The Internet Protocol Journal

September 2005 Volume 8, Number 3

*A Quarterly Technical Publication for Internet and Intranet Professionals*

## In This Issue



You can download IPJ back issues and find subscription information at: **www.cisco.com/ipj**

## From The Editor

Protocol transitions are never easy, particularly not when they involve something so fundamental as the *Internet Protocol* (IP). Organizations considering a move to IPv6 must consider many factors when deciding on the timing for such a deployment. One of the first questions that arises is: "When will the IPv4 address space actually run out, forcing us to use IPv6 instead ?" That question is not a new one; it was being asked in the early 1990s when the IPv6 effort was started. Several factors, such as the deployment of *Classless Interdomain Routing* (CIDR) and *Network Address Translation* (NAT), have "delayed the inevitable," and perhaps led to some complacency on the part of network operators. In this issue we examine the topic of IPv4 address space depletion in more detail. Our main article is by Tony Hain, and it is followed by a response from Geoff Huston and a roundtable discussion with Tony, Geoff, Fred Baker, and John Klensin. We would also like to hear from our readers on this important topic. Please send your comments to **ipj@cisco.com**.

As an old-time network and UNIX user, I am a big fan of tools that allow simple terminal access to remote host computers. My "Internet career" started in Norway in 1976, where I used *Telnet* to access machines in California through the ARPANET. Today, I still access remote servers through a simple terminal interface, but Telnet has been replaced by the *Secure Shell* (SSH) *Protocol* for all the obvious security reasons. SSH is used not just for terminal traffic—it also can be configured to provide secure tunnels to a variety of services such as Webpages and file transfers. Ronnie Angello explains the details in our second article.

In order to better serve our readers, we will be conducting an IPJ Reader Survey in the near future. Details will be available on our Website at **www.cisco.com/ipj**. We appreciate your cooperation in completing the survey.

Finally, let me remind you to visit the IPJ Website and update or renew your subscription.

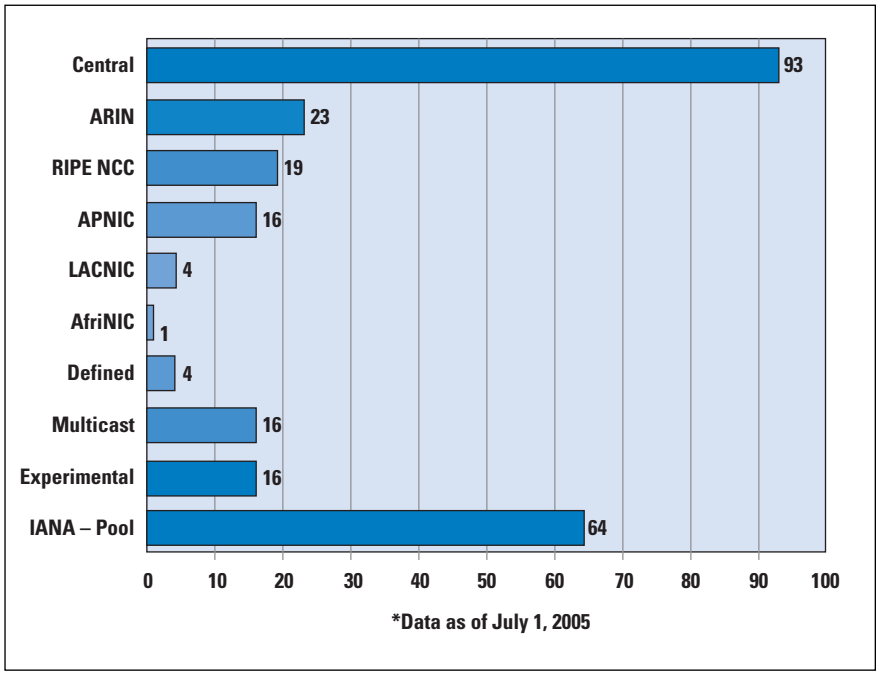
*—Ole J. Jacobsen, Editor and Publisher*
**ole@cisco.com**

# A Pragmatic Report on IPv4 Address Space Consumption

*by Tony Hain, Cisco Systems*

When I interact with people from all around the world discussing IPv6, there continue to be questions about the projected lifetime for IPv4. This article presents consumption rate and lifetime projections based on publicly available *Internet Assigned Numbers Authority* (IANA) data. In addition, there is discussion about why the widely quoted alternative projection may be flawed, thus leading everyone to believe we have much more time than we might.

*Figure 1: IANA /8 Allocations*




## Allocations

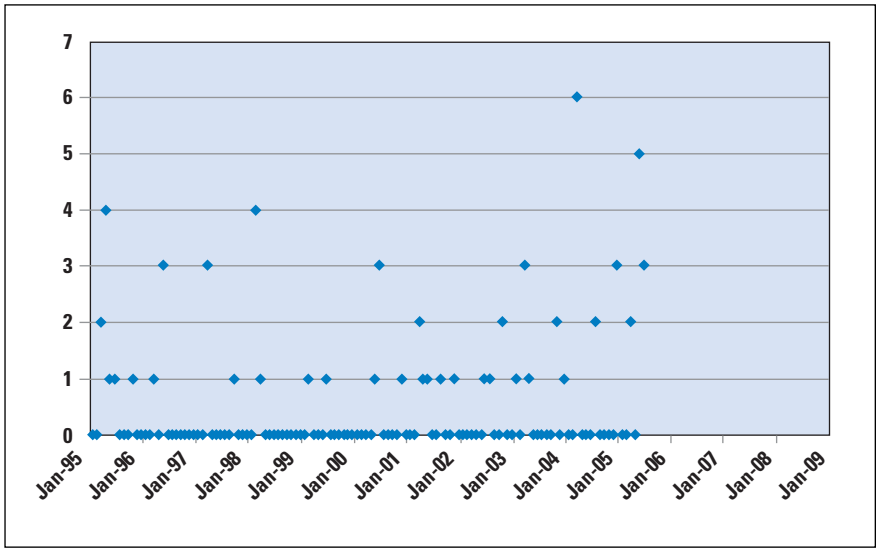
The chart in Figure 1 shows the distribution of all 256 IANA /8 allocation units in IPv4[1] as of July 1, 2005. The Central registry represents the allocations made prior to the formation of the *Regional Internet Registries* (RIRs). ARIN (North America)[2], RIPE NCC (Europe)[3], APNIC (Asia/Pacific)[4], LACNIC (Latin America)[5], and AfriNIC (Africa)[6] are the organizations managing registrations for each of their respective regions. RFC 3330[7] discusses the state of the Defined and Multicast address blocks. The Experimental block (also known as *Class E*—RFC 1700[8]) was reserved, and many widely deployed IPv4 stacks considered its use to be a configuration error. The bottom bar shows the remaining useful global IPv4 pool. To be clear, when the IANA pool is exhausted there will still be space in each of the RIR pools, but by current policy[9] that space is expected to be only enough to last each RIR between 12 and 18 months.

The projection published at **http://bgp.potaroo.net/ipv4**[10] is often quoted as the definitive reference for IPv4 consumption. This report presents a viewpoint consistent with that author's long-standing position that we do not need to change from IPv4 to IPv6 anytime soon, thus showing an extended lifetime for IPv4.

The approach used in the potaroo report is to take the simple exponential fit to the allocation data since 1995. As discussed later in this article, this approach includes the effects of the policy shift to *Classless Interdomain Routing* (CIDR) and subsequent digestion of prior allocations, the lull in IANA allocations to the RIRs for two full years, as well as the fact that the model used does not generate a particularly close fit to the actual run rate over the 10-year period.

Although this author agrees that over very long timeframes (20–50 years) there will be substantial variations in the consumption rate for any number of reasons, the opportunity for events that would reduce the recent rate in the timeframe of the remaining IANA IPv4 pool is not evident. That said, there are numerous things that could increase the consumption rate and exhaust the pool even sooner than this projection.

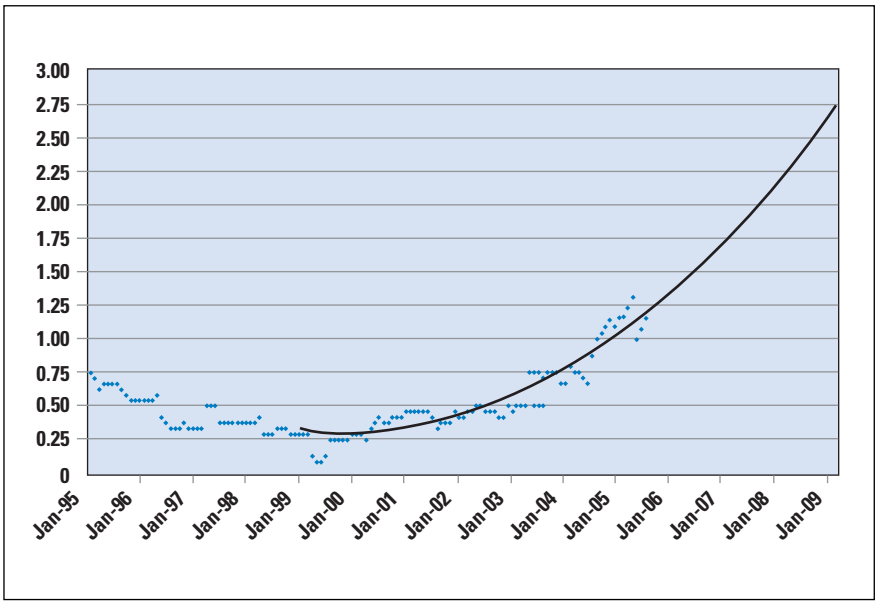
*Figure 2: IANA Allocations to RIRs — Raw /8 Allocations per Month*




The graph in Figure 2 shows the raw per-month IANA allocations since 1995. In raw form it is difficult to discern the trend, or develop an expectation about the overall lifetime of the remaining pool.

Taking a closer look at Figure 3, smoothing the data with a 24-month sliding window (averaging over 12 months back and 12 months forward) exposes the underlying reality that the combined rate and quantity of /8 allocations has been steadily accelerating since 2000 (the graphs for 12-, 18-, and 24-month sliding windows show the same fundamental trend). Though a few of the allocations may arguably have been "one-time" events, those are lost as statistically insignificant in the extended and continuing overall growth rate.

*Figure 3: IANA Allocations to RIRs — Sliding-Window 24-Month Average*




Taken by itself, the most recent allocation rate (22 /8s over the 18 months leading up to July 1, 2005) suggests that the remaining pool of 64 /8s will be exhausted in about 5 years, even if growth abruptly flattens out to hold around 1 /8 per month. Unfortunately at this point there is no reason to believe the allocation rates will slow or that they will turn downward again. All the gain of CIDR absorbing the pre-1995 allocations has already been incorporated, and there is no obvious economic bubble that might burst to lower demand within the time window of the remaining pool.

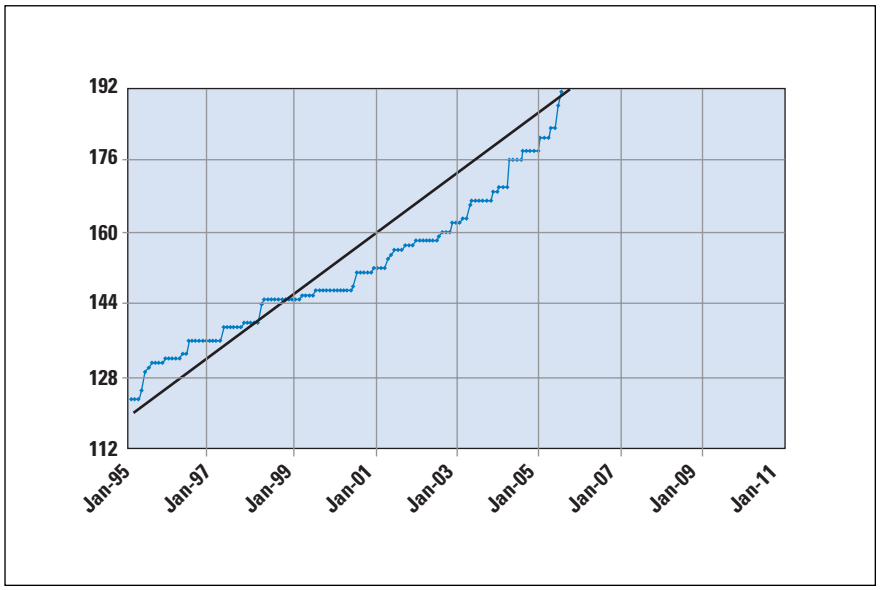
To the contrary, the following URL shows potential demand (to bring developing countries up to just 20-percent connectivity, which is half of what the existing Internet world enjoys today) that will swamp the remaining pool, even in the face of much stricter allocation policies.
**`http://www.nav6tf.org/documents/e-Nations-data.pdf`**

So this view of the sustained trend in allocation growth rate suggests that the lifetime of the remaining central IPv4 pool is 4 years +/-1.

### Projections

Differing from recent articles and section 5 of the report at **`http://bgp.potaroo.net/ipv4`** that hint at linearity in growth, Figure 4 shows that the raw data after 1995 is clearly nonlinear. It starts with a decelerating rate through mid-1998 as the pre-1995 allocations were absorbed (precipitated by the allocation policy shift from class-based to CIDR), followed by a 2-year lull (only 1 /8 per year), then a return to accelerating growth from mid-2000 onward.

*Figure 4: IPv4 Lifetime Projection — Non-Linear Nature of Raw Data*




This suggests that using the past 10-year IANA data is likely to skew the projection toward a much longer period than the recent allocation data would support. Although a longer lifetime projection helps to avoid short-term panic, it can mislead people into believing there is substantial time to worry about this later, resulting in a much bigger problem when reality blindsides everyone sooner than they expected.

*Figure 5: IPv4 Lifetime Projections — Order-N Polynomials, Post-2000 History Basis*

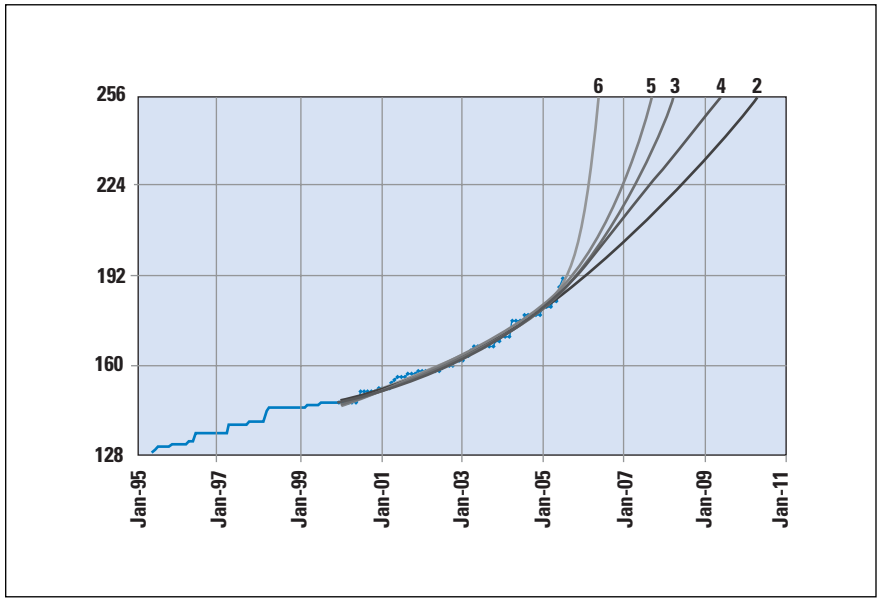
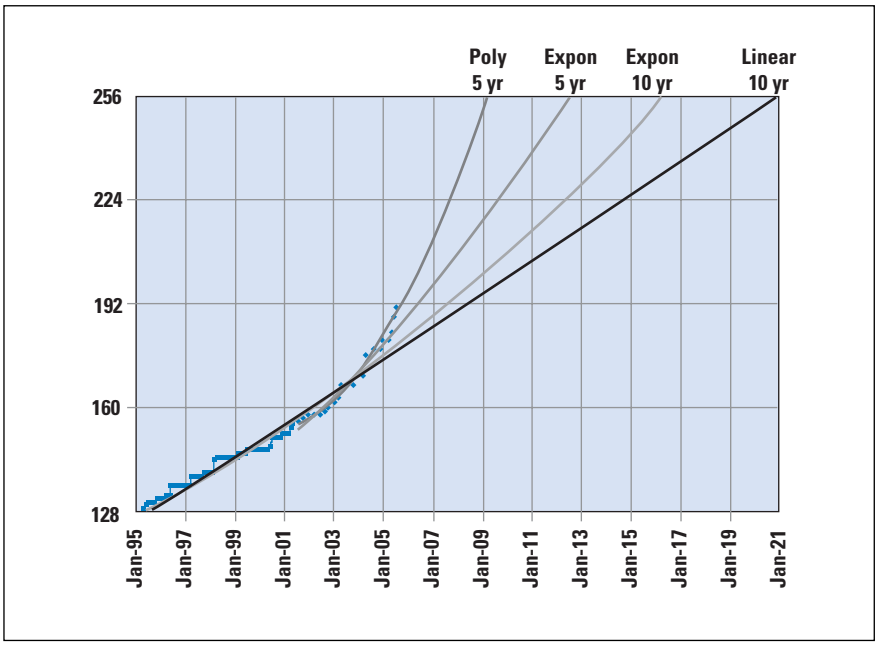


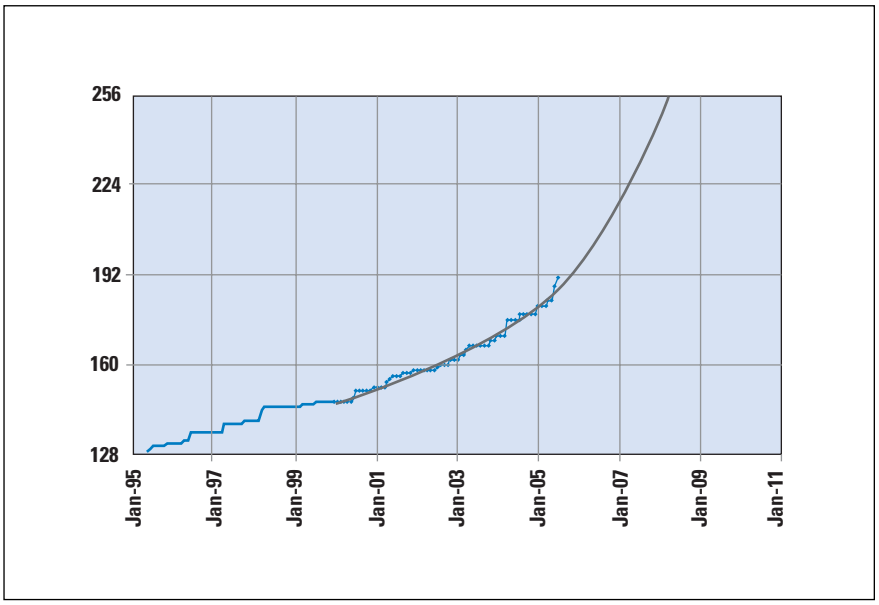
*Figure 6: IPv4 Lifetime Projections — Polynomials and Exponentials*




As in any statistical endeavor there are many ways to evaluate the data. The various projections in Figures 5 and 6 show different mathematical models applied to the same raw data. Depending on the model chosen, the nonlinear historical trends in Figure 6 covering the last 5- and 10-year data show that the remaining 64 /8s will be allocated somewhere between 2009 and 2016, with no change in policy or demand (though as discussed previously there are already reasons to err toward 5-year-based nonlinear models).

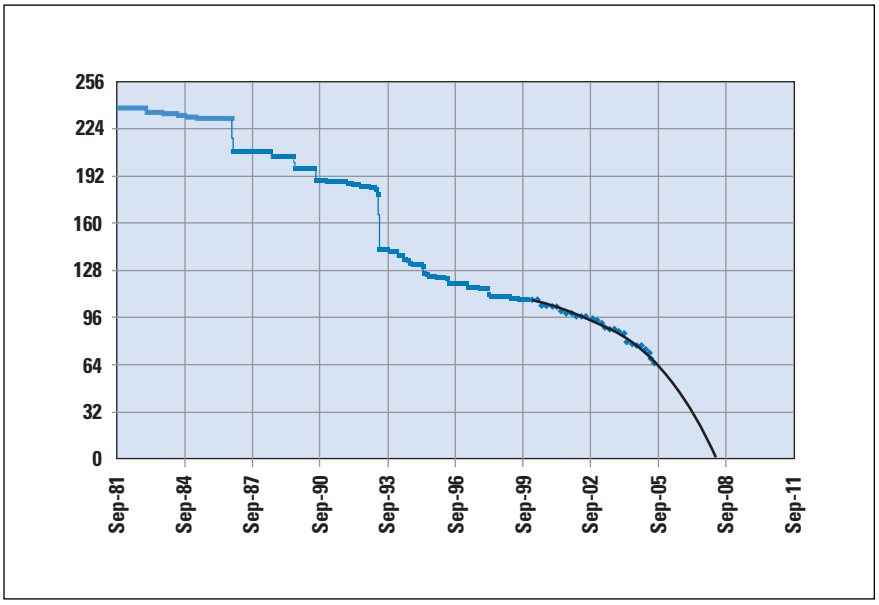
Adding to that, policy is continually changing. ARIN, for example, has recently clarified its policy allowing organizations that demonstrate they have exceeded the capacity of the private space defined in RFC 1918 to acquire IPv4 address blocks from the remaining public pool, even when it is clear these allocations will never be announced to the global Internet. The other regions already have similar policies or are likely to follow suit because the most vocal members of the RIR community have adamantly commented against expanding the private IPv4 range. This policy approach coupled with persistent demand means the actual run rate is going to continue increasing as the large organizations begin consuming public space where they had been using private to support their network growth. For example, one large enterprise has steady growth over 1 percent per month, which currently requires an efficiently managed /12 per year for its expanding network. The enterprise is less than a year from exhausting all the space provided in RFC 1918, so it was very interested in the ARIN policy that allows the enterprise to continue growing through public space. Additionally, multiple commercial service providers expect to reach the capacity of the 1918 space within 12 to 18 months, just supporting management addresses on their existing devices. This does not take into consideration their pending deployment of new services, which they expect will use several new IPv4 addresses per device with marketing targets measured in multiple millions of units.

*Figure 7: IPv4 Lifetime Projection — 5-Year History Basis*




The graph in Figure 7 hints at the likely outcome as word spreads about the perception of policy liberalization and the demonstrable exhaustion of the remaining global IPv4 pool landing within the *return-on-investment* (ROI) period for new equipment. It is based on the same raw historical data as the frequently quoted long-term projection on potaroo's Figure 2.4, but the more aggressive fit on the most recent data set describes a significantly higher consumption rate and shorter lifetime for the remaining pool.

*Figure 8: IPv4 /8 Pool — 5-Year History-Based Projection*




The graph in Figure 8 provides the exhaustion perspective, showing the entire address pool from the publication of IP Version 4[11] (note that data prior to 1995 is accurate as to where it was allocated, but with very coarse granularity as to exactly when). The projection curve is based on the IANA allocations from January 2000 onward.

Only time will tell which projection is correct, but it will already take a fairly significant stalling event to slow consumption and put the actual allocation curve back on the extended track in potaroo's Figure 2.4.

### Reserved Space

There are occasionally arguments that the 16 /8s reserved in the experimental space could be used. Although this is likely to be possible for some IP stack implementations, for others it is not. At a minimum, some quick tests show that Windows 95 through Windows 2003 Server systems consider that block to be a configuration error and refuse to accept it. The operational ability to restrict the space to a select stack implementation is limited, and the amount of space there does not really help even if deployment and operations were trivial. Assuming the sustained growth trend in allocations continues, by the time the remaining 64 /8s in the IANA pool are finished the rate would be approaching 3 /8 allocations per month, so the entirety of the old Class E space would amount to about 6 months of run rate.

### Reclaiming Allocations

Another debate occasionally resurfaces about reclaiming some of the early allocations to further extend the lifetime of IPv4. Hopefully this article has shown that the ROI for that approach is going to be extremely low. Discussions around the Internet community show there is an expectation that it will take several years of substantive negotiation (in multiple court systems around the globe) to retrieve any /8s. Then following that effort and expense, the likelihood of even getting back more than a few /8 blocks is very low. Following the allocation growth trend, after several years of litigation the result is likely to be just a few months of additional resource added to the pool—and possibly not even a whole month. All this assumes IANA does not completely run out before getting any back, because running out would result in pent-up demand that could immediately exhaust any returns.

### Summary

*Network Address Translation* (NAT) and CIDR did their jobs and bought the 10 years needed to get IPv6 standards and products developed. Now is the time to recognize the end to sustainable growth of the IPv4-based Internet has arrived and that it is time to move on. IPv6 is ready as the successor, so the gating issue is attitude. When CIOs make firm decisions to deploy IPv6, the process is fairly straightforward. Staff will need to be trained, management tools will need to be enhanced, routers and operating systems will need to be updated, and IPv6-enabled versions of applications will need to be deployed. All these steps will take time—in many cases multiple years. The point of this article has been to show that the recent consumption rates of IPv4 will not be sustainable from the central pool beyond this decade, so organizations would be wise to start the process of planning for an IPv6 deployment now. Those who delay may find that the IANA pool for IPv4 has run dry before they have completed their move to IPv6. Although that may not be a problem for most, organizations that need to acquire additional IPv4 space to continue growing during the transition could be out of luck.

## References


[1] **http://www.iana.org/assignments/ipv4-address-space**

[2] **http://www.arin.net/**

[3] **http://www.ripe.net/**

[4] **http://www.apnic.net/**

[5] **http://www.lacnic.net/**

[6] **http://www.afrinic.net/**

[7] **http://www.rfc-editor.org/rfc/rfc3330.txt**

[8] **http://www.rfc-editor.org/rfc/rfc1700.txt**

[9] **http://www.rfc-editor.org/rfc/rfc2050.txt**

[10] **http://bgp.potaroo.net/ipv4**

[11] **http://www.rfc-editor.org/rfc/rfc791.txt**

[12] Geoff Huston, "The Myth of IPv6," *The Internet Protocol Journal,* Volume 6, No. 2, June 2003.

[13] Geoff Huston, "IPv4: How long do we have?," *The Internet Protocol Journal,* Volume 6, No. 4, December 2003.


## Another Perspective

*Ed.: We asked Geoff Huston to provide some feedback on this article and he responded with the following:*

Dear Editor,

There are, of course, many ways to undertake predictions, and over the millennia humanity has explored a wide diversity of them. In every case the challenge is to make predictions that end up being closely correlated to the unfolding story, and of course hindsight is always the harshest judge of such predictions.

Tony's work takes a different base point for making the projection from earlier work that I did in this area. Tony looks at the rate of allocation from the IANA to the RIRs, and bases his predictions on the trends visible in that time series of data. By contrast, I used the assumption that assigned addresses are destined for use in the public IPv4 Internet, and I used the trends visible in the amount of advertised address space as the basis for the predictions of consumption.

One of the more interesting data artifacts is the first-order differential of the rate at which the span of addresses announced in the IPv4 public Internet has increased over time.
(Figure 4.4 of **http://bgp.potaroo.net/ipv4/**)

One interpretation of this data is that there are two phases of recent activity: prior to March 2003 and post-March 2003. Prior to March 2003 the longer-term address growth rate was the equivalent of some 3.5 /8 blocks per year.

Post-March 2003 we see a different consumption growth rate, fluctuating between 5 and 8 /8s per year, with a mean value of some 7.5 /8s per year. There is no strongly obvious longer-term compound growth rate visible in this view of the data. Given some 64 /8s remaining in the IANA pool as of July 2005 and a base consumption rate of a mean of 7.5 /8s per year, the simple division yields 8.5 years, or 2014 as the time of forecast exhaustion of the IANA address pool. At that point the RIRs will be holding about 25 /8 blocks in their unallocated pools, and a further two years of allocations could be made from these pools.

So I would offer the view that the post-2003 data offers a perspective of exhaustion of the unallocated address pools in 2016, with the caveat that such a prediction assumes that the current address demand levels will continue, the actions of industry players are invariant, and the current address allocation policies will continue as they are at present.

Of course these three caveats represent relatively major assumptions about the future—and are perhaps unlikely to happen. It is likely that there will be changes in all these factors in the coming years, and these will obviously impact these predictive models.

To summarize, I observe that these different predictive approaches yield slightly different outcomes, but not beyond any reasonable error margin for predictions of this nature. Sometime in the forthcoming 5 to 10 years the current address distribution policy framework for IPv4 will no longer be sustainable for the current industry address consumption model because of effective exhaustion of the unallocated address pool.

When looking at this prediction from the perspective of the service provider enterprise, the prediction can be re-expressed as a problem relating to investment lifecycles. The ISP industry and the enterprise sector have already made considerable investments in IPv4-based infrastructure in equipment, infrastructure, and operational capability, and we are seeing some considerable reluctance to add to this with additional investment into IPv6 capability at this time. The direction of the use of various forms of NAT-based approaches and increasing use of application layer gateways in the public and enterprise environments can be seen as an effort to extend the lifetime of the existing infrastructure investment. In a volume-based market with relatively low revenue margins, this position certainly has some sound rationale from a business management perspective. But I agree with Tony here that such business approaches are ultimately short-term in nature, because they do not allow IPv4 to encompass indefinite further decades of Internet growth in a silicon-dense world.

However, in terms of understanding the next few years of a process of industry transition of protocol infrastructure into IPv6 deployment, perhaps the real issues here are more centered on competitive business factors and sector investment profiles than they are about detailed introspection of trends within various number series.

The numbers all indicate that this is not a matter that can be deferred indefinitely. Tony's call for some timely attention to the need to commence investment in IPv6-based service infrastructure is one that I hope the industry is listening to attentively.

*—Geoff Huston*
**gih@apnic.net**

### A Virtual Roundtable

*Ole:* Let's open this discussion on the point of measurement methods. We invited John Klensin and Fred Baker to join Geoff and Tony in the discussion at our virtual round table. (We often all see each other at IETF meetings, but there is seldom enough time to gather everyone around a real table, hence this discussion took place with a few rounds of e-mail).

*Geoff:* As I said in my response letter, Tony's work takes a different base point for making the projection from the earlier work that I did in this area. My work has focused on the trends from the addresses used in the public IPv4 Internet, and then deriving projections on consumption based on this data. It assumes that the influencing factor for address consumption is the use of addresses in the public IPv4 Internet.

*Tony:* As Geoff noted, he and I have discussed over time that we are looking at different parts of the data set and coming to different conclusions. One specific point that distorts the approaches is the time delay between IANA allocation to the RIRs and the appearance of that space for public use. In particular, his comment about 5 to 8 /8s per year is based on the delayed public use data that will eventually catch up with the fact that IANA has allocated 13 /8s just since the beginning of 2005. If the allocation rates had close to linear growth, the delay would not be a big factor. Another point of distortion is the potential for some of the allocations to never show up as publicly routed.

*Ole:* So when do we actually run out?

*Geoff:* There are many specific milestones that will pass in sequence. The unallocated address pool held by IANA will exhaust first, and then the RIR pools of unallocated data will drain. At that point there is no stream of "new" addresses to fuel further growth, and that is probably a reasonable point in time to say that we have "run out." Assuming that the current business influential factors and allocation policies remain in place, then the projection models from recent data indicate that this "run-out" date is around 2016, or some 11 years from now. Of course these are unlikely assumptions as the prospect of exhaustion draws nearer, and there may be a "last-minute rush" of address allocation requests from the service provider industry that could draw in that projected "run-out" date. Such additional consumption pressures are difficult to factor in to trend-based predictive models, of course. It is also conceivable that the industry could shift its attention almost entirely to IPv6-based protocol infrastructure in the coming years, in which case the "run-out" projection for IPv4 would extend out further in time simply because of the translation of the consumption activity to the IPv6 address pool.

*Tony:* As I noted early on in my article, there will still be pool available at each of the RIRs when the IANA pool that I focused on is exhausted. In the past I have said we would never completely run out because nobody could afford that last address, but in light of the accelerating consumption of IPv4 coupled with the less-than-aggressive deployment of IPv6, I can see how the pool might actually run dry.

*John:* In practical terms, the point at which one has "run out" of address space is not tied to being the last applicant to the RIRs for an address pool. I have suggested that point will never arise: the RIRs (and, to the extent to which the *Internet Corporation for Assigned Names and Numbers* [ICANN] can make decisions, the IANA), will continually recalibrate policies to prevent "running out." Of course the inevitable consequence of those recalibrations is that, although one does not need to worry about approaching an RIR and being told "no space left," the combination of monetary, justification, and general aggravation costs is such that one does not even want to contemplate being the applicant for the next-to-last available block. That reasoning says that looking at the date on which near exhaustion is reached is relatively uninteresting. The more important question is when one enters the end game for IPv4 space because, as soon as the end game begins, the space is essentially exhausted.

I suggest that the criterion for entrance into the end game is not measured statistically but by looking at the point at which one needs to start designing networks and subnets, not in a way that is optimal from a network architecture or network management and growth standpoint, but in order to conserve address space and/or to avoid extended discussions with applicable RIRs (or one's ISP that deals with the RIR). From that point of view, we have already run out, and probably ran out a couple of years ago. Every time someone who has multiple machines is pointed to private address space because of a presumed shortage, it is an indication that we have already run out of space. Every time China manages to make a successful political point—regardless of the country's actual internal dynamics and economics—about its inability to get addresses for its population, it is an indication that we have already run out of address space. Every time an ISP decides to use private space to manage its backbone, it is an indication that we have already run out of address space.

*Fred:* I have made the same point, from a point of view of economics. In essence, when a commodity is common and demand is low, there are calls to squander it because it costs nothing—something one hears a lot of in the IPv6 community. When supply and demand are comparable, a market develops, and I need to tell you that I certainly pay for the IPv4 addresses at *my* house. When demand outstrips supply, we enter a regulated market of some kind, and our current allocation policies certainly reflect a regulated market. The step after a regulated market is a black market, and it is not too hard to find that either.

*John:* Actually, in our present situation, there is an intermediate step before things deteriorate completely into a black market. Although it is unlikely that any significant fraction of the early IPv4 academic, research, or commercial allocations could be recovered and reused, there are governmental allocations that might be recovered under significant political pressures. Unfortunately, in addition to politicizing the allocation process much more than we have seen so far, such moves might push the present users of those allocations toward NATs in ways that would make the ultimate transition to IPv6 more difficult while not gaining very much additional time for the IPv4 space.

*Tony:* Political pressure or not, simple logistics argues against this. Given the rate of growth in consumption, any reclaimed government space would be consumed in substantially less time than it would take to rebuild their network and release it. Even a small network sitting on a /16 would take at least a year to release that much space, and at the current spot on the escalating curve that /16 represents around 2 hours of IANA run rate. Getting back a whole /8 would logistically take several years, and then at that point on the curve the result would be about a week of run rate. If several of these government organizations have a mesh of direct interactions and head down the same path, the resulting overlap in the private address space would require creating a complex NAT system worthy of a Nobel Prize. Reclamation is a nice bar-room debate topic, but the return on investment is extremely low. If an organization were to consider rebuilding its network to release an IPv4 allocation, it would make much more sense for that organization to rebuild it as IPv6 than to move publicly addressed nodes behind a NAT.

*Geoff:* It would be strongly preferred by all, I would suggest, that the "black market" option be avoided. If the consequence of the exhaustion of the unallocated pool of IPv4 addresses is the trading of already-allocated IPv4 addresses, then a responsible way for the industry to support that scenario is to encourage such a market to operate with the support of some form of "clear title" that could legitimate trading transactions. Without structure and stability in a trading market, the value of the trade is meaningless, and in this case the potential for chaos in the network itself is undeniable.

*Fred:* We are in fact starting to see networks designed to be IPv6-only or IPv6-dominant (the latter being a network that might use IPv4 internally but offer only IPv6 services to some or all of its customers) in China, Japan, and other places. The economic argument is the one these operators are primarily giving—they state that they see a roadmap to the number of addresses that they need in IPv6, while in IPv4 they are significantly constrained. This sounds to me a lot like John's comments about network design, but the other way—rather than designing their networks to what they perceive as IPv4 addressing policy limitations, they are choosing a path that they perceive as giving them options.

We also see evidence of networks designing themselves to the limits of address allocation in IPv4, usually using multiple layers of NATs. For quite a while, for example, China Unicom used multiple layers of NAT in order to work around what the company felt was a deficiency in its ability to get IPv4 addresses from its national registry. As I understand it, the company has changed its strategy to include getting IPv4 address allocations directly from APNIC, and at the same time to deploy an IPv6 network in parallel to move away from IPv4 dependence.

*John:* There is another factor at work in this. Transitions are never free. If we are going to design and build out a substantially new network, we are rapidly reaching the point—some would say that we have reached it already—at which it is cheaper to design and build that network for IPv6, making whatever arrangements are needed at its interconnection points with IPv4 networks, than to build in IPv4 and face a transition later. As those decisions are increasingly made, it may both reduce pressure on new IPv4 allocations and create free pools of IPv4 space that could be recovered and reused. For example, the U.S. Department of Defense (DoD) has announced a fairly aggressive schedule for moving to IPv6. If they meet that schedule and were then willing to free up the IPv4 space that they would presumably no longer be using, it would free up the equivalent of several /8s. While I agree with Tony that this hypothetical case would be unlikely to make any significant difference in the long run, it illustrates another difficulty with trying to make assertions about what is happening by statistical projections alone.

*Ole:* It is frequently stated that North America is immune to the address exhaustion problem.

*Tony:* Well despite persistent rumors and press statements to that effect, ARIN continues to consume about 30 percent of the annual allocation from IANA. If the past allocations were sufficient to stave off global exhaustion, why the continued consumption? In any case, when the central pool is exhausted the North American region will be in the same situation as everyone else—unable to expand or acquire new IPv4 addresses.

*Geoff:* We are seeing growth in Internet-based services in all regions of the industry, including North America. And network growth needs to be fueled by network addresses. We are seeing a combination of a continued demand for further addresses, and the use of various forms of network configurations that attempt to make the most efficient use of already-allocated addresses. There is little data to suggest that any region, including that of North America, is in a position of immunity from these growth-related factors.

*Ole:* There is widespread opinion that NAT will solve the problems for a long time to come.

*Geoff:* The ISP industry certainly has made considerable investments there, and many millions of end users today use the Internet behind NAT devices. Given the size of this investment and the factors of inertia in large-scale service markets, it is reasonable to predict that NATs will be around for quite some time. But NATs add cost to network services. If we are talking about a network that is restricted to servicing the communications needs of people, then this is a relatively high-value activity, and the additional costs of the deployment of NATs are being absorbed within the cost base of the network service economy. And for such human activity-based services this may well continue for some time, given the existing levels of industry investment in service infrastructure that includes the use of NATs. Certainly any new application that is adopted by the Internet user population needs to work across a wide variety of NAT configurations. From this perspective it is likely that IPv4 and NATs will continue to be part of the Internet landscape for a long time to come.

But although this approach has the potential to service a portfolio of service markets for some time to come, it cannot service all forms of service markets—not in the future nor even today. It does not solve all the "problems" and certainly does not encompass all the opportunities that the Internet offers. The potential of IPv6 is one that includes an address span designed to match the full potential of the volume-driven silicon industry, both now and in a future that extends out for many decades to come. One likely scenario for IPv6 is in servicing a truly massive device-dense environment. This scenario encompasses far more than services that are primarily directed at human end users. And the associated service market will be more akin to that of a relatively undifferentiated commodity market, where simplicity and low cost are the dominant service provider discriminants. Because of their additional complexity and associated incremental cost, NATs are marginalized in such commodity markets directed at servicing device density, and it is there that the true leverage of the IPv6 address span becomes a major influential factor.

*Tony:* As Geoff notes, NAT has been widely available and deployed globally over the same timeframe as the recent consumption. Yet the accelerating growth trend continues, consuming to the point where only 25 percent of the total IPv4 space remains available. Although NAT does slow the rate of public address consumption from what it might otherwise be, it creates more problems than it solves. Geoff also raises the economic investment in NAT to date, which is an interesting contrast to many complaints I hear about the cost of deploying IPv6. Most people who look at what it will take to deploy IPv6 in their network are very quick to dismiss this investment in the array of costs associated with NAT. Often they insist on a demonstration of value for the IPv6 investment while at the same time they refuse to allow consideration of removing their development, and ongoing operational support costs for IPv4 NAT.

Although I agree that in the interim overlap period the costs are additive, in the long term staying on the IPv4/NAT path those costs only compound, whereas on the IPv6 path they disappear. The duration of that overlap is somewhat self-controlled as a direct trade-off between the costs for running both protocols in parallel versus the costs associated with aggressively moving the end systems and applications to IPv6.

*Ole:* Another area frequently discussed on various lists is that the U.S. DoD and Federal Government mandates for service availability in 2008 are just another instance of the *Government OSI Profile* (GOSIP) and that they too will disappear.

*Tony:* What these discussions miss is that the situation is entirely different now. In the early 1990s the U.S. GOSIP effort was directed by a strong desire to consolidate the array of protocols in use at that time toward a common one. Other governments had similar efforts that led them collectively toward a suite that was developed with international governmental input. IPv4 was an alternative to the mandate with applications already supporting it, while the OSI protocols existed in some router products but did not have many applications available.

At this point the existing government networks are already consolidated, and there is no alternative. Yes, IPv6 still has fledgling application support, but the IPv4 pool is no longer a sustainable resource to draw on, and there is no other option. So the government networks either stop growing or, as the U.S. DoD and Government agencies have announced, they will move to IPv6. This implies preparing the application community to meet the impending reality.

*Geoff:* Although the strategic directions of one single—but relatively large—market player does have some bearing on the direction of the global market in Internet-based service provision, I do not see evidence that this will be sufficient to influence the entire market in any particular direction. This was certainly evident in the case of GOSIP some years ago, and continues to be an aspect of the market today. The global communications sector carries the impetus and burden of massive investment in infrastructure, process, technology, services, and consumer product portfolios. The sector has already undergone a revolutionary change with the advent of the Internet over the past decade. Doubtless there is considerable reluctance on the part of many sector players to continue to invest in further change in the protocol infrastructure of Internet-based services. On the other hand, the upheavals in the service provider sector have also eliminated much historical complacency about the stability of these markets and the adequacy of the associated service portfolio. It is reasonable to suggest that this sector is now very attentive to the prospect of expanded markets and new service opportunities that can take advantage of the existing infrastructure to create new revenue streams. So I think it is the current dynamics of the service provider sector and the potential for new service markets that would be the most persuasive factor for service providers to invest in an IPv6 protocol infrastructure.

*Ole:* Closing thoughts?

*Tony:* As I said at the end of my article, now is the time to recognize that we have reached the end of sustainable growth in IPv4. For most existing organizations that can foretell they have as much space as they will need for the next decade, this is not really an internal problem. Where these organizations will have a concern is when they deal with newcomers or others that have been forced into IPv6 because of exhaustion of the pool. Those organizations that foresee expansion and growth should evaluate Geoff's analysis as well as mine and weigh their plans against the risks of either or both of us being wrong.

In any case it only makes sense to start IPv6 capability discussions with the product vendors now. Product development cycles can be lengthy, and the only way for the vendor community to mesh with an organization's deployment plans is to have sufficient notice about those plans and timeframes. It would also be wise for the organization's network architects to start thinking about the impacts of an IPv6 deployment. Both protocol versions are packet-based and the names start with IP, but there are enough differences in the details that it is worth taking a fresh look to see what might be easier or cheaper than just blindly deploying IPv6 identically to the IPv4 deployment.

*Geoff:* The Internet continues to present challenges to the communications sector, and I would suggest that the underlying influential factor is the combination of the silicon and software industries that continue to fuel the demand side with fascinating, innovative, and compelling uses of communications that continue to surprise us with their continual restatement of the size of the domain in which we operate. We appear to be moving beyond servicing devices that are activated and influenced primarily by direct human activity, such as e-mail and Web use, and we are now looking at various command, control, and monitoring functions that embed themselves deeply in other devices and in other elements of our infrastructure. This encompasses larger concepts such as "smart buildings" and "smart traffic control," and they reach all the way down to the level of embedding into consumer devices and even identification tags. This is not a world that can readily be serviced by an IPv4 protocol infrastructure, and we are already seeing various levels of network indirection in both NATs and various forms of overlay networks to attempt to compress this new scale of basic network addressing demands into the IPv4 environment. This appears to be a complex, and therefore costly task. But the expectation here is that the service industry is heading toward a commodity utility function, where the essential attributes of the underlying network are simplicity and efficiency. These factors suggest that the market characteristics that arise from the propulsion of the silicon and software industries are inexorably tugging the communications service industry to embrace simple, scalable, and efficient networking technologies. It is in this space that the essential attribute of IPv6, that of the size of the address pool, has its most effective leverage. Here the "run out" of IPv4 will inevitably focus our common attention on how best to engage with future needs and roles. And in this perspective the IPv6 technology has a critical and central role.

*John:* Tony, I think we need to assume that, when it comes down to translating the projections into an answer to the "when do we need to get serious about IPv6?" question, both you and Geoff are, to a considerable extent, wrong. Geoff's articles and projections have been interpreted by some people as containing a "there is no problem, we can continue with IPv4 until we all retire" message. Viewed from that direction, yours can be seen as "we cannot be quite *that* complacent." Instead, I think we should all be looking at going directly to IPv6 in newer network installations rather than concentrating on whether we can get enough IPv4 space for them. We also need to be examining—now, not a few years in some projected future—the applications and services for end networks and end users, not just backbone and ISP services and operations. One of my particular concerns is that we have enterprise and customer support people and protocols all over the world who are used to thinking about things in an IPv4 world, including the support advantages of "all NAT-based end networks look the same" architectures. The need to retrain them to think about things differently, and to design and build new tools for their use, may suggest a more time-consuming and expensive transition than changing over the networks themselves.

*Fred:* What is clear to me from this discussion, Geoff's prior analysis, and Tony's analysis here, is that there is a timeline. We are *not* debating whether IPv4 address availability is limited or whether it can be "saved" by address allocation policy, nor are we debating the economic or technical impacts of more or less draconian allocation policies. We *are* debating what constitutes the end game, when and why that end game will become important, and whether perhaps we are already seeing the first steps of it. We are also not debating whether perhaps some new architecture would be preferred over the one in IPv6; if we had an alternative on the table today we could discuss that, but experience tells us that the proposals being considered by the *National Science Foundation* (NSF) and others are sufficiently "researchy" to not be ready for wide-scale deployment in the necessary timeframe.

As such, from my perspective, there is a present call to action.

What U.S. DoD and recent congressional hearings have recommended is in keeping with the IETF's recommendation and with the IPv6 address allocation strategies of the RIRs. The simplest transition strategy involves presently procuring equipment, operating systems, and applications that are IPv6-capable in preference to systems that are limited to IPv4. At some point in the future, perhaps in the 2008–2010 timeframe, we should plan to turn on IPv6 networking capabilities throughout our networks, and this means gaining experience with IPv6 on a smaller scale in 2005–2007 in our networks, in server applications, and in user systems. Turning down IPv4 capabilities, which is the endpoint of such a transition, is a business decision that does not need to be made hastily; we should presume that coexistence will be important for a decade, and probably more.

*Ole:* Thank you, gentlemen!

TONY HAIN is currently the Senior Technical Leader, IPv6 technologies, with Cisco Systems. In addition to providing guidance to the various internal product teams, he was also co-chair of the IETF working group developing IPv6 transition tools. His IETF participation since 1987 includes a term on the Internet Architecture Board from 1997 to 2001. Named an *IPv6 Forum Fellow* in 2004, he is currently serving as Technology Director on the forum's North American IPv6 Task Force steering committee. Prior to joining Cisco in 2001, he spent 5 years at Microsoft, where his roles included Program Manager for IPv6 as well as Network Analyst for the CIO's office. Prior to Microsoft, he was the Associate Network Manager for the U.S. Department of Energy's Internet effort, ESnet. With this range of roles, spanning the space between the implementation technologists and senior management, he brings a real-world viewpoint to the deployment decision process. E-mail: **ahain@cisco.com**

GEOFF HUSTON holds a B.Sc. and a M.Sc. from the Australian National University. He has been closely involved with the development of the Internet for the past decade, particularly within Australia, where he was responsible for the initial build of the Internet within the Australian academic and research sector, and has served his time with Telstra, where he was the Chief Scientist in the company's Internet area. Geoff is currently the Internet Research Scientist at the Asia Pacific Network Information Centre (APNIC). He served as a member of the Internet Architecture Board from 1999 until 2005, and currently co-chairs the Site Multi-homing and Routing Operations IETF Working Groups. He is author of *The ISP Survival Guide,* ISBN 0-471-31499-4, *Internet Performance Survival Guide: QoS Strategies for Multiservice Networks,* ISBN 0471-378089, and co-author of *Quality of Service: Delivering QoS on the Internet and in Corporate Networks,* ISBN 0-471-24358-2, a collaboration with Paul Ferguson. All three books are published by John Wiley & Sons. E-mail: **gih@apnic.net**

JOHN KLENSIN is an independent consultant based in Cambridge, Massachusetts. He has been involved in the design, development, and deployment of ARPANET and Internet applications, and occasionally lower-layer technologies, since the late 1960s and early 1970s. He has also been intermittently involved with Internet administrative and policy issues since the early 1980s. His current work primarily focuses on internationalization of the Internet on both technical and policy dimensions. E-mail: **klensin@jck.com**

FRED BAKER has worked in the data communications industry, building network elements such as switches and routers, since 1978. His involvement with Internet technology started in 1986, and with the IETF in 1989. He has contributed to the development of OSPF, QoS, PPP, SNMP MIBs, and a variety of other technologies. He has also held a variety of management positions, including chairing various working groups, participating in the IAB, and chairing the IETF. He currently serves on the Technical Advisory Board of the U.S. Federal Communications Commission and as the Chairman of ISOC's Board of Trustees. E-mail: **fred@cisco.com**

# Practical Uses of SSH Tunneling in the Internetwork

*by Ronnie Angello*

While the growing popularity of broadband Internet services and elevated concerns with securing *Wireless LANs* (WLANs) have become major concerns for network administrators today, *Secure Shell* (SSH) *Protocol* tunneling has proven to be a secure and effective solution for addressing various needs and concerns of both network users and administrators. Making the transition from traditional dialup remote access to a broadband solution can bring along with it some roadblocks when trying to preserve functions and security. WLANs can be difficult to secure in the enterprise, mainly because of the various client types that must connect to the network. SSH tunneling can help alleviate both of these issues.

SSH tunneling, also known as SSH *port forwarding,* is the process of forwarding selected TCP ports through an authenticated and encrypted tunnel. These tunnels can be constrained to within two points of the company's enterprise network, or it can originate on a small office or home office (SOHO) computer on a given provider's network, and transit the Internet to a server on the enterprise network. Some practical uses for SSH tunneling are outlined in this article.

### A Look Back at Traditional Remote Access

Remote access is the method of connecting from a SOHO computer that resides on a remote foreign network, or has no permanent network connection, to the enterprise network or central office. Usually this involves traversing the Internet. This can be for the purpose of telecommuting, providing on-call support from home, checking e-mail while away from the office, or for the old-fashioned workaholic who must work from home. Remote access used to involve simply accessing a network through an analog phone line or possibly ISDN. In either case, the user was authenticated by an access server that resides on the enterprise network and given authorization to certain resources.

When connected to the access server, users had the feel of being connected to their company's enterprise network. They were free to browse internal Web pages and access various Windows domain resources. They could connect to the network neighborhood and transfer files to and from the work computer. They could connect directly to internal UNIX servers with SSH and use a local X-server application to access UNIX applications from the SOHO.

PC remote-control applications such as VNC, etc. could be used to access files and applications that reside on a host computer on the enterprise network without extensive configuration on the home PC. In addition to the ease of configuration for the administrator or user, fewer applications need to be installed on the home computer to accomplish work tasks from home. This approach saves software licenses in addition to valuable company resources.

Most network administrators cannot let PC configuration consume a great deal of their time because they are busy enough as it is. From a function standpoint, users felt like they were working from their office at work. It was too slow though, so it did not really matter. Then broadband services were introduced, and they offer high bandwidth, but getting the same functions is a bit more challenging. Users benefit from the extra added bandwidth, but of course the administrator has to make sure that everything works as if nothing ever changed.

### Broadband Services Emerge

Many users are now migrating from their traditional dialup connections for Internet access to a technology that offers more bandwidth such as cable or DSL. Broadband wireless services are now emerging in some areas as well. These services may even be cheaper than what the company or individual was previously paying for ISDN service, and it is "always on." Most users are no longer dialing a company access server to access the resources that are vital to their job. They are now permanently connected to a foreign provider's network, and often the only choice for secure remote access to the enterprise is through a VPN. Strict policies, however, may need to be enforced on the remote SOHO computer for it to be a comfortable solution for security administrators to implement.

For those organizations without the time, money, or manpower to implement and support VPN, Linux login servers can be opened up to the Internet to authenticate users that employ SSH to access the enterprise network from these remote networks. These servers are no more than relay points to access internal systems. They should be placed in the DMZ or on a "screened" network protected by a firewall. The other internal systems are not directly accessible from the remote networks. In cases where remote access is considered a valuable resource to the organization, more than one of these servers should be implemented for load sharing and redundancy.

However, certain functions are lost. Initiating an application from a UNIX computer and displaying it to your SOHO computer with a local X server has been proven to be slow and inadequate from some remote networks. In addition, internal domain PCs and network shares are no longer accessible through the network neighborhood, and file transfer is not available without an additional secure, standalone application. The remote-control applications that access the internal PC will no longer work without opening holes in the firewall. There is a simple solution to all this that is free, secure, and effective: SSH tunneling.

### Securing Broadband Remote Access

The functions described in this section can be achieved with any SSH client capable of tunneling, any Web browser that supports HTTP and *Secure Sockets Layer* (SSL) proxies, and any PC remote-control application. The first step is always to connect to the remote login server that has been made accessible to the SOHO user. When connected to this login server, the user can use SSH to access any other internal machine, or take advantage of SSH port forwarding to accomplish their other tasks.

A proxy server may already be configured on your enterprise network. This server is configured to accept connection requests for Web pages and allow the clients to view them with little network overhead. The SSH client on the SOHO computer is configured to forward the specified local source HTTP port (such as 8080) to port 80 on the remote destination HTTP proxy server. It can also be configured to forward the specified local source SSL port (such as 4433) to port 443 on the remote destination SSL proxy server.

The browser on the client machine is configured to use the HTTP or SSL proxy server **`localhost`** on the specified local port(s). When the browser attempts to download a page, the SSH client forwards the request to the specified remote proxy server on your enterprise network through the established tunnel. Internal Web pages that would normally be available only on the enterprise local intranet are available without latency and without compromising security.

The same concept can be followed for tunneling PC remote-control application data through SSH. The remote-control host service is not changed, and it is waiting for a connection attempt from a remote computer as it normally would. A new remote-control connection is configured on the SOHO computer pointing to **`localhost`**. Using any additional encryption offered by the remote-control application is possible, but not necessary. Additional encryption will add latency, and SSH provides strong encryption itself with *Triple Digital Encryption Standard* (3DES), Blowfish, etc. The SSH client is configured to forward the local source ports used for the remote-control data (that is, port 3389 for RDP) to destination ports on the host computer on the enterprise network.

Once again, all the functions that the user had when dialing up the enterprise network directly are now available. With SSH, an additional layer of security is provided. Because the desktop of the internal computer is available on the SOHO computer's desktop, users have access to all applications, files, and network resources that they would if they were physically working from their office at work. No additional software applications need to be installed on the office computer to satisfy requirements of working from home, and minimal software needs to be installed on the users' personal home computers. Some of these remote-control applications also provide a file transfer tool that can be used to transfer or synchronize files between the two PCs.

### SSH Tunneling for WLAN Security

Securing WLANs has become a monumental problem today for most network administrators. Many organizations are resorting to proprietary solutions or are simply avoiding the implementation of WLANs entirely. An entire article could be dedicated to the importance of securing wireless and the details of accomplishing such a feat.

In addition to the uses described in the previous sections, SSH tunneling can also be used to supplement or replace weaker, more vulnerable encryption found in other network applications. Consider *Wired Equivalent Privacy* (WEP) encryption, for example.

Although other alternatives such as *Wi-Fi Protected Access* (WPA) are available, most WLANs have been implemented with either no encryption or with static WEP only. Static WEP has been highly criticized because of vulnerabilities in the protocol that have been discovered and widely documented. Even when implemented at the 128-bit level, there are tools circulating the Internet that exploit a well-known vulnerability that allows a hacker to crack WEP keys. Even with a WPA solution in place, there will be clients that support only static WEP. These traditional clients can be secured in the meantime by restricting network access with an *Access Control List* (ACL) and tunneling insecure protocols through SSH. Once again, the same functions can be achieved with a VPN solution, but some organizations have neither the money nor resources to implement it.

### Summary

In conclusion, SSH tunneling can be used well beyond the scope of the methods explained this article. The particular uses outlined in the previous sections have been practical in my experience and have been very successful implementations. When users decide to change to a provider that offers broadband, I have found that simply providing a procedure for configuring tunneling has been successful for getting them operational from home.

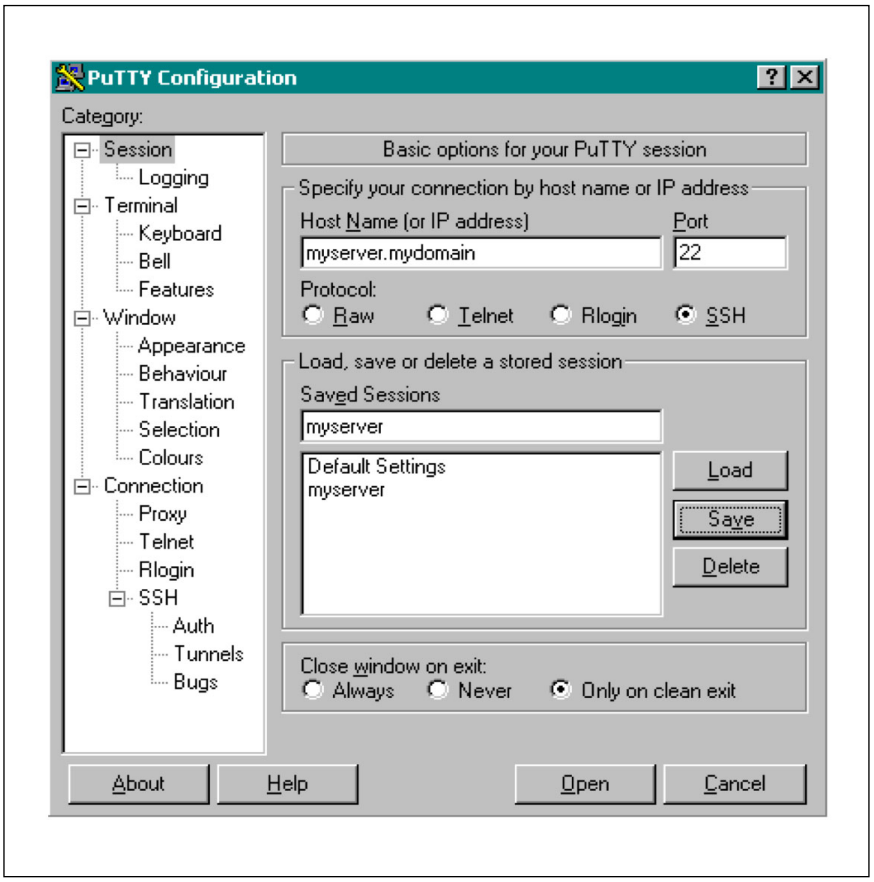
SSH tunneling should be of interest to any organization that wishes to allow its users secure access to all the resources that they may need to accomplish their job functions—especially from a remote location. While exploring possibilities to make a particular application or protocol secure, always consider SSH tunneling an option. SSH provides authentication and encryption that has been proven to be effective for any application.

### Securing Remote Access to Internal PCs, Web Pages, etc.

The following is a short example procedure for configuring tunneling for this specific function. It does not include detailed instructions for configuring specific applications, but it outlines the important steps that must be followed in order for it to work properly.

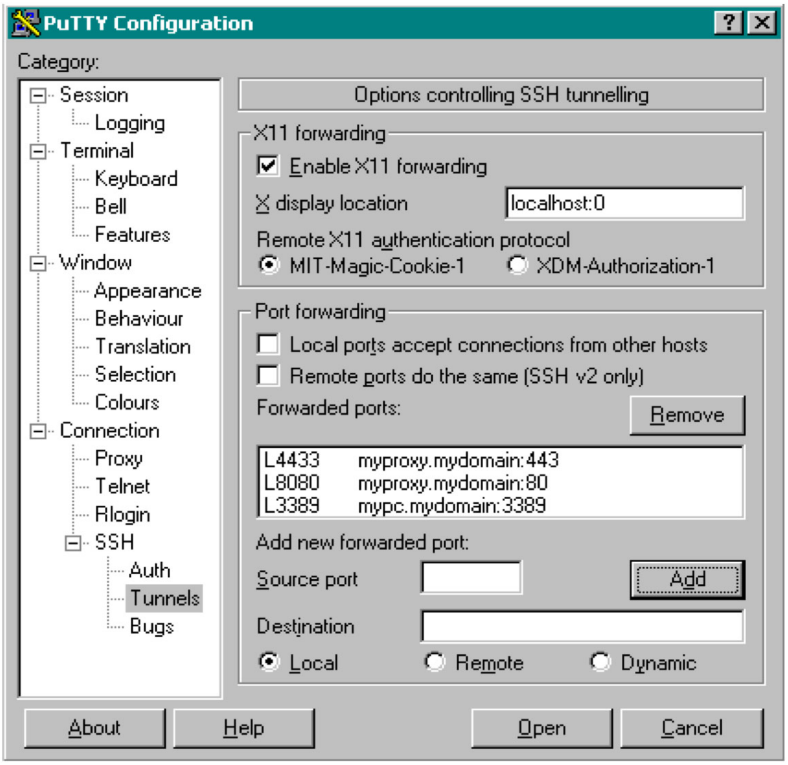
- Any SSH client that supports tunneling can be used. You can download the PuTTY SSH client (**`putty.exe`**) from: **`http://www.chiark.greenend.org.uk/~sgtatham/putty/download.html`**
- Make sure that you select port 22 (SSH). (See Figure 1.)

*Figure 1: PuTTY Configuration Screen — Sessions*




- Choose your preferred encryption cipher; enable compression and X forwarding if desirable. Click "tunnels" in the tree menu. Add the local source port(s) and the remote destination port(s) for the ports that you would like to forward through the tunnel. (See Figure 2.)

*Figure 2: PuTTY Configuration Screen —Tunnels*




- Make sure that the LAN settings in your Web browser are configured to use the HTTP/SSL proxy server **`localhost`** on the local port that you specified.
- Make sure that your remote-control connection is pointing to the computer "LOCALHOST." If you have trouble connecting, make sure that the host service is running on the host PC.

## For Further Reading


[1] The SSH (Secure Shell) Remote Login Protocol, SSH-1 Specification, T. Ylonen, November 1995.

[2] SSH-2 Specifications IETF Secure Shell working group, June 2003.

[3] O'Reilly Network Using SSH Tunneling:
**`http://www.oreillynet.com/pub/a/wireless/2001/02/23/wep.html`**

[4] SSH Tunneling:
**`http://www.ccs.neu.edu/groups/systems/howto/howto-sshtunnel.html`**

[5] SSH Tunnel Tiny HOWTO:
**http://www.frozenblue.net/tools/howtos/?v=ssh-tunnel**

[6] Secure Email Through SSH Tunneling:
**http://www.slac.com/~mpilone/projects/kde/kmailssh/**

[7] Mac OS X SSH Tunneling:
**http://info-center.ccit.arizona.edu/~consult/macx-tunnel.html**

[8] PuTTY Links:
**http://cdot.senecac.on.ca/software/putty/links.html**

[9] William Stallings, "SSL: Foundation for Web Security," *The Internet Protocol Journal,* Volume 1, No. 1, June 1998.

RONNIE ANGELLO, CCNP, CQS-CWLANSS, CCNA, holds an A.A.S. Degree in Information Systems Technology (Specialization in Operating Systems and Network Operations) and is currently completing degree requirements for the Bachelor of Science Degree in Information Science (Concentration in Networking and Communications) at Christopher Newport University in Newport News, Va. He recently passed the CCIE Routing and Switching Qualification Exam and is preparing for the CCIE Lab Exam. E-mail: **angello@jlab.org**

# Book Review

## Network Algorithmics

*Network Algorithmics: An Interdisciplinary Approach to Designing Fast Networked Devices,* by George Varghese, ISBN 0120884771, Morgan Kaufmann, 2004.

This is not a generic algorithms book (that is, it does not overlap much at all with Sedgewick or Coleman as an introduction to algorithms), nor is it a typical introduction to TCP/IP networking book (for example, there is no chapter defining the TCP/UDP/IP header fields, thank goodness). It might best be described as an algorithms analysis book set in the context of networking and also in the context of implementations that mix hardware and software solutions. For those familiar with Radia Perlman's book *Interconnections,* I found aspects of the writing style and approach to be similar. George Varghese—in addition to having been a networking professor for many years—has had a lot of industry experience from licensing algorithms to networking companies, to consulting with Procket Networks in the company's early days of architecting its core router, to starting a security company that was recently acquired by Cisco Systems. I have been doing architecture work at Cisco for several years and can say that George's book has real grounding in how systems are built and analyzed today.

### Organization

Chapter 2 presents abstractions for networking protocols, hardware design, routers, memory technology, and Internet end nodes (servers). This is a great introduction into "systems" thinking. In section 2.2.7, "Final Hardware Lessons," one thing I thought George should have mentioned along with metrics of chip size, speed, I/O, and memory is *power*. Power is becoming a major systems concern in many platforms and deserves mention as an optimization constraint.

Chapters 3 and 4 go through a list of 15 implementation principles to use in approaching algorithmic design in systems and then give examples of these principles in action. What I find interesting about this section is that from working with George in the past, he really does believe and practice "principle"-based architecture thinking. I remember discussing several of the principles with him several years ago, and you can see how his many years of experience working in the networking field have shaped these principles. Many have probably employed some of these, but as George says in the chapter introduction, having them explicitly documented with examples is useful to help clarify our thinking. Some of the principles (and both the short examples in this chapter as well as examples cited in more detail in later chapters) are really fundamental, and I think reading through examples helped clarify in my mind when to use them.

Chapter 5 covers copying data, for example, in a server design. I really like this type of chapter, in which a subject (in this case the effect of packet copying on Web server performance) is explored in detail but with a focus on where algorithms and systems design play an important part.

My biggest question about this chapter is that I was unsure how applicable this is to, say, modern server design using Linux and with latest Gigabit Ethernet *network-interface-card* (NIC) designs. I know there was a lot of interesting work in the late 1990s, but this chapter without any data is more along the lines of an extended example of how to apply implementation principles.

Chapters 6 through 9 are not what I would consider the meat of the book; they treat the topics of implementation and analysis for servers, timers, parsing/classification of packets, and buffer management (memory allocation).

Chapter 10 covers exact match lookups. There is not a lot of meaty algorithmic discussion, but the history of scaling performance of bridges is used to elegantly show an evolution of algorithmic approaches to exact matching.

Chapter 11 is an awesome overview of the state-of-the-art in longest prefix match (used for destination address matching in routers and switches). A good read of this chapter will yield an understanding of the trade-offs in all major published algorithms, although there may be variations or tuned versions of these algorithms in use at companies like Cisco. I believe this chapter covers all the major categories of solutions.

Chapter 12 extends the prior chapter into more general packet classification (which is used in applications like extended access lists). Like the lookup chapter, this chapter addresses one of George's prime core competencies. There is good discussion on leading published approaches (Grid-of-Trie, cross producting, geometric, and decision tree-based approaches). I strongly recommend this chapter.

Chapters 13 and 14 cover packet switching (that is, architecture of fabrics like crossbars for connecting line cards in a router or switch) and then packet scheduling. These topics get a good academic treatment (after all, George is one who introduced *Modified Deficit Round Robin* (MDRR) to the industry as well as academia), and although there are gaps between what many networking markets are defining as requirements for packet scheduling and what is in this chapter, the chapter is still useful.

Chapter 15 is a short chapter that tries to treat at a high analytic level the algorithmic problems involved with routing protocols. It covers this topic without getting very specific into nonrelevant (to the analysis) networking details.

Chapter 16, which addresses measuring network traffic, was probably one of my least favorite chapters. Some of it is academically interesting but requires network level changes that I just do not think will occur. There are some cute tricks relative to counters and such, but I think they are similar to approaches already being used.

Chapter 17 is a network security chapter and seems to serve as an early introduction to the topic of algorithms in network security; this is not a major focus area of the book.

### Areas for Improvement

There is always room for improvement, and I list here three areas in which this book could have been improved:

1. There is a running thread in the book of prefacing technical discussions in some cases with an example from the "normal world," like comparing packets to envelopes in the postal system. I estimate this is less than 1 percent of the content of the book and fairly easy to ignore if it annoys you.
2. I would have enjoyed better (more detailed) figures. A well-done, detailed figure can incorporate multiple concepts in the text around it and make it much clearer. On the positive side, there are numerous figures in the specifications, even if they do tend to be simple and high level.
3. Another area that I would have enjoyed seeing more on is empirical data (tables of data and graphs). I enjoy detailed empirical data of the type that Hennessy and Patterson so effectively use in their *Computer Architecture* book. There are many places (for example, Web server optimizations in Chapter 5) that I think could have benefited from detailed empirical data. However, I think folks often rely on empirical data too much when a simple analysis like the type done throughout the book could be done to help optimize the problem.

### Recommended

Many chapters in this book are directly relevant to the development of networking equipment and software, as well as what is "under the hood" of networking equipment. The book is fun to read and I believe succeeds in trying to convey an organized systems approach to thinking about problems in the networking space.

—*Will Eatherton*
**will@cisco.com**

---

### Read Any Good Books Lately?

Then why not share your thoughts with the readers of IPJ? We accept reviews of new titles, as well as some of the "networking classics." In some cases, we may be able to get a publisher to send you a book for review if you don't have access to it. Contact us at **ipj@cisco.com** for more information.

# Fragments

### Internet Governance Report Available

The *Computer Science and Telecommunications Board* (CSTB) of the National Academies has recently published a report entitled "Signposts in Cyberspace: The Domain Name System and Internet Navigation."

A summary report, as well as links to the full report can be found at:
**http://www.cstb.org/dns/signpost.html**

From the summary: "The *Domain Name System* (DNS) enables user-friendly alphanumeric names to be assigned to Internet sites. Many of these names have gained economic, social, and political value, leading to conflicts over their ownership—especially names containing trademarked terms. Congress, in Public Law 105-305, directed the Department of Commerce to request the *National Research Council* (NRC) to perform a study of these issues. When the study was initiated, steps were already underway to address the resolution of domain name conflicts, but the continued rapid expansion of the use of the Internet had raised a number of additional policy and technical issues. Furthermore, it became clear that the introduction of search engines and other tools for Internet navigation was affecting the DNS. Consequently, the study was expanded to include policy and technical issues related to the DNS in the context of Internet navigation. This report presents the NRC's assessment of the current state and future prospects of the DNS and Internet navigation, and its conclusions and recommendations concerning key technical and policy issues."

The report was produced by the Committee on Internet Navigation and the Domain Name System: Technical Alternatives and Policy Implications, National Research Council.

### First Protocols for Policy Makers Forum to be held October 28

The Internet has achieved the same global economic significance that propelled issues of international trade and finance onto the front pages of newspapers and the forefront of international policy thinking twenty years ago. This change is raising the profile of specialized issues and "obscure" policies for a rapidly expanding circle of public and private-sector stakeholders. Increased general understanding will be vital to assuring that Internet's growth, development, and coordination mechanisms continue to serve important public interests.

In recognition of this growing need for public education, Packet Clearing House is organizing a series of day-long roundtable fora to encourage sharing of technical and institutional know-how between prominent Internet architects, policy makers, and leading opinion leaders from related sectors. With the support of the *American Registry for Internet Numbers* (ARIN), the forum, to be called *Protocols for Policy Makers* (PfP), will meet for the first time on October 28, in conjunction with the NANOG 35 and ARIN XVI Internet operations and policy meetings in Los Angeles, California.
See **http://nanog.org/arinattend.html**

PfP will explore themes of competition, coordination, and possible conflict between new alternative Internet naming and addressing systems which are challenging the status-quo, such as the national registries recently proposed by the International Telecommunications Union and competitive private-sector "alternate roots." What outstanding problems are these new mechanisms intended to solve, and what goals might they achieve? How will these innovations contribute to the advancement of Internet public interests? What risks, costs, and complications may be imposed on the Internet by the emergence of multiple divergent systems? At PfP, these issues will be examined through a day of structured round-table discussions, interspersed with comments from leading experts on the Internet's current naming and addressing systems and prominent advocates of the current restructuring proposals. A complete agenda and list of speakers will be published shortly at **http://www.pch.net**

PfP will be open to the public, but space is very limited. For more information, or to request an invitation, please e-mail **pfp@pch.net**. Expressions of interest from potential speakers, meeting hosts, and institutional co-sponsors are also welcome. Plans for future PfP meetings are already underway, with a second meeting, tentatively titled "When Voice Goes to Bits" to focus on technical, commercial, and regulatory implications of the migration voice telephony to IP. Suggestions for future meeting themes, venues, and contributions should be directed to PfP Forum Chair Tom Vest at **pfp-sponsor@pch.net**

### Jun Murai Recognized with Postel Award

Professor Jun Murai is this year's recipient of the Internet Society's prestigious *Jonathan B. Postel Service Award*. The award recognizes Professor Murai's vision and pioneering work that helped countless others to spread the Internet across the Asia Pacific region.

The Postel Award was presented during the 63rd meeting of the *Internet Engineering Task Force* (IETF) in Paris, France by Daniel Karrenberg, chair of this year's Postel Award committee, and Lynn St. Amour, President and CEO of the Internet Society.

"Jun Murai has always encouraged, inspired and helped others, particularly his students and his colleagues in other parts of the Asia Pacific region," said Karrenberg. "He has also played a key role in creating structures for Internet coordination in the region (particularly the *Asia Pacific Network Information Centre* [APNIC]), and he is widely recognized for his recent pioneering work in IPv6 implementation."

Jun Murai is currently Vice-President at Keio University in Japan, where he is a Professor in the Faculty of Environmental Information. In 1984, he developed the *Japan University UNIX Network* (JUNET), and in 1988 established the WIDE Project (a Japanese Internet research consortium) of which he continues to serve as the General Chairperson. He is President of the *Japan Network Information Center* (JPNIC), a former member of the Board of Trustees of the Internet Society and a former member of ICANN's Board of Directors.

The Jonathan B. Postel Service Award was established by the *Internet Society* (ISOC) to honor those who have made outstanding contributions in service to the data communications community. The award is focused on sustained and substantial technical contributions, service to the community, and leadership. With respect to leadership, the nominating committee places particular emphasis on candidates who have supported and enabled others in addition to their own specific actions.

The award is named after Dr. Jonathan B. Postel, who embodied all of these qualities during his extraordinary stewardship over the course of a thirty-year career in networking. He served as the editor of the RFC series of notes from its inception in 1969, until 1998. He also served as the ARPANET "Numbers Czar" and the *Internet Assigned Numbers Authority* (IANA) over the same period of time. He was a founding member of the *Internet Architecture Board* (IAB) and the first individual member of ISOC, where he also served as a trustee.

Previous recipients of the Postel Award include Jon himself (posthumously and accepted by his mother), Scott Bradner, Daniel Karrenberg, Stephen Wolff, Peter Kirstein and Phill Gross. The award consists of an engraved crystal globe and $20,000.

ISOC is a not-for-profit membership organization founded in 1992 to provide leadership in Internet-related standards, education, and policy. With offices in Washington, DC, and Geneva, Switzerland, it is dedicated to ensuring the open development, evolution and use of the Internet for the benefit of people throughout the world. ISOC is the organizational home of the IETF and other Internet-related bodies who together play a critical role in ensuring that the Internet develops in a stable and open manner. For over 13 years ISOC has run international network training programs for developing countries and these have played a vital role in setting up the Internet connections and networks in virtually every country connecting to the Internet during this time. For more information visit: **`http://www.isoc.org`**

### Internet Root Servers Deployed in India

APNIC recently announced that three new Internet DNS root name servers are now operational in India.

These servers, launched in an official ceremony in New Dehli, India, on 25 August 2005, are the first root name servers deployed in India and South Asia and are already bringing significant improvements in speed and reliability to Internet users in India and the surrounding region.

APNIC has coordinated these deployments with the *Department of Information Technology* (DIT) and the respective root server operators.

F-root, operated by *Internet Software Consortium* (ISC) has been installed in Chennai; I-root, operated by Autonomica, has been installed in Mumbai; and K-root, operated by RIPE NCC, has been installed in Noida, near Delhi.

The installation of the root servers in India has been made possible by DIT, the *National Internet Exchange of India* (NIXI), and the *Internet Service Provider Association of India* (ISPAI), with financial and logistical support from APNIC. The three deployments in India bring the total number of root DNS servers in the Asia Pacific region to 24, 16 of which have been made possible with APNIC's support.

"We are pleased that India is able to contribute to the deployment of the first root name servers in South Asia," said Mr Pankaj Agrawala, Joint Secretary of DIT. "These three root servers will not only benefit the Indian Internet community, but also Internet communities in the surrounding region."

Paul Wilson, Director General of APNIC, added, "The deployment of these three root name servers in India is a positive example of Internet community coordination. The installation has involved the private sector, not-for-profit organizations, and government bodies working together to improve DNS stability and Internet response times for developing countries in South Asia."

Amitabh Singhal, Acting CEO of NIXI, said, "India is among the top ten countries in Internet usage, with over 35 million current subscribers and a five year target for 40 million, translating into more than 200 million total users by 2010. Sustainable infrastructure capacity building is imperative. As a budding intellectual capital of the world, with conducive socio-economic and political environments, India is justifiably proud of hosting three root servers, visibly putting our country, as well as the South Asian region, firmly on the world Internet route map."

More information about the participants can be found below.

- *APNIC* is one of five Regional Internet Registries currently operating in the world. It provides allocation and registration services which support the operation of the Internet globally.
  **http://www.apnic.net**
- *Autonomica AB* is responsible for **i.root-servers.net**, the first root name server to be installed outside the United States of America. **i.root-servers.net** has been operational since 1991 and is now anycast from more than 25 locations around the Internet.
  **http://www.autonomica.se**
- *DIT* operates under the Ministry of Communications and Information Technology, *Government of India* (GOI).
  **http://www.mit.gov.in**
- *ISC* operates one of the 13 root DNS servers as a public service to the Internet. ISC has operated F-root for the IANA since 1993.
  **http://www.isc.org**
- *NIXI* is joint effort between the GOI and the ISP industry to localize Internet traffic in India. NIXI has nodes in Delhi, Mumbai, Chennai and Kolkatta. **http://www.nixi.in**
- The *RIPE NCC* is one of five Regional Internet Registries currently operating in the world. It provides allocation and registration services which support the operation of the Internet globally.
  **http://www.ripe.net**

### IETF Journal Announced

The Internet Society (ISOC) is pleased to announce the *IETF Journal*, a new publication produced in cooperation with the IETF Edu team. Our aim is to provide an easily understandable overview of what is happening in the world of Internet standards, with a particular focus on the activities of the IETF *Working Groups* (WGs). Each issue of the journal will highlight some of the hot issues being discussed in IETF meetings and in the IETF mailing lists.

The focus of this first issue will be a look back at the accomplishments of the recent 63rd meeting of the IETF in Paris.

We trust that this publication will give all those with an interest in the increasingly important Internet standards development process an opportunity to keep abreast of many of the topics being debated by the IETF. Articles will cover issues such as:

- Reports from the IETF and IAB Chair
- News from the IETF Edu Team
- Update from the IASA and the IAD
- Summary of the plenary discussions
- Highlights of IETF developments related to topics such as Routing, DNS, and IPv6
- Recently published RFCs.

The journal will be available shortly at the following URL:
**`http://www.isoc.org/pubs/IETF-Journal`**

### Upcoming Events

The *North American Network Operators' Group* (NANOG) will meet in Los Angeles, October 23–25, 2005. For more information, see:
**`http://nanog.org`**

The *American Registry for Internet Numbers* (ARIN) will meet (jointly with NANOG) in Los Angeles, October 26–28, 2005. For more information, see: **`http://arin.net`**

The *Internet Engineering Task Force* (IETF) will meet in Vancouver, Canada, November 6–11, 2005. For more information, visit:
**`http://ietf.org`**

The *Internet Corporation for Assigned Names and Numbers* (ICANN) will meet in Vancouver, Canada, November 30–December 4, 2005. For more information, see: **`http://www.icann.org`**

The *Asia Pacific Regional Internet Conference on Operational Technologies* (APRICOT) will be held in Perth, Australia, February 22–March 3, 2006. For more information, see: **`http://www.2006.apricot.net`**

# Call for Papers

*The Internet Protocol Journal* (IPJ) is published quarterly by Cisco Systems. The journal is not intended to promote any specific products or services, but rather is intended to serve as an informational and educational resource for engineering professionals involved in the design, development, and operation of public and private internets and intranets. The journal carries tutorial articles ("What is…?"), as well as implementation/operation articles ("How to…"). It provides readers with technology and standardization updates for all levels of the protocol stack and serves as a forum for discussion of all aspects of internetworking.

Topics include, but are not limited to:

- Access and infrastructure technologies such as: ISDN, Gigabit Ethernet, SONET, ATM, xDSL, cable, fiber optics, satellite, wireless, and dial systems
- Transport and interconnection functions such as: switching, routing, tunneling, protocol transition, multicast, and performance
- Network management, administration, and security issues, including: authentication, privacy, encryption, monitoring, firewalls, trouble-shooting, and mapping
- Value-added systems and services such as: Virtual Private Networks, resource location, caching, client/server systems, distributed systems, network computing, and Quality of Service
- Application and end-user issues such as: e-mail, Web authoring, server technologies and systems, electronic commerce, and application management
- Legal, policy, and regulatory topics such as: copyright, content control, content liability, settlement charges, "modem tax," and trademark disputes in the context of internetworking

In addition to feature-length articles, IPJ will contain standardization updates, overviews of leading and bleeding-edge technologies, book reviews, announcements, opinion columns, and letters to the Editor.

Cisco will pay a stipend of US$1000 for published, feature-length articles. Author guidelines are available from Ole Jacobsen, the Editor and Publisher of IPJ, reachable via e-mail at **ole@cisco.com**

**The Internet Protocol Journal**

**Ole J. Jacobsen**, Editor and Publisher



*The Internet Protocol Journal is published quarterly by the Chief Technology Office, Cisco Systems, Inc.*
*www.cisco.com*
*Tel: +1 408 526-4000*
*E-mail: ipj@cisco.com*



The Internet Protocol Journal, Cisco Systems
170 West Tasman Drive, M/S SJ-7/3
San Jose, CA 95134-1706
USA

ADDRESS SERVICE REQUESTED

PRSRT STD
U.S. Postage
**PAID**
**PERMIT No. 5187**
**SAN JOSE, CA**