IETF Journal

Autumn 2006
Volume 2, Issue 2

Published by the Internet Society in cooperation with the Internet Engineering Task Force

Inside this issue:

Editorial

By Mirjam Kühne, Editor, IETF Journal

Welcome to the Autumn 2006 edition of the IETF Journal. In this issue we're pleased to present highlights from last July's IETF meeting in Montreal. While it is not intended as a full report, we believe it provides a fairly comprehensive overview of developments and outcomes from the meeting.

Several important discussions that took place were related to processes, such as the request-for-proposal (RFP) for the RFC Editor function and independent submissions to the RFC Editor. Also discussed were possible changes to the IETF Standards Track. See summaries of those discussions in the plenary report that appears on page 4.

Other updates from the IETF meeting cover the areas of Routing, the Domain Name System (DNS), and Mobility. Note also the special reports from the Internet Research Task Force and the IETF Tools Team.

We are especially pleased to announce a new Internet Society fellowship program that is intended to increase participation at IETF meetings by people from developing countries. The program was launched at IETF 66 with the inclusion of two engineers from Africa, both of whom attended the meeting for the first time. A report on their impressions appears on page 18. Many thanks to the mentors who volunteered to guide the fellows through their first meeting. ISOC plans to continue the fellowship program through the next IETF meeting. We hope to see more participation by engineers from developing regions - both on the mailing lists and at the meetings.

Those of you who are relatively new to IETF meetings might be interested to read about the history of DNS security and the development of DNSSEC, which appears on page 25.

Those who have a longer history of attending IETF might remember the GSE (or 8+8, as it was originally called), a proposal to address the scalability of multihoming for the routing system. This proposal is now being revisited. See the article on this topic on page 29.

Finally, ISOC would like to take this opportunity to recognise the 25th anniversary of the development of the TCP/IP protocol. We offer a special tribute to this breakthrough in communications technology on page 23.

Have fun reading. We welcome your comments and invite contributions to future issues of the IETF Journal.

News from the IETF Chair

By Brian Carpenter, IETF Chair

IETF 66 was held in the Palais des congrès in Montréal, Canada, close to the city's Chinatown and a short walk from the famous old town. We shared this enormous facility with a conference on pediatric pulmonology, and from its attendees we learned that we could be far more imaginative in our naming of meeting rooms. Their exhibit area, for example, was named Thoracic Park. IETF 66 was hosted by Ericsson Canada which, with the help of Combat Networks and a team of dedicated volunteers, provided excellent wireless networking throughout the week. More than 1,200 people from 44 countries attended. Notably, more than 70 people named China as their country of origin. As always, the week was a busy combination of working group (WG) meetings, BoF (birds-of-a-feather) sessions, research groups, and formal and informal meetings of all kinds on the side.

Brian Carpenter
IETF Chair

Since IETF 65, four new WGs were chartered and 13 WGs were closed, leaving approximately 120 WGs currently chartered. Between the meetings, the WGs and their individual contributors produced 463 new drafts, not to mention 852 updates. The Internet Engineering Steering Group (IESG) approved 96 drafts for publication as RFCs. For the first time in several years, this was fewer than the number of RFCs published during the same interval - evidence that the publication backlog is getting smaller.

The IESG has opened a new section on the IETF Web site for material provided by the IESG, (see http://www.ietf.org/IESG/content/**).** Among other things, the site includes a link to the IESG's own wiki, an informal guide to IESG procedures that was developed as a collaboration among Area Directors.

The IESG has decided to start session scheduling for future meetings two weeks earlier to allow extra time for resolving clashes in the draft agenda. Similarly, BoF requests must now be sent earlier to allow extra time for evaluation. Deciding which BoFs to approve is one of an Area Director's most important tasks, and it requires consultation with both the IESG and the Internet Architecture Board (IAB).

Scheduling information for the next IETF meeting may always be found via http://www.ietf.org/meetings/meetings.html**.** I look forward to seeing many of you in San Diego, California, USA, from November 5-10, 2006.

IETF 66 Facts and Figures

1236 registered attendees

from 44 countries

4 new WGs

13 WGs closed

463 new Internet-Drafts

852 updated Internet-Drafts

80 IETF Last Calls

96 approvals

around 138 published RFCs (88 standards and BCPs)

1 appeal

News from the IAB

By Leslie Daigle, IAB Chair

Leslie Daigle
IAB Chair

During the Thursday plenary at the Montreal IETF meeting, I gave an overview of the IAB's activities since the last IETF meeting. The details, including pointers, can be found at http://www3.ietf.org/proceedings/06jul/index.html. In this note, I'd like to draw attention to 3 particular areas you might want to keep an eye on and/or participate in.

First, as voiced at the IAB's BoF sessions at the network operator group meetings during the last year (NANOG, APRICOT, RIPE), as well as many other places, there are growing concerns about the evolution of network addressing and routing architectures. This year, the IAB is planning to hold an invitational workshop to examine the issues. The stated goals of the workshop include producing a concise problem statement of the current concerns about scalable routing and addressing. Concerns we have heard raised, that will be discussed at the workshop, include:

- Difficulty in changing provider due to PA/CIDR addressing schemes
- Lack of effective multi-homing support
- Limited capability to protect against either the spoofing of individual host IP addresses or entire IP address blocks
- Limited ability to secure the routing system itself

Stay tuned for the workshop report.

Second, the IAB recently finalized a document outlining some possible next steps for evolving the Internationalized Domain Name (IDN) standard. This document, written by John Klensin, Patrik Fältström and Cary Karp, describes some of the lessons learned and issues perceived with current IDN usage. The statement proposes some areas for further exploration within IETF work, ICANN bodies and elsewhere. Some of the documents main conclusions include:

- IDNA and its tables need another look in terms of its use of Unicode
- Need more stable normalization
- May need a more restricted, permitted character list
- There is no IDNA/DNS solution to several problems – they can't be solved in this technology, or perhaps any technology
- Time for another look at the "above DNS" approaches?

See http://tools.ietf.org/html/draft-iab-idn-nextsteps-06 for full details and suggestions of where further work may be pursued.

Finally, as one of its non-technical work activities, the IAB has been working to help develop a clearer community definition of both the RFC Series and RFC Editor function. This is in support of this year's IASA RFP process for the RFC Editor function. The IAB has proposed a framework for the RFC Series and an RFC Editor function for the specific purpose of ensuring the RFC Series and RFC Editor role are maintained and supported in ways that are consistent with the stated purpose of the RFC Series and the realities of today's Internet research and engineering community. Details can be found in <draft-iab-rfc-editor>. However, the RFC Series contains more than IETF documents. As part of this effort, the IAB has sponsored the open independent@ietf.org mailing list, which is discussing draft-klensin-rfc-independent as a process document to describe the handling of the independent submissions the RFC Editor receives today.

These are three key areas of active work for the Internet community as a whole - not just for the IAB. There's plenty of engineering work to go around!

New Administrative Process and New Challenges Discussed at IETF 66

By Mirjam Kühne

Please note that this is not a complete report of the plenary sessions, but is instead summarising the highlights of the discussions.

More than 1200 participants from 44 countries convened in Montreal in July for IETF 66. Hosted by Ericsson and held at the Palais des congrès in Montréal, the meeting featured a robust network provided by RISQ and CANARIE. In his welcoming remarks in the first plenary session, IETF Chair Brian Carpenter described the progress made in the past four months since IETF 65.

Since then, four new working groups (WGs) were formalised, 13 closed, and approximately 120 are now chartered. A total of 463 new Internet-Drafts were submitted, 56 percent of them submitted in the four weeks prior to the meeting. Internet-Drafts updated since Dallas total 852, with 67 percent of those being updated in the four weeks prior to the meeting. In addition, there have been 80 IETF Last Calls, 96 approvals, and roughly 138 RFCs published.

Brian described a number of enhancements being made to the IETF meeting process, including a new scheduling procedure for WG and BoF sessions. The IESG is proposing to commence scheduling for working group meetings two weeks earlier than usual so as to allow more time for adjustments to the agenda and to minimise the possibility of clashes. Looking ahead to IETF 67, planned for November 5–10 in San Diego, the change means WG scheduling will open on August 7, with a cut-off date of September 18 for both WG scheduling and BoF proposals. A preliminary agenda will be announced on September 29 and the final agenda will made available on October 16.

Standards-Track Reform

A significant portion of the plenary discussion was dedicated to the future of standards-track reform. The current three-stage standards process, as described by Scott Bradner in RFC 2026, was intended “to provide a fair, open, and objective basis for developing, evaluating, and adopting Internet Standards.” As such, a specification undergoes first a period of development and review by the Internet community, and then revisions based on preliminary implementation. It is then adopted as a Proposed, Draft or Full Standard by the IESG and then published as an RFC.

As mentioned in the previous issue of the IETF Journal, in reality, protocol specifications are often adopted and implemented before they are approved as full standards. RFC 3774, published in May 2004, addressed the erosion of the three-stage process, stating that “in practice, the IETF currently has a one-step standards process that subverts the IETF's preference for demonstrating effectiveness through running code in multiple interoperable implementations. This compresses the process that previously allowed specifications to mature as experience was gained with actual implementations.” This problem has further been identified and described in the newtrk WG, chartered in 2004, but no clear community consensus has emerged as a solution to the problem.

Opening the plenary discussion, Brian offered three possible directions. In option 1, RFC 2026 would be clarified, not just to “reflect the reality” that few standards go through the process as laid out in RFC 2026 but also to acknowledge that “the problems with the existing standards track are not serious enough to justify the effort needed to make substantial changes.” Option 2 focuses on document relationships,

IETF 66 BoF Sessions

Applications Area

dmsp - Distributed Multimodel Synchronizastion Protocol

wae - Web Authentication Enhancement

RAI Area

rtpsec - Securing the Real-Time Transport Protocol

Security Area

hoakey - Handover and Application Keying and Pre-Authentication

nea - Network Endpoint Assessment

Transport Area

offpath - Path-decoupled Signalling for Data or, as the newtrk charter now says, “[creating] a new series of shorter IESG-approved IETF documents to describe and define IETF technology standards.” Option 3 would focus attention on the “maturity” levels of proposed standards, which would mean revising the three-stage IETF standards track described in RFC 2026.

A spirited debate ensued, with a number of participants admitting that large segments of the industry start accepting standards as soon as they have an RFC number, preferring not to wait for specifications to go through the entire three-stage process. However, a critical step in the standards track is one in which interoperability is demonstrated through a number of independent implementations, a part of the process that, if neglected, could lead to problems down the road. Other participants preferred to look at ways to get the “running code” piece back into the IETF, maybe as part of the WG specification writing process and not as part of the standards track.

“Sometimes we spend too much time on process and yet it seems like the problems are not big enough,” said Sam Hartman, who added that “while it would be nice to fix RFC 2026,”it doesn’t seem to be a high-enough priority for the community. “If we turn to it,” he said, “we need to get it right; it would be easy to get this one wrong and that would be very bad.” Bradner agreed, pointing out that “the Internet hasn’t stopped because we are not following RFC 2026.” However, Bradner also said he believes that a good description of the standards process would be useful. “The three-stage process shows maturity levels of the document,” he added.

Long-time IETF participant Dave Crocker suggested that a second label be explored, one in which the market reports its use of particular standards. In this scenario, a document would move to “Proposed Standard” for X number of years. If the market does not come back within that many years, it ceases to be a standard. Dave said that while not moving through the levels doesn’t seem to be hindering the process, he lamented the “lack of feedback about the specs.”

Brian put the issue to an informal vote, resulting in no clear consensus on which of the three options is preferred. There was some agreement, however, that the nature of the problem may have been incorrectly stated and that the discussion would continue on the IETF mailing list.

IAOC Makes Progress, Moves toward the Future

Lucy Lynch, chair of the IETF Administrative Oversight Committee (IAOC), reported that a request-for-information (RFI) had been issued for the RFC Editor function, as were a statement of work (SoW) and a draft IANA service-level agreement (SLA). Both the SoW and the SLA were sent to the IETF community. Shortly after the IETF 66 meeting, an RFP for the RFC Editor function was published.

Lucy gave a brief summary of a two-day IAOC retreat held in May, where members outlined future activities and turned their attention to setting goals. Topics discussed at the retreat included meeting planning, funding models, and IASA goal setting, as well as plans for finalising an IASA work plan for 2006–2007. Lucy announced progress on the newly formed IETF Trust, saying that some issues still await resolution within the Intellectual Property Rights (IPR)-WG, which affect both ISOC and the Trust. “We need to deal with current physical assets transferred from CNRI and develop a document’s retention policy,” said Lucy. The IAOC expects to solicit additional donations of IETF-related IPR in the coming year and the group is addressing the need for inventory tools and archival storage. More information can be found at the IAOC Web site at http://wkoi.uroegon.edu/~iaoc/, which will soon move to http://iaoc.ietf.org.

IASA: The First 180 Days

Reporting on the activities of the IETF Administrative Support Activity (IASA), IETF Administrative Director (AD) Ray Pelletier outlined plans to establish a multievent contract with a hotel chain, which he said, was intended to reduce both costs and risks associated with IETF meetings as well as to improve long-term planning. The group is also considering the possibility of outsourcing the operations of the IETF meeting network, which Ray said would reduce the host's workload and possibly attract organisations that may not possess expertise in running big networks to host an IETF meeting.

According to Ray, the additional resources that have recently been invested in the RFC Editor had met with success, as evidenced by the elimination of copy-edit backlog. An RFP is being issued for the RFC Editor function and Ray anticipates that new contracts will be negotiated in September. A transition is scheduled for December and RFC Editor services are expected to commence in January 2007.

IANA Operations Manager Barbara Roseman announced that while a few positions remained unfilled, the increase in administrative staffing has resulted in measurable service improvements. By introducing redundancy and providing for back-up for staff on leave, single points of failure have been eliminated. In addition, the regularising of request processing has led to increased responsiveness and better collection of statistics. However, according to Roseman, due to legacy data and the fact that highly detailed and accurate statistical reporting requires considerable preparation time, a few issues remain regarding the collection and analysis of statistics. Prior to IETF 67, IANA expects to complete implementation of the SLA within the IAB, migration of historic data, public documenting of IETF-related processes, and the launch of a new web site for IETF assignments.

IETF66 Plenary
Photo: Shane Kerr

IETF 66 Technical Report Draws Attention to DNS and IDN

The publication of draft-iab-idn-nextsteps took center stage at the technical session of the IETF 66 plenary. IAB Chair Leslie Daigle made it clear that the intention of the document is not to propose solutions, but "to identify issues and in some cases possibilities." The document identifies a number of experiences with IDNA and IDN, including concerns about character spoofing and similarities that do not currently have technical fixes. Leslie also pointed out the difficulty in trying to design policies that are helpful in solving such problems but not so restrictive as to make IDNs unappealing.

The position of the IAB is that the time is right to review certain aspects of IDNA - especially the tables. Leslie suggested that a more restricted approach to permitted characters may be needed. There was also general agreement that IDN creates problems for DNS as it exists today. "We could spend a long time debating this," said John Klensin. "But the point is that these things overload the DNS and DNS was never built to deal with these kinds of issues."

Independent Document Submissions

A proposal to change the document draft-iab-rfc-editor-00.txt based on feedback from the community was offered by Leslie. The draft describes the need for an RFC Series that is implemented for the community and that has a number of key characteristics: it needs to be expertly implemented and clearly managed, and it needs appropriate community input and review of activities.

According to Leslie, the proposal would allow for more explicit integration of the RFC Editor, the IAB, and the IETF, while ensuring that each retains clear and distinct responsibilities. RFC Series decisions would thus be more open to community discussion, with the IAB monitoring the discussion and maintaining the coherency of the series and related discussions. Leslie said the IAB believes that the proposed changes would clarify the role of decision-making groups and ensure the right balance of RFC streams. The community is encouraged to participate in the discussion on a new mailing list at https://www1.ietf.org/mailman/listinfo/independent

A fair portion of the plenary session was dedicated to a public discussion about independent submissions to the RFC Editor. IAB member Olaf Kolkman, who led the discussion, is the moderator of the independent@ietf.org mailing list.

Olaf pointed out that it is important to the RFP process that policies and procedures be documented. The IAB is looking for broad community input, which is why the mailing list was announced at the IETF as well as other venues, such as the Regional Internet Registry (RIR) communities. The feedback received on the list thus far indicates that the document :draft-klensin-rfc-independent-02.txt provides a good desription of the current process and procedures. There are a number of open issues, related mostly to different interpretations of RFC 3932 - “The IESG and RFC Editor Documents: Procedures” - and which require broad agreement. Olaf presented a number of issues and encouraged people to raise their voices on the mailing list. He pointed out that now is the time to set up a good working structure for RFCs and independent submissions.

In the discussion that followed, Allison Mankin cautioned participants about making changes to RFC 3932 as it is an approved BCP that was reviewed by the community. In contrast, others felt it is important that the document reflect reality, such as in the editorial process and in the way reviews are accomplished.

It was also suggested that non-IETF-consensus documents should have a different name or label, so they do not get confused with standards-track documents.

This discussion will be continued on the independent@ietf.org mailing list.

The Challenges of Appeals

With formal appeals on the rise, the IAB was asked if having to deal with those types of challenges is consuming too much of the group’s time and, if so, whether the IAB would consider using a separate organisation to handle appeals. Members of the IAB agreed that the appeals process is an important part of the IAB’s work, as well as part of its charter. As Eric Rescorla pointed out, the appeals process may take time, and the IAB should probably think about streamlining the process instead of creating a separate pool of people to handle appeals.”

A key concern expressed by attendees was the potential of one-person appeals to obstruct the process. “The past several months have demonstrated the potential for people who are not actively participating in getting things done to be launching appeals,” said John Klensin. He added that the IAB should consider mechanisms for “applying dampers” for appeals that are submitted repeatedly, before the situation gets out of control.

Of equal concern was the need to handle changes to the appeals process with care. Bradner reminded the group that the appeals process was not only an issue of fairness but also a condition set by the IETF’s insurance company as means for avoiding litigation. “We need to be careful when changing the appeals process,” he said. “We cannot be seen as an unfair organisation by insurance companies, even though we have never had to use the insurance so far.”

Dave Crocker expressed appreciation for the diligence with which appeals have been handled thus far but said he felt “real concern” about the abuses he has seen recently. “The IETF model of rough consensus is about broad-based concern and consensus,” he said, suggesting that the IETF consider requiring that appeals be submitted by multiple people or at least supported by others. “It might also make sense that an appeal should not be resubmitted by the same person,” he said.

While it was agreed that care should be taken when addressing changes to the appeals process, there was general agreement that the number of appeals should be reduced by any means.

Anticipating the Future

When asked what they thought the most pressing problems over the next five years might be, IAB member Lixia Zhang answered that the routing system needs a fundamental change. “Congestion and routing were problems at IETF 1,” she said. “Congestion solved itself but the routing problem is still there.” She also pointed out the value of looking back to see which protocols have been successful and which others took a lot of time but didn’t take off. Conducting this type of review, she said, would help in the future with prioritising tasks.

IAB member Bernard Aboba answered that until the IAB workshop on unwanted traffic, he had underestimated the security problem. “Often we seem to underestimate how difficult it is to apply security mechanisms on a global scale,” he said. Dave Oran, a new IAB member, said that over the past 20 years, there have been dramatic shifts in traffic as a result of e-mail, the Web, and VoIP. “We need to be getting more knowledgeable about applications that have no known upper bound on bandwidth,” he said. “We need to figure out network management without thinking that adding bandwidth is the solution.” It was recommended to work closely with ISOC on issues related to policy and politics.

The discussion gravitated toward nontechnical issues, particularly those involving policy, including whether the IETF should be concerned with Net Neutrality. “There are a lot of important issues out there that are at a different level,” said Bradner, “but we should not overlook them. Routing might not be important anymore if bad things happen at level 8 or 9.”

All presentations given during the IETF 66 plenary session can be found at:
`http://www3.ietf.org/proceedings/06jul/index.html`

Palais des congrès Montréal
Photo: Mirjam Kühne

IETF 66 Review: Routing

By Geoff Huston

Note: This article does not attempt to provide a complete summary of all IETF activities in this area. It reflects the author's personal perspective on some current highlights.

The IETF uses a broad classification of activity into "Areas", and within each one working groups (WGs) are chartered. For most of the IETF Areas, in addition to working group meetings, there is a general area meeting that is intended to review work across the entire Area. This is a report of the Routing Area meeting held at IETF66.

Routing Area Directorate

The Routing Area Directorate is an advisory group of routing experts selected by the Area Directors. The Area Directors use the feedback from the Directorate while making decisions on a range of topics related to the IETF Routing area.

One of the consistent issues with the IETF is that of ensuring that documents receive adequate and timely peer review. The Area Directors will be putting some effort into the Routing Area Directorate to ensure that directorate members will be a resource for WG chairs to undertake early review of routing drafts prior to the final steps of IESG submission. The Area Director review will also use the Routing Area Directorate for comments as part of the Director's review process.

Routing Area WG Reports

Bidirectional Forwarding Detection (BFD)

This WG is not meeting at IETF 66. The WG is close to undertaking a last call on its current document set. The Routing Area Directorate should be involved in review at this stage

Common Control and Measurement Plane (CCAMP)

This WG has a new new co-chair, Deborah Brungard. There was reported to be steady progress on drafts. Considering that the WG currently has some 29 active drafts, this is certainly a very active group. There is currently a strong focus on interior gateway routing protocols with traffic engineering capabilities and automatic mesh capabilities.

The GMPLS-controlled Ethernet Label Switching (GELS) topic, relating to GMPLS control of Ethernet environments, which held a BoF at IETF 64 will now be folded into the CCAMP WG charter, and GMPLS control protocols will be used. There is no new data plane definition in this proposed direction, which is addressing one of the more contentious issues that surfaced in the GELS BoF.

Forwarding and Control Element Separation (FORCES)

The WG is attempting to complete the base model document, which then forms the foundation for the WG's document set. The WG has decided to place a deadline for review comments for this draft, in an attempt to complete the document in the near future. Current work includes consideration of the Transport Mapping Layer, and the potential to use SCTP in this context, as well as a FORCES MIB.

Inter-Domain Routing (IDR)

There has been some progress in terms of moving documents thriough the IESG, and a number of RFCs were published after IETF 65 (RFC 4456 and RFC4486). A number of additional documents are to be passed to the IESG in the weeks immediately following IETF 66. It was noted by the WG chair that there have been various recent efforts to add capabilities to BGP in working groups outside the Routing Area, and a call for IDR involvement in this effort was made. As one example, the Softwires WG in the Internet area will work on some framework documents to attempt to address these issues, and some deliberate effort to use cross-WG postings was made in that case. Some further discussion with the ADs was proposed. Other areas and other WGs do undertake some work on extensions to routing protocols, with the need to manage outcomes to ensure coherence and consistency of the resultant routing protocol extensions. This is asserted to be a matter for AD attention and work management. It was recommended that some coordination effort across WGs should be undertaken as early as possible when work on routing protocols is taken up in other working groups ("early cross-area review" is the procedure being considered here).

IS-IS for IP Internets (ISIS)

There are some further work items in this WG, including support for IPv6, multi-topology routing, policy control mechanisms, and extensions for advertising routing information and HMAC SHA authentication. It is expected that the WG will be rechartered to reflect this intended work agenda.

Layer 1 Virtual Private Networks (L1VPN)

It was reported that this WG is progressing well, with the framework document completed in the working group and applicability description underway. A large part of this WG's agenda is concerned with emulation of edge-to-edge circuit states via GMPLS paths. There has been some cross-working group and cross-area review of these L1VPN drafts.

Mobile Ad-Hoc Networks (MANET)

The three core MANET documents have been updated. The WG is now looking at a common generalized messaging format. Update of this document has been completed, and work on a common neighbourhood discovery protocol is underway. Within MANET there are at present both pro-active and re-active approaches, and some effort to pull these together will be undertaken. The chairs of Autoconf (Internet Area) and MANET are working on a common-architecture document and will publish this following IETF 66. This is a document that would benefit from early cross-area review.

Multiprotocol Label Switching (MPLS)

This working group had a busy agenda at IETF 66. The highlight is point-to-multi-point point traffic engineering-label-switched path (TE LSPs) and a report of the meeting with ITU-T Q12/15 on T-MPLS.

Open Shortest Path First IGP (OSPF)

Most the original OSPF charter documents have been completed by the working group at this stage (MIB on V2 and security on v3). Only the OSPFv3 MIB remains. The WG is currently in the process of re-chartering with work items including Multi-Topology Routing (MTR) in OSPF and OSPF in a MANET environment. The initial

OSPF MANET work has focused on flooding and adjacency reduction optimizations. Some OSPF WG review of the CCAMP documents has been requested.

Path Computation Element (PCE)

The PCE architecture RFC (first WG RFC) has been published: RFC 4655 and two PCE documents are in the RFC editor queue. The base protocol space is now stable, and a call for review has been made. There have been some proposals for further PCE work, and the chair would like to hold off on further specification of requirements until there has been some experience with the base protocol. There was consideration of an experimental track on manageability of the PCE specifications. Policy work is outstanding, as is consideration of the complete requirement set.

Routing Protocol Security Requirements (RPSEC)

The RPSEC Working Group is finishing up with work on the generic threats document, and this document is now back in the RFC Editor's queue. The documents on OSPF vulnerabilities and the BGP attack tree are being reviewed, and appear to be close to completion. The BGP security requirements document is also considered to be almost ready for a working group Last Call. Without further new items, the WG will have completed its current charter with those documents.

Routing Area Working Group (RAWG)

This group did not meet at IETF 66. Currently on the Working Group's agenda is an open question about loop-free/microloop detection algorithms that need to be resolved prior to last call of the IP Fast-Reroute document.

Secure Inter-Domain Routing (SIDR)

This is the first meeting of this working group since it was chartered following IETF 65. Current work is focused on the basic instruments of trust within the routing and addressing environment, examining a profile for X.509 Public Key certificates that would be able to function as 'right-of-use' tokens in a routing context. This would form the basic trust injection function for the work on securing route origination. The next work item is that of an overall architecture for secure inter-domain routing systems. In addition at IETF 66 SIDR gave some time to air the varying proposals for TCP MD5 key rollover.

Virtual Router Redundancy Protocol (VRRP)

The VRRP for V6 spec with the IESG, as it relates to consideration of the SeNd technology. The unified MIB is under MIB doctor review, and the subsecond timer work is under WG review. This working group did not meet at IETF 66.

RFC 1264 Discussion

What should the requirement be for routing area documents that are forwarded to the IESG for publication as RFCs? The Routing Area had previously requested implementation reports, detailing the outcomes of implementation and interoperability of the specification, preferably from at least two independent implementations as a pre-requisite for passing a document to the IESG for publication at the Proposed Standard level. There was strong consensus at the area meeting to have "good" requirements, however the developing picture appears to be visible consensus in the Routing Area to have Proposed Standard publication requirements that are no different than the other IETF areas. At this stage it is proposed that it be a working group matter to determine requirements related to implementation reports, and due attention should be given to quality of WG documents in this process. This would make the requirements specified in RFC 1264 to be considered historic for the Routing Area, particularly in terms of specifying more stringent preconditions for Routing Area Proposed Standard documents.

IP Routing in the Global Information Grid

This was a report to the meeting on the recent U.S. Departement of Defense initiative called the "Global Information Grid". This initiative is projected to have a number of routing objectives, as well as objectives of supporting Quality of Service (QoS) and security. This Global-Internet-Geography (GIG) environment proposes pervasive node and network mobility, implying that the current Internet routing paradigm may not be totally applicable in a number of dimensions. Some potential for "fundamental change" in inter-domain is contemplated.

IAB Routing and Addressing Workshop

Dave Meyer reported on the proposed IAB workshop on routing and addressing. Mid-October is the likely time for this by-invitation-only IAB workshop. Current workshop activity appears to be the definition of a routing problem statement and a requirement list. Ross Callon commented that it would be valuable for a broader consideration on the routing-discuss mailer on the identification of the problems of routing and addressing.

(The routing-discuss mailing list is: `routing-discussion@ietf.org`)

Recent IESG Document and Protocol Actions

A full listing of recent IESG Document and Protocol Actions can be found at:

`http://ietfjournal.isoc.org/DocProtoActions0202.htm`

IETF 66 Review: DNS

By Jaap Akkerhuis and Peter Koch

Note: This article does not attempt to provide a complete summary of all IETF activities in this area. It reflects the author's personal perspective on some current highlights.

For more details about DNS-related working-group meetings, refer to the minutes and Jabber notes for each meeting at http://www3.ietf.org/proceedings/06jul/index.html

DNS Extensions Working Group

The DNS Extensions Working Group deals with both the details of the DNS protocol and its extensions, such as the Security Extension (DNSSEC)[1]. The group reports progress on reducing backlog and advancing documents. To date, the Wild Card Clarify draft, which updated the wildcard definition of RFC 1034, achieved RFC status (RFC 4592: The Role of Wildcards in the Domain Name System). According to the authors, the RFC did not change the essence of the protocol; rather, it refined the definition of RFC 1034 to make it more consistent and to reflect reality.

The dynamic host client identifier draft is currently in the RFC Editor queue and the new Resource Record Type-code has been assigned by IANA (DHCID, Type-code 49). Four other documents are in IETF Last Call; two are in working group Last Call and the others are close to that stage.

NSEC3 Update

The work on NSEC3 is still progressing. A workshop where various implementations and specifications were tested was successful. Several issues were discovered in both the drafts and implementations. David Blacka presented those in detail at the IETF meeting, and a lively discussion followed that is likely to continue on the mailing list. A new workshop test is planned for September. In the meantime, a permanent testbed has been set up for the purpose of testing various ideas. Details of the efforts can be found at http://www.nsec3.org. Discussions will take place on the official dnsext-wg mailing list.

DNS Trust Anchor Management

A couple of competing drafts have emerged proposing a roll-over mechanism for updating the trust anchors in DNSSEC aware resolvers. Several reviews of the various methods have been published on the dnsext mailing list. A few common elements of the drafts include threshold schemata or the use of timers. The timer-based proposal seems to be the most complete, and it was agreed that the proposal should serve as the basis for further work.

DNAME Clarify Effort

Meeting participants expressed concern that RFC2672 (status: Proposed Standard) suffers from omissions and is unclear in terms of how the DNAME interacts with wildcards, EDNS0, compression and similar stuff. Plans are currently under way for the use of DNAME in a wide-scale operation, though the DNSSEC implementers expressed the need for clarification. The plan is to first collect open questions and perceived ambiguities in the DNAME specification in a separate draft and then to ask the working group for feedback in order to create resolutions that can be added to a DNAME Clarifications document. An explicit 'non-Goal' is the creation of a DNAME-2.

[1] See a historical perspective on DNSSEC later in this issue.

New Work: DNS Cookies

Donald Eastlake presented a proposal for a dynamic system that requires neither configuration nor set-up in order to provide weak authentication of queries and responses between servers and resolvers. It can be described as a weaker version of client authentication and is intended to greatly reduce the increasingly popular attacks that use forced source addresses. Although there was some interest in the proposal, it has not been formalised, and the community is encouraged to comment. Feedback on operational requirements will be solicited from the DNSOP WG.

Mark Andrews proposed a method for revealing zone cuts without having negative entries recorded in caches. He wrote an Internet-Draft and asked about achieving Last Call (LC) for the document. Apparently, the method has already been implemented in the latest version of bind.

Milestones

Given the catching up there has been on the backlog and the new work trickling in, there is a need to update the milestones. The chairs will prepare a draft to discuss.

DNS Operations Working Group

The DNS Operations Working Group deals with the daily dross of operating DNS. The WG does not create protocols; the participants discuss the use of protocols in practice. Work on reducing the backlog continues and there are now a couple of Internet-Drafts in Last-Call stage or close to publication, such as the DNSSEC best practices and the server id. The last extension will make it easier to maintain DNS any-cast server clusters.

Open Recursive Servers

It became increasingly popular in the past year to use public recursive name servers as amplification mechanisms in D-DOS attacks. In essence, one spoofs the address of the victim and generates, via such public name servers, massive amounts of traffic to the victim. At the request of the WG chairs, Frederico Neves and João Damas have written an I-D to document this practice for the DNS operators and there have been discussions about the trade-offs when dealing with these attacks. Solutions are being discussed, and a new version of the I-D is expected to be published soon.

Default Local Zones and AS 112

Project AS112 (`http://www.as112.net`) is a loosely organised group of volunteers who operate systems that respond to DNS queries for the reverse mapping of so-called private address space of BCP 5 (RFC 1918). These queries should never meet the public Internet. Therefore a two-stage approach for managing them was suggested. First, nameserver software vendors are encouraged to avoid leaking queries by directly answering those about local zones and, second, as a potential new WG work item, by documenting the setup of AS112 for new team members and organisations or site DNS administrators, which will reduce the pollution.

Other (Non-WG) Internet-Drafts and Discussions

Web server cookie-validation ads often make assumptions about the structure of the Top-Level-Domain (TLD) name space. As Yngve Pettersen presented, even though administrative hierarchy does not imply or follow the hierarchy of the DNS, concerns are being raised about attempts to subvert this principle. A number of suggestions were made, but no conclusions or actions were finalised.

ENUM WG

ENUM is a protocol that links the DNS with the VoIP and PSTN worlds by mapping phone numbers to services such as SIP, instant messaging, multimedia mail messages and, of course, phone calls. Defining and specifying ENUM services has been a major work item in the ENUM WG for some time, so it seems more than reasonable that the WG should now address the issue of providing guidelines and a registration template to aid future service registrants. The prospected multitude of services, and the use of the DNS NAPTR record, will lead to larger DNS responses; even more so when DNSSEC is used in combination with ENUM. Therefore, the WG is currently working on a recommendation to vendors and operators of ENUM-related name servers and clients to support the DNS EDNS0 protocol extension for larger packet sizes. There is a document collecting examples of the ways in which different implementers chose to interpret the ENUM specification. It is expected to serve as a source of information when it comes to clarifying and advancing RFC 3761, which is the big task for the ENUM WG in the next year.

Miscellaneous

The DNS is an attractive data publishing and retrieval mechanism, which explains why so many IETF working groups are working on DNS-based solutions for their particular problems. This offers ample opportunity for cross-working-group discussions that make it possible to share experiences designing DNS extensions while at the same time preserving the benefits that the DNS offers, such as scalability. The dnsext WG is working on an update to RFC 2929 that will make registration of new resource record types easier and that provides some operational and architectural guidance. This is also the intention of a draft initiated by the IAB that discusses trade-offs of several popular approaches for basing new applications on the DNS.

Early consultation with the dnsext and dnsop WGs is encouraged wherever use of the DNS appears on a WG's work plan.

IETF 66 Review: Wireless

By James Kempf

Note: This article does not attempt to provide a complete summary of all IETF activities in this area. It reflects the author's personal perspective on some current highlights.

A new wireless-related working group, 16NG, which is working on IPv6 over 802.16 wireless links, met for the first time during IETF 66. The group previously met twice as a BoF and several times in interim meetings as part of its aggressive agenda to meet the deadlines for WiMax network deployment A design team is currently working on a document describing how to map the IPv6 subnet onto the 802.16 link.

HOKEY, which deals with usage of the extensible authentication protocol (EAP) in emerging mobile networks held its second BoF meeting at IETF 66 to discuss standardizing the EAP application key hierarchy for handover, backend authentication-authorization-accounting (AAA) work for preauthentication, and AAA key distribution for services.

The number of new BoFs related to wireless appears to be slowing. With a number of working groups pursuing wireless topics already formed, particularly in the Internet area, the capacity for IETF to do more work is limited. Therefore, chartering new groups may need to wait until existing working groups are finished and have closed.

Group Providing IP Mobility Support for IPv6 Starts Up

NETLMM, a group recently formed in the Internet Area is currently developing a protocol to provide IP mobility support for IPv6 in the network, rather than as a host-based service like Mobile IP. According to the NETLMM charter, the host is not involved in IP-level movement. The host keeps the same IP address across a span of the wireless network and wired backhaul that is limited topologically to a local area, called a localized mobility management domain. The host keeps the same IP address as it moves around a geographical area that is covered by a collection of wireless access points. The access points are connected through a wired IP backhaul into a topologically limited domain, called a localized mobility management domain. The exact topological extent of the localized mobility management domain depends on particular deployment circumstances. The host never changes its IP address as it moves across the localized mobility management domain, while routing updates to a mobility anchor from the local access router ensure that the host continues to receive its traffic. Existing cellular architectures provide similar support, but are limited to a single family of wireless technologies, such as UMTS, while NETLMM is independent of wireless link technology. NETLMM serves to complement Mobile IP, since Mobile IP is still needed if the host requires session continuity as it moves between localized mobility management domains.

A problem statement document and a requirements document have been sent to the IESG. The IESG has reviewed the problem statement document and returned it for further editing; the requirements document is still in Area Director review.

At the IETF 66 meeting, the primary topic of discussion was the first draft of the design team protocol document. The design team has been meeting since February, and the first draft of its protocol design was published in June. The protocol consists of messages to associate a mobile node with a mobility anchor in the wired backhaul network, called a Localized Mobility Anchor (LMA), when the mobile node first comes on the network. The protocol also provides messages that allow a Mobile Access Gateway (MAG) located at the access router to move the forwarding end point for the mobile node from one access router to another as the mobile node moves across the localized mobility management domain. The working group gave feedback to the design team on the document, recommending that the design team try to simplify the protocol.

The working group also discussed a draft describing the mobile node to access router interface. Prior to the Montreal meeting, the working group was operating on a type of addressing model for the mobile node to access router interface called "multilink subnets", in which the access routers all advertise the same IPv6 subnet prefixes on their wireless interfaces while continuing to function as routers on their wired interfaces. In the past, this model was discussed and rejected by the IETF. Currently, the working group is conducting discussions on its mailing list to determine the type of addressing model it will support. Outcomes are expected by early August.

The protocol design team plans to continue meeting for two months and to issue a final draft in the middle of September. The working group is planning an interim meeting at the end of September to discuss issues with the design. It hopes to have the protocol draft ready for working group last call by early October.

CAPWAP Makes Progress

The Control and Provisioning of Wireless Access Points (CAPWAP) working group, which operates in the Operations and Management Area, is developing a protocol for control and provisioning of wireless access points. The initial target wireless protocol is 802.11, which is designed to allow a centralized Access network Controller (AC) to control a collection of access points, called Wireless Termination Points (WTPs), in the CAPWAP architecture. The AC performs such functions as power management, load balancing, and security, functions that are difficult to perform on the individual access points because they require co-ordination. The working group recently published two RFCs as follows:

RFC 4564: describes objectives for the CAPWAP protocol

RFC 4565: describes a protocol-design evaluation to determine which of several candidate drafts the working group should use as the starting point for the standard

The working group is planning to complete the protocol and MIB by June 2007. The chairs are interested in determining whether there are any implementations of CAPWAP underway, and in organizing an interoperability test.

At their meeting at IETF 66, the editors of the CAPWAP protocol specification discussed issue resolution on the document. One of the main issues was the replacement of the original CAPWAP security protocol with datagram transport layer security (DTLS). The protocol document has now been edited to reflect the design change. Use of DTLS will ensure that CAPWAP would benefit from fixes of any flaws discovered in the DTLS protocol by other applications. The working group decided that the first standardized version of the protocol will reflect only changes to the 802.11 protocol that are incorporated into the 802.11ma specification. 802.11ma is an update of the 802.11 specification, which is due out next year and which will include all of the amendments (such as 802.11i, 802.11e, etc.) that have been put in place since the original 802.11 specification was published in 1999.

The primary topic of discussion at the meeting was a proposal to multiplex control and data traffic between the WTPs and the AC on a single UDP port. This would cause all CAPWAP control messages in addition to all data traffic from the 802.11 terminals to go through a single port on the AC. The reason for this proposal is that it would simplify network-address-translation (NAT) keepalives for deployments in which NATs are positioned between the AC and WTPs. The keepalives on the control traffic between the WTPs and the AC serve to also keep the NAT bindings for the 802.11 terminal data channels active in case the terminals go dormant. The main argument against this proposal is that it would limit scalability. All traffic for terminals on the WTPs would need to go through a single port on a single AC. This would put a reduced upper limit on the number of WTPs that an AC could support, compared with the case in which multiple ports are used. In addition, existing middleboxes between the WTPs and the AC won't recognize the multiplex header, which could cause problems. A consensus call made earlier this year resulted in an almost even split in opinion among working group members. Currently, the working group is undergoing a rather heated debate about this issue, and the chairs have requested that the IESG provide a designated domain expert to provide input.

Impressions of Two IETF Newcomers

By Alain Aina and Michuki Mwangi

The ultimate achievement for a technical engineer is the opportunity to participate at the highest level of Internet development, perhaps even serve as a co-author of an RFC. To most Internet engineers, the IETF is a revered organisation and involvement is regarded as a career high. We felt that our participation at the IETF meeting in Montreal was both a personal achievement and a motivational experience. It gave life to a process that we had experienced only on mailing lists. Having had the opportunity to be at the meeting, we were able to appreciate the passion and the energy that are put into the IETF for the good of the Internet.

The 66th IETF meeting was held at the Palais des congrès in Montreal, Canada. The conference facility was large enough that with a total attendance of 1,257, it was difficult to comprehend a meeting of this magnitude. In Africa, most ICT-related meetings do not draw large numbers of participants, except for the WSIS, ITU, and ATU meetings. However, only at the IETF plenary meetings and the breakfast sessions could one appreciate the sheer number of participants. During the plenary, for example, the wireless network was challenged as a result of the large concentration of people in one room at one time. Nonetheless, being newcomers to the meeting, we must admit that the level of organisation was exceptional despite the numbers. It’s no wonder that the IETF budget runs into the thousands of dollars.

Due to our keen interest in the DNS and IPv6, which is a result of our involvement in the African ccTLD and Internet Registry arena, we were interested in attending the working groups on the Internet operations and routing areas. Some of the issues concerning the deployment of DNS-SEC and IPv6 for our region were of particular interest to us. We learned that the Kenyan (.ke) and the Senegalese (.sn) ccTLD Registries have formal plans to commence DNSSEC trials in the near future. We also learned that AfriNIC, the African Regional Internet Registry, is currently undergoing IPv6 training in the region, with the aim of creating the necessary awareness and expertise for deployment. However, in order to appreciate the protocols functionalities, involvement in IETF discussion groups has helped unearth and clarify the challenges faced by those involved in deployment. By attending the IETF meetings, the reality of the issues is made even clearer through the deliberations on their impacts at the Working Group sessions. Of interest were the discussion on the AS112 draft and the DNS reflector attacks drafts that bring to the fore operational concerns as they apply to the DNS. The two drafts have proposed implementation recommendations that are, in our opinion, worth consideration. Initiating discussions within our region on these two drafts seems like a fair starting point for generating sufficient interest in the IETF activities.

Finally, we noticed a large number of participants from the Asian region and were disappointed to see that, other than the two of us, there were no African participants. Increasing participation at IETF meetings from among African nations will be challenging and possible only through increased awareness of the meetings' activities and role. A similar issue was raised at the plenary meetings in regard to the location of future IETF meetings. There were varied opinions as to why considerations should or should not be given to hosting meetings in developing countries and regions. Ultimately, it was felt that hosting the IETF meetings in a region that draws many participants was of more value than hosting a meeting in a location where there would be little participation. Unfortunately, if that was the primary criterion, it would virtually eliminate any possibility of hosting an IETF meeting in Africa. This makes it even more of a challenge to the communities in those regions to become active contributors to the future of Internet protocols and standards development, and not just consumers of the Internet.

We wish to take this opportunity to thank ISOC for making our participation at the IETF 66 meeting possible. Further, we wish to thank our mentors, Joe Abley, Jaap Akkerhuis, John Crain, Lucy Lynch, Frederico Neves and ISOC staff members Mirjam Kühne and Matthew Shears for ensuring that we settled in without much ado.

Alain Aina at IETF66
Photo: Michuki Mwangi

News from the IRTF

By Aaron Falk, IRTF Chair, and Mirjam Kühne

During the Technical Plenary at IETF 66, Aaron Falk, chair of the Internet Engineering Task Force (IRTF), gave a very informative update of some recent developments at the IRTF. Some highlights follow.

Aaron Falk
IRTF Chair

OFFPATH

A BoF session on path-decoupled signalling for data (OFFPATH) was held to discuss creation of a research group (RG) on signalling between end-systems and components in the network, such as firewalls. The purpose of the group is to develop a flexible framework around a simple protocol. Initial work includes a SIP-based implementation from Cornell University.

Delay-Tolerant Networking (DTN) RG

The DTN RG met in Berkeley, California, in May in addition to conducting a review with the IAB. New work has begun on bundle-in-bundle encapsulation and on defining new bundle headers/blocks.

End-to-End (End2End) RG

The End2End RG is planning a meeting in summer 2006 to discuss, among other things, the re-evaluation of the state of work on congestion control. The re-evaluation will be coordinated with the Internet Congestion Control (ICC) RG.

Host Identity Protocol (HIP) RG

The HIP RG met at the past two IETF meetings. The HIP over NAT problem statement is currently in the RFC Editor queue. The experimental report is progressing. Internet-Drafts on the following topics are in preparation:

- Simultaneous multi-access
- Service discovery
- TCP piggybacking

Internet Measurement (IM) RG

The IM RG is considering two workshops: one on techniques for application identification and the other on IM RG bandwidth estimation, at which the RG will determine whether the techniques are ready for standardisation.

IP Mobility Optimization (MobOpts) RG

The document on “Route Optimization Enhancements” will be published as the first IRTF RFC. The RG is further investigating the effect of mobility on transport protocol performance, with measurements on operational networks as well as simulation. The RG has access to a testbed and has developed a draft on Layer-2 abstractions for Layer-3 handover.

Internet Congestion Control (ICC) RG

In May, Michael Welzl joined S. Keshav as co-chair of the ICC RG. A wiki page has been developed and the RG is now working on an online bibliography. There has also been discussion of one or more “survey” documents. The surveys will catalog the IETF RFCs on congestion control in addition to experimental congestion control protocols.

Network Management (NM) RG

The Network Management RG met during IETF 66 to discuss SNMP trace collection and analysis. A workshop is planned to identify network-management research challenges and to draft a five-year research agenda.

Routing Research Group (RRG)

The Routing Research Goup met in Barcelona in April during InfoCom and is currently searching for a new co-chair. The documents “Routing Requirements” and “History of Routing Requirements” are awaiting final updates. The RG is planning to meet at the IETF in November in San Diego.

Scalable Adaptive Multicast (SAM) RG

The SAM RG had its first meeting at IETF 66 in Montreal. A couple of Internet-Drafts are in preparation (see http://www.samrg.org/bib for a full bibliography). SAM RG co-chair John Buford gave a short review of IP Multicast and explained why it is needed. For example, multicast achieves bandwidth savings over unicast. Certain applications, such as real-time video-streaming, are difficult or impossible to deploy without it. The goal of the SAM RG is to enhance the benefits of multicast by offering flexible and incremental deployment options.

With three different types of Multicast available - Application Layer Multicast (ALM), Overlay Multicast (OM) and Hybrid approaches - the group is attempting to create a unified framework that enables interoperability of different multicast protocols based on network, traffic, and group properties. The group is hoping for a dynamic transition between protocols and mechanisms to optimise performance. There are, however, challenges with this approach. For example, determining multicast support by region may require significant awareness of the network topology. For another example, constructing trees across regions will require mapping between different protocols for tree construction and group membership.

The RG has prepared two Internet-Drafts. As next steps the SAM RG will prepare a problem statement and driving scenarios, requirements for a SAM framework, and a survey of ALM/OM/Hybrid technologies and performance metrics. For more information, see the SAM web site: http://www.samrg.org

New Tools Enhance Meeting Efficiencies

By Henrik Levkowetz, Chair of the Tools Team, and Mirjam Kühne

In an effort to enhance IETF participants' meeting preparations, the IETF Tools Team has made a few different tools related to the IETF meetings avalable during the first part of this year. The working group pages under http://tools.ietf.org/wg now provide HTML versions of all the WG agendas, including links to uploaded slides. This should provide a one-stop-shop for everything associated with an individual WG meeting, and has been made possible by the early access to presentations which the Secretariat's new materials upload tool provides.

In addition, the overall meeting agenda at http://tools.ietf.org/agenda has been enhanced so that IETF meeting attendees are able to view the layout of the meeting venue online and locate WG meeting rooms by clicking on the room number next to the WG meeting time.

IETF Meeting Calendar Generator (http://tools.ietf.org/calendar) Here you willl find agenda for upcoming IETF meeting, complete with check boxes for marking sessions you want to attend. An individualised calendar file for the week of the meeting will be created from this information, which can be downloaded and included in your preferred calendar application. The calendar tool has been tested using iCal, Outlook and Google calendar).

A future addition of the Calendar Generator will allow you to 'change your mind' when creating your individual IETF meeting calendar. After opening http://tools.ietf.org/calendar, you will be presented with your previous choices and given the opportunity to change your selections.

Other tool news, not directly related to the meetings:

Searching for Documents by Name. On the left margin of the tools pages at http://tools.ietf.org/tools/ you will find now a search function that allows you search for any string (including the RFC number) in the title of an RFC or Internet-Draft. The results will offer html versions of the matching documents.

Later this summer the Tools Team hopes to release a ***Notification Service*** ('send me e-mail when this draft or charter changes'), which is intended to make it easier to keep track of changes in documents or WG charters. This tool will provide a selective notification mechanism for general use, complementing the IETF announcement mailing lists. It will include RSS and ATOM feeds from the available XML meta-information about Internet-Drafts, RFCs, and WGs. This format will make it possible for individuals and tool-builders to better interface with information from the IETF standards process in a well-defined manner. Over time, the notification tool will produce a complete history of document and charter changes, WG agendas, and minutes.

The Tools Team is always interested in feedback about current tools or any wishes you may have. The team can be reached at tools-discuss@ietf.org.

A full list of all chartered work items and their status can be found on slides presented by Henrik Levkowetz at http://www3.ietf.org/proceedings/06jul/slides/plenaryw-3.pdf

TCP/IP 25th Anniversary

By Brian Carpenter, IETF Chair

The foundation of the Internet is composed of the basic Internet Protocol (IP) and the generally used Transmission Control Protocol (TCP), which together are known as TCP/IP.

TCP/IP was formally standardised in September 1981 - 25 years ago - by the publication of RFC 791 and RFC 793.

The editor of those documents was the late Jon Postel, then of the Information Sciences Institute (ISI) at the University of Southern California. Postel indicated on the original documents that they were

prepared for

Defense Advanced Research Projects Agency
Information Processing Techniques Office
1400 Wilson Boulevard
Arlington, Virginia 22209

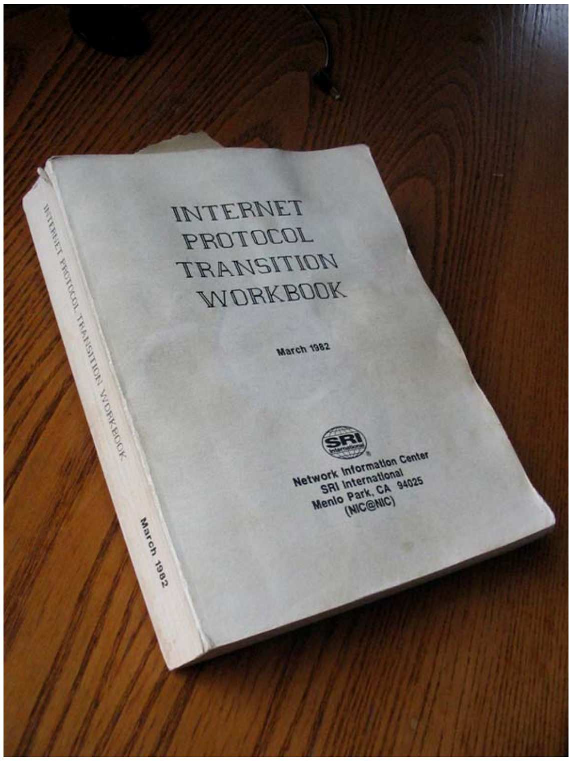
by

Information Sciences Institute
University of Southern California
4676 Admiralty Way
Marina del Rey, California 90291

While Vint Cerf and Robert Kahn were widely credited with the design of TCP/IP, many others involved in the ARPANET project made significant contributions.

According to Vint Cerf, "The core of the documents was RFC 675 from December 1974 authored by Carl Sunshine, Yogen Dalal and me. The subsequent sequence of documents leading up to RFC 791 and 793 had the benefit of quite a few hands, including the participation of Dave Clark, Jon Postel, Bob Braden, Ray Tomlinson, Bill Plummer, Jim Mathis, as well as other early implementers of TCP." Of course, at the time, many other unnamed contributors who participated in the debate.

Since the RFC series was launched in 1969 by Steve Crocker at UCLA, it has continued as the public archive of the Internet's fundamental technology. Since 1977 it has been hosted by the Information Sciences Institute of USC. ARPA funding

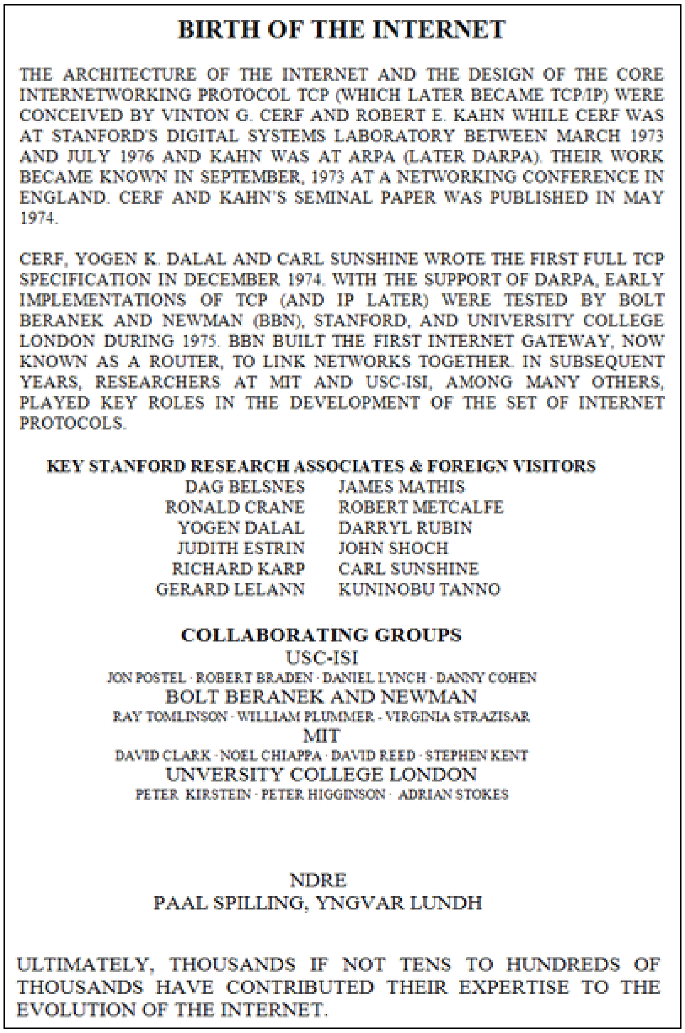

INTERNET PROTOCOL TRANSITION WORKBOOK
March 1982
SRI
Network Information Center
SRI International
Menlo Park, CA 94025
(NIC@NIC)
INTERNET PROTOCOL TRANSITION WORKBOOK
March 1982

Original 1981 IP Transition Workbook
Photo: Daniel Karrenberg

ended in 1998, at which time ISOC took over, as its first major funding effort for Internet standards. Since the end of 2005, the RFC Editor has been supported by the IETF Administrative Support Activity, which is hosted and partly funded by the Internet Society.

The long-serving RFC Editor, Jon Postel, passed away in 1998. His close colleague throughout all these years Joyce Reynolds said: "Operating systems and computers have changed over the years, but Jon's perseverance about the consistency of the RFC style and quality of the documents remained true." Many friends and colleagues remember him at `http://www.postel.org/remembrances/`

"We can't yet say that the Internet is mature, “says Brian Carpenter, chair of the IETF, “but it's a great tribute to the pioneers that the two most basic specifications that were published a quarter of a century ago are still largely valid today. I hope the IP version 6 standard will do as well."



BIRTH OF THE INTERNET

THE ARCHITECTURE OF THE INTERNET AND THE DESIGN OF THE CORE INTERNETWORKING PROTOCOL TCP (WHICH LATER BECAME TCP/IP) WERE CONCEIVED BY VINTON G. CERF AND ROBERT E. KAHN WHILE CERF WAS AT STANFORD'S DIGITAL SYSTEMS LABORATORY BETWEEN MARCH 1973 AND JULY 1976 AND KAHN WAS AT ARPA (LATER DARPA). THEIR WORK BECAME KNOWN IN SEPTEMBER, 1973 AT A NETWORKING CONFERENCE IN ENGLAND. CERF AND KAHN'S SEMINAL PAPER WAS PUBLISHED IN MAY 1974.

CERF, YOGEN K. DALAL AND CARL SUNSHINE WROTE THE FIRST FULL TCP SPECIFICATION IN DECEMBER 1974. WITH THE SUPPORT OF DARPA, EARLY IMPLEMENTATIONS OF TCP (AND IP LATER) WERE TESTED BY BOLT BERANEK AND NEWMAN (BBN), STANFORD, AND UNIVERSITY COLLEGE LONDON DURING 1975. BBN BUILT THE FIRST INTERNET GATEWAY, NOW KNOWN AS A ROUTER, TO LINK NETWORKS TOGETHER. IN SUBSEQUENT YEARS, RESEARCHERS AT MIT AND USC-ISI, AMONG MANY OTHERS, PLAYED KEY ROLES IN THE DEVELOPMENT OF THE SET OF INTERNET PROTOCOLS.

KEY STANFORD RESEARCH ASSOCIATES & FOREIGN VISITORS

DAG BELSNES
RONALD CRANE
YOGEN DALAL
JUDITH ESTRIN
RICHARD KARP
GERARD LELANN

JAMES MATHIS
ROBERT METCALFE
DARRYL RUBIN
JOHN SHOCH
CARL SUNSHINE
KUNINOBU TANNO

COLLABORATING GROUPS

USC-ISI
JON POSTEL · ROBERT BRADEN · DANIEL LYNCH · DANNY COHEN
BOLT BERANEK AND NEWMAN
RAY TOMLINSON · WILLIAM PLUMMER - VIRGINIA STRAZISAR
MIT
DAVID CLARK · NOEL CHIAPPA · DAVID REED · STEPHEN KENT
UNVERSITY COLLEGE LONDON
PETER KIRSTEIN · PETER HIGGINSON · ADRIAN STOKES

NDRE
PAAL SPILLING, YNGVAR LUNDH

ULTIMATELY, THOUSANDS IF NOT TENS TO HUNDREDS OF THOUSANDS HAVE CONTRIBUTED THEIR EXPERTISE TO THE EVOLUTION OF THE INTERNET.

Stanford Plaque commemorating early work done on TCP/IP in 1973-1975
Courtesv of Vint Cerf

DNS Security: A Historical Perspective

By James M. Galvin

Note: This article does not attempt to provide a complete summary of all IETF activities in this area. It reflects the author's personal perspective on some current highlights.

DNS security work began to organize as its own activity in 1993. I do not remember when the first conversations took place, but we met for the first time as a sub-group of what was then the DNS working group during the 28th IETF, November 1993, in Houston, Texas. Today's IETF would call that meeting a BoF, since its principal objectives were to evaluate interest and commitment, and to develop a charter for its own working group. It has been 13 years since those early days and DNS security has undergone many changes since then.

I was the chair of the DNS Security Working Group, which did not conclude until 1999. However, the conclusion of the working group did not end work on DNS security. The consensus of the participants at the time (and of the IESG, of course) was that the work should be continued more directly by DNS experts, where it has continued to this day.

Today, some people wonder whether we will ever be done with DNS security. I believe it has become a slave to changing requirements and an evolving Internet.

The DNS Security working group was first chartered in what was the Service Applications Area with Dave Crocker serving as Area Director. The November 2003 IETF reported the following summary for the DNS Working Group meeting.

> The DNS Security sub-group of the DNS working group met to identify the threats, security services, and requirements of interest to the DNS. The requirements will be distributed to the mailing list for discussion until November 30, 1993. After that time, strawman proposals may be distributed until January 31, 1993. The group will evaluate all proposals with the goal of creating one proposal at the next IETF.
>
> It was decided to create a DNS security working group. In parallel with the activities above a charter will be drafted for review and submission to the IESG.

The working group was officially chartered in March 2004 with the following description.

> The Domain Name System (DNS) security working group (dnssec) will specify enhancements to the DNS protocol to protect the DNS against unauthorized modification of data and against masquerading of DNS data origin. That is, it will add data integrity and authentication capabilities to the DNS. The specific mechanism to be added to the DNS protocol will be a digital signature.
>
> The digital signature service will be added such that the DNS resource records will be signed and, by distributing the signatures with the records, remote sites can verify the signatures and thus have confidence in the accuracy of the records received.
>
> There are at least two issues to be explored and resolved. First, should the records be signed by the primary or secondary (or both) servers distributing the resource records, or should they be signed by the start of authority for the zone of the records. This issue is relevant since there are servers for sites that are not IP connected. Second, the mechanism with which to distribute the public keys necessary to verify the digital signatures must be identified.

> Two essential assumptions have been identified. First, backward compatibility and co-existence with DNS servers and clients that do not support the proposed security services is required. Second, data in the DNS is considered public information. This latter assumption means that discussions and proposals involving data confidentiality and access control are explicitly outside the scope of this working group.

There are two elements of the first summary and the first charter that are important to an understanding of the history of DNS security. First, one of the greatest mistakes we made in those early days was failing to document the actual threat discussions that led to the selection of security services to be added to the DNS protocol. At the time it seemed pretty obvious and straightforward. The scope of work was limited and we estimated we would be done in about one year (an estimation that, unfortunately, has become the DNS Security mantra). In reality it would take more than two years until the first version of the work was completed, resulting in RFC2065 – Domain Name System Security Extensions by Donald Eastlake and Charlie Kaufman – being published in January1997, almost three years from when the working group was first chartered.

Failing to document the threat analysis was wrong for at least two reasons. First, security experts will tell you that adding security without understanding why is like “putting the cart before the horse.” A threat analysis provides a careful study of what is at risk and what needs to be protected. Although it is possible that the lack of this analysis could have been tolerated initially, DNS security is no longer the simple DNS extension it was once imagined it could be. Such a document would have served as an important baseline for reviewing future extensions. Second, with respect to the question of when DNS security will be done, the threat analysis would have established clear goals against which the DNS security specification could have been evaluated. Although a prologue was published in August 2004, the informational RFC3833 “Threat Analysis of the Domain Name System (DNS) by Derek Atkins and Rob Austein,” it has not served this purpose. The conclusion from RFC3833 states:

> Based on the above analysis, the DNSSEC extensions do appear to solve a set of problems that do need to be solved, and are worth deploying.

The document carefully, and probably wisely, does not judge whether the solved problems were the correct problems to solve or whether the solutions are sufficient. Thus, rather than declare "success" for the DNS security work its primary role was to end the almost 10 years of repeating discussions of why the protocol does what it does. Of course, the significance of this should not be underestimated since it has facilitated more focused effort on the issues that do need attention.

The second element of the original charter worthy of special notice is the assertion that data in the DNS is public information. The extant intent of this statement was, as stated in the charter, to ensure that confidentiality and access control services were not considered by the working group, although in principal it is obvious that the information in the DNS is public. The primary purpose of the DNS is to map domain names to IP addresses to facilitate communication between two sites. If the information is not available or is inaccessible then the sites will not be able to communicate. Unfortunately, the assertion later conflicted with a business practice requirement: preventing the transfer of the entire contents of a zone.

Although the data in the DNS must be available to be useful, in ordinary circumstances the DNS protocol inherently limits how quickly any client can access all the data in a zone. If a client knew all the domains in a zone it could query for the data available for each domain individually. Since the label for each domain in a zone could be as long as 256 characters, a brute-force search of the zone for valid domains is impractical. A protocol element for transferring the entire contents of a zone is available but all popular DNS server implementations include mechanisms that restrict access to this functionality. The result is that the entire contents of a zone is frequently unavailable to most clients.

Adding DNS security added functionality that had not previously been present in the DNS. Specifically, if a client queried for a non-existent domain, the response would correctly and securely assert that the domain did not exist but, in addition, it would indicate the lexicographically next valid domain in the zone. Through repeated queries, a client could discover and download the entire contents of a zone. This feature (or mis-feature depending on your point of view) has come to be known in technical circles as "zone walking." The requirement to prevent zone walking has become a gating factor in the deployment of DNS security, particularly with larger zones. Significant technical resources over several years have been focused on this issue. A solution that is approaching broad consensus includes the use of OPT-IN and NSEC3. A second interoperability event is scheduled for the fall of 2006. If it is successful (and all indicators are that it will be) we may see publication of the solution in early 2007.

Moving on, as we neared the end of the first version of the DNS security extensions, dynamic update was getting attention from the DNS community. RFC2065 did include limited coverage of dynamic update issues, but ultimately, the security work for dynamic update was left as a follow on activity of the DNS Security Working Group. Our charter was updated in March 1996 to include dynamic update, as well as a few other technical issues. Of particular note is the fact that the working group moved into the Security Area with Jeff Schiller as Area Director with this update to its charter.

RFC2137 – Secure Domain Name System Dynamic Update by Donald Eastlake – was published in April 1997. Along the way RFC2065 was updated according to implementation and operational experience from developers and early adopters. RFC2535 – Domain Name System Security Extensions by Donald Eastlake – was published in March 1999.

After completion of the issues outlined in the second charter, it was time once again to consider the status of the working group. There was still work to be done. The zone-walking problem had not been resolved and there was a need to provide transaction level authentication by using shared secrets and one-way hashing in the form of TSIG (transaction signature), first published as RFC2845 in May 2000. An obvious choice would have been to update the charter accordingly and press on. However, there were two other IETF changes to consider.

The DNS Security working group was part of the Security Area. At the time its charter was updated, most security work was being done in the Security Area and it was generally accepted that this was a good thing. More recently, there has been more discussion of the question of whether the protocol work was principally about security or whether security was a component of the protocol work to be done. In this regard, the core DNS security work was arguably complete. Implementations were in progress and the work to be done was either an extension or a new requirement based on operational DNS experience. When security work was a component of the protocol work to be done, as it now was with DNS security, there was some preference for the work to progress primarily with those experts. Thus, one change was a suggestion to continue the work in the Internet Area with the DNS work.

Second, at the same time, the DNS IXFR, Notification, and Dynamic Update (DNSIND) working group, which was the only DNS working group at the time, was nearing completion of its charter's stated goals. The question under consideration was whether to create separate working groups for the ongoing work items or to create a working group to manage the work items.

In the spirit of the IETF, a few "hallway" conversations between the IESG, working group chairs, and other interested parties resulted in a proposal to create the DNS

Extensions (DNSEXT) Working Group to manage DNS related work items. The DNSSEC and DNSIND working groups concluded in December 1999 and January 2000, respectively, coincident with the chartering of the DNSEXT working group.

DNS security work began anew with early deployment experiments of RFC2535 by the Swedish and Dutch top-level domain operators, NLnet Labs, and RIPE NCC. They discovered operational problems with the key exchanges between the DNS parents and children. This was one of the principal issues that resulted in a major rewrite that became three specifications - RFC4033, RFC4034, and RFC4035 – published in March 2005. Unfortunately, the zone walking problem had still not been resolved, although this time the working group committed to solving the privacy problem.

During the last few months of the rewrite a few large top-level domain registries came to realize and asserted that the non-requirement for privacy would prohibit the deployment of DNSSEC in their environment. After careful consideration of the issue, the working group believed that a solution could be added after a deployed base of the rewrite existed. Thus, rather than delay the specifications any longer they were published so the working group could lend some focus to the zone walking problem.

DNS security work continues today under the auspices of the DNSEXT working group. Updates have progressed along with some new work. Zone walking will hopefully be resolved soon. Unfortunately, we are still not done but "we will be done soon," or so the story goes. Finally, early adopters are deploying the protocol and the DNS Security Extensions Deployment Intiative http://dnssec-deployment.org working to encourage voluntary adoption of DNS security protocols as part of a global effort to improve the security of the Internet's naming infrastructure. We are much closer to declaring success now than we have been in the 10 years since the first specification was published.

It is worth noting that over the years there has been a shift in the type of people who have worked on DNSSEC development. It started with security people, moved to DNS protocol experts, and finally more operationally inclined experts joined the effort to get their concerns addressed. Each group had its own requirements, and the DNSSEC specification changed accordingly. DNSSEC, like all IETF protocols, is a slave to the members of the working group who have responsibility for it. Although such evolution is arguably rational, perhaps some of the past 13 years could have been spared if more of us could have been working together sooner, rather than one after another. It would be useful to analyze whether and how the IETF could have worked more efficiently. No protocol should ever take more than 13 years to be deployed.

[Editor’s note: For a more detailed technical description of DNSSEC, see the recent publications at http://ispcolumn.isoc.org/]

An Overview of Multihoming and Open Issues in GSE[2]

By Lixia Zhang

Abstract

This draft has three objectives: (1) to discuss the impact of multihoming on the scalability of the global routing system; (2) to provide an overview of GSE, one of the early proposals by Mike O'Dell to address the multihoming scalability problem; and (3) to identify open issues raised by the GSE proposal, which may serve as a first step toward resolving them.

1. Introduction

In its original design IPv4 had a class-based address structure that divided the 2^32 address space into 2^7 large networks (Class-A), 2^14 medium size networks (Class-B), and 2^21 small networks (Class-C)[3]. Each network is represented by a Network ID, also called a network prefix, with the length of 8 bits, 16 bits, and 24 bits for Class A, B, C networks, respectively. Global routing was performed by matching the high order bits of the packet destination address against a table indexed by network prefixes. Each prefix took one entry in the global routing table and the length of the prefix was implied by the address class.

The explosive growth of the Internet during early 1990's brought serious scalability problems to the Internet routing infrastructure: there were too few Class-A address blocks to give out; Class-B blocks were nearly exhausted; and as a result a large number of Class-C blocks were assigned. Because each Class-C network has only 256 addresses, one institution might have to get multiple Class-C address blocks. Since each network ID takes one entry in the global routing table, the table started growing at an alarming rate, until Classless Interdomain Routing (CIDR) was deployed [RFC4632].

At the time it was deployed, CIDR provided an effective way to slow down the growth of the routing table in the Internet backbone, commonly referred to as the Default Free Zone (DFZ). 15 years after CIDR's deployment, however, today's global routing system is facing serious scaling problems again. A rough estimate from the weekly CIDR report [1] shows that the IPv4 DFZ routing table size has gone up by about 36% since September 2004 and doubled since January 2001. The rate of growth also seems to be accelerating over time and, if the current acceleration rate is maintained, the DFZ routing table size would double again in about early 2010[4]. What is the main cause of the rapid routing table growth this time? The problem appears to be customer multihoming and traffic engineering.

The multihoming induced routing scalability problem has long been recognized, and a number of recent IETF efforts have been dedicated to the development of solutions to the problem [2,3]. This draft is intended to help the reader fully understand the importance of the problem, and to describe some alternative solutions in the design space. We first describe the relation between edge

[2] GSE stands for *G*lobal, *S*ite, and *E*nd-system address elements.

[3] In addition, a block of 2^28 addresses was assigned to multicast address, and another 2^28 block was reserved.

[4] On the surface it seems that the Moore's law should be more than adequate to handle the DFZ routing table growth rate. One should understand, however, that routing scalability is a multi-dimension issue, a large table size brings a number of problems in other dimensions which, unfortunately, need to be explained in another article.

multihoming and traffic engineering practice and DFZ routing scalability. We then describe an early proposal, GSE by O'Dell from 1997, and show how it works to resolve the multihoming scalability problem. We also identify some of the open issues that must be resolved before GSE, or similar proposals, can be deployed in practice.

2. Impact of Multihoming on Routing Scalability

The basic idea behind CIDR is simple: the size of an IP address block is allowed to be 2^n, where 0 <= n <= 32. This simple idea helped slow down routing table growth in two ways. First, each organization needs only one address block of the right size, as opposed to multiple Class-C blocks in pre-CIDR days. Second, and perhaps more important, CIDR allows an Internet Service Provider (ISP) to divide an allocated address block into multiple pieces of potentially different sizes, and to assign each piece to a customer according to its need. Each IPv4 address block allocated to an ISP typically has an address prefix 8-21 bits long. The address block allocated to a customer is represented by a prefix longer than the ISP's prefix, with the high order bits being the same as the ISP's prefix. The ISP can announce the prefix of its allocated address block to the global routing system and receive data traffic destined to all of its customers, as long as none of the longer prefixes assigned to individual customers are announced separately. The ISP then distributes the traffic to its customers according to their individual address prefixes. Thus CIDR enables an ISP to support many customers while still announcing only one aggregated prefix to the global Internet. In an ideal CIDR case, the number of routing table entries should be around the same order of magnitude as the number of ISPs. However, in reality, the former has always been much larger than the latter, since each ISP tends to have multiple allocated address blocks, and more important, there exist a large number of provider-independent (PI) prefixes; many of these are legacy allocations that predate the introduction of CIDR.

PI prefixes are the address blocks allocated to customer networks directly. The important property of a PI prefix is that its owner has the freedom to switch providers without renumbering the network. Furthermore, a network with a PI prefix can connect to multiple ISPs simultaneously. This is known as multihoming, which allows the network to stay reachable through whichever providers remain functional when some part of the Internet fails. As Renesys' measurement of the 2003 US East Coast blackout shows, well engineered multihoming can be an effective way to ensure Internet connectivity [4]. In the absence of network failures, a multihomed site can distribute outbound traffic across multiple provider connections to maximize some locally defined goals such as cost, throughput, and/or performance. If routing policy permits, a customer may also subdivide its address allocation, that is, split its prefix into multiple longer ones that are then used for load-balancing the incoming traffic, as shown in Figure-1 below.

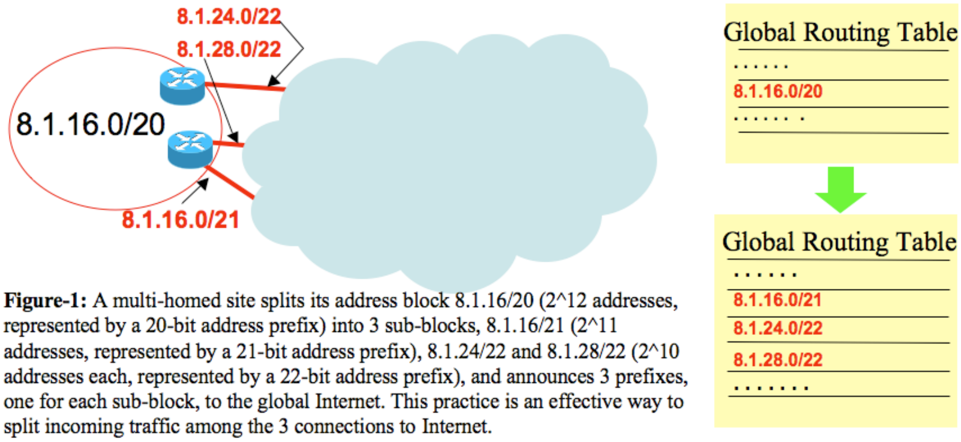


Figure-1: A multi-homed site splits its address block 8.1.16/20 (2^12 addresses, represented by a 20-bit address prefix) into 3 sub-blocks, 8.1.16/21 (2^11 addresses, represented by a 21-bit address prefix), 8.1.24/22 and 8.1.28/22 (2^10 addresses each, represented by a 22-bit address prefix), and announces 3 prefixes, one for each sub-block, to the global Internet. This practice is an effective way to split incoming traffic among the 3 connections to Internet.

The aforementioned advantages of multihomed sites, however, come at the cost of one or possibly multiple entries per site in the global routing table. During the early days of CIDR deployment, the number of customer networks was relatively small, few were multihomed, and most of them got address assignments from their ISPs. Thus CIDR aggregation worked out well. However over time more and more customer networks became multihomed for improved Internet availability and performance. Our recent measurement results indicate that today the majority of customer networks are multihomed [5].

Such pervasive multihoming practice has made a profound impact on the scalability of the current routing and address architecture. Being reachable through any of its providers implies that a customer network must be visible in the global routing table, that is, it must announce a PI prefix, or otherwise make its providers announce a specific prefix for it[5]. Moreover, if a site wants to load-balance incoming traffic, it may also split its prefix into multiple longer ones and announce them to different ISPs. Consequently, both of CIDR's advantages mentioned earlier, one address block per customer site and ISP aggregation of customer prefixes, are lost through current multihoming and traffic engineering practices.

A number of people foresaw the routing scalability problem resulting from multihoming and proposed solutions. Below we describe GSE, one of the earliest proposed solutions suggested by Mike O'Dell in 1997.

3. GSE: An Alternate Addressing Architecture for IPv6: How It Works

The proposed IPv6 address structure inherits from IPv4 the CIDR-style "Provider-based Addressing". Recognizing CIDR's intrinsic limitation in the presence of multi-homed sites, O'Dell proposed to divide IPv6's 16-byte address into three parts, with the lower *N* bytes being the End System Designator (ESD), the middle *M* bytes representing site topology partition (STP) for local routing, and the top (*16-M-N*) bytes being *Routing Goop*[6], or RG, to be used for routing between providers. A Routing Goop signifies where a site attaches to the Global Internet, and a multihomed site will have multiple RGs, one for each of its providers. As the site changes providers, its RGs change but not the remainder of the address structure. When a packet flow moves from one provider connection to another, the RGs in the

[5] Assuming provider **P1** makes a single *aggregated* prefix announcement *P* for multiple customers. If one of the customer networks, *C,* with a **P1** assigned prefix $P_C \subset P$ multihomes with another provider **P2**, **P2** will announce reachability to prefix P_C. This in turn will force **P1** to announce P_C as well - that is, *de-aggregating* its routing announcements. Otherwise it would not get any traffic going to *C*.

[6] *Goop* is American slang for a messy but useful substance.

packets' addresses change as well. Therefore GSE requires that transport and all of the higher level protocols use the ESD portion, instead of the whole IPv6 address as connection identifiers.

The fundamental novelty in the GSE design is to hide a site's RG from its internal hosts and routers, so that they are *insulated* from the external topological connectivity and such changes as multihoming or re-homing (that is, changing providers). This insulation is implemented through the following steps as shown in Figure-2. (1) When generating a packet, the source host fills the destination address with a complete 16-byte IPv6 destination address, including the RG, that it receives from DNS resolution, and fills the upper (*16-M-N*) bytes in the source address with a special "*Site-Local*" prefix[7]. (2) If the destination is not within the local site, the packet will leave the site via one of possibly several site boundary routers, which will insert a proper RG, expected to be used for returning packets of the same end-to-end communication, into the packet's source address. (3) When the packet reaches a site boundary router of the destination network, the router will replace the RG in the destination address with the *Site-Local* prefix. As a result, the internal routers and hosts of a site should never see the value of its own RG.

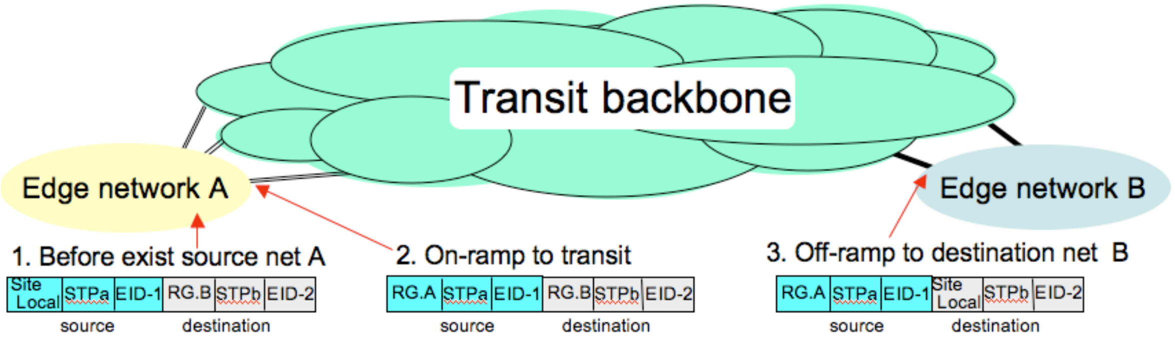

Transit backbone
Edge network A
Edge network B
1. Before exist source net A
Site Local | STPa | EID-1 | RG.B | STPb | EID-2
source
destination
2. On-ramp to transit
RG.A | STPa | EID-1 | RG.B | STPb | EID-2
source
destination
3. Off-ramp to destination net B
RG.A | STPa | EID-1 | Site Local | STPb | EID-2
source
destination

Figure-2 How GSE works.

This insulation provides a site with the flexibility of re-homing and multihoming. Because a site's interior should have no knowledge about the RGs, the site administrator can change providers, and hence change the RGs, whenever needed. At the same time ISPs can also aggressively aggregate RGs as needed for routing scalability.

However, every coin has two sides. Along with its gains GSE also raised a set of new issues that must be fully understood and resolved before it can be put into deployment. In the next section we briefly describe a few of the major ones that have been identified.

Before leaving this section we would like to point out that GSE was not the only proposal in the direction of insulating edge networks from transit providers. In RFC1955 Bob Hinden proposed an ENCAPS scheme that separates providers and customers into two address spaces and uses tunnels to carry packets from source customer networks over the provider space to reach destination customer networks [6]. Here the tunneling plays a role similar that of the RG in the GSE design, hiding the provider space from edge networks.

4. Open Issues in GSE

Before diving into specific open issues in GSE, we would like to stress that the list of issues mentioned in this section is not complete and does not necessarily capture all

[7] "Site-Local" in GSE has no relation to the now-deprecated site-local addresses in the IPv6 specification earlier.

the major ones. Rather we hope that the list can serve as a starting point for future discussions; some of these issues were also mentioned in the GSE proposal [5]. [RFC4218 and RFC4219] provide good sources of information for general threats and considerations in the development of multihoming solutions.

4.1 RGs and DNS Servers

Since hosts learn about destination RGs from DNS lookup, naturally DNS plays a critical role in GSE. One new issue raised by GSE is which RGs to use to reach DNS servers. Even if one may assume that DNS root servers will use host routes that stay relatively stable, other DNS servers may be reachable by using one of multiple RGs. When the hosting sites change providers, the RGs used for reaching the DNS servers also change. Assuming the network hosting one of the example.com DNS servers is multihomed, which one RG or how many RGs should be returned from a DNS server lookup for example.com?

Although GSE strives to insulate a site's internal hosts and routers from RG changes, DNS servers are exceptions. The authoritative DNS servers of a customer site must know the RGs of the site in order to resolve the DNS names for the site, and thus they must be updated with all of the RG changes. Furthermore, whenever a site changes its RGs, all of the DNS servers in the site, both its own and others that it hosts, change their IP addresses. Hence, all of the parents of all of those servers, as well as their owners, must be properly updated.

In addition, GSE also brings up the need for supporting 2-faced DNS. That is, a DNS server must be able to tell whether a query is from a local or remote host, so that it can decide whether to put *Site-Local* or the site's RG(s) in the returned address. For hosts in a multihomed site, the DNS server must also decide which of the multiple RGs to put in the addresses in DNS replies. As we will mention later, one organization may have multiple sites that are interconnected through both a private internal network and the external transit core, thereby adding additional complexity to 2-faced DNS servers.

4.2 The Border Links

As one can see from Figure-2, although GSE insulates edge networks from the transit core, there exist physical links that connect the former to the latter. Let us call them *border links*. On one hand, when packets exit the source site, it is possible to make the source site be aware of the status of its border links and associated routers, so that outbound packets can choose exit routers to avoid any failed border link or router. On the other hand, which border link at the destination end a packet may travel through is determined by the RG in the packet's destination address. In picking a destination RG, the source site has no easy way to tell whether any of the remote edge links may have failed in order to avoid it[8]. The GSE proposal suggested manually configuring all of the routers serving the same site to be aware of each other as a *group*; in case one of the routers loses its connectivity to the site, it can tunnel traffic to the others in the group. Such configuration not only requires close coordination between competing providers, but also must be done for tens of thousands of multihomed edge sites, which posts a big question mark on the feasibility of this proposed solution.

We would like to point out that this issue of handling border link failures is not unique to GSE; the ENCAPS proposal shares a similar problem. In fact any approach in the direction of separating edges from the transit core will find that some special handling is needed to deal with border link failures. Those links along the isolation boundary provide connectivity between the transit core and edge networks. However, they are not covered by routing protocol of the transit core because the

[8] Since sources get the destination RGs through DNS lookup, at least in theory it may be possible to capture the status of the destination edge links through dynamic DNS updates to the destination DNS servers. However the GSE proposal recommended against this approach [2].

edge networks at the other end of those links are now isolated from the core and are no longer routable entities.

4.3 RGs and Tunnels

IP tunneling has been widely used as an simple way to meet various special packet delivery needs. Generally speaking, an IP tunnel can be set up between any two nodes in the same address space. However the GSE design raises new issues in tunneling due to its separation of RGs and the rest of the IP address. In GSE, it is unclear whether tunneling would still be allowed between *any* two IP boxes, or have to be constrained to being between site border routers only. For an IP tunnel across RG boundaries, there are also questions regarding which source and destination RGs should be given to the packets going into the tunnel, and how to handle the packets when they get out of the tunnel and land on a different site.

In light of the extensive use of Virtual Private Networks (VPNs) that has grown up since GSE was proposed, and the use of tunnels at protocol layers below IP, tunneling operations need a thorough examination in the GSE context.

4.4 Traffic Engineering

When an edge network is multihomed, generally speaking it would like to be able to choose exit routers for outbound traffic (*outbound traffic engineering*) and entry routers for incoming traffic (*inbound traffic engineering*). In addition, a transit network may also wish to know how many different paths it has in order to reach a given destination network, so that it can send packets in certain proportion along parallel paths based on some locally defined criteria (*transit traffic engineering*).

GSE was proposed as a scalable way to support site multihoming, but it did not directly address the need for traffic engineering. In particular, the GSE draft mentioned only packets reaching a desired source site exit router, without elaborating on exactly how to direct outbound traffic toward potentially multiple exits. Similarly, the destination RGs are included in the DNS replies, but it is left open as to whether the DNS server, or the sending host, should decide which RG to use among multiple options for inbound traffic engineering. Transit traffic engineering is even more challenging, as a transit network would have no easy way to tell whether packets carrying different destination RGs belong to the same destination site. In short, although it may be possible to enhance GSE for achieving traffic engineering goals, the existing GSE proposal clearly does not solve this problem.

4.5 Other GSE Related Issues

GSE opened a door to decouple edge sites' internal addressing from its connection to the transit core, yet how to take this opportunity to build scalable and robust transit routing operations remains an open issue. In [2] O'Dell sketched out an idea of partitioning the global Internet into a set of tree-shaped regions anchored by "*Large Structures (LS)*". Flat-routing is carried out between LS's and within the regions under each LS. Any two LS's may share a tangency below the top level for "cut-through" paths, but such cut-through paths were considered *controlled circumvention* of otherwise hierarchical paths. Measurement results suggest that, over the past 10 years, the global topology has become more densely connected, and interconnection *below* the top level has become the norm rather than controlled circumventions, suggesting that the originally proposed RG structure and usage may need to be re-evaluated.

Another issue involves routing within large organizations that may have a presence in multiple locations, as well as routing packets between multiple sites of the same organization through the transit core. Each of the sites may be connected through a private internal network, as well as having its own RGs for the connections to the transit core which may also change from time to time. In a GSE setting, how to best utilize both internal and external connectivity for packet delivery between sites seems an entirely open question at this time.

Yet another important issue in GSE deployment concerns the management of End System Designator (ESD) space in order to assure ESD's global uniqueness, as ESDs would be used for end-to-end connection identifications. One must also be prepared to handle ESD collisions in case they occur.

5. A Few Ending Words

It has been nearly 10 years since the GSE proposal was published, yet the problem GSE was set forth to solve is still with us today, and can potentially get much worse when IPv6 starts seeing wide deployment. Despite the IETF's effort in developing multihoming support with provider-allocated addresses [2, 3], regional Internet registries have been under heavy pressure from customers to allocate Provider-Independent IPv6 address blocks, a worrisome sign for IPv6's future routing scalability.

GSE pointed out a brand-new approach to the multihoming support problem. However because it is drastically different from existing practice, at the time it was proposed, a large number of concerns were raised (some of which were captured in [8][9]), and the original proposal was never fully explored to appreciate its advantages, to understand its tradeoffs, and to identify its open issues. In our search for a scalable global routing system design, it seems worthwhile to pay a full revisit to the GSE proposal.

Acknowledgment

First of all, I would like to thank Mirjam Kühne. This article would not have been written without her encouragement and patience. I sincerely thank David Meyer, Brian Carpenter, David Thaler, and other IAB members for their comments. Special thanks go to Elwyn Davies who painstakingly went through an earlier draft and made numerous corrections.

References:

[1] http://www.cidr-report.org/
[2] IETF Site Multihoming in IPv6 Working Group, http://www.ietf.org/html.charters/multi6-charter.html
[3] IETF Site Multihoming by IPv6 Intermediation Working Group, http://www.ietf.org/html.charters/shim6-charter.html
l[4] "Impact of the 2003 Blackouts on Internet Communications", Renesys Corporation, http://www.renesys.com/tech/presentations/blackout_results/, November 2003.
[5] "Observing the Evolution of Internet AS Topology", R. Oliveira et. al., submitted for publication, August, 2006.
[6] "GSE - An Alternate Addressing Architecture for IPv6", Mike O'Dell, http://www.watersprings.org/pub/id/draft-ietf-ipngwg-gseaddr-00.txt, February 1997.
[7] "New Scheme for Internet Routing and Addressing (ENCAPS) for IPNG", R. Hinden, http://www.ietf.org/rfc/rfc1955.txt, June 1996.
[8] Minutes from the two day IPng interim meeting February 27-28, 1997, http://playground.sun.com/pub/ipng/html/minutes/ipng-minutes-feb97.txt.
[9] "Separating Identifiers and Locators in Addresses: An Analysis of the GSE Proposal for IPv6", M. Crawford et. al., http://ietfreport.isoc.org/idref/draft-ietf-ipngwg-esd-analysis/, October 1999.
[RFC4632] "Classless Inter-domain Routing (CIDR): The Internet Address Assignment and Aggregation Plan", V. Fuller and T. Li, http://www.ietf.org/rfc/rfc4632.txt, August 2006.
[RFC4218]
[RFC4219]

Calendar

Autumn 2006 - 67th IETF
November 5 - 10, 2006
Host: Siemens
Location: San Diego, US

Spring 2007 - 68th IETF
March 18 - 23, 2007
Host: TBD
Location: Prague, Czech Republic

Summer 2007 - 69th IETF
July 22 - 27, 2007
Host: TBD
Location: Chicago, US

Autumn 2007 - 70th IETF
December 2 - 7, 2007
Host: TBD
Location: TBD

IETF@20

The IETF is 20 years old in 2006. Check out ongoing activities on http://ietf20.isoc.org

IETF Journal
Autumn 2006
Volume 2, Issue 2

Published three times
per year by the:
Internet Society
4 rue des Falaises
CH-1205 Geneva
Switzerland

Managing Editor:
Mirjam Kühne

Copy Editor:
Wendy Rickard

Editorial Board:
Brian Carpenter
Leslie Daigle
Peter Godwin

E-mail:
ietfjournal@isoc.org

IETF Journal
on the Web
Find us at:
ietfjournal.isoc.org

This Publication has been made possible through the support of the following Platinum Programme Supporters of ISOC: