practice

Knowing where to begin is half the battle.

**BY THOMAS A. LIMONCELLI,
WITH AN INTRODUCTION BY VINTON G. CERF**

Successful Strategies for IPv6 Rollouts. Really.

THE DESIGN OF TCP/IP began in 1973 when Robert Kahn and I started to explore the ramifications of interconnecting different kinds of packet-switched networks. We published a concept paper in May 1974,[2] and a fairly complete specification for TCP was published in December 1974.[1] By the end of 1975, several implementations had been completed and many problems were identified. Iteration began, and by 1977 it was concluded that TCP (by now called Transmission Control Protocol) should be split into two protocols: a simple *Internet Protocol* that carried *datagrams* end to end through packet networks interconnected through *gateways*; and a TCP that managed the flow and sequencing of packets exchanged between hosts on the contemplated Internet. This split allowed for the possibility of real time but possibly lossy and unsequenced packet delivery to support packet voice, video, radar, and other real-time streams.

By 1977, I was serving as program manager for what was then called the Internetting research program at DARPA (U.S. Defense Advanced Research Projects Agency) and was confronted with the question, "How much address space is needed for the Internet?" Every host on every network was assumed to need an address consisting of a "network part" and a "host part" that could uniquely identify a particular computer on a particular network. Gateways connecting the networks of the Internet would understand these addresses and would know how to route Internet packets from network to network until they reached the destination network, at which point the final gateway would direct the Internet packet to the correct host on that network.

A debate among the engineers and scientists working on the Internet ran for nearly a year without a firm conclusion. Some suggested 32-bit addresses (8 bits of network, 24 bits of host), some said 128 bits, and others wanted variable-length addresses. The last choice was rejected by programmers who didn't want to fiddle around finding the fields of an Internet packet. The 128-bit choice seemed excessive for an experiment that involved only a few networks to begin with. By this time, the research effort had reached its fourth iteration (the IP layer protocol was called IPv4), and as program manager, I felt a need to get on with live testing and final design of TCP and IP. In lieu of consensus, I chose 32 bits of address. I thought 4.3 billion potential addresses would be adequate for conducting the experiments to prove the technology. If it worked, then we could go back and design the production version. Of course, it is now 2011, and the experiment never exactly ended.

The Internet Corporation for Assigned Names and Numbers (ICANN) succeeded Jonathan Postel as the op-

The IPV6 Internet
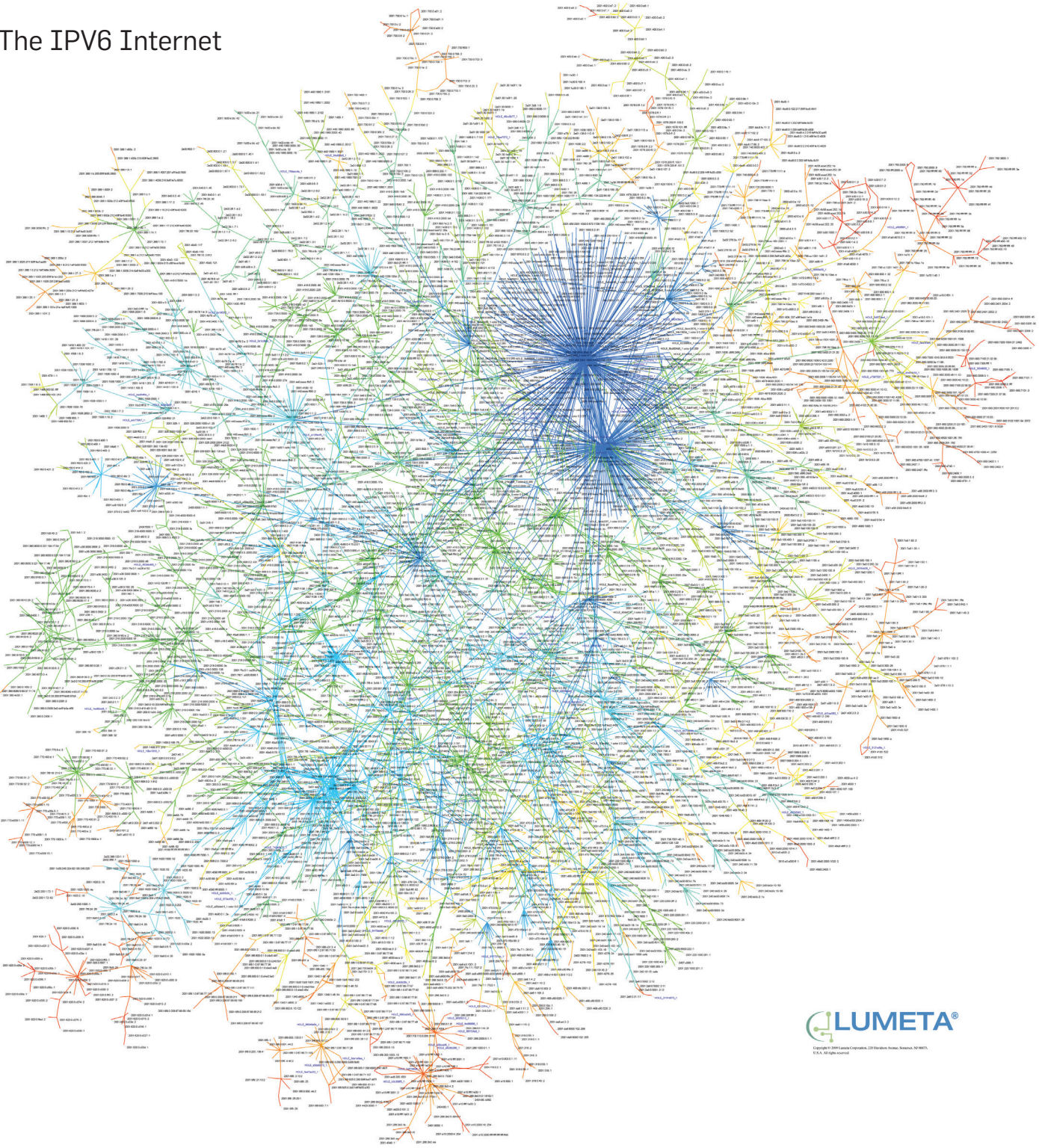
Lumeta Corporation created this map of the IPv6 Internet using data collected on Feb. 5, 2009. Each end node can represent a handful of computers on a small network, or perhaps a large company with hundreds of thousands of hosts. (http://www.lumeta.com/IPv6/)

erator of what was and still is called the IANA (Internet Assigned Numbers Authority). IANA historically allocated large chunks of address space to end users or, after the commercialization of the Internet, to Internet service providers (ISPs). With the creation of the Regional Internet Registries (for

Internet addresses), IANA typically allocated 24-bit subsets of the IP address space (sufficient for 16 million hosts) to one of the five regional registries, which, in turn, allocated space to ISPs or, in some cases, very large end users. As this article was being written, ICANN announced that it had just allo-

cated the last five of these large 24-bit chunks of space.

The Internet Engineering Task Force (IETF) recognized in the early 1990s that there was a high probability that the address space would be exhausted by the rapid growth of the Internet, and it concluded several years

of debate and analysis with the design of a new, extended address format called IPv6. (IPv5 was an experiment in stream applications that did not scale and was abandoned.) IPv6 had a small number of new features and a format intended to expedite processing, but its principal advantage was 128 bits each of source and destination host addresses. This is enough for 340 trillion trillion trillion addresses—enough to last for the foreseeable future.

The IPv6 format is not backwards compatible with IPv4 since an IPv4-only host does not have the 128 bits of address space needed to refer to an IPv6-only destination. It is therefore necessary to implement a dual-stack design that allows hosts to speak to either protocol for the period that both are in use. Eventually, address space will not be available for additional IPv4 hosts, and IPv6-only hosts will become necessary. Hopefully, ISPs will be able to implement IPv6 support before the actual exhaustion of IPv4 addresses, but it will be necessary to allow for dual-mode operation for some years to come.

World IPv6 Day is scheduled for June 8, 2011, at which time as many ISPs as are willing and able will turn on their IPv6 support to allow end users and servers to test the new protocol on a global scale for a day. The move to IPv6 is one of the most significant changes to the Internet architecture since it was standardized in the late 1970s and early 1980s. It will take dedicated effort by many to ensure that users, servers, and Internet service and access providers are properly equipped to manage concurrent operation of the old and new protocols.

Here, Thomas Limoncelli considers steps that can be taken to achieve this objective.

—*Vinton G. Cerf*

Strategies for Moving to IPv6

Someday the U.S. will run out of three-digit telephone area codes and will be forced to add a digit. As Vint Cerf explained, the Internet is facing a similar situation with its address structure. Often predicted and long ignored, the problem is now real. We have run out of 32-bit IP addresses (IPv4) and are moving to the 128-bit address format of IPv6. This section looks at some

Someday the U.S. will run out of three-digit telephone area codes and will be forced to add a digit. As Vint Cerf explained, the Internet is facing a similar situation.

strategies of organizations that are making the transition. The strategies that work tend to be those that focus on specific applications or Web sites rather than trying to convert an entire organization.

The biggest decision for many organizations is simply knowing where to begin. In this article, I consider three possible strategies.

The first scenario we present is a cautionary tale against what might be your first instinct. Though fictional, we've seen this story played out in various forms. The other two examples have proven to be more successful approaches. Knowing this, we would offer the following advice to a business contemplating the transition to IPv6: start with a small, well-defined project that has obvious value to the business.

Story 1: "Upgrade Everything!"

While having a grand plan of upgrading everything is noble and well intentioned, it is a mistake to think this is a good first experiment. There is rarely any obvious value to it (annoys management), it is potentially biting off more than you can chew (annoys you), and mistakes affect people that you have to see in the cafeteria every day (annoys coworkers).

This strategy usually happens something like this: someone runs into the boss's office and says, "Help! Help! We have to convert everything to IPv6." This means converting the network equipment, Domain Name System (DNS), Dynamic Host Configuration Protocol (DHCP) system, applications, clients, desktops, servers. It's a huge project that will touch every device on the network.

These people sound like Chicken Little claiming the sky is falling.

These people are thrown out of the boss's office.

A better approach is to go to the boss and say, "There's one specific thing I want to do with IPv6. Here's why it will help the company."

These people sound focused and determined. They usually get funding.

Little does the boss realize this "one specific thing" requires touching many dependencies. These include the network equipment, DNS, DHCP, and so on—yes, the same list of things

IPv4 & IPv6 INTERNET TOPOLOGY MAP JANUARY 2009
AS-level INTERNET GRAPH
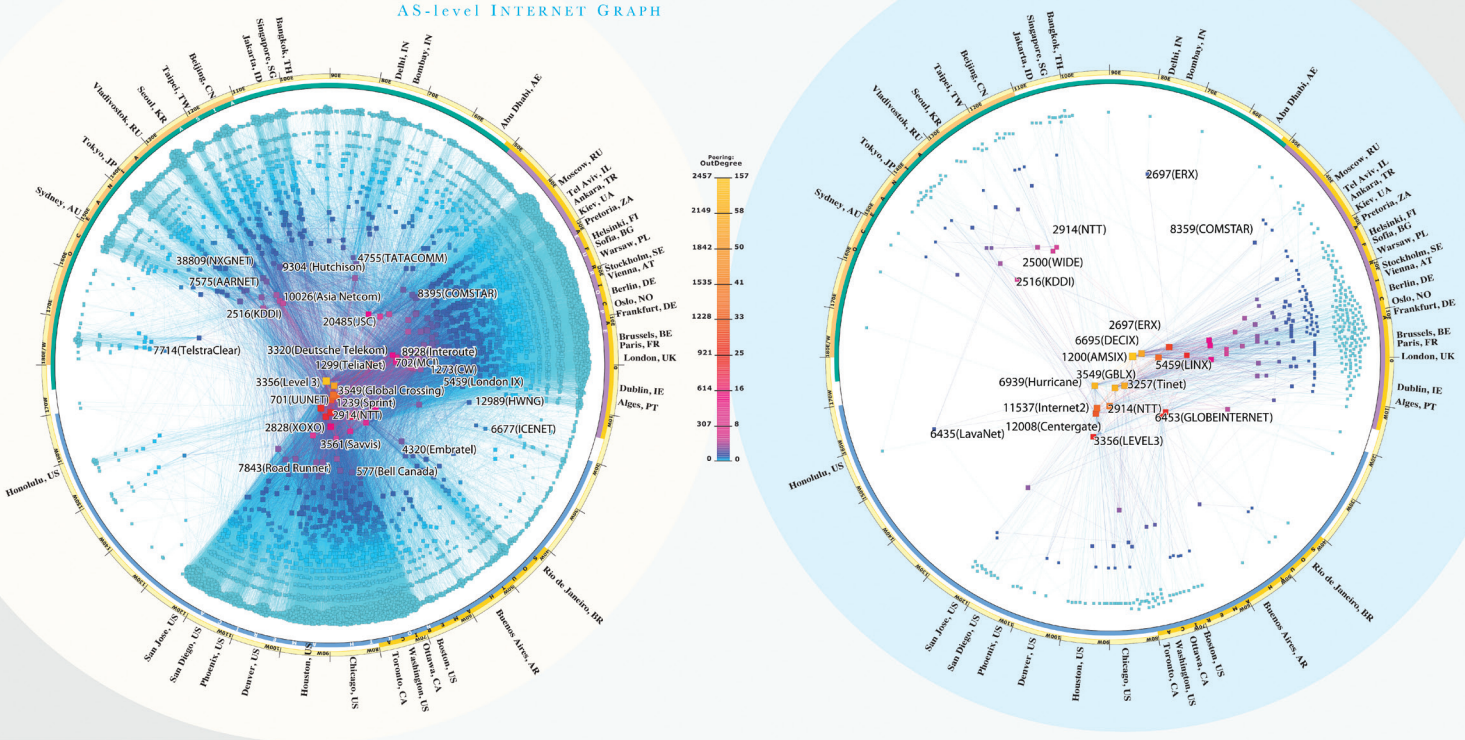
This visualization, from the Cooperative Association for Internet Data Analysis (CAIDA), represents macroscopic snapshots of IPv4 and IPv6 Internet topology samples captured in January 2009.

that Chicken Little was spouting off about.

The difference is these people got permission to do it. Which leads us to...

Story 2: Work from the outside in.

Fundamentally, this second strategy is to start with your organization's external Web presence. Usually an external Web server is hidden behind a hardware device known as a load balancer. When Web browsers connect to your Web site, they are really connecting to the load balancer. It relays the connection (being a "man in the middle") to the actual Web server. While doing that, it performs many functions—most importantly it load-shares the incoming Web traffic among two or more redundant Web servers.

In this strategy the goal is simple: upgrade to IPv6 every component on the path from your ISP to your load balancer and let the load balancer translate to IPv4 for you. Modern load balancers can receive IPv6 connections in the "front" and send IPv4 connections out the "back" to your Web servers. That is, your load balancer can be a translator that permits you to of-

fer IPv6 service without requiring you to change your Web servers. Those upgrades can be phase two.

This is a bite-size project that is achievable. It has a real tangible value that you can explain to management without being too technical: "The coming wave of IPv6-only users will have faster access to our Web site. Without this upgrade, those users will have slower access to our site because of the IPv4/v6 translators that ISPs are setting up as a Band-Aid." That is an explanation that a nontechnical executive can understand.

Management may be unconvinced that there will be IPv6-only users. Isn't everyone "dual stack" as previously described? Most are, but LTE ("4G") phones and the myriad other LTE-equipped mobile devices will eventually be IPv6-only. ARIN (American Registry for Internet Numbers) advised LTE providers that IPv4 depletion is imminent and LTE providers have prepared for a day that new LTE users will be IPv6-only.[10] Obviously this new wave of IPv6-only users will want to access IPv4-only sites, so the carriers are setting up massive farms of servers to do the translation.[4]

There are two problems with this.[3] First, the translation is expected to be slow. Second, geolocation will mistakenly identify users as being where the server farm is, not where the user is. That means if your Web site depends on advertising that is geotargeted, the advertisements will be appropriate for the location of server farm; not where your users are. Since LTE is mostly used in mobile devices, this is particularly pressing.

Therefore, if your company wants to ensure that the next million or so new users have fast access to your Web site,[9] or if revenue depends on advertising, then management should be concerned.

Most CEOs can understand simple, nontechnical, value statements such as "New people coming onto the Internet will have better access to our site," or "It is required to insure that high-paying, geotargeted advertisements, continue to work."

A project like this requires only modest changes: a few routers, some DNS records, and so on. It is also a safe place to make changes because your external Web presence has a good, solid testing regimen for making changes

VISUALIZATION COURTESY OF CAIDA

in a test environment, which gate to a QA environment before hitting the production environment. Right?

Once IPv6 is enabled from the ISP to the load balancer, and the load balancer is accepting IPv6 connections but sending out IPv4 connections to the Web farm, new opportunities present themselves. As each Web server becomes IPv6 ready, the load balancer no longer needs to translate for that host. Eventually your entire Web farm is dual stack. This technique provides a throttle to control the pace of change. You can make small changes, one at a time, testing along the way.

In doing so you will have upgraded the routers, DNS server, and other components. While your boss would shriek if you had asked to change every layer of your network stack, you have essentially done just that.

Of course, once you have completed this and shown that the world didn't end, developers will be more willing to test their code under IPv6. You might need to enable IPv6 to the path to the QA lab. That's another bite-size project. Another path will be requested. Then another. Then the LAN that the developers use. Then it makes sense to do it everywhere. You've now achieved the goal of the person from Story 1, but you've gotten management approval.

During Google's IPv6 efforts we learned that this strategy works really well. Most importantly we learned that it turned out to be easier and less expensive than expected.[8]

Story 3: "One Thing"
This story involves a strategic approach in which an organization picked a single application—its "one thing"—and mounted a focused effort to move it to IPv6. Again, being focused appealed to management and still touched on many of the upgrades requested by our Chicken Little.

Comcast presented a success story at the 2008 Google IPv6 Symposium,[6] demonstrating how it chose one strategic thing to upgrade: the set-top box management infrastructure. Every set-top box needs an IP address so the management system can reach it. That's more IPv4 addresses than Comcast could reasonably get allocated. Instead it used IPv6. If you get Internet service from Comcast, the set-top box

on your TV set is IPv6 even though the cable modem sitting next to it is providing IPv4 Internet service.[7] Comcast had to get IPv6 working for anything that touches the management of its network: provisioning, testing, monitoring, billing. Wait, billing? Well, if you are touching the billing system, you are basically touching a lot of things. Ooh, shiny dependencies. (This is why we put "one thing" in quotes.) The person from Story 1 must be jealous.

At the same symposium Nokia presented a success story that also involved finding "one thing," which turned out to be power consumption. Power consumption, you say? Yes. Its phones waste battery power by sending out pings to keep the NAT (network address translation) session alive. By switching to IPv6, Nokia did not need to send out pings—no NAT, no need to keep the NAT session alive. Its phones can turn off their antennae until they have data to send. That saves power. In an industry where battery life is everything, any executive can see the value. (A video from Google's IPv6 summit details Nokia's success.[5])

In the long term we should be concerned with converting all our networks and equipment to IPv6. The pattern we see, however, is that successful projects have selected one specific thing to convert, and let all the dependencies come along for the ride.

Summary
IPv4 address space is depleted. People who have been ignoring IPv6 for years need to start paying attention. It is real—and really important. IPv6 deployment projects seem to be revealing two successful patterns and one unsuccessful pattern. The unsuccessful pattern is to scream that the sky is falling and ask for permission to upgrade "everything."

The lessons we have learned:

1. Proposals to convert everything sound crazy and get rejected. There is no obvious business value in making such a conversion at this time.

2. Work from the outside in. A load balancer that does IPv6-to-IPv4 translation will let you offer IPv6 to external customers now, gives you a "fast win" that will bolster future projects, and provides a throttle to control the pace of change.

3. Proposing a high-value reason (your "one thing") to use IPv6 is most likely to get management approval. There are no simple solutions, but there are simple explanations. Convert that "one thing" and keep repeating the value statement that got the project approved, so everyone understands why you are doing this. Your success here will lead other projects.

For a long time IPv6 was safe to ignore as a "future requirement." Now that IPv4 address space is depleted, it is time to take this issue seriously. Yes, really. ⓒ

Related articles
on queue.acm.org

DNS Complexity
Paul Vixie
http://queue.acm.org/detail.cfm?id=1242499

A Conversation with Van Jacobson
http://queue.acm.org/detail.cfm?id=1508215

Principles of Robust Timing over the Internet
Julien Ridoux, Darryl Veitch
http://queue.acm.org/detail.cfm?id=1773943

References
1. Cerf, V., Dalal, Y., Sunshine, C. 1RFC 675. Specification Internet Transmission Control Program, 1974.
2. Cerf, V., Kahn, R.E. A protocol for packet network intercommunication. IEEE Transactions on *Communications 22*, 5 (1974), 637–648.
3. Donley, C., Howard, V., Kuarsingh, V., Chandrasekaran, A. and Ganti, V. Assessing the impact of NAT444 on network applications; http://tools.ietf.org/html/draft-donley-nat444-impacts-01/
4. Doyle, J. Understanding carrier-grade NAT (2009); http://www.networkworld.com/community/node/44989.
5. Google IPv6 Conference. IPv6, Nokia, and Google (2008); http://www.youtube.com/watch?v=o5RbyK0m5OY.
6. Google IPv6 Implementors Conference; http://sites.google.com/site/ipv6implementors/2010/agenda.
7. Kuhne, M. IPv6 monitor: an interview with Alain Durand (2009); https://labs.ripe.net/Members/mirjam/content-ipv6-monitor.
8. Marsan, C.D. Google: IPv6 is easy, not expensive (2009); http://www.networkworld.com/news/2009/032509-google-ipv6-easy.html.
9. Miller, R. The billion-dollar HTML tag (2009); http://www.datacenterknowledge.com/archives/2009/06/24/the-billion-dollar-html-tag/.
10. Morr, D. 2010. T-Mobile is pushing IPv6. Hard (2010); http://www.personal.psu.edu/dvm105/blogs/ipv6/2010/06/t-mobile-is-pushing-ipv6-hard.html.

Vinton G. Cerf is Google's vice president and chief Internet evangelist. As one of the "Fathers of the Internet," Cerf is the codesigner of the Internet's TCP/IP protocols and architecture. He holds a B.S. degree in mathematics from Stanford University and M.S. and Ph.D. degrees in computer science from UCLA.

Thomas A. Limoncelli is an author, speaker, and system administrator. His books include *The Practice of System and Network Administration* (Addison-Wesley, 2007) and *Time Management for System Administrators* (O'Reilly, 2005). He works at Google in New York City and blogs at http://EverythingSysadmin.com/