AMT — Automatic IP Multicast without explicit Tunnels

Thomas Kernen and Steve Simlo
Cisco Systems

According to research by Cisco's *Visual Networking Index*, IP traffic will grow by a factor of four between 2009 and 2014 – rising to nearly 64 Exabytes per month, of which 91% will be video (TV, VoD, Internet video and P2P).

One technique which helps to manage this growth is *multicast*. However, today, most static multicast tunnelling uses an encapsulation which results in all multicast traffic having the same source and destination IP address and protocol number, with no Layer 4 information. This forces all traffic to be categorized into the same class by the transit routers and this in turn means that it is impossible to differentiate between streams for the purpose of load balancing or prioritisation.

In contrast, AMT uses UDP encapsulation to provide different source UDP ports for the encapsulated traffic, allowing transit routers to perform flow-based load balancing for more efficient link utilization. This has benefits for broadcasters and content owners, enabling access to an infrastructure with minimal bandwidth requirements per stream and affording an opportunity to further improve the quality of the streams that are delivered.

The problems with realising a multicast service are multidimensional, involving several different players:

- **Users** simply want access to content and do not care whether that content is delivered over a unicast or a multicast infrastructure.
- **Internet Service Providers** seek a way to be compensated for delivering extra services. There has typically been no way of easily monetising these services. In the absence of an easy way to make multicast into a commercial service, there has been slow take-up of the technology from this market sector.
- **Content Providers** love multicast because in effect they pay less money for delivering more, or the same amount of, content.
- **Application Developers** are typically challenged with the new requirements posed by multicast applications. They have to develop an application interface to address the control plane needed in a multicast receiver that is not identical to that found in a unicast receiver. In general this is not something that is well known in Application Development circles even though it is not a hugely complex requirement.

Fig. 1 demonstrates these dynamics in high-level terms:

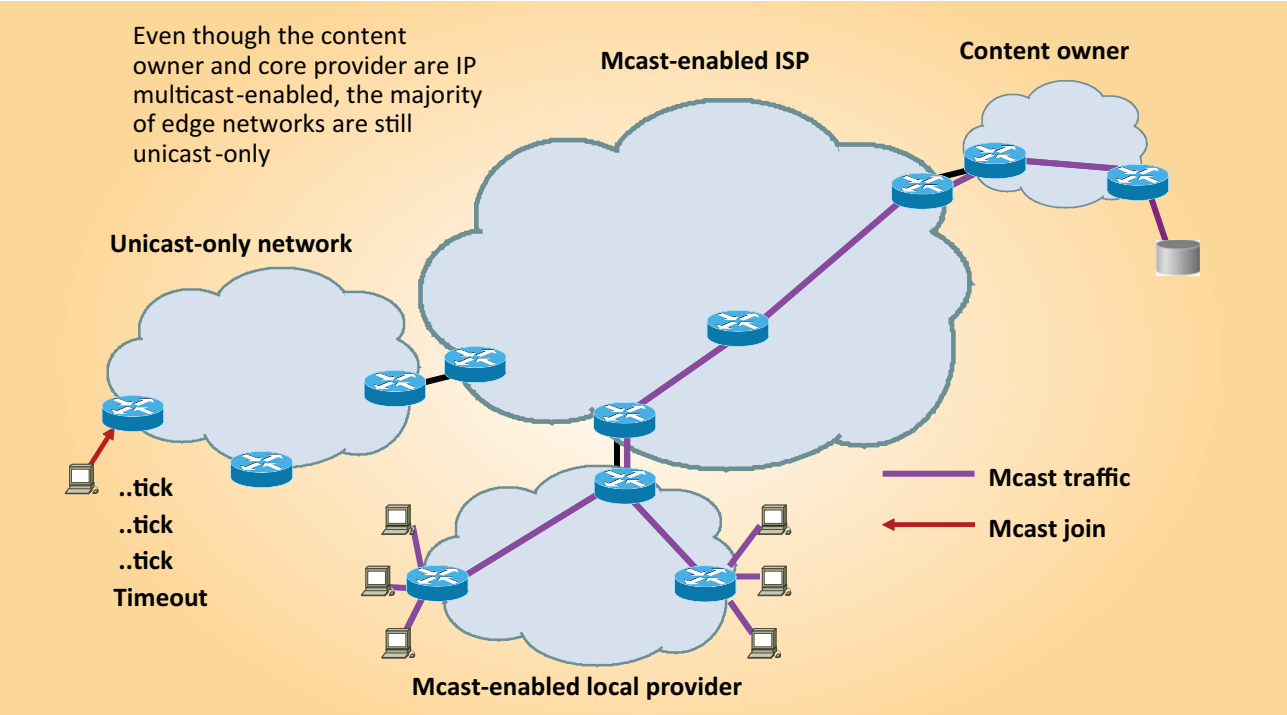

Figure 1
Content delivery in today's IP environment

Multicast delivery of streamed content

The dynamics of content delivery over the Internet have skewed towards the widescale use of unicast. Even a content owner with attachment to a native multicast-enabled ISP is forced to offer both unicast and multicast content. This has occurred since receivers are typically attached to unicast-only last-mile network environments. This results in high bandwidth costs per stream for the content owner and all service providers in the delivery chain *(see Fig. 2)*.

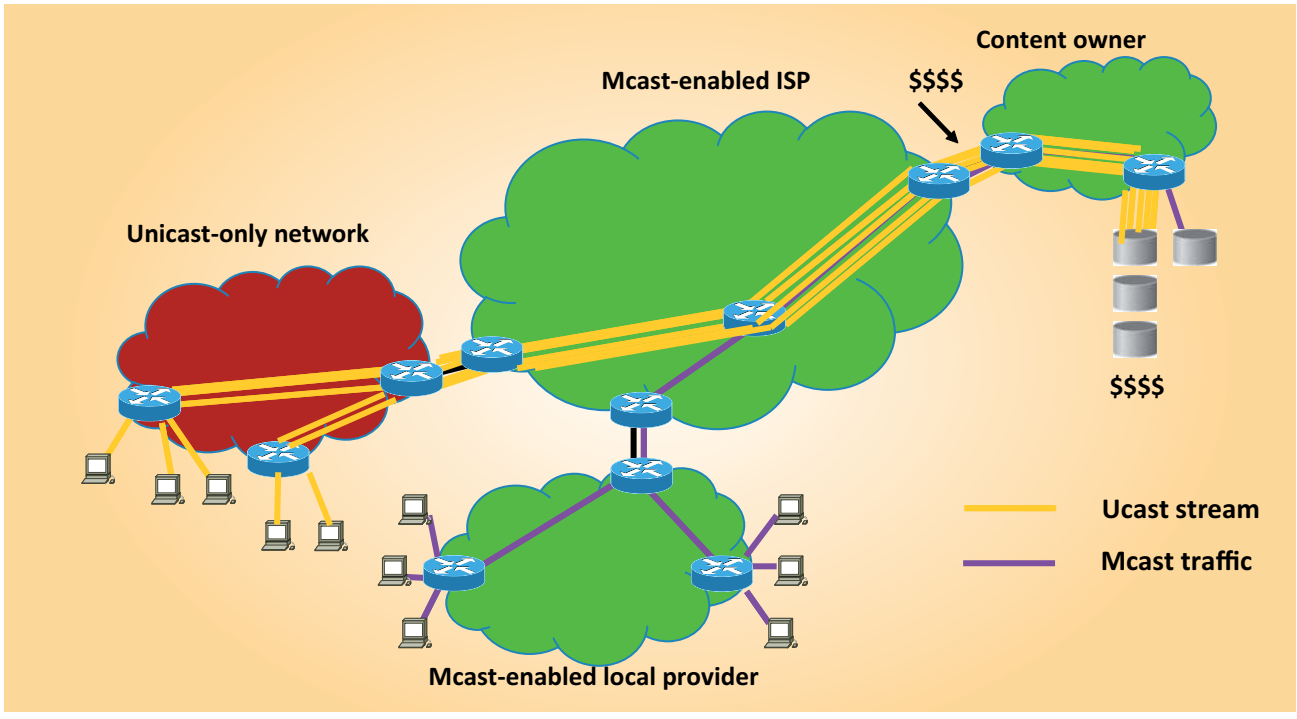

Figure 2
Unicast-only last-mile environment

When hosts request multicast content, they send a message into the network to indicate that they are interested in a specific multicast group. This message is called an IGMP membership report. It is sent to the first hop router adjacent to the receiver. In a unicast-only network, since the network has no multicast capability the router simply discards the IGMP message.

Thus, the receiver application in the host is forced to send unicast requests for the same content.

The consequence of this behaviour was:

- Content owners generated more revenue by providing unicast content because they reached more “eyeballs”.
- The ISP made more money by allowing the content owner to transit multiple copies of the stream.
- The unicast-only network was likely to be either a Tier 1 provider (that didn’t have to pay for transiting traffic to other providers) or a local access provider that did not want to provision multicast.

The following resulted:

1) People grew accustomed to unicast video delivery (through the likes of YouTube and BBC iPlayer).
2) Service providers no longer felt the need to deploy multicast (from a lack of business justification).
3) Further proliferation of unicast delivery models occurred.
4) Multicast became more and more niche.

Further complications resulted from the complexity of maintaining the control plane.

In inter-networks we distinguish between the component that handles packet forwarding – the data plane; and the component that handles the routing and signalling elements – the control plane. The control plane protocol that was used to build the forwarding state was based on the (so-called) “Any Source” Multicast (ASM) model. In this case what was required was the deployment of a special service known as a Rendezvous Point (RP), which acted to connect receivers with active senders. In cases where connections spanned multiple provider networks then it was also required to interconnect these RPs with a further protocol that advertised active senders (known as Multicast Source Discovery Protocol – MSDP). An improvement is available in Source Specific Multicast (SSM) but requires the receiver to be compatible with a new variant of IGMP (IGMPv3) which is not supported in many hosts and applications.

It should be observed that (for example within the UK) there are multicast services available that have been running for sometime. As a result, a small part of the overall Internet audience has, in theory, had access to the services via multicast. For an example see:

http://www.bbc.co.uk/multicast/

All of these factors, rooted in technology (but ultimately dictated by economics), led to a lack of multicast deployment. For these reasons **AMT (Automatic IP Multicast without explicit Tunnels)** has been developed.

Automatic IP Multicast without explicit Tunnels is specified in an Internet draft:

http://datatracker.ietf.org/doc/draft-ietf-mboned-auto-multicast/

This specification is designed to provide a migration path to a fully multicast-enabled backbone and allows multicast to reach unicast-only receivers without the need for any explicit tunnels between the receiver and the source. It is designed to provide the benefits of multicast where multicast is deployed.

AMT provides a hybrid solution where multicast networks get the benefit of multicast and work seamlessly with existing applications. It requires only a client-side shim (somewhere on the client) and router support (in some deployment scenarios). It provides a way of reliably and simply connecting together multicast islands and easing the transition to ubiquitous multicast deployment without requiring wide-scale changes.

AMT allows multicast communication amongst mutually separate multicast-enabled networks/ speakers. It allows them to exchange multicast traffic with senders and receivers in a native multicast environment and requires no manual tunnel configuration. AMT uses an encapsulation interface so that no changes to a host stack or applications are required, all protocols are handled, and there is no additional overhead in core routers. No explicit tunnels are needed, in contrast to existing models which all require static tunnels to be configured.

Other solutions have existed to this same problem for some time, but what makes AMT unique is the way it dynamically establishes the tunnels.

In the longer term, AMT could make multicast available on an Internet-wide basis. If vendors implement AMT functionality in their devices, (for instance as a proxy in receiver applications, home gateways and STBs), then many hosts will be able to reach a much wider variety of multicast content in parallel to the existing unicast-only content.

AMT components

The following terminology is largely adapted from **draft-ietf-mboned-auto-multicast-10**:

AMT Site

An **AMT site** is a multicast network (or host) with an attached / resident gateway served by an AMT Gateway. It could also be a standalone AMT Gateway.

AMT Relay

An **AMT Relay** is typically 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 inter-network, and an AMT pseudo-interface. This device terminates one end of an AMT tunnel and encapsulates multicast packets into those tunnels. While usually a router, it may be a standalone server.

Put more simply, an AMT Relay receives AMT Requests from an AMT Gateway.

Abbreviations

AMT	Automatic IP Multicast without explicit Tunnels	**MAC**	Media Access Control
ASM	"Any Source" Multicast	**MSDP**	Multicast Source Discovery Protocol
IANA	Internet Assigned Numbers Authority http://www.iana.org/	**P2P**	Peer-to-Peer
IGMP	Internet Group Management Protocol	**PIM**	Protocol-Independent Multicast
IP	Internet Protocol	**QQIC**	Querier's Query Interval Code
ISC	Internet Systems Consortium http://www.isc.org/	**RP**	Rendezvous Point
ISP	Internet Service Provider	**SSM**	Source-Specific Multicast
		STB	Set-Top Box
		UDP	User Datagram Protocol
		VoD	Video-on-Demand

AMT Gateway

An **AMT Gateway** is a host, or site gateway router, supporting an AMT Pseudo-Interface. It does not have native multicast connectivity to the multicast backbone infrastructure. This device terminates the other end of an AMT tunnel and de-encapsulates multicast packets from those tunnels.

Put more simply, an AMT Gateway sends AMT Requests to the AMT Relay.

AMT Gateways are expected to be implemented in two ways:

- In a network device (home gateway, router);
- In a host (standalone software or built into an application).

AMT Pseudo-Interface

AMT encapsulation (of multicast packets inside unicast packets) occurs at a point that is logically equivalent to an interface, with the link layer being the unicast-only network. This point is referred to as a **pseudo-interface**. Some implementations may treat it exactly like any other interface and others may treat it like a tunnel end-point. In most (if not all) AMT implementations, the pseudo-interface will be a tunnel end-point.

The distinction between AMT Relay and Gateway is subtle, as it relies on the assumption that there is a static multicast backbone to which some providers are directly connected (whether physically or through a static tunnel) and others are not connected. In many cases, this distinction will be straightforward, but in others, it may be relative to the traffic flow. In other words, think of a relay as having a contiguous multicast-enabled path to the multicast source, whereas the gateway's path to the source by necessity traverses a non-multicast-enabled network.

Using these definitions, we can better understand the picture when we introduce AMT. Initially, let's assume that the multicast-enabled ISP provides the AMT Relay service *(Fig. 3)*.

In this diagram, we see hosts connected to the unicast-only network acting as AMT Gateways. When these hosts want to see content, they try to send an IGMP membership report to the first hop

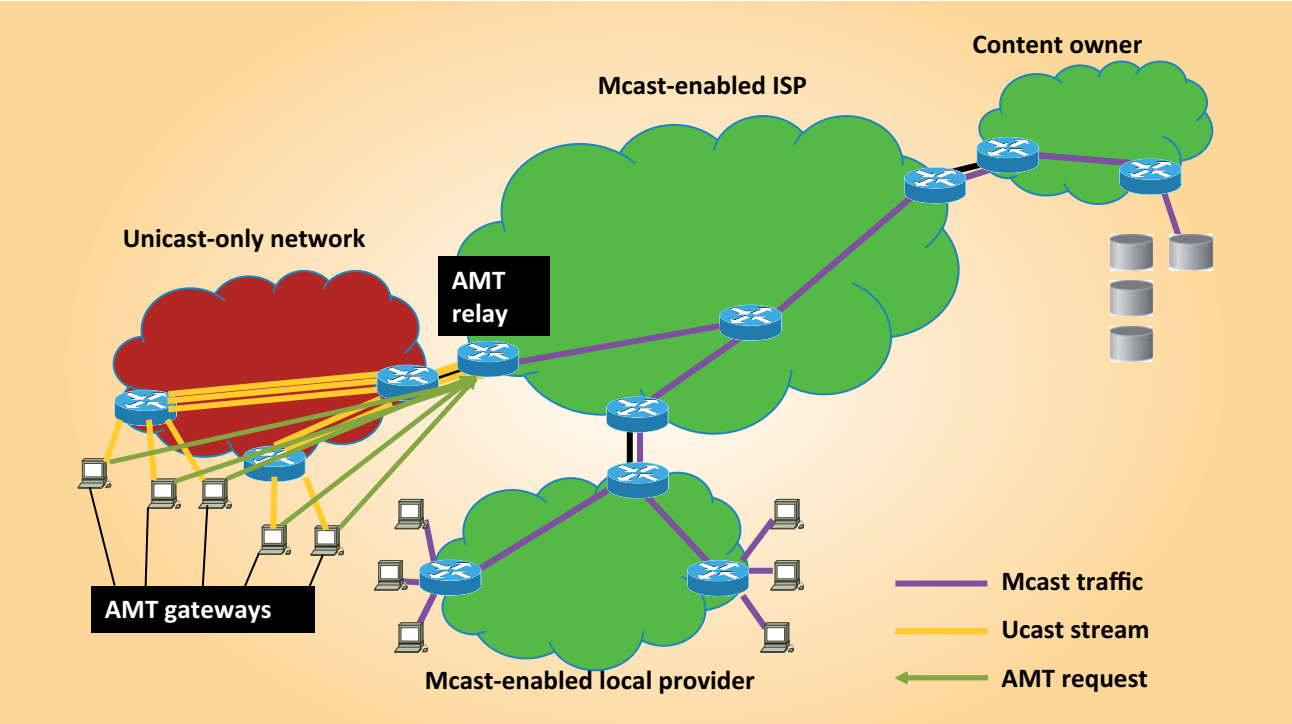

Figure 3
Multicast-enabled ISP providing AMT relay service

router. The AMT Gateway process running on the host will intercept this message and this will then trigger an AMT request towards an AMT Relay. Once the AMT tunnel has been established (by way of a **"3-way handshake"** described in detail in the next section: *AMT Relay discovery*), the host will encapsulate the IGMP membership report into the AMT tunnel. The AMT Relay will then de-encapsulate the IGMP message, which will trigger an upstream PIM join towards the source. In case the AMT tunnel had already been established between the Gateway and the Relay, the IGMP message will be immediately encapsulated into the tunnel.

A clear migration path for the unicast-only network now exists for it to become multicast-enabled. It could start by moving the relay into its network domain and establishing a multicast peering with the upstream ISPs. Then, to further minimize the bandwidth load, it can gradually push multicast capabilities down through the network, into the first-hop routers, removing the need for the host-based AMT Gateways.

AMT Relay discovery

We now explain how the AMT Gateway finds the ISP's AMT Relay. We need an address that is recognized throughout the Internet. In an IP network, one way of providing this function is via an **Anycast Address**. Eventually it is expected there will be an **IANA** Anycast address allocated for the AMT Anycast prefix. Currently the prefix is provided by **ISC** (via 154.17.0.0/16). Each ISP with an AMT Relay needs to advertise this address as reachable throughout the Internet. Within a private network not connected to the Internet, it is possible to use a non-global IP address that is advertised within a particular organization or provider network.

The AMT Gateway sends a special message called an **AMT Relay Discovery messag**e to the AMT Anycast Address. Messages to that address are only responded to by AMT Relays. The gateway will thus rely on the routing table to find the closest Relay. Relays that are either overloaded or in some other way unreachable, will be expected to not advertise the prefix for that period of time. In this way some dynamic resiliency is provided to the AMT architecture.

The message is sent to the reserved UDP port 2268 and includes a special code (or Nonce), which is used to secure the setup of the tunnel *(Fig. 4)*.

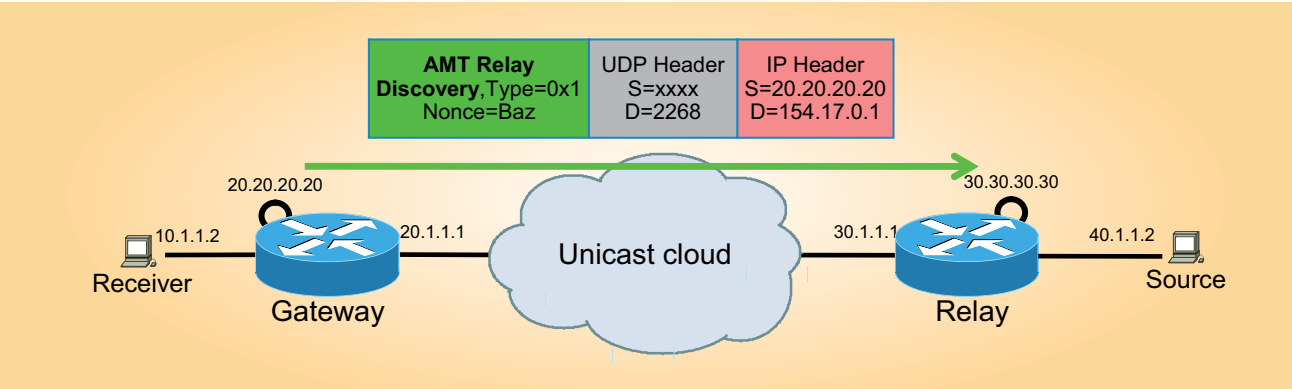

Figure 4
AMT Relay Discovery message

On receipt of an **AMT Relay Discovery message**, the Relay will respond to the Gateway with an **AMT Relay Advertisement message**, which includes the Relay's unique IP address. The Gateway will then use that unique IP address as the destination of any AMT messages (**AMT Requests and AMT Membership Updates**) sent to this specific Relay.

Once again the reserved UDP port 2268 is used and the reply also contains the same Nonce that was originated by the Gateway. Thus the Gateway knows that this reply was a reply to its Discovery *(Fig. 5)*.

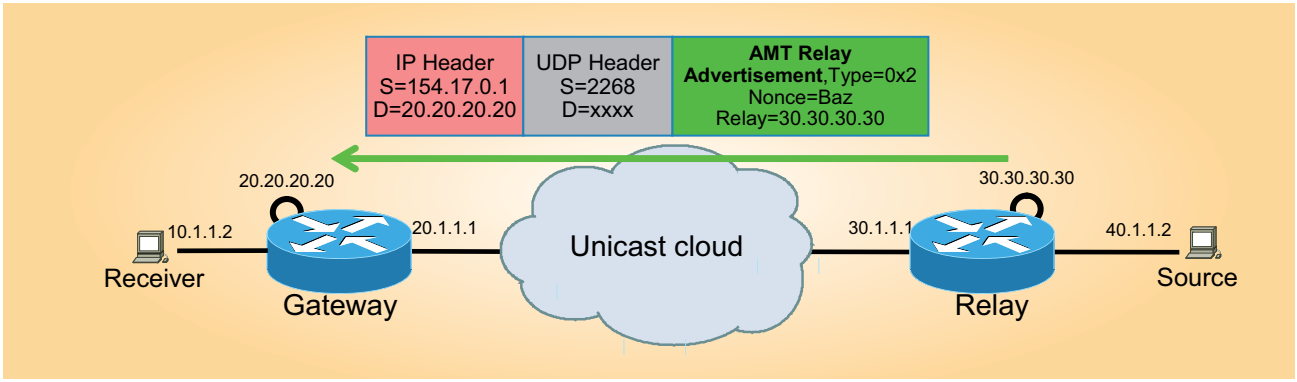

Figure 5
AMT Relay Advertisement message

On receipt of the **AMT Relay Advertisement Message**, the Gateway begins the "**3 way handshake"** by sending an **AMT Request message** to the Relay using the relay's unique IP address as the destination (again along with a new Nonce) *(Fig. 6).*

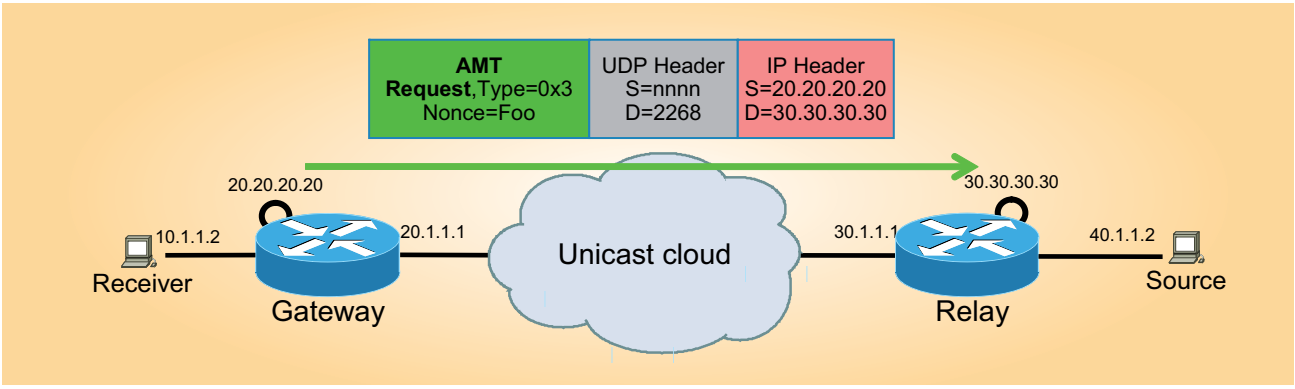

Figure 6
AMT Request message

The Relay responds with an **AMT Query** that includes the new Nonce from the AMT Request, as well as an opaque security code (MAC) that it will expect in any future messages from the Gateway. The AMT query in fact encapsulates the underlying IGMP membership query and includes the **Querier's Query Interval Code (QQIC)**, which specifies the Query Interval used by the querier *(Fig. 7)*.

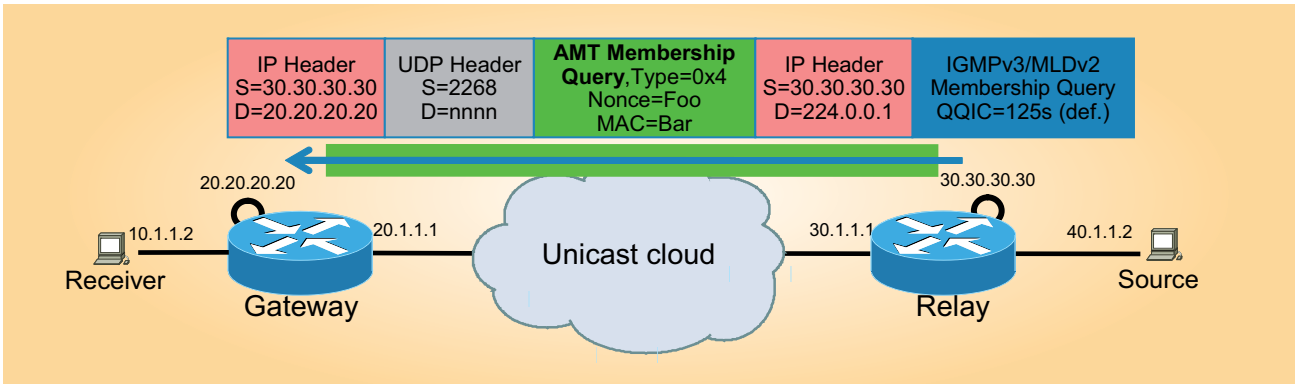

Figure 7
AMT Membership Query message

To join any upstream sources, the Gateway responds with an **AMT Membership Update** that includes the opaque security code, the original nonce from the AMT Request, and an encapsulated IGMPv3 packet *(Fig. 8)*.

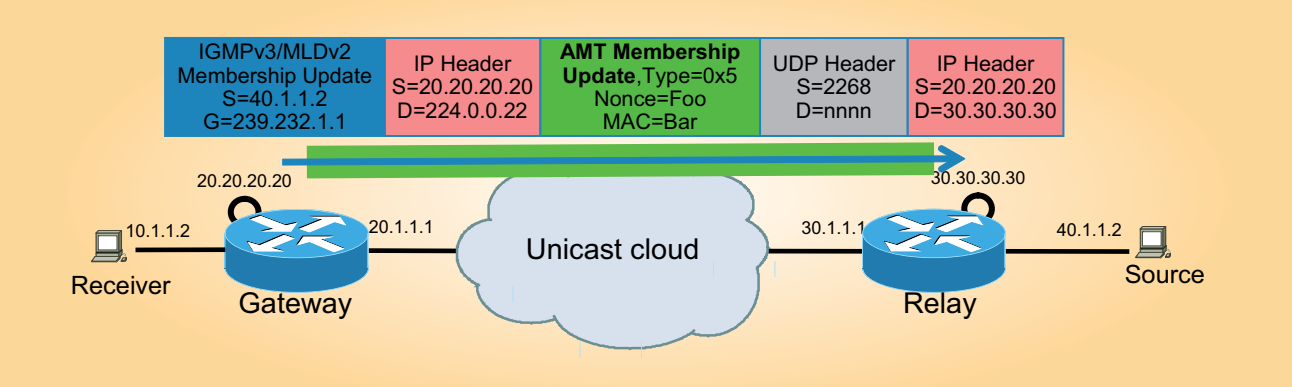

Figure 8
AMT Membership Update message

By validating the security code and Nonce, the Relay finalizes the tunnel setup and begins using it for multicast traffic. The Relay adds the appropriate pseudo/tunnel interface to the multicast route for that particular stream and begins replicating and encapsulating packets to the Gateways *(Fig. 9)*.

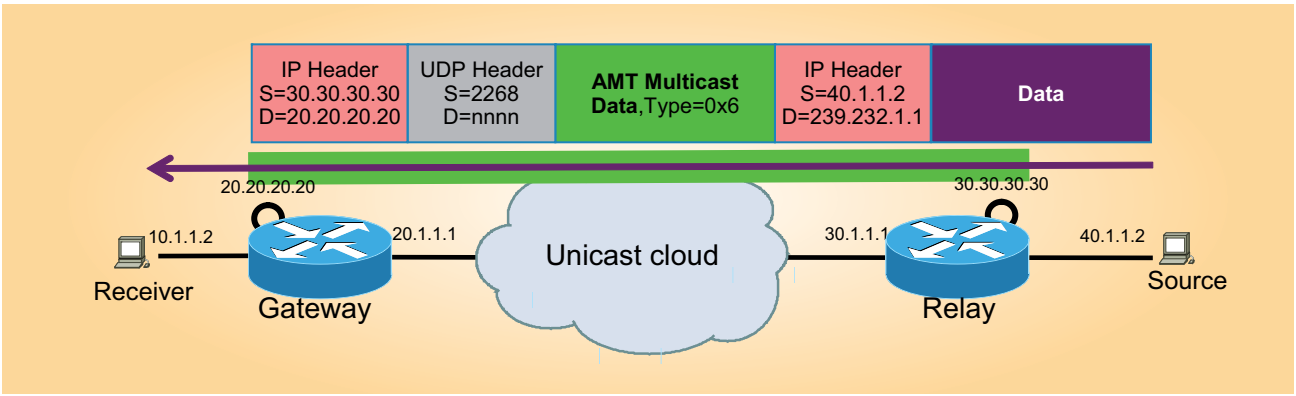

Figure 9
AMT Multicast Data transfer

Any further streams will use the same **Request/Query/Update "3-way handshake"** (but will not need to use the **Discovery/Advertisement** process since the tunnel will already have been established). If any Request does not receive a Query in response, the Gateway will then use the **Discovery/Advertisement** mechanism to find the next available Relay.

Once the tunnel has been established, the communication is effectively identical to a normal router-host IGMPv3 relationship. The Gateway (host) sends periodic AMT Membership Updates to refresh the state on the Relay (router), sending the appropriate update to leave the group when the traffic is no longer desired. Once the tunnel is no longer required by any more receivers it is maintained by the Gateway / Relay for a further time-out period. In that way a new receiver does not need to build a new tunnel if that receiver becomes active again shortly afterwards.

AMT deployment examples

Several deployment scenarios are enabled by AMT

The most obvious deployment scenario of direct relevance to IP broadcasters is that of residential broadband subscribers requiring access to dynamic multicast content. Using AMT Gateway software on a local PC and connecting to well-known AMT Relays (either within the broadband provider's network or outside), these users can receive multicast streams without requiring an expansive up-grade of their provider's network.

In this scenario no dedicated network device functions as a gateway. The receivers have an integrated gateway function, and establish a tunnel directly with the relay. In every receiver a gateway proxy function builds a separate tunnel *(Fig. 10)*.

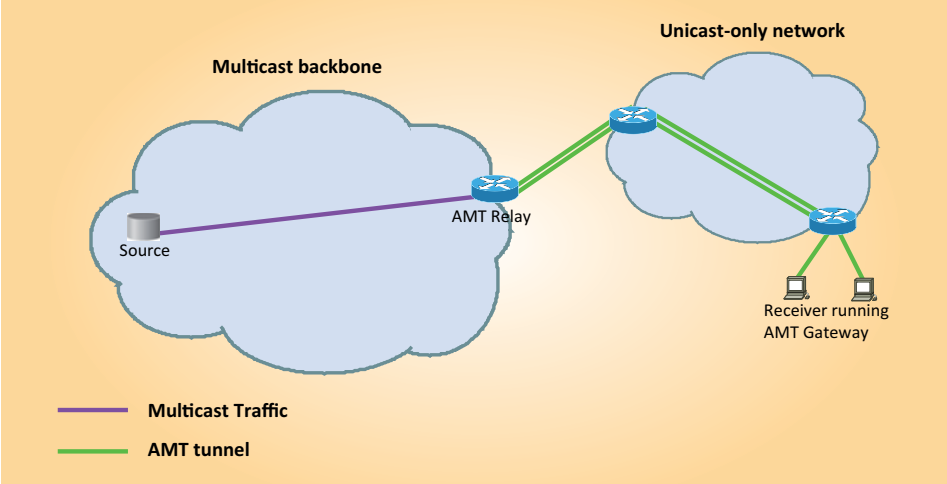

Figure 10
AMT deployment with gateway in receiver

A variant on this topology is when there is a network device functioning as the AMT Gateway. The AMT Gate-way has a directly connected receiver. When the receiver wants to join a multicast group, the Gateway receives the IGMP report and initiates the tunnel establishment process, and maintains it for as long as it is necessary. This solution requires a multicast capable receiver and a Gateway on the receiver side, and needs no multicast support in the network infrastructure. One tunnel is built for every gateway with attached active receivers *(Fig. 11)*.

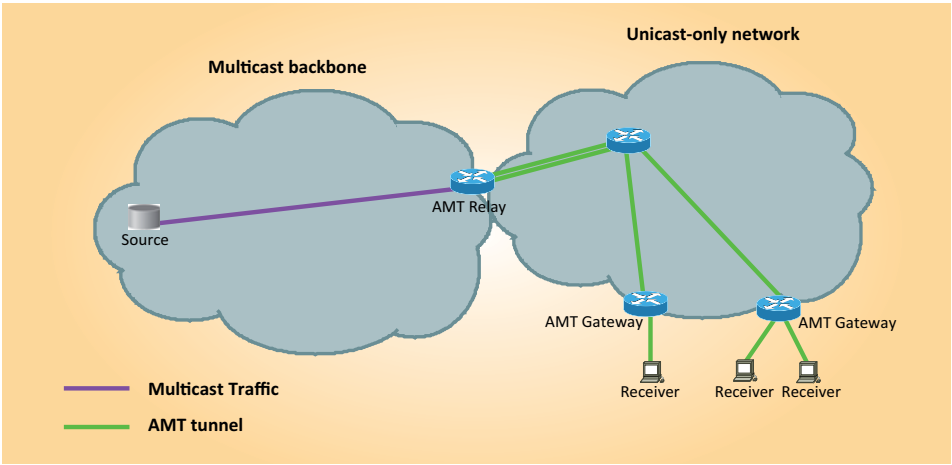

Figure 11
AMT deployment with separate gateway

AMT benefits

AMT offers a number of benefits to the IP broadcasting industry for delivering content. In particular: **Simplicity, Resiliency and Efficiency**:

Simplicity: Instead of the overhead of manually provisioning, establishing and maintaining GRE tunnels between two locations, the receiving network simply sends **AMT Advertisements** to a well-known **Anycast Prefix**. The rest of the tunnel establishment is done automatically without the need for additional configuration

Resiliency: Since the Relay discovery uses an **Anycast address**, Gateways will automatically find

Steve Simlo is a Consulting Engineer with Cisco Systems. He has worked on the design and deployment of IP multicast networks for most of the last 10 years. In that time he has been involved in the rollout of some large campus / enterprise networks and some significant service provider deployments as well.

Most recently Mr Simlo has been working on multicast VPNs with some large service providers and has been investigating IP multicast security. He is engaged with broadcasters looking to integrate multicast technologies into their contribution and distribution infrastructure. He is also a participant in Cisco efforts in the "Sustainable Technologies Space".

Before working for Cisco, Steve Simlo was involved with video, ATM, SNA, X.25 and high-speed modem networks in the 1980s and early 90s.

Thomas Kernen is a Consulting Engineer working for Cisco's European Innovation Consulting Engineering Team. His main area of focus is in defining video architectures and transmission solutions for content providers, broadcasters, telecom operators and IPTV service providers.

Prior to joining Cisco, Mr Kernen spent ten years with different telecoms operators, including three years with an FTTH triple play operator, for whom he developed their IPTV architecture. He is a member of the IEEE, SMPTE and is active in the TM-AVC group within the DVB Project.

the closest Relay. If that Relay should become unavailable or unreachable the routing table will reconverge on the next closest Relay.

Efficiency: today most static multicast tunnelling uses an encapsulation which results in all multicast traffic having the same source and destination IP address and protocol number, with no Layer 4 information. This forces all traffic to be categorized into the same class by transit routers and this in turn means that it is impossible to differentiate between streams for the purpose of load balancing or prioritisation. In contrast, AMT uses UDP encapsulation, providing different source UDP ports for the encapsulated traffic, allowing transit routers to perform flow-based load balancing for more efficient link utilization.

Conclusions

A global ubiquitous multicast service leads to the following:

- The content owner no longer pays for bandwidth from multiple identical streams;
- The multicast-enabled ISP no longer carries as much duplicate traffic;
- The multicast-enabled ISP controls the replication of the content within its network and in the AMT Relay;
- The unicast-only network absorbs the impact of the unicast streams.

By leveraging multicast infrastructures where available and interconnecting them using AMT, content can be distributed by content owners and broadcasters in the most efficient manner possible. To better handle the bandwidth growth generated by video-based services, service providers can therefore work on improving their support for multicast delivery in an incremental fashion. For broadcasters and content owners this enables access to an infrastructure with minimal bandwidth requirements per stream and affords an opportunity to further improve the quality of the streams that are delivered.

This version: 20 December 2010

Published by the European Broadcasting Union, Geneva, Switzerland

ISSN: 1609-1469

Editeur Responsable: Lieven Vermaele

Editor: Mike Meyer

E-mail: tech@ebu.ch

The responsibility for views expressed in this article rests solely with the authors