

MULTIMOB Working Group  
INTERNET-DRAFT  
Intended Status: Proposed Standard  
Expires: April 18, 2013

Luis M. Contreras  
Telefonica I+D  
Carlos J. Bernardos  
Universidad Carlos III de Madrid  
Juan Carlos Zuniga  
InterDigital  
February 25, 2013

Extension of the MLD proxy functionality to support multiple  
upstream interfaces  
draft-contreras-multimob-multiple-upstreams-01

## Abstract

This document presents different scenarios of applicability for an MLD proxy running more than one upstream interface. Since those scenarios impose different requirements on the MLD proxy with multiple upstream interfaces, it is important to ensure that the proxy functionality addresses all of them for compatibility.

The purpose of this document is to define the requirements in an MLD proxy with multiple interfaces covering a variety of applicability scenarios, and to specify the proxy functionality to satisfy all of them.

## Status of this Memo

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at  
<http://www.ietf.org/lid-abstracts.html>

The list of Internet-Draft Shadow Directories can be accessed at  
<http://www.ietf.org/shadow.html>

## Copyright and License Notice

Copyright (c) 2012 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	4
2.	Terminology	4
3.	Problem statement	4
4.	Scenarios of applicability	7
4.1	Fixed network scenarios	7
4.1.1	Multicast wholesale offer for residential services	7
4.1.1.1	Requirements	7
4.1.2	Multicast resiliency	8
4.1.2.1	Requirements	8
4.1.3	Load balancing for multicast traffic in the metro segment	8
4.1.3.1	Requirements	8
4.1.4	Summary of the requirements needed for mobile network scenarios	9
4.2	Mobile network scenarios	9
4.2.1	Applicability to multicast listener mobility	10
4.2.1.1	Single MLD proxy instance on MAG	10
4.2.1.1.1	Requirements	10
4.2.1.2	Remote and local multicast subscription	10
4.2.1.2.1	Requirements	11
4.2.1.3	Dual subscription to multicast groups during handover	11
4.2.1.3.1	Requirements	12
4.2.2	Applicability to multicast source mobility	12
4.2.2.1	Support of remote and direct subscription in basic source mobility	12
4.2.2.1.1	Requirements	13
4.2.2.2	Direct communication between source and listener associated with distinct LMAs but on the same MAG	13

4.2.2.3.1	Requirements . . . . .	14
4.2.2.3	Route optimization support in source mobility for remote subscribers . . . . .	14
4.2.2.3.1	Requirements . . . . .	14
4.2.3	Summary of the requirements needed for mobile network scenarios . . . . .	15
5	Functional specification of an MLD proxy with multiple interfaces . . . . .	17
6	Security Considerations . . . . .	17
7	IANA Considerations . . . . .	17
8	Conclusions . . . . .	17
9	Acknowledgements . . . . .	17
10	References . . . . .	17
10.1	Normative References . . . . .	17
10.2	Informative References . . . . .	17
	Appendix A. Basic support for multicast listener with PMIPv6 . .	18
	Authors' Addresses . . . . .	20

## 1 Introduction

The aim of this document is to define the functionality that an MLD proxy with multiple upstream interfaces should have in order to support different scenarios of applicability in both fixed and mobile networks. This compatibility is needed in order to simplify node functionality and to ensure an easier deployment of multicast capabilities in all the use cases described in this document.

## 2. Terminology

This document uses the terminology defined in [3]. Specifically, the definition of Upstream and Downstream interfaces, which are reproduced here for completeness.

Upstream interface:

A proxy device's interface in the direction of the root of the tree. Also called the "Host interface".

Downstream interface:

Each of a proxy device's interfaces that is not in the direction of the root of the tree. Also called the "Router interfaces".

## 3. Problem statement

The concept of MLD proxy with several upstream interfaces has emerged as a way of optimizing (and in some cases enabling) service delivery scenarios where separate multicast service providers are reachable through the same access network infrastructure. Figure 1 presents the conceptual model under consideration.

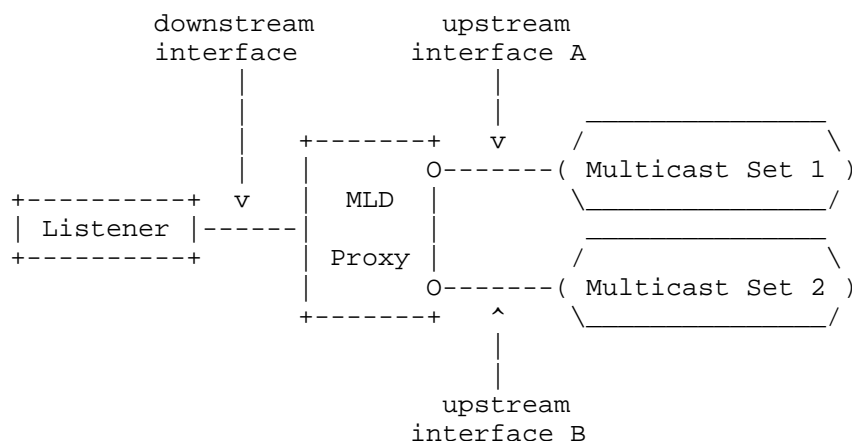


Figure 1. Concept of MLD proxy with multiple upstream interfaces

For illustrative purposes, two applications for fixed and mobile networks are here introduced. They will be elaborated later on the document.

In the case of fixed networks, multicast wholesale services in a competitive residential market require an efficient distribution of multicast traffic from different operators, i.e. the incumbent operator and a number of alternative ones, on the network infrastructure of the former. Existing proposals are based on the use of PIM routing from the metro network, and multicast traffic aggregation on the same tree. A different approach could be achieved with the use of an MLD proxy with multiple upstream interfaces, each of them pointing to a distinct multicast router in the metro border which is part of separated multicast trees deep in the network. Figure 2 graphically describes this scenario.

In the case of mobile networks, IP mobility services guarantee the continuity of the IP session while a Mobile Node (MN) changes its point of attachment. Proxy Mobile IPv6 (PMIPv6) [1] standardized a protocol that allows the network to manage the MN mobility without requiring specific support from the mobile terminal. The traffic to the MN is tunneled from the Home Network making use of two entities, one acting as mobility anchor, and the other as Mobility Access Gateway (MAG). Multicast support in PMIPv6 [2] implies the delivery of all the multicast traffic from the Home Network, via the mobility anchor. However, multicast routing optimization [4] could take advantage of an MLD proxy with multiple upstream interfaces by supporting the decision of subscribing a multicast content from the Home Network or from the local PMIPv6 domain if it is locally available. Figure 3 presents this scenario.

Informational text is provided in Appendix A summarizing how the basic solution for deploying multicast listener mobility with Proxy Mobile IPv6 works.

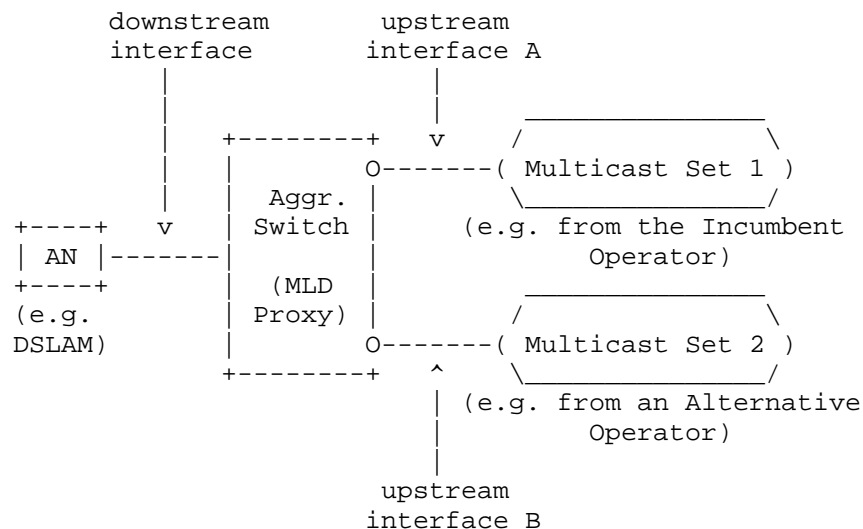


Figure 2. Example of usage of an MLD proxy with multiple upstream interfaces in a fixed network scenario

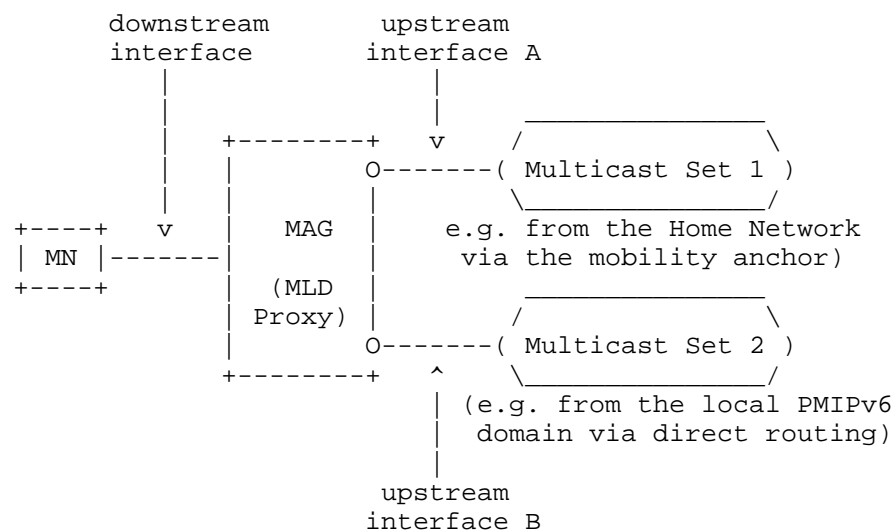


Figure 3. Example of usage of an MLD proxy with multiple upstream interfaces in a mobile network scenario

Since those scenarios can motivate distinct needs in terms of MLD proxy functionality, it is necessary to consider a comprehensive approach, looking at the possible scenarios, and establishing a minimum set of requirements which can allow the operation of a versatile MLD proxy with multiple upstream interfaces as a common entity to all of them (i.e., no different kinds of proxies depending on the scenario, but a common proxy applicable to all the potential scenarios).

#### 4. Scenarios of applicability

This section describes in detail a number of scenarios of applicability of an MLD proxy with multiple upstream interfaces in place. A number of requirements for the MLD proxy functionality are identified from those scenarios.

##### 4.1 Fixed network scenarios

Residential broadband users get access to multiple IP services through fixed network infrastructures. End user's equipment is connected to an access node, and the traffic of a number of access nodes is collected in aggregation switches.

For the multicast service, the use of an MLD proxy with multiple upstream interfaces in those switches can provide service flexibility in a lightweight and simpler manner if compared with PIM-routing based alternatives.

###### 4.1.1 Multicast wholesale offer for residential services

This scenario has been already introduced in the previous section, and can be seen in Figure 2. There are two different operators, the one operating the fixed network where the end user is connected (e.g., typically an incumbent operator), and the one providing the Internet service to the end user (e.g., an alternative Internet service provider). Both can offer multicast streams that can be subscribed by the end user, independently of which provider contributes with the content.

Note that it is assumed that both providers offer distinct multicast groups. However, more than one subscription to multicast channels of different providers could take place simultaneously.

###### 4.1.1.1 Requirements

- The MLD proxy should be able to deliver multicast control messages sent by the end user to the corresponding provider's multicast router.

- The MLD proxy should be able to deliver multicast control messages sent by each of the providers to the corresponding end user.

#### 4.1.2 Multicast resiliency

In current PIM-based solutions, the resiliency of the multicast distribution relays on the routing capabilities provided by protocols like PIM and VRRP. A simpler scheme could be achieved by implementing different upstream interfaces on MLD proxies, providing path diversity through the connection to distinct leaves of a given multicast tree.

It is assumed that only one of the upstream interfaces is active in receiving the multicast content, while the other is up and in standby for fast switching.

##### 4.1.2.1 Requirements

- The MLD proxy should be able to deliver multicast control messages sent by the end user to the corresponding active upstream interface.
- The MLD proxy should be able to deliver multicast control messages received in the active upstream to the end users, while ignoring the control messages of the standby upstream interface.
- The MLD proxy should be able of rapidly switching from the active to the standby upstream interface in case of network failure, transparently to the end user.

#### 4.1.3 Load balancing for multicast traffic in the metro segment

A single upstream interface in existing MLD proxy functionality typically forces the distribution of all the channels on the same path in the last segment of the network. Multiple upstream interfaces could naturally split the demand, alleviating the bandwidth requirements in the metro segment.

##### 4.1.3.1 Requirements

- The MLD proxy should be able to deliver multicast control messages sent by the end user to the corresponding multicast router which provides the channel of interest.
- The MLD proxy should be able to deliver multicast control messages sent by each of the multicast routers to the corresponding end user.
- The MLD proxy should be able to decide which upstream interface is selected for any new channel request according to defined criteria

(e.g., load balancing).

#### 4.1.4 Summary of the requirements needed for mobile network scenarios

Following the analysis above, a number of different requirements can be identified by the MLD proxy to support multiple upstream interfaces in fixed network scenarios. The following table summarizes these requirements.

	Fixed Network Scenarios		
Functionality	Multicast Wholesale	Multicast Resiliency	Load Balancing
Upstream Control Delivery	X	X	X
Downstr. Control Delivery	X	X	X
Active / Standby Upstream		X	
Upstr i/f selection per group			X
Upstr i/f selection all group		X	

Table I. Functionality needed on MLD proxy with multiple upstream interfaces per application scenario in fixed networks

#### 4.2 Mobile network scenarios

The mobile networks considered in this document are supposed to run PMIPv6 protocol for IP mobility management. A brief description of multicast provision in PMIPv6-based networks can be found in Appendix A.

The use of an MLD proxy supporting multiple upstream interfaces can improve the performance and the scalability of multicast-capable PMIPv6 domains.

#### 4.2.1 Applicability to multicast listener mobility

Three sub-cases can be identified for the multicast listener mobility.

##### 4.2.1.1 Single MLD proxy instance on MAG

The base solution for multicast service in PMIPv6 [2] assumes that any MN subscribed to multicast services receive the multicast traffic through the associated LMA, as in the unicast case. As standard MLD proxy functionality only supports one upstream interface, the MAG should implement several separated MLD proxy instances, one per LMA, in order to serve the multicast traffic to the MNs, according to any particular LMA-MN association.

A way of avoiding the multiplicity of MLD proxy instance in a MAG is to deploy a unique MLD proxy instance with multiple upstream interfaces, one per LMA, without any change in the multicast traffic distribution.

##### 4.2.1.1.1 Requirements

- The MLD proxy should be able of delivering the multicast control messages sent by the MNs to the associated LMA.
- The MLD proxy should be able of delivering the multicast control messages sent by each of the connected LMAs to the corresponding MN.
- The MLD proxy should be able of routing the multicast data coming from different LMAs to the corresponding MNs according to the MN to LMA association.
- The MLD proxy should be able of maintaining a 1:1 association between an MN and LMA (or downstream to upstream).

##### 4.2.1.2 Remote and local multicast subscription

This scenario has been already introduced in the previous section, and can be seen in Figure 3. Standard MLD proxy definition, with a unique upstream interface per proxy, does not allow the reception of multicast traffic from distinct upstream multicast routers. In other words, all the multicast traffic being sent to the MLD proxy in

downstream traverses a concrete, unique router before reaching the MAG. There are, however, situations where different multicast content could reach the MLD proxy through distinct next-hop routers.

For instance, the solution adopted to avoid the tunnel convergence problem in basic multicast PMIPv6 deployments [4] considers the possibility of subscription to a multicast source local to the PMIPv6 domain. In that situation, some multicast content will be accessed remotely, through the home network via the multicast tree mobility anchor, while some other multicast content will reach the proxy directly, via a local router in the domain.

#### 4.2.1.2.1 Requirements

- The MLD proxy should be able of delivering the multicast control messages sent by the MNs to the associated upstream interface based on the location of the source, remote or local, for a certain multicast group.
- The MLD proxy should be able of delivering the multicast control messages sent either local or remotely to the corresponding MNs.
- The MLD proxy should be able of routing the multicast data coming from different upstream interfaces to a certain MN according to the MN subscription, either local or remote. Note that it is assumed that a multicast group can be subscribed either locally or remotely, but not simultaneously. However more than one subscription could happen, being local or remote independently.
- The MLD proxy should be able of maintaining a 1:N association between an MN and the remote and local multicast router (or downstream to upstream).
- The MLD proxy should be able of switching between local or remote subscription for per multicast group according to specific configuration parameters (out of the scope of this document).

#### 4.2.1.3 Dual subscription to multicast groups during handover

In the event of an MN handover, once an MN moves from a previous MAG (pMAG) to a new MAG (nMAG), the nMAG needs to set up the multicast status for the incoming MN, and subscribe the multicast channels it was receiving before the handover event. The MN will then experience a certain delay until it receives again the subscribed content.

A generic solution is being defined in [5] to speed up the knowledge of the ongoing subscription by the nMAG. However, for the particular case that the underlying radio access technology supports layer-2

triggers (thus requiring extra capabilities on the mobile node), there could be inter-MAG cooperation for handover support if pMAG and nMAG are known in advance.

This could be the case, for instance for those contents not already arriving to the nMAG, where the nMAG temporally subscribes the multicast groups of the ongoing MN's subscription via the pMAG, while the multicast delivery tree among the nMAG and the mobility anchor is being established.

A similar approach is followed in [6] despite the solution proposed there differs from this approach (i.e., there is no consideration of an MLD proxy with multiple interfaces).

#### 4.2.1.3.1 Requirements

- The MLD proxy should be able of delivering the multicast control messages sent by the MNs to the associated upstream interface based on the handover specific moment, for a certain multicast group.
- The MLD proxy should be able of delivering the multicast control messages sent either from pMAG or the multicast anchor to the corresponding MNs, based on the handover specific moment.
- The MLD proxy should be able of handle the incoming packet flows from the two simultaneous upstream interfaces, in order to not duplicate traffic delivered on the point-to-point link to the MN.
- The MLD proxy should be able of maintaining a 1:N association between an MN and both the remote multicast router and the pMAG (or downstream to upstream).
- The MLD proxy should be able of switching between local or remote subscription for all the multicast groups (from pMAG to multicast anchor) according to specific configuration parameters (out of the scope of this document).

#### 4.2.2 Applicability to multicast source mobility

A couple of sub-cases can be identified for the multicast source mobility.

##### 4.2.2.1 Support of remote and direct subscription in basic source mobility

In the basic case of source mobility, the multicast source is connected to one of the downstream interfaces of an MLD proxy. According to the standard specification [3] every packet sent by the

multicast source will be forwarded towards the root of the multicast tree.

However, linked to the mobility listener problem, there could be the case of simultaneous remote subscribers, subscribing to the multicast content through the home network, and local subscribers, requesting the contents directly via a multicast router residing on the same PMIPv6 domain where the source is attached to.

Then, in order to provide the co-existence of both types of subscribers, an MLD proxy with two upstream interfaces could simultaneously serve all kind of multicast subscribers.

Basic source mobility is being defined in [7] but the solution proposed there does not allow simultaneous co-existence of remote and local subscribers (i.e., the content sent by the source is either distributed locally to a multicast router in the PMIPv6 domain, or remotely by using the bi-directional tunnel towards the mobility anchor, but not both simultaneously).

#### 4.2.2.1.1 Requirements

- The MLD proxy should be able of forwarding (replicating) the multicast content to both upstream interfaces, in case of simultaneous remote and local distribution.
- The MLD proxy should be able of handling control information incoming through any of the two upstream interfaces, providing the expected behavior for each of the multicast trees.
- The MLD proxy should be able of routing the multicast data towards different upstream interfaces for both remote and local subscriptions that could happen simultaneously.
- The MLD proxy should be able of maintaining a 1:N association between an MN and both the remote and local multicast router (or downstream to upstream).

#### 4.2.2.2 Direct communication between source and listener associated with distinct LMAs but on the same MAG

In a certain PMIPv6 domain can be MNs associated to distinct LMAs using the same MAG to get access to their corresponding home networks. For multicast communication, according to the base solution [2], each MN <-> LMA association implies a distinct MLD proxy instance to be invoked in the MAG.

In these conditions, when a mobile source is serving multicast content to a mobile listener, both attached to the same MAG but each of them associated to different LMAs, the multicast flow must traverse the PMIPv6 domain from the MAG to the LMA where the source maintains an association, then from that LMA to the LMA where the listener is associated to, and finally come back to the same MAG from where the flow departed. This routing is extremely inefficient.

An MLD proxy with multiple upstream interfaces avoids this behavior since it allows to invoke a unique MLD proxy instance in the MAG. In this case, the multicast source can directly communicate with the multicast listener, without need for delivering the multicast traffic to the LMAs.

#### 4.2.2.3.1 Requirements

- The MLD proxy should be able of forwarding (replicating) the multicast content to different upstream or downstream interfaces where subscribers are present.
- The MLD proxy should be able of handling control information incoming through any of the upstream or downstream interfaces requesting a multicast flow being injected in another downstream interface.
- The MLD proxy should be able of maintaining a 1:N association between an MN and any of the upstream or downstream interfaces demanding the multicast content.

#### 4.2.2.3 Route optimization support in source mobility for remote subscribers

Even in a scenario of remote subscription, there could be the case where both the source and the listener are attached to the same PMIPv6-Domain (for instance, no possibility of direct routing within the PMIPv6, or source and listener pertaining to distinct home networks). In this situation there is a possibility of route optimization if inter-MAG communication is enabled, in such a way that the listeners in the PMIPv6 domain are served through the tunnels between MAGs, while the rest of remote listeners are served through the mobility anchor.

A multi-upstream MLD proxy would allow the simultaneous delivery of traffic to such kind of remote listeners.

A similar route optimization approach is proposed in [8].

#### 4.2.2.3.1 Requirements

- The MLD proxy should be able of forwarding (replicating) the multicast content to both kinds of upstream interfaces, inter-MAG tunnel interfaces and MAG to mobility anchor tunnel interface.
- The MLD proxy should be able of handling control information incoming through any of the two types of upstream interfaces, providing the expected behavior for each of the multicast trees (e.g., no forwarding traffic on one inter-MAG link once there are not more listeners requesting the content).
- The MLD proxy should be able of routing the multicast data towards different upstream interfaces for both remote and route optimized subscriptions that could happen simultaneously.
- The MLD proxy should be able of maintaining a 1:N association between an MN and both the remote and local MAGs (or downstream to upstream).

#### 4.2.3 Summary of the requirements needed for mobile network scenarios

After the previous analysis, a number of different requirements can be identified by the MLD proxy to support multiple upstream interfaces in mobile network scenarios. The following table summarizes these requirements.

Functionality	Mobile Network Scenarios					
	Multicast Listener			Multicast Source		
	Single MLD Proxy	Remote & local subscr.	Dual subscr. in HO	Direct & remote subscr.	Listener & source on MAG	Route optimi.
Upstream Control Delivery	X	X	X	X	X	X
Downstr. Control Delivery	X	X	X		X	
Upstream Data Delivery				X		X
Downstr. Data Delivery	X	X	X		X	
1:1 MN to upstream assoc.	X					
1:N MN to upstream assoc.		X	X	X	X	X
Upstr i/f selection per group		X				
Upstr i/f selection all group			X			
Upstream traffic replicat.				X		X

Table II. Functionality needed on MLD proxy with multiple upstream interfaces per application scenario in mobile networks

## 5 Functional specification of an MLD proxy with multiple interfaces

<To be completed>.

## 6 Security Considerations

<To be completed>.

## 7 IANA Considerations

<IANA considerations text>.

## 8 Conclusions

<To be completed>.

## 9 Acknowledgements

The authors thank Stig Venaas for his valuable comments and suggestions.

The research of Carlos J. Bernardos leading to these results has received funding from the European Community's Seventh Framework Programme (FP7-ICT-2009-5) under grant agreement n. 258053 (MEDIEVAL project), being also partially supported by the Ministry of Science and Innovation (MICINN) of Spain under the QUARTET project (TIN2009-13992-C02-01).

## 10 References

## 10.1 Normative References

- [1] Gundavelli, S., Leung, K., Devarapalli, V., Chowdhury, K., and B. Patil, "Proxy Mobile IPv6", RFC 5213, August 2008.
- [2] T.C. Schmidt, M. Waehlich, and S. Krishnan, "A Minimal Deployment Option for Multicast Listeners in PMIPv6 Domains", RFC6224, April 2011.
- [3] B. Fenner, H. He, B. Haberman, and H. Sandick, "Internet Group Management Protocol (IGMP) / Multicast Listener Discovery (MLD)-Based Multicast Forwarding ("IGMP/MLD Proxying")", RFC 4605, August 2006.

## 10.2 Informative References

- [4] J.C. Zuniga, L.M. Contreras, C.J. Bernardos, S. Jeon, Y. Kim, "Multicast Mobility Routing Optimizations for Proxy Mobile IPv6", work in progress, draft-ietf-multimob-pmipv6-ropt-01, September 2012.
- [5] L.M. Contreras, C.J. Bernardos, I. Soto, "PMIPv6 multicast handover optimization by the Subscription Information Acquisition through the LMA (SIAL)", work in progress, draft-ietf-multimob-fast-handover-01, July 2012.
- [6] T.C. Schmidt, M. Waehlich, R. Koodli, G. Fairhurst, "Multicast Listener Extensions for MIPv6 and PMIPv6 Fast Handovers", work in progress, draft-schmidt-multimob-fmipv6-pfmipv6-multicast-06, May 2012
- [7] T.C. Schmidt, S. Gao, H. Zhang, M. Waehlich, "Mobile Multicast Sender Support in Proxy Mobile IPv6 (PMIPv6) Domains", work in progress, draft-ietf-multimob-pmipv6-source-01, July 2012.
- [8] J. Liu, W. Luo, "Routes Optimization for Multicast Sender in Proxy Mobile IPv6 Domain", work in progress, draft-liu-multimob-pmipv6-multicast-ro-02, July 2012.

#### Appendix A. Basic support for multicast listener with PMIPv6

This section briefly summarizes the operation of Proxy Mobile IPv6 [1] and how multicast listener support works with PMIPv6 as specified in [2].

Proxy Mobile IPv6 (PMIPv6) [1] is a network-based mobility management protocol which enables the network to provide mobility support to standard IP terminals residing in the network. These terminals enjoy this mobility service without being required to implement any mobility-specific IP operations. Namely, PMIPv6 is one of the mechanisms adopted by the 3GPP to support the mobility management of non-3GPP terminals in future Evolved Packet System (EPS) networks.

PMIPv6 allows a Media Access Gateway (MAG) to establish a distinct bi-directional tunnel with different Local Mobility Anchors (LMAs), being each tunnel shared by the attached Mobile Nodes (MNs). Each mobile node is associated with a corresponding LMA, which keeps track of its current location, that is, the MAG where the mobile node is attached. IP-in-IP encapsulation is used within the tunnel to forward traffic between the LMA and the MAG. Figure 4 (taken from [1]) shows the architecture of a PMIPv6 domain.

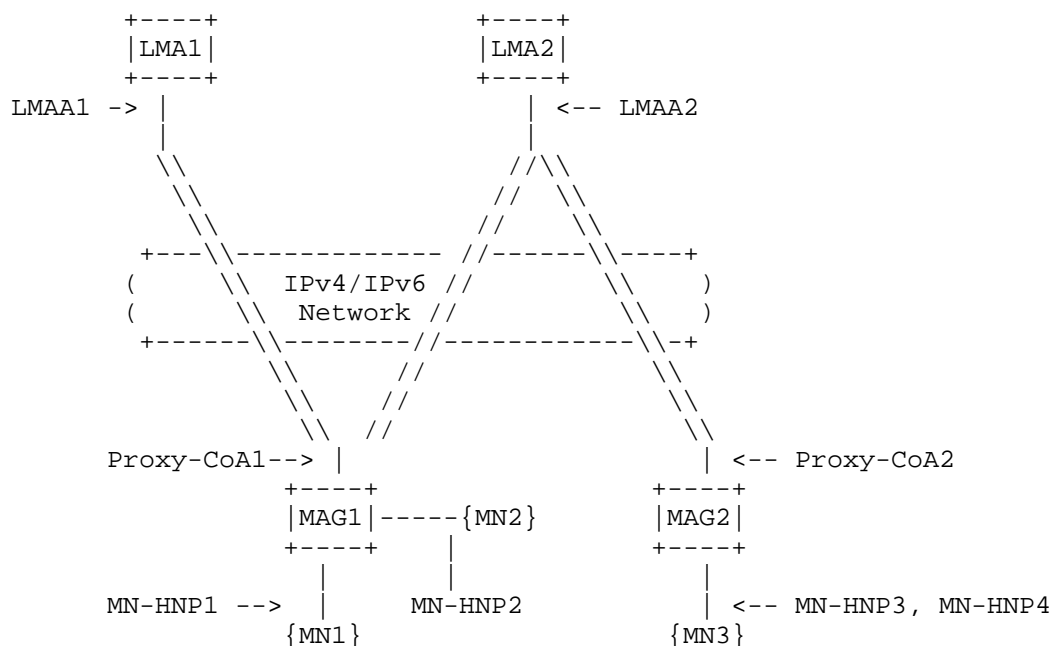


Figure 4. Proxy Mobile IPv6 Domain

The basic solution for the distribution of multicast traffic within a PMIPv6 domain [2] makes use of the bi-directional LMA-MAG tunnels. The base solution follows the so-called remote subscription model, in which the subscribed multicast content is delivered from the Home Network. By doing so, an individual copy of every multicast flow is delivered through the tunnel connecting the mobility anchor to any of the access gateways in the domain. In many cases, these individual copies traverse the same routers in the path towards the access gateways, incurring in an inefficient distribution, equivalent to the unicast distribution of the multicast content in the domain.

The reference scenario for multicast deployment in Proxy Mobile IPv6 domains is illustrated in Figure 5 (taken from [2]).

This fact leads to distribution inefficiencies and higher per-bit delivery costs, incurred by the PMIPv6 domain operator offering transport capabilities to the Home Network operator for serving their MNs when attached to the PMIPv6 domain. As long as the remotely subscribed multicast service is not affected, it seems worthy to explore more optimal ways of distributing such content within the PIMIPv6 domain.

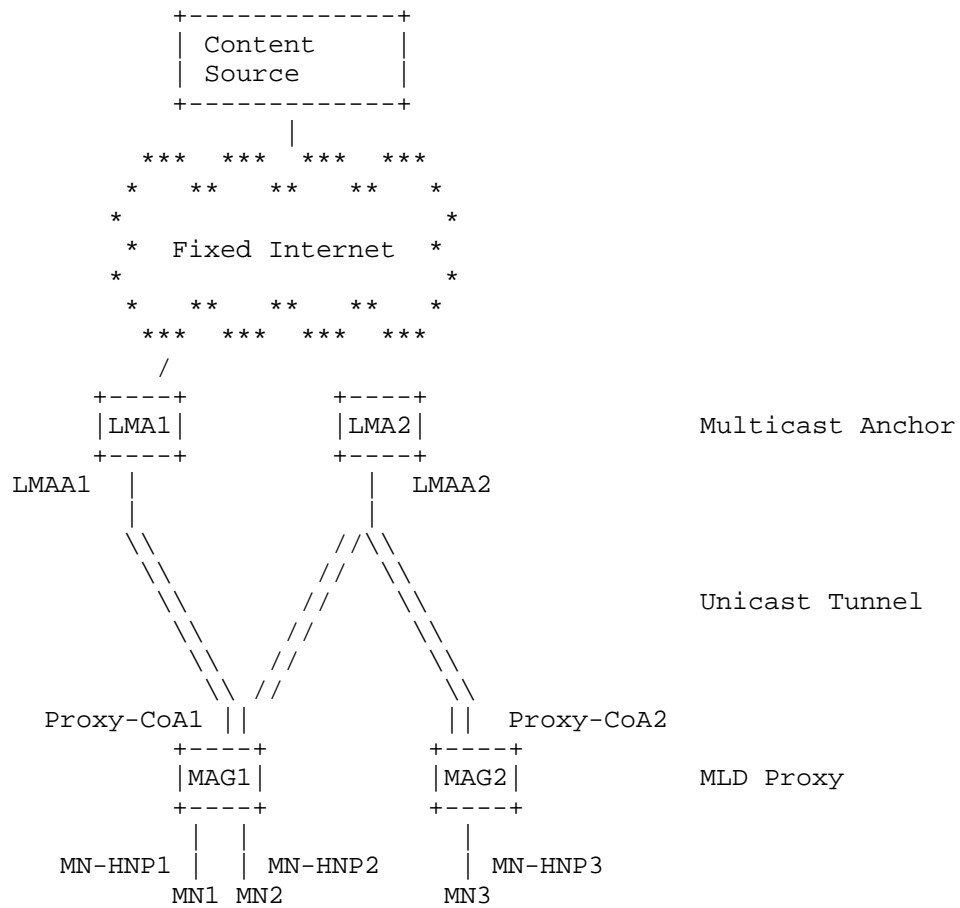


Figure 5. Reference Network for Multicast Deployment in PMIPv6

## Authors' Addresses

Luis M. Contreras  
 Telefonica I+D  
 EMail: lmcm@tid.es

Carlos J. Bernardos  
 Universidad Carlos III de Madrid  
 EMail: cjb@it.uc3m.es

INTERNET DRAFT

MLD proxy with multiple upstream

February 25, 2012

Juan Carlos Zuniga  
InterDigital Communications, LLC  
EMail: JuanCarlos.Zuniga@InterDigital.com

Network Working Group  
Internet-Draft  
Intended status: BCP  
Expires: April 21, 2014

J. Abley  
Dyn, Inc.  
October 18, 2013

DNS Reverse Mapping for Multicast Addresses  
draft-jabley-multicast-ptr-00

Abstract

The mapping of IPv4 and IPv6 addresses to names using the Domain Name System (DNS) is colloquially known as "Reverse Mapping". Reverse Mapping support for registered multicast address assignments in IPv4 is currently incomplete and ad-hoc; in IPv6 there is no support at all.

This document describes procedures to be followed that will result in more systematic and predictable support for Reverse Mapping for IPv4 multicast address assignments, and introduces analogous support for IPv6.

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 April 21, 2014.

Copyright Notice

Copyright (c) 2013 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. General Approach . . . . .	4
3. Naming Scheme . . . . .	5
3.1. IPv4 Multicast Addresses . . . . .	5
3.2. IPv6 Multicast Addresses . . . . .	5
4. Use of MCAST.ARPA . . . . .	6
5. IAB Considerations . . . . .	7
6. IANA Considerations . . . . .	8
6.1. Registry Changes . . . . .	8
6.1.1. IPv6 Multicast Scope Registry . . . . .	8
6.1.2. IPv4 Multicast Address Space Registries . . . . .	8
6.1.3. IPv6 Multicast Address Space Registries . . . . .	9
6.2. Delegation of MCAST.ARPA and MCAST6.ARPA . . . . .	9
6.3. Initial Zone Contents . . . . .	9
6.4. Process Changes . . . . .	11
6.5. Ongoing Support for MCAST.NET . . . . .	11
7. Security Considerations . . . . .	13
8. Acknowledgements . . . . .	14
9. References . . . . .	15
9.1. Normative References . . . . .	15
9.2. Informative References . . . . .	15
Appendix A. Editorial Notes . . . . .	16
A.1. Change History . . . . .	16
Author's Address . . . . .	17

## 1. Introduction

The Domain Name System (DNS), as originally specified in [RFC1034] and [RFC1035], provides support for the mapping of IPv4 addresses to names using a namespace convention within the IN-ADDR.ARPR domain and the PTR resource record type.

The analogous mapping of IPv6 address to names is specified in [RFC3596], adopting a similar namespace convention within the IP6.ARPA domain.

Multicast addresses are assigned by the IANA, and assignments are documented in various IANA registries.

For IPv4, Reverse Mapping of assigned multicast addresses to names has historically been provided in an ad-hoc and incomplete fashion, without tight coordination with IANA multicast address assignment processes. Names assigned to IPv4 multicast addresses have been chosen somewhat arbitrarily within the MCAST.NET domain. For IPv6, no Reverse Mapping is provided.

This document describes procedures to be followed by the IANA to support predictable and consistent Reverse Mapping for registered multicast addresses in IPv4 and IPv6.

## 2. General Approach

This document specifies extensions to existing IPv4 and IPv6 multicast registries to include a mandatory column "DNS Label". This field is required to be populated with a unique, valid DNS label for all future multicast address assignments except in the case where reverse mapping for an address is explicitly not desirable.

The procedures at the IANA relating to multicast address assignment are extended to include the provisioning of appropriate changes in the DNS at the time of registration or de-registration of any multicast addresses. Specific actions requested of the IANA are described in Section 6.

Names for multicast addresses are assigned under MCAST.ARPA for IPv4 addresses, and MCAST6.ARPA for IPv6 addresses. The naming schemes to be used in each case are described in Section 3. The use of MCAST.ARPA rather than MCAST.NET is discussed in Section 4.

### 3. Naming Scheme

#### 3.1. IPv4 Multicast Addresses

Each assigned IPv4 multicast address has an accompanying DNS Label. The name associated with an IPv4 multicast address with DNS Label SOMENAME is SOMENAME.MCAST.ARPA.

For example, suppose the assigned IPv4 multicast address 224.0.1.1 has the DNS Label "NTP". The address 224.0.1.1 maps to the name "NTP.MCAST.ARPA"; the name "NTP.MCAST.ARPA" maps to the address 224.0.1.1.

#### 3.2. IPv6 Multicast Addresses

Each assigned IPv6 multicast address has an accompanying DNS Label ("Address Label").

IPv6 multicast addresses may be of fixed or variable scope. The naming scheme for these addresses incorporates a scope identifier using an additional DNS label ("Scope Label"), specified in a dedicated registry (see Section 6.1.1). Both fixed and variable scope multicast addresses use the same naming scheme.

The name associated with an IPv6 multicast address with Scope Label SCOPE and Address Label SOMENAME is SOMENAME.SCOPE.MCAST6.ARPA.

For example, suppose ff01::1 is an assigned IPv6 multicast address with Scope Label "NODE-LOCAL" and Address Label "ALL-NODES". The address ff01::1 maps to the name "ALL-NODES.NODES-LOCAL.MCAST6.ARPA"; the name "ALL-NODES.NODES-LOCAL.MCAST6.ARPA" maps to the address ff01::1.

The variable-scope multicast address ff0x::fb will have different Reverse Mapping depending on the scope specified in the address (i.e. the value of x), although the Address Label in each case will be the same ("MDNSV6"). The Site-Local address ff05::fb has an associated Scope Label "SITE-LOCAL", and is therefore named MDNSV6.SITE-LOCAL.MCAST6.ARPA. The Link-Local address ff02::fb has an associated Scope Label "LINK-LOCAL" and hence is named MDNSV6.LINK-LOCAL.MCAST6.ARPA.

#### 4. Use of MCAST.ARPA

Use of the MCAST.ARPA domain rather than MCAST.NET for IPv4 multicast addresses is specified for the same reasons that led IP6.INT to be superceded by "IP6.ARPA" [RFC3152].

It is prudent to assume that hard-coded assumptions about names in MCAST.NET exist, and will persist for some time. This document specifies that names in the MCAST.ARPA domain also be available in the MCAST.NET domain, to provide support for software with those assumptions. Ongoing support for the MCAST.NET zone is described in Section 6.5.

It is possible that in the future empirical measurement will confirm that the use of names under MCAST.NET is no longer required and that provisioning of the MCAST.NET domain can safely cease. This document provides no such measurement and makes no such recommendation, however.

## 5. IAB Considerations

This document proposes a delegation within the ARPA domain, and, in accordance with [RFC3172], IAB review and approval of the delegation of MCAST.ARPA and MCAST6.ARPA as described in Section 6.2 is required.

Once IAB approval has been obtained, this section may be removed prior to publication or updated to include text that confirms the IAB's decision, at the IAB's discretion.

## 6. IANA Considerations

### 6.1. Registry Changes

#### 6.1.1. IPv6 Multicast Scope Registry

The IANA is directed to create a new registry as follows:

Registry Name: IPv6 Multicast Address Scopes

Registration Procedure: Standards Action

Reference: This document

Schema: See initial contents, below. Note that the Scope Value and DNS Label fields are mandatory for all rows, and that values chosen for future DNS Label fields are required to be unique within this registry.

The initial contents of this new registry should be:

Scope Value	DNS Label	Scope Name	Reference
0x0	(none)	Reserved	[RFC4291]
0x1	NODE-LOCAL	Node-Local Scope	[RFC4291]
0x2	LINK-LOCAL	Link-Local Scope	[RFC4291]
0x3	(none)	Reserved	[RFC4291]
0x4	ADMIN-LOCAL	Admin-Local Scope	[RFC4291]
0x5	SITE-LOCAL	Site-Local Scope	[RFC4291]
0x8	ORG-LOCAL	Organisation-Local Scope	[RFC4291]
0xE	GLOBAL	Global Scope	[RFC4291]

The IANA may add "Date Registered" and "Last Revised" columns to the schema at its discretion.

#### 6.1.2. IPv4 Multicast Address Space Registries

The IANA is directed to add a mandatory "DNS Label" column to all IPv4 Multicast Address Space registries. The initial contents of the

DNS Label field for each row should be taken from the corresponding MCAST.NET zone owner names where available; addresses with no existing mapping in MCAST.NET should have DNS Labels assigned by the IANA at their discretion.

All existing assignments should have a DNS Label assigned. A DNS Label should be mandatory for all future registrations. DNS Labels are required to be unique for all IPv4 multicast address assignments.

#### 6.1.3. IPv6 Multicast Address Space Registries

The IANA is directed to add a mandatory "DNS Label" column to all IPv6 Multicast Address Space registries. The initial contents of the DNS Label field for each row should be assigned by the IANA at their discretion.

All existing assignments should have a DNS Label assigned. A DNS Label should be mandatory for all future registrations. DNS Labels are required to be unique for all IPv6 multicast address assignments.

#### 6.2. Delegation of MCAST.ARPA and MCAST6.ARPA

The IANA is directed to create and host the MCAST.ARPA and MCAST6.ARPA zones on name servers of their choosing. The MCAST.ARPA and MCAST6.ARPA zones should be signed using DNSSEC, with DNSSEC parameters chosen by the IANA. The initial zone contents should be as described in Section 6.3.

The IANA is directed to provision secure delegations for the MCAST.ARPA and MCAST6.ARPA zones from the ARPA zone (i.e. delegations with accompanying DS RRsets).

#### 6.3. Initial Zone Contents

The IANA is directed to populate the MCAST.ARPA and MCAST6.ARPA zones, and the corresponding reverse mapping zones under IN-ADDR.ARPA and IP6.ARPA, directly from the IPv4 and IPv6 Multicast Address Registries, amended as described in Section 6.1.

As an example, if the IPv6 Variable Scope Multicast Addresses sub-registry contained the following entry:

Address(es)	Description	DNS Label
FF0X::FB	mDNSv6	MDNSV6



```
$ORIGIN MCAST.ARPA.  
;  
; SGI-Dogfight address  
;  
SGI-DOG                A          224.0.1.2  
  
$ORIGIN 224.IN-ADDR.ARPA.  
;  
; SGI-Dogfight address  
;  
2.1.0                  PTR        SGI-DOG.MCAST.ARPA.
```

#### 6.4. Process Changes

The IANA is directed to require a valid and unique DNS Label to be specified within the existing processes of multicast address assignment in IPv4 and IPv6.

The IANA is further directed to maintain the MCAST.ARPA, MCAST6.ARPA and related domains under IP6.ARPA and IN-ADDR.ARPA such that any additions, changes or deletions from the corresponding address registries are reflected accurately in the DNS.

#### 6.5. Ongoing Support for MCAST.NET

IANA is directed to remove all non-apex resource records from the MCAST.NET zone and to add an apex DNAME [RFC6672] with target MCAST.ARPA. The intention is to provide backwards compatibility for software that has hard-coded assumptions about naming conventions for IPv4 multicast addresses.

For example, the following describes the result of this change for MCAST.NET SOA serial 2012123836, with DNSSEC resource records omitted for clarity:

\$ORIGIN MCAST.NET.

; beginning of zone

```
@      SOA      SNS.DNS.ICANN.ORG. NOC.DNS.ICANN.ORG. (
                                2012123836
                                7200
                                3600
                                604800
                                3600 )
```

```
NS      A.IANA-SERVERS.NET.
NS      B.IANA-SERVERS.NET.
NS      C.IANA-SERVERS.NET.
NS      NS.ICANN.ORG.
```

```
DNAME   MCAST.ARPA.
```

; end of zone

## 7. Security Considerations

This document presents no known additional security concerns to the Internet.

## 8. Acknowledgements

Your name here, etc.

## 9. References

### 9.1. Normative References

- [RFC1034] Mockapetris, P., "Domain names - concepts and facilities", STD 13, RFC 1034, November 1987.
- [RFC1035] Mockapetris, P., "Domain names - implementation and specification", STD 13, RFC 1035, November 1987.
- [RFC3596] Thomson, S., Huitema, C., Ksinant, V., and M. Souissi, "DNS Extensions to Support IP Version 6", RFC 3596, October 2003.
- [RFC4291] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", RFC 4291, February 2006.

### 9.2. Informative References

- [RFC3152] Bush, R., "Delegation of IP6.ARPA", BCP 49, RFC 3152, August 2001.
- [RFC3172] Huston, G., "Management Guidelines & Operational Requirements for the Address and Routing Parameter Area Domain ("arpa")", BCP 52, RFC 3172, September 2001.
- [RFC6672] Rose, S. and W. Wijngaards, "DNAME Redirection in the DNS", RFC 6672, June 2012.

## Appendix A. Editorial Notes

This section (and sub-sections) to be removed prior to publication.

### A.1. Change History

00 Initial draft, circulated for the purposes of entertainment.

Author's Address

Joe Abley  
Dyn, Inc.  
470 Moore Street  
London, ON N6C 2C2  
Canada

Phone: +1 519 670 9327  
Email: jabley@dyn.com



Internet Engineering Task Force  
Internet-Draft  
Intended status: Standards Track  
Expires: August 9, 2014

H. Jeng  
AT&T  
J. Haas  
Y. Rekhter  
J. Zhang  
Juniper Networks  
February 5, 2014

Multicast Geo-Distribution Control  
draft-rekhter-geo-distribution-control-04

Abstract

Consider a content provider that wants to deliver a particular content to a set of customers/subscribers, where the provider and the subscribers are connected by an IP service provider. This document covers two areas needed to accomplish this:

1. Providing the content provider with the information of whether it can use the multicast connectivity service provided by the IP service provider to deliver a particular content to a particular set of subscribers, and
2. Providing the content provider with a mechanism to restrict delivery of a given content to a particular set of the subscribers.

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 9, 2014.

Copyright Notice

Copyright (c) 2014 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. Specification of Requirements . . . . .	3
1.1. Introduction . . . . .	3
1.2. Multicast Content Distribution Zones . . . . .	4
1.3. A Brief Overview of Multicast Distribution Reachability Signaling . . . . .	4
1.4. A Brief Overview of Multicast Distribution Control Signaling . . . . .	5
1.4.1. An example of configuration on ERs . . . . .	5
2. Overview of Operations . . . . .	6
3. IANA Considerations . . . . .	7
4. Security Considerations . . . . .	7
5. Acknowledgements . . . . .	7
6. References . . . . .	7
6.1. Normative References . . . . .	7
6.2. Informative References . . . . .	8
Authors' Addresses . . . . .	8

## 1. Specification of Requirements

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

### 1.1. Introduction

Consider a content provider that wants to deliver a particular content to a set of customers/subscribers, where the provider and the subscribers are connected by an IP service provider. This document covers two areas needed to accomplish this:

1. Providing the content provider with the information of whether it can use the multicast connectivity service provided by the IP service provider to deliver a particular content to a particular set of subscribers, and
2. Providing the content provider with a mechanism to restrict delivery of a given content to a particular set of the subscribers.

For the purpose of this document we assume that a content provider consists of one or more Content Servers, and one or more Content Distribution Controllers. While this document assumes communication between Content Servers and Content Distribution Controllers, the procedures for implementing such communication is outside the scope of this document.

Content Servers are connected to one or more IP service providers (ISP) that can offer both multicast and unicast connectivity service to the subscribers of the content provider. The content provider uses this ISP(s) to deliver content to its subscribers.

Subscribers are connected to the Edge Routers (ERs) of the ISP. Note that the multicast connectivity service provided by the ISP extends all the way to the ERs. Such service could be provided by either deploying IP multicast natively, or with some tunneling mechanism like AMT, or by a combination of both within the ISP. However, between the ERs and the subscribers there may, or may not be multicast connectivity.

In the case where a particular subscriber of a given content provider does not have multicast connectivity to its ER, the content provider would use IP unicast service provided by the ISP to transmit the particular content to that subscriber.

A subscriber may want to access a particular content that is not

available to that subscriber due to policy reasons. When that subscriber would have received that content via unicast connectivity, the Content Distribution Controller, or the Content Servers, or both may enforce the policy to not deliver the content. However, when the content would be delivered via multicast connectivity it may be possible for the subscriber to receive the content by illicitly participating in the multicast signaling for that content.

To prevent a subversion of the intent of this content delivery policy, a mechanism is provided to make this policy available to devices participating in multicast signaling.

## 1.2. Multicast Content Distribution Zones

For each item of content provided by a content provider, the content provider maintains a list of subscribers who are either excluded or allowed to receive the content. For the purpose of maintaining this list this document assumes that subscribers are grouped into "zones" based on IP addresses, so that exclusion/inclusion uniformly applies to all the subscribers within a given zone. Procedures by which subscribers are grouped into zones are outside the scope of this document. However, this document assumes that this grouping is done consistently by both the content provider and the ISP(s) that the content provider uses for delivering its content.

One example of an implementation of such a zone is based on the geographic location of the subscribers. Such zones may be used to implement broadcast "blackout" of some content such as a sporting event that may not be allowed to air in certain regions due to regulatory reasons.

## 1.3. A Brief Overview of Multicast Distribution Reachability Signaling

Content providers and Content Distribution Controllers need to know the transport mechanism that a subscriber can use to receive some content. Since not all subscribers may be capable of receiving content via IP multicast, Multicast Distribution Reachability Signaling [MDRS] is used to permit the subscriber's ISP to provide this information.

MDRS permits advertises a BGP prefix with a new SAFI, MCAST-REACH, to indicate subscribers in the accompanying AFI of the MCAST-REACH SAFI can receive multicast traffic.

Please see the MDRS document for more details.

#### 1.4. A Brief Overview of Multicast Distribution Control Signaling

A content provider or a service provider may need to enforce policies to exclude access to an item of content that is delivered by IP multicast. Multicast Distribution Control Signaling [MDCS] permits this such enforcement to be distributed as BGP flowspec filters with a new SAFI, MCAST-FLOWSPEC. These filters are used by the multicast control plane to determine whether access to multicast content may be made available to downstream routers, including ERs.

These flowspec filters are distributed in BGP with Route Targets [RFC4360] that identify the include or exclude policy for a zone. Multicast routers receiving these filters maintain an ordered list of these Route Target policies to install the filters. The multicast control plane then makes use of the filter database to implement the desired policy.

##### 1.4.1. An example of configuration on ERs

Consider an ER in Manhattan that has a port that is provisioned with the following import RTs:

```
<include-manhattan, exclude-manhattan, include-nyc, exclude-
nyc, include-east, exclude-east, include-usa, exclude-usa>
```

When the ER receives a Flow Spec route with <exclude-nyc, include-manhattan, include-usa> RTs, the ER first try to match "include-manhattan" or "exclude-manhattan" (the first ones on the list) - and the result is "include-manhattan". Therefore, the (S, G) carried in the Flow Spec route is allowed on that port of the ER.

Consider another ER in Boston that has a port that is provisioned with the following import RTs:

```
<include-cambridge, exclude-cambridge, include-bos, exclude-
bos, include-east, exclude-east, include-usa, exclude-usa>
```

The above mentioned Flow Spec route will be imported (due to the include-usa RT), and will result in the (S, G) carried in the flow Spec route to be allowed on that port of the ER.

Now consider a different Flow Spec route with the <exclude-usa, include-bos, include-nyc, exclude-manhattan> RTs. The (S, G) carried in the route will be disallowed in Manhattan, allowed in Boston, and allowed in Queens (as the route will match the "include-nyc" RT).

## 2. Overview of Operations

An ISP, using the procedures described in Multicast Distribution Reachability Signaling [MDRS], provides a content provider, and specifically Content Distribution Controller(s) of that content provider, with the information of whether a particular subscriber of that content provider has multicast connectivity to an ER of that ISP with the information of whether a particular group of subscribers can receive multicast content.

To enforce the exclusion/inclusion policies, the content provider uses procedures described in Multicast Distribution Control Signaling [MDCS].

For each content provided by a content provider, the content provider selects a particular multicast channel (S, G) for distributing this content using multicast connectivity service. Procedures by which the content provider selects a particular multicast channel, and maintains the mapping are outside the scope of this document.

Subscribers are connected to the Edge Routers (ERs) of the ISP. Note that when multicast connectivity service provided is by the ISP, that service extends all the way to the ERs. Such service could be provided by either deploying IP multicast natively, or with some tunneling mechanism like AMT, or a combination of both within the ISP. However, between the ERs and the subscribers there may, or may not be multicast connectivity.

When a subscriber wants to receive the particular content from its content provider, the subscriber issues a request for this content to the Content Distribution Controller of the provider. When the Content Distribution Controller receives the request, the Content Distribution Controller uses the information carried in the request (e.g., IP address of the subscriber) to determine the zone of the subscriber, and based on that zone to determine whether the subscriber can receive this content.

If the Content Distribution Controller determines that the subscriber can receive the content, then based on the information provided by the multicast distribution reachability signaling the Content Distribution Controller determines whether the subscriber can receive this content using multicast connectivity service, and if yes, then returns to the subscriber the multicast channel selected for distributing the content.

If the Content Distribution Controller determines that the subscriber can receive the content, but can not receive the content using multicast connectivity service, the Content Distribution Controller

returns to the subscriber the information needed to receive this content using unicast connectivity service.

If the content would have been delivered to the subscriber via multicast connectivity, but the Content Distribution Controller had determined the subscriber was not permitted access to this content, then this policy may need to be enforced by the Edge Routers or upstream multicast routers to prevent illicit access of this content. This policy is enforced by utilizing filtering information distributed using Multicast Distribution Control Signaling [MDCS].

Specification of the procedures for communication between subscribers and Content Distribution Controllers are outside the scope of this document.

### 3. IANA Considerations

This document introduces no IANA Considerations.

### 4. Security Considerations

TBD

### 5. Acknowledgements

The authors would like to thank Han Nguyen for his contributions to this document.

### 6. References

#### 6.1. Normative References

- [MDCS] Jeng, H., Haas, J., Rekhter, Y., and J. Zhang, "Multicast Distribution Control Signaling", draft-ietf-idr-mdcs-00.txt (work in progress), 2014.
- [MDRS] Jeng, H., Haas, J., Rekhter, Y., and J. Zhang, "Multicast Distribution Reachability Signaling", draft-ietf-idr-mdrs-00.txt (work in progress), 2014.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

## 6.2. Informative References

[RFC4360] Sangli, S., Tappan, D., and Y. Rekhter, "BGP Extended Communities Attribute", RFC 4360, February 2006.

### Authors' Addresses

Huajin Jeng  
AT&T

Email: [hj2387@att.com](mailto:hj2387@att.com)

Jeffrey Haas  
Juniper Networks  
1194 N. Mathida Ave.  
Sunnyvale, CA 94089  
US

Email: [jhaas@juniper.net](mailto:jhaas@juniper.net)

Yakov Rekhter  
Juniper Networks  
1194 N. Mathida Ave.  
Sunnyvale, CA 94089  
US

Email: [yakov@juniper.net](mailto:yakov@juniper.net)

Jeffrey (Zhaohui) Zhang  
Juniper Networks  
1194 N. Mathida Ave.  
Sunnyvale, CA 94089  
US

Email: [zzhang@juniper.net](mailto:zzhang@juniper.net)



MBONED Working Group  
Internet Draft  
Intended status: BCP  
Expires: April 27, 2015

Percy S. Tarapore  
Robert Sayko  
AT&T  
Greg Shepherd  
Toerless Eckert  
Cisco  
Ram Krishnan  
Brocade  
October 27, 2014

Multicasting Applications Across Inter-Domain Peering Points  
draft-tarapore-mboned-multicast-cdni-07.txt

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 April 27, 2015.

Copyright Notice

Copyright (c) 2014 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.

This document may contain material from IETF Documents or IETF Contributions published or made publicly available before November 10, 2008. The person(s) controlling the copyright in some of this material may not have granted the IETF Trust the right to allow modifications of such material outside the IETF Standards Process. Without obtaining an adequate license from the person(s) controlling the copyright in such materials, this document may not be modified outside the IETF Standards Process, and derivative works of it may not be created outside the IETF Standards Process, except to format it for publication as an RFC or to translate it into languages other than English.

## Abstract

This document examines the process of transporting applications via multicast across inter-domain peering points. The objective is to describe the setup process for multicast-based delivery across administrative domains and document supporting functionality to enable this process.

## Table of Contents

1. Introduction.....	3
2. Overview of Inter-domain Multicast Application Transport.....	4
3. Inter-domain Peering Point Requirements for Multicast.....	5
3.1. Native Multicast.....	5
3.2. Peering Point Enabled with GRE Tunnel.....	7
3.3. Peering Point Enabled with an AMT - Both Domains Multicast Enabled.....	8
3.4. Peering Point Enabled with an AMT - AD-2 Not Multicast Enabled.....	9
3.5. AD-2 Not Multicast Enabled - Multiple AMT Tunnels Through AD-2.....	11
4. Supporting Functionality.....	13
4.1. Network Interconnection Transport and Security Guidelines	14
4.2. Routing Aspects and Related Guidelines.....	15
4.2.1 Native Multicast Routing Aspects.....	15
4.2.2 GRE Tunnel over Interconnecting Peering Point.....	16
4.2.3 Routing Aspects with AMT Tunnels.....	16
4.3. Back Office Functions - Billing and Logging Guidelines...	19
4.3.1 Provisioning Guidelines.....	19
4.3.2 Application Accounting Billing Guidelines.....	20
4.3.3 Log Management Guidelines.....	21
4.3.4 Settlement Guidelines.....	21
4.4. Operations - Service Performance and Monitoring Guidelines	22
4.5. Client Reliability Models/Service Assurance Guidelines...	24

5. Security Considerations.....	25
6. IANA Considerations.....	25
7. Conclusions.....	25
8. References.....	26
8.1. Normative References.....	26
8.2. Informative References.....	26
9. Acknowledgments.....	26

## 1. Introduction

Several types of applications (e.g., live video streaming, software downloads) are well suited for delivery via multicast means. The use of multicast for delivering such applications offers significant savings for utilization of resources in any given administrative domain. End user demand for such applications is growing. Often, this requires transporting such applications across administrative domains via inter-domain peering points.

The objective of this Best Current Practices document is twofold:

- o Describe the process and establish guidelines for setting up multicast-based delivery of applications across inter-domain peering points, and
- o Catalog all required information exchange between the administrative domains to support multicast-based delivery.

While there are several multicast protocols available for use, this BCP will focus the discussion to those that are applicable and recommended for the peering requirements of today's service model, including:

- o Protocol Independent Multicast - Source Specific Multicast (PIM-SSM) [RFC4607]
- o Internet Group Management Protocol (IGMP) v3 [RFC4604]
- o Multicast Listener Discovery (MLD) [RFC4604]

This BCP is independent of the choice of multicast protocol; it focuses solely on the implications for the inter-domain peering points.

This document therefore serves the purpose of a "Gap Analysis" exercise for this process. The rectification of any gaps identified - whether they involve protocol extension development or otherwise - is beyond the scope of this document and is for further study.

## 2. Overview of Inter-domain Multicast Application Transport

A multicast-based application delivery scenario is as follows:

- o Two independent administrative domains are interconnected via a peering point.
- o The peering point is either multicast enabled (end-to-end native multicast across the two domains) or it is connected by one of two possible tunnel types:
  - o A Generic Routing Encapsulation (GRE) Tunnel [RFC2784] allowing multicast tunneling across the peering point, or
  - o An Automatic Multicast Tunnel (AMT) [IETF-ID-AMT].
- o The application stream originates at a source in Domain 1.
- o An End User associated with Domain 2 requests the application. It is assumed that the application is suitable for delivery via multicast means (e.g., live streaming of major events, software downloads to large numbers of end user devices, etc.)
- o The request is communicated to the application source which provides the relevant multicast delivery information to the EU device via a "manifest file". At a minimum, this file contains the {Source, Group} or (S,G) information relevant to the multicast stream.
- o The application client in the EU device then joins the multicast stream distributed by the application source in domain 1 utilizing the (S,G) information provided in the manifest file. The manifest file may also contain additional information that the application client can use to locate the source and join the stream.

It should be noted that the second administrative domain - domain 2 - may be an independent network domain (e.g., Tier 1 network operator domain) or it could also be an Enterprise network operated by a single customer. The peering point architecture and requirements may have some unique aspects associated with the Enterprise case.

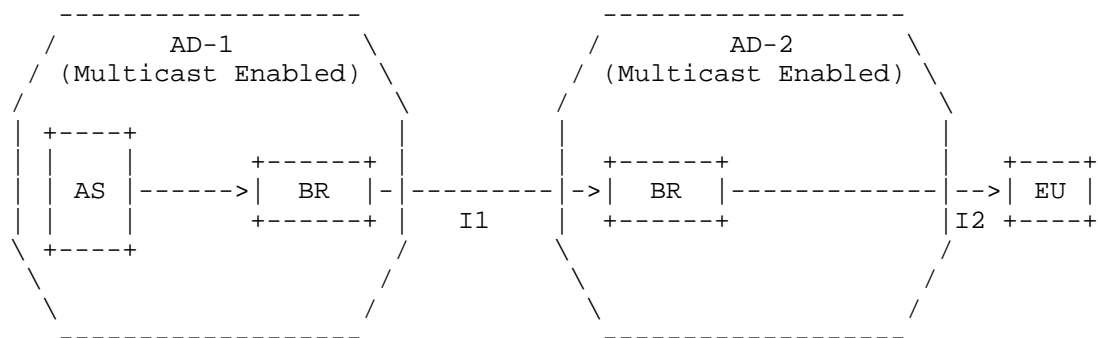
The Use Cases describing various architectural configurations for the multicast distribution along with associated requirements is described in section 3. Unique aspects related to the Enterprise network possibility will be described in this section. A comprehensive list of pertinent information that needs to be exchanged between the two domains to support various functions enabling the application transport is provided in section 4.

### 3. Inter-domain Peering Point Requirements for Multicast

The transport of applications using multicast requires that the inter-domain peering point is enabled to support such a process. There are three possible Use Cases for consideration.

#### 3.1. Native Multicast

This Use Case involves end-to-end Native Multicast between the two administrative domains and the peering point is also native multicast enabled - Figure 1.



AD = Administrative Domain (Independent Autonomous System)  
AS = Application (e.g., Content) Multicast Source  
BR = Border Router  
I1 = AD-1 and AD-2 Multicast Interconnection (MBGP or BGMP)  
I2 = AD-2 and EU Multicast Connection

Figure 1 - Content Distribution via End to End Native Multicast

Advantages of this configuration are:

- o Most efficient use of bandwidth in both domains
- o Fewer devices in the path traversed by the multicast stream when compared to unicast transmissions.

From the perspective of AD-1, the one disadvantage associated with native multicast into AD-2 instead of individual unicast to every EU in AD-2 is that it does not have the ability to count the number of End Users as well as the transmitted bytes delivered to them. This information is relevant from the perspective of customer billing and operational logs. It is assumed that such data will be collected by the application layer. The application layer mechanisms for generating this information need to be robust enough such that all pertinent requirements for the source provider and the AD operator are satisfactorily met. The specifics of these methods are beyond the scope of this document.

Architectural guidelines for this configuration are as follows:

- o Dual homing for peering points between domains is recommended as a way to ensure reliability with full BGP table visibility.
- o If the peering point between AD-1 and AD-2 is a controlled network environment, then bandwidth can be allocated accordingly by the two domains to permit the transit of non-rate adaptive multicast traffic. If this is not the case, then it is recommended that the multicast traffic should support rate-adaption.
- o The sending and receiving of multicast traffic between two domains is typically determined by local policies associated with each domain. For example, if AD-1 is a service provider and AD-2 is an enterprise, then AD-1 may support local policies for traffic delivery to, but not traffic reception from AD-2.
- o Relevant information on multicast streams delivered to End Users in AD-2 is assumed to be collected by available capabilities in the application layer. The precise nature and formats of the collected information will be determined by directives from the source owner and the domain operators.

### 3.2. Peering Point Enabled with GRE Tunnel

The peering point is not native multicast enabled in this Use Case. There is a Generic Routing Encapsulation Tunnel provisioned over the peering point. In this case, the interconnection I1 between AD-1 and AD-2 in Figure 1 is multicast enabled via a Generic Routing Encapsulation Tunnel (GRE) [RFC2784] and encapsulating the multicast protocols across the interface. The routing configuration is basically unchanged: Instead of BGP (SAFI2) across the native IP multicast link between AD-1 and AD-2, BGP (SAFI2) is now run across the GRE tunnel.

Advantages of this configuration:

- o Highly efficient use of bandwidth in both domains although not as efficient as the fully native multicast Use Case.
- o Fewer devices in the path traversed by the multicast stream when compared to unicast transmissions.
- o Ability to support only partial IP multicast deployments in AD-1 and/or AD-2.
- o GRE is an existing technology and is relatively simple to implement.

Disadvantages of this configuration:

- o Per Use Case 3.1, current router technology cannot count the number of end users or the number bytes transmitted.
- o GRE tunnel requires manual configuration.
- o GRE must be in place prior to stream starting.
- o GRE is often left pinned up

Architectural guidelines for this configuration include the following:

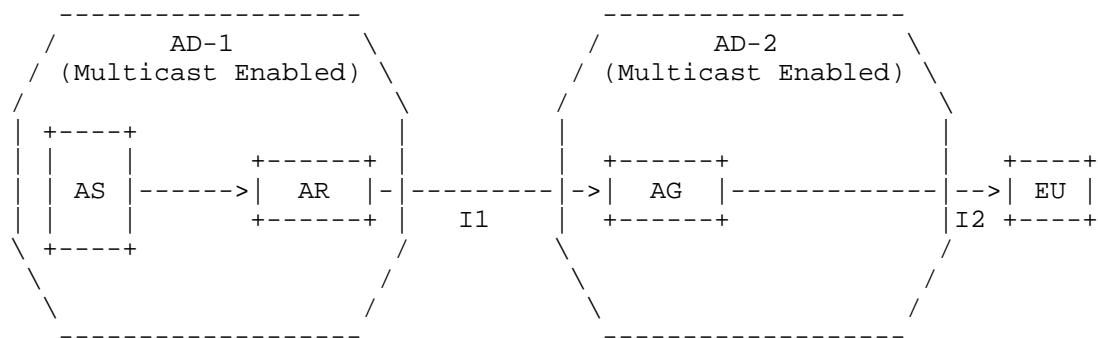
Guidelines (a) through (d) are the same as those described in Use Case 3.1.

- o GRE tunnels are typically configured manually between peering points to support multicast delivery between domains.

- o It is recommended that the GRE tunnel (tunnel server) configuration in the source network is such that it only advertises the routes to the application sources and not to the entire network. This practice will prevent unauthorized delivery of applications through the tunnel (e.g., if application - e.g., content - is not part of an agreed inter-domain partnership).

### 3.3. Peering Point Enabled with an AMT - Both Domains Multicast Enabled

Both administrative domains in this Use Case are assumed to be native multicast enabled here; however the peering point is not. The peering point is enabled with an Automatic Multicast Tunnel. The basic configuration is depicted in Figure 2.



AR = AMT Relay  
 AG = AMT Gateway  
 I1 = AMT Interconnection between AD-1 and AD-2  
 I2 = AD-2 and EU Multicast Connection

Figure 2 - AMT Interconnection between AD-1 and AD-2

Advantages of this configuration:

- o Highly efficient use of bandwidth in AD-1.

- o AMT is an existing technology and is relatively simple to implement. Attractive properties of AMT include the following:
  - o Dynamic interconnection between Gateway-Relay pair across the peering point.
  - o Ability to serve clients and servers with differing policies.

Disadvantages of this configuration:

- o Per Use Case 3.1 (AD-2 is native multicast), current router technology cannot count the number of end users or the number bytes transmitted.
- o Additional devices (AMT Gateway and Relay pairs) may be introduced into the path if these services are not incorporated in the existing routing nodes.
- o Currently undefined mechanisms to select the AR from the AG automatically.

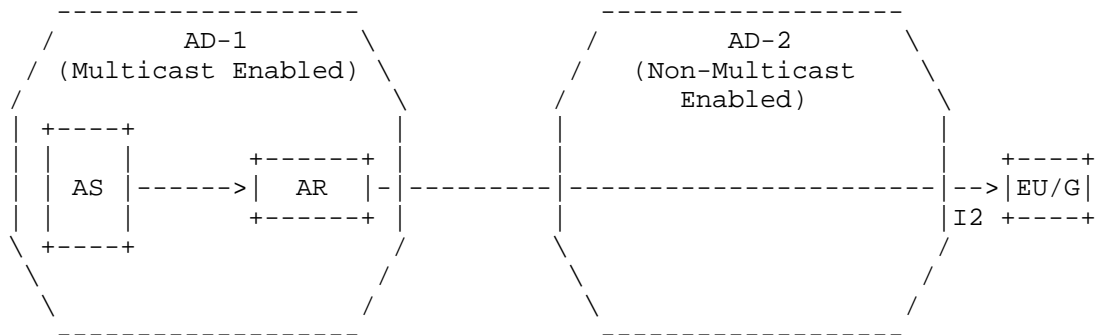
Architectural guidelines for this configuration are as follows:

Guidelines (a) through (d) are the same as those described in Use Case 3.1.

- e. It is recommended that AMT Relay and Gateway pairs be configured at the peering points to support multicast delivery between domains. AMT tunnels will then configure dynamically across the peering points once the Gateway in AD-2 receives the (S, G) information from the EU.

#### 3.4. Peering Point Enabled with an AMT - AD-2 Not Multicast Enabled

In this AMT Use Case, the second administrative domain AD-2 is not multicast enabled. This implies that the interconnection between AD-2 and the End User is also not multicast enabled as depicted in Figure 3.



AS = Application Multicast Source  
 AR = AMT Relay  
 EU/G = Gateway client embedded in EU device  
 I2 = AMT Tunnel Connecting EU/G to AR in AD-1 through Non-Multicast Enabled AD-2.

Figure 3 - AMT Tunnel Connecting AD-1 AMT Relay and EU Gateway

This Use Case is equivalent to having unicast distribution of the application through AD-2. The total number of AMT tunnels would be equal to the total number of End Users requesting the application. The peering point thus needs to accommodate the total number of AMT tunnels between the two domains. Each AMT tunnel can provide the data usage associated with each End User.

Advantages of this configuration:

- o Highly efficient use of bandwidth in AD-1.
- o AMT is an existing technology and is relatively simple to implement. Attractive properties of AMT include the following:
  - o Dynamic interconnection between Gateway-Relay pair across the peering point.
  - o Ability to serve clients and servers with differing policies.
- o Each AMT tunnel serves as a count for each End User and is also able to track data usage (bytes) delivered to the EU.

Disadvantages of this configuration:

- o Additional devices (AMT Gateway and Relay pairs) are introduced into the transport path.
- o Assuming multiple peering points between the domains, the EU Gateway needs to be able to find the "correct" AMT Relay in AD-1.

Architectural guidelines for this configuration are as follows:

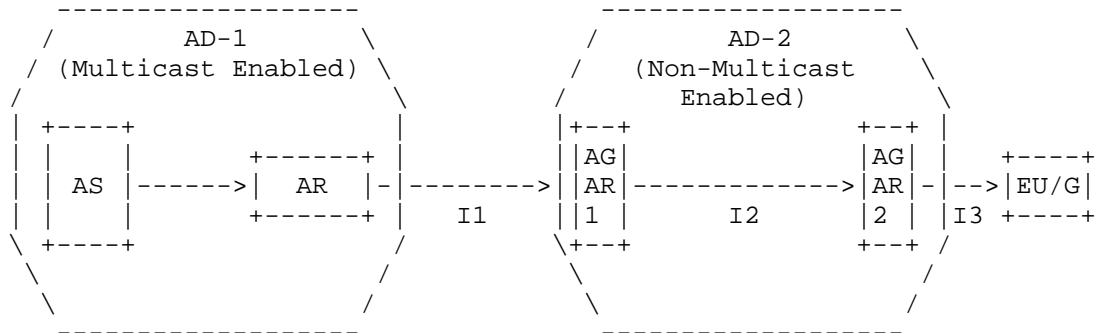
Guidelines (a) through (c) are the same as those described in Use Case 3.1.

d. It is recommended that proper procedures are implemented such that the AMT Gateway at the End User device is able to find the correct AMT Relay in AD-1 across the peering points. The application client in the EU device is expected to supply the (S, G) information to the Gateway for this purpose.

e. The AMT tunnel capabilities are expected to be sufficient for the purpose of collecting relevant information on the multicast streams delivered to End Users in AD-2.

### 3.5. AD-2 Not Multicast Enabled - Multiple AMT Tunnels Through AD-2

This is a variation of Use Case 3.4 as follows:



(Note: Diff-marks for the figure have been removed to improve viewing)

AS = Application Source  
AR = AMT Relay in AD-1  
AGAR1 = AMT Gateway/Relay node in AD-2 across Peering Point  
I1 = AMT Tunnel Connecting AR in AD-1 to GW in AGAR1 in AD-2  
AGAR2 = AMT Gateway/Relay node at AD-2 Network Edge  
I2 = AMT Tunnel Connecting Relay in AGAR1 to GW in AGAR2  
EU/G = Gateway client embedded in EU device  
I3 = AMT Tunnel Connecting EU/G to AR in AGAR2

Figure 4 - AMT Tunnel Connecting AD-1 AMT Relay and EU Gateway

Use Case 3.4 results in several long AMT tunnels crossing the entire network of AD-2 linking the EU device and the AMT Relay in AD-1 through the peering point. Depending on the number of End Users, there is a likelihood of an unacceptably large number of AMT tunnels - and unicast streams - through the peering point. This situation can be alleviated as follows:

- o Provisioning of strategically located AMT nodes at the edges of AD-2. An AMT node comprises co-location of an AMT Gateway and an AMT Relay. One such node is at the AD-2 side of the peering point (node AGAR1 in Figure 4).
- o Single AMT tunnel established across peering point linking AMT Relay in AD-1 to the AMT Gateway in the AMT node AGAR1 in AD-2.
- o AMT tunnels linking AMT node AGAR1 at peering point in AD-2 to other AMT nodes located at the edges of AD-2: e.g., AMT tunnel

I2 linking AMT Relay in AGAR1 to AMT Gateway in AMT node AGAR2 in Figure 4.

- o AMT tunnels linking EU device (via Gateway client embedded in device) and AMT Relay in appropriate AMT node at edge of AD-2: e.g., I3 linking EU Gateway in device to AMT Relay in AMT node AGAR2.

The advantage for such a chained set of AMT tunnels is that the total number of unicast streams across AD-2 is significantly reduced thus freeing up bandwidth. Additionally, there will be a single unicast stream across the peering point instead of possibly, an unacceptably large number of such streams per Use Case 3.4. However, this implies that several AMT tunnels will need to be dynamically configured by the various AMT Gateways based solely on the (S,G) information received from the application client at the EU device. A suitable mechanism for such dynamic configurations is therefore critical.

Architectural guidelines for this configuration are as follows:

Guidelines (a) through (c) are the same as those described in Use Case 3.1.

d. It is recommended that proper procedures are implemented such that the various AMT Gateways (at the End User devices and the AMT nodes in AD-2) are able to find the correct AMT Relay in other AMT nodes as appropriate. The application client in the EU device is expected to supply the (S, G) information to the Gateway for this purpose.

e. The AMT tunnel capabilities are expected to be sufficient for the purpose of collecting relevant information on the multicast streams delivered to End Users in AD-2.

#### 4. Supporting Functionality

Supporting functions and related interfaces over the peering point that enable the multicast transport of the application are listed in this section. Critical information parameters that need to be exchanged in support of these functions are enumerated along with guidelines as appropriate. Specific interface functions for consideration are as follows.

#### 4.1. Network Interconnection Transport and Security Guidelines

The term "Network Interconnection Transport" refers to the interconnection points between the two Administrative Domains. The following is a representative set of attributes that will need to be agreed to between the two administrative domains to support multicast delivery.

- o Number of Peering Points
- o Peering Point Addresses and Locations
- o Connection Type - Dedicated for Multicast delivery or shared with other services
- o Connection Mode - Direct connectivity between the two AD's or via another ISP
- o Peering Point Protocol Support - Multicast protocols that will be used for multicast delivery will need to be supported at these points. Examples of protocols include eBGP, BGMP, and MBGP.
- o Bandwidth Allocation - If shared with other services, then there needs to be a determination of the share of bandwidth reserved for multicast delivery.
- o QoS Requirements - Delay/latency specifications that need to be specified in an SLA.
- o AD Roles and Responsibilities - the role played by each AD for provisioning and maintaining the set of peering points to support multicast delivery.

From a security perspective, it is expected that normal/typical security procedures will be followed by each AD to facilitate multicast delivery to registered and authenticated end users. Some security aspects for consideration are:

- o Encryption - Peering point links may be encrypted per agreement if dedicated for multicast delivery.
- o Security Breach Mitigation Plan - In the event of a security breach, the two AD's are expected to have a mitigation plan for shutting down the peering point and directing multicast traffic

over alternated peering points. It is also expected that appropriate information will be shared for the purpose of securing the identified breach.

#### 4.2. Routing Aspects and Related Guidelines

The main objective for multicast delivery routing is to ensure that the End User receives the multicast stream from the "most optimal" source [INF\_ATIS\_10] which typically:

- o Maximizes the multicast portion of the transport and minimizes any unicast portion of the delivery, and
- o Minimizes the overall combined network(s) route distance.

This routing objective applies to both Native and AMT; the actual methodology of the solution will be different for each. Regardless, the routing solution is expected to be:

- o Scalable
- o Avoid/minimize new protocol development or modifications, and
- o Be robust enough to achieve high reliability and automatically adjust to changes/problems in the multicast infrastructure.

For both Native and AMT environments, having a source as close as possible to the EU network is most desirable; therefore, in some cases, an AD may prefer to have multiple sources near different peering points, but that is entirely an implementation issue.

##### 4.2.1 Native Multicast Routing Aspects

Native multicast simply requires that the Administrative Domains coordinate and advertise the correct source address(es) at their network interconnection peering points(i.e., border routers). An example of multicast delivery via a Native Multicast process across two administrative Domains is as follows assuming that the interconnecting peering points are also multicast enabled:

- o Appropriate information is obtained by the EU client who is a subscriber to AD-2 (see Use Case 3.1). This is usually done via an appropriate file transfer - this file is typically known as the manifest file. It contains instructions directing the EU

client to launch an appropriate application if necessary, and also additional information for the application about the source location and the group (or stream) id in the form of the "S,G" data. The "S" portion provides the name or IP address of the source of the multicast stream. The file may also contain alternate delivery information such as specifying the unicast address of the stream.

- o The client uses the join message with S,G to join the multicast stream [RFC2236].

To facilitate this process, the two AD's need to do the following:

- o Advertise the source id(s) over the Peering Points
- o Exchange relevant Peering Point information such as Capacity and Utilization (Other??)

#### 4.2.2 GRE Tunnel over Interconnecting Peering Point

If the interconnecting peering point is not multicast enabled and both ADs are multicast enabled, then a simple solution is to provision a GRE tunnel between the two ADs - see Use Case 3.2.2. The termination points of the tunnel will usually be a network engineering decision, but generally will be between the border routers or even between the AD 2 border router and the AD 1 source (or source access router). The GRE tunnel would allow end-to-end native multicast or AMT multicast to traverse the interface. Coordination and advertisement of the source IP is still required.

The two AD's need to follow the same process as described in 4.2.1 to facilitate multicast delivery across the Peering Points.

#### 4.2.3 Routing Aspects with AMT Tunnels

Unlike Native (with or without GRE), an AMT Multicast environment is more complex. It presents a dual layered problem because there are two criteria that should be simultaneously meet:

- o Find the closest AMT relay to the end-user that also has multicast connectivity to the content source and
- o Minimize the AMT unicast tunnel distance.

There are essentially two components to the AMT specification:

- o AMT Relays: These serve the purpose of tunneling UDP multicast traffic to the receivers (i.e., End Points). The AMT Relay will receive the traffic natively from the multicast media source and will replicate the stream on behalf of the downstream AMT Gateways, encapsulating the multicast packets into unicast packets and sending them over the tunnel toward the AMT Gateway. In addition, the AMT Relay may perform various usage and activity statistics collection. This results in moving the replication point closer to the end user, and cuts down on traffic across the network. Thus, the linear costs of adding unicast subscribers can be avoided. However, unicast replication is still required for each requesting endpoint within the unicast-only network.
- o AMT Gateway (GW): The Gateway will reside on an on End-Point - this may be a Personal Computer (PC) or a Set Top Box (STB). The AMT Gateway receives join and leave requests from the Application via an Application Programming Interface (API). In this manner, the Gateway allows the endpoint to conduct itself as a true Multicast End-Point. The AMT Gateway will encapsulate AMT messages into UDP packets and send them through a tunnel (across the unicast-only infrastructure) to the AMT Relay.

The simplest AMT Use Case (section 3.3) involves peering points that are not multicast enabled between two multicast enabled ADs. An AMT tunnel is deployed between an AMT Relay on the AD 1 side of the peering point and an AMT Gateway on the AD 2 side of the peering point. One advantage to this arrangement is that the tunnel is established on an as needed basis and need not be a provisioned element. The two ADs can coordinate and advertise special AMT Relay Anycast addresses with each other - though they may alternately decide to simply provision Relay addresses, though this would not be a optimal solution in terms of scalability.

Use Cases 3.4 and 3.5 describe more complicated AMT situations as AD-2 is not multicast enabled. For these cases, the End User device needs to be able to setup an AMT tunnel in the most optimal manner. Using an Anycast IP address for AMT Relays allows for all AMT Gateways to find the "closest" AMT Relay - the nearest edge of the multicast topology of the source. An example of a basic delivery via an AMT Multicast process for these two Use Cases is as follows:

- o The manifest file is obtained by the EU client application. This file contains instructions directing the EU client to an ordered list of particular destinations to seek the requested stream and, for multicast, specifies the source location and the group (or stream) ID in the form of the "S,G" data. The "S" portion provides

the URI (name or IP address) of the source of the multicast stream and the "G" identifies the particular stream originated by that source. The manifest file may also contain alternate delivery information such as the address of the unicast form of the content to be used, for example, if the multicast stream becomes unavailable.

- o Using the information in the manifest file, and possibly information provisioned directly in the EU client, a DNS query is initiated in order to connect the EU client/AMT Gateway to an AMT Relay.
- o Query results are obtained, and may return an Anycast address or a specific unicast address of a relay. Multiple relays will typically exist. The Anycast address is a routable "pseudo-address" shared among the relays that can gain multicast access to the source.
- o If a specific IP address unique to a relay was not obtained, the AMT Gateway then sends a message (e.g., the discovery message) to the Anycast address such that the network is making the routing choice of particular relay - e.g., closest relay to the EU. (Note that in IPv6 there is a specific Anycast format and Anycast is inherent in IPv6 routing, whereas in IPv4 Anycast is handled via provisioning in the network. Details are out of scope for this document.)
- o The contacted AMT Relay then returns its specific unicast IP address (after which the Anycast address is no longer required). Variations may exist as well.
- o The AMT Gateway uses that unicast IP address to initiate a three-way handshake with the AMT Relay.
- o AMT Gateway provides "S,G" to the AMT Relay (embedded in AMT protocol messages).
- o AMT Relay receives the "S,G" information and uses the S,G to join the appropriate multicast stream, if it has not already subscribed to that stream.
- o AMT Relay encapsulates the multicast stream into the tunnel between the Relay and the Gateway, providing the requested content to the EU.

Note: Further routing discussion on optimal method to find "best AMT Relay/GW combination" and information exchange between AD's to be provided.

#### 4.3. Back Office Functions - Billing and Logging Guidelines

Back Office refers to the following:

- o Servers and Content Management systems that support the delivery of applications via multicast and interactions between ADs.
- o Functionality associated with logging, reporting, ordering, provisioning, maintenance, service assurance, settlement, etc.

##### 4.3.1 Provisioning Guidelines

Resources for basic connectivity between ADs Providers need to be provisioned as follows:

- o Sufficient capacity must be provisioned to support multicast-based delivery across ADs.
- o Sufficient capacity must be provisioned for connectivity between all supporting back-offices of the ADs as appropriate. This includes activating proper security treatment for these back-office connections (gateways, firewalls, etc) as appropriate.
- o Routing protocols as needed, e.g. configuring routers to support these.

Provisioning aspects related to Multicast-Based inter-domain delivery are as follows.

The ability to receive requested application via multicast is triggered via the manifest file. Hence, this file must be provided to the EU regarding multicast URL - and unicast fallback if applicable. AD-2 must build manifest and provision capability to provide the file to the EU.

Native multicast functionality is assumed to be available in across many ISP backbones, peering and access networks. If however, native multicast is not an option (Use Cases 3.4 and 3.5), then:

- o EU must have multicast client to use AMT multicast obtained either from Application Source (per agreement with AD-1) or from AD-1 or AD-2 (if delegated by the Application Source).

- o If provided by AD-1/AD-2, then the EU could be redirected to a client download site (note: this could be an Application Source site). If provided by the Application Source, then this Source would have to coordinate with AD-1 to ensure the proper client is provided (assuming multiple possible clients).
- o Where AMT Gateways support different application sets, all AD-2 AMT Relays need to be provisioned with all source & group addresses for streams it is allowed to join.
- o DNS across each AD must be provisioned to enable a client GW to locate the optimal AMT Relay (i.e. longest multicast path and shortest unicast tunnel) with connectivity to the content's multicast source.

Provisioning Aspects Related to Operations and Customer Care are stated as follows.

Each AD provider is assumed to provision operations and customer care access to their own systems.

AD-1's operations and customer care functions must have visibility to what is happening in AD-2's network or to the service provided by AD-2, sufficient to verify their mutual goals and operations, e.g. to know how the EU's are being served. This can be done in two ways:

- o Automated interfaces are built between AD-1 and AD-2 such that operations and customer care continue using their own systems. This requires coordination between the two AD's with appropriate provisioning of necessary resources.
- o AD-1's operations and customer care personnel are provided access directly to AD-2's system. In this scenario, additional provisioning in these systems will be needed to provide necessary access. Additional provisioning must be agreed to by the two AD-2s to support this option.

#### 4.3.2 Application Accounting Billing Guidelines

All interactions between pairs of ADs can be discovered and/or be associated with the account(s) utilized for delivered applications. Supporting guidelines are as follows:

- o A unique identifier is recommended to designate each master account.
- o AD-2 is expected to set up "accounts" (logical facility generally protected by login/password/credentials) for use by AD-1. Multiple

accounts and multiple types/partitions of accounts can apply, e.g. customer accounts, security accounts, etc.

#### 4.3.3 Log Management Guidelines

Successful delivery of applications via multicast between pairs of interconnecting ADs requires that appropriate logs will be exchanged between them in support. Associated guidelines are as follows.

AD-2 needs to supply logs to AD-1 per existing contract(s). Examples of log types include the following:

- o Usage information logs at aggregate level.
- o Usage failure instances at an aggregate level.
- o Grouped or sequenced application access performance/behavior/failure at an aggregate level to support potential Application Provider-driven strategies. Examples of aggregate levels include grouped video clips, web pages, and sets of software download.
- o Security logs, aggregated or summarized according to agreement (with additional detail potentially provided during security events, by agreement).
- o Access logs (EU), when needed for troubleshooting.
- o Application logs (what is the application doing), when needed for shared troubleshooting.
- o Syslogs (network management), when needed for shared troubleshooting.

The two ADs may supply additional security logs to each other as agreed to by contract(s). Examples include the following:

- o Information related to general security-relevant activity which may be of use from a protective or response perspective, such as types and counts of attacks detected, related source information, related target information, etc.
- o Aggregated or summarized logs according to agreement (with additional detail potentially provided during security events, by agreement)

#### 4.3.4 Settlement Guidelines

Settlements between the ADs relate to (1) billing and reimbursement aspects for delivery of applications, and (2) aggregation, transport, and collection of data in preparation for the billing and

reimbursement aspects for delivery of applications for the Application Provider. At a high level:

- o AD-2 collects "usage" data for AD-1 related to application delivery to End Users, and submits invoices to AD-1 based on this usage data. The data may include information related to the type of content delivered, total bandwidth utilized, storage utilized, features supported, etc.
- o AD-1 collects all available data from partner AD-2 and creates aggregate reports pertaining to responsible Application Providers, and submits subsequent reports to these Providers for reimbursements.
- o AD-1 may convey charging values or charging rules to the AD-2, proactively or in response to a query, especially in cases where these may change.
- o AD-2 may convey prices/rates to AD-1, proactively or in response to a query, especially in cases where these may change.
- o Usage data may be collected per end user or on an aggregated basis; the method of collection will depend on the application delivered and/or the agreements with the source provider. In all cases, usage volume is expected to be in terms of delivered packet bits or bytes.

#### 4.4. Operations - Service Performance and Monitoring Guidelines

Service Performance refers to monitoring metrics related to multicast delivery via probes. The focus is on the service provided by AD-2 to AD-1 on behalf of all multicast application sources (metrics may be specified for SLA use or otherwise). Associated guidelines are as follows:

- o Both AD's are expected to monitor, collect, and analyze service performance metrics for multicast applications. AD-2 provides relevant performance information to AD-1; this enables AD-1 to create an end-to-end performance view on behalf of the multicast application source.
- o Both AD's are expected to agree on the type of probes to be used to monitor multicast delivery performance. For example, AD-2 may permit AD-1's probes to be utilized in the AD-2 multicast service footprint. Alternately, AD-2 may deploy its own probes and relay performance information back to AD-1.

- o In the event of performance degradation (SLA violation), AD-1 may have to compensate the multicast application source per SLA agreement. As appropriate, AD-1 may seek compensation from AD-2 if the cause of the degradation is in AD-2's network.

Service Monitoring generally refers to a service (as a whole) provided on behalf of a particular multicast application source provider. It thus involves complaints from End Users when service problems occur. EU's direct their complaints to the source provider; in turn the source provider submits these complaints to AD-1. The responsibility for service delivery lies with AD-1; as such AD-1 will need to determine where the service problem is occurring - its own network or in AD-2. It is expected that each AD will have tools to monitor multicast service status in its own network.

- o Both AD's will determine how best to deploy multicast service monitoring tools. Typically, each AD will deploy its own set of monitoring tools; in which case, both AD's are expected to inform each other when multicast delivery problems are detected.
- o AD-2 may experience some problems in its network. For example, for the AMT Use Cases, one or more AMT Relays may be experiencing difficulties. AD-2 may be able to fix the problem by rerouting the multicast streams via alternate AMT Relays. If the fix is not successful and multicast service delivery degrades, then AD-2 needs to report the issue to AD-1.
- o When problem notification is received from a multicast application source, AD-1 determines whether the cause of the problem is within its own network or within the AD-2 domain. If the cause is within the AD-2 domain, then AD-1 supplies all necessary information to AD-2. Examples of supporting information include the following:
  - o Kind of problem(s)
  - o Starting point & duration of problem(s).
  - o Conditions in which problem(s) occur.
  - o IP address blocks of affected users.
  - o ISPs of affected users.

- o Type of access e.g., mobile versus desktop.
- o Locations of affected EUs.
- o Both AD's conduct some form of root cause analysis for multicast service delivery problems. Examples of various factors for consideration include:
  - o Verification that the service configuration matches the product features.
  - o Correlation and consolidation of the various customer problems and resource troubles into a single root service problem.
  - o Prioritization of currently open service problems, giving consideration to problem impact, service level agreement, etc.
  - o Conduction of service tests, including one time tests or a series of tests over a period of time.
  - o Analysis of test results.
  - o Analysis of relevant network fault or performance data.
  - o Analysis of the problem information provided by the customer (CP).
- o Once the cause of the problem has been determined and the problem has been fixed, both AD's need to work jointly to verify and validate the success of the fix.
- o Faults in service could lead to SLA violation for which the multicast application source provider may have to be compensated by AD-1. Subsequently, AD-1 may have to be compensated by AD-2 based on the contract.

#### 4.5. Client Reliability Models/Service Assurance Guidelines

There are multiple options for instituting reliability architectures, most are at the application level. Both AD's should work those out with their contract/agreement and with the multicast application source providers.

Network reliability can also be enhanced by the two AD's by provisioning alternate delivery mechanisms via unicast means.

#### 5. Security Considerations

DRM and Application Accounting, Authorization and Authentication should be the responsibility of the multicast application source provider and/or AD-1. AD-1 needs to work out the appropriate agreements with the source provider.

Network has no DRM responsibilities, but might have authentication and authorization obligations. These though are consistent with normal operations of a CDN to insure end user reliability, security and network security

AD-1 and AD-2 should have mechanisms in place to ensure proper accounting for the volume of bytes delivered through the peering point and separately the number of bytes delivered to EUs.

If there are problems related to failure of token authentication when end-users are supported by AD-2, then some means of validating proper working of the token authentication process (e.g., back-end servers querying the multicast application source provider's token authentication server are communicating properly) should be considered. Details will have to be worked out during implementation (e.g., test tokens or trace token exchange process).

#### 6. IANA Considerations

#### 7. Conclusions

This Best Current Practice document provides detailed Use Case scenarios for the transmission of applications via multicast across peering points between two Administrative Domains. A detailed set of guidelines supporting the delivery is provided for all Use Cases.

For Use Cases involving AMT tunnels (cases 3.4 and 3.5), it is recommended that proper procedures are implemented such that the various AMT Gateways (at the End User devices and the AMT nodes in AD-2) are able to find the correct AMT Relay in other AMT nodes as appropriate. Section 4.3 provides an overview of one method that finds the optimal Relay-Gateway combination via the use of an Anycast IP address for AMT Relays.

## 8. References

### 8.1. Normative References

[RFC2784] D. Farinacci, T. Li, S. Hanks, D. Meyer, P. Traina, "Generic Routing Encapsulation (GRE)", RFC 2784, March 2000

[IETF-ID-AMT] G. Bumgardner, "Automatic Multicast Tunneling", draft-ietf-mboned-auto-multicast-13, April 2012, Work in progress

[RFC4604] H. Holbrook, et al, "Using Internet Group Management Protocol Version 3 (IGMPv3) and Multicast Listener Discovery Protocol Version 2 (MLDv2) for Source Specific Multicast", RFC 4604, August 2006

[RFC4607] H. Holbrook, et al, "Source Specific Multicast", RFC 4607, August 2006

### 8.2. Informative References

[INF\_ATIS\_10] "CDN Interconnection Use Cases and Requirements in a Multi-Party Federation Environment", ATIS Standard A-0200010, December 2012

## 9. Acknowledgments

Authors' Addresses

Percy S. Tarapore  
AT&T  
Phone: 1-732-420-4172  
Email: tarapore@att.com

Robert Sayko  
AT&T  
Phone: 1-732-420-3292  
Email: rs1983@att.com

Greg Shepherd  
Cisco  
Phone:  
Email: shep@cisco.com

Toerless Eckert  
Cisco  
Phone:  
Email: eckert@cisco.com

Ram Krishnan  
Brocade  
Phone:  
Email: ramk@brocade.com

