

6MAN
Internet-Draft
Obsoletes: [6874](#) (if approved)
Updates: [3986](#) (if approved)
Intended status: Standards Track
Expires: 18 February 2022

B. Carpenter
Univ. of Auckland
S. Cheshire
Apple Inc.
R. Hinden
Check Point Software
17 August 2021

Representing IPv6 Zone Identifiers in Address Literals and Uniform
Resource Identifiers
draft-carpenter-6man-rfc6874bis-02

Abstract

This document describes how the zone identifier of an IPv6 scoped address, defined as <zone_id> in the IPv6 Scoped Address Architecture ([RFC 4007](#)), can be represented in a literal IPv6 address and in a Uniform Resource Identifier that includes such a literal address. It updates the URI Generic Syntax specification ([RFC 3986](#)) accordingly, and obsoletes [RFC 6874](#).

Discussion Venue

This note is to be removed before publishing as an RFC.

Discussion of this document takes place on the 6MAN mailing list (ipv6@ietf.org), which is archived at <https://mailarchive.ietf.org/arch/browse/ipv6/> (<https://mailarchive.ietf.org/arch/browse/ipv6/>).

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 <https://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 18 February 2022.

Internet-Draft

IPv6 Zone IDs in URIs

August 2021

Copyright Notice

Copyright (c) 2021 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the [Trust Legal Provisions](#) and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1.	Introduction	2
2.	Issues with Implementing RFC 6874	4
3.	Specification	4
4.	Web Browsers	6
5.	Security Considerations	7
6.	Acknowledgements	7
7.	References	7
7.1.	Normative References	8
7.2.	Informative References	8
Appendix A.	Options Considered	9
Appendix B.	Change log	10
	Authors' Addresses	11

[1.](#) Introduction

The Uniform Resource Identifier (URI) syntax specification [[RFC3986](#)] defined how a literal IPv6 address can be represented in the "host" part of a URI. Two months later, the IPv6 Scoped Address Architecture 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 specification. Zone identifiers are especially useful in contexts in which 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 node in which they are defined, often being the same as IPv6 interface names. They

are completely meaningless for any other node. Today, they are meaningful only when attached to addresses with less than global scope, but it is possible that other uses might be defined in the future.

The IPv6 Scoped Address Architecture specification [[RFC4007](#)] does not specify how zone identifiers are to be represented in URIs. Practical experience has shown that this feature is useful or necessary, in at least three use cases:

1. When using a web browser for simple debugging actions involving link-local addresses on a host with more than one active link interface.
2. When using a web browser to reconfigure a misconfigured device which only has a link local address and whose only configuration tool is a web server, again from a host with more than one active link interface.
3. When using an HTTP-based protocol for establishing link-local relationships, such as the Apple CUPS printing mechanism [[CUPS](#)].

It should be noted that whereas some operating systems and network APIs support a default zone identifier as described in [[RFC4007](#)], others do not, and for them an appropriate URI syntax is particularly important.

In the past, some browser versions directly accepted the IPv6 Scoped Address syntax [[RFC4007](#)] for scoped IPv6 addresses embedded in URIs, i.e., they were coded to interpret a "%" sign following the literal address as introducing a zone identifier [[RFC4007](#)], instead of introducing two hexadecimal characters representing some percent-encoded octet [[RFC3986](#)]. Clearly, interpreting the "%" sign as introducing a zone identifier is very convenient for users, although it formally breaches the established URI syntax [[RFC3986](#)]. This 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 the URI syntax specification [[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 contexts other than a user interface, a zone identifier is mapped into a numeric zone index or interface number. The MIB textual convention `InetZoneIndex` [[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. [Appendix A](#) briefly describes the various options and their advantages and disadvantages.

This document obsoletes its predecessor [[RFC6874](#)] by greatly simplifying its recommendations and requirements for web browsers. Its effect on the formal URI syntax [[RFC3986](#)] is exactly the same as that of [RFC 6874](#).

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [BCP 14](#) [[RFC2119](#)] [[RFC8174](#)] when, and only when, they appear in all capitals, as shown here.

2. Issues with Implementing [RFC 6874](#)

Several issues prevented [RFC 6874](#) being implemented in browsers:

1. There was some disagreement with requiring percent-encoding of the "%" sign preceding a zone identifier. This requirement is retained in the present document.
2. The requirement to delete any zone identifier before emitting a URI from the host in an HTTP message was considered both too complex to implement and in violation of normal HTTP practice [[RFC7230](#)]. This requirement has been dropped from the present document.

3. The suggestion to pragmatically allow a bare "%" sign when this would be unambiguous was considered both too complex to implement and confusing for users. This suggestion has been dropped from the present document.

3. Specification

According to IPv6 Scoped Address syntax [[RFC4007](#)], 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, the IPv6 Scoped Address Architecture specification 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. Also, %25 would be valid.

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. According to [Section 2.4 of \[RFC3986\]](#), "%" must be percent-encoded to be used as data within a URI, so any occurrences of literal "%" symbols in a URI MUST be percent-encoded and represented in the form "%25". Thus, the scoped address fe80::abcd%en1 would appear in a URI as http://[fe80::abcd%25en1].

- * Open Issue 1: This choice needs to be re-discussed as there is an argument that URI parsers could be coded to avoid percent-encoding here if so directed by the ABNF syntax. This depends on the exact interpretation of [Section 2.4 of \[RFC3986\]](#)
- * Open Issue 2: Depending on the outcome of the previous issue, there is an argument that an alternative separator (specifically "-") would be preferable to "%25".

A <zone_id> SHOULD contain only ASCII characters classified as "unreserved" for use in URIs [[RFC3986](#)]. This excludes characters such as "]" or even "%" that would complicate parsing. However, the syntax described below does allow such characters to be percent-encoded, for compatibility with existing devices that use them.

If an operating system uses any other characters in zone or interface identifiers that are not in the "unreserved" character set, they MUST be represented using percent encoding [[RFC3986](#)].

We now present the necessary formal syntax.

The URI syntax specification [[RFC3986](#)] formally defined the IPv6 literal format 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 syntax allows flexibility for unknown future uses. The rule quoted above from the previous URI syntax specification [[RFC3986](#)] is replaced by three rules:

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

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

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

Alternative rules for Issue 1 above:

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

```
ZoneID = 1*( unreserved )
```

```
IPv6addrz = IPv6address "-" ZoneID
```

Alternative rules for Issue 2 above:

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

```
ZoneID = 1*( unreserved )
```

```
IPv6addrz = IPv6address "-" ZoneID
```

This syntax fills the gap that is described at the end of

[Section 11.7](#) of the IPv6 Scoped Address Architecture specification [[RFC4007](#)].

The established rules for textual representation of IPv6 addresses [[RFC5952](#)] SHOULD be applied in producing URIs.

The URI syntax specification [[RFC3986](#)] 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 at which they originate, since the ZoneID is of local significance only.

The IPv6 Scoped Address Architecture specification [[RFC4007](#)] offers guidance on how the ZoneID affects interface/address selection inside the IPv6 stack. Note that the behaviour of an IPv6 stack, if it is passed a non-null zone index for an address other than link-local, is undefined.

4. Web Browsers

This section discusses how web browsers might handle this syntax extension. Unfortunately, there is no formal distinction between the syntax allowed in a browser's input dialogue box and the syntax allowed in URIs. For this reason, no normative statements are made in this section.

Due to the lack of defined syntax, web browsers have been inconsistent in providing for ZoneIDs. Most have no support, but there have been examples of ad hoc support. For example, some versions of Firefox allowed the use of a ZoneID preceded by a bare "%" character, but this feature was removed for consistency with established syntax [[RFC3986](#)]. As another example, some versions of Internet Explorer allowed use of a ZoneID preceded by a "%" character

encoded as "%25", still beyond the syntax allowed by the established rules [[RFC3986](#)]. This syntax extension is in fact used internally in the Windows operating system and some of its APIs.

It is desirable for all browsers to recognise a ZoneID according to the above syntax.

URIs including a ZoneID have no meaning outside the originating HTTP

client node. However, in some use cases, such as CUPS mentioned above, the URI will be reflected back to the client.

The normal diagnostic usage for the ZoneID syntax will cause it to be entered in the browser's input dialogue box. Thus, URIs including a ZoneID are unlikely to be encountered in HTML documents. However, if they do (for example, in a diagnostic script coded in HTML), it would be appropriate to treat them exactly as above.

5. Security Considerations

The security considerations from the URI syntax specification [[RFC3986](#)] and the IPv6 Scoped Address Architecture specification [[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 the Scoped Address Architecture specification.

To limit this risk, implementations MUST NOT allow use of this format except for well-defined usages, such as sending to link-local addresses under prefix fe80::/10. At the time of writing, this is the only well-defined usage known.

6. Acknowledgements

The lack of this format was first pointed out by Margaret Wasserman and later by Kerry Lynn. A previous draft document by Martin Duerst and Bill Fenner [[LITERAL-ZONE](#)] discussed this topic but was not finalised. Michael Sweet and Andrew Cady explained some of the difficulties caused by [RFC 6874](#).

Valuable comments and contributions were made by Karl Auer, Carsten Bormann, Benoit Claise, Stephen Farrell, Brian Haberman, Ted Hardie, Philip Homburg, Tatuya Jinmei, Yves Lafon, Barry Leiba, Radia Perlman, Tom Petch, Michael Richardson, Tomoyuki Sahara, Juergen Schoenwaelder, Nico Schottelius, Dave Thaler, Martin Thomson, Ole Troan, and others.

7. References

7.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), DOI 10.17487/RFC2119, March 1997, <<https://www.rfc-editor.org/info/rfc2119>>.
- [RFC3986] Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform Resource Identifier (URI): Generic Syntax", STD 66, [RFC 3986](#), DOI 10.17487/RFC3986, January 2005, <<https://www.rfc-editor.org/info/rfc3986>>.
- [RFC4007] Deering, S., Haberman, B., Jinmei, T., Nordmark, E., and B. Zill, "IPv6 Scoped Address Architecture", [RFC 4007](#), DOI 10.17487/RFC4007, March 2005, <<https://www.rfc-editor.org/info/rfc4007>>.
- [RFC5234] Crocker, D., Ed. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", STD 68, [RFC 5234](#), DOI 10.17487/RFC5234, January 2008, <<https://www.rfc-editor.org/info/rfc5234>>.
- [RFC5952] Kawamura, S. and M. Kawashima, "A Recommendation for IPv6 Address Text Representation", [RFC 5952](#), DOI 10.17487/RFC5952, August 2010, <<https://www.rfc-editor.org/info/rfc5952>>.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in [RFC 2119](#) Key Words", [BCP 14](#), [RFC 8174](#), DOI 10.17487/RFC8174, May 2017, <<https://www.rfc-editor.org/info/rfc8174>>.

7.2. Informative References

- [CUPS] Apple, "CUPS open source printing system", 2021, <<https://www.cups.org/>>.
- [LITERAL-ZONE] Fenner, B. and M. Duerst, "Formats for IPv6 Scope Zone Identifiers in Literal Address Formats", Work in Progress, October 2005.
- [RFC3493] Gilligan, R., Thomson, S., Bound, J., McCann, J., and W. Stevens, "Basic Socket Interface Extensions for IPv6", [RFC 3493](#), DOI 10.17487/RFC3493, February 2003, <<https://www.rfc-editor.org/info/rfc3493>>.

- [RFC4001] Daniele, M., Haberman, B., Routhier, S., and J. Schoenwaelder, "Textual Conventions for Internet Network Addresses", [RFC 4001](#), DOI 10.17487/RFC4001, February 2005, <<https://www.rfc-editor.org/info/rfc4001>>.
- [RFC6874] Carpenter, B., Cheshire, S., and R. Hinden, "Representing IPv6 Zone Identifiers in Address Literals and Uniform Resource Identifiers", [RFC 6874](#), DOI 10.17487/RFC6874, February 2013, <<https://www.rfc-editor.org/info/rfc6874>>.
- [RFC7230] Fielding, R., Ed. and J. Reschke, Ed., "Hypertext Transfer Protocol (HTTP/1.1): Message Syntax and Routing", [RFC 7230](#), DOI 10.17487/RFC7230, June 2014, <<https://www.rfc-editor.org/info/rfc7230>>.

[Appendix A](#). Options Considered

This section will be updated as necessary after the open issues are resolved.

The syntax defined above allows a ZoneID to be added to any IPv6 address. 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 solution presented in this document offers more flexibility for future uses and is more straightforward to implement.

The various syntax options considered are now briefly described.

1. Leave the problem unsolved.

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

```
ping fe80::abcd%n1
```

Advantage: works today.

Disadvantage: less convenient than using a browser. Leaves some use cases unsatisfied.

2. Simply use the percent character:

```
http://[fe80::abcd%n1]
```

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

Disadvantage: invalid syntax under [RFC 3986](#); not acceptable to URI community.

3. Simply use an alternative separator:

```
http://[fe80::abcd-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.

4. Simply use the "IPvFuture" syntax left open in [RFC 3986](#):

```
http://[v6.fe80::abcd_en1]
```

Advantage: allows use of browser.

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

5. Retain the percent character already specified for introducing zone identifiers for IPv6 Scoped Addresses [[RFC4007](#)], and then percent-encode it when it appears in a URI, according to the already-established URI syntax rules [[RFC 3986](#)]:

```
http://[fe80::abcd%25en1]
```

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

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

This is the option chosen for standardisation.

[Appendix B](#). Change log

This section is to be removed before publishing as an RFC.

* [draft-carpenter-6man-rfc6874bis-02](#), 2021-08-12:

Carpenter, et al. Expires 18 February 2022 [Page 10]

Internet-Draft IPv6 Zone IDs in URIs August 2021

- Give details of open issues
- Update authorship
- Editorial fixes

* [draft-carpenter-6man-rfc6874bis-01](#), 2021-07-11:

- Added section on issues with [RFC6874](#)
- Removed suggested heuristic for bare % signs
- Editorial fixes

* [draft-carpenter-6man-rfc6874bis-00](#), 2021-07-05:

- Initial version

Authors' Addresses

Brian Carpenter
School 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
United States of America

Email: cheshire@apple.com

Robert M. Hinden
Check Point Software
959 Skyway Road
San Carlos, CA 94070
United States of America

Email: bob.hinden@gmail.com