

MBONED WG
Internet-Draft
Intended status: Standards Track
Expires: April 29, 2018

M. McBride
C. Perkins
Huawei
October 26, 2017

Multicast Wifi Problem Statement
draft-mcbride-mboned-wifi-mcast-problem-statement-01

Abstract

There have been known issues with multicast, in an 802.11 environment, which have prevented the deployment of multicast in these wifi environments. IETF multicast experts have been meeting together to discuss these issues and provide IEEE updates. The mboned working group is chartered to receive regular reports on the current state of the deployment of multicast technology, create "practice and experience" documents that capture the experience of those who have deployed and are deploying various multicast technologies, and provide feedback to other relevant working groups. As such, this document will gather the problems of wifi multicast into one problem statement document so as to offer the community guidance on current limitations.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <https://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on April 29, 2018.

Copyright Notice

Copyright (c) 2017 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. Multicast over WiFi Problems	2
2.1. Low Reliability	3
2.2. Low Data Rate	4
2.3. High Interference	4
2.4. High Power Consumption	4
3. Common remedies to multicast over wifi problems	4
4. State of the Union	5
5. IANA Considerations	6
6. Security Considerations	6
7. Acknowledgments	6
8. Normative References	6
Authors' Addresses	6

1. Introduction

Multicast over wifi has been used to low levels of success, usually to a point of being so negative that multicast over wifi is not allowed. In addition to protocol use of broadcast/multicast for control messages, more applications, such as push to talk in hospitals, video in enterprises and lectures in Universities, are streaming over wifi. And many end devices are increasingly using wifi for their connectivity. One of the primary problems multicast over wifi faces is that link local 802.11 doesn't necessarily support multicast, it supports broadcast. To make make multicast over wifi work successfully we often need to modify the multicast to instead be sent as unicast in order for it to successfully transmit with useable quality. Multicast over wifi experiences high packet error rates, no acknowledgements, and low data rate. This draft reviews these problems found with multicast over wifi. While this is not a solutions draft, common workarounds to some of the problems will be listed, along with the impact of the workarounds.

2. Multicast over WiFi Problems

802.11 is a wireless broadcast medium which works well for unicast and has become ubiquitous in its use. With multicast, however, problems arise over wifi. There are no ACKs for multicast packets,

for instance, so there can be a high level of packet error rate (PER) due to lack of retransmission and because the sender never backs off. It is not uncommon for there to be a packet loss rate of 5% which is particularly troublesome for video and other environments where high data rates and high reliability are required. Multicast, over wifi, is typically sent on a low data rate which makes video negatively impacted. Wifi loses many more packets than wired due to collisions and signal loss. There are also problems because clients are unable to stay in sleep mode due to the multicast control packets continuing to unnecessarily wake up those clients which subsequently reduces energy savings. Video is becoming the dominant content for end device applications, with multicast being the most natural method for applications to transmit video. Unfortunately, multicast, even though it is a very natural choice for video, incurs a large penalty over wifi.

One big difference between multicast over wired versus multicast over wifi is that wired links are a fixed transmission rate. Wifi, on the other hand, has a transmission rate which varies over time depending upon the clients proximity to the AP. Throughput of video flows, and the capacity of the broader wifi network, will change and will impact the ability for QoS solutions to effectively reserve bandwidth and provide admission control.

The main problems associated with multicast over WiFi are as follows:

- o Low Reliability
- o Lower Data Rate
- o High interference
- o High Power Consumption

These points will be elaborated separately in the following subsections.

2.1. Low Reliability

Because of the lack of acknowledgement for packets from Access Point to the receivers, it is not possible for the Access Point to know whether or not a retransmission is needed. Even in the wired Internet, this characteristic commonly causes undesirably high error rates, contributing to the relatively slow uptake of multicast applications even though the protocols have been available for decades. The situation for wireless links is much worse, and is quite sensitive to the presence of background traffic.

2.2. Low Data Rate

For wireless stations associated with an Access Points, the necessary power 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. For this purpose, 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. 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.

2.3. High Interference

As mentioned in the previous subsection, multicast transmission to the stations associated to an Access Point typically proceeds at a much higher power level than is required for unicast to many of the receivers. High power levels directly contribute to stronger interference. The interference due to multicast may extend to effects inhibiting packet reception at more distant stations that might even be associated with other Access Points. Moreover, the use of lower data rates implies that the physical medium will be occupied for a longer time to transmit a packet than would be required at high data rates. Thus, the level of interference due to multicast will be not only higher, but longer in duration.

Depending on the choice of 802.11 technology, and the configured choice for the base data rate for multicast transmission from the Access Point, the amount of additional interference can range from a factor of ten, to a factor thousands for 802.11ac.

2.4. High Power Consumption

One of the characteristics of multicast transmission is that every station has to be configured to wake up to receive the multicast, even though the received packet may ultimately be discarded. This process has a relatively large impact on the power consumption by the multicast receiver station.

3. Common remedies to multicast over wifi problems

One common solution to the multicast over wifi problem is to convert the multicast traffic into unicast. This is often referred to as multicast to unicast (MC2UC). Converting the packets to unicast is beneficial because unicast packets are acknowledged and retransmitted as needed to prevent as much loss. The Access Points (AP) is also

able to provide rate limiting as needed. The drawback with this approach is that the benefit of using multicast is defeated.

Using 802.11n helps provide a more reliable and higher level of signal-to-noise ratio in a wifi environment over which multicast (broadcast) packets can be sent. This can provide higher throughput and reliability but the broadcast limitations remain.

4. State of the Union

In discussing these issues over email and, most recently, in a side meeting at IETF 99, it is generally agreed that these problems will not be fixed anytime soon primarily because it's expensive to do so and multicast is unreliable. The problem of v6 neighbor discovery saturating the wifi link is only part of the problem. A big problem is that the 802.11 multicast channel is an afterthought and only given 100th of the bandwidth. Multicast is basically a second class citizen, to unicast, over wifi. Unicast may have allocated 10mb while Multicast will be allocated 1mb. There are many protocols using multicast and there needs to be something provided in order to make them more reliable. Wifi traffic classes may help. We need to determine what problem should be solved by the IETF and what problem should be solved by the IEEE.

Apple's Bonjour protocol, for instance, provides service discovery (for printing) that utilizes multicast. It's the first thing operators drop. Even if multicast snooping is utilized, everyone registers at once using Bonjour and the network has serious degradation. There is also a lot of work being developed to help save battery life such as the devices not waking up when receiving a multicast packet. If an AP, for instance, expresses a DTIM of 3 then it will send a multicast packet every 3 packets. But the reality is that most AP's will send a multicast every 30 packets. For unicast there's a TIM. But because multicast is going to everyone, the AP sends a broadcast to everyone. DTIM does power management but clients can choose to wake up or not and whether to drop the packet or not. Then they don't know why their Bonjour doesn't work. The IETF may just need to decide that broadcast is more expensive so multicast needs to be sent wired. 802.1ak works on ethernet and wifi. 802.1ak has been pulled into 802.1Q as of 802.1Q-2011. 802.1Q-2014 can be looked at here: <http://www.ieee802.org/1/pages/802.1Q-2014.html>. If we don't find a generic solution we need to establish guidelines for multicast over wifi within the mboned wg. A multicast over wifi IETF mailing list is formed (mcast-wifi@ietf.org) and more discussion to be had there. This draft will serve as the current state of affairs.

This is not a solutions draft, but to provide an idea going forward, a reliable registration to Layer-2 multicast groups and a reliable multicast operation at Layer-2 could provide a generic solution. There is no need to support 2^{24} groups to get solicited node multicast working: it is possible to simply select a number of trailing bits that make sense for a given network size to limit the amount of unwanted deliveries to reasonable levels. We need to encourage IEEE 802.1 and 802.11 to revisit L2 multicast issues. In particular, Wi-Fi provides a broadcast service, not a multicast one. In fact all frames are broadcast at the PHY level unless we beamform. What comes with unicast is the property of being much faster (2 orders of magnitude) and much more reliable (L2 ARQ).

5. IANA Considerations

None

6. Security Considerations

None

7. Acknowledgments

The following people have contributed information and discussion in the meetings and on the list which proved helpful for the development of the latest version this Internet Draft:

Dave Taht, Donald Eastlake, Pascal Thubert, Juan Carlos Zuniga, Mikael Abrahamsson, Diego Dujovne, David Schinazi, Stig Venaas, Stuart Cheshire, Lorenzo, Greg Shephard, Mark Hamilton

8. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <<https://www.rfc-editor.org/info/rfc2119>>.

Authors' Addresses

Mike McBride
Huawei
2330 Central Expressway
Santa Clara CA 95055
USA

Email: michael.mcbride@huawei.com

Charlie Perkins
Huawei
2330 Central Expressway
Santa Clara CA 95055
USA

Email: charlie.perkins@huawei.com

Internet Area
Internet-Draft
Intended status: Informational
Expires: January 20, 2018

C. Perkins
Futurewei
D. Stanley
HPE
W. Kumari
Google
JC. Zuniga
SIGFOX
July 19, 2017

Multicast Considerations over IEEE 802 Wireless Media
draft-perkins-intarea-multicast-ieee802-03

Abstract

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, 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 choices that can be taken to improve the performance of the network. Finally, it provides some recommendations 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 <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 January 20, 2018.

Copyright Notice

Copyright (c) 2017 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 (<http://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 mulitcast issues	4
3.1. Issues at Layer 2 and Below	4
3.1.1. Multicast reliability	4
3.1.2. Lower Data Rate	4
3.1.3. Power-save Effects on Multicast	5
3.2. Issues at Layer 3 and Above	5
3.2.1. IPv4 issues	5
3.2.2. IPv6 issues	5
3.2.3. MLD issues	6
3.2.4. Spurious Neighbor Discovery	6
4. Multicast protocol optimizations	7
4.1. Proxy ARP in 802.11-2012	7
4.2. IPv6 Address Registration and Proxy Neighbor Discovery	8
4.3. Buffering to improve Power-Save	9
4.4. IPv6 support in 802.11-2012	10
4.5. Conversion of multicast to unicast	10
4.6. Directed Multicast Service (DMS)	10
4.7. GroupCast with Retries (GCR)	11
5. Operational optimizations	11
5.1. Mitigating Problems from Spurious Neighbor Discovery	12
6. Multicast Considerations for Other Wireless Media	14
7. Recommendations	14
8. Security Considerations	14
9. IANA Considerations	14
10. Acknowledgements	14
11. Informative References	14
Authors' Addresses	16

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 minimizing media occupancy 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 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 wireless 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 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 [REF 11-15-1261-03]. In addition to protocol design features, certain operational and configuration enhancements can ameliorate the network performance issues created by multicast traffic.

This Internet Draft details various problems caused by multicast transmission over wireless networks. It also explains some 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. Recommendations about how to use and combine these enhancements are also provided.

2. Terminology

This document uses the following definitions:

AP

IEEE 802.11 Access Point.

STA

802.11 station (e.g. handheld device).

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 recipients.

3.1.2. Lower Data Rate

Because more robust MCSs have longer range but also lower data rate, multicast / broadcast traffic is generally transmitted at the lowest common denominator rate, also known as the basic rate. On IEEE 802.11 networks (aka WiFi), 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 longer occupancy of the wireless medium and thus take away from the airtime of other communications and degrade the overall capacity.

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 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. Power-save Effects on Multicast

Multicast can work poorly with the power-save mechanisms defined in IEEE 802.11.

- o Both unicast and multicast traffic can be delayed by power-saving mechanisms.
- o A unicast packet is delayed until a STA wakes up and requests it. Unicast traffic may also be delayed to improve power save, efficiency and increase probability of aggregation.
- o 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.
- o 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:

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

3.2.1. IPv4 issues

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

- o ARP
- o DHCP
- o mDNS

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

3.2.2. IPv6 issues

IPv6 makes much more extensive use of multicast, including the following:

- o DHCPv6
- o IPv6 Neighbor Discovery Protocol (NDP) is not very tolerant of packet losses. In particular, the Duplicate Address Detection

- (DAD) process fails when the owner of an address does not receive the multicast DAD message from another node that wishes to own that same address. This can result in an address being duplicated in the subnet, breaking a basic assumption of IPv6 connectivity.
- o IPv6 NDP Neighbor Solicitation (NS) messages used in 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 multiplies the impact of multicast messages that are associated to the mobility of a Node. Router advertisement (RA) messages are also periodically multicast over the Link.
 - o Neighbors 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 routers very often receive packets destined for machines whose IP addresses may or may not be 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, 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 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 BSS. Proxy ARP is easy to implement at the AP, and offers the following advantages:

- o Reduced broadcast traffic (transmitted at low MCS) on the wireless medium
- o STA benefits from extended power save in sleep mode, as ARP requests for STA's IP address are handled instead by the AP.
- o ARP frames are kept off the wireless medium.
- o 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 Duplicate Address Detection (DAD) mechanism found in the mainstream IPv6 Neighbor Discovery Protocol (NDP) [RFC4861][RFC4862].

The 6lo Working Group is now completing an update [I-D.ietf-6lo-rfc6775-update] to RFC6775. The update enables the registration to a Backbone Router [I-D.ietf-6lo-backbone-router], which proxies for the registered addresses with the mainstream IPv6 NDP running on a high speed aggregating backbone. The update also enables a proxy registration 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 in a variety of Wireless Local Area Networks (WLANs) and Wireless Personal Area Networks (WPANs), the broadcast/multicast domain should be controlled, and 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.

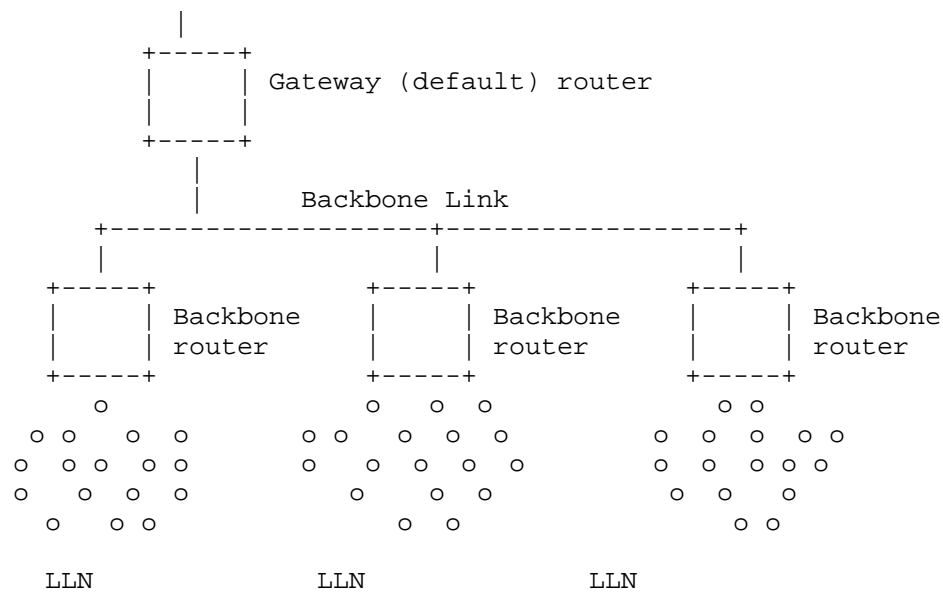


Figure 1: Backbone Link and Backbone Routers

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 that information across the Backbone Link in the reactive fashion of mainstream IPv6 Neighbor Discovery.

RFC6775 and follow-on work are designed to address the needs of LLNs, but the 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 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.

4.4. IPv6 support in 802.11-2012

IPv6 uses Neighbor Discovery Protocol (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

- o Maximum Transmission Unit
- o Router Solicitation
- o 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.

4.5. Conversion of multicast to unicast

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

4.6. Directed Multicast Service (DMS)

There are situations where more is needed than simply converting multicast to unicast. 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]. Here are some characteristics of DMS:

- o Requires 802.11n A-MSDUs
- o Individually addressed frames are acknowledged and are buffered for power save clients
- o The requesting STA may specify traffic characteristics for DMS traffic
- o DMS was defined in IEEE Std 802.11v-2011
- o DMS requires changes to both AP and STA implementation.

DMS is not currently implemented in products.

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 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:

- o unicast a Block Ack Request (BAR) to a subset of members.
- o wait for the corresponding Block Ack (BA).
- o retransmit any missed frames.
- o 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.

- o BAR is sent using downlink MU-MIMO (note that downlink MU-MIMO is already specified in 802.11-REVmc 4.3).
- o BA is sent using uplink MU-MIMO (which is a .11ax feature).
- o Additional 802.11ax extensions are under consideration; see [mc-ack-mux]
- o 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 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 (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 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 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. Acknowledgements

This document has benefitted from discussions with the following people, in alphabetical order: Pascal Thubert

11. 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.
- [I-D.ietf-6lo-ap-nd]
Sarikaya, B., Thubert, P., and M. Sethi, "Address Protected Neighbor Discovery for Low-power and Lossy Networks", draft-ietf-6lo-ap-nd-02 (work in progress), May 2017.
- [I-D.ietf-6lo-backbone-router]
Thubert, P., "IPv6 Backbone Router", draft-ietf-6lo-backbone-router-04 (work in progress), July 2017.
- [I-D.ietf-6lo-rfc6775-update]
Thubert, P., Nordmark, E., and S. Chakrabarti, "An Update to 6LoWPAN ND", draft-ietf-6lo-rfc6775-update-06 (work in progress), June 2017.
- [I-D.ietf-6tisch-architecture]
Thubert, P., "An Architecture for IPv6 over the TSCH mode of IEEE 802.15.4", draft-ietf-6tisch-architecture-11 (work in progress), January 2017.
- [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>>.

- [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, <<http://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, <<http://www.rfc-editor.org/info/rfc4862>>.
- [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, <<http://www.rfc-editor.org/info/rfc6282>>.
- [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, <<http://www.rfc-editor.org/info/rfc6775>>.
- [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
Labège 31670
France

Email: j.c.zuniga@ieee.org

MBONED WG
Internet-Draft
Intended status: Standards Track
Expires: August 31, 2016

Zheng. Zhang
Cui. Wang
ZTE Corporation
February 28, 2016

Multicast Service YANG
draft-zhang-mboned-multicast-service-yang-00

Abstract

This document proposes a general and all-round multicast service YANG model, which provides explanations and guidelines for the deployment of multicast service in all kinds of multicast scenarios. The multicast technologies include BIER multicast, PIM multicast, MPLS multicast and so on. And also, there defines several possible RPCs about how to interact between multicast service model and multicast device model.

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 August 31, 2016.

Copyright Notice

Copyright (c) 2016 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 (<http://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. Design of the multicast service model	2
2.1. Model Structure	3
2.2. Multicast overlay layer	4
2.3. Multicast transport layer	4
2.4. Multicast underlay layer	4
3. Notifications	4
4. Multicast service YANG model	5
5. Normative References	14
Authors' Addresses	14

1. Introduction

This document intents to provide a general and all-round multicast service model, and guides the deployment of multicast technology. Additionally, this model covers many existed multicast technologies. But this document does not define any specifical and detailed YANG models for multicast protocol, such as PIM, MLD, and BIER and so on. Instead, this document assists with them to implement multicast service.

2. Design of the multicast service model

This model includes multicast overlay, the transport layer and the possible underlay information.

Multicast overlay defines the feature of multicast flow, such as (vpnid, multicast source and multicast group) information, (ingress-node, egress-nodes) nodes information. Additionally, BIER information including (Subdomain, ingress-node BFR-id, egress-nodes BFR-id) is also included to provide BIER multicast service. In data center network, for fine-grained to gather the nodes belonging to the same virtual network, there may need VNI-related information to assist.

Multicast transport defines the transport technologies that may be used to forward multicast flow, including BIER forwarding, mpls forwarding or pim forwarding and so on.

Multicast underlay defines the possible technologies that may be used to interaction the necessary information, such as OSPF, ISIS, and BGP and so on.

2.1. Model Structure

```

module: ietf-multicast-service
  +--rw multicast-service
    +--rw multicast-overlay
      +--rw (feature-type)
        +--:(pure-multicast)
          +--rw vpn-id                uint32
          +--rw source-address        inet:ip-address
          +--rw source-wildcard?      uint8
          +--rw group-address         inet:ip-address
          +--rw group-wildcard?       uint8
        +--:(nvo3)
          +--rw vni-type              virtual-type
          +--rw vni-value              uint32
      +--rw nodes-information
        +--rw ingress-node            inet:ip-address
        +--rw egress-nodes* [number]
          +--rw number                uint32
          +--rw egress-node            inet:ip-address
      +--rw bier-information
        +--rw sub-domain              sub-domain-id
        +--rw ingress-node            bfr-id
        +--rw egress-nodes* [number]
          +--rw number                uint32
          +--rw egress-node            bfr-id
      +--rw overlay-technology
        +--rw (overlay-tech-type)
          +--:(mld)
          +--:(mvpn)
    +--rw multicast-transport
      +--rw (transport-type)
        +--:(bier)
          +--rw sub-domain              sub-domain-id
          +--rw (encap-type)
            +--:(mpls)
              +--rw bitstringlength?  uint16
              +--rw set-identifier?    si
              +--rw ecmp?              boolean
              +--rw frr?               boolean
          +--:(cisco-mode)
            +--rw p-group              inet:ip-address
            +--rw graceful-restart?    boolean
            +--rw bfd?                boolean
          +--:(mpls)
            +--rw (mpls-tunnel-type)?
              +--:(mldp)
                +--rw tunnel-id?      uint32

```

```

|         |         |   +--rw frr?                boolean
|         |         |   +--rw backup-tunnel?       boolean
|         |         |   +--:(p2mp-te)
|         |         |       +--rw tunnel-id?        uint32
|         |         |       +--rw frr?              boolean
|         |         |       +--rw backup-tunnel?     boolean
|         |         |   +--:(pim)
|         |         |       +--rw graceful-restart?  boolean
|         |         |       +--rw bfd?               boolean
+--rw multicast-underlay
|   +--rw underlay-requirement?  boolean
|   +--rw (underlay-type)
|       +--:(bgp)
|       +--:(ospf)
|       |   +--rw topology-id?        uint16
|       +--:(isis)
|       |   +--rw topology-id?        uint16
|       +--:(pim)

```

2.2. Multicast overlay layer

This layer defines the feature of multicast service, and the possible overlay protocol may be used. The feature of multicast includes the source and group information. And specific for nvo3, the feature of multicast service may be virtual network identifier.

The ingress and egress nodes information include the IP address of nodes and PEs. In BIER scenario, the nodes information may be BFR-ids.

The overlay technology, until now, MVPN and MLD are necessary.

2.3. Multicast transport layer

BIER is the transport layer technology. MPLS and PIM also are transport layer technologies. The choice of transport layer protocol can be flexible.

2.4. Multicast underlay layer

This layer has a tight connection with the underlay protocol.

3. Notifications

TBD.

4. Multicast service YANG model

```
<CODE BEGINS> file "ietf-multicast-service.yang"
module ietf-multicast-service {

    namespace "urn:ietf:params:xml:ns:yang:ietf-multicast-service";

    prefix multicast-service;

    import ietf-routing {
        prefix "rt";
    }
    import ietf-yang-types {
        prefix "yang";
    }
    import ietf-inet-types {
        prefix "inet";
    }

    organization " IETF MBONED( MBONE Deployment ) Working Group";
    contact
        "WG List:  <mailto:bier@ietf.org>
        WG Chair:  Greg Shepherd
                  <mailto:gjshep@gmail.com>
        WG Chair:  Leonard Giuliano
                  <mailto:lenny@juniper.net>

        Editor:    Zheng Zhang
                  <mailto:zhang.zheng@zte.com.cn>
        Editor:    Cui Wang
                  <mailto:wang.cuil@zte.com.cn>
        ";

    description
        "This module contains a collection of YANG definitions for
        managing multicast service.";

    revision 2016-02-29 {
        description
            "Initial version.";
        reference "https://tools.ietf.org/html/draft-zhang-mboned-mservice-yang"
    }
}
/*feature*/
    grouping general-multicast {
        description "The general multicast address information.";
        leaf source-address {
            type inet:ip-address;
            mandatory true;
        }
    }
}
```

```
        description "The address of multicast source. The value set to zero
            means that the receiver interests in all source that relevant to
            one group.";
    }
    leaf source-wildcard {
        type uint8;
        description "The wildcard information of source.";
    }
    leaf group-address {
        type inet:ip-address;
        mandatory true;
        description "The address of multicast group.";
    }
    leaf group-wildcard {
        type uint8;
        description "The wildcard information of group.";
    }
}

grouping m-addr {
    description "The vpn multicast information.";
    leaf vpn-id {
        type uint32;
        mandatory true;
        description "The vpn-id of the multicast flow.
            If there is global instance, the vpnid value should be zero.";
    }
    uses general-multicast;
}

typedef virtual-type {
    type enumeration {
        enum "vxlan" {
            description "The vxlan type.";
        }
        enum "virtual subnet" {
            description "The nvgre type";
        }
        enum "vni" {
            description "The geneve type";
        }
    }
    description "The collection of virtual network type.";
}

grouping multicast-nvo3 {
    description "The nvo3 multicast information.";
    leaf vni-type {
```

```

        type virtual-type;
        mandatory true;
        description "The type of virtual network identifier. Such as the Vxlan
an      NVGRE and Geneve.";
    }
    leaf vni-value {
        type uint32;
        mandatory true;
        description "The value of Vxlan network identifier, virtual subnet I
D      or virtual net identifier.";
    }
}

grouping multicast-feature {
    description
        "This group describes the different multicast information
        in various deployments.";
    choice feature-type {
        mandatory true;
        case pure-multicast {
            uses m-addr;
        }
        case nvo3 {
            uses multicast-nvo3;
        }
        description "The collection of all possible multicast feature.";
    }
}

grouping ip-node {
    description "The IP information of multicast nodes.";
    leaf ingress-node {
        type inet:ip-address;
        mandatory true;
        description "The ingress node of multicast flow. Or the ingress
        node of MVPN and BIER. In MVPN, this is the address of ingress
        PE; in BIER, this is the BFR-prefix of ingress node.";
    }

    list egress-node {
        key "number";
        description "This ID information of one adjacency.";
        leaf number {
            type uint32;
            mandatory true;
            description "The number of egress nodes.";
        }
        leaf egress-node {

```

```
        type inet:ip-address;
        mandatory true;
        description
            "The egress multicast node of multicast flow.
            Or the egress node of MVPN and BIER. In MVPN, this is the
            address of egress PE; in BIER, this is the BFR-prefix of
            egress node.";
    }
}
/* should import from BIER yang */
typedef bfr-id {
    type uint16;
    description "The BFR id of node.";
}

typedef si {
    type uint16;
    description
        "The type for set identifier";
}

typedef sub-domain-id {
    type uint16;
    description
        "The type for sub-domain-id";
}

typedef bit-string {
    type uint16;
    description
        "The bit mask of one bitstring.";
}

grouping bier-node {
    description "The BIER information of multicast nodes.";
    leaf sub-domain {
        type sub-domain-id;
        mandatory true;
        description "The sub-domain that this multicast flow belongs to.";
    }
    leaf ingress-node {
        type bfr-id;
        mandatory true;
        description "The ingress node of multicast flow. This is the
            BFR-id of ingress node.";
    }
    list egress-node {
```



```
    key "number";
    description "This ID information of one adjacency.";
    leaf number {
        type uint32;
        mandatory true;
        description "The number of egress nodes.";
    }
    leaf egress-node {
        type bfr-id;
        mandatory true;
        description
            "The egress multicast node of multicast flow.
             This is the BFR-id of egress node.";
    }
}

grouping overlay-tech {
    description "The possible overlay technologies for multicast service.";
    choice overlay-tech-type {
        mandatory true;
        case mld {
            description "MLD technology is used for multicast overlay";
        }
        case mvpn {
            description "MVPN technology is used for multicast overlay";
        }
    }
    description "The collection of multicast overlay technology";
}

grouping multicast-overlay {
    description "The node information that connect the ingress multicast
        flow, and the nodes information that connect the egress multicast
        flow.";
    uses multicast-feature;
    container nodes-information {
        description "The ingress and egress nodes information.";
        uses ip-node;
    }
    container bier-information {
        description "The ingress and egress BIER nodes information.";
        uses bier-node;
    }
    container overlay-technology {
        description "The possible overlay technologies for multicast service
        .";
        uses overlay-tech;
    }
}
```

```
    }

/*transport*/

typedef bier-encap-type {
    type enumeration {
        enum "mpls" {
            description "The mpls forwarding function is used in BIER.";
        }
    }
    description "The encapsulation type of BIER transportation.";
}
grouping transport-bier {
    description "The BIER transport information.";
    leaf sub-domain {
        type sub-domain-id;
        mandatory true;
        description "The subdomain id that this multicast flow belongs to.";
    }
    choice encap-type {
        mandatory true;
        case mpls {
            description "The BIER forwarding depends on mpls.";
        }
        description "The encapsulation type in BIER.";
    }
    leaf bitstringlength {
        type uint16;
        description "The bitstringlength used by BIER forwarding.";
    }
    leaf set-identifier {
        type si;
        description "The set identifier used by this multicast flow, especially in BIER TE.";
    }
    leaf ecmp {
        type boolean;
        description "The capability of ECMP.";
    }
    leaf frr {
        type boolean;
        description "The capability of fast re-route.";
    }
}

grouping transport-pim {
    description "The requirement information of pim transportation.";
    leaf graceful-restart {
```

```
        type boolean;
        description "The capability of graceful restart.";
    }
    leaf bfd {
        type boolean;
        description "The capability of bfd.";
    }
}

grouping tunnel-feature {
    description "The tunnel feature.";
    leaf tunnel-id {
        type uint32;
        description "The tunnel id that corresponds this flow.";
    }
    leaf frr {
        type boolean;
        description "The capability of fast re-route.";
    }
    leaf backup-tunnel {
        type boolean;
        description "The capability of backup tunnel.";
    }
}

grouping transport-mpls {
    description "The mpls transportation information.";
    choice mpls-tunnel-type {
        case mldp {
            uses tunnel-feature;
            description "The mldp tunnel.";
        }
        case p2mp-te {
            uses tunnel-feature;
            description "The p2mp te tunnel.";
        }
    }
    description "The collection types of mpls tunnels";
}

grouping cisco-multicast {
    description "The Cisco MDT multicast information in RFC6037.";
    leaf p-group {
        type inet:ip-address;
        mandatory true;
        description "The address of p-group.";
    }
}
```

```
    grouping transport-cisco-mode {
      description "The transport information of Cisco mode, RFC6037.";
      uses cisco-multicast;
      uses transport-pim;
    }

    grouping multicast-transport {
      description "The transport information of multicast service.";
      choice transport-type {
        mandatory true;
        case bier {
          uses transport-bier;
        }
        case cisco-mode {
          uses transport-cisco-mode;
        }
        case mpls {
          uses transport-mpls;
        }
        case pim {
          uses transport-pim;
        }
      }
      description "The collection of all possible multicast transport tech
nology.";
    }
  }

/*underlay*/
  grouping underlay-bgp {
    description "Underlay information of BGP.";
  }

  grouping underlay-ospf {
    description "Underlay information of OSPF.";
    leaf topology-id {
      type uint16;
      description "The topology id of ospf instance.";
    }
  }

  grouping underlay-isis {
    description "Underlay information of ISIS.";
    leaf topology-id {
      type uint16;
      description "The topology id of isis instance.";
    }
  }

  grouping underlay-pim {
```

```
    description "Underlay information of PIM.";
    /* If there are some necessary information should be defined? */
  }

  grouping multicast-underlay {
    description "The underlay information relevant multicast service.";
    leaf underlay-requirement {
      type boolean;
      description "Whether the underlay technology should be required.";
    }
    choice underlay-type {
      mandatory true;
      case bgp {
        uses underlay-bgp;
      }
      case ospf {
        uses underlay-ospf;
      }
      case isis {
        uses underlay-isis;
      }
      case pim {
        uses underlay-pim;
      }
      description "The collection of all possible multicast underlay technology.";
    }
  }

  container multicast-service {
    description "The model of multicast service. Include overlay, transport and underlay.";
    container multicast-overlay {
      description "The overlay information of multicast service.";
      uses multicast-overlay;
    }
    container multicast-transport {
      description "The transportation of multicast service.";
      uses multicast-transport;
    }
    container multicast-underlay {
      description "The underlay of multicast service.";
      uses multicast-underlay;
    }
  }
}
<CODE ENDS>
```

5. Normative References

[I-D.chh-bier-bier-yang]

Chen, R., hu, f., Zhang, Z., dai.xianxian@zte.com.cn, d., and M. Sivakumar, "YANG Data Model for BIER Protocol", draft-chh-bier-bier-yang-02 (work in progress), November 2015.

[I-D.ietf-bier-architecture]

Wijnands, I., Rosen, E., Dolganow, A., P, T., and S. Aldrin, "Multicast using Bit Index Explicit Replication", draft-ietf-bier-architecture-03 (work in progress), January 2016.

[I-D.ietf-pim-yang]

Liu, X., McAllister, P., and A. Peter, "A YANG data model for Protocol-Independent Multicast (PIM)", draft-ietf-pim-yang-00 (work in progress), February 2016.

[RFC6020] Bjorklund, M., Ed., "YANG - A Data Modeling Language for the Network Configuration Protocol (NETCONF)", RFC 6020, DOI 10.17487/RFC6020, October 2010, <<http://www.rfc-editor.org/info/rfc6020>>.

[RFC6037] Rosen, E., Ed., Cai, Y., Ed., and IJ. Wijnands, "Cisco Systems' Solution for Multicast in BGP/MPLS IP VPNs", RFC 6037, DOI 10.17487/RFC6037, October 2010, <<http://www.rfc-editor.org/info/rfc6037>>.

[RFC6087] Bierman, A., "Guidelines for Authors and Reviewers of YANG Data Model Documents", RFC 6087, DOI 10.17487/RFC6087, January 2011, <<http://www.rfc-editor.org/info/rfc6087>>.

[RFC6513] Rosen, E., Ed. and R. Aggarwal, Ed., "Multicast in MPLS/BGP IP VPNs", RFC 6513, DOI 10.17487/RFC6513, February 2012, <<http://www.rfc-editor.org/info/rfc6513>>.

[RFC7223] Bjorklund, M., "A YANG Data Model for Interface Management", RFC 7223, DOI 10.17487/RFC7223, May 2014, <<http://www.rfc-editor.org/info/rfc7223>>.

Authors' Addresses

Zheng(Sandy) Zhang
ZTE Corporation
No. 50 Software Ave, Yuhuatai Distinct
Nanjing
China

Email: zhang.zheng@zte.com.cn

Cui(Linda) Wang
ZTE Corporation
No. 50 Software Ave, Yuhuatai Distinct
Nanjing
China

Email: wang.cuil@zte.com.cn