Internet Area [intarea]
Internet-Draft
Expires: September 22, 2016

C. Perkins
Futurewei
D. Stanley
HPE
W. Kumari
Google
JC. Zuniga
InterDigital

March 21, 2016

Multicast Considerations over IEEE 802 Wireless Media
draft-perkins-intarea-multicast-ieee802-00.txt

Abstract

This document describes some performance issues that have been observed when multicast packet transmission is attempted over IEEE 802 wireless media. Multicast features specified for IEEE 802 wireless media related to multicast are also described, along with explanations about how these features can help ameliorate the observed performance issues. IETF protocols that are likely to be affected by the observed performance issues are identified, and workarounds are proposed in some cases. The performance of multicast over wireless media often can be quite different than the performance of unicast. This draft describes the nature of the differences and the effects on representative IETF protocols. We also describe some efforts that have been made by IEEE 802 Wireless groups to ameliorate the performance differences.

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 September 22, 2016.

Table of Contents

1. Introduction

Many IETF protocol designs depend upon multicast or broadcast for delivery of control messages to multiple receivers. Multicast is

used for various purposes such as neighborhood discovery, network flooding, address resolution, as well as reduction in media access for data traffic.

IETF protocols typically expect to rely on network protocol layering in order to reduce or eliminate any dependence of higher level protocols on the specific nature of the MAC layer protocols or the physical media. In the case of multicast transmission, higher level protocols may be designed as if transmitting a packet to an IP address has the same cost in interference and network media access, regardless of whether the destination IP address is a unicast address or a multicast or broadcast address. This model of operation was reasonable for networks where the physical medium was like an Ethernet.

Unfortunately, for many wireless media, the costs can be quite different. It is the purpose of this Internet Draft to identify the ways in which the costs can be different. Using this information, we then proceed to identify some possible effects on the actual operation of IETF protocols over wireless media.

IEEE 802 Wireless working groups, especially 802.11, have made a number of attempts to improve the performance of multicast transmissions at layer 2. In this draft we also include a description of some of these efforts. This information is closely related to material presented at IETF 94 [cite 11-15-1261-03]

2. Terminology

This document defines the following terminology:

basic rate
: a "lowest common denominator" rate at which multicast and broadcast traffic is generally transmitted.

MCS
: Modulation and Coding Scheme.

3. Identified Issues at Layer 2

In this section we list some of the issues arising at layer 2 surrounding the use of multicast in IETF protocols over wireless media.

- Multicast traffic is typically much less reliable than unicast traffic.
- Multicast / broadcast traffic is generally sent at a lowest common denominator rate, known as a basic rate. This might be as low as 6 Mbps, when unicast links are operating at 600 Mbps. Transmission at a lower rate requires more occupancy of the wireless medium and thus less airtime for everything else.

- Wireless multicast affects wired LANs because the AP extends the wired segment.
  - All broadcast frames on LAN side are copied to WLAN.
  - In WLAN, broadcast messages transmitted at most robust MCS.
  - Most robust MCS implies large frames sent at slow rate.
- Multicast can work poorly with the power-save mechanisms in 802.11.
  - Both unicast and multicast traffic can be delayed by power-saving mechanisms.
  - Unicast is delayed until a STA wakes up and asks for it. Additionally, unicast traffic may be delayed to improve power save, efficiency and increase probability of aggregation.
  - Multicast traffic is delayed in a wireless network if any of the STAs in that network are power savers. All STAs have to be awake at a known time to receive multicast traffic.
  - Packets can also be discarded due to buffer limitations in the AP and non-AP STA.

4. Some Possible Effects on Representative IETF protocols

In this section we list some of the issues arising at layer 3 surrounding the use of multicast in IETF protocols over wireless media. We mention a few representative IETF protocols, and describe some possible effects due to performance degradation when using multicast transmissions for control messages. Common uses include:

- Control plane for IPv4 and IPv6

- ARP and Neighbor Discovery
- Service discovery
- Applications (video delivery, stock data etc)
- Other L3 protocols (non-IP)

4.1. IPv4 uses

The following list contains a few representative IPv4 protocols using multicast.

- ARP
- DHCP
- mDNS

After initial configuration, ARP and DHCP occur much less commonly.

4.2. IPv6 uses

The following list contains a few representative IPv6 protocols using multicast. IPv6 makes much more extensive use of multicast.

- DHCPv6
- Liveness detection (NUD)
- Some control plane protocols are not very tolerant of packet loss, especially neighbor discovery.
- Services may be considered lost if several consecutive packets fail.

Address Resolution

Service Discovery

Route Discovery

Decentralized Address Assignment

Geographic routing

4.3. Disabling Multicast on WiFi

Multicast Listener Discovery(MLD) [RFC4541] is often used to identify members of a multicast group that are connected to the ports of a switch. Forwarding multicast frames into a WiFi-enabled area can use such switch support for hardware forwarding state information. However, since IPv6 makes heavy use of multicast, each STA with an IPv6 address will require state on the switch for several and possibly many multicast solicited-node addresses. Multicast addresses that do not have forwarding state installed (perhaps due to hardware memory limitations on the switch) cause frames to be flooded on all ports of the switch.

4.4. Spurious Neighbor Discovery

On the Internet there is a "background radiation" of scanning traffic (people scanning for vulnerable machines) and backscatter (responses from spoofed traffic, etc). This means that the router is constantly getting packets destined for machines whose IP addresses may or may not be in use. In the cases where the IP is assigned to a machine, the router broadcasts an ARP request, gets back an ARP reply, caches this and then can deliver traffic to the host. In the cases where the IP address is not in use, the router broadcasts one (or more) ARP requests, and never gets a reply. This means that it does not populate the ARP cache, and the next time there is traffic for that IP address it will broadcast ARP requests again. The rate of these

ARP requests is proportional to the size of the subnets, the rate of scanning and backscatter, and how long the router keeps state on non-responding ARPs. As it turns out, this rate is inversely proportional to how occupied the subnet is (valid ARPs end up in a cache, stopping the broadcasting; unused IPs never respond, and so cause more broadcasts). Depending on the address space in use, the time of day, how occupied the subnet is, and other unknown factors, on the order of 2000 broadcasts per second have been observed at the IETF NOCs.

On a wired network, there is not a huge difference amongst unicast, multicast and broadcast traffic; but this is not true in the wireless realm. Wireless equipment often is unable to send this amount of broadcast and multicast traffic. Consequently, on the wireless networks, we observe a significant amount of dropped broadcast and

multicast packets. This, in turn, means that when a host connects it is often not able to complete DHCP, and IPv6 RAs get dropped, leading to users being unable to use the network.

5. Layer 2 optimizations

This section lists some optimizations that have been specified for use with 802.11 that are aimed at reducing or eliminating the causes of performance loss discussed in section Section 3.

5.1. Proxy ARP in 802.11-2012

The AP knows all associated STAs MAC address and IP address; in other words, the AP acts as the central "manager" for all the 802.11 STAs in its BSS. Proxy ARP is easy to implement at the AP, and offers the following advantages:

- Reduced broadcast traffic (transmitted at low MCS) on the wireless medium
- STA benefits from extended power save in sleep mode, as ARP requests are replied to by AP.
- Keeps ARP frames off the wireless medium.

Here is the specification language from clause 10.23.13 in [2] as described in [dot11-proxyarp]:

> When the AP supports Proxy ARP "[...] the AP shall maintain a Hardware Address to Internet Address mapping for each associated station, and shall update the mapping when the Internet Address of the associated station changes. When the IPv4 address being resolved in the ARP request packet is used by a non-AP STA currently associated to the BSS, the proxy ARP service shall respond on behalf of the non-AP STA"

5.2. Buffering to improve Power-Save

The AP acts on behalf of STAs in various ways. In order to improve the power-saving feature for STAs in its BSS, the AP buffers frames for delivery to the STA at the time when the STA is scheduled for reception.

5.3. IPv6 support in 802.11-2012

IPv6 uses Neighbor Discovery Protocol (NDP) instead Every IPv6 node subscribes to special multicast address Neighbor-Solicitation message replaces ARP

Here is the specification language from-10.23.13 in [2]:

> "When an IPv6 address is being resolved, the Proxy Neighbor Discovery service shall respond with a Neighbor Advertisement message [...] on behalf of an associated STA to an [ICMPv6] Neighbor Solicitation message [...]. When MAC address mappings change, the AP may send unsolicited Neighbor Advertisement Messages on behalf of a STA."

NDP may be used to request additional information

- Maximum Transmission Unit
- Router Solicitation
- Router Advertisement, etc.

NDP messages are sent as group addressed (broadcast) frames in 802.11. Using the proxy operation helps to keep NDP messages off the wireless medium.

5.4. Directed Multicast Service (DMS)

DMS enables a client to request that the AP transmit multicast group addressed frames destined to the requesting clients as individually addressed frames [i.e., convert multicast to unicast].

- DMS Requires 802.11n A-MSDUs
- Individually addressed frames are acknowledged and are buffered for power save clients
- Requesting STA may specify traffic characteristics for DMS traffic
- DMS was defined in IEEE Std 802.11v-2011

DMS is not currently implemented in products.

5.5. GroupCast with Retries (GCR)

GCR (defined in [dot11aa]) provides greater reliability by using either unsolicited retries or a block acknowledgement mechanism. GCR increases probability of broadcast frame reception success, but still does not guarantee success.

For the block acknowledgement mechanism, the AP transmits each group addressed frame as conventional group addressed transmission. Retransmissions are group addressed, but hidden from non-11aa clients. A directed block acknowledgement scheme is used to harvest reception status from receivers; retransmissions are based upon these responses.

GCR is suitable for all group sizes including medium to large groups. As the number of devices in the group increases, GCR can send block acknowledgement requests to only a small subset of the group.

GCR may introduce unacceptable latency. After sending a group of data frames to the group, the AP has do the following:

- unicast a Block Ack Request (BAR) to a subset of members.
- wait for the corresponding Block Ack (BA).
- retransmit any missed frames.
- resume other operations which may have been delayed.

This latency may not be acceptable for some traffic.

There are ongoing extensions in 802.11 to improve GCR performance.

- BAR is sent using downlink MU-MIMO (note that downlink MU-MIMO is already specified in 802.11-REVmc 4.3).
- BA is sent using uplink MU-MIMO (which is a .11ax feature).
- Additional 802.11ax extensions are under consideration; see [mc-ack-mux]
- Latency may also be reduced by simultaneously receiving BA information from multiple clients.

6. Higher Layer Optimizations and Mitigations

This section lists some optimizations that have been specified for use with 802.11 that are aimed at reducing or eliminating the causes of performance loss discussed in section Section 6.

6.1. Mitigating Problems from Spurious Neighbor Discovery

ARP Sponges

ARP Sponges sit on a network and learn what IPs addresses are actually in use. They also listen for ARP requests, and, if it sees an ARP for an IP address which it believes is not used, it will reply with its own MAC address. This means that the router now has an IP to MAC mapping, which it caches. If that IP is later assigned to an machine (e.g using DHCP), the ARP sponge will see this, and will stop replying for that address. Gratuitous ARPs (or the machine ARPing for its gateway) will replace the sponged address in the router ARP table. This technique is quite effective; but, unfortunately, the ARP sponge daemons were not really designed for this use (the standard one [arpsponge], was designed to deal with the disappearance of participants from an IXP) and so are not optimized for this purpose. We have to run one daemon per subnet, the tuning is tricky (the scanning rate versus the population rate versus retires, etc.) and sometimes the daemons just seem to stop, requiring a restart of the daemon and causing disruption.

Router mitigations

Some routers (often those based on Linux) implement a "negative ARP cache" daemon. Simply put, if the router does not see a reply to an ARP it can be configured to cache this information for some interval. Unfortunately, the core routers which we are using do not support this. When a host connects to network and gets an IP address, it will ARP for its default gateway (the router). The router will update its cache with the IP to host MAC mapping learnt from the request (passive ARP learning).

Firewall unused space

The distribution of users on wireless networks / subnets changes from meeting to meeting (e.g the "IETF-secure" SSID was renamed to "IETF", fewer users use "IETF-legacy", etc). This utilization is difficult to predict ahead of time, but we can monitor the usage as attendees use the different networks. By configuring multiple DHCP pools per subnet, and enabling them sequentially, we can have a large subnet, but only assign addresses from the lower portions of it. This means that we

can apply input IP access lists, which deny traffic to the upper, unused portions. This means that the router does not

attempt to forward packets to the unused portions of the subnets, and so does not ARP for it. This method has proven to be very effective, but is somewhat of a blunt axe, is fairly labor intensive, and requires coordination.

Disabling/filtering ARP requests

In general, the router does not need to ARP for hosts; when a host connects, the router can learn the IP to MAC mapping from the ARP request sent by that host. This means that we should be able to disable and / or filter ARP requests from the router. Unfortunately, ARP is a very low level / fundamental part of the IP stack, and is often offloaded from the normal control plane. While many routers can filter layer-2 traffic, this is usually implemented as an input filter and / or has limited ability to filter output broadcast traffic. This means that the simple "just disable ARP or filter it outbound" seems like a really simple (and obvious) solution, but implementations / architectural issues make this difficult or awkward in practice.

NAT

The broadcasts are overwhelmingly being caused by outside scanning / backscatter traffic. This means that, if we were to NAT the entire (or a large portion) of the attendee networks, there would be no NAT translation entries for unused addresses, and so the router would never ARP for them. The IETF NOC has discussed NATing the entire (or large portions) attendee address space, but a: elegance and b: flaming torches and pitchfork concerns means we have not attempted this yet.

Stateful firewalls

Another obvious solution would be to put a stateful firewall between the wireless network and the Internet. This firewall would block incoming traffic not associated with an outbound

request. The IETF philosophy has been to have the network as open as possible / honor the end-to-end principle. An attendee on the meeting network should be an Internet host, and should be able to receive unsolicited requests. Unfortunately, keeping the network working and stable is the first priority and a stateful firewall may be required in order to achieve this.

7. Multicast Considerations for Other Wireless Media

Many of the causes of performance degradation described in earlier sections are also observable for wireless media other than 802.11.

For instance, problems with power save, excess media occupancy, and poor reliability will also affect 802.15.3 and 802.15.4. However, 802.15 media specifications do not include similar mechanisms of the type that have been developed for 802.11. In fact, the design philosophy for 802.15 is more oriented towards minimality, with the result that many such functions would more likely be relegated to operation within higher layer protocols. This leads to a patchwork of non-interoperable and vendor-specific solutions. See [uli] for some additional discussion, and a proposal for a task group to resolve similar issues, in which the multicast problems might be considered for mitigation.

8. Security Considerations

This document does not introduce any security mechanisms, and does not have any impact on existing security mechanisms.

9. IANA Considerations

This document does not specify any IANA actions.

10. Informative References

[arpsponge] Arien Vijn, Steven Bakker, , "Arp Sponge", March 2015.

[dot11] P802.11, , "Part 11: Wireless LAN Medium Access Control

(MAC) and Physical Layer (PHY) Specifications", March 2012.

[dot11-proxyarp]
P802.11, , "Proxy ARP in 802.11ax", September 2015.

[dot11aa] P802.11, , "Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 2: MAC Enhancements for Robust Audio Video Streaming", March 2012.

[mc-ack-mux]
Yusuke Tanaka et al., , "Multiplexing of Acknowledgements for Multicast Transmission", July 2015.

[mc-prob-stmt]
Mikael Abrahamsson and Adrian Stephens, , "Multicast on 802.11", March 2015.

[mc-props]
Adrian Stephens, , "IEEE 802.11 multicast properties", March 2015.

[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, <http://www.rfc-editor.org/info/rfc4541>.

[uli] Pat Kinney, , "LLC Proposal for 802.15.4", Nov 2015.

Authors' Addresses

Charles E. Perkins
Futurewei Inc.
2330 Central Expressway
Santa Clara, CA 95050
USA

Phone: +1-408-330-4586

Email: charliep@computer.org

Dorothy Stanley
Hewlett Packard Enterprise
2000 North Naperville Rd.
Naperville, IL 60566
USA

Phone: +1 630 979 1572
Email: dstanley@arubanetworks.com

Warren Kumari
Google
1600 Amphitheatre Parkway
Mountain View, CA 94043
USA

Email: warren@kumari.net

Juan Carlos Zuniga
InterDigital
1000 Sherbrooke W, 10th Floor
Montreal, QC H3A 3G4
Canada

Email: j.c.zuniga@ieee.org