Workgroup: Network Working Group
Internet-Draft:
draft-ietf-pim-zeroconf-mcast-addr-alloc-ps-00
Published: 14 September 2023
Intended Status: Standards Track
Expires: 17 March 2024
Authors: N. Karstens (Garmin International)
D. Farinacci (lispers.net)
M. McBride (Futurewei)

Zeroconf Multicast Address Allocation Problem Statement and Requirements

Abstract

This document describes a network that requires unique multicast addresses to distribute data. Various challenges are discussed, such as the use of multicast snooping to ensure efficient use of bandwidth, limitations of switch hardware, problems associated with address collisions, and the need to avoid user configuration. After all limitations were considered it was determined that multicast addresses need to be dynamically-assigned by a decentralized, zero-configuration protocol.

Requirements and recommendations for suitable protocols are listed and specific considerations for assigning IPv4 and IPv6 addresses are reviewed. The document closes with several solutions that are precluded from consideration.

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 17 March 2024.

Copyright Notice

Table of Contents

1. Introduction

Marine networks contain a combination of sensors, controls, and displays. Installations vary widely depending on the design and intended purpose of the boat and the amount of redundancy required. Sensors on these networks can be a mix of low-cost, low-bandwidth devices, like temperature or fluid sensors, and high-bandwidth devices, like radar, sonar, and video cameras. In most cases, these networks use a single subnet and therefore require layer-2 switches to be deployed.

The most optimal way to distribute sensor data to all displays on the network is multicast. However, use of traditional switches can be problematic when both high-bandwidth and low-bandwidth devices are installed. Low-bandwidth devices are commonly designed with a low-speed link to reduce cost, and the multicast stream from the high-bandwidth device can overwhelm this link. Switch hardware at the low price points that are acceptable to the market do not support source-specific multicast. Instead, multicast streams are differentiated by destination address and switches with multicast snooping [RFC4541] in a default-block configuration are used to isolate multicast streams to the ports with devices that request the data.

This technique presents several challenges. First, defining an industry-standard set of pre-allocated addresses is not practical due to the wide variety of network designs. Manually configuring addresses for each device is not a user-friendly solution. MADCAP [RFC2730] could be used to dynamically assign addresses, but its reliance on a dedicated server results in a single point of failure for the system, which is not acceptable for the target environment. Finally, this method is susceptible to link-layer address collisions (see Section 2 for further discussion).

The desired solution needs to be a decentralized, zero-configuration protocol for dynamically assigning multicast addresses. This document serves as a basis for developing suitable protocols by defining the problem, discussing constraints, and listing requirements.

1.1. Requirements Language

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. Address Collisions

Link-layer address collisions are a concern in two cases.

First, many Ethernet chips include the ability to filter out unwanted traffic. This is typically configured by the network stack in response to an application joining a multicast group. Any link-layer address collision would require that the network stack use CPU time to filter out traffic by its IPv6 multicast address, which may cause poor performance.

Networks that use multicast snooping switches are also susceptible to address collisions. According to Section 4 of [RFC4541], most switch vendors forward multicast traffic based only on the link-layer address (see the results for Q2 and Q3). This means that unwanted data will be transmitted over the link and, depending on the nature of the data, may result in a low-bandwidth link being saturated by a high-bandwidth stream.

3. Protocol Requirements

A decentralized, zero-configuration protocol for dynamic multicast address assignment MUST have the following characteristics:

1. Does not rely on a single point of failure

2. Does not depend on user configuration

3. Coexists with other multicast address assignment protocols

4. Supports operation on a single subnet

5. Does not require an Internet connection

6. Supports multiple applications on the same host

7. Detects and resolves address collisions

Note that an extreme case of address collision may occur after a network partition, when intermittent link failure temporarily divides the network into multiple segments.

A protocol SHOULD ideally have the following characteristics:

1. Supports operation across multiple subnets

2. Does not require significant changes to existing standards

3. Uses functionality commonly available on a variety of platforms

4. Uses capabilities commonly provided to unprivileged applications

5. Avoids depending on configuration data loaded during device manufacture

6. Minimizes network traffic

4. IPv6 Considerations

The IPv6 multicast address guidelines specified in [RFC3307] are well-structured and robust. Section 2 defines the lower 32 bits of the IPv6 address, which are mapped directly to the link-layer, as the group ID, and then assigns ranges of group ID values based on how they are allocated. Section 4.3 describes dynamic assignment of group ID values and lists two different approaches (server allocation and host allocation). However, both approaches are assigned the same range of group ID values, which means they cannot coexist without risking an address collision. Also concerning is that the range for dynamic assignment overlaps with the range used for solicited-node multicast addresses (see Section 2.7.1 of [RFC4291]).

5. IPv4 Considerations

Section 6.4 of [RFC1112] recognizes that more than one IPv4 multicast address can be mapped to the same Ethernet multicast address. This is because the lowest 23 bits are mapped to the Ethernet multicast address. A 32-bit IPv4 multicast address has a 4-bit prefix, which leaves 5 bits inconsequential to the operation, or 32 addresses.

The guidelines for allocating IPv4 multicast addresses in [RFC5771] did not anticipate a need to avoid address collisions. As such, the recommendation for all new designs using dynamic assignment is to use IPv6. If this is not feasible, then the recommendation is for the protocol to assign addresses from a suitable range in the Administratively Scoped Block (239.0.0.0/8) and be aware of other applications on the network using addresses it may collide with.

6. Excluded Solutions

The prefix for IPv4 and IPv6 multicast messages being transmitted on Ethernet are specified in [RFC1112], Section 6.4 and [RFC2464], Section 7, respectively. Allowing a different prefix would support at least two solutions that are being excluded from consideration.

First, reducing the size of the prefix would increase the size of the group ID, thereby reducing the probability of an address collision.

Because link-layer addresses are only relevant on the local subnet, it would also be possible to develop a new protocol to dynamically map network-layer multicast addresses to link-layer multicast addresses in an operation somewhat analogous to DHCP. Multicast packets routed from outside the network could have the address mapped at ingress without any assignment protocol.

Ultimately, using a different prefix seemed like a significant change that would only gain widespread platform support after significant delay.

With IPv4, reserving 32 separate address ranges in the registry could prevent address collisions. However, [RFC5771] cautions that IPv4 multicast address space is limited and this approach seemed excessive.

7. Security Considerations

Security considerations will be discussed by any proposed zero-configuration multicast address allocation algorithm.

8. IANA Considerations

This document has no IANA actions.

9. Acknowledgement

Special thanks to the National Marine Electronics Association for their contributions in developing marine industry standards and their support for this research.

Thanks also to the members of the PIM working group for their early brainstorming sessions and review of this draft.

10. References

10.1. 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>.

[RFC3307] Haberman, B., "Allocation Guidelines for IPv6 Multicast Addresses", RFC 3307, DOI 10.17487/RFC3307, August 2002, <https://www.rfc-editor.org/info/rfc3307>.

[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>.

10.2. Informative References

[RFC1112] Deering, S., "Host extensions for IP multicasting", STD 5, RFC 1112, DOI 10.17487/RFC1112, August 1989, <https://www.rfc-editor.org/info/rfc1112>.

[RFC2464] Crawford, M., "Transmission of IPv6 Packets over Ethernet Networks", RFC 2464, DOI 10.17487/RFC2464, December 1998, <https://www.rfc-editor.org/info/rfc2464>.

[RFC2730] Hanna, S., Patel, B., and M. Shah, "Multicast Address Dynamic Client Allocation Protocol (MADCAP)", RFC 2730, DOI 10.17487/RFC2730, December 1999, <https://www.rfc-editor.org/info/rfc2730>.

[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>.

[RFC4541] Christensen, M., Kimball, K., and F. Solensky, "Considerations for Internet Group Management Protocol (IGMP) and Multicast Listener Discovery (MLD) Snooping Switches", RFC 4541, DOI 10.17487/RFC4541, May 2006, <https://www.rfc-editor.org/info/rfc4541>.

[RFC5771] Cotton, M., Vegoda, L., and D. Meyer, "IANA Guidelines for IPv4 Multicast Address Assignments", BCP 51, RFC 5771, DOI 10.17487/RFC5771, March 2010, <https://www.rfc-editor.org/info/rfc5771>.

Authors' Addresses

Nate Karstens
Garmin International

Email: nate.karstens@gmail.com

Dino Farinacci
lispers.net

Email: farinacci@gmail.com

Mike McBride
Futurewei

Email: michael.mcbride@futurewei.com