SAM Overlay Protocol
draft-buford-irtf-sam-overlay-protocol-01.txt

John Buford, Avaya Labs Research
IETF 71

March 12, 2008

Abstract

- The previously proposed Hybrid Overlay Multicast Framework uses a structured peer-to-peer overlay to connect peers in different types of multicast regions of the Internet.

- We use AMT (Automatic IP Multicast Without Explicit Tunnels) to connect peers in ALM (Application Layer Multicast) regions with peers in native multicast regions.

- Here we define the overlay messaging needed to support the multicast tree operations.
  - Intended to be <u>algorithm agnostic</u>

- **One goal** of this approach is to ensure that the SAM framework is compatible with the <u>P2P-SIP overlay</u>, which is still being defined.

- **Another goal** is to be <u>overlay agnostic</u> so that different overlay algorithms can interoperate in a single SAM (Scalable Adaptive Multicast) session.
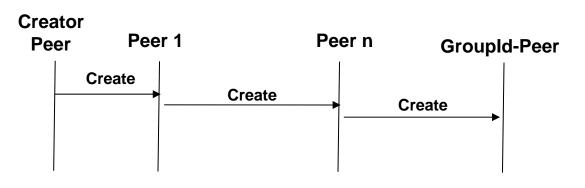
Topics

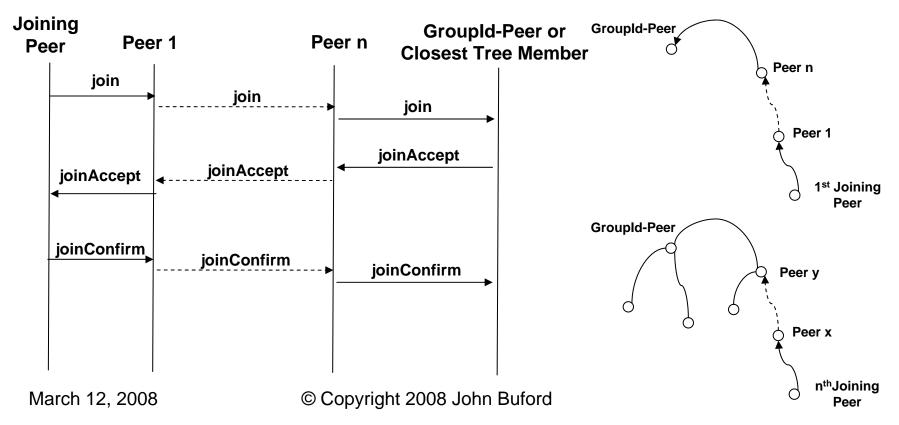March 12, 2008   © Copyright 2008 John Buford

Create(PeerId, SessionKey, GroupId, Options)

- A new ALM tree is created in the overlay with the identity specified by GroupId. The usual interpretation of GroupId is that the peer with peer id closest to and less than the GroupId is the root of the tree. The tree has no children at the time it is created.

- The GroupId is generated from a well-known session key to be used by other Peers to address the multicast tree in the overlay. The generation of the GroupId from the SessionKey MUST be done using the overlay's id generation mechanism.

- Example: Create(peer-123,"John's meeting 2008-03-10", GroupId, [])
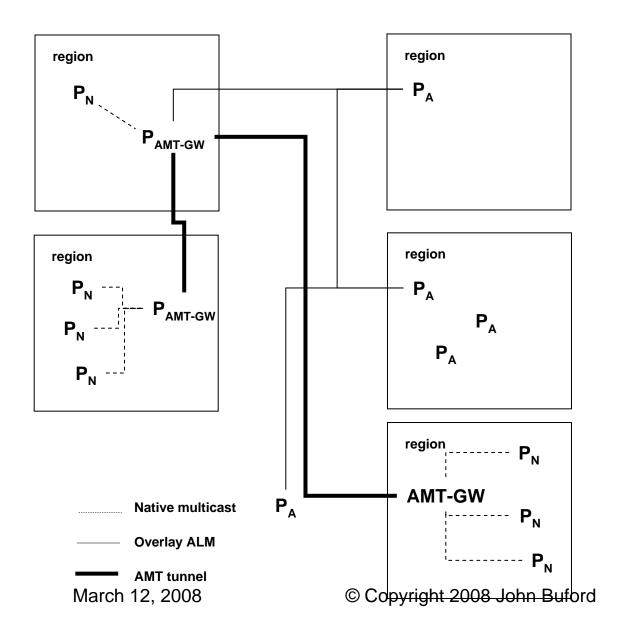
Creator Peer   **Peer 1**   **Peer n**   **GroupId-Peer**

Create

Create

Create

Join(PeerId, GroupId, Options)
JoinAccept(ParentPeerId, ChildPeerId, GroupId, Options)
JoinConfirm(ChildPeerId, ParentPeerId, GroupId, Options)
JoinDecline(PeerId, ParentPeerId, GroupId)

- Causes the distributed algorithm for peer join of a specific ALM group to be invoked. If successful, the PeerId is notified of one or more candidate parent peers in one or more JoinAccept messages.
- The particular ALM join algorithm is not specified in this protocol.
- Multiple step approach permits joining peer to identify multiple candidates and then select one or more parents (could use multiple parents for redundancy)

ALM vs NM Tree Topology

- P_N – Peer in native multicast region
- P_A – Peer in ALM region
- AMT-GW – host which acts as NM gateway to other AMT supported regions
- P_A in same region in same group use some overlay-based mechanism to find closest PA parent node
- Note: AMT tunnels and NM paths are outside overlay routing

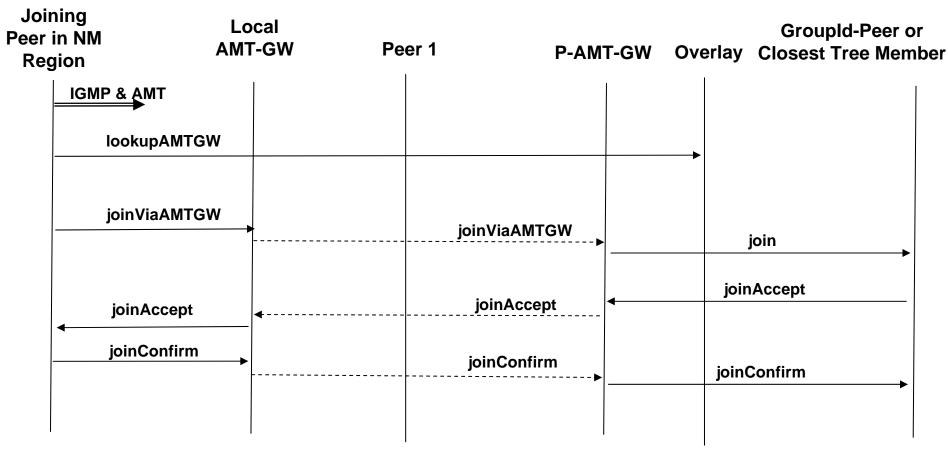March 12, 2008

© Copyright 2008 John Buford

JoinViaAMTGateway(PeerId, AMT-GW, GroupId, Options)

- A request to create a hybrid native multicast connection for the specified PeerId peer to join the tree identified by the GroupId.
- The request is transmitted to one or more parent peer candidates and/or rendezvous peers for the specified group id, according to the usual join protocol in this overlay.
- If the parent peer is a P-AMT-GW (a peer which supports the AMT-GW interface), then after JoinAccept and JoinConfirm steps, instead of an overlay parent-child link, an AMT tunnel is formed using the AMT protocol from the P-AMT-GW to the specified AMT-GW to which the Peer is associated.
- If parent peer is a peer P-NM in native multicast region, then after JoinAccept and JoinConfirm steps, the tunnel is created between P-NM's AMT-GW and the specified AMT-GW, using the AMT protocol.
- If parent peer is a P-ALM, then the requested is propagated to other peers in the tree according to the join processing rules.

JoinViaAMTGateway(PeerId, AMT-GW, GroupId, Options)

- Request is transmitted to one or more parent peer candidates and/or rendezvous peers for the specified group id, according to the usual join protocol in this overlay
  - If the parent peer is a P_{AMT-GW}, then a tunnel is formed using the AMT protocol from the P_{AMT-GW} to the specified AMT-GW
  - If parent peer is a peer P_{NM} in native multicast region, then the tunnel is created between P_{NM}'s AMT-GW and the specified AMT-GW, using the AMT protocol
  - If parent peer is a P_{ALM}, then the requested is propagated to other peers in the tree according to the join rules
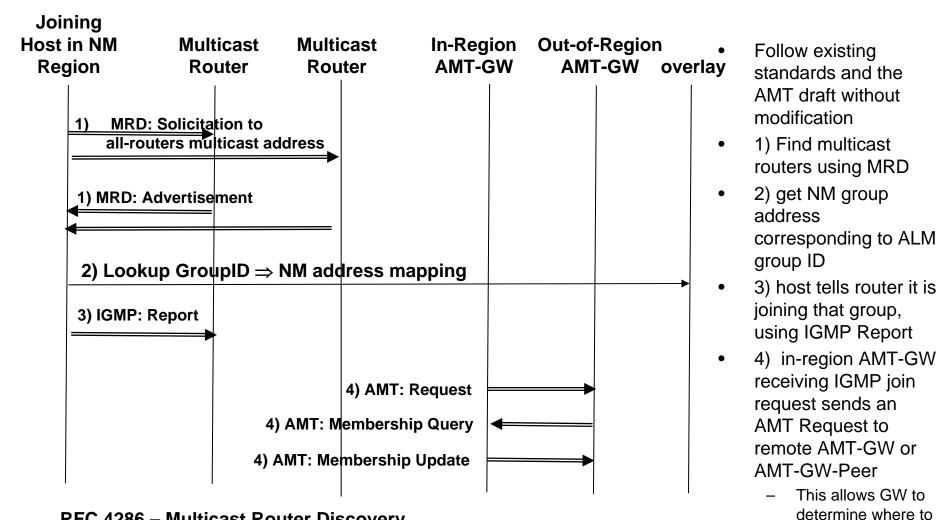
Join via AMT GW

Joining Peer in NM Region **Local AMT-GW** **Peer 1** **P-AMT-GW** **Overlay** **GroupId-Peer or Closest Tree Member**

IGMP & AMT

lookupAMTGW

joinViaAMTGW

joinViaAMTGW

join

joinAccept

joinAccept

joinAccept

joinConfirm

joinConfirm

joinConfirm

Notes:
- Assume P-AMT-GW not already part of this session
-

March 12, 2008   © Copyright 2008 John Buford

Native Layer View (IPv4)

Joining Host in NM Region | **Multicast Router** | **Multicast Router** | **In-Region AMT-GW** | **Out-of-Region AMT-GW** | **overlay**

1) **MRD: Solicitation to all-routers multicast address**

1) **MRD: Advertisement**

2) **Lookup GroupID ⇒ NM address mapping**

3) **IGMP: Report**

4) **AMT: Request**

4) **AMT: Membership Query**

4) **AMT: Membership Update**

RFC 4286 – Multicast Router Discovery
RFC 3376 – Internet Group Management Protocol

- Follow existing standards and the AMT draft without modification
- 1) Find multicast routers using MRD
- 2) get NM group address corresponding to ALM group ID
- 3) host tells router it is joining that group, using IGMP Report
- 4) in-region AMT-GW receiving IGMP join request sends an AMT Request to remote AMT-GW or AMT-GW-Peer
  - This allows GW to determine where to forward joins to

March 12, 2008 © Copyright 2008 John Buford

leaveViaAMTGateway

- leaveViaAMTGateway(peer, AMT-GW, group_id)
  - Peer is the peer requesting to leave the ALM group identified by group_id.
  - AMT-GW is the ip address of the AMT gateway that the peer uses in its native multicast region
  - Request is transmitted the parent peer which is associated with the AMT-GW or provides that role
    - If the parent peer is a P_{AMT-GW}, then it removes the child from its AMT children list and may tear down the AMT tunnel P_{AMT-GW} to the specified AMT-GW if no other children are using it
    - If parent peer is a peer P_{NM} in native multicast region, then the tunnel is created between P_{NM}'s AMT-GW and the specified AMT-GW, using the AMT protocol

joinWithNativeLink

- joinWithNativeLink(ChildPeer,ParentPeer,Group, Options)
  - ParentPeer and ChildPeer are either in same NM region or in two different NM regions with capability for AMT
  - Since media is passed via NM path, parent-child relationship is for control and membership management
  - Typical use is for a peer in NM region to access the local native multicast path for this group
  - This allows child to select specific parent peer, overriding selection based on the basic join method

Reform(GroupId, { PeerId }, Options)

- Re-Form or Optimize Tree

- This triggers a reorganization of either the entire tree or only a sub-tree.

- It MAY include hints to specific peers of recommended parent or child peers to reconnect to.

- A peer receiving this message MAY ignore it, MAY propagate it to other peers in its subtree, and MAY invoke local algorithms for selecting preferred parent and/or child peers.

Heartbeat(PeerId1, PeerId2, GroupId)

- A node signals to its adjacent nodes in the tree that it is alive.

- If a peer does not receive a Heartbeat message within N heartbeat time intervals, it MUST treat this as an explicit Leave message from the unresponsive peer.

  - N is configurable.

Topics

AMT GW Advertisement and Discovery

- Allows peer to disclose to other peers in the overlay their ability to act as a native-multicast gateway (as in AMT) for peers in a given region.
- We expect to use the P2P Publish and Lookup messages for this purpose.
- But to avoid collision with the semantics of those operations, we temporarily define shadow versions within the SAM extension.
- Publish stores an advertisement object for a peer with is an AMT gateway in the DHT for the overlay, under a given key.
- Operations
  - PublishAMTGateway(PeerId, Key, Region, Options)
  - LookupAMTGateway(PeerId, Key)

  - PeerId: the peer which is the AMT gateway
  - Key: the key by which other peers lookup the advertisement, details TBD. There can be more than one key.
  - Region: an id for a region, using some region identification scheme TBD. For example, the Autonomous System Number (ASN) could be used.[RFC 1930]

Peer Region and Multicast Properties Messages

- Peers can advertise the network region that they belong to and its native multicast properties if any. Similar to AMT Gateway advertisement and discovery, uses the DHT for lookup and publish.

- Operations
  - PublishPeerNMInfo(PeerId, Key, Options)
  - LookupPeerNMInfo(PeerId, Key)

Topics

March 12, 2008   © Copyright 2008 John Buford

Message Format

- We propose a new message be added to the 18 existing message types:
  - P2PP-Msg = ... | Experimental

- Experimental Message
  - Experimental = Common-header
  - Peer-Info
  - Certificate-Sign-Request
  - Password
  - Extension-Name
  - Extension-Body

- Extension-Name = SAM | ...
- Extension-Body = extension-specific

SAM Extension Body

- The extension-body used by the SAM extension to P2PP is as follows.
- Extension-Body = Extension-Message-Type | Arguments | Options
- Extension-Message-Type =
  - CreateTree | Join | JoinAccept | JoinConfirm | JoinDecline | JoinViaAMTGateway | JoinWithNativeLink | Leave | LeaveViaAMTGateway | Heartbeat | PublishAMTGateway | LookupAMTGateway | PublishPeerNMInfo | LookupPeerNMInfo | Reform
- Arguments = PID | GroupId | Key | SessionKey | AMT-GW
  - PID = PeerId | AdjacentPeerId | ChildPeerId | ParentPeerId
- Options = {Option}*
- Option = Name, Value

Discussion