IPv6 Operations
Internet-Draft
Intended status: Informational
Expires: April 30, 2018

I. Lubashev
E. Nygren
Akamai Technologies

October 27, 2017

A Recommendation for IPv6 Address/Mask Notation
draft-lubashev-ipv6-addr-mask-01

Abstract

Since network operators are commonly assigned at least /48 IPv6 address prefixes, the operators and standards occasionally find opportunities to devise addressing schemes that further assign operational semantics to less significant bit ranges. There is currently no standard or interoperable textual representation of addresses sharing bit patterns that are not prefixes. This RFC introduces IPv6 Address/Mask notation that allows one to represent address groupings beyond "all addresses that share a single prefix". The representation is similar to the IPv4 address/mask notation in its expressiveness, but it is derived from the familiar address/prefix-length notation for clarity and compatibility with existing parsers.

For example, using this representation, both 2001:db8::/32 and 2001:db8::/ffff:ffff:: have the same meaning. However, a group of addresses having the first 32 bits "2001:0db8::" and the last 16 bits "::1234" requires the new representation: 2001:db8::1234//ffff:ffff::ffff or, equivalently, 2001:db8::1234//32+::ffff.

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 April 30, 2018.

Table of Contents

1. Introduction

We have learned to think of IPv4 address groupings in terms of CIDR blocks, because virtually all logical address groupings fit that model well: IP address allocations, subnets, routing announcements, etc.

With the move to IPv6, the primary mechanism for address grouping remains matching by prefix length, albeit with longer prefix lengths. This only allows for strictly hierarchical address groupings. The longer address lengths, however, provide opportunities for assigning operator-specific semantics to bit strings within addresses beyond the prefix, especially when allocating addresses for virtual services.

Numerous systems (see Appendix A for examples) have been assigning semantics to IPv6 bits that come after IANA prefix bits. Developers of these systems attempted to communicate address patterns underlying their system semantics both in documentation and in machine-readable configurations accompanying the systems. Due to the lack of a standard textual representation, the documentation often resorted to pictographs and verbose English descriptions. The configuration syntax and parsers were invariably ad hoc and incompatible with other systems.

Here we define a syntax for representing groupings (matching rules) of IPv6 addresses, where a set of less significant bits have a particular value. For example, 2001:db8::1234//ffff:ffff::ffff matches all addresses whose 32 most significant bits are 2001:0db8 and whose 16 least significant bits are 1234.

This document only concerns itself with the textual representation of address groups that cannot be expressed as CIDR blocks. Our goal is standardizing on a consistent representation to remove a hindrance to interoperability of systems that wish to express rules and policies that apply to such address groups (see Appendix A for examples). Guidance for the applicability of such address groupings is outside the scope of this document.

1.1. Netmask and Prefix-Length Notations

There are two common textual representations for identifying groups of addresses (networks, subnets, internet routing blocks). These representations can also be used to identify an individual address and its subnet.

The netmask notation described by [RFC0950] is commonly used for IPv4. It consists of a tuple of a network address and a network mask. For example: 198.51.100.4 netmask 255.255.255.0.

The address/prefix-length notation described by [RFC4632] is commonly used for both IPv4 and IPv6. It consists of a tuple of a network address and a prefix length. For example: 198.51.100.4/24 or 2001:db8::1234/32.

Depending on the context, netmask and prefix length notations can specify either a "group of addresses" or "an individual address and a group of addresses to which it belongs". If the network address contains one or more set bits not selected by the network mask or prefix length, then network address specifies an individual address in addition to the subnet. For example: 198.51.100.4/24 means "address 198.51.100.4 within a group of addresses 198.51.100.0 - 198.51.100.255".

2. Problem Description

The problem with the prefix length notation for IPv6 is that it is not sufficiently expressive of IPv6 address groupings for a growing number of applications.

IPv6 address allocation guidelines [RFC6177] guarantee at least a /48 allocation to network operators and strongly recommend a multi-/64 allocation to end sites. Because these address blocks are orders of magnitude larger than any imaginable number of physical hosts, network operators are managing those addresses in new and creative ways.

Sometimes, useful address grouping are not "all addresses that share a prefix of a certain length". Additionally, within an administrative scope, there are use-cases where semantics are assigned to individual bit ranges.

Consider these examples:

1. Allocating a block of addresses to each host and using the least significant bits to indicate a TLS certificate. These operators may need a way to express a rule that applies to all traffic that uses a particular TLS certificate.

2. Network operators managing multiple similar data centers may have different prefixes routed to those data centers but desire a unified set of rules for assigning, managing, and routing IPv6 addresses within those data centers. These operators need to

express rules that do not depend on the prefixes of the addresses to which the rules apply.

3. 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].

4. IPv6 Address/Mask Textual Representation

This RFC extends address/prefix-length notation of [RFC4632] in a way that is reminiscent of the IPv4 netmask notation of [RFC0950]. The address/mask notation allows specifying IPv6 mask instead of the prefix length.

The address/mask notation is defined as:

```
ADDRESS // MASK
```

Both ADDRESS and MASK are IPv6 addresses. The MASK indicates which bits of the ADDRESS are relevant for the address grouping. Note that the MASK may be sparse and is not strictly a prefix. For example: 2001:db8::1234//ffff:ffff::ffff.

The "//" was chosen as a separator for address/mask notation, since it is similar to the "/" separator used by address/prefix-length (and hence is readily recognizable) but prevents incorrectly parsing address/mask as address/prefix-length.

4.1. Constraints and Validation

To be a valid definition for just a group of addresses, the ADDRESS part MUST NOT have any bits set outside of the MASK. Otherwise, the ADDRESS // MASK represents an individual address and a group of addresses it belongs to.

4.2. Examples: groups of addresses

1. 2001:db8::1234//ffff:ffff::ffff

   This specifies IPv6 addresses that look like 2001:db8::1234 when you ignore bits 16-95.

2. ::aa00:1234//::ff00:ffff

   This specifies IPv6 addresses that have "aa" in bits 24-31 and "1234" in bits 0-15.

3. 2001:db8::////ffff:ffff::

   This is equivalent to 2001:db8::/32.

4.3. Examples: specific addresses and groups to which they belong

1. 2001:db8::1:1234//ffff:ffff::ffff

   This specifies IPv6 address 2001:db8::1:1234 that belongs to a group of addresses that look like 2001:db8::1234 when you ignore bits 16-95.

2. 2001:db8::aa00:1234//::ff00:ffff

   This specifies IPv6 address 2001:db8::aa00:1234 that belongs to a group of addresses that have "aa" in bits 24-31 and "1234" in bits 0-15.

3. 2001:db8::1//ffff:ffff::

   This is equivalent to 2001:db8::1/32.

4.4. Textual representation of the Address and Mask

When IPv6 mask is used after "//", both the network address and mask parts MUST be formatted as IPv6 addresses and, therefore, their canonical textual representation is dictated by [RFC5952].

4.5. Use prefix length instead of mask

The canonical representation of a group of IPv6 addresses MUST use a prefix length instead of a mask if possible. That is, if the mask has all its most significant bits set, up to some bit, followed by all clear bits, then the canonical representation MUST use a prefix length.

5. Scoped Mask/Value notation

Since assigning operator-specific semantics to bit ranges is only possible within the address space assigned to the operator by IANA, a common use-case is to specify an address/mask within an IANA-assigned prefix scope. For example, all addresses ending with ::1234 within 2001:db8::/32 can be specified as 2001:db8::1234//ffff:ffff::ffff.

To make these representations easier to manage and validate, it helps to have an explicit convention for representing prefixes within address groups. For example, 2001:db8::1234//ffff:ffff::ffff can be represented as 2001:db8::1234//32+::ffff.

This is specified as:

    ADDRESS // PFX_LEN + SCOPED_MASK

Scoped Mask/Value notation representation can be canonicalized using a ADDRESS // MASK notation. The canonical MASK is constructed by performing the bitwise-or of SCOPED_MASK and the mask derived from an address with the PFX_LEN most significant bits set.

The PFX_LEN most significant bits MUST NOT be set in SCOPED_MASK.

6. Compatibility and Parser guidelines

Only parsers that wish to support address groupings that cannot be represented using address/prefix-length are required to support address/mask notation.

Systems that support communicating address grouping in address/mask notation to other systems SHOULD communicate such grouping in canonical address/prefix-length notation, if possible. This ensures compatibility with systems that do not support address/mask notation, if all configured address groupings are proper CIDR prefixes.

Address groupings that cannot be expressed using address/prefix-length notation MAY be communicated using Scoped Mask/Value (Section 5) notation, as long as the PFX_LEN (semantic prefix scope) has been configured via external means (i.e. PFX_LEN SHOULD NOT be automatically derived from the MASK bitmap by the system itself).

7. Security Considerations

This document only defines textual representation for IPv6 address groupings. It does not intend to recommend when assigning semantics to specific bit ranges and matching based on bit substrings is applicable or appropriate.

IP addresses can be spoofed or attacker-controlled. This is especially true of IPv6 addresses differing only in less significant bits and belonging to different administrative domains. When used in policies applied to incoming traffic, the MASK part of the address/mask notation SHOULD have as many set bits as the semantics of the policy would allow.

Operators wishing to assign semantics to bit ranges should be aware that these semantics may be guessable or leaked outside the organization. Hence, there is a risk of privacy/information leakage.

8. Address Utilization Considerations

Since IPv6 allocation guidelines [RFC6177] guarantee at least a /48 allocation to network operators, it would be an enormous waste of the address space to assign IPv6 addresses only to physical hosts or network interfaces.

On the other hand, a gratuitous use of lower address bits can lead to a premature address space exhaustion and difficulties in adapting to the future needs of the organization within the assigned address space. An example of such gratuitous use is designating large parts of the address space for a bitmask, where only a small fraction of all possible bit combinations is utilized.

9. IANA Considerations

This document has no actions for IANA.

10. Change Log

10.1. Since draft-lubashev-ipv6-addr-mask-00

- Changed separator from "/" to "//"
- Addressed privacy in Section 7
- Added Section 6
- Added Section 8
- Added Appendix A

11. Acknowledgments

The Acknowledgments will come here.

12. References

12.1. Normative References

[RFC0950] Mogul, J. and J. Postel, "Internet Standard Subnetting Procedure", STD 5, RFC 950, DOI 10.17487/RFC0950, August 1985, <https://www.rfc-editor.org/info/rfc950>.

[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>.

[RFC4632] Fuller, V. and T. Li, "Classless Inter-domain Routing (CIDR): The Internet Address Assignment and Aggregation Plan", BCP 122, RFC 4632, DOI 10.17487/RFC4632, August 2006, <https://www.rfc-editor.org/info/rfc4632>.

[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>.

[RFC6177] Narten, T., Huston, G., and L. Roberts, "IPv6 Address Assignment to End Sites", BCP 157, RFC 6177, DOI 10.17487/RFC6177, March 2011, <https://www.rfc-editor.org/info/rfc6177>.

12.2. Informative References

[GeoAddressPatent] Chen, L., Steenstra, J., and K. Taylor, "US 7929535: Geolocation-based addressing method for IPv6 addresses", April 2011, <http://patft1.uspto.gov/netacgi/nph-Parser?patentnumber=7929535>.

[I-D.jiang-semantic-prefix] Jiang, S., Qiong, Q., Farrer, I., Bo, Y., and T. Yang, "Analysis of Semantic Embedded IPv6 Address Schemas", draft-jiang-semantic-prefix-06 (work in progress), July 2013.

[IncognitoRoutingPlan] Kostur, A., "How to Plan Routing for IPv6", July 2015, <https://www.incognito.com/tips-and-tutorials/how-to-plan-routing-for-ipv6>.

[OpenFlow] Open Networking Foundation, "OpenFlow Switch Specification, Version 1.2", December 2011, <https://3vf60mmveq1g8vzn48q2o71a-wpengine.netdna-ssl.com/wp-content/uploads/2014/10/openflow-spec-v1.2.pdf>.

[RFC4380] Huitema, C., "Teredo: Tunneling IPv6 over UDP through Network Address Translations (NATs)", RFC 4380, DOI 10.17487/RFC4380, February 2006, <https://www.rfc-editor.org/info/rfc4380>.

[RFC6066] Eastlake 3rd, D., "Transport Layer Security (TLS) Extensions: Extension Definitions", RFC 6066, DOI 10.17487/RFC6066, January 2011, <https://www.rfc-editor.org/info/rfc6066>.

[SURFnetAddrPlan] SURFnet, "Preparing an IPv6 Address Plan", September 2013, <https://www.ipv6forum.com/dl/presentations/IPv6-addressing-plan-howto.pdf>.

[TeraStream] Lothberg, P. and M. Abrahamsson, "TeraStream - IPv6 Addressing Format", October 2013, <https://ripe67.ripe.net/presentations/251-ripe2-4.pdf>.

Appendix A. Examples of Semantic Use of Lower Address Bits

Assigning semantics to lower bits of IPv6 addresses and defining policies based on such address groupings have been done for many years. Documents describing such policies and configurations for equipment implementing these policies do not use a consistent notation. Most documents resort to pictographs and verbose English, while the configuration syntax and parsers are invariably ad hoc.

A.1. A Framework for Semantic IPv6 Prefix and Gap Analysis

Internet draft [I-D.jiang-semantic-prefix] describes the need for adding semantics to lower IPv6 address bits to define address groups that cannot be expressed as CIDR blocks and analyzes some implications of this practice. This draft describes uses of such address groups for creating routing policies as well as configuring such policies on hosts and routers both statically and dynamically.

A.2. Teredo

Teredo protocol [RFC4380] uses four bit ranges past Teredo IANA prefix bits to encode server and client IPv4 addresses, a flags bitmap, and a port (section "4. Teredo Addresses").

A.3. OpenFlow Switch Configuration

OpenFlow Switch Specification [OpenFlow] describes OpenFlow switch configuration API that can match flows based on an arbitrary IPv6 bitmask applied to IPv6 source (OXM_OF_IPV6_SRC) or destination (OXM_OF_IPV6_DST) addresses. Version 1.2 was the first version to introduce such IPv6 address/bitmask flow match rules (chapter A.2.3.7) in 2011.

A.4. TeraStream IPv6 Addressing

TeraStream [TeraStream] system is using bit ranges to encode service type in IPv6 address bits that come after IANA prefix bits. The system was launched in 2012.

A.5. SURFnet IPv6 Address Plan and Incognito Routing Plan

SURFnet published a white paper [SURFnetAddrPlan] advocating that ISPs use bit ranges past their IANA prefix to encode geo-location and address use types. The white paper is giving examples of a sample address allocation that uses one nibble (bits 68-71) for encoding geo-location and another nibble (bits 64-67) for the use type.

IPv6 Routing Plan [IncognitoRoutingPlan] by incognito is advocating allocating bit ranges past an IANA prefix to designate various address attributes, including "subnet types".

A.6. Geolocation-based addressing method for IPv6 addresses

Qualcomm US patent 7,929,535 [GeoAddressPatent] describes a method of embedding geo-location information, such as latitude, longitude, altitude, in predefined ranges of bits of an IPv6 address past their IANA prefix.

A.7. Customer IDs in less significant bits

CDNs and hosting providers host web sites belonging to multiple customers using shared servers. Due to the lack of support for SNI TLS extension [RFC6066] by some user agents active on the Internet, CDNs resort to using unique IP addresses to identify specific customer domains and, hence, certificates for TLS negotiation. In case of IPv6 addresses, at least some CDNs use the less significant bits of an IPv6 address to identify customer domains (while the more significant bits carry internal routing information). The configuration of systems matching lower bits of IPv6 addresses to individual customer domains must use ad hoc syntax due to the lack of a standard way to express semantics of matching on bit ranges other than address prefixes.

Authors' Addresses

Igor Lubashev
Akamai Technologies

EMail: igorlord@alum.mit.edu

Erik Nygren
Akamai Technologies

EMail: erik+ietf@nygren.org
URI: http://erik.nygren.org/