

Hybrid Overlay Multicast Framework
draft-irtf-sam-hybrid-overlay-framework-02.txt

John Buford, Avaya Labs Research
IETF 71

Topics

- Changes v1 to v2
 - Conform to new ID boilerplate
 - Update section 6.5
- Open Issues
- Targeting an Informational RFC for the SAM Framework?
- Other drafts prepared or planned

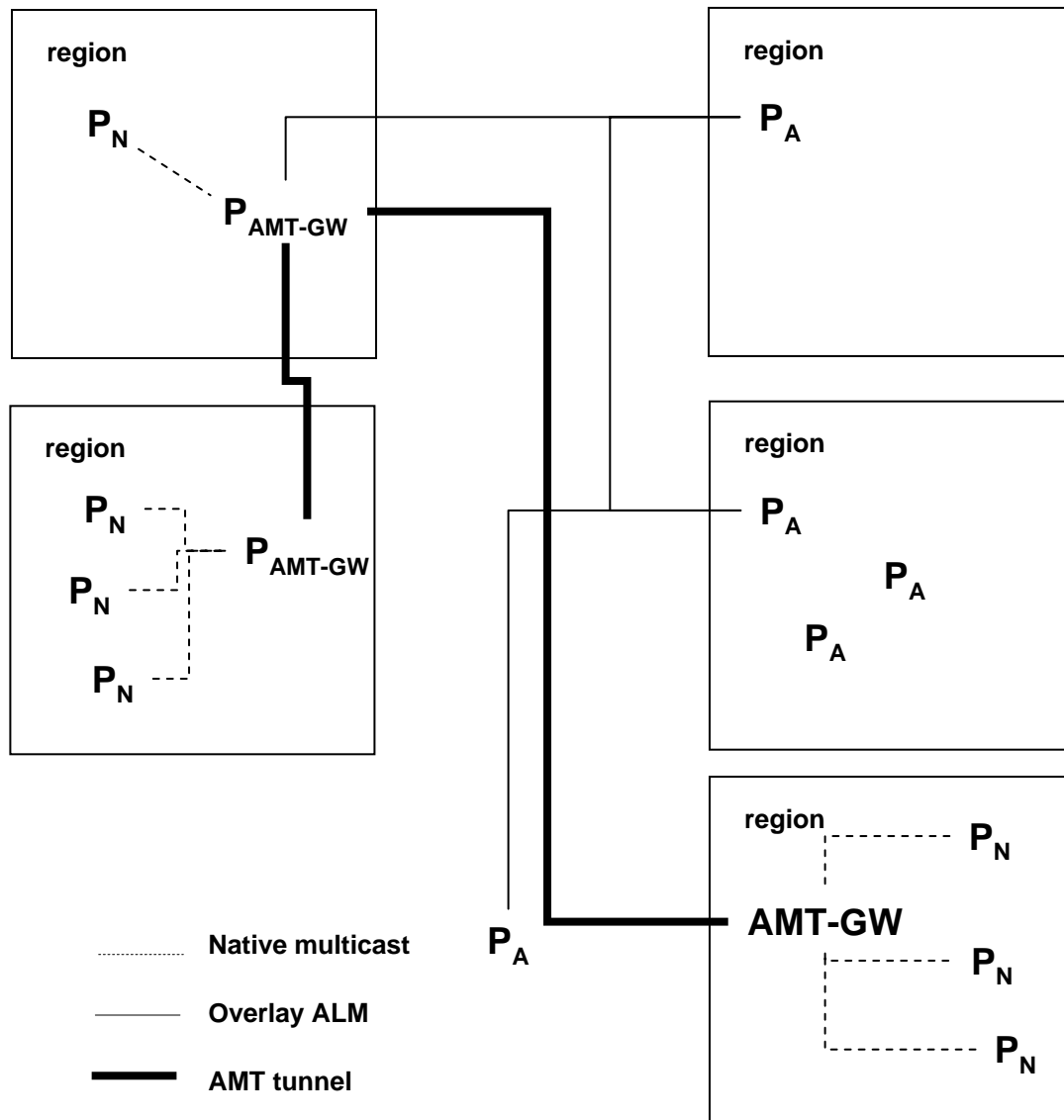
Recap of Open Issues (Sec. 6.5 v1)

- ALM tree topology vs NM topology and NM-ALM edges
 - NM edges might be formed between peers that are not necessarily those parent-child links that would be selected by the ALM join algorithm
- Single NM-ALM edge nodes vs multi NM peers from same region in the tree
 - If the first P-NM in NM region joins ALM tree using ALM edges, then when is the switch-over to the AMT-GW path and how is it done;
 - Likewise in reverse direction, when peers are leaving multicast group
- Initial tree membership is ALM vs initial tree membership is NM
 - Does it matter if the tree is first formed with NM only nodes or ALM only nodes and then expanded to hybrid membership?

ALM vs NM Tree Topology (1/3)

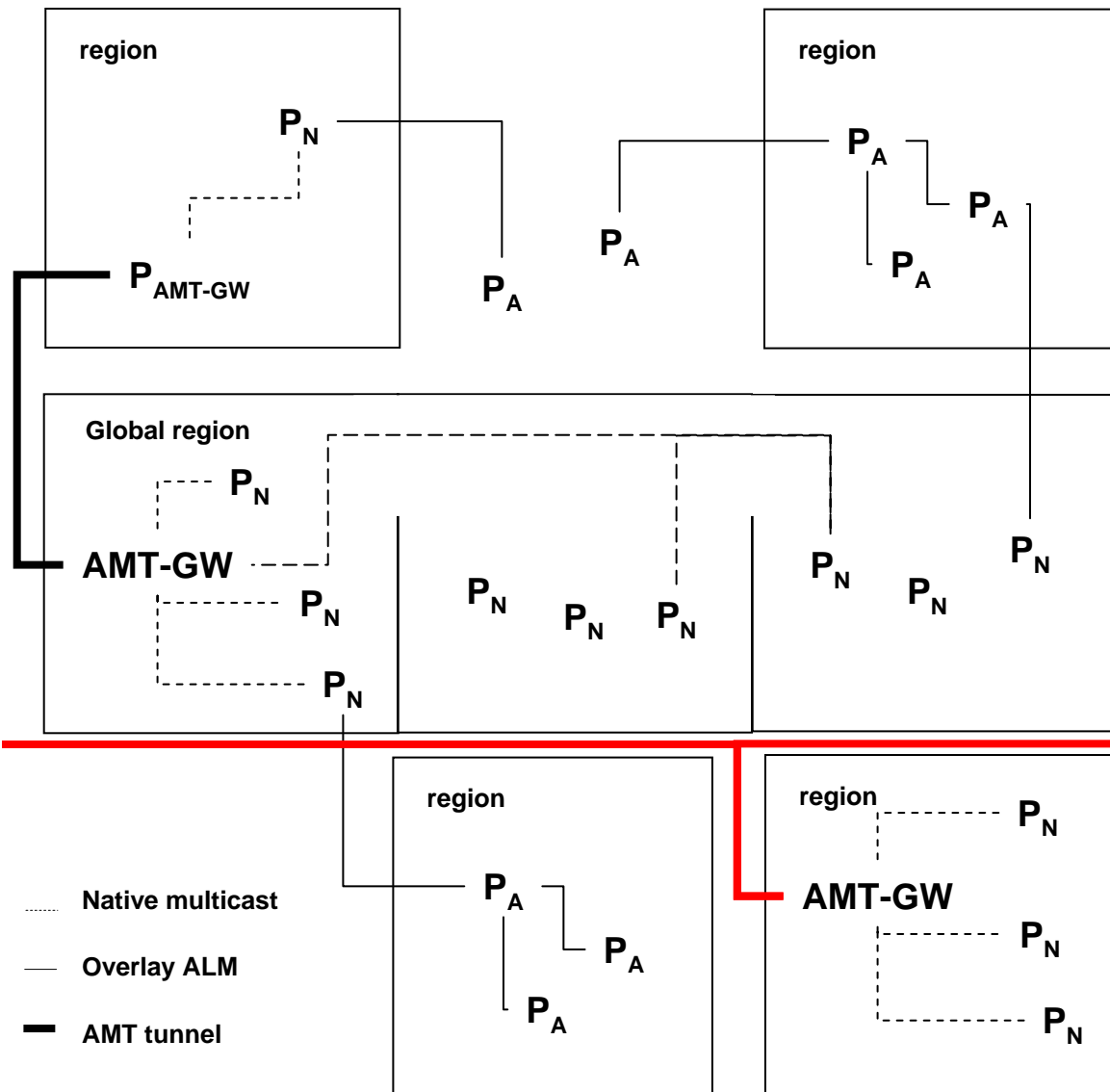
- Issue:
 - How do Peers know when to use NM edge when joining a group?
 - They know what region they are in
 - If they are in an NM region, then they know whether the group is already subscribed to by any AMT-GW-Peer in the region (is there a race condition when 2 peers in the region try to be the first ones which join via different AMT-GW-Peers- this is ok since it creates redundancy although perhaps inefficient)
- Alternatives
 - Peer is in NM region and knows AMT-GW-Peer
 - Propagates overlay join to AMT-GW-Peer in same region (this tells it to expect a NM-join request for that group
 - AMT-GW-Peer joins group using overlay join, if not already
 - Receives ACK with corresponding NM group
 - Peer joins the NM group corresponding to the overlay group (This creates a NM tree in the region)
 - Peer is in NM region and knows AMT Gateway
 - AMT-GW either joins group using overlay join (becoming AMT-GW-Peer), connects to another AMT-GW already in this group, or connects to an AMT-GW-Peer using AMT
- Problems
 - If most endpoints are in separate NM regions, is the overlay efficient way to inter connect them?
 - An inter-region link can specify tunnel or ALM format
 - Two PN form tunnel link; two PA form ALM link; otherwise could be either
 - Makes the AMT-GW-Peer a point of resource contention if it is used for all sessions in the region
 - There can be more than one AMT-GW-Peer in a region

ALM vs NM Tree Topology (2/3)



- 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 P_A parent node
- Note: AMT tunnels and NM paths are outside overlay routing

ALM vs NM Tree Topology (3/3)



- Global “region”
 - E.g., global enterprise network
 - Is region’s AMT GW the most effective NM path? This effects not only peers in the global region, but also peers which could or do form subtrees from these peers

New Join Operators

- **joinViaAMTGateway**
 - Data path to in-region AMT Gateway or P-AMT-GW is formed
 - Control path could be different
- **joinViaNativeLink**
 - Allows peer in NM region to select parent in the same NM region, overriding the ALM join procedure

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.

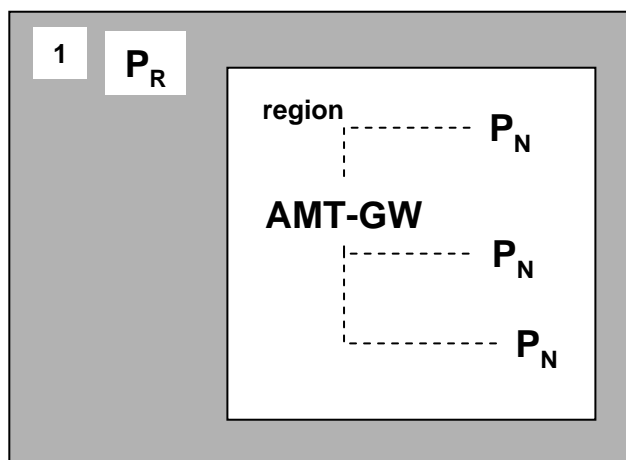
leaveViaAMTGateway

- `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

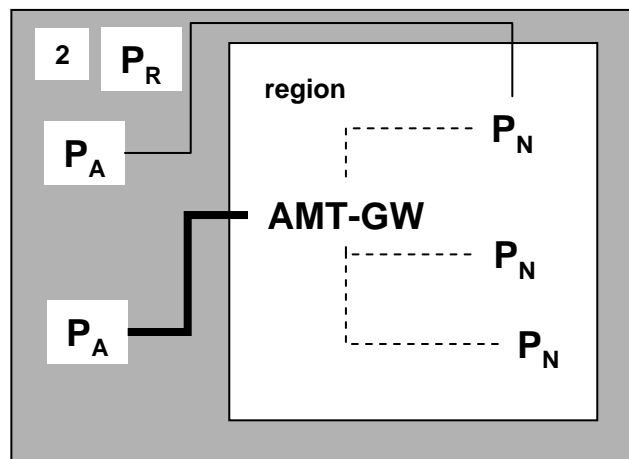
joinWithNativeLink

- `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

Membership Ordering Scenarios (NM first)

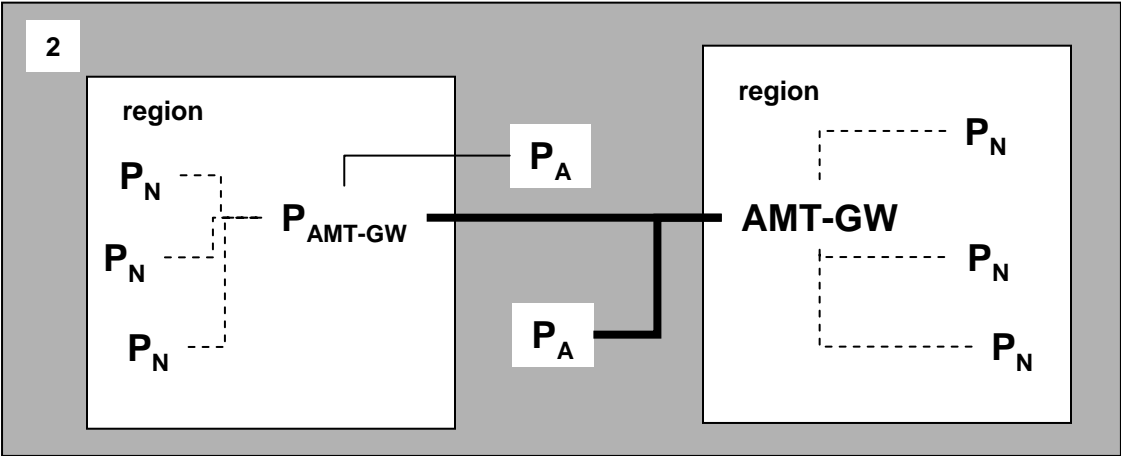
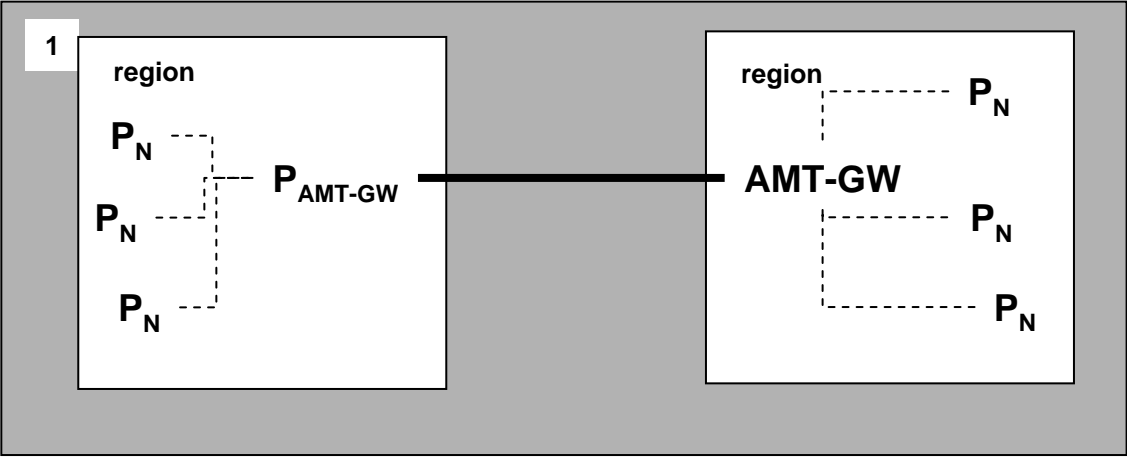


- (1) Peers form NM tree in single NM region
 - Since session is intended for intra-region use, AMT-GW is notified/configured
 - Group ID is created in overlay using some well-known session key
 - (a) Peer at Group ID is the root
 - Each P_N is a child of the root P_R in the ALM tree, used only for control
 - (b) Alternately could select one of the P_N as the root P_R and treat the session key as an indirect index to find this root

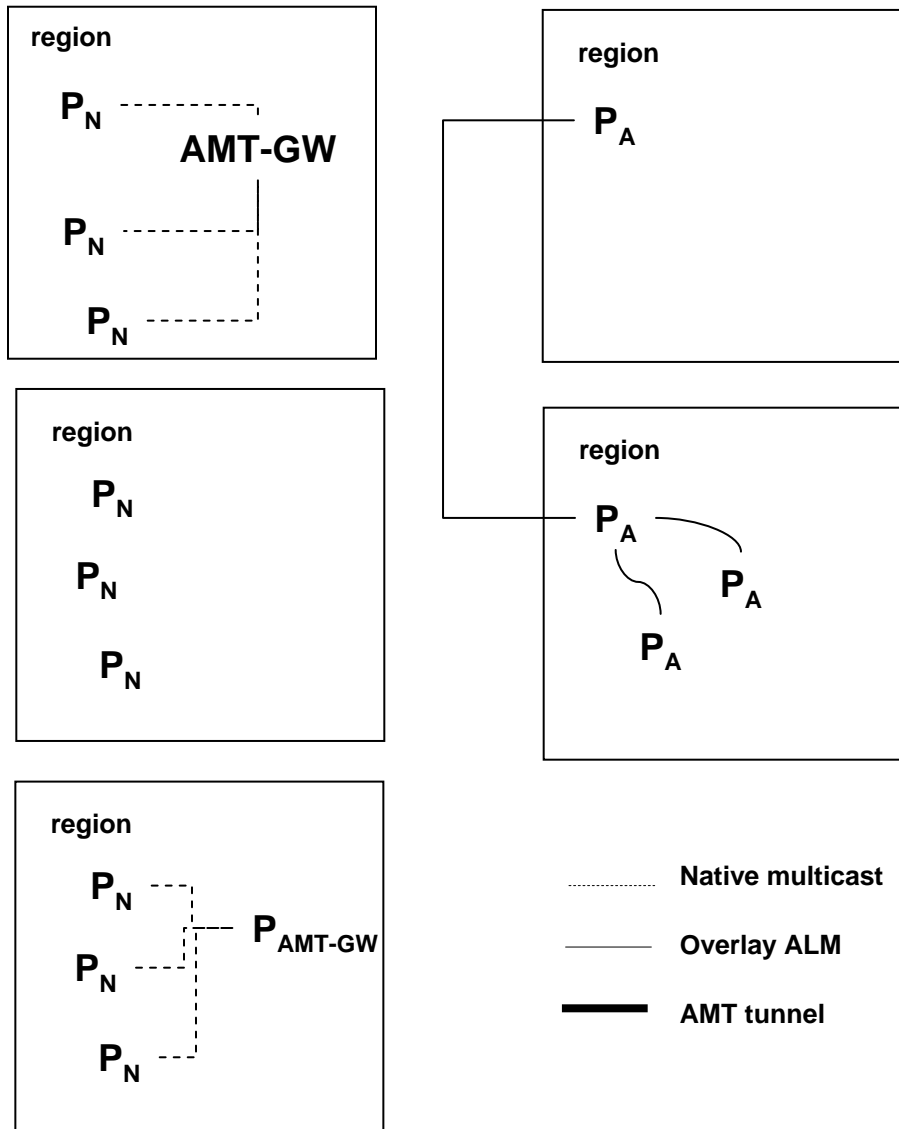


- (2) P_A joins tree
 - Determine Group ID using session key
 - Send join request to P_R, propagates join message to its P_N children which send response to P_A including AMT-GW attachment info (IP address, multicast IP address). P_A selects “best” P_N parent of positive responses.
 - (a) P_A supports AMT tunneling, creates tunnel to AMT-GW using info provided by parent P_N
 - (b) P_A doesn't support AMT tunneling, creates ALM connection directly with selected P_N parent

Membership Ordering Scenarios (NM first)



Membership Ordering Scenarios (ALM first)



- Note ALM topology may not follow in-region links
 - Depends on overlay topology organization
- 1) P_N joins via **AMT-GW**
 - Needs to find P -**AMT-GW** to tunnel through
- 2) P_N joins but no **AMT-GW** in region
 - Has to join as P_A
- 3) P_N joins via P -**AMT-GW**
 - P -**AMT-GW** will have ALM link to P_A root

Open Issues

- If a large tree has two or more independent AMT operated regions, do these AMT regions use the same NM address for the session or different addresses?
 - If there is a requirement for the same, need some way to coordinate the assignment of the NM group address
- Region partitioning scenarios
 - Two regions that were connected via AMT tunnel might be separated due to failure of the tunnel or need of greater capacity due to media change mid-session
 - AMT GWs still operate and now tunnel to separate parents in the ALM tree
- Region merging scenarios
 - Two different AMT-GWs are connecting during tree formation
 - Due to tree-reorganization such that two different AMT regions now have same parent peer

Open Issues: AMT GW as Bottleneck

- Single AMT-GW in region could become bottleneck
 - Should be addressed by AMT-spec
 - Provisioning policy could control this
 - More GWs, larger GW capacity
- $P_{\text{AMT-GW}}$ can be viewed as a P2P media relay
 - Avoid bottleneck by allowing any superpeer to be $P_{\text{AMT-GW}}$

Open Issues (From Charter & Requirements)

- Should the framework be targeted as an informational RFC from the SAM RG?
 - Should the requirements be rolled in to this document? (Requirements document is currently expired)
 - Do we need experimental validation first?
- Other drafts prepared or planned
 - Experimental plan for evaluating SAM framework
 - SAM overlay protocol
 - Specification of a P-AMT-GW node and P-NM node