

Scaling the Internet

Andrew Lippman

MIT Media Laboratory

July 15, 2002, modified: 7/22/02, Printed: July 22, 2002

Abstract

The internet reaches more than half the population of the country and is now a critical communications backbone for government and society. We are at a crossroads today where we are faced with the option of improving its speed and performance or solving the problem of scalable, universal access. The choice is whether we want to mimic the central ownership model of broadcasting versus an egalitarian model of telephony or the press. In this note, I suggest two things: (1) that the build-out of the network's last mile may come via industries other than the communications industry, and (2) that technology allows us to produce a network that scales economically and practically while retaining the end-to-end openness that made it innovative in the first place. The thesis is that if we allow it to develop *incrementally*, evolution can occur at the discretion of each individual without placing bets on large systems or winners and losers.

1.	Introduction	2
2.	Speed versus Symmetry	4
	2.1. Symmetry in Wires	4
	2.2. Symmetry in Wireless	5
	2.3. Open Access	6
3.	Incrementalism versus Installationism	7
4.	Summary	8

1. Introduction

The Speculative bubble of 1999 has diverted people from the simple fact that the internet is part of the backbone of society. The beneficiaries are normal people doing normal things. Particularly, non-profits and small businesses: The Essex National Heritage area, for example, has a simple site that is used 270,000 times per month by people who visit the region and get information about the non-profit agency programs (grants and such). Inns and B&B's who could never afford to compete with hotel chains in mass media are readily accessible, as are restaurants, maps, route guides. E-Bay alone has changed the business landscape – there are now learned articles in refereed journals about selling and bidding techniques that help people distribute everything from antique autos to mass market items. Many schools now require internet references in student papers.

The original dream of the internet as an engine of democracy is becoming realized. Penetration is past 50% of American households, far more if you count email addresses and shared, or family users. Like the Colt 45, it is the communications equalizer of the century. (God made big sites and small stores and ARPA made the internet to make them all equal.)

No one is a greater beneficiary of this than the government itself. The savings afforded by people renewing their automobile registrations on line or filing taxes could underwrite wiring schools and libraries in America if we simply allocated internet-based savings to its extension.

The challenge facing us is how to best scale this opportunity without sacrificing the open, innovative nature that the internet brought to us. When the protocols were first outlined, designers were free to argue various options on their merits alone, without regard for ultimate economic value. There were no business plans based on the network, and commerce was prohibited.

Today that situation is reversed. There are many stakeholders, many businesses whose operation is destabilized by the internet, and many attempts to restrict the innovative potential of the network in favor of existing models.

None of these arguments have any technical support. We have the technology to make the internet work in any way we wish. Indeed, as I will show, the technology favors segregation of bit delivery from services and argues strongly for a network where the

Insights and Analyses of Online Auctions, Bapna, R., et al, Communications of the ACM, November, 2001

Pingree School, Hamilton Ma, requires at least 2 internet references per paper.

Source, Julia Lippman, 2002

Nielsen//NetRatings, the global standard in Internet audience measurement and analysis, reveals that the largest broadband markets achieved significant audience gains of at least 48 percent year-over-year in their respective high speed populations this past April (see Table 1). Overall, more than 25.2 million surfers at-home accessed the Internet via cable modem, DSL, ISDN or LAN last month as compared to 15.9 million individuals in April 2001, increasing 58 percent year-over-year.

Source: Nielsen//netratings, 5/02

intelligence is at the ends rather than embedded within the network architecture. Most important, such distributed intelligence allows for incremental deployment and greater scalability of the design.

The fundamental irony is that even though the internet is a communications medium, it may well be that future extensions of it will come from parties other than the communications industry. In particular, the system may be more *economically scalable* if it is *incrementally built*. That is to say, rather than extend the network by large-scale, universal service methods, innovation and ultimate extension may be best served by individuals provisioning their own networks piece-by-piece. It may well be that the device manufacturers rather than communications providers are the ultimate constructors of the network.

I restrict comments to the last mile. The backbone probably must be a large-scale, universal infrastructure.

2. Speed versus Symmetry

The fundamental issue is speed versus symmetry. The net was invented when 300 bits per second was a fast modem rate; the backbone itself ran at 56K. The current baseline for broadband service is 1.5 megabits per second, fast enough for real-time VHS-quality television and high fidelity music and immediate enough for telephony and graphics. The next thresholds are DVD rates (5-9 megabits per second) and WiFi (802.11b) rates of 11 megabits. 802.11a is five times that, via radio, and commercial Ethernet installations run anywhere from 100 megabits to a gigabit. These numbers are dizzying and they occlude the point.

2.1. Symmetry in Wires

Symmetry is the issue. Provisioning higher speed usually entails a hidden restriction: the faster the system, the less symmetric it is. That means that access is fast when the bits are flowing from a provider to a subscriber, but not the other way. Cable systems, for example, share anywhere from 5-50 megabits per second from the head end to a neighborhood, but one-tenth of that in reverse. Further the architecture of both cable and DSL prevents bits from going from one user to another directly: All local, person-to-person data go up to the head-end, then back down to another subscriber.

This asymmetry has two critical effects. First, it has the architectural effect of making the central office or head end a gatekeeper. The system owner bundles two independent features together: bit delivery and computing or content services. While it is quite natural that a network provider own the equipment, it doesn't follow that this ownership must be coupled with services that have nothing to do with delivery. *There is no engineering reason why this architecture need be adopted.*

Second, it limits the scale of the system. Modern research demonstrates that by localizing information flow, the system capacity can grow with the number of users. By contrast, these asymmetric architectures provide a fixed, limited capacity that must be shared. The importance of this statement cannot be overemphasized: A symmetric cable system can appear to have an infinite number of television channels, everyone can be a broadcaster or server.

D. Vyzovitis, "An Active :Protocol Architecture for Collaborative Media Distribution", MS Thesis, MIT Media Lab, June 2002.

Actually, the telephone network assumes this design: local calls go only as far as the nearest exchange, thus providing greater capacity on intercity trunks. It scales better than the original Federal Express model that routed all packages through Memphis.

2.2. Symmetry in Wireless

The same situation obtains in wireless networks, although the supporting research and examples are more recent. Historically, the spectrum has been regarded as a limited resource who use must be adjudicated by property models. Broadcasters “own” their channels, and cellular operators “own” the bands in which the phones operate.

This is the model for spectrum allocation used, by the FCC and many regulatory agencies. Allocation restricts spectrum use in order to simplify receiver design. It also subdivides a limited resource. This is not the only allocation strategy, but it is the dominant one for commercial applications of radio. Amateur and citizen's bands leave the spectrum use and interference issues to the community of users; regulations largely defend adjacent spectrum. Unlicensed spectrum likewise has no guarantee of merchantability and some protective restrictions.

Consider, by analogy, a cocktail party. There are parallel conversations occurring in the same room, each at low enough power to reach only the intended listener. There are no booming broadcasters flooding the space (usually), and we could pass simultaneous messages among many of us by relay.

Indeed, this is precisely the model for wireless telephony being explored at the MIT Media Lab, Bell Labs, Mesh Networks, ArrayCom, and others.

The significant result from this work is that the wireless network can be more efficiently used, require less power at each radio (longer battery life) and scale better when the radios relay messages among themselves to reach an endpoint than when they blast a signal all the way to the tower.

However, there are implications for regulation. One model is to mimic the law of the sea: no one owns the sea, but there are rules for safety and navigation that make the oceans scalable.

2.3. Open Access

Symmetry is not merely an issue of whether you want to run a web service from your home, or whether you want to make your music library available to your neighbor. It defines whether we are

Spectrum allocation allows for cost-effective receiver design in that potentially interfering signals are prohibited. This is why there are “taboo” channels in television broadcasting. A more robust and expensive receiver would permit use of all channels. It is useful to realize the interference does not exist in the airwaves, only in the receiver.

G.J. Foschini, M. J. Gans, “On Limits of Wireless Communications in a Fading Environment when Using Multiple Antennas”, *Wireless Personal Communications* 6: p. 311-335, 1998.

D. Reed, “Capacity and Interference in Wireless Networks”, Open Spectrum Meeting, MIT Media Lab, June 2002.

Lippman, Network Tutorial, MIT Media Lab, 2002.

<http://www.meshnetworks.com/>

http://www.arraycomm.com/index_flash.html

legitimizing a many-to-many telephone-like system or a digital, packetized broadcast network, with all of the resource ownership, content control, and programming limitations that implies.

From a policy perspective, the important point is that we can build wired and wireless networks in open and extensible ways that retain the original goals of the internet: intelligence at the terminals, and open field for innovation.

3. Incrementalism versus Installationism

Wireless internet access via WiFi (802.11) captures people's imagination these days for two reasons: it is readily available, and it de-stabilizes the market for high-speed cellular extensions. Witness the turmoil in Europe where those who bid for third generation systems assumed a monopoly on mobile data communications that is being undermined.

But the more important difference is how the two systems are built. 802.11 is incremental and the cost is borne by each end-user who installs the necessary access points. It can spread like a virus, and it can be replaced and upgraded at the discretion of the users.

3G telephony on the other hand, requires large-scale installation before anyone upgrades their handset. In the US, the slow pace of 2.5 G proves the point. It is available in select cities, but few bother to purchase a new phone that offers no widespread improvement.

The irony of some of these systems is that they expose inflexible business plans more than faults in the network architecture. Napster, for example, is used to recover a favorite high school song as much as it is used to pirate Madonna. The music industry, however, is structured around big hits. The technology allows them to make a different album for each purchaser, but the marketing department never got that message. Similarly, television time shifting systems (TiVo today, your PC tomorrow), enhance access to programs potentially as much as they disrupt the already broken model of prime-time. The Germans and the French discovered this in 1980.

From Lessig, *The Future of Ideas*. 2002.

An incremental approach fosters the same innovation that the internet has brought us to date. The future network is not merely about web access and telephone calls. In short order, consumer controls will become wireless devices, as will heart monitors, tourist guides, building controls, drive-throughs, toll gates, street lights and parking meters. None of these will operate by permission of a wireless spectrum-holder, and most will be stillborn if so hobbled.

In France, the automatic timeshifting was called EPEOS, and in Germany it is called VPS and is still in use

The likely providers of these options are not the traditional communications companies, they are the builders of devices and systems. Communications is a backbone transport issue, functionality is a personal one. These are nascent applications that to date have had no voice in the discussion.

4. Summary

The activity of the society today involves moving bits as much as moving atoms. However we tend to view the infrastructure for the two activities with a totally different economic eye. The atoms infrastructure (roadways etc.) is developed gradually with local control and distributed costs. That's very much the way the Internet grew into prominence as a network of networks.

We abjure that model for communications. More often, we build large scale, vertically integrated distribution systems, with one company possibly controlling the bit from production to consumption. This lmites innovation and has little technical support.

The key issues, therefore, are how to amplify the utility of the internet penetration, and what policies will allow the internet to become as integrated with democracy as the post office, the telephone and the voting booth. In terms of policy options, we can assess the issue in the following terms:

- Does the policy enable easier, cheaper and more widespread access?
- Does it let everyone be a server (symmetry)?
- Does it allow the build-out cost to be scaled incrementally, by users versus by large systems providers?

The motivating reason for these issues, beyond the obvious benefits of a connected society, is the simple fact, supported by research and example, that an open, end-to-end architecture supports these policy goals potentially with a new cast of characters.