

The Exabyte Internet

We're being hit with an "exaflood" of data, in part due to expanding distribution of high-definition video. We'll need more than fiber. Here's what it will take to prepare our network infrastructure.

By David P. McClure ■ *US Internet Industry Association*

Recent research reports have suggested that data growth may be greater than the planned Internet infrastructure can handle. While some have challenged the precise numbers, it is clear that the amount of data traffic, especially video, on the Internet is growing at an accelerating rate. This will soon create a situation in which billions of gigabytes of data – exabytes – will transit the Internet each year.

The Exabyte Internet promises to bring new services, applications and opportunities to both consumers and businesses in America, but it will not do so automatically. In addition to investments in additional infrastructure and capacity, there will need to be innovations in core technologies and network management,

flexible strategic plans and a national broadband policy that supports investment and innovation over regulation and a continuation of the status quo.

This paper examines the concerns as we move toward an Internet transporting billions of gigabytes of data; the likely effects on consumers, businesses and network operators; and the ramifications on Internet public policy as we grow from a Megabyte Internet to an Exabyte Internet.

It suggests four policy concerns:

- A better understanding of the data.
- A more realistic view of the Internet infrastructure.
- A "light hand" regulatory approach.
- A focus on network investment as the national priority.

The Exabyte Flood

When Discovery Institute Fellow Bret Swanson noted in a *Wall Street Journal* op-ed that a coming "exaflood" of data could exceed available bandwidth capacity of the Internet, he gave focus to growing concern that the Internet might be at risk of failure.¹ Swanson's research was not new, and his predictions will be familiar to readers of this magazine. In 2003, researchers at the University of California at Berkeley estimated that all of the data created from the dawn of humanity to that date amounted to 5 exabytes, or 37,000 times

the total information stored in the Library of Congress.²

Other studies suggest that even this enormous amount of data is being surpassed every single year by the growth of technologies for data creation, distribution, transit and storage.

The global market research firm International Data Corporation released a study in early 2007 entitled "The Expanding Digital Universe: A Forecast of Worldwide Information Growth Through 2010,"³ indicating that in 2006 some 161 exabytes of digital information were created and copied. That amount of information is equal to three million times all the books ever written. IDC predicts the amount of information created and copied in 2010 will surge more than sixfold to 988 exabytes – a compound annual growth rate of 57 percent.

Similar concerns appeared in other research. "Telecommunications Trends: TMT Trends 2007" from the Deloitte Touche Tohmatsu Technology, Media and Telecommunications Industry Group concluded:

"One of the key possibilities for 2007 is that the Internet could be approaching its capacity. The twin trends causing this are an explosion in demand, largely fueled by the growth in video traffic, and the lack of investment in new, functioning capacity. Bottlenecks are likely to become apparent in some of the Internet's backbones, the terabit-capable pipes exchanging traffic between continents. Investment, either in laying new cable or lighting existing fiber, may be stifled by continuing falls in wholesale capacity prices. Similar capacity con-

**IDC released
a study in early 2007
... indicating that in 2006
some 161 exabytes of digital
information were created
and copied. That amount of
information is equal to three
million times all the books
ever written.**

straints may well appear in the ISP and telecommunications networks that provide broadband connectivity to consumers. The impact may be most noticeable in the form of falling quality of service.”⁴

A January 18, 2007, research report from UBS stated, “The growth in online video has significant implications for bandwidth demand on the Internet. Standard-definition video requires roughly 450 times the bandwidth per second of a web-browsing session. Taking it a step further, high-definition video requires 2,700 times the bandwidth per second of a web-browsing session. Based on these numbers, Level 3 and other participants in the industry have suggested that existing Internet backbone capacity is insufficient to transport high-quality commercial video.”

Nor is this strictly a US-focused issue. Netcom CEO Bill Barney has noted that the total daily intra-Asia Internet usage grew by 13,857 percent between 1998 and 2005. Barney forecasts that at the rate that demand for high-bandwidth applications is growing in the region, the 15,360 Gbps of capacity currently serving Asia will be completely used up by 2012.⁵

Though there was some quibbling over the methodology used in each of these reports, and of the final numbers derived to quantify the size of the growth in data, there is consensus that the amount of data created – data that will ultimately be transmitted across the Internet – will grow at a sufficient pace to warrant serious policy consideration.

We Have Some Time

Still, the numbers do not yet mean an Internet in imminent danger of collapse. In 1995, at the dawn of the modern Internet, the total amount of data that traversed the Internet backbone was 1.5 million gigabytes.⁶ By 2006, this had grown to 700 million gigabytes.⁷ While this is significant growth in Internet traffic, it still represents only 0.5 percent of the 161 exabytes of data that were created that year.

There are about 5 million bytes, or five megabytes, in the complete works of Shakespeare, and a pickup truck full of books might amount to one billion bytes, or a gigabyte. One billion of those book-filled pickup trucks, or one billion gigabytes, is an exabyte.

Whatever the exact dimensions of the exaflood, it will impel us to significantly enhance network capabilities to maximize the promised benefits of new and innovative applications.

Swanson himself noted: “I don’t equate the exaflood with a crisis. I think it is a great sign of advance. It will be good for hardware, software, network and applications companies. It will be good for consumers. It will happen if we let it.”⁸

Defining an Exabyte

To get a sense of the discussion, it’s helpful to define the units of measure – the bytes – that make up the digital content world. A single digital character, a letter or number, is a single byte. A typewritten page is about 2,000 bytes, or two

Hear David McClure speak about the coming exaflood at our Broadband Summit

September 10 – 12, 2007

Hyatt Regency DFW

Dallas, Texas



877.588.1649

www.bbpmag.com

Concerns about Traffic Growth

1. Growth in the amount of Internet traffic will continue to accelerate. As more data is created, the human desire for communication will drive an increase in the amount that is communicated across the Internet via e-mail, instant messaging, file sharing and other technologies. According to network infrastructure company CacheLogic, more than 60 percent of Internet traffic is being taken up by peer-to-peer swaps, and about 60 percent of those swaps involve video content.¹³ (This doesn't even count video and music downloading traffic from sites like iTunes and MovieLink.) This surging traffic is driving record sales of routers worldwide as network operators struggle to keep pace.¹⁴

Worse yet, current video traffic is at best in HDTV format, which is already obsolete. The newer "4K" video images have roughly 4,000 horizontal pixels – offering approximately four times the resolution of the most widely used HD television format, and 24 times that of a standard broadcast TV signal.

2. The number of data streams will increase substantially in the coming decade due to the transition from IPv4 (Internet Protocol Version 4) to IPv6. Internet Protocol version 6 is a network layer protocol that will replace IPv4, the current version of the Internet Protocol, for general use on the Internet. The main improvement brought by IPv6 is the increase in the number of addresses available for networked devices.

IPv4 supports 2³² (about 4.3 billion) addresses, which is inadequate for giving even one address to every living person, let alone for supporting embedded and portable devices. IPv6, however, supports 2¹²⁸ addresses; this is approximately 5 x 10²⁸ addresses for *each* of the roughly 6.5 billion people alive today.

With such a large address space available, IPv6 nodes can have as many universally scoped addresses as they desire.¹⁵ Of course, each of these addresses, if attached to a person or device, will begin generating its own stream of data – even if only to acknowledge its existence on the Internet.

3. The evolving business models of internetworking will continue to shift traffic from the Internet backbone to a peered system in which content is streamed directly to consumers. This shift substantially alters the dynamics of infrastructure growth in that it drives a need to invest not only on the Internet backbones but also on the more expensive buildout of the "first mile" connection to consumers.

4. There may be insufficient investment in Internet infrastructure to accommodate the growth in data because government policies inhibit or interfere with the ability of networking companies to attract investment capital from the marketplace. National and state policies toward broadband innovation and deployment directly affect the ability of network operating companies to raise investment funding in the capital markets. To the extent that these policies inject uncertainty into the marketplace, threaten the potential return on infrastructure investments, or interfere with the ability to operate the networks in a way that maximizes capacity and efficiency, such policies negatively limit the ability of network operating companies to support expanding data traffic on the Internet.

kilobytes, and a small, low-resolution image is about 100,000 bytes, or 100 kilobytes. There are about 5 million bytes, or five megabytes, in the complete works of Shakespeare, and a pickup truck full of books might amount to one billion bytes, or a gigabyte. One billion of those book-filled pickup trucks, or one billion gigabytes, is an exabyte.

Opportunities of an Exabyte Internet

The opportunities for new services, new applications and enhancements in lifestyles are apparent in both the recent growth of the Internet and in plans already underway for the next generation of services.

For consumers, this should mean more Internet options at increasingly lower costs. Yahoo has announced that it will allow its mail system users to store an unlimited amount of data on its servers for an unlimited amount of time.⁹

Wal-Mart, Netflix and Amazon have all announced plans to transmit DVD-quality video online. A growing number of universities are making course lectures available online, often in real time. And telemedicine programs are transmitting medical images and linking patients with distant specialists for real-time consultations.

Added to this are new and pending applications for IPTV, video on demand, security, personal communications and telecommuting.

The impact on businesses also should be a net positive. While requirements for local data storage and longer-term data archiving will increase, the costs associated with this storage will decrease overall because of the ability to replace file cabinets and paper storage with less-costly electronic data storage. Data storage itself will become cheaper, more efficient and more secure.

Businesses small and large will experience an increased ability to place content on the Internet for use by their managers and by current and potential customers.

Realizing these opportunities, however, will be virtually impossible using only the Internet structure that exists today. The flood of data, applications

Capacity Growth Is Just the Beginning

- **The Internet is not a single entity.** For the sake of simplicity, it is easy to think of the Internet as a single, integrated network. In reality, it is a loose confederation of thousands of private networks that have agreed to use the same protocols. There is not even a single “Internet backbone,” but instead dozens of backbones operated by companies of every size. Simply “adding capacity” means coordinating the funding, logistics, rights of way, purchases and installations of thousands of updates in virtually the same time frame – a nearly impossible task.

- **The problem is at the fringes, not the center.** Companies that carry the highest volume of Internet traffic are capable of managing their growth, given sufficient investment capital. It is at the outer edges of the Internet, where users connect, that the biggest problems will be experienced. This first-mile loop of connectivity is already struggling to finalize the deployment of basic broadband to all of the interested users, many of who choose to live in remote geographic areas that are expensive and difficult to reach with today’s technologies. Others serve urban centers where available rights of way and existing fiber networks are already under duress, making them expensive and difficult to reach. For these companies, investment capital for capacity may be very difficult to acquire.

- **Even at present growth rates, data will eventually outgrow capacity.** Market research group Teleography notes that average traffic across the net increased 75 percent last year, while capacity grew 47 percent¹⁷ – a trend that is likely to continue, even as capacity increases.

and services will demand innovations in network management, strong and increasing investments, and effective business models that have been tested in the marketplace.

As noted by David Farber, the Carnegie Mellon engineer considered to be the “grandfather” of the Internet: *“An updated Internet could offer a wide range of new and improved services... to provide these services, both the architecture of the Internet and the business models through which services are delivered will probably have to change.”*¹⁰

Concerns About an Exabyte Internet

Network operating companies are pushing hard to deploy and expand the broadband infrastructure of the United States. According to the National Telecommunications and Information Agency (NTIA), a unit of the US Department of Commerce, America leads the world in the number of broadband

subscribers, with 64.6 million as of June 2006.¹¹ It is true that the definition of “broadband” used by the Federal Communications Commission is, at 200 Kbps, lower than that used by most other nations, but minimum broadband connections in the US are generally three or four times higher.

Between June 2001 and June 2006, the number of homes with broadband in the United States increased sixfold, from 9.2 million high-speed lines to 64.6 million. During the same period, satellite and wireless broadband grew by a factor of sixty. As of June 30, 2006, DSL connections were available to 79 percent of households with telephone service available, and cable modem service was available to 93 percent of households with cable service available.¹²

The primary concern about the amount of data being generated is that an increasing proportion of it must eventually traverse the Internet, if only on its way to a storage facility. And while that

Internet traffic today is only a tiny percentage of the total, this inexorable migration of information and the growth in use of broadband by both businesses and consumers raises concerns about the ultimate ability of the operators of thousands of networks that make up the Internet to build infrastructure quickly enough to keep pace.

There are four specific concerns (see box) involving traffic growth and technology issues.

Network operators are today investing heavily in infrastructure, in terms of both capacity and innovations to better manage the flow of traffic across the Internet. According to CMP Technology’s research unit Heavy Reading, worldwide sales of long-haul DWDM (dense wavelength division multiplexing) products grew by more than 30 percent in 2006, to \$1.8 billion, as network operators scrambled to add capacity in the wake of surging demand for bandwidth.¹⁶

Some have suggested that simply adding capacity will resolve the issues of explosive Internet traffic growth. Though additional capacity is an important part of the overall strategy, it will not in itself resolve all of these issues, for three reasons (see box).

In addition to investment in infrastructure and capacity, network operation companies will need four other critical assets as they move forward:

- **Scalable core technologies capable of faster, more effective throughput.**

Though we have been able to extract ever-greater efficiencies in transmission of data over the Internet through compression and innovative routing, some core technologies will simply not offer sufficient throughput to meet the demands of the future.

- **Innovative technologies and solutions for network traffic management.**

In an age in which unsolicited spam comprises 80 percent of the world’s e-mail volume, it is not enough to simply make more capacity available. Continued management to route around bottlenecks, efficiently meet applications’ technical requirements, and ensure that delay-sensitive packets get through will remain key to smooth operations. Ef-

Policy Needs

• A better understanding of the data.

Limited views of Internet broadband growth and penetration, such as the annual OECD rankings and data based on single perspectives of Internet use, are not particularly useful foundations for public policy, and may in fact lead to erroneous or unintended conclusions. More comprehensive approaches to broadband data collection, including the Notice of Proposed Rulemaking promulgated by the Federal Communications Commission¹⁹ in April 2007, will provide a more suitable policy foundation.

• A more realistic view of the Internet infrastructure.

As noted, the Internet is not a single, monolithic entity. No single company or group of companies controls it or should, and no single company or group of companies can meet all of its needs. Just as no single set of Internet applications will benefit every consumer, no “one-size-fits-all” regulation will enable every broadband network to meet consumer needs in an efficient and affordable way.

• A “light hand” regulatory approach.

The most successful changes in Internet laws of the past decade have been those that advanced simple, light-handed deregulation with a minimum of unintended consequences. This approach has left the Internet with a minimal tax burden, fewer legacy telecommunications regulations from the 19th and 20th centuries and fewer barriers to growth. The result has been a vibrant, fast-growing and innovative Internet industry that is well positioned to meet the demands of the future.

• A focus on investment as the national priority.

The most pressing broadband need in America today is to craft an infrastructure capable of reaching every American business and household with the bandwidth and services need by those businesses and households. The goal of any national broadband policy must be to find innovative and effective ways to encourage the private investment to build that infrastructure.

efficiency techniques, such as those already in use by major e-commerce vendors to ensure the delivery of their content through companies such as Akamai, will be necessary to ensure that priority content gets where it needs to go without transmission delays that disrupt performance.

• Flexible, prescient strategic plans.

Juggling growing data streams, changing consumer needs and tastes, rapidly evolving technologies and the uncertainties of the capital markets will place strains on even the best of network operating companies. The ones that are most successful in meeting the demands of their customers will be those with a clear vision, strong strategic plans, effective business models and the ability to understand and react positively to emerging trends.

• National, state and local policies that stimulate continued investment in broadband networks.

Estimates of what it will cost to build networks capable of meeting an ever-increasing flow of data – and then deliver that capability to every American business and household – range from \$300 billion¹⁸ to as much as \$1 trillion. Public policy should first focus on enabling network operating companies to secure and use the investment capital needed to meet these estimates and to provide consumers with affordable broadband services.

Policy Implications for the Exabyte Internet

This latter point speaks to the heart of the policy implications for the Exabyte Internet. Four policy platforms are needed to support the needs for broadband growth in America (see box).

Only then can we be assured that our networks will not constrain economic growth. **BBP**

About the Author

Dave is president and CEO of the US Internet Industry Association, in Alexandria, Virginia. He can be reached by phone at 703-647-7440 or Info.3@usiia.org. The Web site is at www.usiia.org.

References

1. “The Coming Exaflood,” Bret

Swanson, *The Wall Street Journal*, January 20, 2007

2. <http://www2.sims.berkeley.edu/research/projects/how-much-info/datapowers.html>
3. http://www.emc.com/about/destination/digital_universe/
4. <http://www.deloitte.com/dtt/research0,1015,sid%3D1012&cid%3D108298,00.html>
5. “Asia to run out of bandwidth by 2012 - Asia Netcom” Total Telecom, April 25, 2007
6. University of Minnesota Digital Technology Center, as cited in the *Chicago Tribune*
7. University of Minnesota Digital Technology Center, as cited in the *Chicago Tribune*
8. “Exaflood Debate,” Bret Swanson, March 24, 2007, http://www.disco-tech.org/2007/03/exaflood_debate.html
9. <http://www.techcrunch.com/2007/03/27/yahoo-mail-announces-unlimited-storage/>
10. “Hold Off On Net Neutrality,” By David Farber and Michael Katz, *Washington Post*, Friday, January 19, 2007; Page A19
11. NTIA Fact Sheet, May, 2007, at http://www.ntia.doc.gov/ntia-home/press/2007/ICTleader_042407.html
12. FCC Data cited by NTIA, May, 2007, at http://www.ntia.doc.gov/ntiahome/press/2007/ICTleader_042407.html
13. http://news.com.com/Net+video+xplosion+triggers+traffic+jam+worries/2100-1025_3-6042300.html
14. <http://www.itnews.com.au/newstory.aspx?CIaNID=46129>
15. <http://www.en.wikipedia.org/wiki/IPv6>
16. <http://www.heavyreading.com>
17. http://www.baselinemag.com/print_article2/0,1217,a=205078,00.asp
18. “National Imperative: Broadband Everywhere by 2010,” Cisco Systems, January 15, 2002
19. Notice of Proposed Rulemaking, Docket 07:37, at http://hraunfoss.fcc.gov/edocs_public/attachmatch/FCC-07-17A1.pdf