6MAN
Internet-Draft
Updates: 3986 (if approved)
Intended status: Standards Track
Expires: May 10, 2013

B. Carpenter
Univ. of Auckland
S. Cheshire
Apple Inc.
R. Hinden
Check Point

November 6, 2012

Representing IPv6 Zone Identifiers in Address Literals and Uniform Resource Identifiers
draft-ietf-6man-uri-zoneid-05

Abstract

This document describes how the Zone Identifier of an IPv6 scoped address can be represented in a literal IPv6 address and in a Uniform Resource Identifier that includes such a literal address. It updates RFC 3986 accordingly.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on May 10, 2013.

Table of Contents

1. Introduction

The Uniform Resource Identifier (URI) syntax [RFC3986] defined how a literal IPv6 address can be represented in the "host" part of a URI. A subsequent specification [RFC4007] extended the text representation of limited-scope IPv6 addresses such that a zone identifier may be concatenated to a literal address, for purposes described in that RFC. Zone identifiers are especially useful in contexts where literal addresses are typically used, for example during fault diagnosis, when it may be essential to specify which interface is used for sending to a link local address. It should be noted that zone identifiers have purely local meaning within the host where they are defined, and they are completely meaningless for any other host. Today, they are only meaningful when attached to addresses with less than global scope, but it is possible that other uses might be defined in the future.

RFC 4007 does not specify how zone identifiers are to be represented in URIs. Practical experience has shown that this feature is useful, in particular when using a web browser for debugging with link local addresses, but as it is undefined, it is not implemented consistently in URI parsers or in browsers.

Some versions of some browsers accept the RFC 4007 syntax for scoped IPv6 addresses embedded in URIs, i.e., they have been coded to interpret the "%" sign according to RFC 4007 instead of RFC 3986. Clearly this approach is very convenient for users, although it formally breaches the syntax rules of RFC 3986. The present document defines an alternative approach that respects and extends the rules of URI syntax, and IPv6 literals in general, to be consistent.

Thus, this document updates [RFC3986] by adding syntax to allow a zone identifier to be included in a literal IPv6 address within a URI.

It should be noted that in other contexts than a user interface, a zone identifier is mapped into a numeric zone index or interface number. The MIB textual convention [RFC4001] and the socket interface [RFC3493] define this as a 32 bit unsigned integer. The mapping between the human-readable zone identifier string and the numeric value is a host-specific function that varies between operating systems. The present document is concerned only with the human-readable string.

Several alternative solutions were considered while this document was developed. The Appendix briefly describes the alternatives and their advantages and disadvantages.

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].

2. Specification

According to RFC 4007, a zone identifier is attached to the textual representation of an IPv6 address by concatenating "%" followed by <zone_id>, where <zone_id> is a string identifying the zone of the address. However, RFC 4007 gives no precise definition of the character set allowed in <zone_id>. There are no rules or de facto standards for this. For example, the first Ethernet interface in a host might be called %0, %1, %en1, %eth0, or whatever the implementer happened to choose.

In a URI, a literal IPv6 address is always embedded between "[" and "]". This document specifies how a <zone_id> can be appended to the address. A <zone_id> SHOULD contain only ASCII characters classified in RFC 3986 as "unreserved", which conveniently excludes "]" in order to simplify parsing.

Unfortunately "%" is always treated as an escape character in a URI, and according to RFC 3986 it MUST therefore itself be escaped in a URI, in the form "%25". Thus, the scoped address fe80::a%en1 would appear in a URI as http://[fe80::a%25en1].

If an operating system uses any other characters in zone or interface identifiers that are not in the "unreserved" character set, they MUST be escaped with a "%" sign according to RFC 3986.

We now present the necessary formal syntax.

In RFC 3986, the IPv6 literal format is formally defined in ABNF [RFC5234] by the following rule:

```
IP-literal = "[" ( IPv6address / IPvFuture  ) "]"
```

To provide support for a zone identifier, the existing syntax of IPv6address is retained, and a zone identifier may be added optionally to any literal address. This allows flexibility for unknown future uses. The rule quoted above from RFC 3986 is replaced by three rules:

```
IP-literal = "[" ( IPv6address / IPv6addrz / IPvFuture  ) "]"

ZoneID = 1*( unreserved / pct-encoded )
```

```
IPv6addrz = IPv6address "%25" ZoneID
```

This syntax fills the gap that is described at the end of Section 11.7 of RFC 4007.

The rules in [RFC5952] SHOULD be applied in producing URIs.

RFC 3986 states that URIs have a global scope, but that in some cases their interpretation depends on the end-user's context. URIs including a ZoneID are to be interpreted only in the context of the host where they originate, since the ZoneID is of local signifance only.

The 6man WG discussed and rejected an alternative in which the existing syntax of IPv6address would be extended by an option to add the ZoneID only for the case of link-local addresses. It was felt that the present solution offers more flexibility for future uses and is more straightforward to implement.

RFC 4007 offers guidance on how the ZoneID affects interface/address selection inside the IPv6 stack. Note that the behaviour of an IPv6 stack if passed a non-zero zone index for an address other than link-local is undefined.

3. Web Browsers

Due to the lack of a standard in this area, web browsers have been inconsistent in providing for ZoneIDs. Many have no support, but there are examples of ad hoc support. For example, older versions of Firefox allowed the use of a ZoneID preceded by an unescaped "%" character, but this was removed for consistency with RFC 3986. As another example, recent versions of Internet Explorer allow use of a ZoneID preceded by a "%" character escaped as "%25", still beyond the syntax allowed by RFC 3986. This syntax extension is in fact used internally in the Windows operating system and some of its APIs.

This document implies that all browsers should recognise a ZoneID preceded by an escaped "%". In the spirit of "be liberal with what you accept", we also recommend that URI parsers accept bare "%" signs (i.e., a "%" not followed by two valid hexadecimal characters). This makes it easy for a user to copy and paste a string such as "fe80::a%en1" from the output of a "ping" command and have it work.

4. Security Considerations

The security considerations of [RFC3986] and [RFC4007] apply. In

particular, this URI format creates a specific pathway by which a deceitful zone index might be communicated, as mentioned in the final security consideration of RFC 4007. It is emphasised that the format is intended only for debugging purposes, but of course this intention does not prevent misuse.

To limit this risk, implementations SHOULD NOT allow use of this format except for well-defined usages such as sending to link local addresses under prefix fe80::/10.

An HTTP server or proxy MUST ignore any ZoneID attached to an incoming URI, as it only has local significance at the sending host.

5. IANA Considerations

This document requests no action by IANA.

6. Acknowledgements

The lack of this format was first pointed out by Margaret Wasserman some years ago, and more recently by Kerry Lynn. A previous draft document by Martin Duerst and Bill Fenner [I-D.fenner-literal-zone] discussed this topic but was not finalised.

Valuable comments and contributions were made by Karl Auer, Carsten Bormann, Brian Haberman, Ted Hardie, Tatuya Jinmei, Tom Petch, Tomoyuki Sahara, Juergen Schoenwaelder, Dave Thaler, and Ole Troan.

Brian Carpenter was a visitor at the Computer Laboratory, Cambridge University during part of this work.

This document was produced using the xml2rfc tool [RFC2629].

7. Change log [RFC Editor: Please remove]

draft-ietf-6man-uri-zoneid-05: tuned ABNF, clarified RFC 4007 text, 2012-11-06.

draft-ietf-6man-uri-zoneid-04: additional author, 2012-09-21.

draft-ietf-6man-uri-zoneid-03: reverted to percent-encoded model following WGLC, 2012-09-10.

draft-ietf-6man-uri-zoneid-02: additional WG comments, 2012-07-11.

draft-ietf-6man-uri-zoneid-01: use "-" instead of %25, listed alternatives in Appendix, according to WG debate, added suggestion for browser developers, 2012-05-29.

draft-ietf-6man-uri-zoneid-00: adopted by WG, fixed syntax to allow for % encoded characters, 2012-02-17.

draft-carpenter-6man-uri-zoneid-01: chose Option 2, removed 15 character limit, added explanation of ID/number mapping and other clarifications, 2012-02-08.

draft-carpenter-6man-uri-zoneid-00: original version, 2011-12-07.

8. References

8.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC3986] Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform Resource Identifier (URI): Generic Syntax", STD 66, RFC 3986, January 2005.

[RFC4007] Deering, S., Haberman, B., Jinmei, T., Nordmark, E., and B. Zill, "IPv6 Scoped Address Architecture", RFC 4007, March 2005.

[RFC5234] Crocker, D. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", STD 68, RFC 5234, January 2008.

[RFC5952] Kawamura, S. and M. Kawashima, "A Recommendation for IPv6 Address Text Representation", RFC 5952, August 2010.

8.2. Informative References

[I-D.fenner-literal-zone]
Fenner, B. and M. Duerst, "Formats for IPv6 Scope Zone Identifiers in Literal Address Formats", draft-fenner-literal-zone-02 (work in progress), October 2005.

[I-D.iab-identifier-comparison]
Thaler, D., "Issues in Identifier Comparison for Security Purposes", draft-iab-identifier-comparison-03 (work in progress), July 2012.

[RFC2629] Rose, M., "Writing I-Ds and RFCs using XML", RFC 2629, June 1999.

[RFC3493] Gilligan, R., Thomson, S., Bound, J., McCann, J., and W. Stevens, "Basic Socket Interface Extensions for IPv6", RFC 3493, February 2003.

[RFC4001] Daniele, M., Haberman, B., Routhier, S., and J. Schoenwaelder, "Textual Conventions for Internet Network Addresses", RFC 4001, February 2005.

[chrome] Google, "Use the address bar (omnibox)", 2012, <http://support.google.com/chrome/bin/answer.py?answer=95440>.

Appendix A. Alternatives Considered

1. Leave the problem unsolved.

   This would mean that per-interface diagnostics would still have to be performed using ping or ping6:

   ping fe80::a%en1

   Advantage: works today.

   Disadvantage: less convenient than using a browser.

2. Simply using the percent character.

   http://[fe80::a%en1]

   Advantage: allows use of browser, allows cut and paste.

   Disadvantage: invalid syntax under RFC 3986; not acceptable to URI community.

3. Escaping the escape character as allowed by RFC 3986:

   http://[fe80::a%25en1]

   Advantage: allows use of browser, consistent with general URI syntax.

   Disadvantage: somewhat ugly and confusing, doesn't allow simple cut and paste.

4. Alternative separator

   http://[fe80::a-en1]

   Advantage: allows use of browser, simple syntax

   Disadvantage: Requires all IPv6 address literal parsers and generators to be updated in order to allow simple cut and paste; inconsistent with existing tools and practice.

   Note: the initial proposal for this choice was to use an underscore as the separator, but it was noted that this becomes effectively invisible when a user interface automatically underlines URLs.

5. With the "IPvFuture" syntax left open in RFC 3986:

   http://[v6.fe80::a_en1]

   Advantage: allows use of browser.

   Disadvantage: ugly and redundant, doesn't allow simple cut and paste.

Authors' Addresses

Brian Carpenter
Department of Computer Science
University of Auckland
PB 92019
Auckland, 1142
New Zealand

Email: brian.e.carpenter@gmail.com

Stuart Cheshire
Apple Inc.
1 Infinite Loop
Cupertino, CA 95014
US

Email: cheshire@apple.com

Robert M. Hinden
Check Point Software Technologies, Inc.
800 Bridge Parkway
Redwood City, CA 94065
US

Email: bob.hinden@gmail.com