

Internet Area  
Internet-Draft  
Intended status: Informational  
Expires: 29 January 2022

C.E. Perkins  
Blue Meadow Networks  
M. McBride  
Futurewei  
D. Stanley  
HPE  
W. Kumari  
Google  
JC. Zuniga  
SIGFOX  
28 July 2021

Multicast Considerations over IEEE 802 Wireless Media  
draft-ietf-mboned-ieee802-mcast-problems-15

Abstract

Well-known issues with multicast have prevented the deployment of multicast in 802.11 (wifi) and other local-area wireless environments. This document describes the known limitations of wireless (primarily 802.11) Layer-2 multicast. Also described are certain multicast enhancement features that have been specified by the IETF, and by IEEE 802, for wireless media, as well as some operational choices that can be taken to improve the performance of the network. Finally, some recommendations are provided about the usage and combination of these features and operational choices.

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 29 January 2022.

## Copyright Notice

Copyright (c) 2021 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 Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

## Table of Contents

1. Introduction . . . . .	3
2. Terminology . . . . .	4
3. Identified multicast issues . . . . .	5
3.1. Issues at Layer 2 and Below . . . . .	5
3.1.1. Multicast reliability . . . . .	5
3.1.2. Lower and Variable Data Rate . . . . .	6
3.1.3. Capacity and Impact on Interference . . . . .	7
3.1.4. Power-save Effects on Multicast . . . . .	7
3.2. Issues at Layer 3 and Above . . . . .	7
3.2.1. IPv4 issues . . . . .	8
3.2.2. IPv6 issues . . . . .	8
3.2.3. MLD issues . . . . .	9
3.2.4. Spurious Neighbor Discovery . . . . .	9
4. Multicast protocol optimizations . . . . .	10
4.1. Proxy ARP in 802.11-2012 . . . . .	10
4.2. IPv6 Address Registration and Proxy Neighbor Discovery . . . . .	11
4.3. Buffering to Improve Battery Life . . . . .	12
4.4. Limiting multicast buffer hardware queue depth . . . . .	13
4.5. IPv6 support in 802.11-2012 . . . . .	13
4.6. Using Unicast Instead of Multicast . . . . .	14
4.6.1. Overview . . . . .	14
4.6.2. Layer 2 Conversion to Unicast . . . . .	14
4.6.3. Directed Multicast Service (DMS) . . . . .	14
4.6.4. Automatic Multicast Tunneling (AMT) . . . . .	15
4.7. GroupCast with Retries (GCR) . . . . .	15
5. Operational optimizations . . . . .	16
5.1. Mitigating Problems from Spurious Neighbor Discovery . . . . .	16
5.2. Mitigating Spurious Service Discovery Messages . . . . .	18
6. Multicast Considerations for Other Wireless Media . . . . .	18
7. Recommendations . . . . .	19
8. On-going Discussion Items . . . . .	19
9. Security Considerations . . . . .	20

10. IANA Considerations . . . . .	20
11. Acknowledgements . . . . .	21
12. Informative References . . . . .	21
Authors' Addresses . . . . .	25

## 1. Introduction

Well-known issues with multicast have prevented the deployment of multicast in 802.11 [dot11] and other local-area wireless environments, as described in [mc-props], [mc-prob-stmt]. Performance issues have been observed when multicast packet transmissions of IETF protocols are used over IEEE 802 wireless media. Even though enhancements for multicast transmissions have been designed at both IETF and IEEE 802, incompatibilities still exist between specifications, implementations and configuration choices.

Many IETF protocols depend on multicast/broadcast for delivery of control messages to multiple receivers. Multicast allows sending data to multiple interested recipients without the source needing to send duplicate data to each recipient. With broadcast traffic, data is sent to every device regardless of their expressed interest in the data. Multicast is used for various purposes such as neighbor discovery, network flooding, address resolution, as well minimizing media occupancy for the transmission of data that is intended for multiple receivers. In addition to protocol use of broadcast/multicast for control messages, more applications, such as push to talk in hospitals, or video in enterprises, universities, and homes, are sending multicast IP to end user devices, which are increasingly using Wi-Fi for their connectivity.

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 wired, like Ethernet. Unfortunately, for many wireless media, the costs to access the medium can be quite different. Multicast over Wi-Fi has often been plagued by such poor performance that it is disallowed. Some enhancements have been designed in IETF protocols that are assumed to work primarily over wireless media. However, these enhancements are usually implemented in limited deployments and not widespread on most wireless networks.

IEEE 802 wireless protocols have been designed with certain features to support multicast 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 occupy the medium longer; they hamper efficient transmission of traffic using higher order modulations to nearby stations. For these and other reasons, IEEE 802 working groups such as 802.11 have designed features to improve the performance of multicast transmissions at Layer 2 [ietf\_802-11]. In addition to protocol design features, certain operational and configuration enhancements can ameliorate the network performance issues created by multicast traffic, as described in Section 5.

There seems to be general agreement that these problems will not be fixed anytime soon, primarily because it's expensive to do so and due to multicast being unreliable. Compared to unicast over Wi-Fi, multicast is often treated as somewhat of a second class citizen, even though there are many protocols using multicast. Something needs to be provided in order to make them more reliable. IPv6 neighbor discovery saturating the Wi-Fi link is only part of the problem. Wi-Fi traffic classes may help. This document is intended to help make the determination about what problems should be solved by the IETF and what problems should be solved by the IEEE (see Section 8).

This document details various problems caused by multicast transmission over wireless networks, including high packet error rates, no acknowledgements, and low data rate. It also explains some enhancements that have been designed at the IETF and IEEE 802.11 to ameliorate the effects of the radio medium on multicast traffic. Recommendations are also provided to implementors about how to use and combine these enhancements. Some advice about the operational choices that can be taken is also included. It is likely that this document will also be considered relevant to designers of future IEEE wireless specifications.

## 2. Terminology

This document uses the following definitions:

ACK

The 802.11 layer 2 acknowledgement

AP

IEEE 802.11 Access Point

**basic rate**

The slowest rate of all the connected devices, at which multicast and broadcast traffic is generally transmitted

**DTIM**

Delivery Traffic Indication Map (DTIM): An information element that advertises whether or not any associated stations have buffered multicast or broadcast frames

**MCS**

Modulation and Coding Scheme

**NOC**

Network Operations Center

**PER**

Packet Error Rate

**STA**

802.11 station (e.g. handheld device)

**TIM**

Traffic Indication Map (TIM): An information element that advertises whether or not any associated stations have buffered unicast frames

### 3. Identified multicast issues

#### 3.1. Issues at Layer 2 and Below

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

##### 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 reception at all recipients. And since there are no ACKs for multicast packets, it is not possible for the Access Point (AP) to know whether or not a retransmission is needed. Even in the wired Internet, this characteristic often causes undesirably high error rates. This has contributed to the relatively slow uptake of multicast applications even though the protocols have long been available. The situation for wireless links is much worse, and is quite sensitive to the presence of background traffic. Consequently, there can be a high packet error rate (PER) due to lack of retransmission, and because

the sender never backs off. PER is the ratio, in percent, of the number of packets not successfully received by the device. It is not uncommon for there to be a packet loss rate of 5% or more, which is particularly troublesome for video and other environments where high data rates and high reliability are required.

### 3.1.2. Lower and Variable Data Rate

Multicast over wired differs from multicast over wireless because transmission over wired links often occurs at a fixed rate. Wi-Fi, on the other hand, has a transmission rate that varies depending upon the STA's proximity to the AP. The throughput of video flows, and the capacity of the broader Wi-Fi network, will change with device movement. This impacts the ability for QoS solutions to effectively reserve bandwidth and provide admission control.

For wireless stations authenticated and linked with an Access Point, the power necessary for good reception can vary from station to station. For unicast, the goal is to minimize power requirements while maximizing the data rate to the destination. For multicast, the goal is simply to maximize the number of receivers that will correctly receive the multicast packet; generally the Access Point has to use a much lower data rate at a power level high enough for even the farthest station to receive the packet, for example as briefly mentioned in section 2 of [RFC5757]. Consequently, the data rate of a video stream, for instance, would be constrained by the environmental considerations of the least reliable receiver associated with the Access Point.

Because more robust modulation and coding schemes (MCSs) have longer range but also lower data rate, multicast / broadcast traffic is generally transmitted at the slowest rate of all the connected devices. This is also known as the basic rate. The amount of additional interference depends on the specific wireless technology. In fact, backward compatibility and multi-stream implementations mean that the maximum unicast rates are currently up to a few Gbps, so there can be more than 3 orders of magnitude difference in the transmission rate between multicast / broadcast versus optimal unicast forwarding. Some techniques employed to increase spectral efficiency, such as spatial multiplexing in MIMO systems, are not available with more than one intended receiver; it is not the case that backwards compatibility is the only factor responsible for lower multicast transmission rates.

Wired multicast also affects wireless LANs when the AP extends the wired segment; in that case, multicast / broadcast frames on the wired LAN side are copied to the Wireless Local Area Network (WLAN). Since broadcast messages are transmitted at the most robust MCS, many large frames are sent at a slow rate over the air.

### 3.1.3. Capacity and Impact on Interference

Transmissions at a lower rate require longer occupancy of the wireless medium and thus take away from the airtime of other communications and degrade the overall capacity. Furthermore, transmission at higher power, as is required to reach all multicast STAs associated to the AP, proportionately increases the area of interference with other consumers of the radio spectrum.

### 3.1.4. Power-save Effects on Multicast

One of the characteristics of multicast transmission over wifi is that every station has to be configured to wake up to receive the multicast frame, even though the received packet may ultimately be discarded. This process can have a large effect on the power consumption by the multicast receiver station. For this reason there are workarounds, such as Directed Multicast Service (DMS) described in Section 4, to prevent unnecessarily waking up stations.

Multicast (and unicast) can work poorly with the power-save mechanisms defined in IEEE 802.11e, for the following reasons.

- \* Clients may be unable to stay in sleep mode due to multicast control packets frequently waking them up.
- \* A unicast packet is delayed until an STA wakes up and requests it. Unicast traffic may also 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 associated to the AP 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

This section identifies some representative IETF protocols, and describes possible negative effects due to performance degradation when using multicast transmissions for control messages. Common uses of multicast include:

- \* Control plane signaling
- \* Neighbor Discovery

- \* Address Resolution
- \* Service Discovery
- \* Applications (video delivery, stock data, etc.)
- \* On-demand routing
- \* Backbone construction
- \* Other L3 protocols (non-IP)

User Datagram Protocol (UDP) is the most common transport layer protocol for multicast applications. By itself, UDP is not reliable -- messages may be lost or delivered out of order.

### 3.2.1. IPv4 issues

The following list contains some representative discovery protocols, which utilize broadcast/multicast, that are used with IPv4.

- \* ARP [RFC0826]
- \* DHCP [RFC2131]
- \* mDNS [RFC6762]
- \* uPnP [RFC6970]

After initial configuration, ARP (described in more detail later), DHCP and uPnP occur much less commonly, but service discovery can occur at any time. Some widely-deployed service discovery protocols (e.g., for finding a printer) utilize mDNS (i.e., multicast) which is often dropped by operators. Even if multicast snooping [RFC4541] (which provides the benefit of conserving bandwidth on those segments of the network where no node has expressed interest in receiving packets addressed to the group address) is utilized, many devices can register at once and cause serious network degradation.

### 3.2.2. IPv6 issues

IPv6 makes extensive use of multicast, including the following:

- \* DHCPv6 [RFC8415]
- \* Protocol Independent Multicast (PIM) [RFC7761]
- \* IPv6 Neighbor Discovery Protocol (NDP) [RFC4861]
- \* multicast DNS (mDNS) [RFC6762]
- \* Router Discovery [RFC4286]

IPv6 NDP Neighbor Solicitation (NS) messages used in Duplicate Address Detection (DAD) and Address Lookup make use of Link-Scope multicast. In contrast to IPv4, an IPv6 node will typically use multiple addresses, and may change them often for privacy reasons. This intensifies the impact of multicast messages that are associated to the mobility of a node. Router advertisement (RA) messages are also periodically multicasted over the Link.



Neighbors may be considered lost if several consecutive Neighbor Discovery packets fail.

### 3.2.3. MLD issues

Multicast Listener Discovery (MLD) [RFC4541] is used to identify members of a multicast group that are connected to the ports of a switch. Forwarding multicast frames into a Wi-Fi-enabled area can use 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. A solicited-node multicast address is an IPv6 multicast address used by NDP to verify whether an IPv6 address is already used by the local-link. 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. Some switch vendors do not support MLD, for link-scope multicast, due to the increase it can cause in state.

### 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 routers very often receive packets destined for IPv4 addresses regardless of whether those IP addresses are in use. In the cases where the IP is assigned to a host, the router broadcasts an ARP request, gets back an ARP reply, and caches it; then traffic can be delivered to the host. When 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 the router will rebroadcast the ARP requests.

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, thousands of broadcasts per second have been observed. Around 2,000 broadcasts per second have been observed at the IETF NOC during face-to-face meetings.

With Neighbor Discovery for IPv6 [RFC4861], nodes accomplish address resolution by multicasting a Neighbor Solicitation that asks the target node to return its link-layer address. Neighbor Solicitation messages are multicast to the solicited-node multicast address of the

target address. The target returns its link-layer address in a unicast Neighbor Advertisement message. A single request-response pair of packets is sufficient for both the initiator and the target to resolve each other's link-layer addresses; the initiator includes its link-layer address in the Neighbor Solicitation.

On a wired network, there is not a huge difference between unicast, multicast and broadcast traffic. Due to hardware filtering (see, e.g., [Deri-2010]), inadvertently flooded traffic (or excessive ethernet multicast) on wired networks can be quite a bit less costly, compared to wireless cases where sleeping devices have to wake up to process packets. Wired Ethernets tend to be switched networks, further reducing interference from multicast. There is effectively no collision / scheduling problem except at extremely high port utilizations.

This is not true in the wireless realm; wireless equipment is often unable to send high volumes of broadcast and multicast traffic, causing numerous broadcast and multicast packets to be dropped. Consequently, 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 the MAC address and IP address for all associated STAs. In this way, the AP acts as the central "manager" for all the 802.11 STAs in its basic service set (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 for STA's IP address are handled instead by the AP.
- \* ARP frames are kept off the wireless medium.
- \* No changes are needed to STA implementation.

Here is the specification language as described in clause 10.23.13 of [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. IPv6 Address Registration and Proxy Neighbor Discovery

As used in this section, a Low-Power Wireless Personal Area Network (6LoWPAN) denotes a low power lossy network (LLN) that supports 6LoWPAN Header Compression (HC) [RFC6282]. A 6TiSCH network [I-D.ietf-6tisch-architecture] is an example of a 6LoWPAN. In order to control the use of IPv6 multicast over 6LoWPANs, the 6LoWPAN Neighbor Discovery (6LoWPAN ND) [RFC6775] standard defines an address registration mechanism that relies on a central registry to assess address uniqueness, as a substitute to the inefficient DAD mechanism found in the mainstream IPv6 Neighbor Discovery Protocol (NDP) [RFC4861][RFC4862].

The 6lo Working Group has specified an update [RFC8505] to RFC6775. Wireless devices can register their address to a Backbone Router [I-D.ietf-6lo-backbone-router], which proxies for the registered addresses with the IPv6 NDP running on a high speed aggregating backbone. The update also enables a proxy registration mechanism on behalf of the registered node, e.g. by a 6LoWPAN router to which the mobile node is attached.

The general idea behind the backbone router concept is that broadcast and multicast messaging should be tightly controlled in a variety of WLANs and Wireless Personal Area Networks (WPANs). Connectivity to a particular link that provides the subnet should be left to Layer-3. The model for the Backbone Router operation is represented in Figure 1.

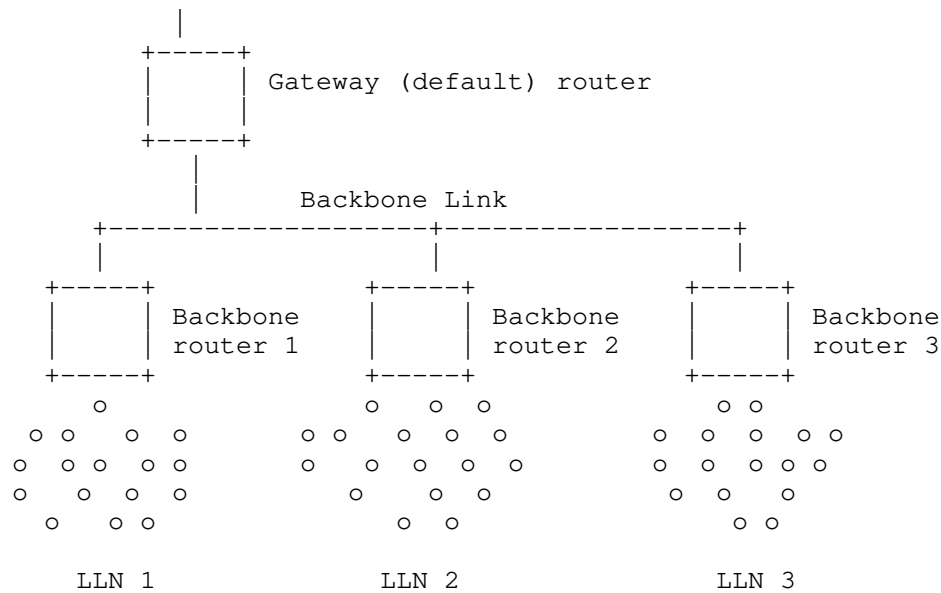


Figure 1: Backbone Link and Backbone Routers

LLN nodes can move freely from an LLN anchored at one IPv6 Backbone Router to an LLN anchored at another Backbone Router on the same backbone, keeping any of the IPv6 addresses they have configured. The Backbone Routers maintain a Binding Table of their Registered Nodes, which serves as a distributed database of all the LLN Nodes. An extension to the Neighbor Discovery Protocol is introduced to exchange Binding Table information across the Backbone Link as needed for the operation of IPv6 Neighbor Discovery.

RFC6775 and follow-on work [RFC8505] address the needs of LLNs, and similar techniques are likely to be valuable on any type of link where sleeping devices are attached, or where the use of broadcast and multicast operations should be limited.

#### 4.3. Buffering to Improve Battery Life

Methods have been developed to help save battery life; for example, a device might not wake up when the AP receives a multicast packet. The AP acts on behalf of STAs in various ways. To enable use of 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. If an AP, for instance, expresses a DTIM (Delivery Traffic Indication Message) of 3 then the AP will send a multicast packet every 3 packets. In fact, when any single wireless STA associated with an access point has 802.11 power-save mode enabled,

the access point buffers all multicast frames and sends them only after the next DTIM beacon.

In practice, most AP's will send a multicast every 30 packets. For unicast the AP could send a TIM (Traffic Indication Message), but for multicast the AP sends a broadcast to everyone. DTIM does power management but STAs can choose whether or not to wake up and whether or not to drop the packet. Unfortunately, without proper administrative control, such STAs may be unable to determine why their multicast operations do not work.

#### 4.4. Limiting multicast buffer hardware queue depth

The CAB (Content after Beacon) queue is used for beacon-triggered transmission of buffered multicast frames. If lots of multicast frames were buffered, and this queue fills up, it drowns out all regular traffic. To limit the damage that buffered traffic can do, some drivers limit the amount of queued multicast data to a fraction of the beacon\_interval. An example of this is [CAB].

#### 4.5. IPv6 support in 802.11-2012

IPv6 uses NDP instead of ARP. Every IPv6 node subscribes to a special multicast address for this purpose.

Here is the specification language from clause 10.23.13 of [dot11-proxyarp]:

"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.

#### 4.6. Using Unicast Instead of Multicast

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

##### 4.6.1. Overview

In many situations, it's a good choice to use unicast instead of multicast over the Wi-Fi link. This avoids most of the problems specific to multicast over Wi-Fi, since the individual frames are then acknowledged and buffered for power save clients, in the way that unicast traffic normally operates.

This approach comes with the tradeoff of sometimes sending the same packet multiple times over the Wi-Fi link. However, in many cases, such as video into a residential home network, this can be a good tradeoff, since the Wi-Fi link may have enough capacity for the unicast traffic to be transmitted to each subscribed STA, even though multicast addressing may have been necessary for the upstream access network.

Several technologies exist that can be used to arrange unicast transport over the Wi-Fi link, outlined in the subsections below.

##### 4.6.2. Layer 2 Conversion to Unicast

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

Although there is not yet a standardized method of conversion, at least one widely available implementation exists in the Linux bridging code [bridge-mc-2-uc]. Other proprietary implementations are available from various vendors. In general, these implementations perform a straightforward mapping for groups or channels, discovered by IGMP or MLD snooping, to the corresponding unicast MAC addresses.

##### 4.6.3. Directed Multicast Service (DMS)

There are situations where more is needed than simply converting multicast to unicast. For these purposes, DMS enables an STA to request that the AP transmit multicast group addressed frames destined to the requesting STAs as individually addressed frames [i.e., convert multicast to unicast]. Here are some characteristics of DMS:

- \* Requires 802.11n A-MSDUs

- \* Individually addressed frames are acknowledged and are buffered for power save STAs
- \* The requesting STA may specify traffic characteristics for DMS traffic
- \* DMS was defined in IEEE Std 802.11v-2011
- \* DMS requires changes to both AP and STA implementation.

DMS is not currently implemented in products. See [Tramarin2017] and [Oliva2013] for more information.

#### 4.6.4. Automatic Multicast Tunneling (AMT)

AMT[RFC7450] provides a method to tunnel multicast IP packets inside unicast IP packets over network links that only support unicast. When an operating system or application running on an STA has an AMT gateway capability integrated, it's possible to use unicast to traverse the Wi-Fi link by deploying an AMT relay in the non-Wi-Fi portion of the network connected to the AP.

It is recommended that multicast-enabled networks deploying AMT relays for this purpose make the relays locally discoverable with the following methods, as described in [I-D.ietf-mboned-driad-amt-discovery]:

- \* DNS-SD [RFC6763]
- \* the well-known IP addresses from Section 7 of [RFC7450]

An AMT gateway that implements multiple standard discovery methods is more likely to discover the local multicast-capable network, instead of forming a connection to a non-local AMT relay further upstream.

#### 4.7. 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 STAs. 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 implementations.

GCR may introduce unacceptable latency. After sending a group of data frames to the group, the AP has to 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 that 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 STAs.

## 5. Operational optimizations

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

### 5.1. Mitigating Problems from Spurious Neighbor Discovery

#### ARP Sponges

An ARP Sponge sits on a network and learns which IP addresses are actually in use. It also listens for ARP requests, and, if it sees an ARP for an IP address that 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 a 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 (one of the most widely deployed arp sponges [arpsponge], was designed to deal with the disappearance of participants from an IXP) and so are not optimized for this purpose. One daemon is needed



per subnet, the tuning is tricky (the scanning rate versus the population rate versus retires, etc.) and sometimes daemons just stop, requiring a restart of the daemon which causes disruption.

#### Router mitigations

Some routers (often those based on Linux) implement a "negative ARP cache" daemon. 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 in use often do not support this. Instead, when a host connects to a 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 learned from the request (passive ARP learning).

#### Firewall unused space

The distribution of users on wireless networks / subnets may change in various use cases, such as conference venues (e.g SSIDs are renamed, some SSIDs lose favor, etc). This makes utilization for particular SSIDs difficult to predict ahead of time, but usage can be monitored as attendees use the different networks. Configuring multiple DHCP pools per subnet, and enabling them sequentially, can create a large subnet, from which only addresses in the lower portions are assigned. Therefore input IP access lists can be applied, which deny traffic to the upper, unused portions. Then 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. Consequently it should be possible 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

Broadcasts can often be caused by outside wifi scanning / backscatter traffic. In order to reduce the impact of broadcasts, NAT can be used on the entire (or a large portion) of a network. This would eliminate NAT translation entries for unused addresses, and the router would never ARP for them. There are, however, many reasons to avoid using NAT in such a blanket fashion.

#### 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. But this conflicts with the need and desire of some organizations to have the network as open as possible and to honor the end-to-end principle. An attendee on a 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.

### 5.2. Mitigating Spurious Service Discovery Messages

In networks that must support hundreds of STAs, operators have observed network degradation due to many devices simultaneously registering with mDNS. In a network with many clients, it is recommended to ensure that mDNS packets designed to discover services in smaller home networks be constrained to avoid disrupting other traffic.

## 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. Unfortunately, 802.15 media specifications do not yet include mechanisms similar to those developed for 802.11. In fact, the design philosophy for 802.15 is oriented towards minimality, with the result that many such functions are 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.

Similar considerations hold for most other wireless media. A brief introduction is provided in [RFC5757] for the following:

- \* 802.16 WIMAX
- \* 3GPP/3GPP2
- \* DVB-H / DVB-IPDC
- \* TV Broadcast and Satellite Networks

## 7. Recommendations

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

Future protocol documents utilizing multicast signaling should be carefully scrutinized if the protocol is likely to be used over wireless media.

The use of proxy methods should be encouraged to conserve network bandwidth and power utilization by low-power devices. The device can use a unicast message to its proxy, and then the proxy can take care of any needed multicast operations.

Multicast signaling for wireless devices should be done in a way compatible with low duty-cycle operation.

## 8. On-going Discussion Items

This section suggests two discussion items for further resolution.

First, standards (and private) organizations should develop guidelines to help clarify when multicast packets would be better served by being sent wired rather than wireless. For example, 802.1ak (<https://www.ieee802.org/1/pages/802.1ak.html>) works on both ethernet and Wi-Fi and organizations could help with deployment decision making by developing guidelines for multicast over Wi-Fi including options for when traffic should be sent wired.

Second, reliable registration to Layer-2 multicast groups, and a reliable multicast operation at Layer-2, might provide a good multicast over wifi solution. There shouldn't be a need to support  $2^{24}$  groups to get solicited node multicast working: it is possible to simply select a number of bits that make sense for a given network size to limit the number of unwanted deliveries to reasonable levels. IEEE 802.1, 802.11, and 802.15 should be encouraged to revisit L2 multicast issues and provide workable solutions.

## 9. Security Considerations

This document does not introduce or modify any security mechanisms. Multicast deployed on wired or wireless networks as discussed in this document can be made more secure in a variety of ways. [RFC4601], for instance, specifies the use of IPsec to ensure authentication of the link-local messages in the Protocol Independent Multicast - Sparse Mode (PIM-SM) routing protocol. [RFC5796] specifies mechanisms to authenticate the PIM-SM link-local messages using the IP security (IPsec) Encapsulating Security Payload (ESP) or (optionally) the Authentication Header (AH).

When using mechanisms that convert multicast traffic to unicast traffic for traversing radio links, the AP (or other entity) is forced to explicitly track which subscribers care about certain multicast traffic. This is generally a reasonable tradeoff, but does result in another entity that is tracking what entities subscribe to which multicast traffic. While such information is already (by necessity) tracked elsewhere, this does present an expansion of the attack surface for that potentially privacy-sensitive information.

As noted in [group\_key], the unreliable nature of multicast transmission over wireless media can cause subtle problems with multicast group key management and updates. When WPA (TKIP) or WPA2 (AES-CCMP) encryption is in use, AP to client (From DS) multicasts have to be encrypted with a separate encryption key that is known to all of the clients (this is called the Group Key). Quoting further from that website, "... most clients are able to get connected and surf the web, check email, etc. even when From DS multicasts are broken. So a lot of people don't realize they have multicast problems on their network..."

This document encourages the use of proxy methods to conserve network bandwidth and power utilization by low-power devices. Such proxy methods in general have security considerations that require the proxy to be trusted to not misbehave. One such proxy method listed is an Arp Sponge which listens for ARP requests, and, if it sees an ARP for an IP address that it believes is not used, it will reply with its own MAC address. ARP poisoning and false advertising could potentially undermine (e.g. DoS) this, and other, proxy approaches.

## 10. IANA Considerations

This document does not request any IANA actions.

## 11. Acknowledgements

This document has benefitted from discussions with the following people, in alphabetical order: Mikael Abrahamsson, Bill Atwood, Stuart Cheshire, Donald Eastlake, Toerless Eckert, Jake Holland, Joel Jaeggli, Jan Komissar, David Lamparter, Morten Pedersen, Pascal Thubert, Jeffrey (Zhaohui) Zhang

## 12. Informative References

### [arpsponge]

Wessel, M. and N. Sijm, "Effects of IPv4 and IPv6 address resolution on AMS-IX and the ARP Sponge", July 2009, <<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.182.4692>>.

### [bridge-mc-2-uc]

Fietkau, F., "bridge: multicast to unicast", January 2017, <<https://github.com/torvalds/linux/commit/6db6f0eae6052b70885562e1733896647ec1d807>>.

### [CAB]

Fietkau, F., "Limit multicast buffer hardware queue depth", 2013, <<https://patchwork.kernel.org/patch/2687951/>>.

### [Deri-2010]

Deri, L. and J. Gasparakis, "10 Gbit Hardware Packet Filtering Using Commodity Network Adapters", RIPE 61, 2010, <[http://ripe61.ripe.net/presentations/138-Deri\\_RIPE\\_61.pdf](http://ripe61.ripe.net/presentations/138-Deri_RIPE_61.pdf)>.

### [dot11]

"IEEE 802 Wireless", "802.11-2016 - IEEE Standard for Information technology--Telecommunications and information exchange between systems Local and metropolitan area networks--Specific requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification (includes 802.11v amendment)", March 2016, <<http://standards.ieee.org/findstds/standard/802.11-2016.html>>.

### [dot11-proxyarp]

Hiertz, G. R., Mestanov, F., and B. Hart, "Proxy ARP in 802.11ax", September 2015, <<https://mentor.ieee.org/802.11/dcn/15/11-15-1015-01-00ax-proxy-arp-in-802-11ax.pptx>>.

[dot11aa] "IEEE 802 Wireless", "Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 2: MAC Enhancements for Robust Audio Video Streaming", March 2012, <[https://standards.ieee.org/standard/802\\_11aa-2012.html](https://standards.ieee.org/standard/802_11aa-2012.html)>.

[group\_key] Spiff, "Why do some WiFi routers block multicast packets going from wired to wireless?", January 2017, <<https://superuser.com/questions/730288/why-do-some-wifi-routers-block-multicast-packets-going-from-wired-to-wireless>>.

[I-D.ietf-6lo-backbone-router] Thubert, P., Perkins, C. E., and E. Levy-Abegnoli, "IPv6 Backbone Router", Work in Progress, Internet-Draft, draft-ietf-6lo-backbone-router-20, 23 March 2020, <<https://www.ietf.org/archive/id/draft-ietf-6lo-backbone-router-20.txt>>.

[I-D.ietf-6tisch-architecture] Thubert, P., "An Architecture for IPv6 over the Time-Slotted Channel Hopping Mode of IEEE 802.15.4 (6TiSCH)", Work in Progress, Internet-Draft, draft-ietf-6tisch-architecture-30, 26 November 2020, <<https://www.ietf.org/archive/id/draft-ietf-6tisch-architecture-30.txt>>.

[I-D.ietf-mboned-driad-amt-discovery] Holland, J., "DNS Reverse IP Automatic Multicast Tunneling (AMT) Discovery", Work in Progress, Internet-Draft, draft-ietf-mboned-driad-amt-discovery-13, 20 December 2019, <<https://www.ietf.org/archive/id/draft-ietf-mboned-driad-amt-discovery-13.txt>>.

[ietf\_802-11] Stanley, D., "IEEE 802.11 multicast capabilities", November 2015, <<https://mentor.ieee.org/802.11/dcn/15/11-15-1261-03-0arc-multicast-performance-optimization-features-overview-for-ietf-nov-2015.ppt>>.

[mc-ack-mux] Tanaka, Y., Sakai, E., Morioka, Y., Mori, M., Hiertz, G., and S. Coffey, "Multiplexing of Acknowledgements for Multicast Transmission", July 2015, <<https://mentor.ieee.org/802.11/dcn/15/11-15-0800-00-00ax-multiplexing-of-acknowledgements-for-multicast-transmission.pptx>>.

- [mc-prob-stmt] Abrahamsson, M. and A. Stephens, "Multicast on 802.11", March 2015, <<https://www.iab.org/wp-content/IAB-uploads/2013/01/multicast-problem-statement.pptx>>.
- [mc-props] Stephens, A., "IEEE 802.11 multicast properties", March 2015, <<https://mentor.ieee.org/802.11/dcn/15/11-15-1161-02-0arc-802-11-multicast-properties.ppt>>.
- [Oliva2013] de la Oliva, A., Serrano, P., Salvador, P., and A. Banchs, "Performance evaluation of the IEEE 802.11aa multicast mechanisms for video streaming", 2013 IEEE 14th International Symposium on "A World of Wireless, Mobile and Multimedia Networks" (WoWMoM) pp. 1-9, June 2013.
- [RFC0826] Plummer, D., "An Ethernet Address Resolution Protocol: Or Converting Network Protocol Addresses to 48.bit Ethernet Address for Transmission on Ethernet Hardware", STD 37, RFC 826, DOI 10.17487/RFC0826, November 1982, <<https://www.rfc-editor.org/info/rfc826>>.
- [RFC2131] Droms, R., "Dynamic Host Configuration Protocol", RFC 2131, DOI 10.17487/RFC2131, March 1997, <<https://www.rfc-editor.org/info/rfc2131>>.
- [RFC4286] Haberman, B. and J. Martin, "Multicast Router Discovery", RFC 4286, DOI 10.17487/RFC4286, December 2005, <<https://www.rfc-editor.org/info/rfc4286>>.
- [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>>.
- [RFC4601] Fenner, B., Handley, M., Holbrook, H., and I. Kouvelas, "Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)", RFC 4601, DOI 10.17487/RFC4601, August 2006, <<https://www.rfc-editor.org/info/rfc4601>>.
- [RFC4861] Narten, T., Nordmark, E., Simpson, W., and H. Soliman, "Neighbor Discovery for IP version 6 (IPv6)", RFC 4861, DOI 10.17487/RFC4861, September 2007, <<https://www.rfc-editor.org/info/rfc4861>>.

- [RFC4862] Thomson, S., Narten, T., and T. Jinmei, "IPv6 Stateless Address Autoconfiguration", RFC 4862, DOI 10.17487/RFC4862, September 2007, <<https://www.rfc-editor.org/info/rfc4862>>.
- [RFC5424] Gerhards, R., "The Syslog Protocol", RFC 5424, DOI 10.17487/RFC5424, March 2009, <<https://www.rfc-editor.org/info/rfc5424>>.
- [RFC5757] Schmidt, T., Waehlich, M., and G. Fairhurst, "Multicast Mobility in Mobile IP Version 6 (MIPv6): Problem Statement and Brief Survey", RFC 5757, DOI 10.17487/RFC5757, February 2010, <<https://www.rfc-editor.org/info/rfc5757>>.
- [RFC5796] Atwood, W., Islam, S., and M. Siami, "Authentication and Confidentiality in Protocol Independent Multicast Sparse Mode (PIM-SM) Link-Local Messages", RFC 5796, DOI 10.17487/RFC5796, March 2010, <<https://www.rfc-editor.org/info/rfc5796>>.
- [RFC6282] Hui, J., Ed. and P. Thubert, "Compression Format for IPv6 Datagrams over IEEE 802.15.4-Based Networks", RFC 6282, DOI 10.17487/RFC6282, September 2011, <<https://www.rfc-editor.org/info/rfc6282>>.
- [RFC6762] Cheshire, S. and M. Krochmal, "Multicast DNS", RFC 6762, DOI 10.17487/RFC6762, February 2013, <<https://www.rfc-editor.org/info/rfc6762>>.
- [RFC6763] Cheshire, S. and M. Krochmal, "DNS-Based Service Discovery", RFC 6763, DOI 10.17487/RFC6763, February 2013, <<https://www.rfc-editor.org/info/rfc6763>>.
- [RFC6775] Shelby, Z., Ed., Chakrabarti, S., Nordmark, E., and C. Bormann, "Neighbor Discovery Optimization for IPv6 over Low-Power Wireless Personal Area Networks (6LoWPANs)", RFC 6775, DOI 10.17487/RFC6775, November 2012, <<https://www.rfc-editor.org/info/rfc6775>>.
- [RFC6970] Boucadair, M., Penno, R., and D. Wing, "Universal Plug and Play (UPnP) Internet Gateway Device - Port Control Protocol Interworking Function (IGD-PCP IWF)", RFC 6970, DOI 10.17487/RFC6970, July 2013, <<https://www.rfc-editor.org/info/rfc6970>>.
- [RFC7450] Bumgardner, G., "Automatic Multicast Tunneling", RFC 7450, DOI 10.17487/RFC7450, February 2015, <<https://www.rfc-editor.org/info/rfc7450>>.



- [RFC7761] Fenner, B., Handley, M., Holbrook, H., Kouvelas, I., Parekh, R., Zhang, Z., and L. Zheng, "Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)", STD 83, RFC 7761, DOI 10.17487/RFC7761, March 2016, <<https://www.rfc-editor.org/info/rfc7761>>.
- [RFC8415] Mrugalski, T., Siodelski, M., Volz, B., Yourtchenko, A., Richardson, M., Jiang, S., Lemon, T., and T. Winters, "Dynamic Host Configuration Protocol for IPv6 (DHCPv6)", RFC 8415, DOI 10.17487/RFC8415, November 2018, <<https://www.rfc-editor.org/info/rfc8415>>.
- [RFC8505] Thubert, P., Ed., Nordmark, E., Chakrabarti, S., and C. Perkins, "Registration Extensions for IPv6 over Low-Power Wireless Personal Area Network (6LoWPAN) Neighbor Discovery", RFC 8505, DOI 10.17487/RFC8505, November 2018, <<https://www.rfc-editor.org/info/rfc8505>>.
- [Tramarin2017] Tramarin, F., Vitturi, S., and M. Luvisotto, "IEEE 802.11n for Distributed Measurement Systems", 2017 IEEE International Instrumentation and Measurement Technology Conference (I2MTC) pp. 1-6, May 2017.
- [uli] Kinney, P., "LLC Proposal for 802.15.4", November 2015, <<https://mentor.ieee.org/802.15/dcn/15/15-15-0521-01-wng0-llc-proposal-for-802-15-4.pptx>>.

#### Authors' Addresses

Charles E. Perkins  
Blue Meadow Networks

Phone: +1-408-330-4586  
Email: [charliep@computer.org](mailto:charliep@computer.org)

Mike McBride  
Futurewei Technologies Inc.  
2330 Central Expressway  
Santa Clara, CA 95055  
United States of America

Email: [michael.mcbride@futurewei.com](mailto:michael.mcbride@futurewei.com)

Dorothy Stanley  
Hewlett Packard Enterprise  
2000 North Naperville Rd.  
Naperville, IL 60566  
United States of America

Phone: +1 630 979 1572  
Email: dstanley1389@gmail.com

Warren Kumari  
Google  
1600 Amphitheatre Parkway  
Mountain View, CA 94043  
United States of America

Email: warren@kumari.net

Juan Carlos Zuniga  
SIGFOX  
425 rue Jean Rostand  
31670 Labège  
France

Email: j.c.zuniga@ieee.org

Network Working Group  
Internet-Draft  
Intended status: Experimental  
Expires: January 9, 2020

K. Rose  
J. Holland  
Akamai Technologies, Inc.  
July 08, 2019

Asymmetric Loss-Tolerant Authentication  
draft-krose-mboned-alta-01

Abstract

Establishing authenticity of a stream of datagrams in the presence of multiple receivers is naively achieved through the use of per-packet asymmetric digital signatures, but at high computational cost for both senders and receivers. Timed Efficient Stream Loss-Tolerant Authentication (TESLA) instead employs relatively cheap symmetric authentication, achieving asymmetry via time-delayed key disclosure, while adding latency to verification and imposing requirements on time synchronization between receivers and the sender to prevent forgery. This document introduces Asymmetric Loss-Tolerant Authentication (ALTA), which employs an acyclic graph of message authentication codes (MACs) transmitted alongside data payloads, with redundancy to enable authentication of all received payloads in the presence of certain patterns of loss, along with regularly paced digital signatures. ALTA requires no time synchronization and enables authentication of payloads as soon as sufficient authentication material has been received.

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 January 9, 2020.

## Copyright Notice

Copyright (c) 2019 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 Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

## Table of Contents

1. Introduction . . . . .	2
2. Conventions and Definitions . . . . .	4
3. Protocol Overview . . . . .	4
4. Protocol Details . . . . .	6
4.1. ALTA Payload . . . . .	6
4.1.1. Authentication Tag . . . . .	7
4.2. Digital Signature . . . . .	8
4.2.1. Application Data . . . . .	8
4.3. Scheme Construction . . . . .	8
5. ALTA Configuration . . . . .	9
5.1. Performance Considerations . . . . .	9
5.1.1. MAC selection . . . . .	9
5.1.2. Digital signature selection . . . . .	9
5.2. Out-of-band Metadata . . . . .	9
6. Operational Considerations . . . . .	9
7. Security Considerations . . . . .	9
7.1. Parsing an ill-formed or inconsistent payload . . . . .	9
7.2. Index overflow . . . . .	9
7.3. Truncated MACs . . . . .	9
8. IANA Considerations . . . . .	9
9. References . . . . .	9
9.1. Normative References . . . . .	9
9.2. Informative References . . . . .	9
Acknowledgments . . . . .	10
Authors' Addresses . . . . .	10

## 1. Introduction

Authenticity of streaming data may be inexpensively established via symmetric message authentication codes (MACs) using keys pre-shared exclusively between two parties, as the receiver knows it did not

originate the data and that only one other party has access to the key. In the presence of multiple receivers, however, this is not possible because all receivers must have access to the same key, giving any one of them the ability to forge messages. Consequently, authentication must be made asymmetric, such that only the sender has the ability to produce messages that correct receivers will verify as authentic.

Naively, a sender may sign individual datagrams using an asymmetric digital signature algorithm, such as RSA or Ed25519, but this carries high computational cost for both the sender and receivers. In the case of streaming video delivery, while the sender's computational load may be dominated by CPU-intensive video encoding, the receiver is often a device with hardware dedicated to efficient video decoding and with limited general purpose computing hardware and/or battery available for high-rate digital signature authentication.

Timed Efficient Stream Loss-Tolerant Authentication (TESLA) [RFC4082] addresses this problem through the use of symmetric authentication by delaying the release of keying material to a deadline at which any packets protected by said key that are subsequently received must be discarded by a receiver. While this reintroduces asymmetry between sender and receiver, it requires the sender and each receiver to (loosely) synchronize clocks and imposes authentication latency relative to RTT and to a pre-declared upper bound on clock skew.

Clock synchronization is not as trivial as it appears: internet-connected hosts often have significant clock skew relative to stratum 0 NTP servers [timeskew], and anyway enterprises serving valuable assets do not regard NTP as a reliable interdomain security protocol. Together with the need to avoid attacks that delay packets required for synchronization, this implies the need for an interactive unicast authenticated clock synchronization protocol, which is complicated by the need to maintain clock synchronization across both the stream publisher and multiple geographically-distributed nodes in a content delivery network (CDN).

This document introduces Asymmetric Loss-Tolerant Authentication (ALTA), which eschews time synchronization for an application of digital signatures to an acyclic graph of symmetric message authentication codes with redundancy sufficient to tolerate certain patterns of loss, and with digital signature authentication load greatly reduced relative to the naive approach. This algorithm is based on research by Golle and Modadugu, as published in [STRAUTH]. Live multicast streaming over an unreliable transport is the intended application for ALTA: object-based integrity solutions or transport security may be more appropriate for unicast transmission or for static objects pulled on-demand.

## 2. Conventions and Definitions

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.

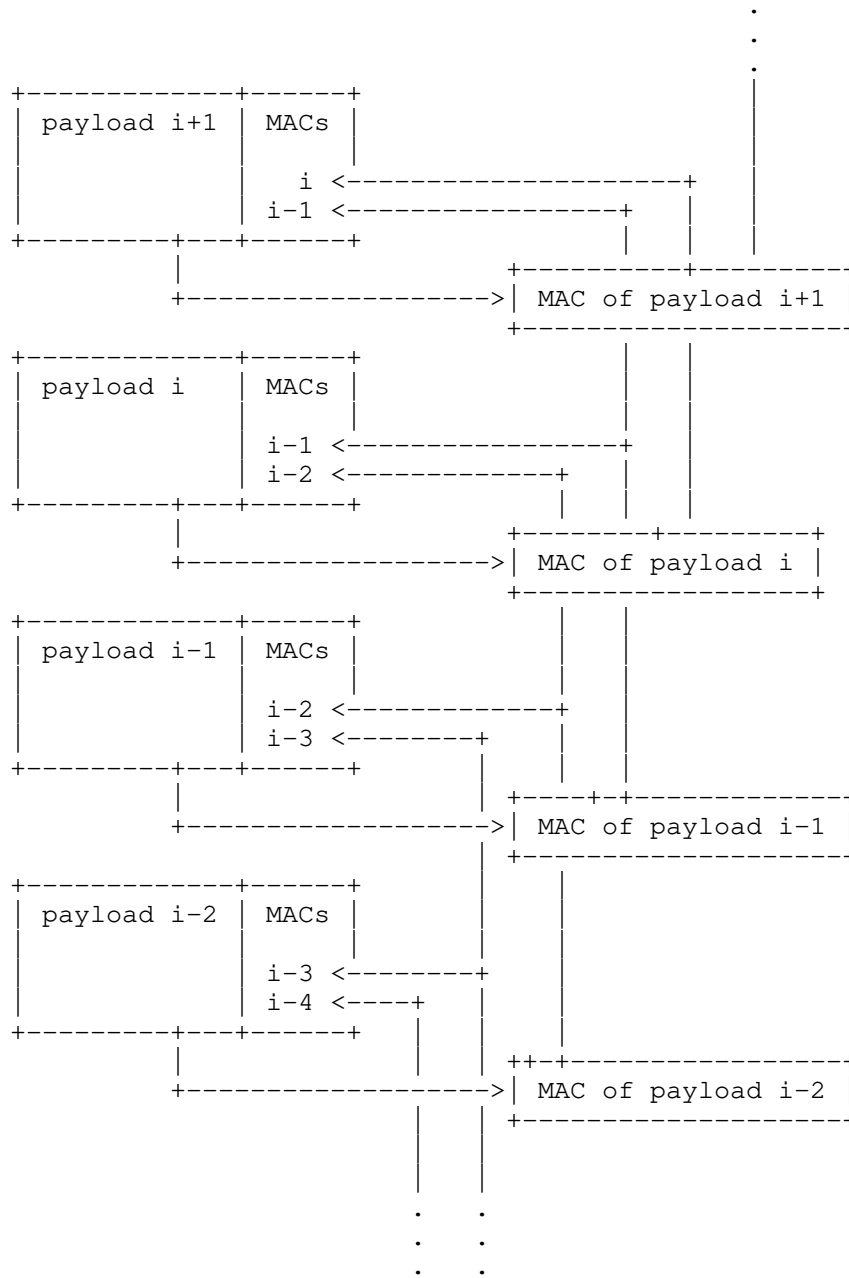
## 3. Protocol Overview

ALTA is intended for streaming datagram use cases in which the receiving application has a deadline for the utility of received data and can tolerate a degree of random packet loss. It combines a segment of application data with a variable-length authentication tag into an ALTA payload to be sent as a unit in a single datagram, with the authentication tags constructed in such a way that a receiver will be able to authenticate nearly all such ALTA payloads received by the deadline under certain patterns of random packet loss.

An authentication tag is a combination of zero or more symmetric message authentication codes (MACs) and either zero or one digital signature. Each MAC is of another ALTA payload in the stream, while the digital signature is of the containing ALTA payload with the signature field itself replaced by all zeroes.

The MACs included in a given authentication tag are determined by a scheme, as defined in section 3 of [STRAUTH]. Conceptually, a scheme is a mostly backward-looking directed acyclic graph of ALTA payloads such that the MAC of a given payload is contained in two or more other payloads in the stream, enabling the loss of one of these to be tolerated without losing the ability to authenticate the given payload.

For purposes of illustration, a simple example scheme is one in which the  $i$ th ALTA payload's authentication tag contains MACs for the  $(i-1)$ th and  $(i-2)$ th payload:



The recommended scheme is more complex and will be covered in detail in Section 4.3.

Encoding a scheme relies on ALTA payloads being addressable deterministically by an index even in the presence of reordering or loss. This index may be deduced from the application data (e.g., making use of an existing sequence number) or by a payload index explicitly encoded in the authentication tag. Two modes are supported:

- o If the index starts at zero and increments by exactly one for each payload in the stream, and if the scheme is known to both sender and receiver, then indices are not required to be encoded for each MAC in an authentication tag as they can be deduced from a given payload's index and from the DAG associated with the scheme. Hereafter, this is referred to as `_implicit offset mode_`.
- o If the index increments unpredictably, or if the scheme is not known to the receiver, then each MAC in an authentication tag must be paired with the explicit index of the ALTA payload from which the MAC is computed. For compactness, this index will be encoded as an offset relative to the index of the containing payload. Hereafter this is referred to as `_explicit offset mode_`.

Authenticity of a payload is established by a chain of MACs rooted in an ALTA payload whose authentication tag contains a digital signature created by a key in which trust has been established out-of-band. Delivery of application data must be delayed until a payload has been authenticated. Note that a given payload may be authenticated by a digital signature as well as by one or more MAC chains; within authentication deadline constraints, receivers should prefer to authenticate by MAC, minimizing the computational load imposed by digital signature authentication.

The variable length of authentication tags in ALTA has implications for application data segmentation when constant-length datagrams are desired (e.g., to maximize data per UDP packet with a given path MTU while avoiding fragmentation).

## 4. Protocol Details

### 4.1. ALTA Payload

An ALTA payload comprises the following elements (defined below) concatenated in-order:

- o Authentication tag
  - \* Options octet
  - \* Optional payload index



- \* Sequence of chained MACs
- \* Optional digital signature
- o Application data

#### 4.1.1.1. Authentication Tag

The authentication tag is the metadata emitted by an ALTA-compliant sender that is required, in combination with other out-of-band metadata, by an ALTA-compliant receiver to authenticate a stream of packets in a manner tolerant to loss and reordering.

##### 4.1.1.1.1. Options Octet

```

  0 1 2 3 4 5 6 7
+---+---+---+---+
|MACct|S| rsvd |
+---+---+---+---+
```

Figure 1: Options Octet

The first octet of the authentication tag contains the count of MACs included ("MACct") as well as a flag "S" indicating whether the tag also contains a digital signature. It also contains four reserved bits which MUST be set to 0 by senders and ignored by receivers.

##### 4.1.1.1.2. Payload Index

```

      0                                     1
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 ...
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               payload index ...
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
```

Figure 2: Payload Index

If the payload index cannot be deduced from the application data in this payload, it must be specified explicitly in the authentication tag as an unsigned quantity of a fixed length specified by out-of-band metadata.

Whether explicit or deduced, the payload index uniquely identifies a single ALTA stream payload within a rollover window of size  $2^N$  for some "N" specified in out-of-band metadata. The payload index MUST start at zero and increment by one for each payload transmitted, with rollover to zero on overflow.

## 4.1.1.3. Chained MACs

```

      0                               1
    0 1 2 3 4 5 6 7 8 9 0 1 ...
+---+---+---+---+---+---+---+---+
|   offset   | MAC ...
+---+---+---+---+---+---+---+---+

```

Figure 3: Example MAC with explicit index

In explicit offset mode, each MAC encoded in the payload comprises an offset from the payload's index, expressed as a signed octet in two's complement, followed by a fixed-length MAC. The length and semantics of the MAC are a function of the MAC algorithm, which is specified by out-of-band metadata. The offset space given in the example in Figure 3 is one octet, ranging from -128 to 127, but may be of any number of whole octets, as specified by out-of-band metadata.

In implicit offset mode, the receiver knows the scheme being employed and so can deduce the indices of the chained MACs from the current payload's index. Consequently, the MACs are simply concatenated in ascending order of source index according to the scheme.

## 4.2. Digital Signature

```

    0 1 2 3 4 5 6 ...
+---+---+---+---+---+---+
| signature ...
+---+---+---+---+---+---+

```

Figure 4: Digital Signature

If "S=1" in the options octet, then a digital signature is included in the tag. The length and content of this digital signature are a function of the signature algorithm, which is specified by out-of-band metadata.

## 4.2.1. Application Data

The application data is opaque, with the exception of the payload index if not specified explicitly in the authentication tag.

## 4.3. Scheme Construction

In the ALTA context, a scheme describes the directed acyclic graph of payload MACs embedded in other payloads for purposes of chained authentication. The recommended scheme is that described in section 3.2 of [STRAUTH], with "a=3" and "p=5".

FIXME: Describe how to construct this scheme in pseudocode.

## 5. ALTA Configuration

### 5.1. Performance Considerations

#### 5.1.1. MAC selection

#### 5.1.2. Digital signature selection

### 5.2. Out-of-band Metadata

## 6. Operational Considerations

As ALTA requires an out-of-band channel for provisioning of metadata, including digital signature keys and cryptographic algorithms, versioning of the protocol to support a future ALTA revision may be performed there and acted upon by the application.

## 7. Security Considerations

### 7.1. Parsing an ill-formed or inconsistent payload

### 7.2. Index overflow

### 7.3. Truncated MACs

## 8. IANA Considerations

This document has no IANA actions.

## 9. References

### 9.1. Normative References

[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>>.

### 9.2. Informative 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>>.

[RFC4082] Perrig, A., Song, D., Canetti, R., Tygar, J., and B. Briscoe, "Timed Efficient Stream Loss-Tolerant Authentication (TESLA): Multicast Source Authentication Transform Introduction", RFC 4082, DOI 10.17487/RFC4082, June 2005, <<https://www.rfc-editor.org/info/rfc4082>>.

[STRAUTH] Modadugu, N., "Authenticating Streamed Data in the Presence of Random Packet Loss", 2001, <<https://crypto.stanford.edu/~pgolle/papers/auth.pdf>>.

ISOC Network and Distributed System Security Symposium

[timeskew]

"FIXME reference for how bad time sync is", n.d..

#### Acknowledgments

The author wishes to acknowledge Eric Rescorla, who introduced the author to the paper describing the loss-tolerant symmetric authentication scheme used as the basis for ALTA.

#### Authors' Addresses

Kyle Rose  
Akamai Technologies, Inc.  
150 Broadway  
Cambridge, MA 02144  
United States of America

Email: [krose@krose.org](mailto:krose@krose.org)

Jake Holland  
Akamai Technologies, Inc.  
150 Broadway  
Cambridge, MA 02144  
United States of America

Email: [jakeholland.net@gmail.com](mailto:jakeholland.net@gmail.com)