

The Internet Protocol Journal

December 2011

Volume 14, Number 4

A Quarterly Technical Publication for
Internet and Intranet Professionals

In This Issue

From the Editor	1
Port Control Protocol	2
Challenges to DNS Scaling	9
Networking @ Home	15
IETF Tools	21
Fragments	25
Call for Papers	31

FROM THE EDITOR

Depletion of the IPv4 address space and the transition to IPv6 has been a “hot topic” for several years. In 2011, interest in this topic grew considerably when the *Asia Pacific Network Information Centre* (APNIC) became the first *Regional Internet Registry* (RIR) to start allocating addresses from its final /8 IPv4 address pool. Although depletion dates are difficult to predict accurately, there is no question that the day will come when it will no longer be possible to obtain IPv4 space from the RIRs. News stories about IP addresses being sold for considerable sums of money are becoming more common.

Numerous organizations have been working diligently to promote, test, and deploy IPv6 through efforts such as the *World IPv6 Day*, while the *Internet Engineering Task Force* (IETF) continues to develop solutions to aid in the transition. One such effort, the *Port Control Protocol* (PCP), is described in our first article by Dan Wing.

The *Internet Corporation for Assigned Names and Numbers* (ICANN) will soon begin accepting applications for new *Top-Level Domains* (TLDs). It is not yet known how many new TLDs will eventually be deployed, but the plans have prompted several studies focused on the resiliency and scalability of the *Domain Name System* (DNS). Bill Manning discusses some of the technical challenges associated with a vastly expanded TLD space.

The *IETF Homenet Working Group* “...focuses on the evolving networking technology within and among relatively small ‘residential home’ networks. For example, an obvious trend in home networking is the proliferation of networking technology in an increasingly broad range and number of devices. This evolution in scale and diversity sets some requirements on IETF protocols.” Geoff Huston gives an overview of some of the challenges facing this Working Group.

The product of the IETF is a set of documents, mainly protocol specifications and related material. These documents start life as *Internet Drafts* and proceed through a series of iterative refinements toward eventual publication as *Request For Comments* (RFCs). Over time, several *tools* have been developed to aid in the document development process, and they are now organized at the IETF Tools webpage. We asked Robert Sparks to give us an overview of some of the most important tools and the process involved in their development.

—Ole J. Jacobsen, Editor and Publisher

ole@cisco.com

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

ISSN 1944-1134

Port Control Protocol

by Dan Wing, Cisco Systems

After the transition to *Internet Protocol Version 6* (IPv6), hosts will often be behind IPv6 firewalls. But before the transition, mobile wireless devices will want to reduce their keepalive messages, and hosts of all sorts will share IPv4 addresses using a variety of address-sharing technologies. To meet these needs, the IETF formed the *Port Control Protocol Working Group* in August 2010 to define a new protocol for hosts to communicate with such devices. The initial output of this Working Group is the *Port Control Protocol* (PCP)^[1]. Interoperability between two independently developed implementations of PCP was demonstrated at the IETF meeting in July 2011, highlighting the importance of this protocol to the industry. After it becomes a standard, PCP is expected to be deployed in various operating systems, IPv6 home gateways, IPv4 home gateways (*Network Address Translators* [NATs]), mobile third- and fourth-generation (3G and 4G, respectively) gateways (*Gateway GPRS Support Nodes* [GGSNs]), and *Carrier-Grade NATs* (CGNs).

Introduction to PCP

PCP performs two major functions: It allows packets to be received from the Internet to a host (such as to operate a server), and allows a host to reduce keepalive traffic of connections to a server. PCP can be extended in two ways: with new *OpCodes* or with new *Options*. The base PCP specification defines two OpCodes: MAP and PEER, and defines several Options that can be carried with those OpCodes.

To operate a server, packets are sent from a host on the Internet to a server. The IP model expects devices to be connected to a network and be able to exchange packets with each other. However, few deployed networks actually permit hosts to receive packets from the Internet because of business needs (for example, to protect wireless spectrum from malicious or accidental packets originated on the Internet) or because of technology restrictions (for example, IPv4 address-sharing devices such as *Network Address and Port Translators* [NAPT]). To operate a server, a host uses the MAP OpCode.

To reduce keepalives, a host needs to send traffic before a middlebox will destroy an idle connection. Many middleboxes, such as firewalls or NATs, maintain state and will destroy mappings if the connection has been idle. Today, in order to prevent destruction of mappings, hosts send keepalive traffic to keep those mappings alive. The keepalive traffic has several disadvantages, including reduction of battery lifetime, network chatter, and server scalability (servers have to discard the keepalive traffic). PCP allows a host to determine how aggressively a middlebox will destroy an idle connection, allowing the host to reduce its keepalive traffic with the PEER OpCode.

PCP is encoded in binary and carried over the *User Datagram Protocol* (UDP), which eases implementation on clients and servers. The client is responsible for retransmitting messages, and all messages are idempotent. The PCP client can be part of the operating system (much like a *Dynamic Host Configuration Protocol* [DHCP] client or a *Universal Plug and Play* [UPnP] *Internet Gateway Device Protocol* [IGD] client) or the PCP client can be coded entirely in an application (much like any other application-level protocol such as the *Network Time Protocol* [NTP]). A major feature of PCP is its flexibility and simple messaging, so it can be implemented easily in a variety of systems and at high scale.

Security

When installing an IPv4 NAT on a residential network, the NAT has a side effect: it prevents unsolicited incoming traffic from reaching hosts inside the home. Traffic that originates inside the home can traverse the NAT toward the Internet. This function is expected by many users to such a degree that when IPv6-capable routers were first installed on residential networks, users complained that their IPv6 hosts were seeing traffic from the Internet. This visibility meant that IPv6 printers, webcams, and other hosts had to be protected from malicious traffic from the Internet. Based on this experience, IPv6 *Customer Premises Equipment* (CPE) routers intended for installation in the residential market filter most unsolicited incoming traffic by default^[3]. Thus, IPv6 CPE routers provide filtering similar to what users experience today with IPv4 NAT devices.

With both IPv4 NAT and RFC 6092 IPv6 routers, outgoing traffic from a host creates a mapping that then allows bidirectional traffic to a specific (*Transmission Control Protocol* [TCP] or UDP) port on the internal host, meaning when a host sends a TCP SYN, a SYN ACK can be returned to the host. Neither IPv4 NAT devices nor RFC 6092 IPv6 routers have to do any additional filtering of that mapping, and after that mapping is created will allow traffic from any host on the Internet to reach the internal host—not just traffic from that particular host. This lack of filtering is necessary for certain applications to function.

PCP was built with a security model similar to that deployed on home networks. With PCP, a host can send a PCP packet requesting a mapping so that any host on the Internet can now initiate communications with the internal host. Similarly, without PCP, a host could send a TCP SYN from a specific port (for example, port 80), thereby creating a mapping nearly identical to a PCP mapping. As with sending a TCP SYN, PCP allows a host to open mappings only for itself, unless the network administrator has taken the extra step to enable the PCP THIRD_PARTY option.

You may wish to have additional restrictions for some networks. PCP is extensible to support authorization, and there is ongoing work to support authentication and authorization within PCP^[8].

PCP is extensible and there are already several proposed extensions to the protocol, including a way to control which IP address pool is assigned to a mapping^[5], bulk port allocation to optimize acquiring a large set of ports^[6], and rapid recovery after NAT failure or network renumbering^[7].

PCP Scenarios

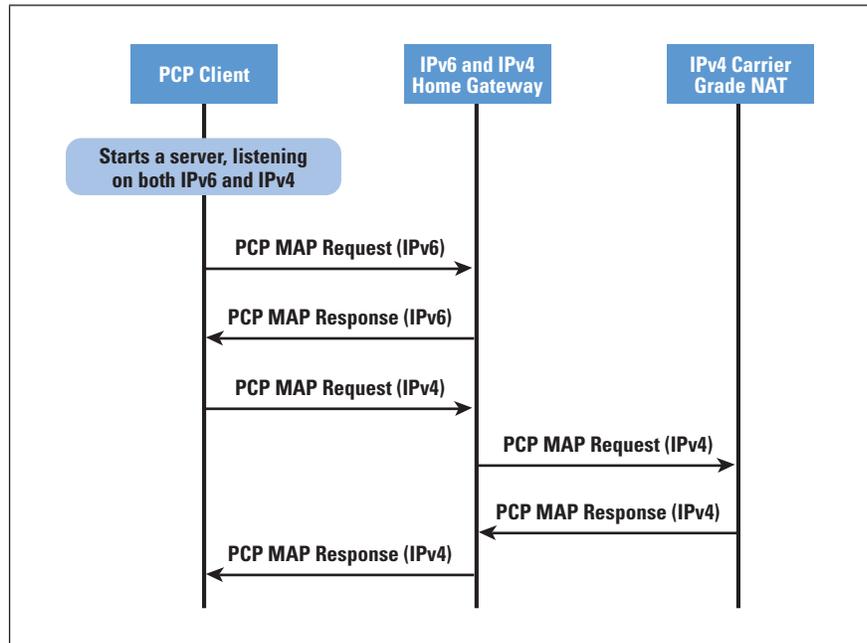
PCP works in all scenarios with IPv4 address sharing (using an IPv4 NAT or using other techniques), an IPv4 or IPv6 firewall, and NATs that translate from IPv6 to IPv4, IPv4 to IPv6, or IPv6 to IPv6. When working with nested NAT, such as a NAT in the home and a NAT operated by the *Internet Service Provider* (ISP), PCP can create the NAT mappings in both devices. When working with IPv6, PCP can create mappings in an IPv6 CPE router. In some networks we expect to see IPv6-only devices that IPv4 clients may need to access. For those devices to work, an IPv6/IPv4 translator (NAT64)^[10, 11] can translate between IPv6 and IPv4. PCP can work with an IPv6/IPv4 translator as well. In other scenarios IPv6/IPv6 translation may be necessary, and although translating IPv6 to IPv6 is far from desirable, PCP can also support IPv6/IPv6 (NPTv6)^[12].

A server, such as a one running on a sensor (for example, thermometer or electric meter), can use PCP to determine its publicly routable IPv4 or IPv6 address and port, and then populate a *Rendezvous* server with that IP address and port. For example, an IPv6-only thermostat might want to be accessible over IPv6 and IPv4, so it can be accessed by both the power company (to push new electricity rate information to the thermostat) and the homeowner (who might have IPv4 access only at work). The thermostat can use PCP to create a TCP mapping in the IPv6 CPE router (necessary because the IPv6 CPE router will, by default, filter unsolicited incoming IPv6 packets) and use PCP to create a TCP mapping in a NAT64 (necessary so the homeowner can access the thermostat). The IPv6 address and its TCP port, and the IPv4 address and its TCP port, can be published to the *Domain Name System* (DNS) (using DNS Server [SRV] records) or published to some other Rendezvous server. Then the power company or the homeowner can use the DNS (or the other Rendezvous server) to communicate directly with the thermostat.

Because PCP can inform the PCP client of address changes, network renumbering can be communicated immediately to hosts—something that cannot be done with most other NAT or firewall control mechanisms. Therefore, devices running on nomadic networks, such as in a connected vehicle, that use PCP will immediately learn when they have connected to a new network. This knowledge can allow them to update information in the DNS or in some other Rendezvous server so they remain accessible from the Internet.

PCP is expected to be implemented in home gateways and Carrier-Grade NATs, which provide value for both IPv6 (to operate a server and learn keepalive timeouts) and IPv4. Figure 1 shows how a dual-stack host would use PCP to operate an IPv6 or IPv4 server.

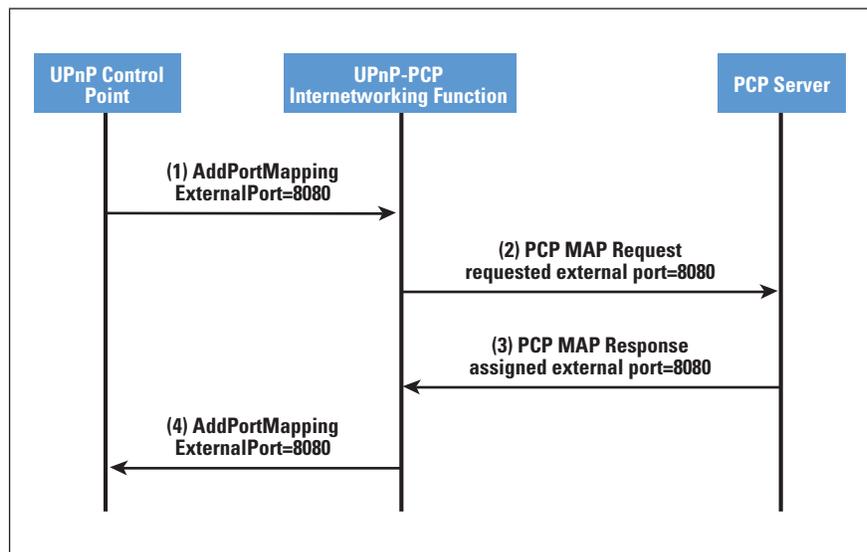
Figure 1: PCP Mapping IPv6 and IPv4



PCP Interworking with UPnP IGD

UPnP IGD Version 1 is widely available on residential-class NAT devices and host operating systems (Windows and OS X). However, because of security concerns it is often disabled by vendors, ISPs, or end users. UPnP IGD itself only works with a single layer of NAT, but it is possible to interwork between UPnP IGD and PCP^[4]. To do this interworking, a home gateway (NAT) processes UPnP IGD messages on its LAN interface and translates those messages to PCP messages on its WAN interface, as depicted in Figure 2.

Figure 2: UPnP-to-PCP Interworking, Showing AddPortMapping Success



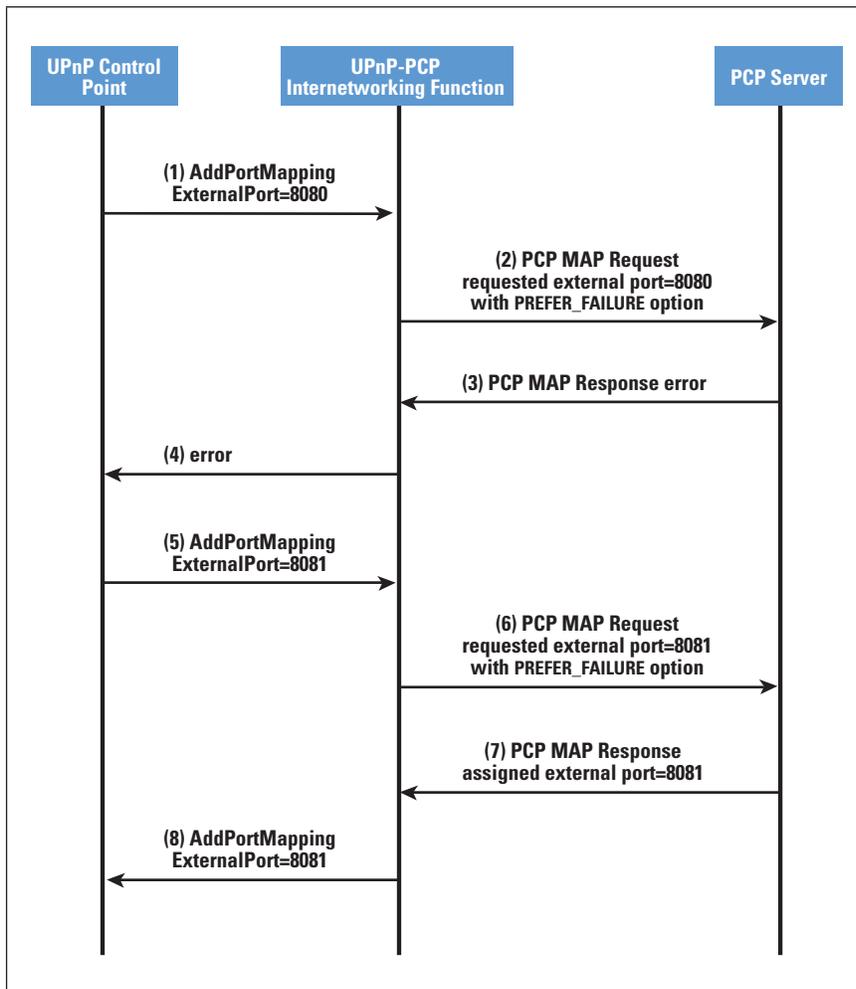
One difficulty with UPnP IGD is its `AddPortMapping` action, which maps a specific port on the home gateway. If that requested port is already mapped to another host, that port cannot be mapped to a new host (because it is already mapped to a different host). This problem exists today with UPnP IGD if two hosts in a home need the same port (for example, TCP port 80) because only one of them can map the port. In a CGN environment, where many subscribers share one IPv4 address, it is almost guaranteed that another subscriber has already mapped a “good” port (for example, 80 for HTTP, 8080 for HTTP, 5001 for Slingbox, 5060 for *Session Initiation Protocol* [SIP], etc.). Today, when a UPnP IGD port mapping is refused, the application may overwrite the first host’s mapping (causing significant problems), “hunt” for an available port, or simply give up and display an error to the user. The “hunting” is often sequential (trying the next-higher port number) but is sometimes random, and is done by the application itself, the operating system UPnP framework, or both.

UPnP IGD Version 2^[2] introduced the *AddAnyPortmapping* action, which avoids the need to “hunt” for an available port and allows the NAT to assign an available port. But UPnP IGD Version 2 is not yet widely available in home gateways, operating systems, or applications. Until IPv6 is ubiquitously available, applications (and users) will need to practice better port agility than has been practiced in the past, because “good” ports will simply not be available when IPv4 addresses are shared.

To ease the interworking with the UPnP IGD `AddPortMapping` action, the base PCP specification includes a `PREFER_FAILURE` option, which avoids creating a mapping if the requested port is unavailable. A message flow of this behavior is shown in Figure 3.

In a *Dual-Stack Lite*^[9] deployment, the home gateway is typically operated without a NAT function. In that configuration, the home gateway is expected to interwork between UPnP IGD (within the home) and PCP (toward the service provider’s CGN). The PCP packets sent by the home gateway will have the source IP address of the home gateway, rather than the IP address of the host that initiated the UPnP IGD action. To accommodate that situation, the home gateway populates the `THIRD_PARTY` option with the IP address of the internal host needing the mapping. The `THIRD_PARTY` option is useful in other scenarios as well, including interworking with other protocols (such as the *NAT Port-Mapping Protocol* [NAT-PMP]^[13]) to PCP, using PCP to create mappings for a device that does not support PCP (for example, an IP-enabled webcam), or using it as the protocol between a web portal operated by the ISP and its CGN.

Figure 3: UPnP-to-PCP Interworking, Showing AddPortMapping Failure



Conclusion

PCP provides functions necessary for IPv6 hosts on home networks; it is a simple, scalable protocol that supports simple firewalling of IPv6 and IPv4 hosts, and to accommodate the transition to IPv6 also supports every conceivable IPv4/IPv6 translation mechanism.

References

- [1] Dan Wing, ed., Stuart Cheshire, Mohamed Boucadair, Reinaldo Penno, and Paul Selkirk, “Port Control Protocol (PCP),” Internet Draft, work in progress, July 2011, [draft-ietf-pcp-base](#)
- [2] “UPnP Gateway committee: IGD:2 improvements over IGD:1,” March 2009, <http://www.upnp.org/resources/documents/UPnPIGD2vsIGD1d10032009.pdf>
- [3] James Woodyatt, ed., “Recommended Simple Security Capabilities in Customer Premises Equipment for Providing Residential IPv6 Internet Service,” RFC 6092, January 2011.

- [4] Mohamed Boucadair, Reinaldo Penno, Dan Wing, and Francis Dupont, “Universal Plug and Play (UPnP) Internet Gateway Device (IGD)-Port Control Protocol (PCP) Interworking Function,” Internet Draft, work in progress, February 2011, **draft-bpw-pcp-upnp-igd-interworking**
- [5] Reinaldo Penno, “PCP Support for Multi-Zone Environments,” Internet Draft, work in progress, June 2011, **draft-penno-pcp-zones**
- [6] Cathy Zhou, Tina Tsou, Xiaohong Deng, Mohamed Boucadair, and Qiong Sun, “Using PCP To Coordinate Between the CGN and Home Gateway Via Port Allocation,” Internet Draft, work in progress, July 2011, **draft-tsou-pcp-natcoord**
- [7] Stuart Cheshire, “PCP Rapid Recovery,” Internet Draft, work in progress, June 2011, **draft-cheshire-pcp-recovery**
- [8] Margaret Wasserman, Sam Hartman, and Dacheng Zhang, “Port Control Protocol (PCP) Authentication Mechanism,” Internet Draft, work in progress, October 2011, **draft-wasserman-pcp-authentication**
- [9] Alain Durand, Ralph Droms, James Woodyatt, and Yiu L. Lee, “Dual-Stack Lite Broadband Deployments Following IPv4 Exhaustion,” RFC 6333, August 2011.
- [10] Congxiao Bao, Christian Huitema, Marcelo Bagnulo, Mohamed Boucadair, and Xing Li, “IPv6 Addressing of IPv4/IPv6 Translators,” RFC 6052, October 2010.
- [11] Marcelo Bagnulo, Philip Matthews, and Iljitsch van Beijnum, “Stateful NAT64: Network Address and Protocol Translation from IPv6 Clients to IPv4 Servers,” RFC 6146, April 2011.
- [12] Margaret Wasserman and Fred Baker, “IPv6-to-IPv6 Network Prefix Translation,” RFC 6296, June 2011.
- [13] Stuart Cheshire, Marc Krochmal, and Kiren Sekar, “NAT Port Mapping Protocol (NAT-PMP),” Internet Draft, (expired), April 2008, **draft-cheshire-nat-pmp-03.txt**

DAN WING is the editor of the Port Control Protocol base specification and co-author of the PCP-UPnP interworking function specification. Dan has co-chaired the IETF’s BEHAVE Working Group since 2006. He is a Distinguished Engineer at Cisco Systems, where he works on IPv6 transition technologies.
E-mail: dwing@cisco.com

Infrastructure Challenges to DNS Scaling

by Bill Manning

This article looks a few steps beyond the *Root Scaling Study* report from 2009.^[1] In 2009, the *Internet Corporation for Assigned Names and Numbers* (ICANN) board commissioned a report to evaluate the effect of scaling the root zone from its current size to an undefined but larger root zone. Attributes considered were *Domain Name System Security Extensions* (DNSSEC), *Internet Protocol Version 6* (IPv6), *Internationalized Domain Names* (IDNs), and a larger number of entries in the zone. The report itself focused on the editorial processes and presentation of the finished root zone to the greater Internet. The report concluded that with prudence and with the addition of some “watch & warn” systems in place, the root zone could accommodate adding IPv6, DNSSEC, and IDNs along with other new *Top-Level Domain* (TLD) entries in a controlled manner. What the report did not consider was the effects of the deployed Internet infrastructure on the ability to get this new information into the rest of the *Domain Name System* (DNS) infrastructures of the Internet. Early experimental evidence^[7, 8] suggests that the current state of infrastructure deployment will create problems for the deployment of these attributes.

Until recently the root zone of the DNS has enjoyed two important stabilizing properties:

- It is relatively small—currently the root zone holds delegation information for 280 generic, country-code, and special-purpose TLDs, and the size of the root zone file is roughly 80,000 bytes.
- It changes slowly—on average, the root zone absorbs less than one change per TLD per year, and the changes tend to be minor.

The root system has therefore evolved in an environment in which information about a small number of familiar TLDs remains stable for long periods of time. However, the type, amount, and volatility of the information that is contained in the root zone are expected to change as a result of the following four recent or pending policy decisions:

- Support for DNSSEC, or “signing the root”
- The addition of IDN TLDs
- Support for the additional larger addresses associated with IPv6
- The addition of new TLDs

These changes are placed in a backdrop of an infrastructure that is fundamentally changing, removing a third attribute of a stable DNS that was the presumption of a common transport protocol with well-defined constraints.

Core Design Principles

The DNS was designed so that queries and responses would have the greatest chance of survival and broadest reachability by using an IPv4 default *User Datagram Protocol* (UDP) packet size of 512 bytes for the initial bootstrapping. Larger packet sizes are supported and the *Transmission Control Protocol* (TCP) was defined and used as an alternate transport protocol—but expected to be infrequently used.

With these core principles intact, the DNS was able to successfully evolve into a highly decentralized dynamic system. The geographic and organizational decentralization of the root system arises from a deliberate design decision in favor of diversity and minimal fate-sharing coordination, which confers substantial stability and robustness benefits on the global Internet.

Simple quantitative extrapolation from a baseline model of the current DNS does not predict realistic future states of the system beyond the very short term, because:

- Each part of the system adapts in different ways to changes in the quantity, type, and update frequency of information, while also responding to changes in the rest of the Internet.
- These adaptations are not—and cannot be—effectively coordinated.
- For some, if not all, of the actors, nonquantifiable considerations dominate their individual adaptation behavior (both strategically, in a planning context, and tactically, in an operations context).

The risks associated with adding DNSSEC and IPv6 addresses to the DNS simultaneously change the basic assumption for DNS Query/Response reachability. Signing DNS data would, by itself, immediately increase the size of any zone by roughly a factor of 4 and increase the size of the response message^[2]. The consequences of the second of these effects could be absorbed by replanning in order to recover lost headroom by adding bandwidth. Adding IPv6 addresses would in addition increase the size of any response. However, simply adding additional bandwidth may be insufficient when there are middleboxes, application layer gateways, or divergent transport options between the query path and the response path.

In these cases more information has to be carried in the packets that are returned in response to a query, meaning that the required amount of network bandwidth needed to support the operations of the server increases. As the DNS messages get bigger, they will no longer fit in single 512-byte packets forwarded by the UDP transport mechanism of the Internet. This situation will lead to clients being forced to resend their queries using UDP “jumbograms” or the TCP transport mechanism—a mechanism that has much more overhead and requires the end nodes to maintain much more state information. It also has much more overhead in terms of “extra packets” sent just to keep things on track. The benefit is, of course, that it can carry much larger pieces of information.

Moving the root system from its default UDP behavior to UDP “jumbograms” or TCP will not only have the undesirable effects mentioned previously, it will also affect the current trend of deploying servers using IP *anycast*^[10]. Anycast works well with single packet transactions (such as UDP), but is much less well suited to handle TCP packet streams. If TCP transactions become more prevalent, the anycast architecture may require changes.

The point of view from the client side is worth mentioning. In certain client configurations, where firewalls are incorrectly configured^[3], the following scenario can occur:

A resolver inside the misconfigured firewall receives a DNS request that it cannot satisfy locally. The query is sent to the root servers, usually over UDP, and a root server responds to this query with a referral, also over UDP. Today, this response fits nicely in 512 bytes. It is also true that for the past 6 years, the *Internet Systems Consortium* (ISC) has been anticipating DNSSEC and has shipped resolver code that, by default, requests DNSSEC data. After the root is signed, the response no longer fits into a 512-byte message. Estimates from the *National Institute of Standards and Technology* (NIST), using standard key lengths, indicate that DNSSEC will push the response to at least 2048 bytes or larger. This larger response will not be able to get past a misconfigured firewall that restricts DNS packets to 512 bytes, not recognizing the more modern extensions to the protocol that allow for bigger packets.

Upon not receiving the answer, the resolver on the inside will then retry the query, setting the buffer size to 512 bytes. The root will resend the response using smaller packets, but because it does not fit in a 512-byte packet, will fragment the response into a series of 512-byte replies, and the root server will set the “fragmented” and “truncated” flags in the packets, indicating to the resolver that the answer was fragmented and truncated, and encouraging the resolver to retry the query once more using TCP transport. The resolver will do so, and the root server will respond using TCP, but the misconfigured firewall also will reject DNS over TCP, because this transport has not been considered a normal or widely used transport for DNS queries.

In this worst case, a node will be unable to get DNS resolution after the root zone is signed, and the DNS traffic will triple, including one round in which TCP state must be maintained between the server and the resolver. There are of course ways around this problem, the most apparent ones being to configure the firewall correctly, or to configure the resolver to not ask for DNSSEC records.

Effect of IPv6 on Priming Queries

The basic DNS protocol specifies that clients, resolvers, and servers be capable of handling message sizes of at least 512 bytes. They may support larger message sizes, but are not required to do so.

The 512-byte “minimal maximum” was the original reason for having only nine root servers. In 1996 Bill Manning, Mark Kosters, and Paul Vixie presented a plan to Jon Postel to change the naming of the root name servers to take advantage of DNS label compression and allow the creation of four more authoritative name servers for the root zone. The outcome was the root name server convention as it stands today.

The use of 13 “letters” left a few unused bytes in the priming response, which were left there to allow for changes—which soon arrived. With the advent of IPv6 addressing for the root servers, it was no longer possible to include both an IPv4 “A” record and an IPv6 “AAAA” record for every root server in the priming response without truncation; AAAA records for only two servers could be included without exceeding the 512-byte limit. Fortunately the root system was able to rely on the practical circumstance that any node asking for IPv6 address information also supported *Extension Mechanisms for DNS (EDNS0)*^[4].

DNSSEC also increases the size of the priming response, particularly because there are now more records in the Resource Record set and those records are larger. In [5] the authors make the following observation: “The resolver MAY choose to use DNSSEC OK^[6], in which case it MUST announce and handle a message size of at least 1220 octets.”

EDNS and MTU Considerations

The changes described will also affect other parts of the Internet, including (for example) end-system applications such as web browsers; intermediary “middleboxes” that perform traffic shaping, firewall, and caching functions; and *Internet Service Providers* (ISPs) that “manage” the DNS services provided to customers.

Although modern DNS server software defaults to using EDNS0, current measurement^[7] collected from several of the RFC 1918^[11] servers suggests that EDNS0 usage has not yet reached generally accepted levels of usefulness. Over the 12-month study, the ratio of EDNS0 queries received at these nodes remained at roughly 65 percent of the total queries received, with about 33 percent being non-EDNS queries. In the “other” camp are queries that set EDNS0 but then restrict packet sizes to 512 bytes. These queries cannot use the larger, negotiable *Maximum Transmission Unit* (MTU) sizes for larger UDP responses and therefore must use TCP to support larger responses. Some evidence suggests that with signed data, there is a pattern of retransmission of queries when responses larger than 512 bytes are generated and blocked. Such retransmissions can take as long as 7 seconds before timing out.

Lack of EDNS0 support in DNS caches suggests that many parts of the Internet will be constrained to using the traditional UDP sizes or will fall back to using TCP. Even where EDNS0 is indicated as being available, there are increased difficulties in knowing or negotiating a consistent *Path Maximum Transmission Unit* (Path MTU)^[8].

The data supports an argument that the expectation of a useful UDP “jumbogram” or enough resources to manage hundreds of thousands or millions of TCP connections is unfounded because of historical expectations on “normal” DNS packet profiles. Clean, clear Internet paths that will allow larger packet sizes are rare, particularly when crossing the Internet. Locally, it is much more likely that larger packet sizes will be found and supported, raising the question for wide-scale deployment of IPv6 or DNSSEC because both attributes require larger packet sizes regardless of transport. If neither larger UDP packets nor TCP will be viable, what other choices are there?

Recent work inside the *Internet Engineering Task Force* (IETF) is exploring the use of the *Hypertext Transfer Protocol* (HTTP) as an alternative transport protocol for DNS messages.^[9] It might be possible to augment the deployed DNS base to understand the addition of a third transport protocol.

The augmentation of the DNS protocol to support multiple transport protocols will require additional logic on the part of the servers to keep track of which transport a query was received on and select that transport when sending back the response. It will also require more complex logic to determine failover selection from one transport to another.

With the efforts going into making the infrastructure of the Internet IPv6-capable, it is possible that the underlying MTU problems may be corrected faster than adoption of a new transport protocol for the DNS. Certainly MTU problems have been considered for many years and for slightly different reasons^[8] principally related to faster signaling rates and changes in the types of data being moved through the Internet. Regardless, this transition will take considerably more time than a simple DNS code refresh. Full support for larger packet sizes in the DNS will require changes in the equipment and code that comprise the baseline Internet infrastructure—and such changes may take decades.

References

- [1] Jaap Akkerhuis, Lyman Chapin, Patrik Fältström, Glenn Kowack, Lars-Johan Liman, and Bill Manning, “Report on the Impact on the DNS Root System of Increasing the Size and Volatility of the Root Zone, Prepared by the Root Scaling Study Team,” Version 1.0, September 2009.
- [2] “DNSSEC and Its Impact on DNS Performance,” 17 August 2009, <http://www.dnsops.gov/dnssec-perform.html>

- [3] Ray Bellis and Lisa Phifer, “Test Report: DNSSEC Impact on Broadband Routers and Firewalls,” SAC035, 16 September 2008,
<http://www.icann.org/en/committees/security/ssac-documents.htm>
- [4] Paul Vixie, “Extension Mechanisms for DNS (EDNS0),” RFC 2671, August 1999.
- [5] Peter Koch and Matt Larson, “Initializing a DNS Resolver with Priming Queries, Internet Draft, expired, July 2008,
<http://tools.ietf.org/id/draft-ietf-dnsop-resolver-priming-01.txt>
- [6] Roy Arends, Rob Austein, Matt Larson, Dan Massey, and Scott Rose, “DNS Security Introduction and Requirements,” RFC 4033, March 2005.
- [7] EDNS Support:
<http://www.ripe.net/data-tools/dns/as112/edns>
- [8] Matt Mathis, “The Case for Raising the Internet MTU,” July 2003, <http://staff.psc.edu/mathis/papers/Cisco200307/index.html>
- [9] Mohan Parthasarathy and Paul Vixie, “Representing DNS Messages Using XML,” Internet Draft, work in progress, September 2011, <http://www.ietf.org/id/draft-mohan-dns-query-xml-00.txt>
- [10] Ted Hardie, “Distributing Authoritative Name Servers via Shared Unicast Addresses,” RFC 3258, April 2002.
- [11] Yakov Rekhter, Robert G Moskowitz, Daniel Karrenberg, Geert Jan de Groot, and Eliot Lear, “Address Allocation for Private Internets,” RFC 1918, February 1996.

BILL MANNING has been in the network field since 1979, most recently with Booz Allen Hamilton. He has been an IETF Working Group chair, RFC author, and an ARIN Trustee, and he has been on numerous ICANN committees. He has worked as part of the teams that run Internet Root name servers, built the first Internet Exchange points, and worked on transitioning from NSFnet to commercial services. Current client work is focused on Internet Policy and Governance, Risk Analysis, and the future of naming systems. E-mail: bmanning@sfc.keio.ac.jp

Networking @ Home

by Geoff Huston, APNIC

One of the more interesting sessions at the *Internet Engineering Task Force* (IETF) meeting in Quebec City in July 2011 was the first meeting of the recently established *Homenet Working Group*^[1]. What is so interesting about networking the home? Well, if you regard challenges as “interesting,” then just about everything is interesting when you look at networking in the home!

It has been a very long time since the state of the art in home Internet involved plugging the serial port of the PC into the dialup modem. The *Asymmetric Digital Subscriber Line* (ADSL) modem, even when combined with some form of Wi-Fi base station, is looking distinctly passé these days. Today, the home network is seeing the intersection of a whole set of interests, including phone service, television service, home security services, energy management, utility service metering, other forms of home device monitoring, and, of course, connecting laptops and mobile devices to the net. The home network is not just a wired *Local-Area Network* (LAN), Wi-Fi home networks are commonplace, and there are also various Bluetooth devices. Maybe sometime soon it will be common for the home network to host some form of *Third-Generation* (3G) femtocell mobile cell phone repeater as well. But these days even that level of network complexity is not enough. Increasingly, the home office is part of the work office, and if numerous residents are at home, then the home network may be an endpoint for several corporate and institutional *Virtual Private Networks* (VPNs)^[2].

Within the home network we want sophisticated security. This security involves not just protecting the network from the neighbors; the security requirements include the ability for individuals to partition off their work-VPN part of the home network from other home users. For resiliency we might want a second network provider, so we might want to add site-based multihoming to the mix. And we need to make all this work for both IPv4 and IPv6.

That set of requirements represents a massive agenda. But to make this situation truly challenging, we cannot expect every home to come with an IT Operational Service Manager to ensure that all the various devices you bring into the home and connect to the network function as required for the particular requirements of the home. Indeed, we cannot expect any home to be so lavishly supported, nor can we afford to support home networking with a bevy of specialized call centers with on-demand support specialists, expert in the panoply of consumer devices that are being sold today.

With today’s home networks, consumers are effectively on their own; and all this equipment better just work straight out of the box. No configuration, no buttons, it just has to work!

Routing @ Home

The evolution of networking at home has progressed from a single computer to a basic LAN, and from there to an Ethernet-bridged network with numerous Wi-Fi and wired LAN segments. All these environments have a single common architecture with a single “boundary” unit that acts as a point of demarcation between the *Internet Service Provider* (ISP) and the home network. This unit is generally called *Customer Premises Equipment* (CPE), and typically encompasses the functions of a modem; an IPv4 *Network Address Translator* (NAT); a *Dynamic Host Configuration Protocol* (DHCP) server for both IPv4 and IPv6; as well as security firewall, bridge, and rudimentary router functions.

But it is unrealistic to assume that home networks will continue to use a centralized model that places all of the management functions of the home network in a single unit. So how should we view home networks? Should home networks be a single bridged LAN, or are we seeing the evolution of home networks into multiple distinct domains with a routing fabric to glue them together? And if that is the case, what routing protocol should be used?

I have noticed in the low end of the CPE market it is not uncommon to see a rudimentary routing function supported by the *Routing Information Protocol* (RIP)^[3]. Thankfully, it is RIP Version 2, so the routing protocol can be configured with variable-length subnet masks, but even so, RIP is a very basic and simple routing protocol. But perhaps in this environment, that might be a positive factor rather than a liability in so far as RIP is simple enough to be auto-configurable. On the other hand, if there is an emergent need for more complex functions, then maybe we need to look a little harder at the available options.

One of these more complex functions is *subnet management*. In IPv6, the CPE will collect an IPv6 address prefix. This process differs from the conventional IPv4 environment where the CPE is typically assigned a single IPv4 address. So the ensuing question is: Is it possible to automate the distribution of IPv6 subnets across the entire home network? What form of management protocol is appropriate for this role?

Of course the situation gets much more complicated if the home network has two (or more) service providers. In the IPv6 environment, this task becomes a challenging one, not only with the distribution of multiple subnets across the home network, but also in the matter of exit path selection. If the home network is exercising due diligence to prevent source address spoofing, it is also necessary for the home routing infrastructure to deliver an outgoing packet to the “right” exit ISP, where the source address of the outgoing packet needs to match the address prefix provided by the corresponding ISP service. In other words, there is a requirement for source address routing in the home.

This challenge was not really addressed by the *Site Multi-Homing by IPv6 Intermediation Working Group* (SHIM6)^[4], despite the best of intentions, and it represents an even greater challenge if the intent is to provide mechanisms that can achieve such routing in an unmanaged home network environment.

I must admit to some concern here. We have managed to keep Internet routing working by using two principles. The first is to try to keep the routing task as simple as possible. Routing propagates a single “best” path to a destination. It does not necessarily do this propagation quickly, nor necessarily does it carry around with it a whole set of alternatives. It does just one job. The second principle is to admit that we have never really succeeded with the first principle of functional simplicity and we have always had expertise at hand to oversee the routing function and apply manual patches as required. The specialized requirements for the home network appear to be breaking both principles. The requirements are certainly not simple, and I see a mix of routing techniques—including various forms of policy-based routing requirements—entering the discussion. Secondly, there is no assurance that if things fail expertise is at hand to mend the failure. Indeed, the more complex the routing environment, the greater the potential for complex forms of failure. As we contemplate ever more complex requirements in the home network, we face a greater risk of encountering failure “by design,” where it is just not possible to design products for this environment that will “just work.”

Names @ Home

What should I call my printer? More to the point, how should I identify my Wi-Fi printer to all those devices at home that want to use it to print? I am sure that I would not like to use a proprietary naming scheme that requires me to add additional name resolution software to every device at home that wants to print something, nor do I want to transcribe IP addresses into everything. I would like my printer to get dynamically assigned IPv4 and IPv6 addresses when the device is plugged in and switched on, and have the name of the printer published via a generic name resolution mechanism, namely the *Domain Name System* (DNS).

But most of the time the rest of the world has no need to know the name of my printer at home, and I am not sure that it is a good move, securitywise, to gratuitously publish information in the public DNS. So what I would like for my printer is some form of “local” or “scoped” DNS, where I can name my printers, my disk servers, and other devices that I have at home in the context of my home and not have this information leak further afield. Is this scoped form of name resolution, split horizon DNS, or split views, possible in the context of the DNS without invoking further elements of configuration management?

Multicast DNS (mDNS) is perhaps one of the strongest candidates for this role. In essence, mDNS replaces the explicit client-server structure of the DNS with a scoped name subdomain of `.local` that is inherently scoped to the associated multicast domain.

This setup allows a client to perform DNS-like name resolution functions on a local network without the need to configure a conventional DNS server environment, and without the need to obtain global delegation of a site name in the global DNS.

An alternative approach is to use a conventional DNS delegation and conventional unicast DNS queries and responses. Clients are able to use DNS *Dynamic Updates*^[5] to provide the local DNS server with their details as they come online. This approach requires either open access from anyone to the nameserver or a security mechanism such as *Transaction SIGnature* (TSIG)^[6]. TSIG generally requires manual configuration, and alternatives are either little used—such as *Transaction KEY* (TKEY)^[7]—or involve further intricacies, such as Microsoft’s *Active Directory*, which uses other user authentication mechanisms to bootstrap the TSIG part using the *Generic Security Service Algorithm for Secret Key Transaction* (GSS-TSIG)^[8]. The DNS server itself can be advertised to all clients via the *Simple Service Discovery Protocol* (SSDP), as part of the larger *Universal Plug and Play* (UPnP) framework.

Sensing and Serving @ Home

Where to go from here? It is certainly the case that electronics has managed to pervade just about every device at home. Electricity meters are morphing into household energy-management systems, and many other household appliances are now controlled by internal processors. But individually configuring each of these devices is a forbidding task. Even adding an interface to allow manual configuration can often be a challenging objective.

The objective here is to define a standard mechanism to allow sensors to sense their local environment when powered up, obtain an IP address, advertise their existence and capabilities to the network, and, as appropriate, rendezvous with the sensor controller or controllers across the home network.

This example is another instance of a more generic class of automating the installation and use of services in “lightly” managed or even unmanaged networks, and it intersects significantly with the objectives encompassed with SSDP and UPnP. The potential volume of such devices places this example more squarely into a class of IPv6-only services, I suspect, which is a significant extension to the existing IPv4-centric UPnP frameworks.

What is needed is a bootstrap protocol that can provide a connecting device with:

- Address configuration
- Routing setup
- Name management and name server discovery
- Discovery of other services and controllers
- Security capabilities

Security @ Home

One of the most significant concerns with home networks lies in the area of security management. Host computers in a home network often want to place a very high level of implicit trust in their immediate network neighbors at the same home. It is not unusual for hosts in a home network to share printers, file servers, data, and even user profiles. Indeed, it is probably commonplace. But beyond this local security domain a host should become paranoid and treat all connection attempts with suspicion. But where does the local trust domain start and stop? What is the “local” security boundary?

This question is difficult to answer in an automated fashion. It is no longer the local LAN, particularly as home networks transition into routed networks. The security boundary is related to the local multicast scope, but this supposition assumes that it is possible to define a multicast scope that encompasses the local trust domain of the home network, and this assumption brings us back to the same question.

Even if you thought you might have a clean answer to the boundary question, you need to remind yourself about telecommuting. With telecommuting, there is a requirement to partition out an entire local network segment from the rest of the home environment and the home security domain and transplant it into the work security domain.

Everything @ Home

Home is certainly the new field of engagement for networked goods and services. However, it is one of the most challenging places to operate in from the perspective of attempting to deliver coherent services in a reliable and secure manner. The components are sourced from various vendors, and constructed incrementally over extended periods of time. It is an environment where older components need to coexist with new devices, and the overall engineering of the environment is at best piecemeal, and perhaps more often not engineered at all. In this environment out-of-the-box interoperability is of paramount importance, and therefore it is an environment where good standards really matter. Perhaps unsurprisingly, given these constraints, networking in the home is one of the environments that appear to raise the most challenges. It is an unforgiving environment where there is no real substitute for simplicity and reliability in a “plug-and-play” world.

The IETF Homenet Working Group has a lot of work to do. The Working Group will have to examine the diverse set of approaches in use today, add IPv6 functions, and produce a coherent set of outcomes in the form of standards that support robust, capable home networks that work in an unmanaged environment.

Ahhh home! There really is no place quite like it!

References

- [1] Homenet Working Group: <http://www.ietf.org/dyn/wg/charter/homenet-charter>
- [2] Paul Ferguson and Geoff Huston, “What Is a VPN?” (Part One and Part 2), *The Internet Protocol Journal*, Volume 1, No. 1 and No. 2, June and September 1998.
- [3] Gary Malkin, “RIP Version 2,” RFC 2453, November 1998.
- [4] Shim6 Working Group (concluded):
<http://wiki.tools.ietf.org/wg/shim6/charters>
- [5] Paul Vixie, ed., Yakov Rekhter, Susan Thomson, and Jim Bound, “Dynamic Updates in the Domain Name System (DNS UPDATE),” RFC 2136, December 1997.
- [6] Paul Vixie, Olafur Gudmundsson, Donald E. Eastlake 3rd, and Brian Wellington, “Secret Key Transaction Authentication for DNS (TSIG),” RFC 2845, May 2000.
- [7] Donald E. Eastlake 3rd, “Secret Key Establishment for DNS (TKEY RR),” RFC 2930, September 2000.
- [8] Stuart Kwan, Praerit Garg, James Gilroy, Levon Esibov, Randy Hall, and Jeff Westhead, “Generic Security Service Algorithm for Secret Key Transaction Authentication for DNS (GSS-TSIG),” RFC 3645, October 2003.

GEOFF HUSTON, B.Sc., M.Sc., is the Chief Scientist at APNIC, the Regional Internet Registry serving the Asia Pacific region. He 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 numerous Internet-related books, and was a member of the Internet Architecture Board from 1999 until 2005; he served on the Board of Trustees of the Internet Society from 1992 until 2001.
E-mail: gih@apnic.net

IETF Tools—Making It Easier to Make the Internet Work Better

by Robert Sparks

Many activities are associated with defining and refining an *Internet Engineering Task Force* (IETF) protocol, and all of them are detail-oriented. As IETF Working Groups are formed, mailing list discussions proceed, documents are written and reviewed, and interoperability is evaluated, participants encounter tasks that can be significantly simplified with the help of software tools. Fortunately, those participants frequently are also skilled software developers, and they create and share these tools as the need arises. A new paradigm has evolved recently: When a pressing need for a tool is identified—particularly one that has a large scope—the *IETF Administrative Oversight Committee* (IAOC) accelerates the creation of the tool by working with the community to gather requirements and financing the development of a solution. Comprehensive lists of available tools are maintained at [1] and at [2]. This article introduces a few important tools and discusses how you can help improve them or develop new ones.

Document Tools

The *Extensible Markup Language to Request For Comments* (XML2RFC)^[0] tool was developed to assist with Internet-Draft composition. Marshall Rose created and maintained the initial versions, capturing its input language and operation instructions in RFC 2629^[18]. This tool simplifies draft creation and maintenance by automatically producing documents that satisfy the RFC Editor's layout requirements, and assists in including the appropriate boilerplate as defined by the *IETF Trust*. It also simplifies the task of the *RFC Production Center*^[19, 20]. Starting with XML input rather than a draft in text form reduces the work required to create the RFC. The IAOC is currently funding a reimplementations of XML2RFC to reflect many years of user feedback, simplify maintenance—particularly of boilerplate handling—and make it easier for volunteers to contribute improvements. This reimplementations is currently available at [3]. Tony Hansen has been very active in gathering the requirements for and evaluating the reimplemented version. Julian Reschke also maintains *Extensible Stylesheet Language Transformations* (XSLT) code at [4] that translates RFC 2629-based input into several output formats.

After a new draft is prepared, Henrik Levkowitz' *Internet-Draft Nit Checker* (idnits) tool at [5] can scan it for any problems with the RFC Editor's checklist and guidelines and for other problems that drafts frequently encounter later in review. There are also tools for verifying sections of the document containing formal languages such as *Augmented Backus-Naur Form* (ABNF) or XML.

When an editor is satisfied that the document is ready to place in the repository, the automated *ID Submission tool*^[6] assists with an easy upload. At any point two versions of a draft can be compared with *rfdiff*^[7], a flexible comparison program created by Henrik Levkowitz.

As a draft progresses, its history and current status can be tracked using the *Internet-Drafts Tracker* (ID Tracker) tool^[8]. This tool provides powerful search capabilities into the entire Internet-Draft repository, and a comprehensive view into the lifecycle of each Internet-Draft. With its roots in a tool to help the *Internet Engineering Steering Group* (IESG) keep track of drafts in IESG evaluation, the ID Tracker has evolved into a portal touching almost all aspects of IETF work. Each step of that evolution has improved efficiency and transparency, and has simplified access to the history of the development of each document.

Recent additions to the tracker allow for an easier capture of the details of Working Group processing. *Work in progress* will provide more visibility into the Working Group chartering and rechartering processes. The tracker is also used by other document streams. Many of the enhancements to the tracker are informed by the views into documents and Working Groups maintained by Henrik Levkowitz at [2]. The tracker continues to evolve through both IAOC-funded development efforts and volunteer contributions. An extension in progress will add visibility into the RFC Editor and *Internet Assigned Numbers Authority* (IANA) actions. When this extension is done the entire lifecycle of a Draft, from -00 submission to RFC publication, can be viewed in a single place.

Working Group and Meeting Tools

At each IETF meeting, a participant can build a custom view of the agenda using the tools at the datatracker and the tools sites. For example, [9] renders an interactive JavaScript-based calendar contributed by Adam Roach showing the *Real-Time Applications and Infrastructure* (RAI) meetings at IETF82. The pages at [10] provide a quick reference to the jabber rooms and audio streams of each Working Group meeting. The meeting materials tool facilitates uploading of agendas, slides, and minutes, which become available immediately through the agenda views.

Each Working Group has a Subversion Repository and an integrated instance of Trac^[21] at its disposal. The Subversion Repository can be used to maintain Working Group draft source, versioned instances of test documents, and even implementation code. IETF-specific customizations of the Trac system are described at [11]. Many Working Groups are already taking advantage of what the wiki Trac provides, and are using its ticketing feature to effectively track major Working Group document problems.

Notable examples are the problem tracking integrated into the *Hypertext Transfer Protocol Bis* (HTTPBIS) document status page at [12], and the summary of DISPATCH activity at [13]. The Trac wiki capability is also used by the Working Group Chairs at [14] and the IESG at [15].

IETF News

Keeping up with all of the activity across the IETF can be a challenge. One of the better tools for seeing what is happening is *The Daily Dose of the IETF*, created by Pasi Eronen, available at [16].

Again, this article is an introduction to just a few important tools. Comprehensive lists of available tools are maintained at [1] and [2].

Many of these tools were created because a person who needed them coded an initial version and contributed it to the community. Volunteers (and when needed, IAOC-funded efforts) then improve these tools over time. For several years, a group of volunteers have been meeting the Saturday before each IETF meeting for a day-long *Code Sprint*. If the existing tools need a minor tweak to make things work much better for you, or if you have an idea for a new tool you would like to start, please consider participating at the next Code Sprint. Between sprints, you can still help with the code. Refer to the sprint pages for an upcoming or recent sprint such as [17] and for information about getting started.

Whether or not you can contribute to the code, please discuss your ideas on the `tools-discuss@ietf.org` mailing list.

Several tool contributors have already been mentioned. Henrik Levkowitz deserves to be mentioned again. His herculean efforts maintaining `tools.ietf.org` and creating many of the tools there are of great benefit to the community.

References

- [0] Marshall T. Rose and Carl Malamud, "Writing Internet Drafts and RFCs Using XML," *The Internet Protocol Journal*, Volume 10, No. 1, March 2007.
- [1] <http://www.ietf.org/tools>
- [2] <http://tools.ietf.org/>
- [3] <http://xml.resource.org/>
- [4] <http://greenbytes.de/tech/webdav/rfc2629xslt/rfc2629xslt.html>
- [5] <http://tools.ietf.org/tools/idnits/>
- [6] <https://datatracker.ietf.org/submit/>
- [7] <http://www.ietf.org/tools/rfcdiff/>

- [8] <http://datatracker.ietf.org/>
- [9] <https://datatracker.ietf.org/meeting/82/agenda.html#RAI>
- [10] <http://tools.ietf.org/agenda/82/>
- [11] <http://trac.tools.ietf.org/misc/venue/wiki/IetfSpecificFeatures>
- [12] <http://tools.ietf.org/wg/httpbis/>
- [13] <http://trac.tools.ietf.org/wg/dispatch/trac/wiki>
- [14] <http://wiki.tools.ietf.org/group/wgchairs/>
- [15] <http://trac.tools.ietf.org/group/iesg/trac/wiki>
- [16] <http://tools.ietf.org/dailydose/>
- [17] <http://trac.tools.ietf.org/tools/ietfdb/wiki/IETF82Sprint>
- [18] Marshall T. Rose, “Writing I-Ds and RFCs using XML,” RFC 2629, June 1999.
- [19] Leslie Daigle, “RFC Editor in Transition: Past, Present, and Future,” *The Internet Protocol Journal*, Volume 13, No. 1, March 2010.
- [20] RFC Editor, “40 Years of RFCs,” RFC 5540, April 2009.
- [21] <http://trac.edgewall.org/about>

ROBERT SPARKS is an Area Director for the Real-Time Applications and Infrastructure Area (RAI) in the IETF. He previously chaired the IETF’s SIMPLE Working Group, which defines extensions to SIP for Presence and Instant Messaging, and the GEORPIV Working Group, which provides tools for applications to carry geographic location information and privacy rules to affect its use. Robert is a co-editor of the core SIP standard (RFC 3261), and several important SIP updates and extensions. He coordinates the premier real-time communications interoperability event, the SIPit. Robert is a Principal Software Engineer at Tekelec, and has held management and research positions at Estacado Systems, Xten (now Counterpath), dynamicsoft, Lucent, MCI Worldcom, and Texas A&M University. Robert holds a Master’s degree in Mathematics and a Bachelor’s degree in Computer Science from Texas A&M University. E-mail: rjsparks@nostrum.com

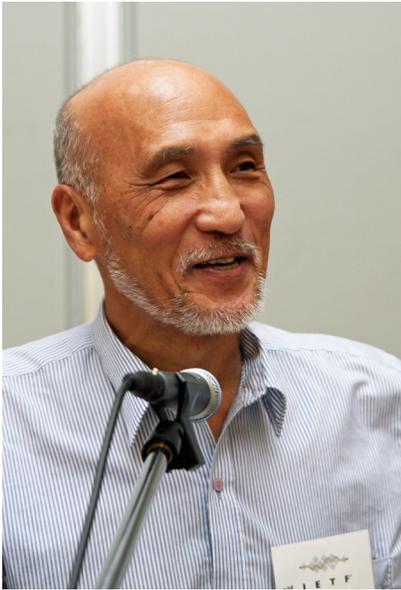


Photo: Peter L othberg

Professor Kilnam Chon Receives 2011 Postel Service Award

The *Internet Society* (ISOC) recently announced that its prestigious *Jonathan B. Postel Service Award* was presented to leading technologist Professor Kilnam Chon for his significant contributions in the development and advancement of the Internet in Asia.

Professor Chon contributed to the Internet’s growth in Asia through his extensive work in advancing Internet initiatives, research, and development. In addition, his pioneering work inspired many others to promote the Internet’s further growth in the region. The international award committee, comprised of former Jonathan B. Postel award winners, noted that Professor Chon was active in connecting Asia, and that his efforts continue today in the advancement of the Internet in other regions.

The Postel Award was established by the Internet Society to honour individuals or organisations that, like Jon Postel, have made outstanding contributions in service to the data communications community.

Lynn St. Amour, President and CEO of ISOC, commented, “I met Professor Chon nearly fifteen years ago. He has long been a pioneer in the advancement of the Internet, striving to ensure its robust development. Beyond the amazing breadth of Professor Chon’s work, perhaps his most remarkable achievement is his ability to inspire others. As a result of his work and the efforts of those he has motivated, Kilnam Chon has helped to ensure the global Internet is truly for everyone.”

ISOC presented the award, including a US\$20,000 honorarium and a crystal engraved globe, during the 82nd meeting of the *Internet Engineering Task Force* (IETF) in Taipei, November 13–18, 2011.

The Internet Society is the world’s trusted independent source of leadership for Internet policy, technology standards and future development. Based on its principled vision and substantial technological foundation, ISOC works with its members and Chapters around the world to promote the continued evolution and growth of the open Internet through dialog among companies, governments, and other organizations around the world. For more information about the Postel Service Award see: <http://www.isoc.org/postel/>

Alexandre Cassen and R emi Despr es Receives 2011 Itojun Service Award

The third *Itojun Service Award* was presented to Alexandre Cassen and R emi Despr es at the *Internet Engineering Task Force* (IETF) meeting held in Taipei, Taiwan in November 2011. The awardees were recognized for their design and implementation of “6rd,” an IETF protocol that aims to speed the transition to global deployment of IPv6, which is critical to ensuring the continued growth and evolution of the Internet.

The 6rd protocol has been implemented by several *Internet Service Providers* (ISPs) around the world, including *Free Telecom*—the second largest ISP in France—as part of their efforts to deploy IPv6.

First awarded in 2009, the Itojun Service Award honors the memory of Dr. Jun-ichiro “Itojun” Hagino, who passed away in 2007 at the age of 37. The award, established by the friends of Itojun and administered by the *Internet Society* (ISOC), recognizes and commemorates the extraordinary dedication exercised by Itojun over the course of IPv6 development.

“Alexandre and Rémi’s efforts have helped to quickly bring a real IPv6 experience to hundreds of thousands of Internet users, demonstrating that IPv6 deployment can be effectively implemented on a large scale by commercial network providers,” said Jun Murai of the Itojun Service Award committee and founder of the WIDE Project. “On behalf of the Itojun Service Award committee, I am extremely pleased to present this award to Alexandre and Rémi for the significant work they have done to advance IPv6 development and deployment.”

The Itojun Service Award is focused on pragmatic contributions to developing and deploying IPv6 in the spirit of serving the Internet. The award, presented annually, includes a presentation crystal, a US\$3,000 honorarium and a travel grant.

Alexandre Cassen said, “It is truly an honor to have been selected to receive the Itojun Service Award. As a software developer myself, It is particularly touching to receive an award created in the memory of a coding legend such as Itojun. I would also like to thank the entire team at Free Telecom who, in 2007, implemented and deployed 6rd, allowing any subscriber who asked for IPv6 to have it with a single click. As I write this, Free Telecom has more than 1,500,000 subscribers using IPv6 every day, and all new subscribers have IPv6 enabled by default. IPv6 is happening Itojun!”

Rémi Després said, “The Itojun Award is the best possible recognition that long efforts to make IPv6 deployment practicable have been useful to the Internet community. Latecomer in IPv6 standardization, I was about to send my first email to Itojun on a technical issue when I heard of his death. I was even sadder since we undoubtedly would have otherwise enjoyed sharing our ideas and our enthusiasm. Sharing the honor of this award with Alexandre Cassen perfectly illustrates the great progress possible when a dynamic network operator with a pioneer spirit and talented engineers adopts an innovative and simple design. Making IPv6 operational on a large scale in only five weeks will be remembered as a milestone of both of our professional lives.”

More information on the Itojun Service Award is available at:
<http://www.isoc.org/itojun>

Internet Society Joins Opposition to Stop Online Piracy Act

The Internet Society Board of Trustees has expressed concern with a number of U.S. legislative proposals that would mandate *Domain Name System* (DNS) blocking and filtering by *Internet Service Providers* (ISPs) to protect the interests of copyright holders. While the Internet Society agrees that combating illicit online activity is an important public policy objective, these critical issues must be addressed in ways that do not undermine the viability of the Internet as a platform for innovation across all industries by compromising its global architecture. The Internet Society Board of Trustees does not believe that the *Protect-IP Act* (PIPA) and *Stop Online Piracy Act* (SOPA) are consistent with these basic principles.

Specifically, the Internet Society is concerned with provisions in both bills regarding DNS filtering. DNS filtering is often proposed as a way to block illegal content consumption by end users. Yet policies to mandate DNS filtering will be ineffective for that purpose and will interfere with cross-border data flows and services undermining innovation and social development across the globe.

Filtering DNS or blocking domain names does not remove the illegal content—it simply makes the content harder to find. Those who are determined to download filtered content can easily use a number of widely available, legitimately-purposed tools to circumvent DNS filtering regimes. As a result, DNS filtering encourages the creation of alternative, non-standard DNS systems.

From a security perspective, DNS filtering is incompatible with an important security technology called *Domain Name System Security Extensions* (DNSSEC). In fact, DNSSEC would be weakened by these proposals. This means that the DNS filtering proposals in SOPA and PIPA could ultimately reduce global Internet security, introduce new vulnerabilities, and put individual users at risk.

Most worrisome, DNS filtering and blocking raises human rights and freedom of expression concerns, and often curtails international principles of rule of law and due process. Some countries have used DNS filtering and blocking as a way to restrict access to the global Internet and to curb free expression.

The United States has been a strong proponent of online Internet freedoms and therefore has an important responsibility to balance local responsibilities and global impact, especially with respect to Internet policy. Given this commitment to global Internet freedom, it would be harmful to the global Internet if the United States were to implement such an approach.

“The Internet Society Board of Trustees is deeply concerned about the ramifications of the PIPA and SOPA bills on the overall stability and interoperability of the Internet,” said Raul Echeberria, Chairman of the Internet Society Board of Trustees.

“The Board recognizes that there can be misuses of the Internet; however, these are greatly outweighed by the positive uses and benefits of the Internet. We believe the negative impact of using solutions such as DNS blocking and filtering to address these misuses, far outweighs any short-term legal or business benefits.”

“The Internet Society believes that sustained, global collaboration amongst all parties is needed to find ways that protect the global architecture of the Internet while combating illicit online activities,” said Internet Society President and CEO Lynn St. Amour. “Mandating DNS blocking and filtering is simply not a viable option for the future of the Internet. We must all work together to support the principles of innovation and freedom of expression upon which the Internet was founded.”

For more details on DNS Filtering, visit:

<http://www.isoc.org/internet/issues/dns.shtml>

See also:

<https://www.eff.org/deeplinks/2011/12/internet-inventors-warn-against-sopa-and-pipa>

APNIC and JPRS Collaborate to Translate DNSSEC Technology Experiment Report

The *Asia Pacific Network Information Centre* (APNIC) has collaborated with *Japan Registry Services* (JPRS) to translate from Japanese into English the documents “DNSSEC Technology Experiment Report – Verification of Functionality and Performance” and “DNSSEC Technology Experiment Report – Operational Design.”

These documents contain the latest information on *Domain Name System Security Extensions* (DNSSEC) implementation, and provides information to those interested in implementing it. These reports are designed to introduce case studies to share knowledge and results gained through experiments conducted in 2010 that JPRS carried out in cooperation with Japanese ISPs, equipment vendors, and hosting providers.

APNIC would like to thank JPRS’s great initiative and all those involved in the process for making such an important contribution to DNSSEC awareness. APNIC also appreciates JPRS for making the documents available in English for wider distribution. The reports are available for download from:

<http://jprs.jp/dnssec/doc/DNSSEC-testbed-report-fpv1.0-E.pdf>

and

<http://jprs.jp/dnssec/doc/DNSSEC-testbed-report-odv1.0-E.pdf>

RFC Series Editor Appointment

The *Internet Architecture Board* (IAB) is pleased to announce the appointment of Heather Flanagan as the *Request For Comments Series Editor* (RSE). Ms. Flanagan will assume the responsibilities from the Acting RSE, Olaf Kolkman, and begin her tenure on January 1, 2012. The contract negotiated by the *IETF Administrative Oversight Committee* (IAOC) includes an initial term of two years and a presumptive renewal of two years.

Ms. Flanagan was selected by the *RFC Series Oversight Committee* (RSOC) based upon her experience, education, skills and energy she will bring to the position.

Ms. Flanagan is currently the Project Coordinator for the *COmanage* project, an effort funded by a grant from the *National Science Foundation* (NSF) and Internet2 to create a collaboration management platform, prior to that she was Director of Systems Administration, IT Services at Stanford University in Palo Alto, California. Her technical background is complemented by a Masters of Science of Library Science from the University of North Carolina, Chapel Hill that will prove invaluable in the accessing and indexing of RFCs.

Ms. Flanagan brings a high degree of energy and enthusiasm to the position. Her interpersonal skills as a facilitator and good listener will enable her to work well with the capable staff at the RFC Production Center and with the community in reaching consensus on a variety of issues facing the RFC Series.

The RSOC selection followed a lengthy process that included announcing the position inside and outside the community, several rounds of interviews, reference checks, and face-to-face interviews in Taipei at IETF 82. More than thirty-five applications were received, two-thirds of which were from outside the community.

We express our congratulations to Ms. Flanagan. We also want to extend our thanks to Ray Pelletier and the RSOC chaired by Fred Baker for their role in bringing the RSE selection process to a successful conclusion; to Olaf Kolkman for his service to the community as Acting RSE; to Joel Halpern for his ongoing work as editor of the “RFC Editor Model v2” document; and to the RFC Production Center for its customary diligence in the editing and publishing of RFCs this year, likely the second most productive in RFC publication history.

We look forward to working with the new RSE; we wish her well; and know that the community will work with Heather for the betterment of the RFC Series.

—For the IAB
Bernard Aboba, IAB Chair

2011 Global IPv6 Survey Results

On October 20, 2011 the *Number Resource Organization* (NRO) announced the publication of the “Global IPv6 Deployment Monitoring Survey 2011 Results,” initially previewed at the *Internet Governance Forum* (IGF) in Nairobi, Kenya, in September.

The findings from the survey drew on data supplied by around 1,600 international respondents, over 350 of which were from the *American Registry for Internet Numbers* (ARIN) region. On behalf of ARIN and GNKS Consulting, we would like to thank all who participated in the survey. Your feedback is crucial to expanding the understanding of where this community is moving, and what can be done to ensure readiness for the widespread adoption of IPv6. We hope you will take this opportunity to review the results at: http://www.nro.net/wp-content/uploads/ipv6_deployment_survey.pdf

The Public Switched Telephone Network in Transition

The United States *Federal Communications Commission* (FCC) recently held two workshops to examine the transition from the *Public Switched Telephone Network* (PSTN) to new technologies. Circuit-switched wireline voice technology has created a high standard for reliability, accessibility, and ubiquity. Consumers will continue to expect and demand these qualities, even as they shift from PSTN services to services provided over different networks. The transition away from the PSTN is already occurring, and is likely to accelerate. Through these workshops, the Commission will seek input on the technical, economic, and policy issues that must be addressed to minimize disruption during this transition, and to protect consumers, public safety, competition, and other important interests. For more information, visit: <http://www.fcc.gov/events/public-switched-telephone-network-transition-0>

Upcoming Events

The *North American Network Operators’ Group* (NANOG) will meet in San Diego, California, February 5–8, 2012. For more information see: <http://nanog.org>

The *Asia Pacific Regional Internet Conference on Operational Technologies* (APRICOT) will meet in New Delhi, India, February 21–March 2, 2012. For more information see: <http://www.apricot2012.net/>

The *Internet Engineering Task Force* (IETF) will meet in Paris, France, March 25–30, 2012. For more information see: <http://www.ietf.org/meeting/>

The *Internet Corporation for Assigned Names and Numbers* (ICANN) will meet in San Jose, Costa Rica, March 11–16, 2012 and in Prague, Czech Republic, June 24–29, 2012. For more information, see: <http://icann.org/>

Call for Papers

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

Topics include, but are not limited to:

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

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

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

This publication is distributed on an “as-is” basis, without warranty of any kind either express or implied, including but not limited to the implied warranties of merchantability, fitness for a particular purpose, or non-infringement. This publication could contain technical inaccuracies or typographical errors. Later issues may modify or update information provided in this issue. Neither the publisher nor any contributor shall have any liability to any person for any loss or damage caused directly or indirectly by the information contained herein.



The Internet Protocol Journal, Cisco Systems
170 West Tasman Drive
San Jose, CA 95134-1706
USA

ADDRESS SERVICE REQUESTED

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

The Internet Protocol Journal

Ole J. Jacobsen, Editor and Publisher

Editorial Advisory Board

Dr. Vint Cerf, VP and Chief Internet Evangelist
Google Inc, USA

Dr. Jon Crowcroft, Marconi Professor of Communications Systems
University of Cambridge, England

David Farber
Distinguished Career Professor of Computer Science and Public Policy
Carnegie Mellon University, USA

Peter Löthberg, Network Architect
Stupi AB, Sweden

Dr. Jun Murai, General Chair Person, WIDE Project
Vice-President, Keio University
Professor, Faculty of Environmental Information
Keio University, Japan

Dr. Deepinder Sidhu, Professor, Computer Science &
Electrical Engineering, University of Maryland, Baltimore County
Director, Maryland Center for Telecommunications Research, USA

Pindar Wong, Chairman and President
Verifi Limited, Hong Kong

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

Copyright © 2011 Cisco Systems, Inc. All rights reserved. Cisco, the Cisco logo, and Cisco Systems are trademarks or registered trademarks of Cisco Systems, Inc. and/or its affiliates in the United States and certain other countries. All other trademarks mentioned in this document or Website are the property of their respective owners.

Printed in the USA on recycled paper.

