DISPATCH working group	D. Wing	
Internet-Draft	A. Yourtchenko	
Intended status: Standards Track	Cisco	
Expires: February 28, 2011	August 27, 2010	

TOC

Migrating SIP to IPv6 Media Without Connectivity Checks
draft-wing-dispatch-v6-migration-00

Abstract

During the migration from IPv4 to IPv6, it is anticipated that an IPv6 path might be broken for a variety of reasons, causing endpoints to not receive RTP data. Connectivity checks would detect and avoid the user noticing such a problem, but there is industry reluctance to implement connectivity checks.
This document describes a mechanism allowing dual-stack SIP endpoints to attempt communications over IPv6 and fall back to IPv4 if the IPv6 path is not working. The mechanism does not require connectivity checks.

Status of this Memo

Copyright Notice

Table of Contents

1. Introduction   TOC

During the deployment of a dual-stack network and dual-stack SIP
endpoints, the IPv4 network is likely to remain more robust and
reliable than the newly-deployed IPv6 network. An IPv6 network might be
a disconnected island (not connected to another IPv6 network) or due to
human error a firewall rule or routing error might prevent two IPv6
networks from communicating. Non-SIP applications and their underlying
hosts typically follow IPv6 address selection (Draves, R., "Default
Address Selection for Internet Protocol version 6 (IPv6),"
February 2003.) [RFC3484] which prefers IPv6 over IPv4, but will try
IPv4 after timing out connection attempts to the IPv6 address. This
timeout obvious negative effects on the user's experience to such a
degree that IPv6 has been avoided by many large content providers on
the Internet [whitelist] (Google, "Google IPv6 DNS Whitelist,"
March 2008.).
To achieve a similar function on the media plane, SIP endpoints are
expected to use connectivity checks (Rosenberg, J., "Interactive
Connectivity Establishment (ICE): A Protocol for Network Address
Translator (NAT) Traversal for Offer/Answer Protocols," April 2010.)
[RFC5245] to ensure there is IPv6 (or IPv4) connectivity. ICE has an
advantage over the technique currently used by web browser, in that ICE

does not wait for a user-noticeable timeout before trying other addresses. However, ICE is considered overkill for the problem of migrating to IPv6 because of the overhead of timers, interaction with existing media state machines, and CPU impact of SHA1 on large devices processing many sessions and on small devices.
The mechanism described in this document avoids these problems by combining the idea of "media-latching" (commonly used by SBCs, [I-D.ietf-mmusic-media-path-middleboxes] (Stucker, B. and H. Tschofenig, "Analysis of Middlebox Interactions for Signaling Protocol Communication along the Media Path," July 2010.)) with the preference for IPv6, while allowing dual-stack hosts the ability to fall back to IPv4. In this way, dual-stack endpoints do not suffer broken media if the IPv6 network between two endpoints is down.

---

2. Notational Conventions

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119] (Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels," March 1997.).

---

3. Operation for connectionless media (RTP over UDP)

Offers and Answers are constructed containing both IPv6 and IPv4 addresses, using an agreed-upon SDP syntax (see Section 7 (Session Description Protocol)). When a dual-stack SIP UA has media to send, it sends the media over IPv6, while at the same time simultaneously sending T milliseconds worth of data (Section 6 (Bandwidth, Delay, and Asymmetric Results)) of the media over IPv4 using the same RTP sequence numbers and timestamps as thats RTP data sent over IPv6.

> Note: an alternative is for y% of packets to be sent over IPv6, and x% of packets to be sent over IPv4 (x+y=100%). For example, 50%/50%, 90%/10%, or 10%/90%. This has an advantage that we don't double the bandwidth used; however, it has the drawback that if IPv6 is broken, it is noticeable to the user.

With two endpoints, Alice and Bob, there are four situations:

1. IPv6 from Alice to Bob is working and IPv6 from Bob to Alice is working. This is the only situation handled by other proposed techniques which do not perform connectivity checks.

2. IPv6 from Alice to Bob is working, but IPv6 from Bob to Alice is failing

3. IPv6 from Alice to Bob is failing, but IPv6 from Bob to Alice is working

4. IPv6 from Alice to Bob is failing, and IPv6 from Bob to Alice is failing

The success scenario (1) is straightforward. To deal with the failure scenarios (2-4), the following procedures are followed by the endpoints:

If a SIP endpoints has received T/2 milliseconds worth of RTP data over IPv4, and no packets over IPv6, it knows the IPv6 path from its peer is not working (it is the receiving peer in scenario 2, 3, or 4, above). It assumes the IPv6 path to its peer is also not working, and assumes the IPv4 path to its peer is working. So, it ceases sending RTP (and RTCP) over IPv6 and sends all subsequent RTP (and RTCP) over IPv4.

If a SIP endpoint is sending RTP over IPv6 and receives ICMP hard or soft errors, it immediately stops sending RTP over IPv6 and starts sending RTP (and RTCP) over IPv4. ICMP errors received from sending IPv4 are processed normally (they are usually ignored).

Both endpoints MUST support [Symmetric RTP and RTCP (Wing, D., "Symmetric RTP / RTP Control Protocol (RTCP)," July 2007.)](#) [RFC4961], which is already widely implemented and also a requirement of [SBC 'media latching' (Stucker, B. and H. Tschofenig, "Analysis of Middlebox Interactions for Signaling Protocol Communication along the Media Path," July 2010.)](#) [I-D.ietf-mmusic-media-path-middleboxes].

---

4. Operation, connection-oriented media (TCP)   [TOC](#)

First completed TCP handshake wins. [[To be completed.]]

---

5. Considerations for recvonly/sendonly/inactive   [TOC](#)

An endpoint only sends RTP when it normally needs to send RTP. Thus, if an endpoint is currently in 'recvonly' or 'inactive', it won't send RTP. However, it may of course be receiving RTP over IPv6 or over IPv4. This provides an indication to the receiver that the (incoming) IPv6 or IPv4 path is working. Just as if it was sending RTP, the receipt of only IPvX indicates the IPvY path is broken. Thus, when it does eventually need to send RTP packets, it SHOULD only send using the working (IPvX) path.

6. Bandwidth, Delay, and Asymmetric Results

If IPv6 is not working and a switchover to IPv4 occurs, there will be an audible artifact (or visible artifact, if using video). This can be avoided by sending data simultaneously over IPv4 and over IPv6 until both ends see IPv6 traffic and both ends stop sending over IPv4. However, this causes twice the bandwidth consumption at the beginning of the phone call and is undesirable. That is why only "T" milliseconds of traffic are proposed to be sent.
It is possible that IPv6 works in one direction but not the other. This can result in a false positive on one endpoint. When this occurs, the other endpoint will have received no IPv6 packets (but will have received IPv4 packets), and will thus switch over to IPv4. So, this is self correcting and the traffic will be sent over IPv4.

7. Session Description Protocol

A new SDP session-level attribute, happy-eardrums, is defined. It contains one value, T, expressed in milliseconds. Presence of this attribute indicates the endpoint supports the mechanism described in this document. The happy-eardrums attribute adheres to the RFC 4566 "attribute" production. The ABNF syntax of happy-eardrums is provided below:

```
he-attribute = "a=happy-eardrums=" he-value
he-value     = 1*5DIGIT
```

Additionally, it is necessary to have SDP which encodes the IPv6 address and port. ANAT (Camarillo, G. and J. Rosenberg, "The Alternative Network Address Types (ANAT) Semantics for the Session Description Protocol (SDP) Grouping Framework," June 2005.) [RFC4091] has poor backwards compatibility with devices that do not understand ANAT, thus it SHOULD NOT be used. Superior to ANAT are the encodings described in [I-D.hutton-mmusic-icemicrolite] (Hutton, A. and J. Elwell, "A mechanism for conveying alternate addresses using ICE syntax," March 2010.), [I-D.boucadair-mmusic-altc] (Boucadair, M., Kaplan, H., Gilman, R., and S. Veikkolainen, "Session Description Protocol (SDP) Alternate Connectivity (ALTC) Attribute," February 2010.), and [I-D.ietf-mmusic-sdp-capability-negotiation] (Andreasen, F., "SDP Capability Negotiation," March 2010.).
However, for the mechanism in this draft to function, the Answer MUST include both IPv6 and IPv4 addresses in its answer. All of the existing mechanisms (enumerated above) do not provide this functionality. For example, ANAT recommends the answerer set the port number to 0 in the answer, and [I-D.hutton-mmusic-icemicrolite] (Hutton, A. and J. Elwell,

[“A mechanism for conveying alternate addresses using ICE syntax,”](#)
[March 2010.)](#) has the answer only include the candidate address family.
Thus, existing O/A procedures for communicating IPv6 address do not
work, as-is, with the mechanism proposed in this document.

> Note: this draft is not dependent on any particular SDP encoding of
> IPv6. But, of course, it does require that all end nodes use the
> same SDP encoding of IPv6. A consensus on backwards-compatible IPv6
> SDP encoding still needs to be reached.

---

8. Sending SIP Re-Invite   [TOC](#)

To provide information to on-path devices that elevate the QoS for
certain flows, and to release resources on the other endpoint, an
endpoint MUST send a SIP re-INVITE after it has decided which IP
address family to use. [[details to be added.]]

---

9. Security Considerations   [TOC](#)

An attacker, sending UDP data on IPv6 to a newly-starting endpoint
within the negotiated T time, could cause the endpoint to think IPv6 is
working when IPv6 is failing.

---

10. Acknowledgements   [TOC](#)

This idea was inspired from discussions with attendees and invitees to
the SIP/IPv6 Bar BoF at IETF78.
Thanks to Simon Perreault for his review. Thanks to Marc Petit-Huguenin
for his review, and for suggesting the 90%/10% media split. Thanks to
Muthu Perumal for suggesting re-invite after deciding which address
family works, and explaining problem with semantics of current
mechanisms to exchange IPv6 address in SDP.
The nickname of this draft is derived from Happy Eyeballs, which
provides quick delivery of IPv6 or IPv4 content to HTTP browsers
[I-D.wing-http-new-tech] (Wing, D., Yourtchenko, A., and P. Natarajan,
“Happy Eyeballs: Trending Towards Success (IPv6 and SCTP),”
August 2010.).

[TOC](#)

11. IANA Considerations

Register new SDP attribute.

---

12. References

12.1. Normative References

[RFC2119]	Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels," BCP 14, RFC 2119, March 1997 (TXT, HTML, XML).
[RFC3484]	Draves, R., "Default Address Selection for Internet Protocol version 6 (IPv6)," RFC 3484, February 2003 (TXT).
[RFC4961]	Wing, D., "Symmetric RTP / RTP Control Protocol (RTCP)," BCP 131, RFC 4961, July 2007 (TXT).

12.2. Informational References

[I-D.boucadair-mmusic-altc]	Boucadair, M., Kaplan, H., Gilman, R., and S. Veikkolainen, "Session Description Protocol (SDP) Alternate Connectivity (ALTC) Attribute," draft-boucadair-mmusic-altc-00 (work in progress), February 2010 (TXT).
[I-D.hutton-mmusic-icemicrolite]	Hutton, A. and J. Elwell, "A mechanism for conveying alternate addresses using ICE syntax," draft-hutton-mmusic-icemicrolite-01 (work in progress), March 2010 (TXT).
[I-D.ietf-mmusic-media-path-middleboxes]	Stucker, B. and H. Tschofenig, "Analysis of Middlebox Interactions for Signaling Protocol Communication along the Media Path," draft-ietf-mmusic-media-path-middleboxes-03 (work in progress), July 2010 (TXT).
[I-D.ietf-mmusic-sdp-capability-negotiation]	Andreasen, F., "SDP Capability Negotiation," draft-ietf-mmusic-sdp-capability-negotiation-13 (work in progress), March 2010 (TXT).
[I-D.wing-http-new-tech]	Wing, D., Yourtchenko, A., and P. Natarajan, "Happy Eyeballs: Trending Towards Success (IPv6 and SCTP)," draft-wing-http-new-tech-01 (work in progress), August 2010 (TXT).
[RFC4091]	

	Camarillo, G. and J. Rosenberg, "The Alternative Network Address Types (ANAT) Semantics for the Session Description Protocol (SDP) Grouping Framework," RFC 4091, June 2005 (TXT).
[RFC5245]	Rosenberg, J., "Interactive Connectivity Establishment (ICE): A Protocol for Network Address Translator (NAT) Traversal for Offer/Answer Protocols," RFC 5245, April 2010 (TXT).
[whitelist]	Google, "Google IPv6 DNS Whitelist," March 2008.

Authors' Addresses

	Dan Wing
	Cisco Systems, Inc.
	170 West Tasman Drive
	San Jose, CA 95134
	USA
Email:	dwing@cisco.com
	Andrew Yourtchenko
	Cisco Systems, Inc.
	6a de Kleetlaan
	Diegem 1831
	Belgium
Email:	ayourtch@cisco.com