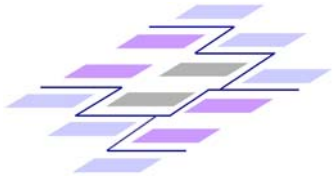


Hybrid Overlay Multicast Framework
draft-irtf-sam-hybrid-overlay-framework-01.txt

John Buford, Avaya Labs Research

Topics

- SAM Charter Recap and Problem Statement
- AMT(Automatic Multicast Tunneling) Overview
- Definitions
- Hybrid Framework
- Scribe-style Overlay Multicast in Hybrid Framework
- Future Work & Open Issues

SAM Charter Recap

- "The expected findings of the RG include characterizing the problem space, including driving scenarios, comparisons and analysis of existing approaches, a SAM framework that supports multiple ALM/OM/ native/hybrid protocols, analysis of network infrastructure impact when multicast traffic becomes a dominant flow in a network, and deployment scenarios which are independent of but can support and evolve with network infrastructure support for native multicast. The findings are expected to be published in technical reports, academic papers, and/or RFCs "
  - Problem statement: draft-irtf-sam-problem-statement-01.txt
  - Comparisons: a survey ID is being drafted
    - See: H. Yu, J. Buford. Advanced Topics in Peer-to-Peer Overlay Multicast. in Encyclopedia of Wireless and Mobile Communications (Ed. B. Fuhrt) CRC Press. To appear
  - SAM Framework: draft-irtf-sam-hybrid-overlay-framework-01.txt
  - Infrastructure analysis: TBD
  - Deployment scenarios: TBD

Hybrid Framework Problem Statement

- See draft-irtf-sam-problem-statement-01.txt
- “Hybrid”
  - Connects native multicast regions and non-multicast (or application layer multicast) regions
- “Framework”
  - allows different overlay algorithms, different ALM control algorithms, and different native multicast protocols to be used

Basic Idea

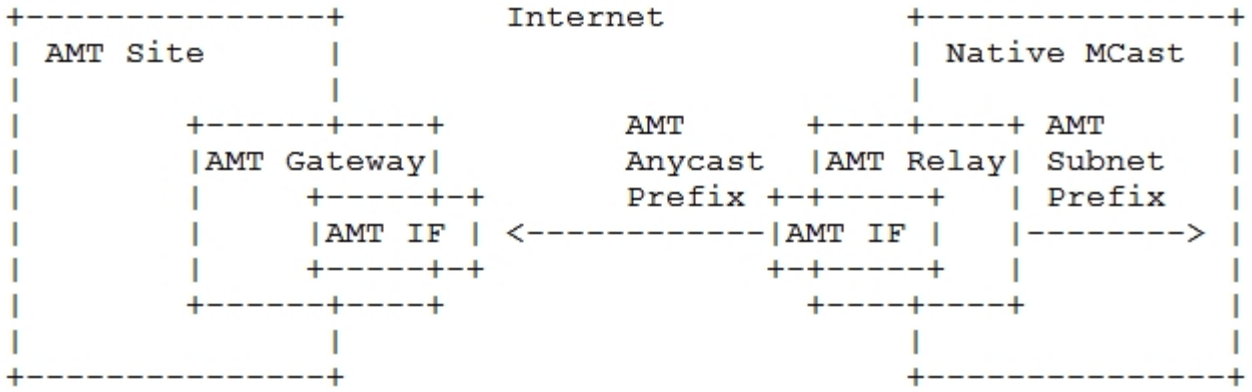
- Assume many peers connect in large overlay
- Some peers have native multicast support in their region, others do not
- There are many multicast sessions running concurrently among the peers, with different group sizes and membership
  - Use ALM/OM for peers which don’t have native multicast support
  - Use AMT relaying and gateways for peers which do have multicast support
  - Use overlay for group membership management and ALM signaling
- Different cases to consider include:
  - Sessions with no or all native mcast support
  - Small groups where most peers are either nmcast capable or not

Review of
Automatic Multicast Tunneling (AMT)
draft-ietf-mboned-auto-multicast-07

Automatic Multicast Tunneling (AMT)
draft-ietf-mboned-auto-multicast-07

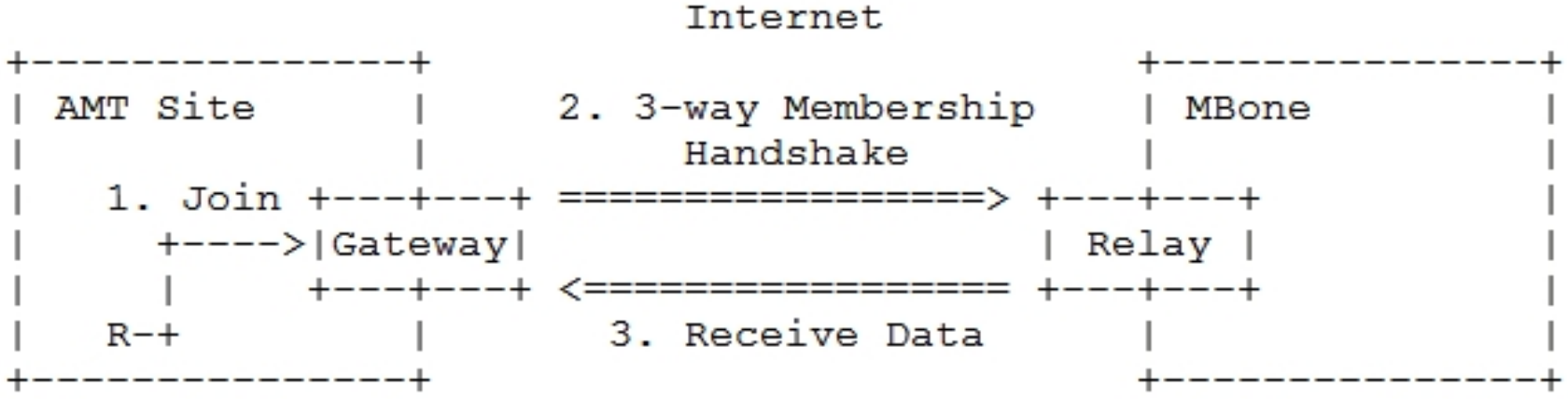
- allows multicast communication among isolated multicast-enabled sites or hosts, attached to a network which has no native multicast support.
- enables them to exchange multicast traffic with the native multicast infrastructure
- no manual configuration required
- uses an encapsulation interface so that no changes to a host stack or applications are required, all protocols (not just UDP) are handled, and there is no additional overhead in core routers.

AMT
draft-ietf-mboned-auto-multicast-07



- AMT Site
  - A multicast-enabled network not connected to the multicast backbone and served by an AMT Gateway. It could also be a stand-alone AMT Gateway.
- AMT GW
  - A host, or a site GW router, supporting an AMT MCAST⇔UDP encapsulation
    - Does not have native multicast connectivity to the native multicast backbone infrastructure.
- AMT Relay Router
  - A multicast router configured to support transit routing between AMT Sites and the native multicast backbone infrastructure. The relay router has one or more interfaces connected to the native multicast infrastructure, zero or more interfaces connected to the non- multicast capable internetwork, and an AMT pseudo-interface.

Receiving Multicast in an AMT
draft-ietf-mboned-auto-multicast-07



- relays receive the traffic natively and unicast-encapsulate it to gateways
- gateways decapsulate the traffic and possibly forward it into the AMT site.
- Each gateway has an AMT pseudo-interface that serves as a default multicast route.
  - Requests to join a multicast session are sent to this interface and encapsulated to a particular relay reachable across the unicast-only infrastructure.
- Each relay has an AMT pseudo-interface
  - Multicast traffic sent on this interface is encapsulated to zero or more gateways that have joined to the relay.
- The AMT recipient-list is determined for each multicast session. This requires the relay to keep state for each gateway which has joined a particular group or (source, group) pair. Multicast packets from the native infrastructure behind the relay will be sent to each gateway which has requested them.

AMT Restrictions

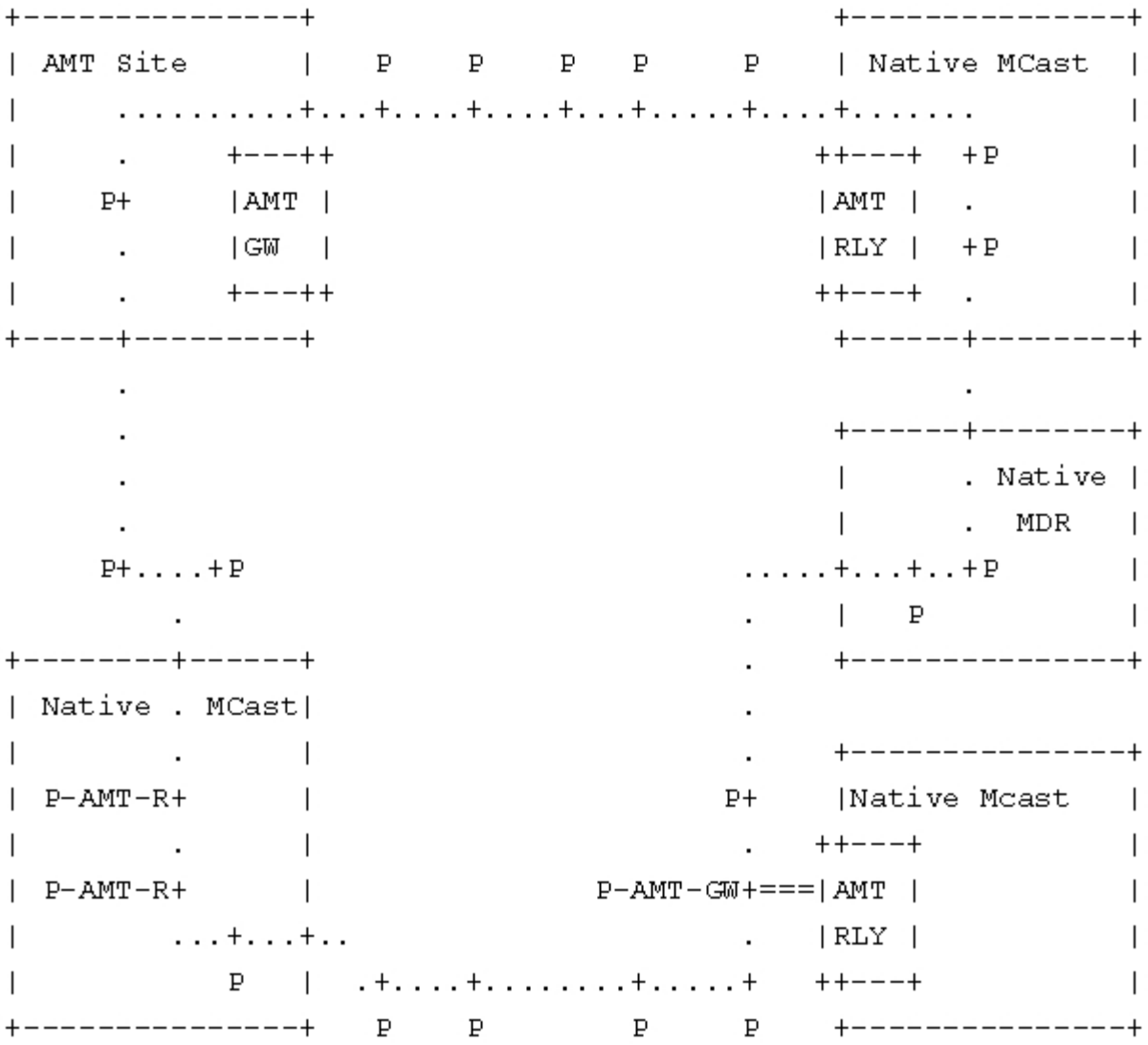
- isolated sites/hosts can receive SSM
- isolated non-NAT sites/hosts can send SSM
- isolated sites/hosts can receive general multicast.
- AMT does not permit isolated sites/hosts to send general multicast.

AMT Implementations

- Reported at IETF 66 in MBONED WG
  - FreeBSD Public Domain (Contact: Tom Pusateri)
  - Microsoft (Contact: Dave Thaler)
  - Cisco (Contact: Dino Farinacci)

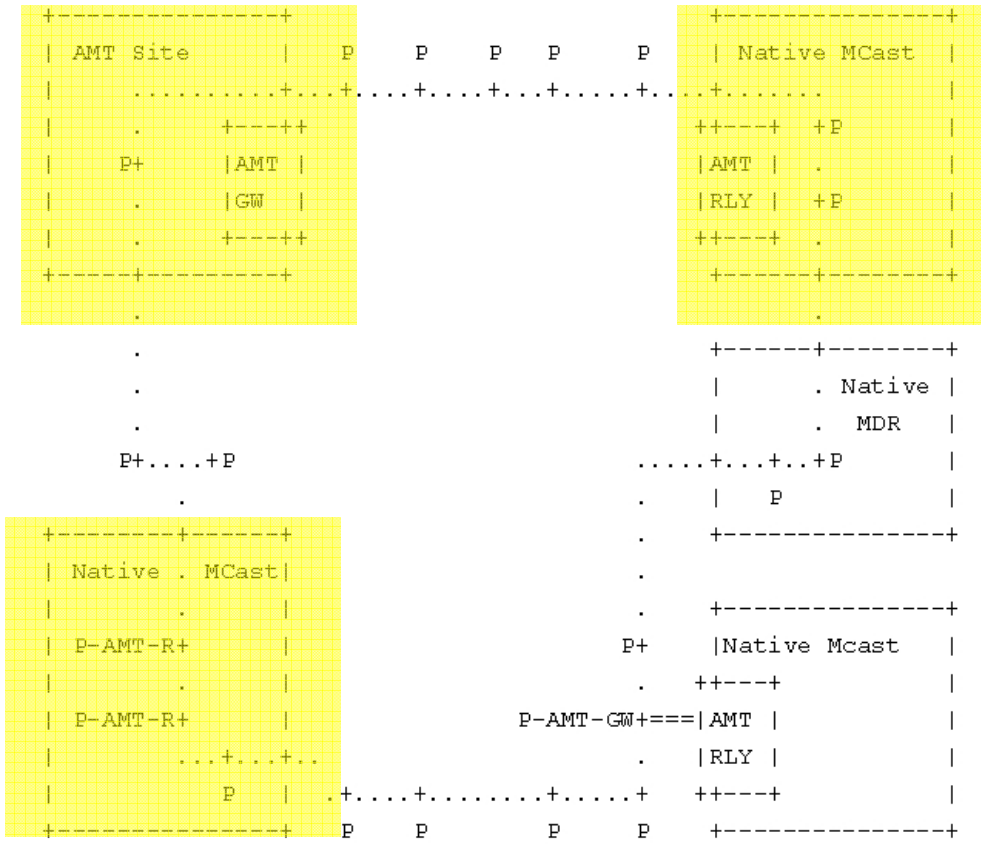
Hybrid Framework

draft-irtf-sam-hybrid-overlay-framework-01.txt



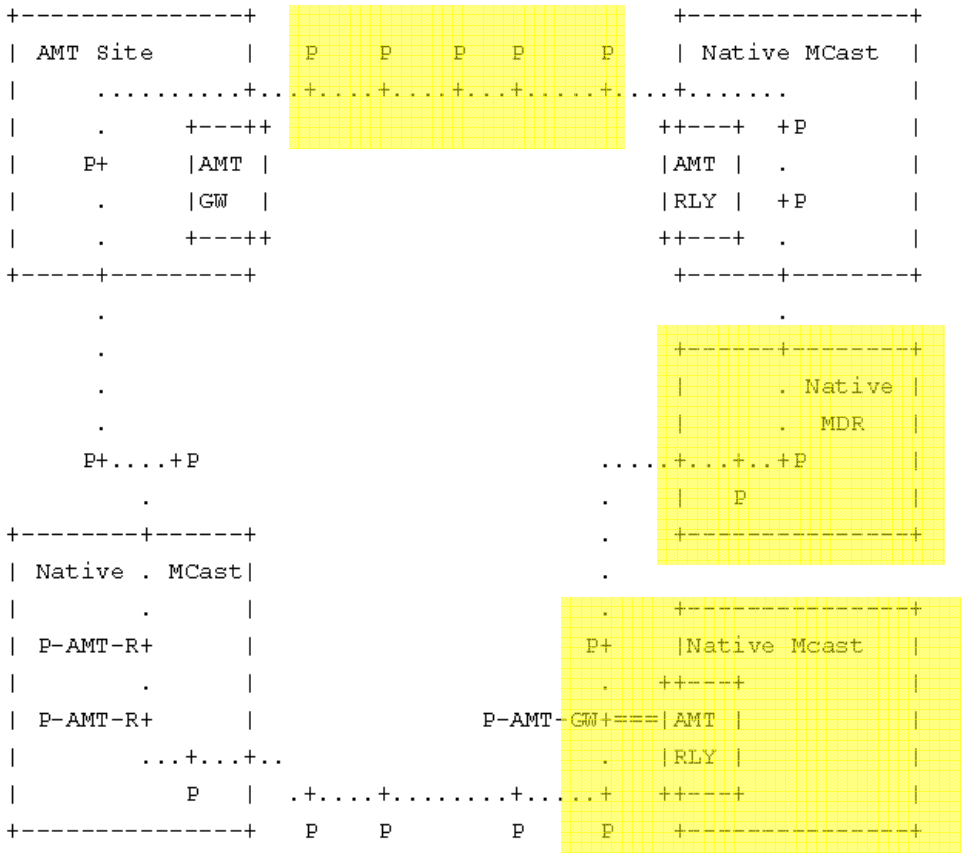
- Hybrid architecture in six regions of the network.
- All peers are connected in an overlay
- Lines are the predecessor/successor links between peers.
- Peers may have other connections in the overlay.

Regions



- Native multicast (NM) with a local AMT gateway (AMT GW). There are one or more peers (P) connected to the overlay in this region.
- Native multicast with a local AMT relay (AMT RLY). There are one or more peers (P) connected to the overlay in this region.
- Native multicast with one or more peers which emulate the AMT relay behavior (P-AMT-R) which also connect to the overlay. There may be other peers (P) which also connect to the overlay.

Regions



- No native multicast: Peers (P) in this region connect to the overlay
- Native MDR is a native multicast region using multi-destination routing, in which one or more peers reside in the region.
- Native multicast with no peers that connect to the overlay, but for which there is at least one peer in the unicast-only part of the network which can behave as an AMT-GW (P-AMT-GW) to connect to multicast sources through an AMT-R for that region.
  - It may be feasible to also allow non-peer hosts in such a region to participate as receivers of overlay multicast; for this version, we prefer to require all hosts to join the overlay as peers.

Definitions

- Overlay network
  - An application layer virtual or logical network in which end points are addressable and that provides connectivity, routing, and messaging between end points. Overlay networks are frequently used as a substrate for deploying new network services, or for providing a routing topology not available from the underlying physical network. Many peer-to-peer systems are overlay networks that run on top of the Internet.
- Overlay Multicast (OM):
  - Hosts participating in a multicast session form an overlay network and utilize unicast connections among pairs of hosts for data dissemination. The hosts in overlay multicast exclusively handle group management, routing, and tree construction, without any support from Internet routers. This is also commonly known as Application Layer Multicast (ALM) or End System Multicast (ESM).

Definitions

- Peer:
  - an autonomous end system that is connected to the physical network and participates in and contributes resources to overlay construction, routing and maintenance.
  - some peers may also perform additional roles such as connection relays, super nodes, NAT traversal, and data storage.
- Region
  - A region is a contiguous internetwork such that if native multicast is available, all routers and end systems can connect to native multicast groups available in that region.

Assumptions

- Peers connect in a large-scale overlay, which may be used for a variety of peer-to-peer applications in addition to multicast sessions.
- Peers may assume additional roles in the overlay beyond participation in the overlay and in multicast trees.
- We assume a single structured overlay routing algorithm is used. Any of a variety of multi-hop, one-hop, or variable-hop overlay algorithms could be used.
- Peers may be distributed throughout the network, in regions where native multicast (NM) is available as well as regions where it is not available.
- AMT-R and AMT-GW can be implemented in peers.
- Peers are able to determine, through configuration or discovery:
  - Can they connect to a NM router
  - Is an AMT gateway accessible
  - Can the peer support the AMT-GW functionality locally
  - Is MDR supported in the region

Assumptions

- Overlay Multicast
  - The overlay supports concurrent multiple multicast trees.
  - Different OM protocols may co-exist in same overlay
    - E.g. centralized for small groups, decentralized for large groups
  - The limit on number of concurrent trees depends on peer and network resources and is not an intrinsic property of the overlay.
  - Some multicast trees will contain peers use ALM only, i.e., the peers do not have NM connectivity. Some multicast trees will contain peers with a combination of ALM and NM.
  - Although the overlay could be used to form trees of NM-only peers, if such peers are all in the same region we expect native mechanisms to be used for such tree construction, and if such peers are in different regions we expect AMT to handle most cases of interest.

Scenarios

- ALM-Only Tree - Algorithm 1
- ALM tree with peer at AMT site (AMT-GW)
- ALM tree with NM peer using AMT-R
- ALM tree with NM peer with P-AMT-R

ALM-only Tree (simple algorithm)

(doesn't consider per node load, admission control, or alternative paths)

Peer which initiates multicast group:

groupID = create(); // allocate a unique groupId

// the root is the nearest peer in the overlay

// out of band advertisement/distribution of groupID,

// perhaps by publishing in DHT

Any joining peer:

// out of band discovery of groupID, perhaps by lookup in DHT

joinTree(groupID); // sends "join groupID" message

The overlay routes the join request using the overlay routing mechanism toward the peer with the nearest id to the groupID. This peer is the root. Peers on the path to the root join the tree as forwarding points.

Leave Tree:

leaveTree(groupID) // removes this node from the tree

Propagates a leave message to each child node and to the parent node. If the parent node is a forwarding node and this is its last child, then it propagates a leave message to its parent. A child node receiving a leave message from a parent sends a join message to the groupID.

Message forwarding:

multicastMsg(groupID, msg);

ALM tree with peer at AMT site (AMT-GW)

- The joining peer connects to the tree using the ALM protocol, or, if the tree includes a peer in an NM region, then the peer can use the AMT GW to connect to the NM peer through the AMT relay.
  - The peer can choose the delivery path based on latency and throughput.
- If the peer is not a joining peer and is on the overlay path of a join request:
  - If its next hop is a peer in an NM region with AMT-R, then it can select either overlay routed multicast messages or AMT delivered multicast messages.
  - If its next hop is a peer outside of an NM region, then it could use either ALM only or use AMT delivery as an alternate path

ALM tree with NM peer using AMT-R

- There are these cases:
- There is no peer in the tree which has an AMT-GW
  - The NM peer uses ALM routing
- There is at least one peer in the tree which can function as P-AMT-GW
  - The NM peer can join the tree using ALM routing and/or connecting to the P-AMT-GW.
- There is at least one peer in the tree which is in an AMT-GW region
  - The NM peer can join the tree using ALM routing and/or connecting to the AMT-GW.

ALM tree with NM peer with P-AMT-R

- Either the NM peer supports P-AMT-R or another peer in the multcast tree in the same region is P-AMT-R capable.
- The three cases earlier apply here, replacing AMT-R with P-AMT-R.

Future Work

- The next version of this document will elaborate:
  - ALM tree topology vs NM topology and NM-ALM edges
  - Single NM-ALM edge nodes vs multi NM peers from same region in the tree
  - Initial tree membership is ALM vs initial tree membership is NM

Future Work

- Simulation
  - Requires P2P overlay simulator with native layer model which supports multicast protocols
    - E.g.: SSFNet, Oversim
  - See “Tools for Peer-to-Peer Network Simulation” draft-irtf-p2prg-core-simulators-00.txt
- Implementation
  - XCAST on PlanetLab for MDR Region and overlay enhancement
  - Modify overlayweaver to support P-AMT-R and P-AMT-GW
  - Deploy OW on PlanetLab
- Federated and hierarchical overlay designs are deferred to future version of the ID

Questions / Discussion