

The ISP Column

An occasional column on things Internet

November 2006

Geoff Huston

Whither Routing?

I attended a Internet Architecture Board workshop in October 2006 on the topic of routing and addressing, and in this column I'd like to report on the impressions on this topic that I took away from this workshop. I should note at the outset that this is not in any way an official report of the meeting, nor even a report of my impressions of the presentations and the related workshop discussion. This article could best be described as my reactions to what I heard at this workshop and what I think they might imply in the area of routing and addressing.

Routing is, of course, a central concept within the operational Internet. Routing provides the essential "how" information that allows IP packets to be sent to their ultimate destination. A routing protocol defines an information exchange that feeds a distributed computation. The local component of this distributed computation maintains a routing information base that is then passed into a local forwarding data structure. The intended outcome of this distributed computation is for every router to be able to maintain a local packet forwarding state that is consistent with both its neighbouring routers' forwarding states and that is also consistent in a larger network-wide context. The result is that each local forwarding decision passes the packet closer to its ultimate destination. Indeed, in a "perfect" routing world each forwarding system would be optimal, in that the local decision would be to forward the packet as close as it possibly can to its ultimate destination.

The tough part, particularly in inter-domain routing is figuring out what "closer" really means. Certainly "closer" implies avoiding routing loops. But what else does it, or should "closer" imply?

Is it minimizing the time to deliver the packet to its destination, or minimizing the potential for jitter (queuing delay) or loss, or using a path with the highest possible bandwidth, or using a path that matches the path taken by other packets in the same "stream", or is "closer" based on some other service attribute of the path?

Right now BGP uses a very simple common base metric, in that a BGP forwarding decision should, by default, prefer paths with fewer domain transits over more, and refinements to this selection algorithm become a local matter. So when an application wants a particular network service response from the inter-domain forwarding system there is some considerable confusion over how this could, or should, be achieved.

The routing protocols in use within the Internet these days are, by Internet standards very mature technologies. The current suite of routing protocols, including BGP 4 for inter-domain routing and the family of interior routing protocols, have enjoyed extensive deployment experience that extends back over some 15 years or more. This family of protocols has seen the Internet's inter-domain routing space grow from its early base of a few thousands of routing entries to the current population of some two hundred thousand entries in October 2006.

One of the more interesting aspects of the Internet Architecture is the decomposition of routing and forwarding. The role of a routing protocol is to populate the forwarding table, and IP in and of itself does not define any particular routing protocol to undertake that role. Over the years a considerable number of protocols have been deployed, and certainly one of the potential degrees of freedom when considering Internet routing protocols is that there is no intrinsic architectural barrier to considering any form of routing protocol.

But as we look forward some rather tough questions tend to reveal themselves. How large and how diverse a network can the current routing protocols and their hardware forwarding platforms support? Are the ways we use routing today making these questions more acute? How long can our current technology base in routing and forwarding last? And can we make changes, either incrementally or at a quite fundamental level, that can provide some surety of cost effective routing viability over a longer timeframe?

Of course the information that routing protocols manipulate are address prefixes, as it is these prefixes that are passed into the forwarding table. The manner in which addresses are distributed through the network has a direct impact on certain critical aspects of the routing function. What characteristics of the address distribution make these routing questions easier or harder? How can addressing and routing work together?

While these questions are of some interest today, they are of very considerable interest when looking forward. When we look at the major shift in the size of the address pool with IP version 6 (IPv6), and make the assumption that our use of addresses within routing will need to encompass not just the computer-dominated network of today, but a truly massive universe of devices. These device addresses may be embedded at time of manufacture, and their ultimate deployment location is unknown, and a significant fraction of these devices may be in pockets, cars or luggage tags. So when considering the impacts of IPv6 on routing and addressing then we also need to wonder about how well other factors, such as device deployment and device mobility, can be factored into the routing question.

And, of course, this is no longer a small network. It supports an entire global communications enterprise, and such an industry does not make fundamental technology changes quickly, nor easily, nor cheaply.

Some Background

Of course this topic of routing, addressing and their inter-dependencies, is not a novel one, as one could argue that routing and addressing form the fundamental inputs in any communications system. In this respect the Internet is no different, and these considerations of routing and addressing have

formed a significant part of the internet's technology history, stretching back at least 30 years. Internet Experimental Note 19 (IEN19), authored in January 1978, includes a discussion of the effects of matching hierarchical addressing to hierarchical routing topology, and the impact of this on network reliability.

IPv6 was one outcome of a relatively intense effort on the part of the IETF through the 1990's in looking at the issue of address exhaustion in the IPv4 network. Another outcome of that same effort was a short-term mediation response to routing scaling with the shift to provider-based classless address distribution and routing (CIDR). The operational Internet commenced deployment of CIDR into the routing system in 1994, and some would argue, the effort still continues.

Of course the background to the routing and addressing question is much larger than could fit into this column. An excellent compilation of material was made available to the IAB workshop (see the Reading List at the end of this article)

What are Today's Burning Routing Topics? Some Perspectives

The IP architectural use of addresses in dropping of the distinction between the "who", "where" and "how" of networking is one of the compromises that I believe made IP such an elegantly simple architecture. In IP an "address" is a token that serves all three purposes. Of course part of the tension in today's routing and addressing world is that it is becoming increasingly apparent that this assumption does not provide the same level of leverage in a truly massive and diverse networked world. "Identity" is a persistent state that wants to remain unchanged irrespective of location, while "Location" is relative to the network's topology, and a forwarding path is a path that is relative to both the sender and the receiver.

The result of our inability to balance these tensions relating to identity, location and path within the "address" is best seen in any time series measurement of the size of the Internet's inter-domain routing table. The size of the routing table inflates inexorable, the number of BGP protocol messages inflates inexorably, and the class of persistent instabilities within the routing system continues unabated (one such time series of routing table growth is shown at <http://bgp.potaroo.net>).

One perspective is that the addition of multiple routed technologies into this mix just adds more fuel to what appears to be a well ignited pyre of routing inflammation. Inside any transit IP service provider the combination of IPv4 reachability routes, Traffic Engineering routes, both internally and advertised by others, various flavours of Virtual Private Network routes, and internal routes, is already daunting in terms of scaling in the near term future. Adding IPv6 routes takes a domain that is populated by low hundreds of thousands of entities and, within some projections, appears to double the routing population size, to some extent or another.

Of course we've all heard this rapid Internet growth stress before, and while there have been some notable routing incidents, a truly widespread routing disaster hasn't happened yet, as far as we understand. The general observation appears to be that in response to scaling pressure in the routing system Moore's Law (see: http://en.wikipedia.org/wiki/Moore's_law) has come to the rescue every time, and there is no particular empirical reason to doubt that once more we will be rescued just in time (see: <http://www.cs.ucl.ac.uk/staff/M.Handley/papers/only-just-works.pdf>).

One workshop presentation, however, I found particularly intriguing. So far we've seen improvements in speed, improvements in efficiency and decreases in unit costs of packet switching. Moore's Law has not just rescued us each time, it's managed to provide continual reinforcement that volume matters, and bigger is not only better, but in unit cost terms, cheaper. "Larger" is "better" in terms of costs and operating efficiencies, or at least that's been our general experience to date. But router hardware is not necessarily mainstream computing technology hardware, and the lower production volumes of

router-specific chips tend to make the manufacturing price higher and the technology cycles longer. Also, Moore's law has applied well, so far, to device density on a silicon wafer, but it does not apply to clock speeds of the chip, and also while DRAM capacity has grown at a rate of 2.4 times every 2 years, the memory cycle time, or speed, appears to have grown by a much smaller factor of around 1.2 every two years. As routing and forwarding applications compute across ever rising populations, they require both more memory and faster memory access speeds just to stay even at a constant cost point, let alone managing to leverage further efficiencies in unit cost. The silicon industry appears to have recently focussed on lower power designs because of the predominance of manufacture of chips for portable battery-powered devices, and making some amount of tradeoff of total processing capacity in the process. Big, fast SRAM devices appear to no longer sit in the mainstream high volume, lower unit cost path of the silicon manufacturing industry.

The forwarding engines in routers moved away from software on generic CPU engines many years ago, and the router designers headed towards ASICs to achieve the necessary performance of this specialized function. Again, the low production volumes here and the short lifespan of each technology cycle of these router-based ASICs lead to the current observation that we may have already moved past the price/performance inflection point, in which case higher performance only comes at higher per packet switching costs. The observation made in this presentation was that while chip device density is increasing by a factor of 2x every 2 years, in line with Moore's Law, the net per chip cost is rising at 1.5x over the same 2 year period. The implication is that constant cost per packet of forwarding in routers implies a growth of no higher than 1.5x every 2 years in the size of the forwarding decision space.. Factoring in additional capital expenditure and operational expenditure costs associated with larger and more power-hungry units may bring that constant real cost point down to some 1.3x every 2 years.

The inference I heard from this consideration is that constant convergence times in terms of router capacity to process routing updates implies a growth in the forwarding space of some 1.2x every two years, while the constant unit cost ceiling is at a maximum growth trajectory of 1.3x every 2 years. So far, over the past 7 years we've seen growth factors of between a peak of 2x every two years and a minimum of 1.3x. The inference here is that continued unbounded growth in routing that implies similar growth in forwarding data table sizes implies that the unit costs of forwarding Internet traffic increases. This enters a situation where bigger and bigger networks become more and more inefficient in terms of unit costs. When this data is coupled with considerations of power supply and heat loads of switching equipment, pin densities, edge-to-edge device physical sizes and clock times, and the serial nature of many critical elements of forwarding function, then the total picture is looking uncomfortable for a future larger network. When we include the logistics of supplying adequate power to the switching unit, and then removing the resultant generated heat from the chips, from the equipment chassis and the building itself, then one emerging view is that we may have already reached some practical air-cooling limits and DC rail power limits here.

It appears that unbounded continued growth of the routing and forwarding system in the Internet appears to trigger off some real limitations relating to hardware design and switching centre infrastructure, particularly relating to the power and cooling subsystems.

What are our Routing Problems?

Well maybe that's not quite the right question, as the perspectives on the problem space of routing, and the broader space of routing and address distribution functions when combined, are indeed many and varied. The IAB workshop posed a somewhat different question about the common perception of the most critical aspects of routing as we see them today.

The most critical aspect of routing was seen as scaling. Could we really support a routing table twice the size of today? Or how about ten times the size? Or even larger? How long before we are

confronted with that size of routing space? What may need to be the engineering trade-offs to cope with such a larger routing space? Obviously it would be wonderful to have a routing system that somehow manages to become larger, faster, cheaper, and more stable all at the same time, but there is a general view that this is not going to be possible. As we scale the network in size, the current properties of routing will need to change, and possibly we may need to revise our expectations about some aspects of speed, cost and performance. Indeed we may need to do more than that and even get to the point of revising our model of routing and the protocols that support it. Or even take a critical look at some of the underlying architectural properties of IP itself.

So scaling is hard, and the common perception is that scaling is the driving impetus behind this entire effort of looking once more at routing and addressing.

What else is a "problem"?

It appears that convergence speed is a problem for many. Slow convergence to a consistent forwarding state implies dynamically shifting forwarding paths, which, in turn, may impact the performance of certain applications, particularly those that are highly sensitive to jitter or packet loss.

Traffic engineering, or the ability to exercise some control over the traffic loads imposed upon various network paths is also a problem for many, particularly so where the level of inter-provider connectivity is dense. Rather than allowing a default routing configuration to direct all traffic to a small set of "best" paths, the desire in a traffic engineering environment is to perform some form of load distribution of the traffic across a broader set of paths. Of course, in the spirit of if the only tool you have is a hammer then everything looks like a nail, then if the only tool you have is routing, then traffic engineering requirements are expressed in terms of dynamic adjustment of routing parameters, normally through increasing the number and diversity of advertised routing entries for each network provider or site.

And, of course, there's the problem of routing security, or, more particularly, the general lack of it in today's routing environment. Attacks on the integrity of application behaviour can start with an initial lever of traffic redirection, and one of the more effective ways to achieve that is through subversion of the routing system.

That's by no means an exhaustive list, of course. The issues of the lack of natural constraint in the routing system, imperfect, or even non-existent feedback systems, Quality of Service routing, or the lack thereof, intrinsic support for various form of mobility, the capability of BGP to converge to unintended routing states, the lack of meaningful path metrics in the inter-domain space, to name but a few more are also "problems" with routing as we know it today.

If the current "problems" in the routing system are phrased as those aspects of the system that fail to meet our expectations or desires, then its clearly the case that the routing system appears to be replete with such problems. However, that's not a very satisfactory point to reach, and it is also necessary to understand which of these problems are seen as being critical to the future Internet, and which fall into the space of being desirable, but perhaps not essential for all.

It was clear at the workshop, and probably clearly evident elsewhere, that if there is a highest ranked "problem" in the routing space then it would be that of scaling the routing system. If we cannot scale a single cohesive consistent routing framework for the Internet, then the most likely outcome is that sooner or later we'll end up with a number of fragmented Internets, all loosely connected in various ways inside application layer gateways. Anyone who remembers the email chaos of the 1980's probably has no desire to re-live the experience. And anyone who is of the view that the most enduring value within the Internet architecture lies within a model of clear edge-to-edge interaction

across a network infrastructure would look at such an outcome and shudder. Indeed, even today's NAT-dense networked world looks preferable from the end user perspective. If Metcalfe's Law (http://en.wikipedia.org/wiki/Metcalfe's_law), that the value of a network is in proportion to the square of the number of its users, is indeed the case, then such a fragmentation of the network would represent a significant devaluation of the value of the Internet, particularly for those applications that are not provisioned in any such inter-Internet gateways. And of course innovation though the introduction of new applications and services would encounter significant adoption barriers. From such a state, ossification of the Internet would not be far away.

So scaling the Internet and its routing system is certainly important. But maybe its more than this. Just dumping more prefixes, and more updates, and more attributes, into the routing system without any form of constraint or mediation, without incremental cost, and without care, looks like a losing proposition. The technology will not get cheaper as the forwarding space gets ever larger, and indeed it may head in the other direction and expose higher incremental unit costs as we continue to scale the network. The question that this exposes is: how can we envisage a vastly larger routing system? How can we achieve a common beneficial outcome across a loosely coupled collection of players who appear to react to various economic and technology signals in very diverse ways?

What are some of the Routing Levers?

What are some of the levers that we can pull to achieve a constrained routing system that does not bloat into an untenable size? As this was workshop within the broad umbrella of the IETF, then it should be no surprise to learn that the approach tended to concentrate on technology levers rather than levers of an economic or regulatory nature. And that's probably as it should be, in that an economic approach is one that often ends up rationalizing outcomes rather than exploring different approaches. The Internet is an activity that is heavily grounded in technology, and it is technology that tends to provide the driving impetus for change. Having said that its also appropriate to recognize that as the Internet grows ever larger it is becoming impervious to even these technology signals.

Addressing appears to be a very central part of the story here. A central proposition here, that I've always heard attributed to Yakov Rekhter, states that addresses can follow topology, or topology can follow addresses, but you can only pick one.

Making addresses follow network topology has proved problematical at the edges of the network. Such a principle of topologically-aligned addresses has become adopted as "provider-based addressing", where customer networks are numbered from the provider's address block. But what if you want to change providers? Then, unfortunately, as a customer you may be forced to renumber into a new address block provided by your new provider, and renumbering is hard. Addresses are configured in hosts, in routers, in firewalls, in caches, and in a myriad of other hard-to-find places. Little wonder that many enterprises have chosen private numbering within their network and interface to the world via NATs and a heavily constricted service interface. But what then about IPv6? What of the concept that NATs are an artefact of address size limitations of IPv4, and that the larger address space of IPv6 makes the entire concept of NATs supposedly historic? At this stage the issues of the cost of renumbering come to the forefront again, and, in spite of the toolkit of dynamic host configuration, router discovery, incremental DNS updates and similar, addresses still creep into static configurations, and as a result, renumbering still remains hard. And of course the associated issue is that of so-called "multi-homing", where an end site is configured with multiple providers. The motivations for such a configuration may be for service resiliency, or for avoiding single provider lock-in, or it may be because the end-site is a widely distributed one. For whatever reason multi-homing is a distinct factor here, and in this case provider-based addressing assumes some very challenging proportions.

If we want addressing to follow topology, but at the same time want end sites to be able to use provider-independent address blocks, and at the same time also want a scaleable routing system,

then perhaps are we asking for too much? This leads into a deeper area of IP architecture, namely that the semantics of an IP address are indeed overloaded., and making addresses jump through the "who", where" and "how" hoops simultaneously is just not scaleable in the long run.

This line of thought has lead to various approaches to disambiguate the "who" from the "where" and "how". This is commonly termed the "ID/LOC split", where identity tokens (IDs) are persistent and remain firmly associated with some form of "endpoint" and locations (LOCs) are used in the routing and addressing system, and are able to follow topology in a high consistent fashion. While this is not a new concept, it is only recently that we've seen some interesting steps in exploring what this may mean in terms of a protocol stack implementation. The work in HIP, and SHIM6 are both recent efforts to create protocol frameworks that allows an application to maintain fixed IDs as the "conversation" identifiers of each party, while allowing the lower levels of the protocol stacks have a more fluid view of the current preferred locator to use to reach the other party. Its also the case that work in mobility protocols in IP start from the basic condition that identity and location are distinct concepts. Its early days for these type of approaches to the ID/LOC, and their ability to leverage some favourable outcomes in routing is clearly an unknown factor at present. Perhaps one could see these efforts as yet another instance of the magic incantation that "there is no problem in computing science that cannot be solved by adding another layer of indirection!"

Its not understood whether the entire routing system is in sore need of an entirely novel routing architecture. There have been a number of proposals to re-examine the routing system and take a different approach to the tokens that are manipulated in the routing protocol, and the way in which forwarding tables are populated and used. However, on the whole, these exercised have been confined to more theoretic approaches so far. A couple of these approaches appear to be very intriguing, however. One is to use yet another layer of routing aggregation above the current prefix block, and treat all prefixes that share a common routing policy as an instance of one of these higher level aggregated entities. The salient observation is that while the Internet's routing table has 200,000 entries and a single eBGP peering session generates of the order of 300,000 prefix updates per day, the true level of effective forwarding diversity at any point of routing is far closer to a number in the thousands, or between 2 to 3 orders of magnitude smaller. Another intriguing approach relates to "stretch" and the observation that routing protocols work hard to present the best possible path to the forwarding system for each routed address prefix. What happens if you stretch this "best path" constraint to something a little more flexible? The theoretic result points to a potential for a dramatic reduction in the amount of routing information exchange to arrive at an acceptable forwarding state.

So where are we?

I suspect that identifying scaling as the major issue for the routing system is correct, and does not deviate from similar exercises undertaken in the past. The IETF reached a similar conclusion in the early 1990s, with the ROAD effort, and, at the time, undertook CIDR as a short term expediency and then worked on developing IPv6 as a longer term addressing approach. But the subsequent intention to look at the longer term options for routing evidently dissipated, perhaps because CIDR was simply sufficiently effective at the time to reverse the more critical aspects of routing inflation.

But CIDR is not as effective any more. Over half the entries in the routing table are more specifics of covering aggregates. The routing update load profile tends to relate to some form of intrinsic instability of these more specific advertisements. Its not that CIDR itself has failed, but the demands of an increasingly rich interconnection environment and the associated desire to perform load balancing and traffic engineering across this dense interconnection mesh has motivated many network operators to walk away from the highly aggregated routing environment that CIDR represented. In addition, there are pressures on addressing. These include the coupling of the desire not to re-introduce NATs as an integral part of the IPv6 architecture, while at the same time recognising that renumbering is a considerable cost imposition to end sites, and relating this with the increasing use of

multi-homing at the edges of the network all point to a network world that views strict provider-based address aggregation in routing as an unattractive imposition.

Its also the case that we are nowhere near a “new” inter-domain routing protocol, nor are the network’s endpoints poised to deploy stack software that supports any form of ID/LOC split. These technologies will take further time to refine, and even longer to deploy across a sufficiently large portion of the network to make a real impact on the routing system.

At the same time we are facing some readily identifiable scaling problems at the router hardware level, and the unconstrained continued bloating of the routing system may lead us inexorably towards a rather undesirable position of forced network-level fragmentation.

Its clear that more work needs to be done here, and also clear that doing nothing is in and of itself a form of decision that may lead to a less desirable outcome of network partitioning and fragmentation. Its also evident that this workshop appears to have raised a level of momentum to undertake some further work in this area.

Personally, I suspect that its often easier to fix the most immediate and pressing problem. I suspect that the hardware scaling issues may be tractable through increased indirection between the routing and forwarding systems, with techniques that can lead to FIB “compression,” to the effect that the same forwarding decisions can be made as at present, but with a much smaller and faster search space in the forwarding information base to get there. On the other hand, we also have to recognise that the Internet has attained a critical role in global communications infrastructure, and there are many stakeholders who would prefer to see some consideration given to a longer term approach to these issues.

This latter option is difficult right now, in that there is much we really do not know about the behaviour of today’s routing system, and longer term approaches are often flawed if they are based on poor or incomplete premises. It’s a shame that the level of funding for research in routing appears to have dropped so dramatically over the past decade. This is definitely not a solved problem, and its going to take more than the research resources of the vendor and service provider sectors to work through this topic. From my understanding of the situation, I would strongly support the proposition that there should be some substantive research input to this problem, and in general the optimal time for such input is in the early stages of the effort, rather than afterwards!

Further Reading:

A report on the IAB workshop will be forthcoming – this document will be updated to include a pointer to that material when the report is published.

Elywn Davies compiled a reading list for the Workshop . This quite extensive reading list is based on that compilation.

30 years on

- **A Note on Inter-Network Naming, Addressing, and Routing**, IEN 19, John Shoch
(<http://ana-3.lcs.mit.edu/~jnc/tech/ien/ien19.txt>)
- **Why the Internet Only Just Works**, Mark Handley
(<http://www.cs.ucl.ac.uk/staff/M.Handley/papers/only-just-works.pdf>)

Background

- **Analysis of IDR requirements and History**
(<http://www.ietf.org/internet-drafts/draft-irtf-routing-history-03.txt>)
- **Endpoints and Endpoint Names: A Proposed Enhancement to the Internet Architecture**, Noel Chiappa
(<http://users.exis.net/~jnc/tech/endpoints.txt>)
- **Will The Real "End-End Principle" Please Stand Up?**, Noel Chiappa
(http://users.exis.net/~jnc/tech/end_end.html)

Previous Sets of Requirements

- **Goals and Functional Requirements for Inter-Autonomous System Routing** RFC1126, Mike Little
(<http://www.ietf.org/rfc/rfc1126.txt>)
- **Requirements for Inter-Domain Routing**, Avri Doria et al.
(<http://draft-irtf-routing-reqs.potaroo.net>)

Problem Statements

- **Some Foundational Problems in Interdomain Routing**, Nick Feamster et al.
<http://nms.lcs.mit.edu/papers/camera.pdf>
- **Beyond Interdomain Reachability**, Olivier Bonaventure et al
<http://www.net.informatik.tu-muenchen.de/wired/position/bonaventure-wired.pdf>

Workshops

- **Workshop on Internet Routing Evolution and Design**, 2003
(<http://www.net.informatik.tu-muenchen.de/wired/>)
- **Workshop on Internet Routing Evolution and Design**, 2006
(<http://www.wired2006.org/>)
- **Inter-Domain Routing Workshop**, 2004
(<http://www.tm.uni-karlsruhe.de/idrws2/submissions.php?year=2004>)
notes by Geoff Huston
(<http://www.potaroo.net/ispcol/2004-03/2004-03-isp.htm>)
- **Clean Slate Network Research Workshop**
(<http://www.cl.cam.ac.uk/research/srg/netos/cleanslate/cleanslate.html>)
Report by John Crowcroft and Peter Kay
(<http://www.cl.cam.ac.uk/research/srg/netos/cleanslate/CCR-report.pdf>)

The State of the Routing World

- **The BGP Report**
(<http://bgp.potaroo.net>)
- **BGP Update Report**
(<http://bgpupdates.potaroo.net>)
- **The BGP Report for 2005**
(<http://www.potaroo.net/ispcol/2006-06/bgpupds.html>)

Routing Schemes

- **Nimrod**, Noel Chiappa
(<http://ana-3.lcs.mit.edu/~jnc/nimrod/docs.html>)
- **Pip**, Paul Francis
(<http://portal.acm.org/citation.cfm?id=190325&coll=portal&dl=ACM>)
(<http://www.ietf.org/rfc/rfc1621.txt>)
(<http://www.ietf.org/rfc/rfc1622.txt>)
- **GSE--An Alternate Addressing Architecture for IPv6**, Mike O'Dell
(<http://draft-ietf-ipngwg-gseaddr.potaroo.net>)
- **The Map & Encap Scheme for scalable IPv4 routing with portable site prefixes**, Steve Deering
(<http://www.cs.ucla.edu/~lixia/map-n-encap.pdf>)
- **New Scheme for Internet Routing and Addressing (ENCAPS) for IPNG**, Bob Hinden
(<http://www.ietf.org/rfc/rfc1955.txt>)
- **An Overview of Multihoming and Open Issues in GSE**, Lixia Zhang
(http://www.cs.ucla.edu/~lixia/0609GSE_Overview.pdf)

More Recent Research Proposals

- **Meta-Routing**, Tim Griffin and Joao Luis Sobrinho
(<http://www.acm.org/sigs/sigcomm/sigcomm2005/paper-GriSob.pdf>)
- **Tim Griffin's Web Page**
(<http://www.cl.cam.ac.uk/~tgg22/>)
- **HLP: A Next Generation Interdomain Routing Protocol**, Lakshminarayanan Subramanian et al.
(<http://www.acm.org/sigs/sigcomm/sigcomm2005/paper-SubCae.pdf>)

Overlay Networks and Policy Distribution Networks

- **Resilient Overlay Networks**, Andersen et al
(<http://nms.lcs.mit.edu/papers/ron-sosp2001.pdf>)
- **PlanetLab PLUTO Routing Overlay Networks Project web page**
(<http://www.planet-lab-jp.org/pluto/routingoverlay.html>)
- **Scalable Routing Overlay Networks**, Akiro Nakao et al.
(<http://portal.acm.org/citation.cfm?id=1113372>)
- **OPCA: Robust Interdomain Policy Routing and Traffic Control**, Sharad Agarwal et al.
(<http://sahara.cs.berkeley.edu/papers/ACK03.pdf>)

Compact Routing

- **Toward Compact Interdomain Routing**, Dmitri Krioukov and kc claffy
(http://arxiv.org/PS_cache/cs/pdf/0508/0508021.pdf)
- **Compact Routing with Minimum Stretch**, Lenore Cowen
(<http://www.eecs.tufts.edu/~cowen/sodasap.ps>)
- **Compact Routing on Internet-Like Graphs**, Dmitri Krioukov, Kevin Fall, Xiaowei Yang
(<http://citeseer.ist.psu.edu/651100.html>)
- **Compact Name-Independent Routing with Minimum Stretch**, Ittai Abraham et.al.
(<http://www.cs.huji.ac.il/~ittaia/papers/AGMNT04.pdf>)

Policy and Configuration Problems in BGP4

- **Towards a Logic for WideArea Internet Routing**, Nick Feamster and Hari Balakrishnan
(<http://www-static.cc.gatech.edu/~feamster/publications/logic-sigcomm2003.pdf>)
- **Verifying the Correctness of Wide-Area Internet Routing**, Nick Feamster and Hari Balakrishnan
(<http://www.lcs.mit.edu/publications/pubs/pdf/MIT-LCS-TR-948.pdf>)
- **Detecting BGP Configuration Faults with Static Analysis**, Nick Feamster and Hari Balakrishnan
(<http://www-static.cc.gatech.edu/~feamster/publications/rcc-nsdi-camera.pdf>)
- **BGP routing policies in ISP networks**, Matthew Caesar and Jennifer Rexford
(<http://www.cs.princeton.edu/~jrex/papers/policies.pdf>)

-
- **Nick Feamster's Web Pages**
(<http://www-static.cc.gatech.edu/~feamster/>)
 - **Cabo Paper**
(<http://www-static.cc.gatech.edu/~feamster/papers/cabo-tr.pdf>)

The Effects on Routing Hardware

- **An Empirical Study of Router Response to Large BGP Routing Table Load**, Di-Fa Chang et. al.
(<http://www.isi.edu/~johnh/PAPERS/Chang02a.pdf>)

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