Internet Area WG
Internet-Draft
Intended status: Informational
Expires: September 14, 2017

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

March 13, 2017

Multicast Considerations over IEEE 802 Wireless Media
draft-perkins-intarea-multicast-ieee802-02

Abstract

Some performance issues have been observed when multicast packet transmissions of IETF protocols are used over IEEE 802 wireless media. Even though enhamcements for multicast transmissions have been designed at both IETF and IEEE 802, there seems to exist a disconnect between specifications, implementations and configuration choices.

This draft describes the different issues that have been observed, the multicast enhancement features that have been specified at IETF and IEEE 802 for wireless media, as well as the operational chioces that can be taken to improve the performace of the network. Finally, it provides some recommendations about the usage and combination of these features and operational choices.

Table of Contents

1. Introduction

Many IETF protocols depend on multicast/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 for reduction in media access for the transmission of data that is intended for multiple receivers.

IETF protocols typically 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 transmissions, higher level protocols have traditionally been designed as if transmitting a packet to an IP address had 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 was reasonable for networks where the physical medium was usually wired, like Ethernet. Unfortunately, for many wireless media, the costs to access the medium can be quite different. Some enhancements have been designed in IETF protocols that are assumed to work primarily over wilress media. However, these enhancements are usually implemented in limited deployments and not widely spread on most wireless networks.

IEEE 802 wireless protocols have been designed with certain fetures to support multicat traffic. For instance, lower modulations are used to transmit multicast frames, so that these can be received by all stations in the cell, regardless of the distance or path attenuation from the base station or access point. However, these lower modulation transmissions take longer on the medium and therefore they reduce the capabilities to transmit more high efficiency traffic with higher order modulations to stations that may be in closer vicinity. Due to these and other reasons, some IEEE 802 working groups like 802.11 have designed several features to improve the performance of multicast transmissions at Layer 2 [REF 11-15-1261-03]. Besides protocol design features, some operational and configuration enhancements can also be applied to overcome the network performance issues created by multicast traffic.

This Internet Draft identifies the problems created by the usage of multicast traffic over wireless networks. It also highlights the different enhancements that have been designed at IETF and IEEE 802, as well as the operational choices that can be taken, to ameliorate the effects of multicast traffic. Some recommendations about the usage and combinations of these enhancements are also provided.

2. Terminology

This document uses the following definitions:

AP
: IEEE 802.11 Access Point.

STA
: IEEE 802.11 station.

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

MCS
: Modulation and Coding Scheme.

3. Identified mulitcast issues

3.1. Issues at Layer 2 and below

In this section we list some of the issues related to the use of multicast transmissions over IEEE 802 wireless technologies.

3.1.1. Multicast reliability

Multicast traffic is typically much less reliable than unicast traffic. Since multicast makes point-to-multipoint communications, multiple acknowledgements would be needed to guarantee the reception on all recepients.

3.1.2. Lower data rate

Because lower MCS have longer range but also lower data rate, multicast / broadcast traffic is generally transmitted at the lowest common denominator rate, also known as a basic rate. On IEEE 802.11 networks (aka Wi-Fi), this rate might be as low as 6 Mbps, when some unicast links in the same cell can be operating at rates up to 600 Mbps. Transmissions at a lower rate require more occupancy of the wireless medium and thus restrict the airtime for all other medium communications and degrade the overall capacity.

Wired multicast affects wireless LANs because the AP extends the wired segment and multicast / broadcast frames on the wired LAN side are copied to WLAN. Since broadcast messages are transmitted at the most robust MCS, this implies that large frames sent at slow rate over the air.

3.1.3. Power-save effects on multicast

Multicast can work poorly with the power-save mechanisms defined in IEEE 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.

3.2. Issues at Layer 3 and above

In this section we mention a few representative IETF protocols, and describe some possible negative effects due to performance degradation when using multicast transmissions for control messages. Common uses of multicast include:

- Control plane for IPv4 and IPv6
- ARP and Neighbor Discovery
- Service discovery
- Applications (video delivery, stock data etc)
- Other L3 protocols (non-IP)

3.2.1. IPv4 issues

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

- ARP
- DHCP
- mDNS

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

3.2.2. IPv6 issues

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

- DHCPv6
- Liveness detection (NUD)

- o Some control plane protocols are not very tolerant of packet loss, especially neighbor discovery.
- o Services may be considered lost if several consecutive packets fail.

Address Resolution

Service Discovery

Route Discovery

Decentralized Address Assignment

Geographic routing

3.2.3. MLD issues

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.

3.2.4. 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.

4. Multicast protocol optimizations

This section lists some optimizations that have been specified in IEEE 802 and IETF that are aimed at reducing or eliminating the issues discussed in Section 3.

4.1. 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.
- Changes are not needed to STA implementation.

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"

4.2. 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.

4.3. 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.

4.4. Conversion of multicast to unicast

It is often possible to transmit control and data messages by using unicast transmissions to each station individually.

4.5. Directed Multicast Service (DMS)

There are situations where more is needed than simply converting multicast to unicast [Editor’s note: citation needed]. For these purposes, 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. DMS does require changes to both AP and STA implementation.

4.6. 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 does require changes to both AP and STA implementation.

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.

5. Operational optimizations

This section lists some operational optimizations that can be implemented when deploying wireless IEEE 802 networks to mitigate the issues discussed in Section 3.

5.1. Mitigating Problems from Spurious Neighbor Discovery

ARP Sponges

An ARP Sponge sits on a network and learn which IPs addresses are actually in use. It 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.

6. 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.

7. Recommendations

This section provides some recommendations about the usage and combinations of the multicast enhancements described in Section 4 and Section 5.

(FFS)

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
SIGFOX
425 rue Jean Rostand
Labege 31670
France

Email: j.c.zuniga@ieee.org