosystem of measurement in crafting its metric-specific interventions. The market-based need of ISPs and the needs of end-users (both mass market and enterprise customers) to engage in significant performance 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 will be provided without the providers’ express consent (by third-party analysts and from overlay measurement platforms). The FCC’s need for such data is likely to be less detailed and granular than that of 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). Also, the FCC should stand ready to compel data collection from ISPs and other industry participants if adequate data is not provided voluntarily or is otherwise unobtainable 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 dialogues with initiatives like M-Lab, and by promoting access to open source measurement platforms and clients.

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, 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 at all. Furthermore, to make interpretation more tractable, there will need to be greater recourse to index metrics (which weigh 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.


61 Bauer, Lehr, and Clark, “Understanding Broadband Speed Measurements.”
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, then 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.

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.62

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 that do not provide detailed discussions of their methods have, at best, only limited value.

Ideally, the code used to generate the measurements would be open source; open source code has the benefit that independent parties can understand and evaluate the experimental methodology. 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 especially 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 the FCC should be more willing to acknowledge that data errors will occur and will need to be corrected. It will not have the sole responsibility for managing performance monitoring and should not assume excessive responsibility

62 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 requirements 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). Jon G. Elberath and Michael Pecht, “IEEE 1413: A Standard for Reliability Predictions,” IEEE Transactions on Reliability 61, no. 1 (Mar. 2012): 125-129.
responsibility. This will place a greater burden on the market, but that is the price of giving market forces greater scope for discretion.

**Shift to Ex Post Enforcement**

Finally, we believe that the increased complexity and dynamism, and the shift of responsibility from command-and-control management to reliance on market forces, will make it harder for the FCC to craft or rely on *ex ante* (specific) 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* 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 should use that authority sparingly. Wherever possible, the FCC should leverage market-based (third-party or voluntary) data for assessing performance. To do this effectively, however, the FCC may need to deploy some in-house infrastructure like the SamKnows service or conduct Commission-sponsored primary research to facilitate its verification and interpretation of data from other sources, and to fill such gaps as exist.

**CONCLUSIONS**

In this article, we have focused on the 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 and its supporting IP infrastructure becomes unavoidably the focus for government regulation. It is no longer merely an application that relies on the PSTN – it is now 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 raise the question as to what portions of that legacy regulation need to be re-mapped into the broadband Internet world. To the extent that 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. (For example, 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 a continuing role for communications policy regulatory enforcement. Of course, in keeping with the secular trend of the last 50 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 of meeting these 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 new 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.


