IETF Journal

Volume 4, Issue 2
October 2008

*A report from IETF 72, July 2008, Dublin, Ireland. Published by the Internet Society in cooperation with the Internet Engineering Task Force**

Inside this issue

Focus on IPv6 at IETF 72

From the Editor's Desk, by Mirjam Kühne

Set against the beautiful backdrop of a golf resort near Dublin, IETF 72 offered an opportunity to revisit many of the same themes discussed at prior meetings.

IPv6 was, once again, a hot topic, most notably during the Wednesday plenary, for which the Internet Architecture Board organized a panel to discuss IPv6. Panelists consisted of a number of IPv6 operators and experts, each of whom described their experiences deploying IPv6 within their networks and organizations. The panel discussion, moderated by Gregory Lebovitz, is described in detail on page 17.

Citywest Hotel Convention Centre near Dublin, site of IETF 72.

In this issue, Shane Kerr takes a look at the IETF response to the DNS vulnerability that was discovered by Dan Kaminsky and that is often referred to as the Kaminsky Attack (see page 31). While the discussions within the IETF took place primarily within working groups, Shane offers a good look at the history of DNS security and a brief review of recent DNS security work as well as a compilation of IETF responses.

Stepping outside the usual technical discussions, the IETF Journal sat down with IETF lawyer Jorge Contreras, who discussed his work with the IETF and, in particular, the latest developments in the field of intellectual property rights. The interview with Jorge appears on page 11.

Once again, the IETF meeting played host to a number of engineers from various parts of the developing world. Their participation at IETF 72 was made possible through generous support by the Internet Society as part of its Fellowship to the IETF programme. This was also the first time that some of the earlier fellows returned to attend their second IETF meeting. See more about the fellows and their experiences on page 14.

Thank you to the contributors to this issue. We wish everyone fun reading. And as always, we welcome both your comments and your contributions for future issues.

IETF Meetings Related to Peer-to-Peer and Bandwidth Management

By Leslie Daigle

As bandwidth management has become an important topic, the IETF has held a handful of meetings to discuss the issues it involves and possible IETF work items that might address them. One of those meetings was a one-day workshop organized by the Real-time Applications and Infrastructure (RAI) area and held at Massachusetts Institute of Technology at the end of May 2008. Following that were two BoF (birds-of-a-feather) sessions, held at IETF 72 in Dublin.

Continued on page 4

Message from the IETF Chair

By Russ Housley

Russ Housley, IETF Chair

Held at City West in Dublin, in July 2008, IETF 72 was by all measures a highly successful meeting. With 1,183 people from 48 different countries in attendance, the week was filled with the usual mix of working group (WG) meetings, BoF (birds-of-a-feather) sessions, research group (RG) meetings, and, as always, many side meetings. Our host, Alcatel-Lucent, certainly made everyone feel welcome, and we had a wonderful time at the Guinness Storehouse on Tuesday evening. The network connecting City West to the rest of the Internet was provided by eircom, and the local network was provided by Alcatel-Lucent, with considerable support from volunteers.

Since IETF 71, 5 new WGs were chartered and 11 WGs were closed, leading to approximately 115 chartered WGs in total. Between the meetings, the WGs and their individual contributors produced 475 new Internet-Drafts and generated 1,071 updated Internet-Drafts. The Internet Engineering Steering Group approved 134 Internet-Drafts for publication as RFCs and the RFC Editor published 88 new RFCs.

One of the hot topics during IETF 72 was the coexistence of IPv4 and IPv6. The discussions about requirements for NAT-PT (network address translation–protocol translation) in the Internet Area were especially lively. To aid in the discussion, an IPv6-only network was available throughout the week so people could see what the Internet would be like without IPv4. While the topic was not resolved, an interim meeting is being organized for early October in Montreal to continue the discussions. The meeting will cover topics that affect work ongoing in a number of WGs, including SOFTWIRE, V6OPS, and BEHAVE.

At IETF 73, we expect to conduct a scheduling experiment that we hope will lead to the creation of additional meeting time for WGs. Instead of ending at 11:30 on Friday, we will end at 15:15. The after-lunch meeting slots will mean more than 16 additional hours of session time for WGs. Following the meeting, the IETF will evaluate the results of the experiment and determine whether a longer meeting is something we want to continue in the future.

I look forward to IETF 73 in Minneapolis, which is scheduled for 16–21 November 2008, and will be hosted by Google, and to IETF 74 in San Francisco on 22–27 March 2009, which will be hosted by Juniper Networks. Scheduling information for future IETF meetings may always be found at http://www.ietf.org/meetings/meetings.html. I look forward to seeing you at the meetings.

New BoF Meetings

Descriptions and agendas for all BoF meetings can be found at http://www.ietf.org/meetings/past.meetings.html.

Applications Area
morg: Message Organization

Internet Area
explisp: Experimentation in LISP

Routing Area
alto: Application-Layer Traffic Optimization

Transport Area
ippm: IP Performance Metrics, Next Steps
tana: Techniques for Advanced Networking Applications


Olaf Kolkman, IAB Chair

IETF 72 Facts and Figures

Registered attendees 1,183
Countries 48
New WGs 5
Closed WGs 11
WGs Chartered 115
New Internet-Drafts 475
Updated Internet-Drafts 1,071
IETF Last Calls 105
Approvals 134

(March–June 2008)

97 RFCs published of which

- 44 standards tracks
- 4 BCP

107 Internet–Drafts submitted for publication

- 79 submitted by the IETF

IANA Actions (March–June 2008)

Processed 1,422 IETF-related requests of which:

- 775 Private Enterprise Numbers
- 66 Port Numbers
- 67 TRIP ITAD Numbers
- 11 media-type requests

Words from the IAB Chair

By Olaf Kolkman

"To the universal deployment of IPv6" is the toast to which some of our colleagues have raised their glasses for nearly 10 years. That toast can be heard during unofficial events and gatherings at IETF meetings that have taken place since IETF 43, when, after an IPng working group session, some folks retreated to empty a few bottles of Scotch. The relevance of this factoid is linked to the technical plenary at IETF 72, during which the Internet Architecture Board (IAB) expressed interest in IPv6 deployment issues.

In the context of the emerging completion of the IANA IPv4 registry, the IAB asked itself a number of questions such as, What are the actual deployment barriers in various environments ranging from Internet service providers (ISPs) to application service providers, to business enterprises, to end users? What are the approaches toward IPv4 completion contingency planning? What are the success factors inherent in the actual deployment of IPv6? and, What can the IAB do to hasten IPv6's deployment?

With those questions in mind, the IAB organized a plenary discussion to which we invited a number of folks who have key roles in the deployment of IPv6 in their organizations. Our intention was not only to try to address the questions asked here but also to inspire participants to go back home and assess how they could play their part in an IPv6 rollout. A report by the moderator, Gregory Lebovitz, appears on page 17.

On a personal note, it was difficult to find engineers who could shine a light on the deployment issues within business enterprises. I wonder how much that has to do with the fact that IPv6 has not yet made it onto those corporations' agendas. To a certain extent, I understand how deployment of a new IP address family might not make it onto the agenda. It shouldn't have to; IP infrastructure should just work. However, I find it hard to believe that at the chief technology officer (CTO) level, no conscious decision has been made about how to move forward with IPv4. It seems clear to me that as soon as IANA's IPv4 registry is completed, the landscape of IPv4 allocation and assignment is going to change. If I were a CTO, it would make sense to me to have already made an analysis of how to move forward within that landscape. Could it be that CTOs are not tuned in to the issue? An entirely different and somewhat plausible explanation for the difficulty in finding engineers with experience in enterprise-scale IPv6 deployment is that those engineers are not participating in the IETF. The root cause for both of those explanations may be that IP addressing is a typical ISP issue and that engineers at ISPs may be more attuned to IPv4 address completion than are their counterparts in the business enterprises sector.

These personal musings aside, the IAB is interested in hearing more adoption stories, specifically those in which barriers were encountered and overcome. We also want to hear about barriers that persist—especially those cases in which some IETF-related work might help, whether it's a protocol, a best-current-practices document, informational material, or experimental work. In a future plenary we plan to report on both the lessons and the outstanding issues that we learned from those stories. Please send your stories to ipv6-adoption@iab.org. If your story speaks best to universal adoption of IPv6, then the tradition described in the first paragraph will be honoured and you will be provided with either a bottle of single-malt Scotch or a nonalcoholic beverage of your choice.

Peer-to-Peer, continued from page 1

RAI Peer-to-Peer Infrastructure Workshop

The one-day workshop began with a discussion that was designed to frame the issue and that included both the Internet service provider (ISP) and peer-to-peer (P2P) service operator viewpoints, as well as input with regard to additional scoping.

Jason Livingood and Rich Woundy provided some technical perspective from their vantage point at Comcast, the cable service provider. Participants learned that Internet service providers are observing enough P2P application traffic in their networks to impact customers' delay-sensitive applications and services, such as VoIP. The unacceptable customer experience occurs when one or more of a subscriber's delay-sensitive applications are noticeably degraded in performance because other subscribers' P2P applications are consuming all available bandwidth—that is, in excess of network planners' expectations. While the ideal solution might be to ramp up the bandwidth for each customer, there are both technical and business reasons that such a solution is not generally feasible. For one, some P2P systems literally know no bounds, thereby making aggressive use of all available bandwidth.

Jason and Rich's goals for a better future include the following:

- Optimize to provide the best possible network experience for the broadest set of customers, which means minimizing or eliminating cross-customer service quality impacts.
- Enable continued Internet evolution in a manner that avoids a game of cat and mouse—that is, detection and mitigation of specific protocols.

On the other side of the issue were Stanislav Shalunov and Eric Klinker of BitTorrent, who offered the P2P perspective. After walking through some of the complexities they face in order to deliver the reliable and quick file sharing for which BitTorrent is known, they said there is a point where P2P and network operations sometimes run at odds with each other. When attempting to locate sources of file chunks, P2P networks look for the least common chunks, and they figure out where to get them. These are the less available chunks of what you're downloading. However, from a network perspective, the location of these chunks may not be the most desirable. Stanislav and Eric observed that efficient choosing of peers could reduce transit traffic by more than 50 percent.

The afternoon presentations showcased proposals for moving forward. Some of the proposals focused on P2P technologies and oracles, such as P4P: Provider Portal for (P2P) Applications, by Haiyong Xie and Laird Popkin; Traffic Localization, by Yu-Shun Wang; and ISP-Aided Neighbour Selection in P2P Systems, by Vinay Aggarwal. Others focused on network and bandwidth management approaches, such as Bob Briscoe's Solving This Traffic Management Problem . . . and the Next, and the Next, which can be found at http://www.cs.ucl.ac.uk/staff/B.Briscoe/pubs.html#p2pi-solutions>.

The end of the day brought no immediate conclusions for IETF action items, but it did increase awareness and set the stage for BoF sessions to be held two months later at IETF 72.

ALTO BoF

The RAI area held the Application-Layer Traffic Optimization (ALTO) BoF in Dublin at IETF 72. Chaired by Enrico Marocco and Vijay K. Gurbani, the discussion focused on (1) possible approaches to exposing certain aspects of network topology to applications and (2) how P2P technologies might be enhanced to take advantage of that information.

According to draft-marocco-alto-problem-statement-02.txt, "Many of the existing overlay networks are built on top of connections between peers that are established regardless of the underlying network topology. In addition to simply achieving suboptimal performance, such networks can lead to congestions and cause serious inefficiencies. . . . [T]raffic generated by popular P2P applications often crosses network boundaries multiple times, overloading links which are frequently subject to congestion."

Both the presentations and the discussion at the meeting focused on issues surrounding caching and peer selection (P2P application questions), as well as on bandwidth costs. A draft charter for a working group was proposed and discussed. However, while there was support for moving forward with this topic area, the sense of the room was that the draft charter did not adequately capture a clear problem statement that partici-

Comic BoF

Plenary Report

By Mirjam Kühne


Mirjam Kühne

Note: This is not a complete report of the plenary sessions; rather, it is a summary of the highlights of the discussions. All IETF 72 presentations can be found at http://www.ietf.org/meetings/past.meetings.html.

"Even though Irish is the native language of Ireland, English has become the dominant language, just like IP is the dominant language of networking," said Kevin O'Callaghan, Ireland's leader of Alcatel Ireland, which served as host organization for IETF 72. Kevin welcomed participants to Dublin and said he was honoured to address the meeting. "It is fundamental to allow Nets around the world to communicate," he said. "The impacts on society have been truly beneficial."

On behalf of Alcatel-Lucent, which played a significant role in shaping the telecom infrastructure in Ireland, Kevin expressed his delight in having the opportunity to host and sponsor the IETF in Dublin. He was impressed by the number of people attending, the variety of organizations represented, and the number of languages being spoken.

Ireland's Green Party's minister of energy Eamon Ryan, who oversees communications, energy, and natural resources, addressed the group, saying that for politicians, the objective is to provide access and ubiquitous connectivity that society can use for critical issues, such as health care, education, and enterprise. There have been difficulties in achieving that goal in Ireland, and the move to ubiquitous connectivity has been slower than desired. In fact, in the past few years, the country has been playing catch-up in the areas of domestic and public use of the Internet. However, a new era of investment has been ushered in, and Ireland is becoming recognized as a place where companies can more easily invest in connectivity. Mobile broadband pickup has accelerated at a comparatively high rate, and an open-access system has allowed flexible development of new applications. This has clearly given the country an advantage in the area of digital services.

In his address, Minister Ryan echoed the sentiment of Vint Cerf, who suggested at a recent conference of the Organisation for Economic Co-operation and Development (OECD) that the development of the open-access system might necessitate a newer regulatory system than the one that grew up under the old phone systems. Minister Ryan said Ireland has challenged itself to apply new technologies with a view toward reduction of energy use. He said he's very encouraged by this type of meeting, because the IETF has a model that uses democratic processes at its very core. Peter O'Connell, director of strategy and regulation at eircom, the company that provided connectivity at the IETF in Dublin, said it was impressive to see so many people of so many backgrounds agreeing on standards.

Compared with similar companies worldwide, eircom may be small, but it is big by Irish standards. The company used to focus on voice, but now it focuses on broadband. Peter described the government as supportive of investment. "A policy platform that encourages investment is essential to the development of the Internet and to making both the Internet and the content that is available on the Internet accessible and affordable to the users," he said.

Continued on next page

Peer-to-Peer, continued

pants could support as IETF work. It was sent to the mailing list for refinement.

TANA BoF

The Transport Area held the Techniques for Advanced Networking Applications (TANA) BoF, chaired by Stanislav Shalunov and Gorry Fairhurst. The session focused on some of the transport-layer possibilities for supporting bandwidth-intensive applications.

According to http://www.ietf.org/internet-drafts/draft-shalunov-tana-problem-statement-01.txt:

"The TANA BoF is held to explore the problem space, to gauge the interest in the problems within the Transport area, and to see if the community and the area directors believe that it makes sense to form a TANA working group within the Transport area chartered to work on

1. standardizing end-to-end congestion control that enables advanced application to minimize the delay they introduce into the network and a protocol using it and

2. a document describing the current practice of peer-to-peer apps' use of multiple transport connections and recommendations in this space."

Discussion between those in the group included Jason's review of ISP requirements and Laird's review of P2P application requirements. A number of issues and angles were also discussed, and at the end of the meeting there was clear support for moving forward with work on the first item.

Plenary, continued from page 5

IAB Update

Following Internet Research Task Force (IRTF) chair Aaron Falk's update covering recent developments in the IRTF (details on page 33), Internet Architecture Board (IAB) chair Olaf Kolkman gave an update of IAB activities as follows:

What Makes for a Successful Protocol has been published as RFC 5218 (see a related article in the *IETF Journal,* Volume 3, Issue 3, December 2007). Design Choices While Expanding the DNS is technically approved, but it needs one more round of editorial review before sending it to the RFC Editor.

Photo by Peter Löthberg

Alcatel–Lucent managing director Kevin O'Callaghan.

A number of documents are now works in progress, including Principles of Internet Host Configuration, for which a call for comments will be issued shortly. The document titled Headers and Boilerplates is related to the work on RFC 3932bis (Procedures for IESG and RFC Editor documents) and the IRTF stream definition. It is expected that these documents will reduce the necessity for IESG statements by providing clearer guidelines for document authors.

There was quite a bit of activity in the area of Internet architecture. This past April, the IAB held a retreat in Stockholm, hosted by Netnod and Arceo. The retreat was designed as a workshop at which everyone brought up topics they wanted to discuss and after which a work plan was developed. A discussion on evolution of the IP model, including assumptions regarding which IP models are valid and which are not, was led by Dave Thaler. A discussion on peer-to-peer architecture was led by Gonzalo Camarillo. Gregory Lebovitz is leading an ongoing discussion on IPv6 deployment.

In addition, there have been a number of organizational activities. Bert Wijnen stepped down as IEEE 802.1 liaison and has been replaced by Eric Gray. John Klensin has been appointed liaison within ISO/TC46. Lars Eggert is succeeding Mark Twonsley as IESG liaison to the IAB. The IAB thanks Bert for his many years of good service. IAB liaison shepherds are successfully helping track and communicate the activities of the IAB liaisons.

Olaf showed a diagram of the structure of the joint working team on multiprotocol-label-switching (MPLS) extensions from the plenary session at IETF 71. He then reported on the joint working team of the ITU-T and the IETF, which is discussing issues related to MPLS. The team has issued a statement saying a transport profile for MPLS (MPLS-TP) will be developed that will take into account the ITU-T transport network requirements. The ITU-T will integrate MPLS-TP into the transport network and will align the current transport MPLS (T-MPLS) ITU recommendation with MPLS-TP. Further work on T-MPLS will be terminated.

Another new organizational development is the IETF's involvement with the OECD in cooperation with the Internet Society (ISOC). Together with ISOC and 15 other technical organizations, the IAB cosigned a memorandum on the future of the Internet in a global economy. The memorandum can be found at http://www.isoc.org/pubpolpillar/docs/oecd-technical-community-memorandum.pdf.

The IAB also participated in a technical forum on the future of the Internet Economy, which was organized prior to the OECD Ministerial meeting.

Olaf reminded the audience that the IAB is responsible for maintaining and defining the RFC Editor model, whereas the IAOC is responsible for the implementation of the agreement between the IETF and the RFC Editor. The RFC Editor contract will be up for bids in 2009. To guarantee continuity, a comprehensive model is needed. To that end, the RFC Editor function will be split into four functions:

- Independent Stream Approver
- RFC Editor
- Production House
- Publisher

The RFC Editor and Independent Stream Approver roles are new components. They may all be part of one vendor, or they could be separate. The question is, How do we select the functions or vendors? One possibility is through a request for proposals; another is through a NomCom process. The IAB welcomes suggestions.

The discussion took place on the RFC interest list, and a conclusion was planned for end of August. More details can be found on the IAB Web site.

Finally, Olaf pointed out that the IAB has a new logo, which was designed by IAB executive director Dow Street.

Typically, at IETF meetings an open-microphone session is part of each plenary. At IETF 72, the open-mic session was replaced by a technical panel titled IPv6 Experiences from the Field, in which five panellists described their experiences with IPv6 deployment in

their particular environment. (See article page 17.)

Administrative Updates

IETF Trust administrative procedures have been revised, reviewed by the community, and adopted. The RFC series has been assigned an ISSN (International Standard Serial Number), which will make it easier for libraries and other archives to identify them.

The Intellectual Property Rights (IPR) working group (WG) has been working on legal provisions for IETF documents. Even though a licence for code was previously developed by volunteers—the Nonprofit Open Source Licence 3.0 (OSL)—several issues have been raised by employees at for-profit enterprises. In the end, the trustees decided to replace OSL with the Berkeley Software Licence for volunteer code. (Additional details on this subject can be found on page 11.)

Jonne Soininen, chair of the IETF Administrative Oversight Committee, and Ray Pelletier, IETF administrative director, gave an update on the financial status of the IETF and announced hosts of future IETF meetings. IETF 74 will be hosted by Juniper and take place in San Francisco. IETF 75 will be hosted by Swedish country code top-level domain .se and take place in Stockholm. IETF 76 will be hosted by the WIDE project and take place in Hiroshima, Japan.

The high-level financial overview for 2008 looks fairly positive. A total of USD 650,000 in meeting sponsors has been secured by the Internet Society. Those three meetings will contribute USD 1.1 million to be invested in other secretariat activities, IETF Administrative Support Activity (IASA) activities, and RFC editor activities. ISOC will contribute USD 1.55 million to IASA's 2008 budget from its organizational member contributions and other sources. Expenses are projected to be slightly under budget for 2008, and the contingency budget of USD 50,000 remains intact.

Another new organizational development is the IETF's involvement with the OECD in cooperation with the Internet Society (ISOC). Together with ISOC and 15 other technical organizations, the IAB cosigned a memorandum on the future of the Internet in a global economy.

Jonne also announced that during the remainder of 2008 and in 2009, a number of RFPs would be announced, including those for the meeting network contract, the RFC editor, and the IETF secretariat.

Edu Team Report

The Edu team is responsible for organizing the tutorial sessions on the Sunday prior to each IETF meeting. Its mission is to manage the internal education activities of the IETF and to offer training and other educational material that "improves the effectiveness of the IETF operations." While the Newcomers Tutorial might be the best-known, there are three different types of tutorials:

- Process-oriented topics: bringing new work into the IETF and document life cycle
- Training on tools: XML2RFC and IETF tools
- Technical topics such as security, DNS, routing, and IPv6

The Edu team also organizes topical trainings for WG chairs during each IETF meeting.

Recently, the Edu team Web site was transitioned to a new wiki site, which makes it more stable and easier to update.

One open issue that comes up from time to time among Edu team members is the role of the technical tutorials and what the tutorials should cover: Should they be introductory-level cross-trainings for IETF participants or should they cover in-depth training on specific technologies? Should there also be training on timely or controversial issues or should the training focus on the technical knowledge needed to produce high-quality IETF specifications?

The Edu team is seeking feedback on those questions and would welcome e-mail sent to edu-discuss@ietf.org.

IESG Open Mic

During the open-mic session on Thursday, a lengthy discussion focused on both the process and the usefulness of so-called PROTO write-ups—a particular way of providing feedback for Internet-Drafts submitted to the IESG. The IESG said the PROTO write-ups provide feedback that is helpful to IESG members, especially in areas where an IESG member is not expert in the subject area. A discussion followed about the practice of having the IESG make decisions during telephone chats and via other means—a practice that may not be as transparent to the community as it could be.

According to Russ Housley, the IESG's use of Data Tracker has made

Continued on next page

Photo by Peter Löthberg

IETF 72 participants meet in hotel lobby.

Plenary, continued from page 7

the entire process much more transparent. "Now, anyone can see the status of the review and see what comments need to be resolved for the document to progress," he said.

While most participants agreed that the tracker is a valuable tool, some suggested it could be enhanced so that it can be used more consistently. Russ confirmed that review of the tool is already on the to-do list for the IESG.

Photo by Peter Löthberg

Internet pioneer Vint Cerf chats with IETF 72 attendees.

IESG member Magnus Westerlund pointed out the need for more bottom-up review. "Cross-area reviews need to be happening," he said. "Many of the issues that come up in a DISCUSS should be addressed early on in the review process." A DISCUSS is a certain way for the IESG to provide feedback for an author of an Internet-Draft.

IAB member and liaison to the IETF Loa Andersson said the problems that were being discussed stem from a relatively small number of DISCUSSes. He said he felt that, overall, the work is being done well. "I have been a WG chair for some time, and I have experienced a number of area directors," he said. "Usually, the DISCUSS comments have helped improve the document."

At the end of the IESG open mic session, the suggestion was made that the IETF consider extending the meetings to Friday afternoon. Most of those who commented on the subject were opposed to that suggestion and made some other suggestions instead. For instance, one participant suggested the meetings start earlier in the mornings or that there be more calls between meetings. In general, it was felt that more work needs to be done outside the three meetings that are held each year.

The discussion continued on mailing lists following IETF 72, and in the meantime, the IESG has decided to proceed with the suggestion at IETF 73 in Minneapolis and to provide meeting slots until 15:15 on that Friday.

IAB Open Mic

During the IAB open mic session, a participant asked what the IAB thinks about adding new congestion-control algorithms to TCP: Do we view an aggregated UDP/IP header as just the layer 3 datagram layer over which we run this TCP implementation, as opposed to sticking with TCP? Or are we using another congestion-controlled transport like SCTP or DCCP? This issue came up in the Techniques for Advanced Networking Applications (TANA) BoF that met during IETF 72.

A variety of views were expressed by the IAB in response. On one hand, Stuart Cheshire said he thought it would be a good approach to "experiment with this at the user level running over UDP. Then, when the algorithm is worked out, it can be standardized and then gradually made into the mainstream TCP implementations over a longer time frame, like five years." This would recognize the tension that exists between the IETF, which makes long-term standards, and the companies that want to ship products.

On the other hand, Dave Oran suggested that the community proceed with caution with regard to congestion control algorithms. "There is a balance to be struck," he said. "Involvement of the sponsoring ADs and the relevant ICCRG [Internet Congestion Control Research Group] is important. Let's move with much speed and low haste."

Then should one use UDP for peer-to-peer communication between peers that are stuck behind NAT gateways, and TCP for everything else?

No, said Stuart, who answered that one can have NAT-to-NAT peer-to-peer communication equally well with either TCP or UDP. "The reason TANA wants a new congestion control algorithm is that they want something less aggressive than today's TCP," he said. "It has nothing to do with NAT gateways."

Scott Bradner remembered that when he was a transport AD, many people wanted to use UDP, that usually, it was a way to say that TCP was too heavy and too slow. He cautioned that one should "be very concerned about the underlying excuse."

Photo by Peter Löthberg

Shuttle takes IETF participants to Dublin.

Real-Time Text

By Arnoud van Wijk

When we want to communicate electronically, most of us use voice and, at an increasing rate, video. When we do, such communications occur in real time[1]; that means that we send and receive audio and video *continuously as we communicate,* and we consider this as the normal way to converse with each other.

For most of us, text is a static medium. We use it to read newspapers and Web sites; we exchange text messages by using mobile phones; and we use instant messaging on our computers to communicate with each other while doing other tasks. When we need efficient conversation, we pick up the phone and call the person.

But what do we do if we're unable to use a telephone because we can't hear or speak or because we're in an environment or a situation where the use of voice is inappropriate, such as in a restaurant or during a meeting? What if we find ourselves in danger and need to contact the police without being heard?

The solution is real-time text. For the majority, this will be a valuable additional communication medium besides audio and video. Real-time text can be used either as the only communication mode or together with audio and video, which is called total conversation.[2]

For those who are deaf or hard of hearing or who have a speech impairment, real-time text is an essential communication capability. This is especially true in the current era. Every day, we all depend more and more on the telephone for immediate contact. Social contacts are maintained, business is conducted, and even our safety depends on the telephone. With real-time text as a mainstream communication feature, Internet telephony is for everyone.

The Technology Behind Real-Time Text

Real-time text works by sending and receiving text on a character-by-character basis. The characters are sent immediately (within a fraction of a second) once typed and are then displayed immediately to the receiving person(s). This allows text to be used in the same conversational manner as voice. It's like talking by using text.

Real-time text that runs over IP networks is designed around the ITU T.140 real-time text presentation layer protocol. T.140 allows real-time editing of text, even in cases of backspacing and retyping. T.140 is based on the ISO 10646-1 character set, which is used by most IP text specifications, and it uses the UTF-8 format. This allows any language to be used with real-time text, including English, Chinese, and Russian.

Real-time text uses the same real-time transport protocol (RTP) as voice over IP (VoIP) and video over IP. The text is encoded according to IETF RFC 4103 (RTP Payload for Text Conversation), which supports an optional error-correction scheme based on redundant transmission (as described in RFC 2198). This results in a very low end-to-end character loss across IP networks that have moderately high packet loss. (It also makes it very good for wireless accesses.)

To improve efficiency, the text is buffered for 300–500 milliseconds before it is sent while still meeting the real-time text performance requirements of RFC 4103. The traffic load of real-time text at 30 characters per second is between 2 and 3 kilobits per second depending on the language used (including the overheads for RFC 4103 with the maximum level of redundancy, RTP, UDP and IP).

Real-time text uses the standard session initiation protocol (SIP) (RFC 3261) and the session description protocol (SDP) (RFC 4566). SIP is used without any alteration; there is no difference between real-time text and VoIP for SIP. The real-time text encoding is identified by using the SDP media definition m=text.

To the deaf or hard of hearing, real-time text feels more like conversation.

To ensure proper technical implementation and use of real-time text, RFC 5194[3] lists the essential requirements for real-time text and defines a framework for implementation of all required functions based on SIP and RTP. This includes interworking between real-time text and existing text telephony on the public switched telephone network and other networks.

The ECRIT IETF working group defines real-time text as one medium in the access to emergency services (see RFC 5012, draft-ietf-ecrit-phonebcp and draft-ietf-ecrit-framework). With the growing number of people with hearing and/or speech impairments, it would be prudent for multimedia emergency public-safety answering points in Europe (which uses 112) as well as in the United States and Canada (which use 911) to support real-time text.

Mainstream Use of Real-Time Text

The use of real-time text will not be limited to those who cannot use speech. Like captioning on TV, real-time text will be used more by people without disabilities than by those who have them. On phone calls, people can type in phone numbers, addresses, names,

Continued on next page

Real-Time Text, continued from page 9

and other information better passed in text than dictated—especially when the two communicators have different accents. When using one line or otherwise occupied in a conference, a person can answer a second line in text only and receive a quick message or have a quick text conversation. When talking to an elderly parent, for example, a person can use text to supplement voice to make sure important information has been understood. In interactions with an interactive voice response system, instead of having to wait while the voice slowly reads out all the choices, real-time text can provide an almost instant list of the choices visually so users can immediately read and select the number they want to press. These and a myriad of other uses will become common as real-time text gets deployed as a natural and always present parallel communication mode on any voice phone call.

The Real-Time Text Taskforce

Launched on 30 July 2008, the Real-Time Text Taskforce (R3TF) is an independent open forum for engineers, motivated individuals, experts, companies, and organizations that wish to help test, implement, and advance the widespread adoption of the real-time text framework.[4] Its goal is to ensure that real-time text is as readily available as voice for all users. The Internet Society is assisting in the effort by serving as an incubator of the R3TF.

Having a single real-time text standard that is used everywhere would make access to communication services easier, and it would eliminate any potential interworking issues. Unfortunately, with the diverse types of communication networks and devices that are in use today, this is not possible.

The R3TF will promote real-time text as the real-time text standard that most terminals and networks can either use native or easily interconnect with via gateways between different network borders. This means that alternative real-time text protocols may be used, but they must be able to interconnect via gateways with real-time text to ensure full interoperability. This will make possible the goal of real-time text being available everywhere.

The R3TF will help facilitate the development of interworking test beds that will enable implementers to test how well their solutions comply with the standard. Moreover, the task force will facilitate the distribution of information about the technology, about the technology's user requirements, and about the technology's implementation, and it will act as an educator on related issues.

The Web site of the R3TF is http://www.realtimetext.org.

The R3TF Is for Everyone

What can you do to help the R3TF?

- Add your knowledge and expertise to those of the R3TF so that the task force can grow real-time text and remove barriers to its implementation.
- Help prevent SIP networks from blocking real-time text traffic, even if they block VoIP traffic for internal reasons. In SIP networks and products, real-time text support, as described in RFC 4103, should be regarded as normal mainstream and possible now. Real-time text not only works now; it also is part of next-generation-network system specifications.
- For non-SIP network and products, ensure that the real-time text protocol/service used does interoperate with RFC4103.
- Build on the open-source client that supports real-time text (as well as VoIP and video) to implement clients for mobile/cellular terminals as well as computers.
- Include real-time text in your design and development of new services, such as voice and videoconference services, answering machine services, call centre services, language interpretation services, gateway services, network interconnection services, and emergency services. All of those services are enriched to higher usability via support of real-time text together with voice and sometimes video. Open-source components are available and should be continually improved when possible.
- Include real-time text in 112/911 emergency services when they move to IP networks. Authorities are already pushing for this in the European Union and the United States.
- Become an R3TF sponsor and encourage employees to participate in projects.

References

1. The International Telecommunication Union (ITU) defines real time in ITU-T F.700 Section 2.1.2.1. Real-time text is defined in ITU-T F.700 Annex A.3 and ITU-T F.703 Section 5.3.2.3.

2. Defined in ITU F.703 Section 7.2 and specified for the next-generation-network IP Multimedia Subsystem in 3rd Generation Partnership Project TS 26.114, Multimedia Telephony, Media Handling and Interaction.

3. RFC 5194 was published by the IETF as an informational document.

4. The Real-Time Text Taskforce is a project group separate from the IETF.

Arnoud van Wijk is disability projects coordinator for the Internet Society. Along with other researchers, Arnoud documented the technique for real-time text described in this article, combining existing IETF standards to facilitate text streaming over IP networks. He and Guido Gybels, director of New Technologies at RNID (http://www.rnid.org.uk), with contributions from other experts in communication and accessibility for people with disabilities, edited and coauthored Framework for Real-Time Text over IP Using the Session Initiation Protocol, *which the IETF recently published as information document RFC 5194.*

Talking with Jorge L. Contreras

The attorney for the IETF talks with the IETF Journal about intellectual property, licensing, and what makes the IETF so unusual in the world of copyrights and copylefts.

Photo by Mirjam Kühne

IETF attorney Jorge L. Conteras

IETF Journal: How did you get involved in the IETF?

Jorge L. Contreras: My firm, WilmerHale, has provided legal advice for the IETF from the early days. I took over from a partner who left about 10 years ago and have been working with IETF since then.

IJ: Did you ever think you would be interested in Internet standards?

JLC: I have an undergraduate degree in electrical engineering, so I was always interested in technical issues. But the law that applies to standards organizations was much less developed when I began practicing law in 1991. The first case to draw significant attention to intellectual property issues in standards organizations involved Dell Computer. In 1998, Dell participated in an SDO [standards development organization] that was creating a video bus standard. They did not comply with the IP policy of the SDO and then later tried to enforce their patents against other SDO participants. The U.S. Department of Justice brought an action against Dell, with the result that Dell's patents were no longer enforceable. It was a significant result and one that really shook up the world of standards law.

Today, most companies that come to the IETF are aware of the Dell case and a number of more recent cases that have elaborated on the issues first raised in Dell.

IETF participants often have patents that affect standards developed at IETF, and so long as they disclose these patents in compliance with the IETF IPR [intellectual property rights] rules, they are fine. If they violate the rules, however, they risk the loss of patent rights, just as Dell did.

IJ: Has the IETF ever been sued under patents?

JLC: No, the IETF has never been sued for patent infringement. An SDO like IETF is just a forum in which participants can develop standards. IETF itself does not make or sell products, so it cannot actually infringe a patent. It is the implementers of a standard, who sell a product that implements the standard, who can be accused of infringing.

That is not to say, though, that IETF is never involved in patent litigation. Because patents are increasingly important to the companies that send participants to IETF meetings, there are sometimes lawsuits between those companies. If the lawsuits involve patents covering standards that were developed at IETF, the IETF is often called upon to provide evidence regarding the IPR policies that were in effect at the time, the participants in certain working groups, and the contributions that various parties made to a standard. In such cases, the IETF is not a defendant but merely acts as the provider of information that is essential to resolution of the case.

IJ: What other legal matters do you handle for IETF?

JLC: I advise IETF on a wide range of legal issues. One issue that has gotten attention recently is the licensing of IETF software tools to the community. The tools can be created by companies contracted by the IETF—like the IETF Secretariat—or by volunteers. There is a strong desire to release the tools on an open-source basis, and I have worked with the tools team, the IAOC and the IETF Trust to evaluate various approaches that might satisfy the many constituencies involved with the IETF. For example, certain open-source developers favour the General Public License, or GPL, but commercial enterprises often have an aversion to the GPL because it is viewed as risky when distributed with proprietary software. As part of the process, we helped develop a new variant of the Open Software License [OSL] that could be used with nonprofit enterprises. There were objections to the license from for-profit enterprises. Finally, we settled on using the open-source BSD licence, but it took many rounds of discussion to get there.

IJ: Is working with the IETF different from working with other organizations?

JLC: Oh, yes, the IETF is unique. Often, it is not actually clear who the client is. The IETF is more a concept than an entity. The Internet Society [ISOC] is the organizational home of the IETF, and ISOC signs most of the contracts for services that the IETF needs, including everything from the RFC Editor and IETF Secretariat to contracts for hotel venues and network services.

The IETF Trust is a different legal entity altogether. It is the custodian of the IPR owned by IETF, including software code, the IETF trademarks and logo, written records, and, most important, the RFC series itself.

Continued on next page

Jorge L. Contreras, continued from page 11

Now that the IETF Trust is up and running, it is also available to administer rights in older RFCs. Those rights are generally still with the RFC authors, but the Trust has set up a lightweight process for authors to license their old RFCs to the Trust. A number of companies and individual authors have already done that, and we are encouraging other active IETF participants to follow suit.

IJ: Outside of the IETF, do you work with other standards bodies?

JLC: I work for numerous companies that are involved with other standards groups, both large SDOs like IEEE and smaller consortia and special interest groups. I also get involved in projects that affect Internet standards groups, like ISOC's establishment of the Public Interest Registry to become the domain name registry for the .org top-level domain. That was an interesting project, and I continue to advise organizations in that area.

IJ: What's been happening in the IETF's IPR Working Group recently?

JLC: The IPR WG has nearly completed a revision of the IETF IPR policy related to copyrights [currently at BCP 78]. The revised policy will make a lot of things easier to manage. For example, if someone wants to evolve an IETF RFC in an SDO other than IETF, there's no way for IETF to grant that permission today. The other SDO must seek permission directly from the original document authors. In some cases, that is very difficult, because authors have left their companies, moved on to other projects, or, in some cases, passed away. For example, in one case, we transferred some of the 802.11 MIB work to IEEE. That took some doing, because some of the RFCs were quite old. In that case, IEEE had to go and contact the authors, so things took a little while. Under the new policy, the IETF Trust will be empowered to grant licences—subject to specific guidance that has been given by the IETF community.

IJ: How about the IETF's patent policy?

JLC: The patent policy is now codified at BCP 79. The IETF has a process whereby participants have to disclose their patents if the patents are related to work that is ongoing in an IETF working group. You can also disclose other people's patents if you know they affect the work in a WG. If you know of a third-party patent that could affect the development of a certain protocol or standard, you can disclose it.

In this way, the IETF is unusual: It does not require a patent holder to license its patent. The patent holder has to disclose it, but it is not obligated to license it. This has not discouraged companies from sending participants to IETF meetings, however. And in fact, many companies license their patents covering IETF standards for free. In some areas, however, there are lots of patents, and people do pay royalties on them.

Not long ago, some IETF participants said they believed that a certain company's patent would cover all of IPv6. It caused quite a stir in the community. I had several discussions with the company and tried to get them to commit to licence it for free. Eventually, they never specified what they thought the patent actually covers, and to my knowledge, they never licensed it. But I'm also unaware that they have ever tried to enforce the patent against implementers of IPv6, so that's where it stands today.

IJ: It sounds as if a lot of what you do is to manage people's expectations about the IETF, is that right?

JLC: Yes, the IETF is an unusual organization, and even though it's very well-known, not many people—especially other lawyers—really understand how it all works. So I spend a lot of time talking to other lawyers about the IETF processes. These include both the in-house lawyers of IETF participants and lawyers for litigants that involve IETF standards, parties that IETF contracts with, and others.

IJ: What about your book? Will there be a version 2?

JLC: I'm chair of the Technical Standardization Committee of the American Bar Association's Section of Science and Technology Law. In that capacity, I served as editor of our committee's Standards Development Patent Policy Manual, which was published last year and is intended to offer a set of annotated, policy-neutral language that SDOs, consortia, and others can use in developing their own patent policies. A number of IETFers worked on the book project, and I am grateful for all of their help and insights.

As is often the case, right after the publication there were changes to laws and regulations, new cases came up, and so on. So, yes, it will need an update, and I'm currently looking for volunteers to help with that project!

For more information about Jorge L. Contreras, please see http://www.wilmerhale.com/jorge_contreras/. For a list of publications, visit http://www.wilmerhale.com/publications/whPubsList.aspx?Attorney=685d7e20-d9c6-4a01-82f3-5712e6929245

Jorge is the author of Standards Development Patent Policy Manual *(paperback), which can be found at http://www.amazon.com/Standards-Development-Patent-Policy-Manual/dp/1590319281.*

Note: Standards Development Patent Policy Manual *can also be found on Google Books.*

The Internet and Bandwidth-Intensive Activities

By Leslie Daigle

Scattered across the globe, a number of networking issues are being noted that can loosely be classed as stemming from bandwidth-intensive activities. These are applications and services that cause traffic that is higher in bandwidth, that follows paths neither anticipated nor desired by the network architect and operator, and that has an impact on other network activities traffic. As a result of the degradation in service the situation causes for other users, network providers often turn to bandwidth management in an attempt to curtail or restrict the excessive flow.

In the United States, one such situation emerged when network service provider Comcast and P2P file sharing service BitTorrent wound up at odds over whether Comcast's bandwidth management activities were unduly restricting the BitTorrent service. The case was not unique, but it shed light on the situation and prompted a healthy discussion at the Real-time Applications and Infrastructure area's peer-to-peer (P2P) Infrastructure Workshop, which was held in May 2008 (see page 1).

As the IETF considers the work it might undertake to address these types of issues, some of the questions on people's minds are, What is the basic problem? Is it an issue with a particular application (P2P)? Or is it a question of network providers' inability or lack of interest to reprovision network segments?

The answer is not simple. From a technology perspective, the Internet is activity agnostic. Therefore, in principle, any popular application or service could run traffic in directions that have nothing to do with the way actual networks are built. In fact, more data-intensive (and network-topology-independent) applications and services are rolling out and causing these types of bandwidth issues. For example, virtual realities, such as Second Life, along with other video and audio streaming applications are also having noticeable impacts on networks. Nevertheless, some of the issues are exacerbated by the nature of P2P technology. By definition, peers and peer connectivity are not tied to network topology, so traffic does not necessarily follow expected network paths. Furthermore, P2P file-sharing services are designed to store and share large chunks of data, so they tend to use all available bandwidth in data transfer between peers.

These are not new problems. The Internet, in its global reach and local reality, has been dealing with, at least, pockets of exorbitant demand for bandwidth since its inception. Typically, the demand has been dealt with as a network operational issue, not a technological issue. To use an example from the all-but-forgotten past, FTP traffic was, at one time, one of the largest sources of expensive, transoceanic link traffic from Europe to the United States. Paying for a European Archie anonFTP archive index for the purpose of giving precedence to European sites mirroring the same files made sense as a way of reducing that traffic demand.[1]

The best path forward appears to be a dual-pronged approach—in other words, finding ways to allow such applications as P2P technology-based ones to (1) fine-tune their use of network bandwidth availability and (2) identify more-palatable options for managing bandwidth in the face of overwhelmed network links. These are the types of activities the IETF considered in two BoF sessions at IETF 72 in Dublin (see page 1).

As we engage in IETF activities on the topic—including BoF sessions and, eventually, working groups—it's important to keep a few simple realities in mind. First, it's not strictly about P2P technology itself. Second, the network impact might be local (as in, "My P2P participation is wrecking my neighbour's VoIP [voice over Internet protocol]") or transit (such as spikes in peering costs due to unexpected load). Third, there are different classes of reasons that the network operator may have no reasonable incentive or ability to adjust the network to meet demand. These include the constraint of dealing with significant, unmatched expenses (such as no additional revenue to offset the capital cost) or dynamically changing bandwidth demands. As an example of the latter, network operators have little or no control over which peer becomes a supernode in a P2P network.

Note that none of this is to be confused with traffic unintended by any local customer (unwanted traffic, denial of service, etc.), which does not need accommodation so much as remediation.

In today's Internet, the broader impact of dealing with existing problems through traffic shaping, with its unintended consequences, or through tiered Internet access has the potential for a chilling effect on openness and innovation.

The long view is that this sort of stretch in network demand is normal and, on a global level, healthy. By making a network to ship packets around, the model is about packet shipping; it's not about making and maintaining highly specialized connections. To address the current issues, we need to consider approaches that not only solve the immediate problem but also are applicable beyond any particular application technology and that do not introduce such architectural complexity as to limit aspirant applications.

1. http://www.savetz.com/articles/ibj_bunyip.php.

IETF 72 Welcomes ISOC Fellows

By Wendy Rickard

Six technology specialists and researchers from Africa, Asia, and South America journeyed to Dublin, Ireland, for their first IETF meeting as part of the Internet Society's (ISOC's) Fellowship to the IETF programme. Guided by their mentors and supported by ISOC staff, the fellows had been selected from among dozens of applicants and given opportunities to sharpen their technical skills, indulge their interests, and meet face-to-face with colleagues and others whom they'd known only by reputation or by way of time spent on working group (WG) mailing lists.

Photo by Gerard Ross

ISOC Fellowship to the IETF programme fellows and mentors at IETF 72 in Dublin.

By all accounts, the fellows benefited considerably from the experience, bringing with them—and taking away—their own unique perspectives.

Meet the Fellows

Born and raised in a small city a few kilometers from Santiago, Chile, Hugo Salgado developed a passion for mathematics as a child. Although he began his studies by pursuing a degree in physics at the Universidad de Chile, in his third year he discovered the Internet and changed his major to computer science. Today he works for NIC Chile, the .cl country code top-level domain registry where he develops software written in Perl, maintains a few Web sites, and implements DNS-related technologies, such as IDN. Even though Internet security has become a passion of late, Hugo writes that it is the IDN work that interests him most. "We were the first Spanish TLD with IDN," he writes, "so we are very concerned about the changes that came with IDNAbis."

Attending an IETF meeting was an eye-opener for Hugo. "Everyone was friendly and open-minded," he said. "That makes for a very rich environment for developing ideas and to be creative in." Following the trip, Hugo planned to participate in mailing list discussions and to spread the work of the IETF. "We are currently preparing our first informational RFC submission on the .cl extensions for EPP registration," he added.

Ali Hammad Akbar is an assistant professor in the department of computer science and engineering at the University of Engineering & Technology (UET), the largest engineering university in Pakistan. He also consults to UETs Al-Khawarizmi Institute of Computer Sciences, where he helped establish a research group called WATCHNETs (wireless and ad hoc actuator and sensor networks).

No stranger to travel, the Pakistani native recently earned a Ph.D. in electrical engineering from Ajou University in South Korea, a master of science in electrical engineering from the University of New South Wales in Australia, and a bachelor's degree in electrical engineering from the National University of Sciences and Technology in Pakistan. He is interested primarily in 6LoWPANs, particularly in mobility, security, and routing issues. (See page 16.)

Mohamad Dikshie Fauzie conducts research in Internet measurement analyses in dual-stack IPv4/IPv6 environments at Keio University in Japan. Born and raised in Indonesia, Mohamad supplements his research with work as a School on Internet–Asia (SOI-Asia) operator at Bandung Institute of Technology, where he operates dual-stack IPv4 and IPv6 networks for real-time distance learning using IP multicast operation. He's also actively involved in Indonesia's IPv6 task force, a group consisting of academics, government representatives, and representatives from telecommunications companies and Internet service providers that are working to implement IPv6 in Indonesia.

Attending IETF 72 helped Mohamad gain a better understanding of the

IETF 72 Fellows and Mentors

Alejandro Acosta (Venezuela)
Mentor: Elwyn Davies
Ali Akbar (Pakistan)
Mentor: Carsten Bormann
Tamrat Bayle (Ethiopia)
Mentor: Hesham Soliman
Mohamad Dikshie Fauzie (Indonesia)
Mentor: Erik Nordmark
Hugo Salgado (Chile)
Mentor: Patrik Fältström
Kumar Saurabh (India)
Mentor: Hadriel Kaplan

IETF 72 Returning Fellows

Alberto Castro (Uruguay)
Martin German (Uruguay)
Subramanian Moonesamy (Mauritius)

problems associated with IPv6 deployment and Protocol Independent Multicast (PIM), an area that is critical to his research. He plans to use the knowledge he gained in Dublin as the basis for a report to Indonesia's IPv6 task force about the latest developments. He also plans to write a report to SOI-Asia about the status of PIM–Sparse Mode.

University professor and researcher Tamrat Bayle possesses a deeply held belief in the power of a robust information infrastructure to support his country's IT industries. He also says he's a strong believer in the IETF's contribution to the success of the Internet. This native Ethiopian is an assistant professor and head of the Department of Information Technology at the College of Telecommunications and Information Technology of the Ethiopian Telecommunications Corporation. There he teaches graduate-level courses in advanced computer networks, data communications, mobile ad hoc networking, and wireless IP networks. He also advises on master's degree theses and serves as a principal investigator on a project whose aim is to increase classroom interaction with mobile wireless technology.

Photo by Gerard Ross

Fellow Mohamad Dikshie (right) with mentor Erik Nordmark.

Attending IETF 72 offered Tamrat a unique opportunity to enhance his knowledge and to contribute more integrally to the growth of the emerging IT industries in his country. "The Internet is the new frontier in our country," he writes. "It is my strong belief that [attending] the IETF meeting will assist my country in general—and the Ethiopian Telecommunications Corporation in particular—in finding solutions for efficient and scalable internetworking needs."

In his role as internetworking coordinator with British Telecom in Venezuela, Alejandro Acosta was recently involved in the implementation of IPv6 within the company's network running BGP, firewalls, Linux, and IP services, including SSH, Apache, and DNS. "My responsibilities in the company include everything related to the IP protocol," Alejandro writes, "including devices, services, connections, troubleshooting, and VoIP using Open Source."

Since he was in high school, Alejandro has been interested in "everything related to computers and specifically with the Internet." His main areas of interest today include IP, BGP, DNS, and routing protocols. "The IETF community has had a big impact on those fields and, therefore, on me," he writes. Attendance at an IETF meeting not only brings Alejandro face-to-face with technologists working on the issues that interest him most; it also enables him to bring the knowledge he gains to the two national universities with which he works closely, as well as to the groups, such as LACNIC.

Kumar Saurabh says the IETF is one of the main forces behind the evolution of the Internet and its associated standards. The software engineer at Sonus Networks in India is helping design and implement Border Gateway Function (BGF) by using H.248 protocol. He also works on Session Border Controllers (SBCs), which involves extensive application of the session initiation protocol (SIP) and which explains his interest in the IETF's SIP working group.

At Sonus, Kumar is involved in developing next-generation telecom products, such as SBC, BGF, and softswitch. "My work involves developing applications using IETF protocols like SIP and Megaco," he writes. Kumar hopes to use his experience at the meeting to further enrich his knowledge base of those protocols. "No one in Sonus Bangalore has ever attended an IETF meeting," he writes. "I hope to be able to tell them how important the IETF meeting is to our work."

As part of a new initiative, ISOC launched its Returning ISOC Fellowship to the IETF programme at IETF 72. For the first time, fellows who attended a previous IETF meeting were able to return for IETF 72. ISOC invites applications from alumni of the fellowship programme who wish to actively participate and contribute to the work of the IETF and who feel that attending another meeting in person is essential to their own professional development and local community. For more information, see http://www.isoc.org/educpillar/fellowship/returningfellows.shtml.

ISOC extends opportunities for organizations to become sponsors of this important programme. Sponsorship demonstrates a commitment to technical capacity building in less-developed regions and shows support for extending participation in the IETF to those in developing countries. It also creates opportunities to build contacts with technologists and potential regional leaders who are highly knowledgeable about conditions in developing countries. Those who are interested in becoming an ISOC Fellowship to the IETF sponsor should contact ISOC at fellow-sponsor@isoc.org.

Joining the IETF Fold

Ali Hammad Akbar

An IETF 72 fellow reflects on working group politics, the culture of leadership, and the IETF dress code.

I'd like to take this opportunity to express my sincere thanks and appreciation to the Internet Society and its sponsors for their sponsorship of the ISOC Fellowship to the IETF program and to the organizers at the IETF, who made it possible for me to participate in the 72nd meeting of the IETF. All of the travel, lodging, and hospitality arrangements extended during the stay were splendid and were handled very professionally. My indebtedness is to Leni Nazare, Martin Kupres, and Mirjam Kühne for their efforts.

Though I was by and large familiar with the workings of IETF, I was indeed a newbie to its proceedings. In the past, I would wonder how working group (WG) ideas popped up: was it the WG chair who would spell out the seemingly Utopian ideas? Or was it the companies that would steer the chartering of the agenda items? It was only through firsthand experience at an IETF meeting that I became able to fully comprehend the working style: a clearly chalked-out agenda is thoroughly deliberated and firmed up through consensus on the mailing lists. Afterward, it gets floated for consideration at the IETF meeting. Objective synoptic presentations and question-and-answer sessions after each Internet-Draft expose the participants to a therapeutic forum wherein they reflect upon and make appropriate amends to their proposals. The ideas evolve through debate and candid comments, though at times these might seem like forays into an Internet war zone! Undoubtedly, the two consequent features of IETF meetings—synergy and proactivity—arise from honest debate that sometimes relegates courtesy to a lower priority. These are the people for whom the maxim holds true: "If violence is a sin, silence is a felony."

My doubts were further cleared up when I met my mentor, Prof. Carsten Bormann, a pleasant, modest, and knowledgeable person who maintained a characteristic silence unless provoked and who chaired the 6LoWPAN working group. While attending the meeting as a first timer, I found myself awkwardly overdressed—not at all in the getaway style of the majority of participants. Professor Bormann promptly commented in a lighthearted manner that of all of the participants, only two gentlemen were clad very formally—a dig on me, I realized!

> I realized that people here at the IETF accept being led only by those who are focused, doers, and team players—and who are not bothered by carefree attire and easy style.

The meeting also afforded me interaction with other fellows from all over the globe. While exchanging notes with Hugo Salgado from South America and SM from Mauritius, I realized that such people from diverse backgrounds and of a multitude of ethnicities, who have will and potential, very often remain untapped contributors. Though their passions might find expression through an open call for participation on the IETF mailing lists, it is only through bursaries and fellowships that they can get recognition for their efforts. It was indeed very rewarding for us all to be part of such an activity.

At times, my admiration for academic stalwarts has bordered on infatuation. And there could be no better place than the IETF meeting to get them all together. To my immense pleasure, I could meet Prof. David Culler and Samita Chakrabarty and attend talks by Pascal Thubert and Jean-Philippe Vasseur—people who work in subjects closely related to my interests and who are highly regarded. Likewise, it was quite motivating to see young scholars like Jonathan Hui and Phil Kevin at 6LoWPAN and ROLL working groups to present their works. Their success stories bespeak of their professional approach and relentless pursuits.

With limited leisure time at my disposal, a city tour seemed like a good extracurricular activity, so I set out to hitchhike. While I went to the city centre in a more hop-on, hop-off mood, my touring spree turned out to be more academic. My visit to the highly acclaimed Trinity College and the National University of Ireland enabled me to meet some very good researchers in Dublin. Prof. Murphy from University College Dublin visiting faculty from Washington State University, and Prof. Sumit Roy were among the professors I met who work on IP-based wireless networks. I plan to take those initial interactions on to subsequent levels of rapport.

As a whole, the experience of spending just about a week in the IETF folds proved remarkably eventful. It could, however, result in more than that. For instance, pursuit of an Internet-Draft together with my mentor—in order to explore ways and means of opening a new ISOC chapter in Lahore (my hometown) and offering talented minds for ISOC—all are merely the initial contemplations we wish to pursue.

IPv6 Deployment: Lessons from the Trenches

An IETF 72 panel looks at experiences with IPv6 deployment.

Photo by Peter Löthberg

IETF 72 panel discusses IPv6 deployment.

During the IETF 72 technical plenary, the Internet Architecture Board (IAB) hosted a panel on the subject of IPv6 deployment. The five-member panel was composed of Internet community members who have firsthand experience with operational IPv6 deployments. They represent the perspectives of Regional Internet Registries, or RIRs; network operations teams; broadband services; content delivery services; and host applications. The panelists communicated their IPv6 adoption successes and hurdles, as well as their IPv4-depletion contingency plans, including carrier-grade network address translations (NATs), or CGNs, a concept that is currently under debate.

Motivation and Background

Studies suggest that the completion of IPv4 address allocations by IANA to the RIRs could occur as early as 2011 (see http://www.potaroo.net/tools/ipv4/). Regardless of the exact date, the idea of an empty storehouse will, without question, change the networking landscape. In an effort to prepare for that eventuality, various players are working toward widening and hastening IPv6 deployment.

Some content providers are planning ahead by readying their content for IPv6 endpoint hits. Some of the service providers and broadband providers that saw the need started their efforts some time ago, and they're now moving toward infrastructure and service delivery that can and will run on IPv6. Both face a bit of a chicken-and-egg problem: the content providers would move faster if they knew the operators had the services fully baked to deliver IPv6 eyeballs to the content. And the operators would invest more in their IPv6 services and infrastructure deployments if the content the end customer demanded were available and abundant.

Enterprises are, in some way, the swing votes. Objectively, few enterprises have moved toward wide-scale deployments of IPv6. Some have IPv6 pilots; and a few have partial deployments, yet they may hit the v4 allocation cliff harder than either of the other two, aforementioned communities. This will leave enterprises in a costly scurry to rectify the situation at the last minute.

IPv6 deployment in operational networks across the Internet is a work in progress. Transition mechanisms have existed and been deployed for years, including dual stacks and various translation and tunnelling mechanisms. Over time, more hosts and networks are moving to native IPv6 operations. Collectively, we now have several years of experience with both.

The IAB established this plenary topic to provide background and input from those experienced with deployments and issues in order to empower the IETF community to do its part to encourage and facilitate the universal deployment of IPv6.

View from the RIRs

Mark Kosters, Chief Technology Officer, ARIN

RIR Allocations

Currently there are 39 IPv4 "/8" address blocks (2^24, or 16,777,216 addresses) remaining in the IANA free pool that can be allocated by IANA to RIRs. In turn, the five RIRs that together cover the globe assign portions of the /8s to ISPs in their regions according to their local policies and practices. To give a sense of the allocation pace, in December 2004 there were 76 /8s remaining; in December 2005 there were 65; in December 2006 there were 55; in December 2007 there were 42; and in June 2008 there were 39. While the absolute number of allocations per year has remained somewhat constant, the percentage of remaining /8s being allocated is growing steadily each year.

If, as one model predicts (http://www.potaroo.net/tools/ipv4/), the IANA free pool runs dry in December 2011, then IP address blocks that have never been allocated to RIRs will be allocated. That won't mean end users or corporations will no longer be able to get IPv4 public addresses from their ISPs. It means the currently defined IANA free pool allocations will be completed. Rather than receiving this news as Chicken Little would—as an indication that the sky is falling—Mark urged us to look at it like the westward expansion of the United States in the 1800s. At that time it was possible to go out and get land from the government for only a small registration fee—land that had not been previously

Continued on next page

IPv6 Deployment, continued from page 17

farmed. Today, you can still acquire a farm or a large piece of land; it just costs a lot more, and it will have been owned by others—whether or not they used it—thereby giving it a definite market value. And today there exists a complex and robust market for titling, brokering, and owning the land. Likewise, the days of IPv4 address "homesteading" are coming to an end, and many more stringent policies, and, perhaps, markets, will emerge for dealing with IP address use and propriety.

The IPv6 address space, on the other hand, is a 128-bit space, of which IANA has given out very little. The homesteading of IPv6 has barely begun. The RIRs have been fairly active in the past few years in assigning IPv6 address blocks to their Local Internet Registries (LIRs) and ISPs, with RIPE NCC and APNIC being by far the most active. ARIN and APNIC were making similarly sized allocations since 2003—around the 40 to 50 mark per year—until 2007, when ARIN handled over a hundred IPv6 allocations. To date, RIPE NCC has dealt almost half of all of the IPv6 allocations—1,208—which equates to 33,041 /32s. The next closest is APNIC, with 580, accounting for 23,233 /32s, followed by ARIN with 496, LACNIC with 97, and AfriNIC with 47. These allocations are from the RIRs to providers. However, only ARIN, APNIC, and AfriNIC have actually assigned provider-independent (PI) address blocks to end sites directly. The RIPE NCC and LACNIC have not yet assigned PI IPv6 addresses; both have policy proposals in the works for doing so. PI address grants are held in tension because backbone routing for IPv6 has not yet reached the same scale and stability as those present in IPv4 BGP, and the community questions how much route bloat from PIs can be reasonably handled.

> Both [service and broadband providers] face a bit of a chicken-and-egg problem: the content providers would move faster if they knew the operators had the services fully baked to deliver IPv6 eyeballs to the content. And the operators would invest more in their IPv6 services and infrastructure deployments if the content the end customer demanded were available and abundant.

RIR Policy Development

With IPv4 allocation completion on the horizon, the volume of policy work by RIRs on this subject has ballooned. Fourteen policy proposals—almost half of the open proposals—concern IPv4 depletion policies. One of the proposals concerns global policy, which states that as IANA's free allocations of /8s approach exhaustion, every RIR will get a /8. Another proposal deals with liberalizing the transfer policy for moving blocks of addresses between RIRs and LIRs and ISPs and organizations. In the IPv6 policy realm, proposals exist to make it easier to transition to and use IPv6. In the autonomous-system (AS) realm, the RIRs currently allocate a lot of 2-byte AS numbers. These would be exhausted in 2011 as well if the current allocation rate continues. There are 4-byte AS numbers available, but many parties have returned them because the network as a whole does not support them very well (e.g., BGP implementations and OSS systems). We need to encourage our vendors and ISPs to better support 4-byte ASs, because this is another hole—like the IPv4 depletion—looming on the Internet highway.

Building Awareness

In addition to policy and allocations, the RIRs conduct a fair amount of awareness and promotion work for IPv6. The RIRs have taken an aggressive stance on trying to move IPv6 adoption forward. They have issued position statements, delivered awareness-raising and educational workshops, conducted research on IPv4 allocations and IPv6 deployment, provided education and advice for governments and nongovernmental organizations, and hosted IPv6 networks during meetings, such as those of the IETF—all to help drive IPv6 adoption.

"This is an important time of transition, and people need to participate," Mark said in closing. "The registries' policy development process is very bottom-up. So come join the effort."

Lessons Learned from IPv6 Deployment

Alain Durand, Director of Internet Governance and IPv6 Architecture, Comcast

Comcast, a large cable operator in the United States, began its IPv6 deployments several years ago with a two-phase approach. First, it is readying the infrastructure by using IPv6 for management of cable modem devices and network infrastructure. This paves the way for the second phase, wherein Comcast will offer home users IPv6 service to their endpoints. That second phase is now in planning stages, with some lab trials under way.

The cable industry's Data over Cable Service Interface Specification (DOCSIS) model assigns an IP address to each customer's cable modem. This differs from DSL, wherein an individual modem does not have its own IP address. Cable operators thus use a lot of IP addresses and are very involved in allocation policy. Their move to IPv6 was somewhat motivated by the depleting IPv4 address space. They foresaw that in a few years an organization requesting a few IPv4 addresses for Web and mail servers might still easily get them, whereas a request to IANA for a large-scale allocation is likely to be

turned down much sooner. Moving the modems' management interfaces to IPv6 means that they each get a globally unique and routable IP address, thereby avoiding messy address overlaps or NATs.

The company learned important lessons from its initial experiments with IPv6 deployment. The first lesson taught that starting early saves money. For example, including IPv6 early in the DOCSIS 3.0 specifications in early 2005 ensured that products rolling out today include IPv6 at no extra cost. The second lesson was that the network buildout for IPv6 was easier than expected. Comcast started by adding IPv6 addresses on router interfaces—and nothing else. To its surprise, nothing bad happened. So Comcast left the network like that for several months. Next, it enabled IPv6 routing, IS-IS, and BGP. Again, nothing bad happened. Not one outage occurred that was traced back to the IPv6 pieces. This was a critical step in building the confidence of the operations team, proving that IPv6 could run stably in a production dual stack network. Currently, IPv6 exists in both Comcast's access and backbone networks.

For Comcast the problem with IPv6 was not in the networking layers but in the application and services ecosystem around the layers, like their operations and support, management, billing, and third-party systems. Not only do many of those systems lack the required IPv6 support, but getting that support is not high on most vendors' priority lists. Adapting the code doesn't happen overnight. It takes implementation, testing, debugging, deploying, scaling, and stabilizing. It's a linear process that can be accelerated only so much by applying additional resources. "Nine women can't make a baby in one month," quipped Alain, urging application vendors to get moving. Hosted or outsourced services from third-party vendors are in a spot similar to applications, wherein many do not yet support IPv6.

Seeing the IPv4 depletion ahead, the Comcast team is plotting a transition course that includes IPv4 access for quite some time. Once scarcity hits, the team imagines, a wide range of IPv4-only computers, devices, and operating systems will still exist in customers' homes (e.g., Win95, 98, XP, PlayStations, Xboxes, and some consumer devices). Though more and more IPv6-ready equipment is entering homes, customers will not jettison the old IPv4-only equipment; they keep using the older devices. Also, the content on the Internet is almost exclusively IPv4. Recently, thanks to Google, we have some IPv6 content to look at, but not much. IPv4 hosts and content will need reachability for some time yet.

"Following IPv4 completion, if you give all new customers IPv6 only, with no IPv4 support, then the IPv4-only devices can't get out of the home, and new, IPv6-only devices can't get to the predominant IPv4-only content on the Internet."— Alain Durand

Comcast's plan therefore involves a dual-stack core, which Alain says should work well—at least until it's no longer possible to get any more IPv4 addresses. "Following IPv4 completion, if you give all new customers IPv6 only, with no IPv4 support, then the IPv4-only devices can't get out of the home, and new, IPv6-only devices can't get to the predominant IPv4-only content on the Internet." To counter that hitch, one might add the practice of NATing new customers with overlays of private addresses from the RFC 1918 blocks. This brings undesirable results, though, such as NATs piled one on top of the other, with multiple overlapping addresses to customers. While not impossible, NATing creates a lot of complexity—especially in management—and it makes the troubleshooting of customer issues much more difficult. In addition, traffic-engineering gyrations, such as source-based routing, might be necessary to handle the overlaps. Outages are usually linked to complexity in the network, so this increased complexity in the network is less than desirable.

Photo by Peter Löthberg

The Comcast team is looking at using tunnels and provisioning IPv6 to customers' home gateways. This would enable Comcast to offer both IPv6 and IPv4 services to endpoints behind those gateways. IPv6 would run natively from hosts to IPv6 targets. IPv4 is delivered by first having the gateway speak IPv4 internally, and second, by transporting the IPv4 packet over the IPv6 infrastructure network via a tunnel. The IPv4 packet would be detunnelled inside the ISP network at a large v4-to-v4 NAT box, a CGN. That box translates the IPv4 packet to a globally routable IPv4 address and sends it off to the IPv4 target. Alain refers to this as a dual-stack-lite service. (Work on this is occurring in the Softwire working group.) The advantage is that the infrastructure itself requires only IPv6 addresses for

Continued on next page

IPv6 Deployment, continued from page 19

operated devices while allowing either IPv4-only or dual-stack hosts to reach IPv4-only content on the Internet. "This dual-stack-lite concept makes IPv6 services incrementally deployable," claims Alain. "You don't have to wait for the rest of the Internet—the content and applications—to move to IPv6 in order to roll out the service. You can deploy IPv6 in your own world and get some immediate benefit out of it." The IETF's most successful protocols over the past 15 years have been incrementally deployable, and that is Alain's goal for networks transitioning to IPv6.

A Step-by-Step Transition Plan

Shin Miyakawa, Director of IP Technology Development at NTT Communications

While sharing much of Alain's view of the need for incremental deployment of IPv6 in the face of depleted IPv4 space, Shin cautioned against accepting a common misunderstanding about carrier-grade NAT. "Please do not get comfortable with the idea that a carrier-grade NAT will 'solve' our problems," he warned. "Do not believe that if we have a CGN, we need not move to IPv6." On the contrary: according to Shin, the CGN concept has significant and serious restrictions as well as implications for the Internet's end-to-end design principle. "The last thing we would want is the existence of a CGN to slow IPv6 adoption," he said. Instead, Shin suggests employing CGN as a backward-compatible mechanism for the purpose of speeding along IPv6 adoption. "It will do so by ensuring an incremental deploy-ability mechanism that does not exclude or alienate v4 hosts [nor v4 content] once deployed," he said. Once such a model is supported, network operators will be able to see a clear transition path to IPv6 that contains neither product cliffs nor flagship upgrades.

Shin warned that the dual-stack lite's CGN component may necessarily limit each customer's number of concurrent sessions. TCP and UDP allow for only 65,535 total source ports per single IP address. As IPv4 addresses become scarce, a single public IPv4 address must support many CPE routers, each with several hosts behind it. How many hosts can be supported by one IPv4 address depends on the number of concurrent TCP and UDP connections being made by the average host.

To illustrate that point, Shin offered the example of browser connections to a Google Maps display of San Francisco International Airport, an application that employs AJAX technology. One characteristic of AJAX is that it simultaneously makes many discreet HTTP connections from the host to the server in order to pull down different small "pieces" of the content all at the same time (see Figure 1). As Shin demonstrated, if the session count is limited to 30 concurrent connections, the map loads appropriately. If it is limited to 20, one block of the picture is missing. If it is limited to 15, only 35 percent of the picture's blocks are successfully received. At a limit of 5 connections, the browser throws an error message. If each user needs only one such application connection at a time, then one IPv4 address would serve approximately 2,100 hosts. This 30-concurrent-connections number is for Google Maps only. The NTT Communications research team has observed iTunes opening 230–270, Amazon grabbing 90, YouTube pulling 90, and OCN's Photo Friend consuming 170–200.

If one estimates an average of 500 open connections from a host—allowing for no safety buffer—one could infer for a CGN deployment a user-to-IPv4-address ratio of about 130:1. An 8:1 ratio would allow for a worst case, of about 8,200 simultaneous per-user connections. And for a case where multiple computers are all connecting from the same customer premise, a 20:1 ratio would allow for a worst case of about 3,300. What is the right ratio to use?

After warning of the CGN issues, Shin proceeded to offer a graphical, time-lapsed, step-by-step progression of how an operator might transition to an IPv6 network by using a dual-stack-lite service architecture. This transition plan offers incremental deployment and IPv4 backward compatibility. The operator starts by enabling IPv6 on its peering routers and backbone. Then the CGN element/function is added, which logically sits north of the operator's access concentrator in the POP. The operator may then place a private IPv4 address on the CPE router's provider-edge-facing (PE) interface. This IP will be NAT'ed by the CGN sitting in the operator infrastructure. This step addresses the decreasing availability of routable IPv4 addresses.

Step two introduces IPv6 on the customer side of the CPE, a client-based Softwire tunnelling solution from the customer's hosts, and a tunnelling con-

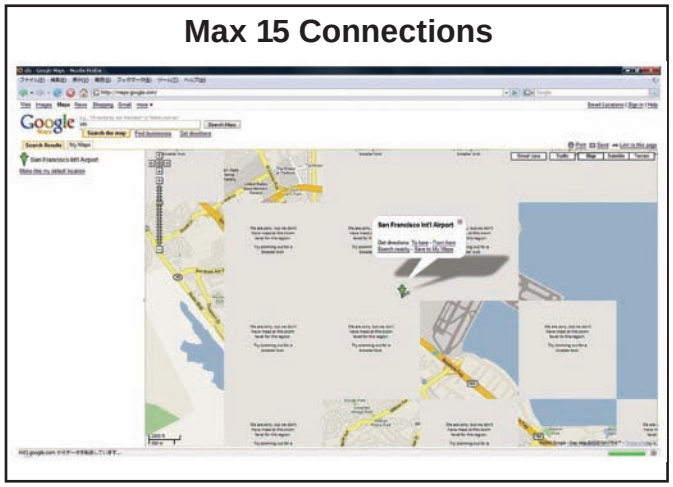

Figure 1: When concurrent connections are reduced, an image may not be able to load appropriately.

centrator in the POP. This solution encapsulates IPv6 over IPv4 by using L2TP as the encapsulating protocol. A client-side software component sits on the customer's host, and a networking device sitting north of the CGN terminates the L2TP tunnel in the operator's network. At this point, the deployed network supports dual-stack customer hosts and delivers to those hosts connectivity to both v4 and v6 content.

As new customer deployments occur, Step 3 involves moving the Softwire v6-in-L2TP-over-v4 client function off the hosts and into the CPE router. This allows the clients sitting on the CPE's private side to contain any or all of IPv4-only, IPv6-only, or IPv4/IPv6 dual-stack hosts while still providing access to either native v4 or native v6 Internet content.

Step 4 upgrades the provider edge (PE) access concentrator to be IPv4/IPv6 dual stack. Note that the Softwire tunnelling mechanism is required only until such time as both the PE access concentrator and the CPE support v4/v6 dual stack. Once both can run IPv6 natively, the operator has no further need for the tunnelling mechanism. This captures the attractiveness of the proposed transition plan: any of the key pieces—PE access concentrator, CPE router, or end-host system—can be combined in any combination of their v4-only or dual-stack support, and still, the operator will be able to deliver a v4/v6 service. This provides true incremental deployability for operators, saving them from costly forklift upgrades (for end hosts, CPE gear, or access concentrators) and providing them with a grow-as-you-go transition. The CGN remains in the network until the provider was prepared to end the life of the IPv4 service.

NTT is considering putting in place by spring 2010 a service as described earlier—before the IPv4 address completion. NTT is working with other ISPs on the same architecture, and such was discussed extensively at a recent Internet Area interim meeting on the v4-to-v6 topic. In the assistance of various enterprises, application service providers, schools, and governmental agencies with dual-stack deployments, Shin notes, externally facing systems (e.g., Web, e-mail, and DNS) must be IPv6 capable first; then the customers' other, internal systems follow.

In addition to the aforementioned dual-stack proposal, the NTT group in Japan has already deployed a commercialized IPv6 service to over 5 million customers. This solution uses a highly controlled (i.e., walled garden), all-IPv6 network, including endpoints, to deliver specialized value-based services to end users.

What can the IETF do to hasten IPv6 adoption? According to Shin, first is an additional IPv4 private address space allocation for carrier/operator access network equipment that sits in the IETF's internal infrastructure devices, behind a CGN. Next would be defining a simple security scheme for IPv6 (see page 23). Implementations of IPv6 DNS deployments need wider dispersion. Multiprotocol label switching (MPLS) with IPv6 support must exist in the network devices. Though commercially available IPv6 supporting firewalls have been available since NetScreen (now Juniper) released its in the early 2000s, other security devices—like IDP, IPS, and Web filtering—are still awaited, as are load balancers.

> "Please do not get comfortable with the idea that a carrier-grade NAT will 'solve' our problems. Do not believe that if we have a CGN, we need not move to IPv6." — Shin Miyakawa

A Simple Call to Action

Lorenzo Colitti, Network Engineer and Researcher, Google

Google has one of the only large content sites deploying IPv6 today. Explaining why the company made the investment, Lorenzo stated, "When the day comes that users have only IPv6, they will need to be able to get to Google. That's the long-term view." He identified several shorter-term reasons for the deployment, including lower latency and packet loss, AJAX applications breaking behind excessive NAT (as described in detail by Shin), and the irritation of NAT traversal mechanisms, such as STUN. The year 2011 is when Google foresees the RIR IPv4 address pool exhaustion to their ISPs and PIs, and they point to IPv6 as the only sensible solution. Starting as early as possible will ensure deployment is ready in time, if not well before. The Google IPv6 effort began as a side (20 percent) project on the part of a few employees. Once others caught wind of ipv6.google.com, they jumped in alongside Lorenzo and team, and now Gmail and News are IPv6 enabled too. They built a pilot network and ran the infrastructure at an IPv6 conference, proving that it worked.

As more and more people wanted to get IPv6 running for their applications, the team grew and they scaled up the pilot. Even though it was a pilot, they dispelled the notion that it was experimental, which could lead to skimping on quality. The key lay in incremental growth. A pilot IPv6 network doesn't have to be as scalable or as capable as an IPv4 network on day one, but it does need the same types of production services as soon as possible, as are found in IPv4 networks, including monitoring, support, documentation, written quality standards, and audits.

Start small, and make steady prog-

Continued on next page

IPv6 Deployment, continued from page 21

ress. In addition, Lorenzo urged that, whenever possible, one should design the IPv6 deployment to be as similar as possible to one's IPv4 network. "Every time I made a design decision that wasn't the same as IPv4, it turned out to bite me down the road," he said.

Lorenzo counselled against the use of tunnels in interdomain routing, citing increased latency and difficulty in debugging. Also, today most IPv6 operators are indiscriminately giving transit to any other IPv6 operator, which slows convergence, increases round-trip times, and creates partial visibility for yours and others' networks. Some transit providers have incomplete IPv6 routing tables that cause black holes for those routes. The solution is direct peering with quality IPv6 networks, yielding direct contacts by which to address and improve routing and connectivity issues. "Don't assume the current IPv6 Internet works," he said. "On the other hand, get involved. The only way to accelerate the progress is to accelerate the involvement."

Lorenzo's IPv6 product feature/capability wish list included MPLS traffic engineering, mature load balancing, IPv4-style multihoming, and hardware processing for both 6to4 and extension header filtering. Lorenzo notes that today they have to drop at the edge of their network every IPv6 packet that is not clearly TCP, UDP or ICMP due to a lack of hardware filtering for extension headers. IPv6-style multiple-address multihoming has not worked well. Failovers break TCP connections. Lorenzo says HIP and SHIM6 do support failovers, but then new connections see timeouts; both lack load balancing or traffic engineering from the network side, which are features present in v4. Without those features it's tough to make large-scale applications deployable. Lorenzo also says multihoming ought to be allowed using /48 prefixes. He expressed the need for /127's on point-to-point links (currently prohibited by RFC 4291) to help avoid denial-of-service attacks, and VRRP for v6 because neighbor unreachability is not fast enough.

> Start small, and make steady progress . . . . [W]henever possible, one should design the IPv6 deployment to be as similar as possible to one's IPv4 network.

NAT-PT is another item big on Lorenzo's list because, he says, it will lead to an all-IPv6 network more quickly than dual-stack lite will. He argues that once IPv4 address allocations slow significantly, parts of the network will be able to serve hosts an IPv6 address only. Of course, all IPv4 content will not be served immediately on v6, so v6-only hosts will, for some time, still need to reach v4-only sites. Using something like NAT-PT transforms the content deployment problem into an application porting problem. However, NAT-PT is deprecated in RFC 4966. "All of the drawbacks of RFC 4966 are really drawbacks of NAT, and they are all present in v4 NAT too," said Lorenzo. Like Shin, Lorenzo doesn't like NATs—or the NAT mechanisms in NAT-PT. However, being an IPv6 champion, he feels that incremental transition mechanisms are key to hastened adoption. "If we want to reclaim the end-to-end principle, then we must do it with IPv6." And in order to do that, Lorenzo called for a bare-bones standard that addresses a v6 host connecting to a v4 server. This would require a DNS application layer gateway. The focus should be on network-based translation scenarios, not on a host-based solution. "[End-to-end connectivity] can happen in v6 but not while we still have v4 around." Lorenzo thinks NAT-PT's presence will lower the value of IPv4 on hosts, and operators and users will opt for IPv6 hosts, with NAT-PT fronting the v4 applications. As soon as the content is v6, the NAT-PT function will be removed, and it's pure v6 from there.

First Impressions Are Everything

Stuart Cheshire, Wizard without Portfolio at Apple

Apple chose IPv6 link local addressing for the AirPort Express wireless base station because of the auto configuration features that require the user to do nothing to get LAN-based printing, streaming music, and network administration running. The solution has been more reliable and works better than IPv4 in a single-network home. The Internet, as Stuart pointed out, is a different story. Five elements interact to make an IPv6 solution work for the customer across the Internet: the operating system, application software (such as a Web browser), home network, ISP, and content (such as Web sites).

Without any one piece, the solution fails. Content is the key. Without IPv6 content, ISPs have no incentive to carry IPv6. As Stuart said, if a customer cannot get IPv6 from the ISP, then there's no reason for content accessible via IPv6. If the browser is not IPv6 enabled, then the customer will not be able to access IPv6 content and therefore will not purchase IPv6 connectivity. If the OS is not v6 ready, then apps can't be either. Incentives must exist for all of the players. Apple's incentive was the so-called coolness factor of IPv6, so now Apple's OS, Apple's browser, and most of Apple's networking apps support it. "Now we must find compelling incentives for the other three pieces to support v6," encouraged Stuart. "Or, at the very least, make sure there are no disincentives."

To drive home the point, Stuart described an application-level issue that Apple faced with regard to its Safari

IPv6 Panel Takes Questions from the Floor

What are the biggest operational issues percolating up from the network administration staff?

Alain Durand: IPv6 addresses are difficult to type. I often ask people to write down an IPv6 address and read it back over the phone. Of all the attempts, only one person has succeeded in repeating the address correctly. So we must avoid IPv6 addresses as much as possible, and we must be more ardent DNS users.

Shin Miyakawa: It's not technical, but the lack of education of the customer-facing support staff. We don't want the front-line staff answering support calls by saying, "What is IPv6? I don't know anything about it." Getting IPv6 into all the training manuals and achieving a basic level of familiarity across all levels of staff are the hardest things NTT has faced.

Lorenzo Colitti: Failure to follow familiar IPv4 design principles. There is muscle memory there, and so people just naturally try to do an operation the v4 way, but then something breaks, or it doesn't work. They don't know why, and they don't know where to find the answer. Training everyone that something must be done differently in this situation is a challenge and takes time. Putting safeguards in place, like deployment templates, helps keep people from such mistakes.

Are you able to get the equipment you need? What are the equipment gaps?

Alain: The back-office applications are the most difficult.

Mark Kosters: The RIRs are putting our money where our mouths are by making our services available on IPv6. Some of the middle boxes such as load balancers are not really ready yet with their IPv6 support. Also, their technical support and best practices are not mature, so we do not get as much consultative assistance as we would expect.

Are you driving IPv6 to your customers, or are your customers driving you to IPv6?

Shin: We are pushing customers to use IPv6. We need IPv6 because we have a lack of resources from which to offer customers advanced services.

Alain: Many customers want to access their e-mail, browse the Web, and download a video from YouTube. I don't think they care whether this gets accomplished over IPv4, IPv6, or avian carrier [RFC 1149]. What matters is that we can keep the services of the Internet growing regardless of whether we have IPv4 or are out of v4 and need IPv6.

Mark: The research is pretty conclusive that people see that IPv4 works now, and so they feel no need for IPv6. They see IPv6 as a capital cost for them. The question arises: Where does the money come from that will help us solve a problem that we don't have right now?

Gregory Lebovitz: It appears unanimous that we are driving IPv6, not the customers, because we need to produce new services that customers do want, and we do not have enough IPv4 addresses going forward to accomplish the task. We need IPv6 as an enabler. If customers are not begging us for IPv6, then the stakes are very high for us to make its presence very transparent to users—or risk its rejection. It has to be invisible and usable by the "grandmother living in the countryside," as Shin says.

NATs in the middle can break things and create walled gardens. Is it possible to demand that DOCSIS put NATs at the ends?

Alain: DOCSIS is layer 2, not layer 3. We are not talking about centralized carrier NAT. The question is, Where is this divide in the network? I agree that this is bad and that it could provide a single point of failure and that NATs need to get close to the customer. I suggest talking to the application developers.

Is it possible to do some analysis to figure out how much it will cost an end user to have a public IPv4 address? Could this be an incentive for using IPv6?

Shin: NTT is currently that kind of research.

Mark: There is a new policy proposal that is related to IPv4 address transfers. One aspect of transfer is market-based transfer, which currently is not allowed by RIR policy. ARIN asked a lawyer to look into this, and the answer is that an open market would be best.

What is the state of readiness of IPv6 DHCP [Dynamic Host Configuration Protocol]?

Alain: The DHCP community has worked quite a bit with vendors recently, and the results are quite promising. (Editor's note: See articles on IPv6 DHCP back-offs in recent editions of the *IETF Journal*).

Applications developers are waiting for the peer-to-peer readiness of IPv6. Do you see inbound connectivity to the home as compelling-enough motivation?

Stuart: I agree that inbound connectivity is the only compelling reason to use IPv6. Everything else is currently also available on IPv4. So, the only reason is that it gives me unhindered peer-to-peer connectivity.

IPv6 Deployment, continued from page 23

Web browser and certain performance delays associated with dual-stack clients. When a dual-stack-capable Web browser connects to the Internet, it first conducts an AAAA record lookup, then it does an A (Address) record lookup and tries an IPv6 connection. If that fails, it tries an IPv4 connection. Everything works as long as this series of events occurs quickly, including any failures. Unfortunately, lack of AAAA responses and IPv6 connection failures are common.

Apple faced an issue regarding a big-name Web site. Customers complained that the site was really slow. Upon investigation, Apple discovered that the site's DNS servers did not send valid responses to AAAA queries. Either such sites do not respond to an AAAA query, or they send back a mangled, malformed response. The connection sequence blocks are waiting for an answer that isn't coming. The user perceives this as application slowness, which started occurring when the user's computer first got an IPv6 address.

The root issue is the getaddrinfo() API, which blocks waiting for an IPv6 address query that may never be answered. Because the problems first appeared when the ISP started offering IPv6 service, the issue landed in the ISPs' laps. This is bad for the industry because both the customer and the ISP get a bad impression about IPv6. These thorny disincentives must be removed if we are to hasten deployment.

Apple's implementation now disassociates the IPv4 and IPv6 tracks. The tracks perform the AAAA and A queries in parallel, and they try the IPv4 and IPv6 http connections in parallel, taking the first response and resetting the second. Failures or hangs at either step, on either track, no longer hang the user experience.

Five elements interact to make an IPv6 solution work for the customer across the Internet: the operating system, application software (such as a Web browser), home network, ISP, and content (such as Web sites).

The getaddrinfo() API was the problem because it exposes addresses to the application when it shouldn't. Applications don't need to know about Ethernet addresses, and they don't need to map IP addresses to Ethernet addresses; the kernel ought to do that for them with ARP. Likewise, an app should not be involved in mapping DNS names to addresses. The app should just tell the system, "Here are a name and an application/service. Please connect me to it." Apple uses a connect-by-name API so that applications don't even have to know or care whether the underlying connection is IPv4 or IPv6. Both Java and Windows have similar APIs.

The moral, according to Stuart, is that as we move toward IPv6, when you build an app, avoid getaddrinfo() and such APIs. Instead, use concurrency and asynchrony. Yes, these new mechanisms will send extra packets and initially open more connections than will actually be used. The first one that succeeds will proceed, while the others will be reset. This is the right design decision for IPv6. "We are trading off a few extra packets and connections to vastly improve the user experience," said Stuart. "And that's the point: let's make sure to remove the barriers to transition that might otherwise make people regret their first tentative steps into IPv6 deployment and use."

Related Links

Introduction Email from the IAB

http://www.ietf.org/mail-archive/web/ietf/current/msg52686.html

IAB's follow up email

http://www.ietf.org/mail-archive/web/ietf-announce/current/msg05138.html

Audio stream archive

http://limestone.uoregon.edu/ftp/pub/videolab/media/ietf72/

Select "ietf72-ch3-wed-plenary.mp3"

The IPv6 panel starts at time 01:13:00 on the stream.

Q&A (only 1/3 appears in the article) at 02:29:00

Presentation archive

http://www.ietf.org/proceedings/08jul/plenaryw.html

http://www.nttv6.jp/~miyakawa/IETF72/

http://www.stuartcheshire.org/IETF72/

One hypothetical model of an address allocation timeline

http://www.potaroo.net/tools/ipv4/

Related drafts

draft-nishitani-cgn-00.txt

draft-shirasaki-shared-adrs-00.txt

draft-ietf-v6ops-cpe-simple-security-02

IPv6 Transition at IETF 72

By Geoff Huston

The developmental work of the Internet Engineering Task Force (IETF) on IPv6 has, from the outset, included the study of the particular issues associated with transition to IPv6. The first effort to explore the transition space was at IETF 29 in March 1994, and it was termed TACIT, an acronym of Transition and Coexistence including Testing. While it was admittedly a forced acronym, it was illustrative of the IETF's desire to include consideration of transition issues as part of the design of IPv6 itself. The underlying consideration here is a study of how a diverse amalgam of applications, hosts, and network elements that collectively make up the Internet and the related collection of enterprise networks can be upgraded, selectively augmented, or replaced in order to support IPv6 and, ultimately, to deprecate all further use of IPv4 while at the same time preserving all of the essential, "any-to-any" end-to-end property of the Internet Protocol (IP) through the transition. From the TACIT birds-of-a-feather sessions, the baton was then passed to the NGTRANS working group in July 1995 at IETF 33. This working group was active until IETF 55 in mid-2002, when the baton was again passed—this time to the V6OPS working group, which met first in early 2003 at IETF 56. The study of transition to IPv6 has now broadened in scope, and today a number of IETF working groups are examining aspects of transition to IPv6, including the SOFTWIRE, BEHAVE, and INTAREA working groups, in addition to V6OPS.

Given that this study now encompasses a period of 14 years, what exactly are the issues with respect to the transition to IPv6, and why is this transition taking such a long time?

Backwards Compatibility

This is not the only transition we've faced at the basic level of protocol infrastructure, and the conventional approach is to make the changes "backward compatible" Backward compatibility can take many forms, but typically involves some form of initial negotiation between communicating parties that establishes whether both parties are capable of recognizing and using some extension or new attribute. For example, the Border Gateway Protocol (BGP) is in the process of transitioning from 16-bit Autonomous System (AS) numbers to 32-bit AS numbers. This BGP transition uses a combination of translation and tunnelling that allows a BGP speaker configured to use the longer AS numbers to be backward compatible with the existing installed base of BGP that uses the 16-bit AS number format. The protocol specification of BGP includes an initial capability negotiation when BGP is first started up, allowing a "new" BGP speaker to establish whether its BGP neighbour is also capable of supporting longer format AS numbers or not. As a result, upgraded versions of BGP can coexist with older versions of BGP, so that the overall transition of BGP to use 32 bit AS numbers can be undertaken on a piecemeal basis. This particular backward-compatible translation technique relies on a combination of capability negotiation and the properties of hop-by-hop interpretation of tokens, where AS-number values are interpreted in a strictly local context.

IP is an end-to-end protocol, as distinct from a hop-by-hop protocol, and an IP packet's destination address needs to have meaningful context at all points in the network. IP itself is a connectionless datagram protocol, without any form of capability negotiation. Its also a very challenging exercise to equip a network with intermediaries that attempt to change the IP packet header mid-flight. This implies that the use of translation and substitution to create backward compatibility has limited applicability in the context of IP itself.

A Classical Transition

The original approach to IPv6 transition could be termed a "classical" view of transition. Because IPv6 is not a backward-compatible augmentation of IPv4, it is not possible to deploy new hosts and network infrastructure with support for only IPv6 and have these networks, devices, and applications exchange IP packets with their IPv4 counterparts. An application that is equipped with IPv6 requires its host to have IPv6 support in its protocol stack, and for the host to be able to communicate, the network is required to have IPv6 support. And if an application wishes to communicate with another application, all the networks on the path between the two hosts also must be configured to support the transmission of IPv6 packets. In other words, a "complete" deployment of IPv6 requires all applications, hosts, and network infrastructure and middleware to be aware of IPv6 and explicitly configured to handle IPv6 packets. In this classical form of transition, the major constraint is to avoid any flag day, or any other form of synchronized or orchestrated common activity across the entire

The study of transition to IPv6 has now broadened in scope, and today a number of IETF working groups are examining aspects of transition to IPv6, including the SOFTWIRE, BEHAVE, and INTAREA working groups, in addition to V6OPS.

Continued on next page

IPv6 Transition at IETF72, continued from page 25

network. Individual elements of the network should be able to undertake their part of the transition without requiring any action to be performed on any other element. The transition should be a piecemeal activity. This classical approach, in general terms, assumes that each application, host device, and network element is able to make an independent decision as to when to enable support for IPv6. To preserve connectivity of the network as a whole, then, as and when each network element or end device is configured with support for IPv6, it would not "cut over" and remove all IPv4 support from the device, but, instead, it would support the operation of both IPv4 and IPv6 for an extended period. This was termed the dual-stack transition approach. This mode of progressive shift of the elements of the Internet to a dual-stack operation would continue for as long as there were essential components of the overall environment—from applications to Internet infrastructure—that support only IPv4. Only when the entire connectivity domain was supporting comprehensive dual-stack operation would it be possible to deprecate IPv4 from the network and remove all support for this protocol (see Figure 1).

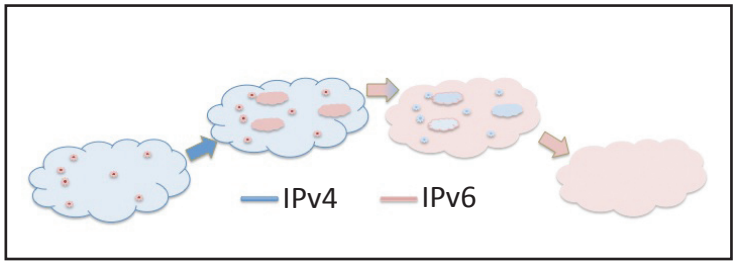

Figure 1. The Progressive Stages of IPv6 Transition.

The issue with this approach to IPv6 transition was that it relied on a strong mix of altruism, common purpose, and shared motivation, as well as a high level of technical capability from everyone: from suppliers and vendors through to network operators and even end users. For early adopters of IPv6, whether it was application designers, suppliers of host operating systems or routers, or network operators and system administrators, the investment in dual-stack capability in their area of responsibility would generate the greatest extent of resultant benefit only when the transitional dual-stack phase was complete. In other words, there was no immediate reward for those early adopters of IPv6, and late adopters did not experience any detrimental side effects, because the full benefits of an outcome of IPv6 adoption would be realized only once the entire environment adopted IPv6 in a dual-stack configuration with IPv4, at which point IPv4 could be deprecated from the operational network.

This approach assumes that all parties are equally motivated to undertake this transition, and that each party will do so as quickly as possible. It also assumes that all applications, all connected devices, and all components of the network's infrastructure are capable of being configured to operate in dual-stack mode. Perhaps those assumptions may have been feasible in practical terms if IPv6 had been in a position to offer very significant cost, performance, or functionality improvements over IPv4. In such a case the superior characteristics of the new technology would have propelled the transition process. However, any such major relative improvement in performance, cost, and utility is not the case in a comparison of IPv6 with IPv4, because IPv6 represents only a marginal change in the underlying network design. Following a further decade of incremental refinement in both IPv4 and IPv6 we have the current situation where, apart from the larger address fields in the packet header, there is no significant relative change in IPv6 from a performance or benefit perspective. In addition, the Internet itself is now so much larger and so much more diverse that commonality of purpose is difficult to sustain. These days, altruism often takes a backseat to business interests as the Internet now operates as a collection of quite conventional business enterprises. Indeed, since the bursting of the Internet bubble at the start of this decade, this sector of business is relatively conservative as well, and far greater emphasis is placed on securing immediate returns on invested capital over and above the undertaking of longer-term investments and with less-certain outcomes. This implies that any such commonality of purpose and a vision of a longer-term outcome is extremely challenging to sustain in the face of shorter-term considerations.

The combination of these factors creates a situation that has been incapable of sustaining the operation of this "classical" transition process. So the IETF was motivated to look at transition in slightly different terms—to see whether this approach could be refined to offer some more-immediate benefits to early adopters and not to stall the entire process while awaiting completion of the late adopters of dual stack.

Transition with Incremental Outcomes: Tunnelling

The initial refinement to this original transition model, explored in the NGTRANS working group, was intended to allow various IPv6-only and dual-stack applications to support IPv6 from the outset, so that any benefits related to IPv6 could be realized immediately and not be forced to await the actions of the slowest adopters to also make their moves. The motivation involved the restoration of simple application programming interfaces for applications, the restoration of coherent end-to-end packet delivery in an IPv6 network, and the

benefits that this clear and simple application architecture offers to applications that operate in an over-the-top mode. Such an end-to-end packet transport environment offers strong end-to-end channel security as well as restoration of the uniform binding of IP address to end-point identity in the IP architecture.

The objective of the attempt to operate in an end-to-end IPv6-only mode over a largely IPv4 substrate network led to the development of a number of approaches to IPv6 transition that relied on tunnelling techniques, wherein IPv6 packets are encapsulated in an IPv4 packet wrapper, allowing these IPv6 "islands" to treat the IPv4 network as a form of transmission media, or a non-broadcast multicast network. That led to the development of the general technique of carrying IPv6 packets in IPv4 by treating IPv6 as an IPv4 protocol—namely protocol 41 (see Figure 2).

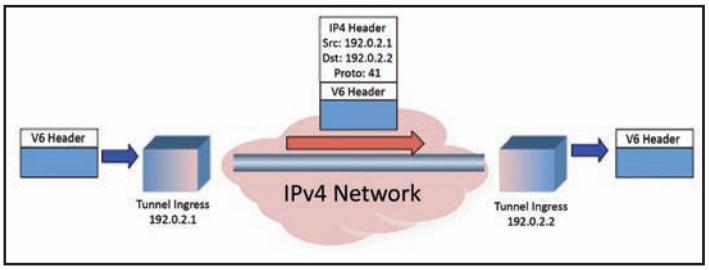

Figure 2. IPv6 in IPv4 Tunnelling.

The general characterization of this approach to this form of dual-stack transition was to allow the initial "islands" of IPv6 adoption to connect to each other via these tunnels, essentially creating an IPv6 connected network from the outset. As more of the infrastructure adopted the same form of dual-stack support, these islands would start to directly interconnect, making the islands larger and the tunnelled gaps shorter. As these gaps shrink to the point of general dual-stack support, it may be an option to then tunnel the remaining IPv4 traffic over IPv6, but perhaps that's getting well ahead of ourselves right now (see Figure 3).

While the motive and logic for the use of tunnels in this transition scenario are certainly sound, the overhead here is that tunnels normally require explicit configuration of both ends of the tunnel, and any form of tunnel topology that attempts a fully meshed interconnection of the IPv6 islands runs into an N-squared scaling problem in tunnel configuration almost immediately.

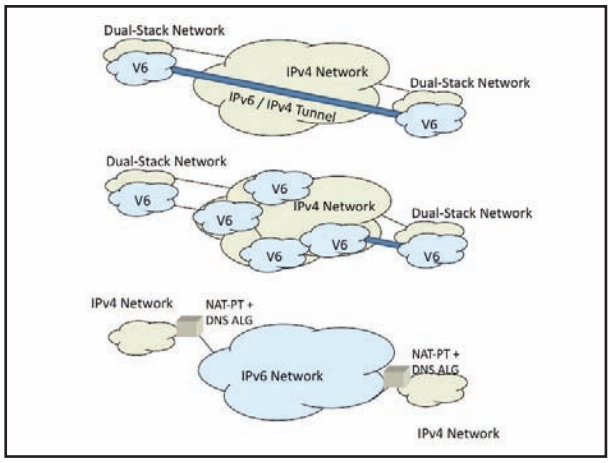

Figure 3. Transition Using Tunnels.

This, in turn, has led to exploration of approaches that supported the concept of fully meshed tunnels—but with an extremely simple single end configuration. This is achieved by associating an IPv4 tunnel endpoint in an endpoint IPv6 address. When such a packet is passed to a tunnel ingress, the IPv4 tunnel egress address is defined by the original IPv6 destination address, so that the tunnel does not have to be explicitly configured at both ends. One of these is the 6to4 technique, which generates an IPv6 48-bit prefix by prepending 2002::/16 to the front of the 32-bit IPv4 address. This allows a dual-stack gateway to double as an IPv6 tunnel egress, serving a local network of IPv6 hosts with tunnel services. Each 6to4 gateway, or 6to4 individual host, needs only to configure its end of the tunnel. All IPv6 packets between 6to4 sites are passed directly from 6to4 gateway to gateway. To complete the picture, each local 6to4 network needs to provide 6to4 gateway service for IPv6 packets from non-6to4 IPv6 networks (see Figure 4).

A related form of embedding IPv4 in IPv6 addresses to aid in autotunnelling is ISATAP, the Intra-Site Automatic Addressing Protocol, which embeds the IPv4 address in the interface identifier field of the IPv6 address to support a local scope automated IPv6 over IPv4 tunnelling approach. These approaches can be combined, so that an enterprise can construct an IPv6 network with a single infrastructure gateway that creates the prefix and tunnels over the wide area network by using 6to4 while tunnelling over the local area network by using ISATAP.

The shortcoming of the 6to4 approach is that it assumes a general avail-

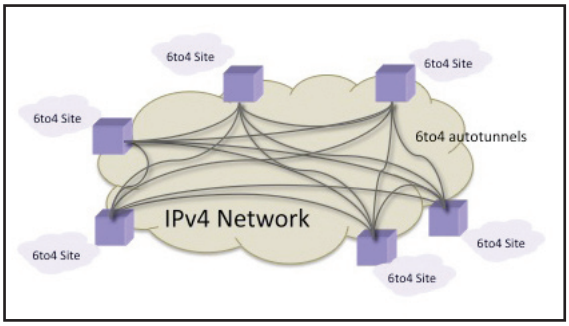

Figure 4. 6to4 Tunnelling.

Continued on next page

IPv6 Transition at IETF72, continued from page 27

ability and use of public IPv4 addresses. A single host behind a network-address-translation (NAT) gateway cannot use this approach given that the implicit IPv4 tunnel endpoint is drawn from a private address pool and is therefore not visible outside the IPv4 private address scope. It also requires firewalls to be aware of protocol 41 and apply the IPv6 filter rules to the inner IPv6 packet.

The Teredo approach addresses both of these concerns by using explicit support for NAT traversal, and embedding the IPv6 packet inside an IPv4 UDP transport session rather than as an IP transport. Teredo takes a relatively conventional approach to NAT traversal, using a simplified version of the STUN active probing approach to determine the type of NAT, and uses concepts of "clients," "servers," and "relays." A Teredo client is a dual-stack host that is located in the IPv4 world, possibly behind a NAT. A Teredo server is an address and reachability broker that is located in the public IPv4 Internet. A Teredo relay is a Teredo tunnel endpoint that connects Teredo clients to the IPv6 network.

The tunnelling protocol used by Teredo is not the simple IPv6-in-IPv4 protocol 41 used by 6to4. IPv4 NATs are sensitive to the transport protocol and generally pass only Transmission Control Protocol (TCP) and User Datagram Protocol (UDP) transport protocols. In Teredo's case, the tunnelling is UDP, so all IPv6 Teredo packets are composed of an IPv4 packet header and a UDP transport header, followed by the IPv6 packet as the tunnel payload. Teredo represents a different set of design trade-offs compared with 6to4. In its desire to be useful in an environment that includes NATs in the IPv4 path, Teredo is a per-host connectivity approach, compared with 6to4's approach, which can support both individual hosts and end sites within the same technology. Also, Teredo is now a host centric multiparty rendezvous application, and Teredo clients require the existence of dual-stack Teredo servers and relays that exist in both the public IPv4 and IPv6 networks. From Teredo's host-centric perspective, it could be said that Teredo is more a connectivity tool than a service solution (see Figure 5).

The common feature of all of these transition approaches is the use of tunnels. Tunnels are extremely convenient in terms of their ability to interconnect diverse islands of IPv6 without requiring any change to the intervening IPv4 infrastructure. However, tunnels are not without their attendant problems. Tunnels can be fragile, unstable, and challenging to diagnose. The issue of Internet Control Message Protocol (ICMP) treatment within tunnels is a good example, where a return ICMP error notice is sent not to the original source host, as intended, but to the tunnel ingress point that is the source address of the outer tunnel packet. The inner payload, which contains the initial fragment of the original packet, also includes the tunnel header. The critical point here is the interplay between end-to-end signalling and Maximum Transmission Unit (MTU) discovery. Where there is a tunnel MTU mismatch coupled with an ICMP handling problem, the situation often manifests itself as a TCP "hang", where the initial SYN handshake succeeds, but the first large data packet is never transmitted. A typical dual-stack implementation will lock into IPv6 or IPv4 at the point of completion of the initial TCP handshake completion, and the data payload problem then causes the user's application to hang. The name to protocol family association is now locked into the user's cache, so that resetting the connection and forcing the application to use IPv4 rather than IPv6 is invariably beyond the user's direct control.

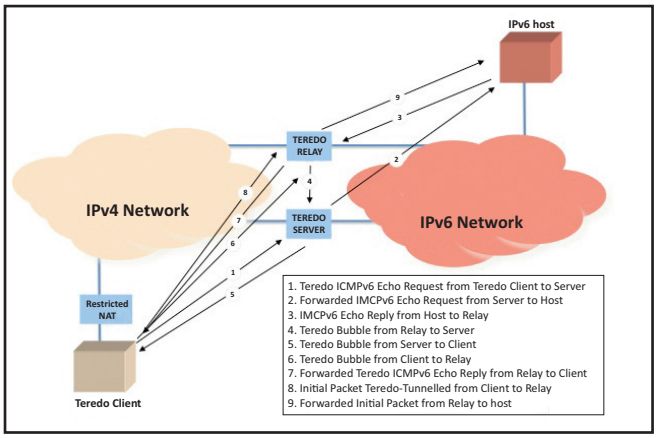

Figure 5. An Example of a Teredo Rendezvous.

So, is it possible to avoid tunnels and still achieve incremental outcomes for early adopters of dual stack? Behind all of the transition scenarios so far lies the assumption that IPv4 and IPv6 support distinct universes of connectivity. However, both protocols present much the same set of functions to the upper-level transport protocols, and the header fields of the protocol are similar. Just how bad is this backward incompatibility of IPv6 with respect to IPv4? Is it completely impossible for an IPv4-only host to initiate, maintain, and close a conversation with an IPv6-only host and vice versa? If one allowed various forms of intermediaries, including protocol-translating NATs and various permutations of Domain Name System (DNS) servers, is this still impossible? Probably not impossible, but it would go well beyond the conventional mode of packet protocol header manipulation and would call upon protocol header translation, cross-protocol NAT bindings, DNS manipulation, and various forms of application level gateways.

An approach to this form of translation was described in RFC2766, "Network Address Translation—Protocol Translation (NAT-PT)." The approach creates a number of security vulnerabili-

ties and appears to operate with a high level of assumption about application behaviours, making its operation extremely fragile. The NAT-PT approach was subsequently deprecated in RFC 4966 which consigned NAT-PT from Proposed Standard to Historic status, with the comment: "Accordingly, we recommend that: the IETF no longer suggest its usage as a general IPv4-IPv6 transition mechanism in the Internet, and RFC 2766 is moved to Historic status to limit the possibility of it being deployed inappropriately."

IPv4 Exhaustion and IPv6 Transition

The one common assumption in all of these transition scenarios is that this dual-stack transition will take place across the period when there are still sufficient IPv4 addresses to address the entire Internet across the entire transition phase and that the event that IPv6 was primarily intended to avert, the exhaustion of the supply IPv4 addresses from the unallocated pool, would not occur during the transition process.

It is generally anticipated that this transition will take up to a further decade to complete from the current time, while depletion of the unallocated IPv4 address pool may occur within the next two to three years. On one hand, while the overall transition toolbox always assumed a wide array of deployment approaches, this forecast shortage of IPv4 will shift the scaling trade-offs for transition approaches in ways that will be more complex and more expensive to operate than the simpler, dual-stack approach would have been. On the other hand, this is a forced scenario because there is no opportunity to go back in time to try this transition again under different circumstances.

Whatever scenario of IPv6 transition we contemplate, it now has to be one that will take into account the forthcoming acute shortage of public IPv4 addresses, which implies an environment that is heavily reliant on various forms of NATs and possibly some further extensions to NAT behaviours and NAT deployment models, including the possibility of augmenting the NAT-at-the-edge deployment model with various forms of NAT in the middle, as the industry contemplates the potential of so-called carrier-grade NATs and related approaches.

The challenge as we undertake these new technical approaches will be to not lose sight of the fact that short-term cost pressures need to be balanced against the collective long-term desirable outcome of an achievable exit strategy from the ever more complex environment of keeping IPv4 operating.

IETF 72 Activity

In IETF 72, the issues that we have been confronting, with this combination of dual-stack transition to IPv6 and IPv4 address depletion, were discussed in a number of working groups, as well as the Technical Plenary session. What follows is a brief summary of the relevant activity in each of those working groups. While these brief summaries provide a general overview of current activities, the brevity of the description here can get in the way of precision, and the reader is referred to the proceedings of the IETF 72 meeting and of course the associated Internet Drafts for a more complete description of these technical contributions (http://www.ietf.org).

At the Technical Plenary, the IETF was shown some of the underlying metrics of address allocation and the current predictions of depletion of the unallocated IPv4 address pool in 2011. The prospect of broadening the domain of NAT deployment from the edges of the network to parts of the interior boundaries using carrier-grade NATs was also foreshadowed at that session. A report of the experience gathered at Google pointed to a pragmatic approach to dual-stack deployment that advocated undertaking IPv6 support designed to the same production quality standard as IPv4. It was reported that Google was not in a position to dual stack its major service point at present, given that IPv6 today still represents lower reliability and higher latency for some users as compared with IPv4 connectivity to the same service point. A presentation by Apple pointed to consumer products that already make use of IPv6 link-local addressing. The presentation also looked at a dataflow model of connection establishment in a dual-stack environment, where both IPv4 and IPv6 connections are initiated in parallel, and the first path to successfully complete the DNS and initial packet exchange to complete the connection is the protocol that is associated with the application's original connection request (see Figure 6).

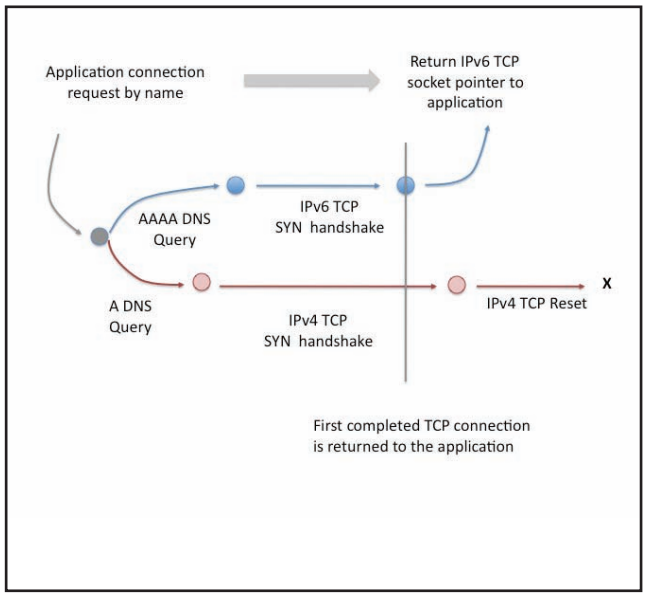

Figure 6. A Dataflow Model of Protocol Selection [Adapted from: Stuart Cheshire, Apple, IETF 72 Plenary].

Continued on next page

IPv6 Transition at IETF72, continued from page 29

The V6OPS working group is looking at some of the basic operational tools to support transition, and, in particular, a reexamination of the requirements for V4-V6 translation mechanisms to see whether there may be viable approaches to provide the original NAT-PT function that might address some of the shortcomings in the original specification. The basic problem being addressed by that effort can be envisaged in a scenario where there are no more IPv4 addresses and a network domain is deployed using only IPv6, and this domain wants to be able to communicate with a domain that is still operating in IPv4 only and has not deployed IPv6. At IETF 72, the working group reviewed a set of goals to see whether there could be a viable set of requirements that could be refined from such a set. While this approach of first defining a set of requirements and then working on potential solutions is a conventional mode of operation for the IETF, the consideration relating to market timing, where deployment of a solution is anticipated to be needed by 2010, is a very sobering call to focus the effort here. A related effort concerns the evaluation of modified NAT behaviour, where the conventional binding space of a vector of inner and outer addresses and ports and an associated protocol is replaced by an outer-side address and port and an inner-side tunnel identifier and an address and port that refer to the NAT device at the other end of the tunnel. The essential concept here is that the NAT function is then a distributed function across a common outer-facing-edge device and the set of inner NATs that are used as a customer-premises-equipment (CPE) device. Other work presented at IETF 72 included a review of proposed refinement of the Teredo specification that would improve its NAT behaviour discovery function from the simple two-mode discovery in the current specification to a mode that discovers up to eight different NAT types. The motivation here is that the more Teredo traffic that can be off-loaded from the Teredo relay to an optimized peer-to-peer connection, the more reliable the Teredo performance. Related work has been reexamining the security issues that are exposed by the use of tunnelling and the potential for disruption and hostile attack on the tunnel.

The BEHAVE working group started out with a charter to provide some standard specifications for the behaviour of IPv4 to IPv4 NAT units, but in recent times this has been expanding to encompass examination of the role of NATs in various IPv6 transition scenarios, including examination of NATs that perform protocol translation. The current agenda of contributions to review includes the IVI scheme—a proposal to use bidirectional address mapping between subsets of IPv6 and IPv4 addresses to allow a form of stateless transition wherein the binding of the translation is carried in the address fields of the packet itself. Another approach to NAT-PT is also being studied. In this case, the asymmetric nature of conventional 4-to-4 NATs is exploited and a proposal for a 6-to-4 NAT was made to the working group. In this contribution, the communication is initiated by the IPv6 host, and the synthesized view of the remote IPv4 world is provided by embedding the IPv4 address in the synthesized IPv6 address. The NAT64 host performs a protocol translation by extracting the IPv4 address out of the IPv6 destination address and providing one of its own addresses as the source address of the IPv4 packet. A NAT binding state is maintained—indexed by the IPv4 address values. The reverse packet performs a binding lookup, allowing the IPv6 destination address to be substituted, and the source IPv4 address is again wrapped up in the synthesized IPv6 packet. BEHAVE also reviewed a contribution calling for specification of the carrier-grade NATs, whereby the NAT translation function is provided at the interior boundary of an Internet service provider (ISP) network in conjunction with NATs being performed at the CPE edge.

The SOFTWIRES working group has also been involved in aspects of the IPv6 transition with regard to consideration of Softwires NAT, or SNAT. SNAT combines IPv4 NAT and IPv4-in-IPv6 softwires to carry IPv4 traffic through the ISP network that uses only IPv6 service. In essence, this approach creates a split NAT whereby the inner NAT is connected to the outer NAT via an IPv6 software tunnel. Multiple CPE NATs are multiplexed through a single external NAT, thereby reducing the total number of IPv4 addresses in use by the ISP (see Figure 7).

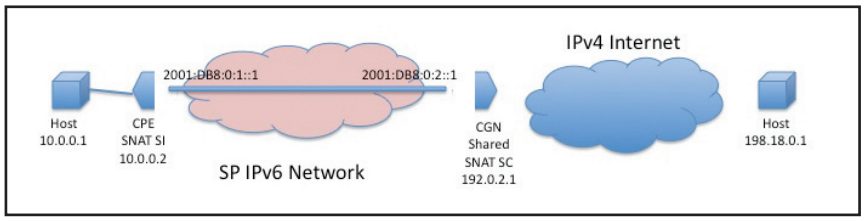

Figure 7. Softwires SNAT.

The INTAREA meeting considered a proposal that calls for standard handling of MTU negotiation, fragmentation, and signalling for tunnels. Given that tunnels appear to be major components of this piecemeal IPv6 transition model, the consistent treatment of tunnelled traffic appears to be an emerging, near-term imperative for the transitioning Internet and for the IPv6 Internet as well. The impending exhaustion of the IPv4 address pool has caused another critical-use address proposal to emerge. In this

IETF Response to the Kaminsky DNS Vulnerability

By Shane Kerr

If you follow the news about information technology, you probably have heard about a new DNS vulnerability discovered by Dan Kaminsky, which is often referred to as the Kaminsky Attack. Since the DNS is a protocol of the IETF—and certainly one of the most successful IETF protocols—let's have a look at how the IETF is dealing with the issue.

Understanding the Vulnerability

There are descriptions of the attack on the Internet. Google can find many, but the Illustrated Guide to the Kaminsky DNS Vulnerability is quite good, even if you know almost nothing about DNS. It can be found at http://www.unixwiz.net/techtips/iguide-kaminsky-dns-vuln.html.

Dan Kaminsky's official talk about the vulnerability can be found on the Black Hat site at http://www.blackhat.com/html/webinars/kaminsky-DNS.html.

Briefly, the Kaminsky attack allows an attacker to put incorrect data into the cache of a recursive resolver, which allows the attacker to return any answer it wants to future DNS queries for a given domain. The attack is quite clever, since almost every piece of the attack has been known for a long time. However, in this case, the pieces have been put together in a new and unexpected way.

DNS Security (Pre)History

The DNS protocol was originally designed without any security, which was not unusual for protocols designed in the early 1980s. Any protocol will likely have security problems, and this is especially true for one that was designed before security was a concern—and for one that has been in use during most of the evolution of the Internet.

The list of security problems with the DNS protocol and its implementations is long and varied, but here are a few:

- Using reverse DNS to impersonate hosts
- Software bugs (buffer overflows, bad pointer handling, and so on)
- Bad crypto (predictable sequences, forgeable signatures)
- Information leaks (exposing cache contents or authoritative data)
- Cache poisoning (putting inappropriate data into the cache)

Work on securing the DNS began in the mid-1990s.

Continued on next page

IPv6 Transition at IETF72, continued

case, it's a call for reservation of IPv4 unicast address space to be used within a carrier's infrastructure for bridging the gap between the carrier-grade NAT at the boundary of the carrier's network and the CPE devices at the boundary to the customer. Given the often protracted debates such calls for reservation of address space often engender—and the relatively short time frame left for exhaustion of the remaining pool of IPv4 addresses—it's not clear whether the IETF will be able to reach a clear consensus on this proposal in the remaining time available.

Summary

There is no doubt that the impending exhaustion of the IPv4 unallocated address pool adds some level of urgency as well as an element of complexity to the IPv6 transition agenda, and the work of the IETF will no doubt increase in intensity in future meetings. It appears we're now working under strict time pressures to develop the standard specification for tools and protocol mechanisms that need to be fielded into production networks within a very compressed time frame; and having every vendor, every network operator, and every operating system suppler devise distinct approaches runs the risk of making the situation more difficult than it would otherwise be.

The challenge for the IETF is to ensure some clarity of focus on the work concerning transition tools for IPv6 that also would assist in increasing the address utilization efficiency of IPv4 addresses and to be mindful of the increasingly strident call for standardization of these hybrid technologies that couple tunnelling, address mapping, NATs, and protocol transformation in ways that application designers, operating system vendors and providers, and operators of networking infrastructure can use in simple and effective ways.

Disclaimer

The foregoing views do not necessarily represent the views or positions of either the Asia Pacific Network Information Centre or the Internet Society.

Geoff Huston is chief scientist at APNIC, the Regional Internet Registry serving the Asia Pacific region. He holds both a B.Sc. and an M.Sc. in computer science from Australian National University. Geoff has been closely involved with development of the Internet for many years—particularly within Australia, where he oversaw initial build of the Internet within the Australian academic and research sectors. He is author of a number of Internet-related books; was a member of the Internet Architecture Board from 1999 until 2005; and served on the Board of Trustees of the Internet Society from 1992 until 2001. See http://www.potaroo.net.

IETF Response to the Kaminsky DNS Vulnerability, continued from page 31

1997: RFC 2065, adding security extensions, was published several years after work began.

1999: After early implementation, it was determined that RFC 2065 needed work, so RFC 2535 was created to fix the flaws (see Appendix B of RFC 2535 for the full list of the differences).

2005: At workshops for implementers, yet more problems were discovered, including extreme difficulty getting secret-key signing material to and from a zone parent. In response, a set of three RFCs were published with the new DNS Security Extensions (DNSSEC) standards: RFC 4033, RFC 4034, and RFC 4035. RFC 4033 was the first time the requirements for DNS security were documented!

2008: DNSSEC gave users the ability to see the entire contents of a zone, and it was by looking at signed records that read no such domain. Privacy concerns made this unacceptable to some domain operators, so a new type of DNSSEC record was created. (RFC 5155 defines the extensions.)

What appears here is the history of DNSSEC proper, but it is not the only work in the area. For example, TSIG (RFC 2845) is a way to authenticate DNS operations by using a shared secret, a procedure that has been widely deployed.

As of this writing, several top-level domains and part of the reverse tree have been secured with classic DNSSEC, as defined in RFCs 4033, 4034, and 4035. The Public Interest Registry (PIR) has announced that in 2009 it intends to sign .ORG with the RFC 5155 extensions.

Recent DNS Security Work

The IETF currently has two working groups (WGs) dedicated to DNS issues:

- The DNS Operations WG, which is an ongoing group dedicated to non-protocol aspects of the DNS, such as DNSSEC best practices and root server recommendations
- The DNS Extensions WG, which is in theory a conventional IETF working group chartered for a specific goal, but in practice it is an ongoing group that has been active since 1999: changes to the DNS protocol, such as the DNSSEC RFCs and explanations of how wildcards work, come out of this working group.

Both groups recently have done security-related work. For example, draft-ietf-dnsop-reflectors-are-evil-06.txt has been approved for publication as best current practices, and work on DNSSEC trust anchor configuration and maintenance is going on in the dnsop working group. Refinements of the DNSSEC protocol (new hash algorithms in the wake of recently discovered weaknesses in existing hash algorithms, clarifications of DNSSEC) are being evaluated, and a draft describing how to make the DNS more resilient against forged answers (draft-ietf-dnsext-forgery-resilience-*.txt) was already being discussed before Kaminsky revealed the problems he discovered.

IETF's Response

Until 7 August 2008, when the details were made public, there were IETF participants who had inside knowledge of the Kaminsky attack. There also were participants whose information came only from press reports or other public sources, such as by looking at patches to open-source resolvers. Initial discussions had slightly surreal qualities, as those who had in-depth knowledge of the exploit discussed the ramifications with those who could only guess.

On the dnsop list, the exploit was announced, and the follow-up mails included a study of the mathematics of the exploit (estimating how long a typical exploit would take), links to and discussions of vulnerability checkers, posts of relevant news articles (such as exploits both in labs and "in the wild"), and links to studies of how quickly resolvers were patched into various environments.

The dnsext list is the WG whereby any changes to protocols would have to be standardized. A number of proposals on the list were discussed, as were a number of nonproposals: for instance, DJ Bernstein has an alternative to DNSSEC that was mentioned, but djb is unlikely to put it forward as an Internet draft.

Most of the proposals centred on ways to make it even more difficult for an attacker to spoof packets. Another common theme was for recursive servers to detect when someone was trying to spoof traffic and therefore, put the resolver into a mode that makes spoofs more difficult.

In the end, the dnsext WG chairs decided to set deadlines for semiformal proposals at the end of September 2008. The proposals would be discussed in October, and in November a recommended forgery resistance approach would be selected. As of the time of this writing, the process is still under way.

Current Status

Some people think the Kaminsky attack will accelerate DNSSEC adoption. The .gov domain will now be signed in 2009, and there have been rumours that the root zone will be signed soon as a result of the attack. However, no changes are expected in the DNSSEC protocols.

The dnsext WG is going to make specific suggestions for a new forgery-resistance mechanism based on the outcome of the selection process. This should help harden resolvers enough to make the Kaminsky attack unfeasible in the current Internet.

The IETF is a great organization for reviewing the details of the DNS on a technical level. A considerable amount of good information and discussion ensued as a result of the attack. It is our hope that everyone involved will benefit from the recommendations that will follow.


Aaron Falk, IRTF Chair

IRTF Report

By Aaron Falk

What follows are summaries of several updates on the Internet Research Groups (RGs), some of which were reported during the Technical Plenary at IETF 72.

Since IETF 71, three Internet Research Task Force (IRTF)-stream RFCs have been published, including RFC 5166 (TMRG), RFC 5184 (MOBOPTS RG), and RFC 5207 (HIPRG). Three drafts are in the RFC Editor's queue, and two are in process toward publication.

A document is being developed that proposes to formally establish an IRTF RFC document stream. The publication is entwined with draft-iab-streams -headers-boilerplates as well as the revision to RFC 3932.

There is still continued interest in establishing a research group (RG) on unwanted traffic mitigations and another one on network virtualization.

During IETF 72, four RGs met. Following is a summary of recent developments as well as of developments reported by RGs during the IETF 72 technical plenary.

Anti-Spam RG (asrg)

The document that describes the mechanisms used for DNS blacklists will be in the standards track. The other ASRG document (best current practices on blacklist operations) is still a draft.

The RG has set up a wiki on spam mitigation techniques that is being further populated and that may evolve into a document analysing why some of those techniques should not be used. The wiki is located at http://wiki.asrg.sp.am.

Delay-Tolerant Networking Research Group (dtnrg)

There are currently three drafts in the RFC Editor queue on delay-tolerant networking (DTN) for deep-space communication (draft-irtf-dtnrg-ltp-*).

In addition, the DTN code base been moved from Intel to Sourceforge. The Networking for Communications Challenged Communities (N4C) project is helping with the code maintenance, and the Defense Advanced Research Projects Agency is funding a phase 3 effort on DTN.

Host Identity Protocol Research Group (hip)

Network address translation and Firewall Traversal Issues of Host Identity Protocol (HIP) Communication has been published as RFC 5207.

Current topics of discussion are:

- Migration of HIP certificate draft to HIP WG
- Using HIP for RFID
- Middlebox authentication extensions to HIP

There are a number of ongoing HIP experiments as follows:

- Boeing is using HIP to build secure overlay networks over untrusted wireless and wired infrastructure.

Continued on next page

IRTF Report, continued from page 33

- HIIT: A video/voice/chat communication system on Linux PDAs utilizing Peer-to-Peer Session Initiation Protocol (P2PSIP) with SPAM prevention over HIP
- ICSALabs: IPv6-only HIP connectivity for gaming and SIP applications using Teredo

Internet Congestion Control Research Group (iccrg)

Three proposals on congestion control are currently under review. The review on Compound TCP is nearly completed, and the reviews on CUBIC and HTCP are just beginning.

CUBIC is an extension to the current TCP standards. The protocol differs from the current TCP standards only in the congestion window adjustment function in the sender side.

HTCP stands for TCP Congestion Control for High-Bandwidth-Delay Product Paths.

The ICCRG Slow Start design team continues to work on characterizing issues with slow start.

Two surveys are currently under discussion: one on open congestion control research issues and one on the current congestion control RFCs (in IRSG [Internet Research Steering Group] review at the moment).

The RG is planning to meet at the next IETF meeting in Minneapolis.

Mobility Optimizations Research Group (moboptsrg)

Unified Layer 2 (L2) Abstraction for Layer 3 (L3)—Driven Fast Handover has been published as RFC 5184.

Current topics include:

- IP Mobility Location Privacy Solutions (revising based on IRSG review)
- Media-Independent Pre-Authentication Framework (revising based on RG Last Call done)
- Multicast Mobility (Problem Statement and Brief Survey being discussed)

Network Management Research Group (nmrg)

The document specifying SNMP trace exchange formats and specifying a format for aggregation of SNMP messages is currently in IRSG review. Another document on SNMP trace analysis definitions has been published in the AIMS (Autonomous Infrastructure, Management and Security) 2008 conference.

A planning meeting to discuss NETFLOW/IPFIX data analysis is scheduled for 30 October in Munich, Germany. The group is looking for new chairs for the RG.

Routing Research Group (rrg)

The focus in RRG has moved from advocating proposals to discussing tradeoffs of different architectural approaches. There is a lot of interest, and many discussions (1,250 messages since IETF 71) have been held. A recommendation is expected to be published by March 2009.

Scalable, Adaptive Multicast Research Group (samrg)

The SAMRG is meeting in conjunction with the Consumer Communications and Networking Conference (CCNC) 2009, which is scheduled for 10–13 January 2009 in Las Vegas. There will be a special session on Scalable Adaptive Multicast in P2P Overlays. More information can be found at http://www.samrg.org under Meetings.

The RG might meet at IETF 73 in Minneapolis.

Transport Modelling Research Group (tmrg)

Metrics for the Evaluation of Congestion Control Mechanisms has been published as RFC 5166. The group is now working on a new draft on the Common TCP Evaluation Suite (L. Andrew and S. Floyd, editors, draft-irtf-tmrg-tests-00.txt).

For more information about the Internet Research Task Force, visit http://www.irtf.org/.

Recent IESG Document and Protocol Actions

A full list of recent IESG Document and Protocol Actions can be found at http://www.isoc.org/ietfjournal/DocProtoActions0402.shtml.

IETF Meeting Calendar

IETF 73
16–21 November 2008
Host: Google
Location: Minneapolis, MN, USA

IETF 74
22–27 March 2009
Host: Juniper Networks
Location: San Francisco, CA, USA

IETF 75
26–31 July 2009
Host: .SE
Location: Stockholm, Sweden

IETF 76
9–13 November 2009
Host: WIDE
Location: Hiroshima, Japan

Register now for

IETF 73

16–21 November 2008

Minneapolis, MN, USA
http://ietf.org/meetings/73-IETF.html

Early bird registration: USD 635 (through Friday, 7 November 2008)
Regular registration: USD 785
Full-time students: USD 150 with on-site proof of ID

IETF 73 is being hosted by Google

Special thanks to Alcatel·Lucent for hosting IETF 72

Special thanks to Google for hosting IETF 73

The ISOC Fellowship to the IETF is sponsored by

Afilias Google intel Microsoft

This publication has been made possible through the support of the following Platinum Programme supporters of ISOC

Afilias Alcatel·Lucent ARIN Cisco Comcast RIPE NCC

IETF Journal
IETF 72
Volume 4, Issue 2
October 2008

Published three times a year by the Internet Society

4 rue des Falaises
CH–1205 Geneva
Switzerland

Managing Editor
Mirjam Kühne

Associate Editor
Wendy Rickard

Editorial and Design
The Rickard Group, Inc.

E-mail
ietfjournal@isoc.org
Find us on the Web at
http://ietfjournal.isoc.org

Internet Society