

To NAT or IPV6 – That is the question

Geoff Huston
December 2000

Last month we looked at the issues of policy management of the IP address space and the way in which IP addresses are managed at present. Lets now take a look forward at the various address policy issues we are likely to confront in the near future and the approaches we may use to address them, if you will excuse the pun!

Every graph that shows some aspect of the growth of the Internet has the same overall shape. Every graph has the same curve rising steeply to the right, showing the impact of exponential growth. Some graphs show doubling every year, some triple, and some rise at even faster rates. Across all of these graphs the overall story remains pretty constant – we are still in a period of very rapid growth and there are no visible signs of a slowing down of the expansion of the Internet.

How big can the Internet grow? Is it like the theory of the ever expanding cosmos, expanding for all eternity, or are there fundamental limits to the technology of the Internet which will inevitably cause this growth to slow down?

When the Internet engineering community first looked at this issue some nine years ago the conclusion was that the most pressing problem of potential exhaustion was that of the IP address space. While 32 bits of address space can mathematically encompass some 4 billion addressed devices, in the real world we cannot use the address space so efficiently, and a more realistic limit to the number of addressable devices in a 32 bit Internet address space is somewhere between 40 million and 1 billion devices. At the time of the initial study we were using address relatively inefficiently, and the lower 40 million limit looked like being reached some time in the mid to late nineties. A number of outcomes of the study were adopted by the Internet community, including the use of *classless inter-domain addressing* (CIDR), a technique intended to improve the address utilization efficiency by making routing more flexible. The major outcome of this effort was the specification of a new version of the internet protocol, IP version 6. The major change in this version of the IP protocol was the expansion of the IP address fields in the header from 32 bits per address to 128 bits. There's no doubt that 128 bits is a very large number, and if there was an address space shortfall in IP version 4, 128 bits of address is certainly a very large compensatory measure.

So IPv6 gets defined, disaster is averted in the nick of time, the Internet is saved and we can all head to the bar to tell tall stories of the heroic struggle with the technology beast. Right? Wrong! In the meantime something changed in the Internet, and some eight years later we are still using IP version 4, and address exhaustion still appears to be some time off. IPv6 remains in the wings, patiently awaiting the day of widespread adoption.

Were we wrong when we thought that address space was the problem, and is the 128bit address space unnecessary? There is a body of opinion to support this view as, in the meantime, we've seen the widespread introduction of Network Address Translation gateways (NATs) and Application Level Gateways (ALGs) into the Internet. These units are typically part of a security firewall system, and are intended to provide a level of managed isolation between a private address realm and the public Internet. The advantage of this approach is that of incremental deployment, where each private network can make a local decision to use private address space rather than public IP address space, make a local decision to use a NAT interface to the internet, and the impact of the deployment of the NAT is purely local to that private address realm. There no doubt that local decisions can be made far more readily than

global decisions, and much of the Internet is now sitting behind such translating gateways. This NAT approach is not only a solution for corporate networks connecting to the Internet. We have also seen Internet Dial-UP ISPs use private address space for their dial-in customers, ensuring that thousands of individual dial-up Internet connections can be supported using a very small number of public IP addresses. With NATs the number of devices requiring public IP address space is closer to the number of active devices actually communicating at any point in time, rather than the number of connected devices. This way the demand on public IP addresses is reduced substantially.

So are NATs the answer to the scaling problem, and can we build an ever larger Internet based on ever increasing deployment of NAT units between various network address realms? Maybe not, as there are some real concerns about the limited functionality of a NAT gateway. NATs can work well in a client-server environment where the client lies behind the NAT and the set of associated servers is well defined. But its not all good news. Because NATs store dynamic address translation state within the unit, NATs do introduce a single point of failure for its set of associated clients. The current deployment model of NATs is a single level of attached networks to a core public Internet. Scaling will introduce more complex technologies, and multi-step NAT architectures are beyond the simple NAT model. More fundamentally, NATs break the end-to-end transparency architecture of the Internet, and this can cause NATs to provide a very restricted view of the Internet. Any form of end-to-end authentication and security becomes a problem with a NAT box, and the concept of what unique identifier space is used to replace the now defunct global address space model is also left dangling. In short, NATs do not support the original Internet promise of true peer-to-peer and any-to-any communication. While NATs can support the client interactions with web servers today, is the future Internet application model to be restricted to web transactions simply because NATs cannot support much else?

Which brings us back to IPv6. If NATs cannot support a suitably rich functional communications model, and if we believe that the Internet is so much more than the web, then why aren't we all using IPv6 already? What's holding us back from making the transition to a protocol which does promise to solve the address scaling issue and also provide a level of functional richness that appears to be missing from NATs. Part of the problem that is holding us back is one of transition. IPv6 is not IP version 4, and the two protocols do not interoperate directly. If you want an IPv6 device to communication with an IP version 4 device then you will need some form of NAT function to translate IP packet headers between the two protocol realms. But if you need to use NAT, then what is the value in using IPv6 behind the NAT unit? Why not simply use IP version 4 as a private network like all other NAT deployments. IPv6 will only be of value in such a situation if there are a large number of users located in the version 6 network. As Metcalf's law puts it, the value of a network is proportional to the square of the number of users of the network, and right now the value of the IP version 4 network with its hundreds of millions of users is many times greater than the value of the current version 6 network with its thousands of users Over a billion times greater in Metcalf's terms.

So if we are going to see IPv6 deployed as the mainstay of tomorrow's Internet, then we are going to need to identify a collection of IP version 6 users who number in the millions or more. The Metcalf *value* of this many users will then start to come closer to that of the existing version 4 installed base, and there will then be some value and reason for ISPs and their customers supporting IP version as part of their service portfolio. There are a number of potential candidates for this new wave of IPv6 users. Some would see the massive numbers of devices that are projected to use the 3rd generation of mobile network technology as being an ideal candidate for IPv6 deployment, while others would see the mass produced consumer electronic industry as being the catalyst for an uptake in IPv6.

But right now we are still guessing, and there are no clear answers as to when and how IPv6 will achieve significant value in its own right. Equally, we are aware of a sustaining view that NATs cannot truly scale into an Internet of billions or more devices with a rich communications capability. It seems that as an industry we are poised on the brink of the collective decision to

move away from NATs and use IPv6 instead. But it appears that starting the significant adoption of IPv6 will require more than just words of warning about NATs impacting peer-to-peer models of network functionality and creating a complex end-to-end communications architecture. The collective decision to get serious about IPv6 deployment may well require a kick in the pants from an industry sector that cannot be accommodated from IP version 4 space, using an application model that cannot sit within an environment of NATs or ALGs. But don't hold your breath just yet, as we may be waiting for some time for this kick into IPv6, and in the meantime NATs do have a role to play in the Internet of today.
