

The ISP Column

An occasional column on things Internet

January 2006

Geoff Huston

IPv6 – Extinction, Evolution or Revolution?

For some years now the general uptake of IPv6 has appeared to be “just around the corner”. Yet the Internet industry has so far failed to pick up and run with this message, and it continues to be strongly reluctant to make any substantial widespread commitment to deploy IPv6. Some carriers are now making some initial moves in terms of migrating their internet infrastructure over to a dual protocol network, but for many others it’s a case of still watching and waiting for what they think is the optimum time to make a move.

So when should we be deploying IPv6 services? At what point will the business case for IPv6 have a positive bottom line? It’s a tough question to answer, and while advice of “sometime, probably sooner than later” is certainly not wrong, it’s also entirely unhelpful as well!

I’m not sure that anyone can provide a clearer date in response to that question, but what may be useful is to explore why IPv6 will be useful to have sometime in the near term future and how IPv6 and IPv4 are likely to interact. And then the “when” of IPv6 may be a little clearer – or maybe not!

To start off with this exploration it may be useful to compare where we started with the Internet with where we are today, and then see how this relates to the IPv6 story.

The Evolution of the Internet Architecture

The original architectural model for IP was in many respects a very simple model, but also one that was very powerful. Perhaps, in the spirit of William of Occam, the true strength of IP lay in what had been deliberately omitted from the specification, leaving in the form of the Internet a relatively simple and straightforward packet switching architecture.

William of Occam, (1285-1349), English philosopher and scholastic theologian. Occam was born in Surrey, England. He entered the Franciscan order and studied and taught at the University of Oxford from 1309 to 1319. Denounced by Pope John XXII for dangerous teachings, he was held in house detention for four years (1324-28) at the papal palace in Avignon, France, while the orthodoxy of his writings was examined. Siding with the Franciscan general against the pope in a dispute over Franciscan poverty, Occam fled to Munich in 1328 to seek the protection of Louis IV, Holy Roman emperor, who had rejected papal authority over political matters. Excommunicated by the pope, Occam wrote against the papacy and defended the emperor until the latter's

death in 1347. The philosopher died in Munich, apparently of the plague, while seeking reconciliation with Pope Clement VI.

Occam's Razor, "Pluralitas non est ponenda sine necessitate", has become a basic principle in science and philosophy, stating that entities should not be multiplied needlessly. This principle underlies all scientific modelling and theory building. In any given model Occam's Razor helps to cut away those concepts, variables or constructs that are not really needed to explain the phenomenon. Though such a process there is less chance of introducing inconsistencies, ambiguities and redundancies.

The network implemented an unreliable datagram delivery service. Each datagram (or packet), had information describing its source and intended destination. Each network switch (or router), either moved the packet closer to where it believed the destination was located, or it just dropped the packet. In the latter case the switch may send a control notification packet back to the sender, depending on the reasons for the drop. All the functionality that created various transport services, functionality to support mapping of application-level endpoint names to network addresses, and functionality to distribute available network resources across competing applications resided within the end systems rather than the network. For a network it really doesn't get much simpler than this.

But if you were to look for a faithful implementation of this simple architecture in today's Internet networks you'll be somewhat disappointed. The concept of single packet forwarding plane, with a single addressing model spanning the entire network, and a uniform end-to-end transport level congestion control model, has largely disappeared from most production networks, and the basic concept of 'end-to-end' is now perhaps more of an item of historic interest than a current pillar of networking architecture. These days carrier internet networks come replete with multiple forwarding layers, thanks to MPLS, numerous active network elements, including firewalls NATs and application layer gateways, various forms of NAT traversal agents and of course application level gateways and application level switches, load balancers, dynamic application switches and various forms of context-sensitive dynamic environments. We also have various forms of resiliency mechanisms, including path diversity elements, resource management systems, and QoS response systems. We have active Distributed Denial of Service (DDOS) detection elements embedded in the network and even network level session and application tracking systems as one more level of network defence against the ever-escalating security problem. This is no longer anything remotely similar to the concept of a simple unreliable datagram delivery service, and if you are looking for a simple dumb network with smart edges then you won't find it in production Internets.

What happened to the original Internet model? What was so wrong with a model of data communications that placed most of the functionality of the network into the devices themselves, and cast the network into a role of best effort packet switching? One sneaking suspicion is that the data communications industry itself, or at least the carrier part of the industry, is resisting this path to network simplicity, and in their continual quest to wring out every drop of value out of their networks the carrier ISP sector continues to be seduced by feature-packed network services that are intended to offer their customer higher value network solutions. Another way of looking at this role is that the carrier industry is hooked on the complexity business, and has embarked on a business model of creating networking systems that are sufficiently complex that customers are supposed to balk at doing it themselves. After all any construction enterprise can hang wire on poles, bury wire in the ground, or drop wire to the bottom on the sea. The highly complex operation of the resultant network is supposedly the unique value-adding role of the carrier enterprise. Of course this complexity escalation works only as long as the solutions are not so complex that the carriers themselves start to balk as well! As a carrier industry we may have already crossed this particular complexity line, and we may have already managed to create a technology environment that is sufficiently complex that no

player, not even the carrier, is able to manage the resultant interwoven mesh of disparate systems that make up a carrier Internet platform.

The question in my mind when looking at this rapid progression from architectural simplicity into often mind-boggling, and doubtless eye-wateringly expensive complexity for Internet networks is whether this is the outcome of a disordered process of entropy or one of a more ordered and directed process of evolution of the Internet?

The case for entropy is certainly very strong. What is evident is that the internet is besieged by various forms of local optimizations that intentionally alter the behaviour of parts of the network to suit the desired characteristics of certain classes of application. Such incremental local actions tend to impose a cost on the entire system. Whether the issue is one of adding network level support for mobility, support for various forms of address compression, support for differentiated service outcomes, resilience against various forms of hostile attack, or various forms of enhanced service availability, the typical outcome is one of increased network complexity and increased network cost with increasingly marginal returns in terms of overall service capability. This is a drive to disorder and decay in that local changes are not uniformly adopted, and the network itself starts to alter its overall state from uniform simple order into visible chaotic disorder.

Of course it is also possible to view this change process as one of evolution, where an active system is under constant pressure to adapt in order to survive and thrive in a changing environment. There's no obviously intelligent design here, and the overall evolutionary process follows no particular planned path. The outcomes are often chaotic and invariably unpredictable, but within the process is a driving discipline of a competitive environment where service providers are constantly challenged to adapt their service offering to meet the demands of customers. Here it is the competitive market that imposes the evolutionary pressure to adapt and survive or wither away into commercial bankruptcy .

Herbert Spencer, 1820 – 1903, British philosopher and sociologist, was a major figure in the intellectual life of the Victorian era. He was one of the principal proponents of evolutionary theory in the mid nineteenth century. It was Spencer who invented the phrase "survival of the fittest", and originally applied it to the process of elimination of firms in the rather vicious cut and thrust of Victorian capitalism. Upon the publication of Charles Darwin's "On the Origin of Species" in 1859 Spencer quickly saw the parallels to natural selection and applied the phrase to the process of natural evolution. As a result he became on the a group of philosophers known as "social Darwinists", applying Darwin's principles to human society. It has often been considered a relatively harsh philosophy, espousing in its most extreme form that the fittest members of society naturally survived and prospered, while the weaker members of a society were doomed to perish.

Many of the incremental measures we see in today's networks have been bought about by this reactionary response to market pressures rather than through a distinct planned process of technology development. One could characterize firewalls, Network Address Translators (NATs), Quality of Service (QoS), Application Level Gateways (ALGs), network caches, and a myriad of similar mechanisms as examples of this form of ad hoc response to market pressures for network services. Whether they represent entropy or evolutionary change in the Internet model is perhaps left as a personal perspective.

One area of technology continues to sit outside this process of current technology churn in the Internet, and that's IPv6. IPv6 is not an outcome of a reactive model of technology development, but is instead an example of a centrally planned development that was designed in anticipation of a market situation. Curiously, the very conditions that IPv6 was intended to avoid, namely that of a chronic address shortage in the deployed network, have already manifested themselves in many ways and in many places, and yet the market demand for IPv6 services remains relatively insignificant, and certainly below a threshold for viable commercial services for many operators.

So what's the problem? How will IPv6 services appear in the market? Is this an evolutionary process of orderly migration of IPv4-based services into an IPv6 networking realm? Or is IPv6 going down a path of premature extinction, never to appear as part of the mainstream communications portfolio? Or will IPv6 play for high stakes here and take on IPv4 as its major competitor and win market share through a revolutionary process of defining price and performance points that are simply not sustainable with any other technology, including IPv4?

Lets now look at the potential futures for IPv6, and in particular look at the options of extinction, evolution and revolution in the context of IPv6 and its struggle for market takeup in the coming years.

Extinction

Is IPv6 another case of *OSification*, or another example of a network technology that simply will never attain mainstream adoption?

The Open Systems Interconnection (usually abbreviated to OSI) was a new effort in networking started in 1982 by the International Organization for Standardization (ISO), along with the ITU-T.

Prior to OSI, networking was completely vendor-developed and proprietary, with protocol standards such as SNA and DECnet. OSI was a new industry effort, attempting to get everyone to agree to common network standards to provide multi-vendor interoperability. It was common for large networks to support multiple network protocol suites, with many devices unable to talk to other devices because of a lack of common protocols between them.

However, the actual OSI protocol stack that was specified as part of the project was considered by many to be too complicated and to a large extent unimplementable. Taking the "forklift upgrade" approach to networking, it specified eliminating all existing protocols and replacing them with new ones at all layers of the stack. This made implementation difficult, and was resisted by many vendors and users with significant investments in other network technologies. In addition, the OSI protocols were specified by committees filled with differing and sometimes conflicting feature requests, leading to numerous optional features. Because so much was optional, many vendors' implementations simply could not interoperate, negating the whole effort.

The collapse of the OSI project severely damaged the reputation and legitimacy of the organizations involved, especially ISO. The worst part was that OSI's backers took too long to recognize and accommodate the dominance of the TCP/IP protocol suite.

http://en.wikipedia.org/wiki/Open_Systems_Interconnection

Will IPv6 act as a catalyst to take a step in some completely different technology direction that may be as radical in their nature as previous big leaps of technology in the communications sector? In the same fashion as the industry has already lurched through multiplexing solutions based on Frequency Division Multiplexing, Time Division Multiplexing and then Packet Switching, are we awaiting something far more radical than a realignment of some of the IP packet's header fields? Is IPv6 a rather eloquent demonstration that packet switching has reached some basic set of limitations and that a successor technology to IPv4 needs to take a completely new approach to a shared communications environment?

The original IP architecture, as a very simple adaptation layer between a broad collection of packet switching technologies and a similarly broad collection of services and application, is certainly dying at the moment, if not already dead. The model of coherent and transparent end-to-end packet transmission is disappearing from today's network, and is being replaced with a collection of packet header rewriters, a set of content sensitive packet forwarding systems and even entities that perform session interception and regeneration. Any application that assumes a simple end-to-end model of packet delivery has no role in today's Internet, and any popular internet application has to be able to invent its own identity space, and be able to allow its data streams to pass through NATs, ALGs and other middleware elements with impunity. This may require multi-party interactions to complete the transaction where previously only two parties were necessary. For peer-to-peer environments we are now looking at application mediators and agents to assist in setting up the necessary rendezvous points, as well as assisting in the identification of what forms of middleware behaviour exist in the network path (STUN, ICE and TURN are good examples of this approach of application-level middleware discovery). Efforts to impose overlay topologies, tunnels, virtual circuits, traffic engineering, fast reroutes, protection switches, selective QoS, policy-based switching on IP networks appear to have simply added to the cost and detracted from the end user utility.

So, today, we are engineering applications and services in an environment where NATs, firewalls and ALGs are assumed to be part of the IP plumbing. We now have constrained models of interaction that divide the work into clients and servers, and mandate that all transactions are initiated by clients and are directed to servers. We now have forced applications to invent their own per-application identity realms, and required applications to also require the deployment of active middleware in the form of agents in order to orchestrate multi-party rendezvous and referral. By implication NAT states and other middleware states are now multi-party shared states, and what were considered to be local autonomously functioning entities now are faced with the complexities of supporting a signalling environment that is associated with distributed shared state.

All this complexity is not just a problem in the abstract sense, but a form of architecture that results in more fragile applications and higher operational costs. The Internet, far from becoming simpler and cheaper, is under increasing pressure to take on increasing complexity and operate with escalating costs.

Can IPv6 reverse this trend? We've all heard the observations that IPv6 was a typical standardization conservatism. IPv6 also represents an outcome of engineering compromise between making marginal changes and taking an entirely new approach to packet switching architecture, and the standards

process is invariably one that tends to avoid making radical decision. IPv6 represents a very marginal change in terms of design decisions from IPv4. IPv6 did not manage to tackle the larger issues of overloaded address semantics. IPv6 did nothing to address routing scaling issues. IPv6 has done little in terms of altering the semantics of packet switching, and what we are left with in IPv6 is a slightly larger address field

One could be excused for thinking that the marginal changes in IPv6 over IPv4 represent such a small difference that no one would be interested in paying their share of the rather high price of worldwide transition

Alex Lightman, chairman of the IPv6 Summit, was reported to have raised the question of who will actually pay for the transition to IPv6. As reported by internetnews.com, "There is an unreleased report by the Dept. of Commerce estimating it will take \$25-\$75 billion to pay for the transition, according to one of our speakers," Lightman said. "So what part of that will the U.S. government pay for?"

December 12, 2005,
<http://www.internetnews.com/infra/article.php/3570211>

But if IPv6 is indeed too small a change over IPv4 and its fate is really to be that of extinction, then what other approaches can we take to a successor to IPv6? Is there anything else around today that takes a radically different view of how to multiplex individual transactions within a common communications system? The answer to this question appears to be "no", or at least there appears to be nothing that has been developed beyond the initial conceptual stage, and certainly nothing that has been extensively evaluated for such a role. So, for the near term, there does not appear to be any alternative technology waiting in the wings. If we don't appear to want to adopt IPv6, and are happy to let it lapse into extinction, then we need to design and develop another protocol. In that case how long would such a new design effort take? And if we embarked along such a path what is the likelihood that the effort would encounter precisely the same set of constraints as the IPv4 and IPv4 design efforts and what is the likelihood that the effort would end up in much the same place as IPv6 - taking a slightly different view of a common set of design trade-offs between a common set of basic constraints that were already encountered in IPv4? Of course there is also the option of heading well beyond the current concepts of packet switching and look at entirely different communications architectures, but here the considerations of the design and development timelines become a significant inhibitory factor here.

So if we think that IPv6 is not the answer, and we believe that we should look elsewhere for a successor technology to IPv4, then it is likely that any such effort would take at least a decade, or, more likely longer to generate a workable outcome. And the other nagging consideration here is the question of whether such a design effort would end up as a marginal outcome in any case. Would we be looking at no more than a slightly different set of design trade-offs within a common set of constraints?

So in the near term, and possibly in a longer term of some decades to come "extinction" is not a very likely outcome for IPv6 - there is simply no other option on our horizon, so if we are to move away from IPv4 sometime soon then IPv6 is what we will be using instead.

Evolution

So if the premature extinction of IPv6 is highly unlikely, then can we made do with IP4 indefinitely, or should we be looking for some evolutionary path into IPv6?

Can we continue to use IPv4 indefinitely? There's little doubt that the IPv4 network model is under relatively severe stress in terms of its address and routing scalability, and there is no confidence that IPv4 can be made to scale indefinitely to encompass larger and larger populations of users. As we've already noted the Internet is no longer a simple network, and as it continues to grow then its likely that at some point the cost of scaling the various components and their forms of interaction reach a point where its just no longer a viable proposition to continue to grow. While increased volume usually implies lower unit cost, at come point the cost of complexity starts to become a significant factor in unit cost escalation, and the network reaches a scaling failure point. The possible pressure points include the capability to scale NAT deployment indefinitely, the capability to scale routing systems, the capability to scale network middleware indefinitely, the capability to effectively ward off various forms of hostile attack on the network, and the capability for an ever larger ever more complex network to operate in a stable and useful fashion. Whether this is a failure point of the capability of the technology, where the network itself reaches a size where it just cannot operate in a stable mode, or whether this is a failure point of the underlying economics of the network where the unit costs of the service escalate beyond the point of viability is an open question, but the common factor is that IPv4 is a technology platform with finite scaling bounds, and it cannot fuel an open-ended networking future.

Hopefully we should have evolved the network beyond these limitations well before reaching such a critical failure point, and the major lever here appears to be to head towards a simpler network that performs fewer functions within the network. Simpler networks, simpler applications, simpler operation, better scaling properties. This is certainly the core promise of IPv6.

So if the question is "should we evolve the network to IPv6?", then the general answer appears to be a resounding "yes" for most values of "we".

However the precise motivations vary for each player. IPv6 can allow for the resumption of a network model that uses unique global addresses for each connected endpoint, for endpoint populations that can scale into the hundreds of billions. IPv6 is capable of embracing a device-dense world. The per-address cost can be reduced dramatically through the elimination of various forms of dynamic address translation technologies, as well as the elimination of the scarcity premium factor in IPv4 address mechanisms. Application complexity can also be reduced, and the diversity of application models can be broadened. This model of universal addressing allows for many forms of peer-to-peer networking models as well as supporting communication transaction security models that reply on end-to-end coherence. All these factors point to a networking model that supports simple and ubiquitous communications services which in turn supports utility device deployments. So the desired outcomes appear to point to simpler networks, simpler applications, larger populations of connected devices, more efficient services, and a broader diversity of service models. So the set of potentials presented by ubiquitous adoption of IPv6 presents a very compelling picture of benefits for a diversity of players in the industry.

However none of these potentials has managed to persuade the industry to take the plunge and undertake the transition to IPv6 so far. The potential benefits of IPv6 appear to offer insufficient drive to the industry to get this transition underway. Why is this? Perhaps its because the pressure points of the current IPv4 deployment don't cause uniformly high levels of pain. ISPs are neither application authors nor are they device manufacturers. So ISPs do not directly incur the additional cost of complexity in the application or the cost of additional memory , additional software and additional configuration complexity in the device. So the ISP feels insufficient levels of direct pressure to roll out a new network protocol.

What else would drive an ISP to deploy a new networking protocol? In crude terms there are two very basic business drivers – fear and greed. Greed is the desire to enter new markets in a way that

maximises beneficial outcomes, while fear is a defensive response to emulate the business opposition to defend an existing market position. So in these terms is there an “early adopter reward” for deployment of IPv6? What is the fear or greed driver here that would propel the ISP industry into undertaking this transition? Unfortunately there appear to be no clear “early adopter” rewards for IPv6. Existing players currently have strong motivations to defer expenditure decisions because of strong shareholder pressure to improve the earnings per share position within the carrier industry. This is not the time to support a business case to leap too far ahead of the existing business model and take a somewhat riskier longer term position in the market. There is still some considerable uncertainty over the future of the voice industry as the competition with VOIP becomes more intense, and there is still a basic push by the industry to enter into value-added service markets that entail more complex network architectures, and IPv6 is seen as being a longer term direction that has little of relevance to the current ISP industry position. The return on investment in the IPv6 business case is simply not evident in today’s ISP industry. New players have no compelling motivations to leap too far ahead of their seed capital. All players see no incremental benefit in early adoption. And many players short term interests lie in deferral of additional expenditure. So the short term industry response appears to be to defer expenditure on IPv6-based deployments and await further developments.

So if the question is “when will this transition to IPv6 happen”, the general industry response appears to be “later”. So the real question here is what is the nature of the trigger for change, or, at what point, and under what conditions, does a common position of “later” become a common position of “now”?

So far we have no clear answer from industry on this question

This is not a case of where regulatory initiative would be all that helpful. Our previous experience with OSI and various national and regional GOSIP programs has provided a convincing lesson that technology adoption through regulatory measures or administrative fiat are abject failures. So we are forced to look back at the market interaction between services providers and consumers of the services to see where the leverage may lie. Unfortunately there are few network differentials in the current consumer world that provide any great leverage – after all it’s still email and it’s still the web, and the choice of protocol over which these applications operate should be a matter of supreme indifference to the end consumer. Expecting the consumer to pay more for a supposedly seamlessly invisible network attribute is indeed a bad case of wishful thinking. Indeed it is perhaps worse than this. In recent years we have managed to create a secondary supply industry based on network complexity, address scarcity, and insecurity. The prospect of further revenue erosion from simpler cheaper network models based on IPv6 deployment is one that this industry views with some suspicion and fear. The business obstacles don’t stop here. The concept of simpler networks leads to the concept of revenue erosion for provision of network services. In an industry that has already undergone significant turmoil over the past decade, and where the current incumbents are looking at weak financial figures for their businesses the entire concept of outlaying more capital investment to deploy an IPv6 network is not exactly a glowing proposition. Indeed the industry has already invested large sums in packet-based data communications over the past decade, and there is little investor interest in still further infrastructure investment at present. When you add to this the consideration that IPv6 is a step back to a simpler, cheaper network, then this translates to an incremental investment that will reduce their revenue yield per customer. This is not exactly a business-friendly proposition. So it’s little wonder that the industry has been far more fascinated in the concept of MPLS, QoS and VPNs in an effort to increase the returns on their network investment through the quest for “value added services” and at the same time paid lip service to IPv6 without any major level of investment to match.

Oops!

So evolution, or an ordered migration from IPv4 to IPv6, does not appear to be happening. IPv6 is not seen in a highly positive light. IPv6 promotion may have been too much too early, and these days IPv6 may be seen as tired rather than wired.

“Everything over HTTP” and the client-server model of networking has proved far more viable than perhaps it should have, and these days any decent application that gains popular attention can traverse NATs, ALGs and a myriad of other middleware barriers with consummate ease. If it couldn’t be so agile then it simply would not gain popular attention. So we now have an Internet where the service portfolio appears to be collapsing into a small set of applications that are based on an even more limited set of HTTP transactions between servers and clients.

Maybe it’s just deregulation of the industry, where short term business pressures simply support the case for further deferral of IPv6 infrastructure investment. In this economic view of the Internet industry there is insufficient linkage between the added cost, complexity and fragility of deploying network middleware and associated traversal applications at the edge of the network and the costs of infrastructure deployment of IPv6 in the middle. This leads to the observation that deregulated markets are often not perfect information markets, and the points of pain, or cost, become isolated from potential remedies, or savings.

It would appear that evolution is really not an option for IPv6 either.

Revolution

The transformation of IPv4 from a research experiment to a mainstream public communications environment is an interesting case of technology revolution. IPv4 presented a portfolio of cheaper switching technologies, more efficient network usage, simpler networks with lower operational costs, and structural cost transfer from operational costs within the network to capital costs at the edge. IPv4 represented a compelling and revolutionary business case of stunningly cheaper and more effective services to end customers. This was the silicon revolution at its most effective. The transformation has not been ordered and well planned. Some of the giants of the older telephone world have lost vast amounts of money, some have gone bankrupt with others have been sold off as mere shadows of their former market presence. Workforces are being realigned, investors have had to adjust their expectations and regulators have been confronted with an entirely new set of market behaviours and associated services.

Perhaps the most compelling view of IPv6 is in the same vein of being a revolutionary force with large scale disruptive implications to the industry. The leverage here lies in the observation that IPv6 represents an opportunity to embrace the communications requirements of a device-dense world – an opportunity that is simply lacking in the IPv4 realm. This device dense world is a world that is far larger than that of human-use devices, and encompasses a potential population that is at least some 2 – 3 orders of magnitude larger than today’s Internet. This encompasses a world of embedded communications, smart tags and applications that can encompass many forms of active and passive monitoring.

In and of itself this sounds benign, of not innocuous for the Internet. But how much money would you let your washing machine spend on communications services? Or your luggage tag? Or any one of thousands of chattering devices? The economics of a device-based communications world are vastly different from that of a human-mediated communication. In the voice world the value proposition shifted away from cost-based service tariffs towards value-based tariffs. It wasn’t the cost of allowing two people to speak to each other, but the value people placed in being able to talk to each other. Even the Internet so far has an inherent value in human-based communication. The value of today’s Internet lies in people-to-people messaging, lies in web browsing, lies in downloading entertainment, and lies in other predominately human pastimes. In a device world the value proposition is at a much lower level, and one way to look at the resolution of a device-based Internet is to think of a service environment that reduces the end consumer costs by a further 2 to 3 orders of magnitude. Yes, that implies that the threshold for a device-rich communications world is an industry price benchmark of

megabit per second access tariffs for between 2 to 30 cents a month, or being able to purchase gigabit per second internet access for the same \$30 price benchmark we use today.

How to achieve these revised price benchmarks for Internet services is the critical question. We've already extracted massive improvements in transmission cost efficiencies in the move into wave division multiplexing on fibre cable. We've already extracted massive improvements in the efficiency of switching through the move from time to packet switches and the move from state-based circuit switches into stateless packet-based switches. We've already extracted further cost efficiency in the network by pushing many of the services and functionality out to the edge and attempting to follow a direction of simpler cheaper networks.

So what's left? I suspect that the truly revolutionary message in IPv6 is a message about the extracting efficiencies in the business model of communications. We appear to be looking at a transition from value to volume with IPv6. IPv6's true leverage is about the ability to encompass world of tens of billions of chattering devices. The service industry that provides the networking services to these tens of billions of devices will not be a bloated inefficient relic of a bygone era of monopoly service enterprises. Indeed its likely that there will be nothing in common with the enterprises that operate in this industry today. IPv6 appears to be carrying an implication of a quite dramatic shift in the service enterprise to an industry based on a commodity utility. We are looking at an industry that will operate at a level of single digit operating margins and investment returns similarly phrased. If we want IP to operate from anonymous sockets in the wall, or seamlessly over wireless, then we will be looking at service delivery systems that provide simple lowest common denominator networking service. The search for valued-added services and value-added networks have no logical role in such a commodity utility world. This all sounds quite conventional, and the path to commoditization of many artefacts and services is a well trodden one in many industries and service sectors. So why is this such a revolutionary message for the communications industry? I suppose that the observation here is that this is one industry which is continuing to live the myth that there is a pot of gold out there in value-added networking-land, and that the windfall profits made in successive waves of innovation in the telephone industry over the decades will continue to repeat itself, and there is a pervasive air of denial over a message that says that the value is going to be destroyed by volume. In this industry the words "commodity" and "utility" remain taboo!

The IPv6 Condition

In taking an objective look at IPv6, there are no compelling technical feature or revenue levers in IPv6 that are driving new investments in existing IP service platforms. It does not appear that an industry-wide shift to IPv6 is going to be driven by the current value-added network service model and the associated current set of consumers of today's services. There is just insufficient marginal benefit to the end consumer to create a value proposition that will justify paying an increased tariff for having access to IPv6 as well as IPv4 – after all its still email and its still the web!

The current user base has managed to become wedged in a situation where there is not enough impetus to move away from the networking model of IPv4, and we appear to be stuck within a client-server model of network-mediated relationships. The network operators continues to push the network into undertaking a higher valued role in mediating communications and usage of the network continues with a largely human-directed set of services. One could characterize this as an environment that places extracting maximal value from the network as the prime objective, over serving maximal volume

Interestingly, the underlying engine for digital communications, the silicon chip industry also started in a vein of attempting to place silicon chips in highly-valued devices, but this industry made the switch to a volume industry decades ago. This is an industry that has significant cost differentials between

design and fabrication, so its probably little surprise that they quickly appreciated the longer term value in a general approach to recouping the design cost in very high volume production runs.

It likely that IPv6 sits in this same situation, and will only gain widespread industry acceptance within a broader shift in the communications industry from value to volume. It we are truly looking at an Internet of gadgets, of billions of chattering devices, then what will drive IPv6 deployment in a device rich world is a radical and revolutionary value to volume shift in the IP packet carriage industry. In IPv6 we appear to be looking at a shift in the industry to that of an undistinguished commodity utility service provision industry. An industry that will inevitably take on once more a very conservative profile and one that will no longer be able to afford further extensive and rapid innovation. So if we take this step into such a world then we need to be pretty confident that we are comfortable with this step being a very long term one.

The IPv6 Revolutionary Manifesto

It its going to be unlikely that IPv6 is an evolutionary step for the Internet, but rather that of yet another revolutionary step for the communications industry. It is likely that IPv6 will need to compete for market share with IPv4, and the basic terms of the competition for the consumer will be price-based competition rather than feature or service-based. IPv6's basic potential is that of extraordinary volume, but to achieve this we will need to push down unit cost of packet delivery by orders of magnitude. It appears that the major means of getting there is through commodity volume economics that will direct the industry towards even "thicker" transmission systems, simpler, faster switching systems, lightweight application transaction models, and an industry profile of a commodity utility sector.

This is definitely going to be a painful revolution, as it will be the industry itself that will offer the highest levels of resistance to such a radical agenda.

In June 2003 the following announcement was made by the US Department of Defence:

US Department of Defense adopts IPv6

Implementation of the next-generation Internet protocol that will bring the Department of Defense closer to its goal of net-centric warfare and operations was announced on June 13, 2003 by John P. Stenbit, Assistant Secretary of Defense for networks and information integration and DoD Chief Information Officer.

The new Internet protocol, known as IPv6, will facilitate integration of the essential elements of DoD's Global Information Grid—its sensors, weapons, platforms, information and people. Secretary Stenbit is directing the DoD-wide transition.

The current version of the Internet's operating system, IPv4, has been in use by DoD for almost 30 years. Its fundamental limitations, along with the world-wide explosion of Internet use, inhibit net-centric operations. IPv6 is designed to overcome those limitations by expanding available IP address space, improving end-to-end security, facilitating mobile

communications, enhancing quality of service and easing system management burdens.

"Enterprise-wide deployment of IPv6 will keep the warfighter secure and connected in a fast-moving battlespace," Secretary Stenbit said. "Achievement of net-centric operations and warfare depends on effectively implementing the transition."

Secretary Stenbit signed a policy memorandum on June 9 that outlines a strategy to ensure an integrated, timely and effective transition. A key element of the transition minimizes future transition costs by requiring that, starting in October 2003, all network capabilities purchased by DoD be both IPv6-capable and interoperable with the department's extensive IPv4 installed base.

I was asked to provide a comment on this announcement, and at the time I made the following response:

The enduring value of IPv6 lies in the massive amount of coherent address space that allows literally billions of devices to be uniquely addressed. Address uniqueness is a strong value proposition when you want an identifier space to cover a very large deployment space. As an example of this, one of the two properties of the original Digital-Intel-Xerox Ethernet II specification that remains in today's 10 Gigabit Ethernet specification is unique 48 bit MAC addresses. All of that highly innovative CSMA/CD thinking that at the time we thought was the fundamental property of Ethernet has been dispensed with, and it's the address space that still defines "Ethernet" today.

The general observation is that any communications systems requires any party to be able to uniquely identify any other party in order to initiate a private communication session. If you cannot perform that most basic of communications functions, then you simply do not have a functional peer-to-peer communications network.

But doesn't that mean that the stories of IPv4 address exhaustion have some substance? With the large amount of addressable devices hidden behind NATs, and the associated move to using domain names as the underlying identifier space for many communications applications, the pressure on consumption of IPv4 address space has been reduced considerably, but at the cost of increased network complexity. This has implied that in a world of human-driven screens and keyboards we see some considerable lifetime left in the admittedly comfortable world of IPv4 as we know it. To support this model we've actually moved away from the IP address as the unique identifier token for many applications, and substituted an application model that is largely driven from domain names. As a trivial example, look the virtual hosting mechanism as implemented in web server implementations to see this shift in server identifiers from IP address to domain name. So in the context of the current IP market, as both as consumers of the technology and as an industry, we can live with this identity split for some time yet, because we appear to concentrate our use IP addresses as a routing and

forwarding framework identity and increasingly use the DNS as the identifier realm of applications.

Our world is a world where the device is subservient to the user, and the applications we associate with the Internet of today are applications that are essentially human pastimes, such as e-mail, web browsing, or high-value automated transactions, such as those commonly bracketed into the e-commerce area. And we've now established a highly valuable global industry upon these foundations.

In so doing we should recognize the emergence of a second set of communications realms populated by uniquely identified devices that number in their billions, where the inter-device traffic is not human mediated, and the value of the device transactions are, on an individual transactions value level, far lower than the value of the human-driven realm of IPv4. In other words, in a device rich communications realm, it's likely that the human value we'd ascribe on average to each packet is far lower than our current Internet IPv4 world of human-mediated communications. And it's this extravagantly device-equipped world that we see the U.S. Department of Defense heading. If your stock in trade is one of quite astounding feats of logistical deployment of large numbers of people and large numbers of items of equipment, then the communications requirement is of a different order of scale to that of the retail Internet markets, and, yes, I'm sure that there are entirely effective arguments behind that decision to look forward to a communications realm with a uniform base protocol identifier domain in a scale that is 2 to the power 96 times larger than the entire IP address identifier domain of IPv4.

I would be cautious about high levels of expectation that this immediately translates into an impetus in the market where you and I converse. My host here where I'm typing this message is already IPv6 capable, and if you are running a recent version of host software, then it's a reasonable assumption that yours is too. But I'll send this message over IPv4 and you'll receive it over IPv4, and between my mail sender and your mail receiver the transport channel will also be IPv4. Should we use IPv6 instead? Would I pay my provider additional money to compensate it for part of its additional expenditure to support a simultaneous IPv6 capable network between you and me? To send precisely the same message? In precisely the same time? Along the same path? Using the same transport TCP session? Obviously, to me, as a (hopefully) economically rational consumer of such services, and no doubt to you, in a similar role, there is no value in spending more money to achieve outcomes in IPv6 that are identical to what we can already do today in IPv4. And in the retail Internet world that remains the basic IPv6 conundrum. Why should any provider spend additional resources to service the same market with identical services, and in so doing be unable to raise additional revenue to offset their additional service costs? One interpretation is that there is no natural motivation for such activities in today's market, otherwise it would already be very widespread indeed.

What we've seen in the mainstream Internet world is an emerging mythology about IPv6 that somehow this additional expenditure, ultimately on the part of the consumer, provides some additional benefit for the consumer, motivating them to switch from IPv4-only services to some hybrid of mixed v4 and v6 and ultimately to a v6 world, and thereby funding the additional provider expenditure associated with such a massive transition.

The reality is more sobering in that in the retail Internet world there is so far nothing obvious in the "additional benefit" category. I'm using Network Address Translation (NAT) right now, using an ssh session back to my mail server that drives through NAT boxes to make a secure SMTP session, across a first step of 802.11 wireless in order to pass this message into a mailing list. I've auto-configured my laptop in the wireless world, and for me I'm living in a plug-and-play world that supports my level of roaming access. Would IPv6 make this session any more secure? Any different in terms of Quality of Service (QoS) ? In plug-and-play models of roaming? Would there be any visible difference in terms of my ability to communicate with you? To all of these questions the basic answer is still "no."

So, for you and I, we look inside the IPv6 technology box, and find nothing new there to motivate us to spend more money for our existing Internet-based communications services, and for some time to come it would appear that this limitation will still hold.

On the other hand there are circumstances where there is a need to operate in a much larger base protocol address space. These include situations where one wants to take advantage of Internet applications that operate across a world of literally billions of devices, large and small. The application space may want to gather constant reports on the characteristics of the "thing" it is attached to, from a ration pack to a component of a large naval vessel. You may want to use supply channels for such devices such that the deployment is a plug-and-play world without a massive variety of detailed configuration processes. You may be looking to an architecture that would be stable for many years. In such circumstances you really want take advantage of a uniform set of Internet application technologies that potentially span massive numbers of addressable devices. Here a large base address space is a definite asset. And for such industry sectors in voicing such requirements where there is also a somewhat different ultimate value proposition for the supported communications activity, then it's quite understandable that there can be an attractive proposition offered by immediate adoption of IPv6.

But back in the communications realm where you and I currently exchange our messages, such requirements remain in a future framework that is still waiting for relevant value propositions that allow it to gain traction with you and me.

Maybe we just need to be patient. Steam ships did not halt operation the first day a diesel powered vessel appeared. It was a much slower process that lead to an outcome of the change of the maritime fleet. The

next generation of mechanization of naval vessels offered cheaper services, and, as often happens, market price won in that commodity market.

Market price often wins in competitive commodity markets. And the Internet retail market is, in many parts of the world and in many sectors, a strongly competitive space with all the characteristics of a commodity offering. And there no doubt that if you and I could communicate in precisely the same fashion as we do today, with precisely the same applications and service environment, using precisely the same host devices and operating systems as we do today, but at some attractive fraction of today's price, then I'm sure that neither of us would care in the slightest that our data was encapsulated using a packet framing format and address tokens that used the IPv6 protocol specifications.

Disclaimer

The views expressed are the author's and not those of APNIC, unless APNIC is specifically identified as the author of the communication. APNIC will not be legally responsible in contract, tort or otherwise for any statement made in this publication.

About the Author

Geoff Huston B.Sc., M.Sc., has been closely involved with the development of the Internet for many years, particularly within Australia, where he was responsible for the initial build of the Internet within the Australian academic and research sector. He is author of a number of Internet-related books, and has been active in the Internet Engineering Task Force for many years.

www.potaroo.net