

Workgroup: 6MAN

Internet-Draft: draft-ietf-6man-rfc6874bis-02

Obsoletes: [6874](#) (if approved)

Updates: [3986](#), [3987](#) (if approved)

Published: 5 July 2022

Intended Status: Standards Track

Expires: 6 January 2023

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

## **Representing IPv6 Zone Identifiers in Address Literals and Uniform Resource Identifiers**

### **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 and Internationalized Resource Identifier specifications (RFC 3986, RFC 3987) 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](mailto:ipv6@ietf.org)), which is archived at <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 6 January 2023.

## Copyright Notice

Copyright (c) 2022 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 Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

## Table of Contents

- [1. Introduction](#)
- [2. Issues with Implementing RFC 6874](#)
- [3. Specification](#)
- [4. URI Parsers](#)
- [5. Security Considerations](#)
- [6. IANA Considerations](#)
- [7. References](#)
  - [7.1. Normative References](#)
  - [7.2. Informative References](#)
- [Appendix A. Options Considered](#)
- [Appendix B. Change log](#)
- [Appendix C. Acknowledgements](#)
- [Authors' Addresses](#)

## 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, usually being the same as IPv6 interface names. They are completely meaningless for any other node. Today, they are meaningful only when attached to link-local addresses, 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 various use cases, including the following:

1. A web browser may be used for simple debugging actions involving link-local addresses on a host with more than one active link interface.
2. A web browser must sometimes be used to configure or reconfigure a device which only has a link local address and whose only configuration tool is a web server, again in a host with more than one active link interface.
3. The Apple and open-source CUPS printing mechanism [[CUPS](#)] [[OP-CUPS](#)] uses an HTTP-based protocol [[RFC3510](#)][[RFC7472](#)] to establish link-local relationships, so requires the specification of the relevant interface.
4. The Microsoft Web Services for Devices (WSD) virtual printer port mechanism can generate an IPv6 Link Local URL in which the zone identifier is present and necessary, but is not recognized by any current browser.

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 is not supported by the URI syntax in [[RFC3986](#)] or the Internationalized Resource Identifier (IRI) syntax in [[RFC3987](#)]. Therefore, this document updates RFC 3986 and RFC 3987 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 URI parsers. Its effect on the formal URI syntax [[RFC3986](#)] is different from 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 dropped 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 [[RFC9110](#)]. 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 since it is now irrelevant.

## **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, %25, %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 identifier can be appended to the address. The URI syntax defined by [\[RFC3986\]](#) does not allow the presence of a percent ("%") character within an IPv6 address literal. For this reason, it is backwards compatible to allow the use of "%" within an IPv6 address literal as a delimiter only, such that the scoped address fe80::abcd%en1 would appear in a URI as http://[fe80::abcd%en1] or https://[fe80::abcd%en1].

This use of "%" as a delimiter applies only within an IPv6 address literal, and is irrelevant to and exempt from the percent-encoding mechanism [\[RFC3986\]](#).

A zone identifier MUST contain only ASCII characters classified as "unreserved" for use in URIs [\[RFC3986\]](#). This excludes characters such as "]" or even "%" that would complicate parsing. For the avoidance of doubt, note that a zone identifier consisting of "25" or starting with "25" is valid and is used in some operating systems. A parser MUST NOT apply percent decoding to the IPv6 address literal in a URI, including cases such as http://[fe80::abcd%25] and http://[fe80::abcd%25xy].

If an operating system uses any other characters in zone or interface identifiers that are not in the "unreserved" character set, they cannot be used in a URI.

We now present the corresponding formal syntax.

The URI syntax specification [\[RFC3986\]](#) formally defines 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 [\[RFC3986\]](#) is replaced by three rules:

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

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

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

This change also applies to [\[RFC3987\]](#).

This syntax fills the gap that is described at the end of Section 11.7 of the IPv6 Scoped Address Architecture specification

[[RFC4007](#)]. It replaces and obsoletes the syntax in Section 2 of [[RFC6874](#)].

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 zone identifier are an example of this, since the zone identifier is of local significance only. Such a URI cannot be correctly interpreted outside the host to which it applies.

The IPv6 Scoped Address Architecture specification [[RFC4007](#)] offers guidance on how the zone identifier 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.

In cases where the RFC 6874 encoding is currently used between specific software components rather than between a browser and a web server, such usage MAY continue indefinitely.

#### 4. URI Parsers

This section discusses how URI parsers, such as those embedded in 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.

In practice, although parsers respect the established syntax, they are coded pragmatically rather than being formally syntax-driven. Typically, IP address literals are handled by an explicit code path. Parsers have been inconsistent in providing for zone identifiers. Most have no support, but there have been examples of ad hoc support. For example, some versions of Firefox allowed the use of a zone identifier 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 zone identifier 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 URI parsers to recognise a zone identifier according to the syntax defined in [Section 3](#). An IPv6 address literal never contains percent-encodings. In terms of Section 2.4 of [[RFC3986](#)], the "%" character preceding a zone identifier is acting

as a delimiter, not as data. Any code handling percent-encoding or percent-decoding must be aware of this.

As noted above, URIs including a zone identifier have no meaning outside the originating HTTP client node. However, in some use cases, such as CUPS, the host address embedded in the URI will be reflected back to the client, using exactly the representation of the zone identifier that the client sent.

The various use cases for the zone identifier syntax will usually require it to be entered in a browser's input dialogue box. However, URIs including a zone identifier might occur in HTML documents. For example, a diagnostic HTML page might be tailored for a particular host. Because of such usage, it is appropriate for browsers to treat such URIs in the same way whether they are entered in the dialogue box or encountered in an HTML document.

## **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.

However, this format is only meaningful for link-local addresses under prefix fe80::/10. It is not necessary for web browsers to verify this, or to validate the zone identifier, because the operating system will do so when the address is passed to the socket API, and return an error code if the zone identifier is invalid.

It is conceivable that this format could be misused to probe a local network configuration. In particular, a script included in an HTML web page could originate HTTP messages intended to determine if a particular link-local address is valid, for example to discover and misuse the address of the first-hop router. However, such attacks are already possible, by probing IPv4 addresses, routeable IPv6 addresses or link-local addresses without a zone identifier. Indeed, with a zone identifier present, the attacker's job is harder because they must also guess the zone identifier itself; the zone identifier increases the search space compared to guessing only the interface identifier. Zone identifiers vary widely between operating systems; in some cases they are simple integers or conventional names such as "eth0" but in other cases they contain arbitrary characters derived from MAC addresses. In any case, an attacker must discover them before probing any link-local addresses. This argues against the recommendation of [[RFC4007](#)] to support a default zone identifier.

It should be noted that if a node uses an interface identifier in the outdated Modified EUI format [RFC4291] for its link-local address, the search space for an attacker is very significantly reduced, as discussed in Section 4.1.1.1 of [RFC7707]. The resultant recommendations of [RFC8064] apply to all nodes, including routers, since they ensure that the search space for an attacker is of size  $2^{64}$ , which is impracticably large.

Nevertheless, even a Modified EUI link-local address is significantly harder to guess than typical IPv4 addresses for devices such as home routers, which are often included in published documentation.

## 6. IANA Considerations

This document makes no request of IANA.

## 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>>.
- [RFC3987] Duerst, M. and M. Suignard, "Internationalized Resource Identifiers (IRIs)", RFC 3987, DOI 10.17487/RFC3987, January 2005, <<https://www.rfc-editor.org/info/rfc3987>>.
- [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>>.



**[RFC8064]**

Gont, F., Cooper, A., Thaler, D., and W. Liu,  
"Recommendation on Stable IPv6 Interface Identifiers",  
RFC 8064, DOI 10.17487/RFC8064, February 2017, <<https://www.rfc-editor.org/info/rfc8064>>.

**[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, "Apple CUPS", 2022, <<https://www.cups.org/>>.

**[LITERAL-ZONE]**

Fenner, B. and M. Dürst, "Formats for IPv6 Scope Zone  
Identifiers in Literal Address Formats", Work in  
Progress, October 2005.

**[OP-CUPS]**

Sweet, M., "OpenPrinting CUPS", 2022, <<https://openprinting.github.io/cups/>>.

**[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>>.

**[RFC3510]**

Herriot, R. and I. McDonald, "Internet Printing Protocol/  
1.1: IPP URL Scheme", RFC 3510, DOI 10.17487/RFC3510,  
April 2003, <<https://www.rfc-editor.org/info/rfc3510>>.

**[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>>.

**[RFC4291]**

Hinden, R. and S. Deering, "IP Version 6 Addressing  
Architecture", RFC 4291, DOI 10.17487/RFC4291, February  
2006, <<https://www.rfc-editor.org/info/rfc4291>>.

**[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>>.

**[RFC7472]**

McDonald, I. and M. Sweet, "Internet Printing Protocol  
(IPP) over HTTPS Transport Binding and the 'ipps' URI

Scheme", RFC 7472, DOI 10.17487/RFC7472, March 2015,  
<<https://www.rfc-editor.org/info/rfc7472>>.

[RFC7707] Gont, F. and T. Chown, "Network Reconnaissance in IPv6 Networks", RFC 7707, DOI 10.17487/RFC7707, March 2016,  
<<https://www.rfc-editor.org/info/rfc7707>>.

[RFC9110] Fielding, R., Ed., Nottingham, M., Ed., and J. Reschke, Ed., "HTTP Semantics", STD 97, RFC 9110, DOI 10.17487/RFC9110, June 2022, <<https://www.rfc-editor.org/info/rfc9110>>.

## Appendix A. Options Considered

The syntax defined above allows a zone identifier 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 zone identifier 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%en1
```

Advantage: works today.

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

### 2. Simply use the percent character:

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

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

Disadvantage: requires code changes to all URI parsers.

This is the option chosen for standardisation.

### 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 [[RFC3986](#)]:

`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.

## **Appendix B. Change log**

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

\*draft-ietf-6man-rfc6874bis-02, 2022-07-05:

- Improve discussion of URLs in HTML documents
- Discuss scripting attack and Modified EUI IIDs
- Several editorial clarifications
- Some nits fixed

\*draft-ietf-6man-rfc6874bis-01, 2022-04-07:

- Extended use cases

- Clarified relationship with RFC3986 language
  - Allow for legacy use of RFC6874 format
  - Augmented security considerations
  - Editorial and reference improvements
- \*draft-ietf-6man-rfc6874bis-00, 2022-03-19:
- WG adoption
  - Clarified security considerations
- \*draft-carpenter-6man-rfc6874bis-03, 2022-02-08:
- Changed to bare % signs.
  - Added IRIs, RFC3987
  - Editorial fixes
- \*draft-carpenter-6man-rfc6874bis-02, 2021-18-12:
- 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

## **Appendix C. Acknowledgements**

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

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

#### **Authors' Addresses**

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

Email: [brian.e.carpenter@gmail.com](mailto:brian.e.carpenter@gmail.com)

Stuart Cheshire  
Apple Inc.  
1 Infinite Loop  
Cupertino, CA 95014  
United States of America

Email: [cheshire@apple.com](mailto: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](mailto:bob.hinden@gmail.com)