Universidad del País Vasco
Euskal Herriko Unibertsitatea

Departamento de Arquitectura y Tecnología de Computadores
Konputagailuen Arkitektura eta Teknologia Saila

eman ta zabal zazu

Universidad del País Vasco
Euskal Herriko Unibertsitatea

KONPUTAGAILUEN ARKITEKTURA ETA TEKNOLOGIA SAILA
DEPARTAMENTO DE ARQUITECTURA Y TECNOLOGÍA DE COMPUTADORES

PIM-SM extension for Source-Specific Multicast through non multicast networks

EHU-KAT-IK-04-15

Informe de Investigación
Research report

JM Rivadeneyra

July 2015

Table of contents

Abstract

Deployment of multicast in the open Internet is stagnated, mainly as a result of service provider policies and network limitations. To skip the lack of multicast connectivity between receivers and networks that carry traffic generated by multicast sources, the IETF has developed a proposal, called Automatic Multicast Tunnelling (AMT), supported in routers at least from 2011. Even so, it has not brought the necessary momentum to the expansion of multicast. In this report a similar but simpler than AMT proposal to skip the non-multicast gap is described. The basic idea in the proposal is to remove from multicast routing architecture some elements imposed by ASM model, those elements that are not needed for the SSM applications (e.g. Internet TV), but make multicast an 'all-or-nothing' technology.

1. Introduction

IP multicast was proposed 30 years ago [1] to efficiently support one-to-many and many-to-many communication. Multicast packets are delivered via multicast trees built by multicast-enabled routers. Nodes use IGMP [2] and MLD [3] protocols to become a receiver for a multicast group, reporting about it to a Designated Router (DR) in its local network. DR uses the PIM protocol to interact with other multicast enabled routers outside the local network, to attract traffic from multicast groups required by receivers in the local network. This is, PIM is used by routers to build the multicast distribution tree (MDT) [4].

Originally [5], IP multicast was defined under the Any Source Multicast (ASM) model, where a multicast address is the identification of a multicast group G. In the ASM model, senders can transmit data to G without any registration or authentication, and receivers would receive data from all senders towards G. ASM has brought serious security and management problems into multicast, mainly related to inter-domain multicast. Moreover, the main benefit provided by multicast appears in one-to-many applications that don't have an ASM nature. This is the case of Internet live TV or IPTV. Source-Specific Multicast (SSM) [6] was introduced by IETF more than a decade ago, to overcome the problems of ASM for one-to-many applications like IPTV. SSM endows the receivers with the ability to specify the data source S to receive from, along with the multicast destination address G. The pair (S,G) is called a channel, to differentiate from the term group (G) used in ASM. SSM must be supported in the host stack and applications, and PIM-SSM is just a subset of PIM. SSM has been recognized as a simple, scalable and secure multicast method, even for an inter-domain context.

Multicast has enjoyed success in certain applications inside corporate networks and for IPTV systems in the *walled garden* model. Providers offering IPTV to their customers use IP multicast within their networks, but currently rely on pre-provisioned unicast traffic from the content source into their networks. Such provisioning allows ISPs to deliver content from major content providers to their immediate access customers. Supporting multicast across domains would further allow ISPs to transitively extend this delivery to more viewers and content providers without requiring each content provider to partner with individual ISPs.

However, deployment of multicast in the open Internet is still stagnated, even in the SSM flavour, mainly as a result of service provider policies and network limitations. Only in the context of intra-domain multicast [8] the Sparse Mode (SM) [7] and the SSM variant of PIM (PIM-SM and PIM-SSM, respectively) have been deployed. The problem for inter domain multicast deployment is similar to that of IPv6 deployment: both technologies present an “all or nothing” nature, this is, every router and firewall between source and receiver requires multicast protocols to be enabled. Consequently, the business model for multicast is broken, as those service providers who deploy multicast in their networks don't notice any added value for it, because they can not guarantee the integrity of the multicast chain from source to their customers in the context of the open Internet. Content providers are not interested in transmitting multicast streams that couldn't be received by many end users, and ISP's with many end users are unwilling to invest in deploying multicast in their networks, just to be able to receive and offer to their customers a potential multicast content that doesn't exist. From the user's point of view, to receive a multicast stream or an unicast one is the same, in terms of bandwidth consumption. So, multicast is globally good (less traffic in transient and access networks, better and more services for users), but nobody acts locally to realize it.

2. The AMT solution

Recognizing the problem, the IETF has developed a proposal, called Automatic Multicast Tunneling (AMT) [9], to skip the lack of multicast connectivity between receivers and networks that carry traffic generated by multicast sources. AMT enables sites, hosts, or applications that do not have native multicast access, to request to and receive multicast traffic from a network that does provide multicast connectivity to a source, using UDP-based encapsulation to overcome the lack of multicast connectivity. AMT builds dynamic tunnels from hosts and can support any host application. Although the AMT protocol is defined for both SSM and ASM traffic, it is primarily intended for use in SSM applications.

The AMT protocol employs a client-server model in which a "gateway" sends requests to receive specific multicast traffic to a "relay" that responds by delivering the requested multicast traffic back to the gateway. Gateways are deployed within networks that lack multicast support or lack connectivity to a multicast-enabled network containing multicast sources of interest. The gateway functionality may be directly implemented in the host requesting the multicast service or within an application running on a host, servicing one or more receivers in the same network. Relays are deployed usually in routers within multicast-enabled networks that contain, or have connectivity to, multicast sources. The primary function of AMT is to provide the handshaking, encapsulation, and decapsulation required to transport the IGMP and MLD messages and multicast IP datagrams between gateways and relays. The IGMP and MLD messages that are exchanged between gateways and relays are encapsulated as complete IP datagrams within AMT control messages. Multicast IP datagrams are replicated and encapsulated in AMT data messages. Notice that the bandwidth cost for this replication will be higher than that required if the receivers were reachable via native multicast. All AMT messages are sent via unicast UDP/IP.

To find a relay with which to communicate, AMT calls for a model that uses anycast. Under this approach, one or more relays advertise a route for the same IP address prefix. A gateway sends a message to an anycast IP address within that prefix. This message is routed to the topologically nearest relay that has advertised the prefix. The relay that receives the message responds by sending its unicast address back to the gateway. The gateway uses this address as the destination address for any messages it subsequently sends to the relay. Once the gateway has located an AMT relay, it periodically sends IGMP/MLD messages over a dynamically created UDP tunnel to the relay, in a way similar to a receiver requesting native multicast traffic from a local DR. Relays receive the multicast traffic natively from the sources, and encapsulate it into unicast datagrams to be forwarded to gateways through tunnels. This allows any potential receiver in the Internet to create a dynamic tunnel to download multicast data streams.

AMT is a transition strategy towards the full multicast-enabled Internet, a way to provide much more video content over the Internet in the near future, until all service providers support multicast. Its hybrid multicast-unicast approach reduces the per user cost for content providers, by moving the replication point (the relay) as close to the end user as possible, and making it a part of the existing network infrastructure, providing a more efficient and less costly way to replicate datagrams. Is a way for transit service providers (who can get access to the content, but don’t have many end users) to provide video delivery service to content owners, where it would not be economically feasible otherwise. It can also be a transition strategy for local service providers to afford a partial transition to multicast, not supporting it on all downstream equipment. In this way, they can keep legacy unicast-only Broadband Remote Access Server (BRAS) or digital subscriber line access multiplexer (DSLAM) devices, while offering multicast access to their subscribers by AMT relays.

Despite not being approved by IETF until 2015 as a proposed standard, AMT relay functionality is supported in routers at least from 2011 (in Juniper routers). Even so, it has not brought the necessary momentum to the expansion of multicast.

3. An alternative to AMT: PIM extension for hosts

AMT paves the way for Internet multicast by skipping the non multicast gap between receivers and multicast enabled networks. It seeks to minimize changes that users and networks providers have to do to enable AMT in their equipments. Nevertheless, using AMT requires the multicast enabled network provider to install and manage relays, and gateways have to be deployed within networks containing multicast receivers or bundled into the user applications receiving multicast traffic. A relay discovery protocol is needed, and, finally, as any tunnelling mechanism, AMT is less efficient than native multicast or even unicast traffic replicated in relays. Here is described a similar but simpler method than AMT to skip the non-multicast gap.

Operation

The central idea proposed here is similar to AMT, this is, to replicate multicast traffic as unicast in the border between multicast and unicast Internet, but avoiding the use of tunnelling to get it. The way for it is to simplify the multicast protocol architecture concerning SSM, removing the use of a Designated Router and the protocols needed to communicate hosts and DRs, this is, IGMP/MLD, as they are only necessary when you don't know where is/are the source(s) for the requested traffic. This is a characteristic of ASM, not present in SSM. Doing so, the “all-or-nothing” nature of multicast vanishes, and tunnelling is no longer needed. For a multicast receiver to express its interest in receiving traffic addressed to a multicast group, instead of using IGMP/MLD to report to the closest multicast router, we propose to extend PIM-SM to be used by receiver hosts for directly requesting multicast traffic through the unicast Internet. In fact, PIM-SM definition does not force to use IGMP/MLD (“other mechanisms might also serve this purpose” in [7], p. 7). So, the process to start receiving a multicast channel would be:

- The receiver R sends a new PIM-join/host (S,G) message into an unicast IP datagram with destination S, being R the unicast IP address of the interface claiming to receive the channel.
- This PIM-join/host travels through the unicast route towards the requested source, until it runs into a PIM-extended multicast enabled router (m-router).
- If this m-router is not in the MDT for the requested channel (S,G), it will perform two tasks: (1) it turns the PIM-join/host request into a standard PIM-join and sends it upstream to the source S, to get itself into the MDT for that channel and start receiving it, and, (2) as it does when it gets a standard PIM-join for a new channel, it adds an entry in its multicast forwarding table for (S,G). But in this new case for a PIM-join/host, it also adds a PIM-extended forwarding branch in the (S,G) entry of the multicast forwarding table, with next-hop being R, instead of the next m-router downstream in the MDT that appears in a standard multicast forwarding branch. The PIM-extended nature of this forwarding branch marks that received multicast traffic must be forwarded in unicast fashion in this branch, using R instead of G for the destination address of the forwarded datagram.
- If the m-router is currently in the MDT for (S,G), it only has to add the described extended-PIM branch for (S,G) entry in the multicast forwarding table.
- When the m-router receives a datagram with source S and destination G, besides forwarding it in multicast fashion towards the m-routers downstream in the MDT, it makes an unicast copy and forwards it to R.

To abandon a multicast channel (S,G) two different mechanisms can be used:

- In a hard state fashion, the application in the receiving host sends a PIM-prune/host unicast message with destination S. When this message reaches the PIM-extended router that is forwarding towards the host, this router will remove the branch in the (S,G) entry for that receiver. If the entry becomes empty, a standard PIM-prune message is sent upstream in the MDT.

- In a soft state fashion, the host branches in the forwarding tables need an associated per receiver downstream join timer (S,G, host). The timer starts when a PIM-join/host is received for that branch. If the timer expires, the branch is removed. So, the application in the receiving host must periodically renew its channel subscription sending PIM-join/host messages.

Implementation and deployment issues

- The standard PIM join and prune messages cannot be used by receivers, as these messages are sent to multicast addresses. So, in the protocol side, RFC 4601 has to be updated to define a new type of PIM message, the Join-prune message for hosts, with a unicast destination address. As only 9 codes are currently used from 16 possible for the Type field in the PIM message, this is not a problem.
- In the router side, it requires software changes in the PIM entity, maintaining backwards compatibility with previous versions.
- In transit networks, it can be deployed only in the multicast routers located in the downstream border, those which connect transit providers with non multicast ISPs. So, it doesn't require a whole deployment in current Internet multicast, not even in all m-routers inside a current multicast domain.
- In ISP networks, it can be deployed in routers located in the distribution network by ISPs willing to keep legacy non multicast equipment in their access networks (BRAS and/or DSLAM). This way, they can benefit from reducing the traffic cost of links with transit providers, with a minimal investment and risk in multicast transition. It enables a smooth multicast transition in ISP networks, powered by the demand of multicast traffic by subscribers.
- In the receiver host side, applications should turn to use extended PIM instead of IGMP/MLD to subscribe/unsubscribe multicast channels. To do it, the new socket options and functions specified in [10] provide support of one-to-many type multicast applications. Specifically, it defines socket options to join and leave source-specific channels (MCAST_JOIN_SOURCE_GROUP and MCAST_LEAVE_SOURCE_GROUP).

Main characteristics

- Initially, PIM extension for hosts is only defined for SSM multicast, not for ASM. However, it could be extended for ASM as well, using an anycast addressing model to identify extended-PIM multicast routers, and sending PIM-join/host for ASM traffic to this anycast destination.
- The whole scheme is similar to an AMT system, being a hybrid multicast-unicast approach that moves the replication point as close to the end user as possible in the multicast core network.
- But compared to AMT, this proposal is more efficient, as does not encapsulate/decapsulate datagrams, and does not require the intervention of a gateway between receivers and multicast routers.
- Its implementation in multicast enabled routers is simple. It just require to add the ability to process the new join and prune messages from hosts. It does not require any new data structure, as the same multicast forwarding table is used, and join timers associated to multicast branches are currently used.
- It does not require any extra management task to network administrators, it should work just

enabling extended PIM in selected routers.

4. References

[1] Stephen Deering and David Cheriton. Host groups: A multicast extension to the Internet Protocol, RFC 966, December 1985.

[2] B. Cain, S. Deering, I. Kouvelas, B. Fenner, and A. Thyagarajan. Internet group management protocol, version 3, RFC 3376, October 2002 .

[3] R. Vida and L. Costa. Multicast listener discovery version 2 (MLDv2) for IPv6, RFC 3810, June 2004.

[4] P. Savola. Overview of the Internet Multicast Routing Architecture, RFC 5110, January 2008.

[5] Stephen Deering. Host extensions for IP multicasting, STD 5, RFC 1112, August 1989.

[6] H. Holbrook, B. Cain. Source-Specific Multicast for IP, RFC 4607, August 2006.

[7] B. Fenner et.al.. Protocol Independent Multicast - Sparse Mode (PIM- SM), IETF RFC 4601, Aug. 2006.

[8] L. Zheng, J. Zhang, R. Parekh. Survey Report on Protocol Independent Multicast - Sparse Mode (PIM-SM) Implementations and Deployments, IETF RFC 7063, Dec. 2013.

[9] G. Bumgardner. Automatic Multicast Tunneling, IETF RFC 7450, February 2015.

[10] D. Thaler, B. Fenner, B. Quinn. Socket Interface Extensions for Multicast Source Filters, IETF RFC 3678, Jan. 2004.