

Apuntes
de
Redes de Ordenadores

Tema 12
IP Multicast

Uploaded by

IngTeleco

<http://ingteleco.iespana.es>
ingtelecowed@hotmail.com

La dirección URL puede sufrir modificaciones en el futuro. Si
no funciona contacta por email

12.- Multicast

12.1.- Introducción

La diferencia entre un servicio unicast y uno multicast es que en el primero existe un emisor y un receptor de cada datagrama, mientras que en el segundo hay un emisor pero normalmente son varios los destinatarios del mismo.

Son numerosas las aplicaciones que se basan en un servicio multicast, por ejemplo:

- Teleconferencia (audio, video, pizarra compartida, editor de texto)
- Juegos compartidos y simulaciones distribuidas
- Internet TV
- Aplicaciones distribuidas
- Transferencias fiables de ficheros
- Localización de un Servidor/servicio

Aunque este tipo de envío podría conseguirse con el envío de datagramas unicast (punto a punto) a cada uno de los posibles nodos de destino, hay numerosas razones que hacen aconsejable la capacidad multicast.

El primer modo y más simple de implementar este tipo de aplicaciones es realizar multicast mediante unicast del nivel de aplicación, es decir, que la aplicación emisora envíe N paquetes duplicados unicast, uno a cada receptor.

Esta solución, si bien muy simple plantea el problema de la replicación de paquetes, que aumentan la utilización de la red, y en especial del enlace o enlaces de salida del emisor que está inundado con N paquetes duplicados, tantos como destinatarios posibles. La red no puede hacer nada para evitar esta situación, puesto que las copias no son generadas en la red sino fuera de ésta.

El unicast a nivel de aplicación resulta útil para grupos pequeños y con una utilización baja del ancho de banda, como los juegos de red

Además del problema de replicación de paquetes, hay otros dos problemas añadidos, el de la gestión de la pertenencia a grupos y el de la denominación de estos. La gestión de la pertenencia es complicada porque el emisor debe mantener una lista con todos los nodos pertenecientes a un grupo, con el fin de poder enviar una copia del paquete a cada uno de ellos; teniendo en cuenta que la pertenencia a los grupos es dinámica resulta muy complicado poder mantener esta lista actualizada en todo momento.

En cuanto a la denominación, en el mundo unicast cada estación es identificada por su dirección de red, y es por tanto esta dirección la que se utiliza para encaminar paquetes hacia/desde un nodo. En una implementación multicast de este tipo, no queda más remedio que denominar un grupo mediante la lista de las direcciones IP de sus miembros, lo que nos conduce al mismo problema de mantenimiento anterior, aparte de limitar notablemente la escalabilidad.

La solución a estos problemas es que el nivel de red ofrezca un servicio multicast. La principal ventaja en el uso de envíos multicast es la disminución del tráfico en la red; los datagramas que comparten un grupo de enlaces hasta sus destinos sólo precisan que sean transmitidos una vez y sólo se replica el mensaje cuando es necesario, recayendo la responsabilidad de la duplicación de paquetes en los routers (para hacer llegar una copia a cada miembro del grupo), de modo que estos aseguren que los paquetes viajarán una vez por cada enlace como máximo (basta con asegurar que todos los miembros reciben una copia del paquete). De este modo, la responsabilidad de la gestión de pertenencia a un grupo también reside en la red, y serán los routers quien mediante el uso de protocolos de routing se encarguen de mantener esta información. En cuanto a la denominación de los grupos, cada uno se identificará mediante un nombre lógico o dirección multicast, lo que desacopla la identidad del grupo de la de sus miembros, favoreciendo por lo tanto el escalado.

En cuanto al funcionamiento de los grupos multicast, es importante tener en cuenta que los miembros de los mismos son los receptores, mientras que los emisores pueden ser o no miembros del grupo. Además está permitido que un nodo pertenezca simultáneamente a varios grupos, además de considerar que la pertenencia a los mismos tiene un carácter dinámico.

Otro aspecto muy interesante del envío multicast hace referencia al descubrimiento de recursos. Son varias las aplicaciones que precisan saber si un determinado tipo de servicio está disponible en la red o no; los protocolos Bootstrap Protocol (BOOTP) y Open Shortest path First (OSPF) se encuentran entre ellos. El uso de mensajes multicast para enviar la petición a los nodos que son potencialmente capaces de proporcionar este servicio resultaría muy útil a estas aplicaciones. Otro aspecto importante del envío multicast es su uso por parte de aplicaciones de transmisión multimedia, lo que se ha convertido en un hecho bastante frecuente. Las señales de audio y vídeo son capturadas, comprimidas y transmitidas a un grupo de estaciones receptoras. En lugar de utilizar conexiones punto a punto entre los nodos participantes, el envío multicast permite la distribución de los datos multimedia a los receptores. En el mundo real las estaciones se pueden unir o abandonar en cualquier instante la difusión de audio o vídeo. La flexibilidad de adhesión o abandono de un grupo proporcionado por el envío multicast puede hacer sencillo el manejo de una pertenencia variable.

La noción de grupo resulta esencial para el concepto de multicast. Por definición, un mensaje multicast se envía desde un emisor a un grupo de nodos de destino. Hay tres tipos de direcciones IPv4: unicast, broadcast, y multicast. Una dirección unicast está concebida para transmitir un datagrama a un solo destino. Una dirección de difusión o broadcast se emplea para enviar un datagrama a toda una subred. Una dirección multicast está pensada para permitir el envío de datagramas a un conjunto de nodos que han sido configurados como miembros de un grupo multicast distribuidos por varias subredes.

El envío multicast no es orientado a la conexión; un datagrama se envía a los miembros del grupo de destino con la misma fiabilidad ("best-effort") que un datagrama unicast IP standard. Esto supone que los datagramas multicast no tienen garantizada la entrega a todos los miembros del grupo, ni la llegada en el mismo orden de transmisión de los paquetes.

La única diferencia entre un datagrama multicast IP y uno unicast es la presencia de una dirección de grupo en el campo dirección de destino de la cabecera IP. En lugar de una dirección de Clase A, B, o C, el envío multicast utiliza una dirección de destino de Clase D (224.0.0.0- 239.255.255.255).

Cuando un nodo (un proceso de un nodo) desea recibir mensajes multicast enviados a un grupo concreto, sólo precisa "escuchar" los mensajes enviados al mismo. Si el emisor y el destinatario de un datagrama comparten un bus común (i.e. Bus Ethernet), cada nodo sólo preciso saber qué grupos tienen miembros entre los procesos del mismo. Sin embargo, si el emisor y el destinatario no comparten LAN, el envío de los mensajes multicast resulta más complicado. Para solucionar el problema del envío de datagramas multicast en el ámbito de Internet, los nodos deben unirse a un grupo informando de ello a los routers multicast de su subred, para lo que utilizan el protocolo IGMP (Internet Group Management Protocol). El abandono del grupo se hace también mediante IGMP. De este modo los routers multicast de la red conocen los miembros de los grupos multicast de su red y pueden decidir si encaminar los mensajes multicast a su red o no. Siempre que un router multicast recibe un datagrama multicast comprueba el identificador de grupo del mensaje y encamina el mensaje sólo si hay un miembro de dicho grupo en redes a las que está conectado. IGMP proporciona la información precisa en la última etapa del envío de mensajes multicast a sus destinos.

Sin embargo, para el envío de datagramas multicast desde el emisor hasta los nodos de destinos de otras redes, los routers multicast necesitan intercambiar la información que han reunido acerca de la pertenencia a grupos de los nodos directamente conectados a ellos. Hay varios algoritmos tales como la inundación ("flooding"), árboles de expansión ("spanning tree"), "reverse path broadcasting", y "reverse path multicasting" para intercambiar información de encaminamiento entre routers. Algunos de estos algoritmos se han usado en protocolos de routing multicast dinámicos tales como DVMRP (Distance Vector Multicast Routing Protocol), MOSPF (Multicast extension to Open Shortest Path First), y PIM (Protocol Independent Multicast). Basándose en la información de encaminamiento obtenida a partir de estos protocolos, cuando un datagrama multicast se envía a un grupo multicast, los routers multicast decidirán si encaminar dicho datagrama a sus redes o no. Finalmente, los routers hojas en el árbol de distribución multicast, verán si hay o no miembros de un grupo particular en las redes a las que está físicamente conectado utilizando la información de IGMP y decidiendo si encaminar o no el datagrama.

En las siguientes secciones revisaremos las direcciones multicast y el modo en que se relacionan éstas con las direcciones de nivel MAC. Se estudiará también el protocolo IGMP y las extensiones precisas en los nodos para soportar IP multicasting. Los algoritmos y protocolos de routing constituirán el grueso del capítulo, para finalmente revisar el Internet Multicast Backbone y sus especificaciones.

12.2.- Modelo de Servicio IP Multicast (RFC-1112)

La especificación del Servicio Multicast del protocolo IP aparece en el RFC1112, y detalla la operación de los servicios multicast en este modelo.

Los grupos definidos en este modelo pueden ser de cualquier tamaño y con sus miembros ubicados en cualquier lugar de Internet. La pertenencia a los grupos es dinámica, pudiendo adherirse o abandonar un grupo en cualquier momento. Los emisores de un grupo no necesitan ser miembros del mismo, de modo que un nodo cualquiera puede enviar datagramas a cualquier grupo definido en Internet.

Cada grupo se identifica mediante una sola dirección IP de clase D (224.0.0.0 - 239:255.255.255).

Los routers escuchan las direcciones multicast y utilizan protocolos de routing para gestionar la pertenencia a los grupos (IGMP, RFC2236) y por lo tanto tener información sobre la necesidad de replicación de datagramas para que estos alcancen a cualquier miembro del grupo.

12.2.1.- Direcciones IP Multicast

Las direcciones clase D (entre 224.0.0.0 y 239.255.255.255) están previstas en IP para tráfico multicast.

Se asigna una dirección IP de Clase D a un grupo de nodos que define un grupo multicast. Los cuatro bits más significativos de las direcciones de Clase D se fijan a "1110", y los siguientes números de 28-bit reciben la denominación de identificador del grupo multicast, no estando, por lo tanto estructuradas las direcciones como las direcciones IP unicast. Su formato es el indicado en la figura siguiente:

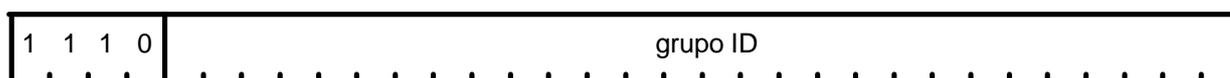


Figura 12.1 Dirección IP de Clase D

Existen dos tipos de grupos multicast: temporales y permanentes. Algunos grupos permanentes han sido predefinidos por la IANA (Internet Assigned Numbers Authority) en el RFC1700 para propósitos especiales. La dirección 224.0.0.0 está reservada y el rango de direcciones desde 224.0.0.1 a 224.0.0.255 está reservado para el uso de protocolos de routing y otros protocolos de descubrimiento y mantenimiento de topologías de bajo nivel; los routers multicast no deberían encaminar datagramas con dirección de destino dentro de este rango independientemente del valor de su TTL.

El resto de los grupos que van desde 224.0.1.0 a 239.255.255.255 están asignados a diversas aplicaciones multicast o permanecen sin ser asignadas. De este rango, las direcciones que van de 239.0.0.0 a 239.255.255.255 están reservadas para ser utilizadas para aplicaciones locales no extensibles a Internet. Existen algunas otras direcciones de Clase D ya reservadas para grupos bien conocidos, por ejemplo:

224.0.0.1	Todos los nodos en una LAN (all-systems.mcast.net)
224.0.0.2	Todos los routers en una LAN (all-routers.mcast.net)
224.0.0.5	Todos los routers OSPF en una LAN
224.0.0.6	Todos los routers OSPF designados en una LAN
224.0.1.11	IETF-1 Audio
224.0.1.12	IETF-1 Vídeo

Las direcciones multicast temporales son asignadas y abandonadas dinámicamente.

Los servicios IP multicast en la emisión y recepción de datagramas presentan algunas características específicas respecto del servicio unicast. El servicio de emisión utiliza la operación normal de emisión de datagramas IP, con una dirección IP multicast como destino, pero debe proporcionar a la aplicación emisora un modo de:

- Especificar la interfaz de red de salida, cuando hay más de uno
- Especificar el tiempo de vida (TTL) del paquete emitido
- Habilitar/deshabilitar el "loopback" si el nodo emisor es miembro del grupo de destino en la interfaz de salida

A la operación normal de recepción de datagramas IP debe incorporar dos operaciones nuevas:

- Unir_a_grupo_multicast (Dirección de grupo, interfaz)
- Abandonar_grupo_multicast (Dirección de grupo, interfaz)
- Recibir datagramas multicast para los grupos a los que pertenece mediante la recepción normal de datagramas IP

12.2.2.- Traducción de direcciones IP de Clase D a direcciones MAC

La transmisión de un paquete multicast supone la traducción de la dirección IP de destino a una dirección MAC multicast.

Un datagrama multicast se envía al grupo de miembros con la misma fiabilidad del mejor intento que los datagramas unicast, siendo posible la pérdida y desordenación de los datagramas. Como en los datagramas unicast, debería existir una dirección MAC a la cual hacer corresponder la dirección IP. La IANA ha reservado un grupo de direcciones MAC IEEE-802 para los paquetes multicast que van desde 01:00:5E:00:00:00 a 01:00:5E:7F:FF:FF (hex). Una dirección IP multicast puede corresponderse a una dirección IEEE-802 ubicando en los 23 bits menos significativos de la dirección IP multicast en los 23 bits menos significativos de la dirección MAC multicast. La correspondencia entre direcciones IP multicast y direcciones MAC IEEE-802 se ilustra en la figura adjunta que muestra cómo la dirección de grupo 224.10.8.5 (E0-0A-08-05) se traduce en una dirección Ethernet Multicast.

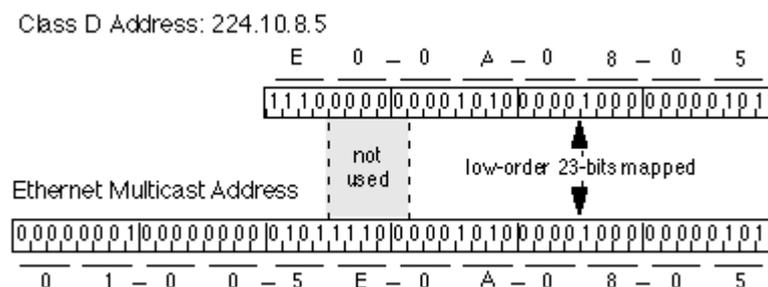


Figura 12.2 Conversión entre Dirección IP y dirección MAC Ethernet

Como puede intuirse, habrá 32 direcciones IP diferentes que se corresponderán con la misma dirección Ethernet, porque los cinco bits del identificador de grupo multicast son ignorados.

12.3.- IGMP

Los nodos que desean recibir datagramas multicast deben informar a los routers vecinos que están interesados en recibir datagramas dirigidos a ciertos grupos multicast. De este modo, cada nodo se convierte en miembro de uno o más grupos multicast y recibe los datagramas dirigidos a dicho grupo. El protocolo mediante el que los nodos comunican esta información a los routers se denomina Internet Group Management Protocol (IGMP). IGMP también es utilizado por los router para comprobar periódicamente si los miembros de los grupos conocidos están todavía activos. En caso de que exista más de un router multicast en una subred (LAN), uno de ellos es elegido para efectuar las consultas y asumir la responsabilidad de cuidar del estado de pertenencia de los grupos multicast con miembros activos en su subred. Basándose en la información obtenida de IGMP el router puede decidir si reenviar los mensajes multicast que recibe a sus subredes o no. Después de recibir un datagrama multicast enviado a un determinado grupo multicast, el router comprobará si existe al menos un miembro del grupo particular en su subred. Si ese es el caso, el router reenviará el mensaje a la subred; en caso contrario descartará el datagrama. Obviamente, esto será la última fase del envío de un datagrama multicast.

El actual estándar en Internet es IGMPv1 (RFC-1112), si bien ya se ha publicado la versión 2 (RFC-2236) que ya se está difundiendo con rapidez y está en proceso de estudio la versión 3 del protocolo. Opera sobre LANs de difusión y enlaces punto a punto (su operación sobre redes no de difusión, multiacceso [NBMA] está todavía en estudio, p.e., MARS para ATM).

La operación del protocolo es la siguiente.

En cada LAN se elige un router (router designado) que será en el encargado de enviar las peticiones a la red para que los nodos informen a qué grupos están adheridos. Este router, envía periódicamente un mensaje de Pregunta de Pertenencia al grupo de “todos los nodos” (224.0.0.1), con TTL = 1

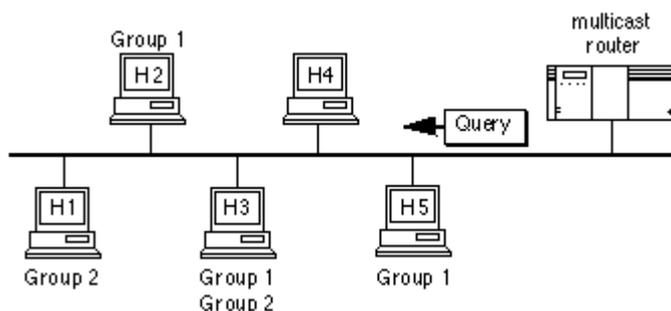


Figura 12.3 Mensaje Query IGMP

Cuando un nodo escucha una de estas peticiones arranca un temporizador aleatorio (entre 0 y 10 segundos) por cada uno de los grupos a los que pertenece. Cuando expira el temporizador para un grupo G, envía una Notificación de Pertenencia (“report”) a dicho grupo, con TTL = 1, por lo que dichas notificaciones nunca saldrán de la subred en que fueron generadas. El resto de los miembros de G escucharán la Notificación y detendrán sus temporizadores para dicho grupo, evitando enviar notificaciones de pertenencia al mismo, puesto que al router sólo le interesa conocer qué grupos tienen miembros, pero no cuántos ni quienes son estos. A partir de estas notificaciones el router construirá una lista con los grupos conectados a cada uno de sus interfaces.

Cuando un nodo se adhiere a nuevo grupo, envía una o dos notificaciones inmediatas, en lugar de esperar una pregunta con el fin de que la pertenencia al grupo sea conocida lo antes posible.

Cuando un router no recibe ninguna respuesta para un determinado grupo después de varias preguntas, asume que no hay ningún miembro del grupo en la red.

El formato de los mensajes IGMP V1 es el siguiente:

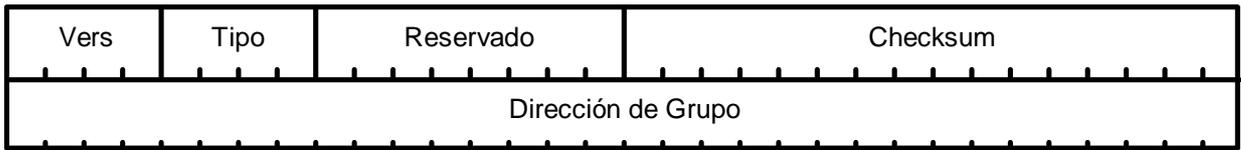


Figura 12.4 Formato del mensaje IGMP v1

El campo Version especifica la versión de IGMP.

El campo Tipo diferencia entre los distintos tipos de mensaje:

- 1 = Pregunta de Pertenencia
- 2 = Notificación de pertenencia

El campo Checksum permite comprobar que el mensaje no contiene errores.

El campo Dirección(es) de Grupo contiene la lista de direcciones IP de Clase D a las que pertenece el nodo que envía la notificación (en las preguntas está vacío).

La principal diferencia de IGMP Versión 2, que es una mejora del original, es la inclusión de un nuevo mensaje para indicar el abandono de un grupo, lo que reduce la latencia de abandono ("leave latency"). Además se especifica un método de elección estándar para el router que genera las preguntas, el de menor dirección IP, algo que había quedado inconcluso en la versión anterior. Finalmente se unen en un solo campo Versión y Tipo, pero manteniendo la compatibilidad con la versión 1.

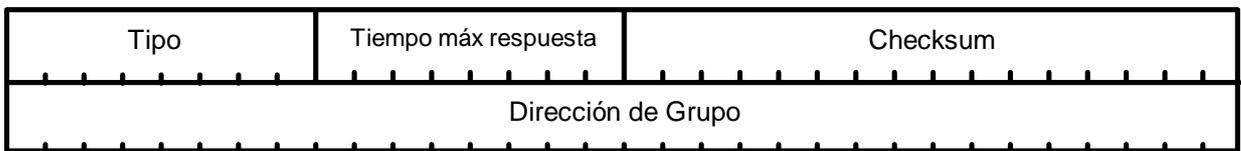


Figura 12.5 Formato del mensaje IGMP v2

IGMP Versión 3, la última versión, posibilita que los nodos se unan a un grupo y especificar un conjunto de emisores de dicho grupo desde los cuales quiere recibir datagramas multicast añadiendo capacidad de filtrado de emisor. Igualmente los mensajes de abandono de grupo de la Versión 2 se han mejorado para soportar el abandono de grupos-emisores.

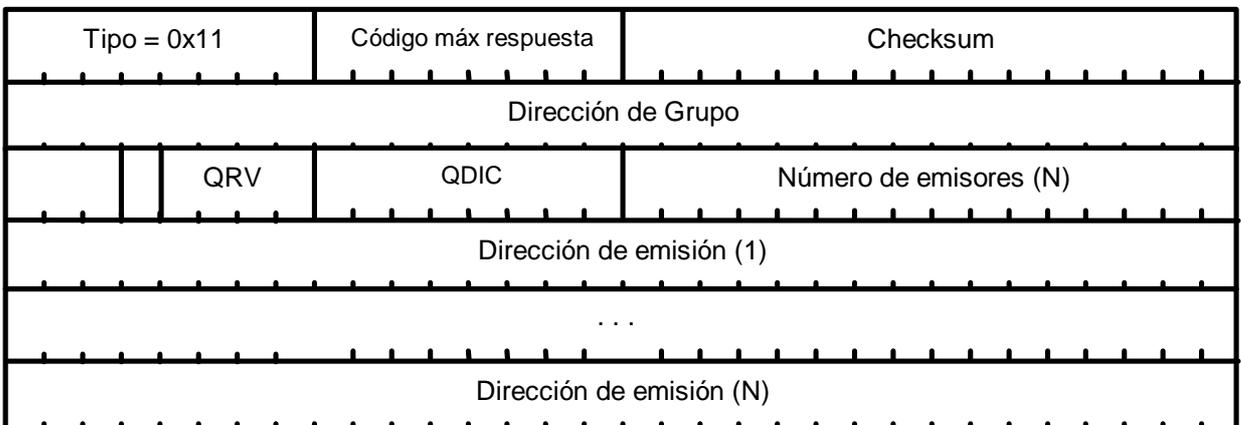


Figura 12.6 Formato del mensaje IGMP v3

12.4.- MBone

En 1992, se escogió un conjunto de subredes interconectadas con routers capaces de encaminar datagramas multicast para experimentar con multicasting. Este entorno de prueba multicast fue denominado Multicast Backbone (MBone) y proporcionó un medio para el desarrollo de aplicaciones multicast. El MBone que comenzó con 40 subredes en cuatro países diferentes, incluye en la actualidad más de 3400 subredes en más de 25 países y continuará creciendo a gran velocidad.

Major MBONE Routers and Links

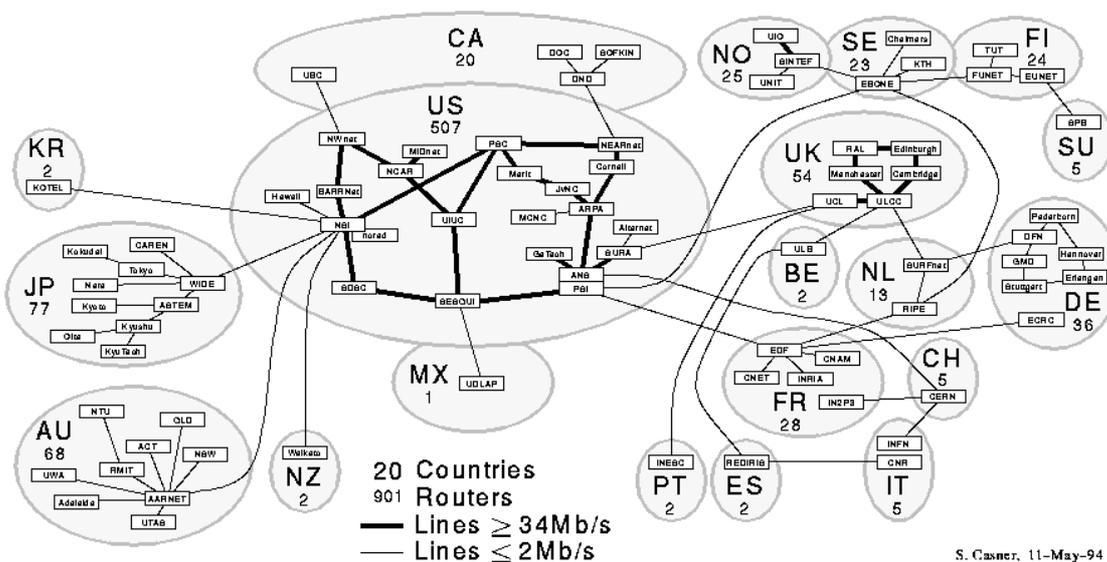


Figura 12.7 Mapa de MBONE en 1994

MBone es esencialmente una red virtual implementada sobre algunas porciones de Internet. En MBone, se conectan entre sí islas de redes con capacidad multicast mediante enlaces virtuales denominados "túneles". Es a través de estos túneles como se encaminan los mensajes multicast a través de las porciones de Internet sin capacidad multicast. Para encaminar los datagramas multicast a través de estos túneles, se encapsulan a su vez en IP (IP-over-IP) de modo que parezcan datagramas unicast normales a los routers involucrados.

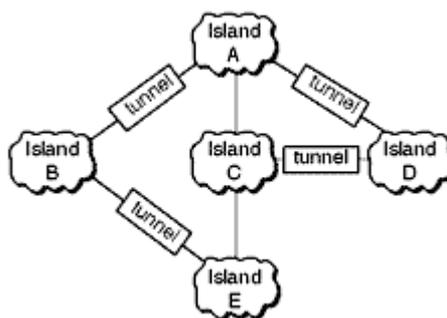


Figure 12.8: Internet Multicast Backbone (MBONE)

Los routers multicast, sus subredes directamente conectadas y los túneles que las interconectan son los componentes de MBone. El único protocolo de routing utilizado en las primeras etapas de MBone fue DVMRP, y aunque en la actualidad se emplean otros tales como MOSPF y PIM, todavía es DVMRP el empleado por la mayoría de los routers MBone. A medida que aumente la disponibilidad de software de routing multicast en todos los routers empleados en Internet, el uso de multicast "nativo" reemplazará gradualmente la necesidad de utilización de túneles.

En la actualidad, Mbone se emplea para transportar multicast de audio y vídeo para las reuniones de las IETF, misiones espaciales de la NASA, Sesiones del Congreso y Senado de los USA así como conferencias y seminarios de carácter técnico.

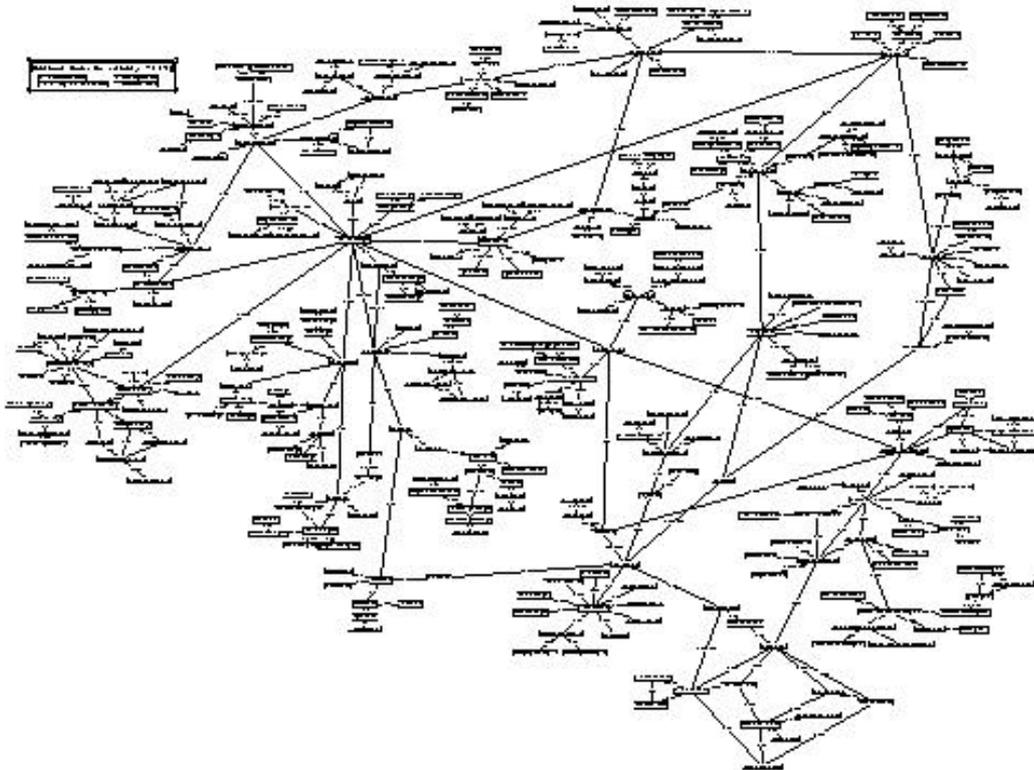


Figura 12.9 Mapa de MBONE en 1998

12.5.- Routing IP Multicast

IGMP proporciona el último paso en el servicio de envío de datagramas multicast, dado que sólo está relacionado con el envío de tráfico multicast desde el router local a los miembros de un grupo directamente conectados a sus subredes. IGMP no contempla el envío de datagramas multicast entre routers vecinos o a través de Internet. Para conseguir un servicio global es preciso definir protocolos de routing multicast, responsables de la construcción de árboles de distribución a través de los cuales se distribuyan los datagramas multicast.

El problema del routing multicast es hacer llegar los datagramas a todos los miembros de un grupo, para ello es preciso que los routers tengan una visión de la ubicación de grupos en la red limitada no a las redes a las cuales están directamente conectados, sino en toda la red.

A la pregunta ¿cómo hacer llegar los paquetes a los miembros de un grupo? pueden proporcionarse varias respuestas:

- Puede utilizarse la difusión de los datagramas, lo que supone un consumo de ancho de banda que no resulta aceptable puesto que los datagramas circulan por redes y segmentos en los cuales no hay ningún nodo perteneciente al grupo.
- Puede utilizarse un árbol de expansión ("Spanning tree"), lo que supone aplicar en algunos casos un retraso muy superior al de un camino unicast. Además puede producirse un problema de concentración de tráfico porque el tráfico no se distribuye adecuadamente por los recursos de la red.
- La mejor solución es resolver la pertenencia dinámica a los grupos, con el consiguiente mantenimiento del árbol de distribución de los datagramas.

Los criterios aplicables para la evaluación de los protocolos de routing multicast son los siguientes:

- Ancho de banda que utilizan para la distribución de datos.
- Volumen del tráfico de control necesario.
- Número de estados necesarios en los routers.
- Latencia de la adhesión (y abandono).
- Adaptabilidad a los cambios en la pertenencia a los grupos y la topología
- Acomodo de emisores dinámicos

12.6.- Algoritmos de routing multicast

Se han propuesto numerosos algoritmos para construir árboles multicast a través de los cuales enviar los datagramas hasta sus destinos y que pueden utilizarse en los protocolos de routing multicast. Revisaremos en esta sección desde los más sencillos, Inundación y Arboles de Expansión hasta los más sofisticados tales como Reverse Path Forwarding (RPF), Truncated Reverse Path Forwarding (TRPF), Steiner Trees (ST), y Core-Based Trees (CBT).

12.6.1.- Inundación

El algoritmo de inundación, que se utiliza en protocolos tales como OSPF es la técnica más sencilla para enviar los datagramas multicast a los routers de una red. En este algoritmo, cuando un router recibe un datagrama multicast, comprueba si ha recibido este mismo datagrama con anterioridad o es la primera vez que lo recibe. Si es la primera vez, el router reenviará el datagrama por todos sus interfaces a excepción de aquel por el que lo recibió; en caso contrario, el router descartará el datagrama, asegurando así que todos los routers de la red recibirán al menos una copia del datagrama.

Aunque este algoritmo resulte muy simple, tiene algunas desventajas importantes. Genera un gran número de datagramas duplicados, desaprovechando por consiguiente una parte importante de ancho de banda. Además, dado que cada router necesita seguir la pista de los datagramas para saber si lo han recibido con anterioridad, deben mantener una entrada diferente en su tabla para cada datagrama recientemente recibido, por lo que este algoritmo conduce a una utilización muy pobre de los recursos de memoria.

12.6.2.- Arboles de expansión

Un algoritmo mejor es el del árbol de expansión ("Spanning Tree"). Este algoritmo, que se utiliza también en los puentes IEEE-802 es potente y fácil de implementar. En él, se selecciona un subconjunto de enlaces para definir una estructura de árbol (un grafo libre de bucles) tal que sólo haya un camino activo entre dos routers cualesquiera. Dado que este árbol se expande a todos los nodos de la red se denomina árbol de expansión. Cuando un router recibe un datagrama multicast lo reenvía por todos los enlaces que pertenecen al árbol de expansión, excepto por el que recibió el datagrama, garantizando que el datagrama alcanza a todos los routers de la red. Obviamente, la única información que el router necesita almacenar es una variable booleana por interfaz de red indicando si pertenece al árbol de expansión o no.

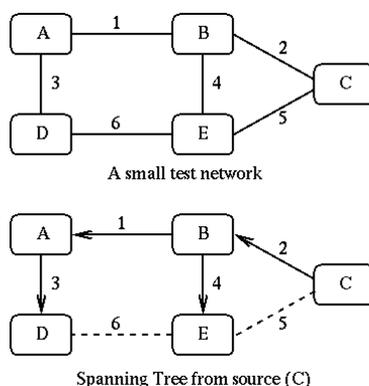


Figura 12.10 Arbol de expansión

En la figura se muestra un ejemplo con cinco nodos y seis enlaces para mostrar la definición de un árbol de expansión.

Los árboles de expansión son muy potentes y fáciles de implementar, sin embargo, tienen dos deficiencias: centralizan todo el tráfico en un reducido número de enlaces y no considera la pertenencia a grupos.

12.6.3.- Reverse Path Broadcasting (RPB)

El algoritmo RPB que se utiliza actualmente en Mbone (Multicast Backbone), es una modificación del algoritmo de árbol de expansión. En este algoritmo, en lugar de construir un árbol de expansión para la red completa, se construye un árbol implícito para cada emisor. Estos árboles de expansión constituirían unos árboles de distribución con la raíz en el emisor emanando desde la red directamente conectada al emisor. Dado que hay muchos emisores (fuentes) potenciales para un grupo, se construye un árbol de expansión para cada par activo (emisor, grupo).

El algoritmo RPB es realmente muy simple. Para cada pareja (emisor, grupo), cuando llega un paquete a través de un enlace, que el router local considera que está en el camino más corto hacia el emisor del mismo, el router los reencamina por todos las interfaces excepto por el de llegada. En caso contrario, el datagrama es descartado. La interfaz por el que el router espera recibir los datagramas multicast de un emisor determinado se conoce como enlace "padre". Los enlaces de salida a través de los cuales el router encamina el datagrama multicast se conocen como enlaces "hijos".

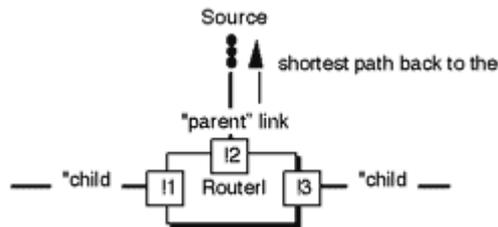


Figure 12.11 Algoritmo RPB

En la figura se muestran a modo de ejemplo los árboles de distribución multicast desde dos emisores A y C respectivamente.

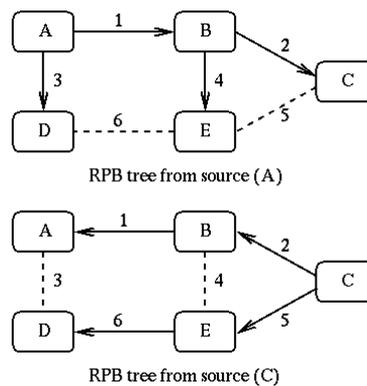


Figure 12.12 Arbol RPB

El algoritmo RPB puede ser mejorado fácilmente para evitar duplicaciones de datagramas teniendo en cuenta el hecho de que si el router local no está en el camino más corto entre el emisor y un vecino, el datagrama será descartado en el router vecino. Por lo tanto, si este es el caso, no sería necesario reenviar el datagrama hacia dicho router vecino. Esta información puede obtenerse con facilidad si se está utilizando un protocolo de routing de estado de enlace. Si se emplea un protocolo vector-distancia, un vecino puede bien advertir su próximo salto para el emisor como parte de los mensajes de actualización de rutas o hacer "poison-reverse" de la ruta.

Este algoritmo es fácil de implementar. Además, dado que los datagramas se envían a través del camino más corto desde el emisor a los nodos de destino, resulta muy rápido. El algoritmo RPB no necesita ningún mecanismo para detener el proceso de envío. Los routers no necesitan tener conocimiento del árbol de expansión completo y puesto que los datagramas son enviados a través de árboles de expansión diferentes (y no un único árbol) el tráfico se distribuye entre varios árboles, aprovechando mejor la red. Sin embargo, el algoritmo RPB adolece de una deficiencia grave: no tiene en cuenta la información acerca de la pertenencia a grupos multicast para construir los árboles de expansión.

12.6.4.- Truncated Reverse Path Broadcasting (TRPB)

El algoritmo TRPB ha sido propuesto para solucionar algunas de las limitaciones de RPB. Como mencionamos anteriormente, utilizando IGMP un router puede determinar si hay presentes miembros de un grupo multicast en una subred o no, evitando en este último caso el envío de datagramas multicast hacia ella. Si esta subred es una "subred hoja", es decir, no tiene ningún otro router conectado a ella, el router truncará el árbol de expansión. TRPB es similar a RPB, no se reenviará un datagrama a un router vecino si el router local no está en el camino más corto desde el router vecino al nodo emisor.

En la figura se muestra la operación del algoritmo TRPB. En este ejemplo el router recibe un datagrama multicast por se enlace "padre" para el par (emisor, G1). El router reencamina el datagrama por I1, dado que la interfaz tiene al menos un miembro de G1. Sin embargo no reencamina el datagrama por la interfaz I3 dado que no tiene miembros en el grupo de destino. El datagrama se reencaminará por la interfaz I4, sólo si un router en este enlace considera la interfaz como su enlace "padre" hacia (emisor,G1).

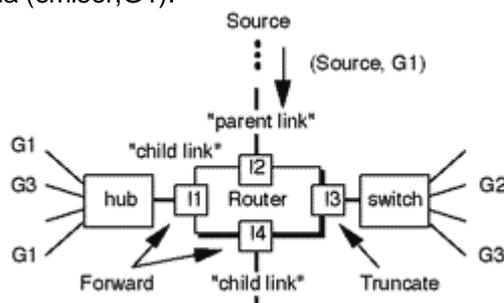


Figure 12.13 Algoritmo TRPB

Aunque utiliza la pertenencia a un grupo, el algoritmo TRPB y las subredes hojas son truncadas del árbol, no se elimina todo el tráfico innecesario en las subredes "no hojas" que no tienen miembros de un grupo.

12.6.5.- Reverse Path Multicasting (RPM)

El algoritmo RPM (también conocido como RPB con podas) es una mejora de RPB y TRPB. RPM construye un árbol de envío que expande sólo:

- 1) subredes con miembros de un grupo
- 2) routers y subredes en el camino más corto a las subredes con miembros de un grupo.

El árbol RPM puede ser podado de modo que los datagramas multicast se envíen por enlaces que conducen a los miembros del grupo destino.

Para un par dado (emisor, grupo) el primer datagrama multicast se envía basándose en el algoritmo TRPB. Los routers que no tienen ningún router más abajo en el árbol TRPB se denominan routers hojas. Si un router hoja recibe un datagrama multicast para un (emisor, grupo) dado y no tiene ningún miembro del grupo en sus subredes, enviará un mensaje de "poda" al router del que recibió el paquete multicast. El mensaje de poda indica que los paquetes multicast correspondientes a dicha pareja (emisor, grupo) no deben ser reenviados por el enlace por que el que se recibe un mensaje de poda.

Es importante resaltar que los mensajes de poda sólo se envían un salto hacia atrás en dirección al emisor. Es preciso que el router por encima registre la información de poda en su memoria. Por otro lado, si el router por encima no tiene ningún miembro local y recibe mensajes de poda de todos sus "hijos" en el árbol TRPB, enviará a su vez un mensaje de poda a su "padre" en el árbol TRPB indicando que los datagramas multicast correspondientes al par (emisor, grupo) no es preciso que sean reenviados hacia él. Los mensajes de poda en cascada truncarán el árbol original TRPB de modo que los datagramas multicast serán enviados sólo por aquellos enlaces que conduzcan a un nodo de destino (miembro de un grupo). En la figura se muestra el árbol obtenido después del intercambio de mensajes de poda en una red.

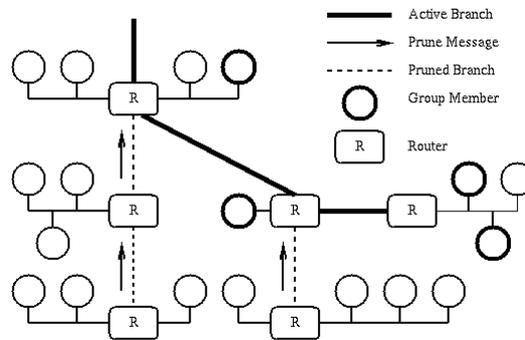


Figure 12.14 Arbol RPM

La pertenencia a grupos y la topología de la red cambia dinámicamente, por lo que el estado de poda de los árboles debe refrescarse a intervalos regulares. Por lo tanto, la información de poda se elimina periódicamente de los routers y el siguiente paquete se reenviará a todos los routers hoja reiniciándose el proceso. RPM precisa una capacidad de almacenamiento relativamente grande para mantener la información de estado de todos los pares (emisor, grupo), lo que constituye un inconveniente que hace al algoritmo no escalable.

12.6.6.- Árboles de Steiner (ST)

En la familia de los algoritmos RPB (RPB, TRPB, y RPM) se utiliza para enviar del datagrama el camino más corto entre el nodo emisor y cada nodo de destino, garantizando que el envío se hace lo más rápidamente posible. Sin embargo, ninguno de estos algoritmos trata de minimizar el uso de los recursos de la red. En la figura se representan un árbol RPB y otro diferente suponiendo que C es el emisor y A y D son los destinatarios.

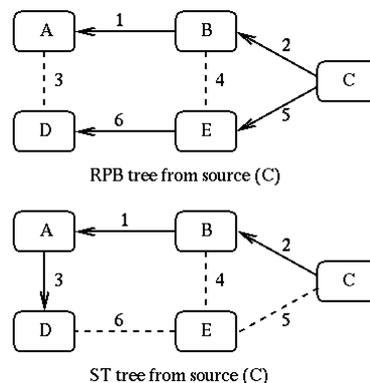


Figure 12.15 Arbol de Steiner

Puede apreciarse que el segundo árbol utiliza menos enlaces. Aunque este árbol es más lento que el RPB (porque los datagramas deben dar tres saltos para llegar a D en lugar de sólo 2 que requiere el árbol RPB), utiliza menos enlaces. Este tipo de árbol es denominado árbol de Steiner. Aunque el árbol de Steiner minimiza el número de enlaces utilizados para construir un árbol de distribución, las

dificultades a la hora de calcularlo han hecho que tenga una importancia muy pequeña. Dada la forma en que el ST cambia con la adhesión o abandono de un nodo de un grupo multicast, estos árboles son muy inestables.

12.6.7.- Core-Based Trees (CBT)

El último algoritmo propuesto para la construcción de árboles de distribución multicast es el Core-Based Tree (CBT). A diferencia de los anteriores, CBT crea un árbol de distribución para cada grupo. En otras palabras, el árbol utilizado para enviar los datagramas multicast de un grupo particular es un árbol independiente de la localización del nodo emisor. Un solo router, o un conjunto de routers son elegidos para constituir el router núcleo ("core") el árbol de distribución. Todos los datagramas para un grupo particular son enviados como mensajes unicast al router "core" hasta que alcanzan un router que pertenece al correspondiente árbol de distribución; entonces el datagrama se envía por todas las interfaces que son parte del árbol de distribución excepto por el que llegó, como se muestra en la figura.

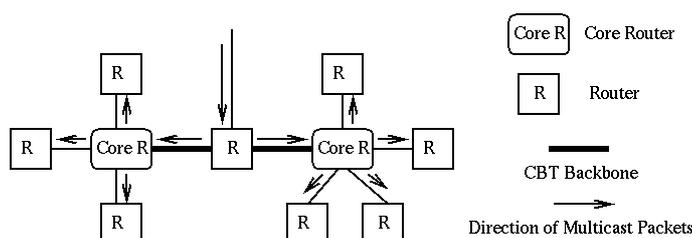


Figure 12.16 Core-Based Tree

Dado que CBT construye un árbol de expansión para cada grupo multicast, los routers multicast necesitan guardar menos información en comparación con los requisitos de los algoritmos anteriores. CBT también conserva el ancho de banda de la red porque no necesita inundar ningún datagrama multicast. Sin embargo, utilizar un árbol por cada grupo puede llevar a una concentración de tráfico y a cuellos de botella en los alrededores de los router "core", además de producir rutas no-óptimas y por lo tanto retraso en el envío de los datagramas.

Los algoritmos discutidos hasta ahora pueden ser utilizados para desarrollar protocolos de routing multicast. Cada uno de ellos tiene ventajas e inconvenientes que los hacen más o menos eficientes en determinadas situaciones.

12.7.- Protocolos de routing multicast

Similares a los protocolos de routing unicast (como RIP y OSPF), hay un conjunto de protocolos de routing multicast que permiten determinar a los routers dónde enviar los datagramas multicast. En este apartado discutiremos los protocolos multicast existentes y veremos cómo dichos protocolos utilizan algunos de los algoritmos presentados anteriormente para intercambiar información de routing multicast. En primer lugar revisaremos tres protocolos de routing: Distance Vector Multicast Routing Protocol (DVMRP), Multicast Extensions to OSPF (MOSPF) y Protocol Independent Multicast - Dense Mode (PIM-DM) que son más eficientes en situaciones donde los miembros de los grupos multicast están densamente distribuidos en la red. A continuación presentaremos Protocol Independent Multicast - Sparse Mode (PIM-SM), que se comporta mejor cuando los miembros de los grupos están distribuidos de modo disperso.

12.7.1.- Clasificación de los protocolos de Routing Multicast

Los protocolos de routing multicast pueden clasificarse atendiendo a diversos criterios. En primer lugar según la clasificación de los árboles de distribución multicast en los que se fundamentan:

- Árboles separados con raíz en cada emisor de datos (DVMRP, MOSPF, PIM-DM, PIM-SM)
- Árboles con raíz compartida en un punto común (CBT, PIM-SM)

El árbol de distribución con raíz en el emisor se construye un árbol para cada emisor cuya raíz se encuentra en el nodo emisor de los datagramas multicast y en cuyas ramas se encuentran los nodos pertenecientes al grupo.

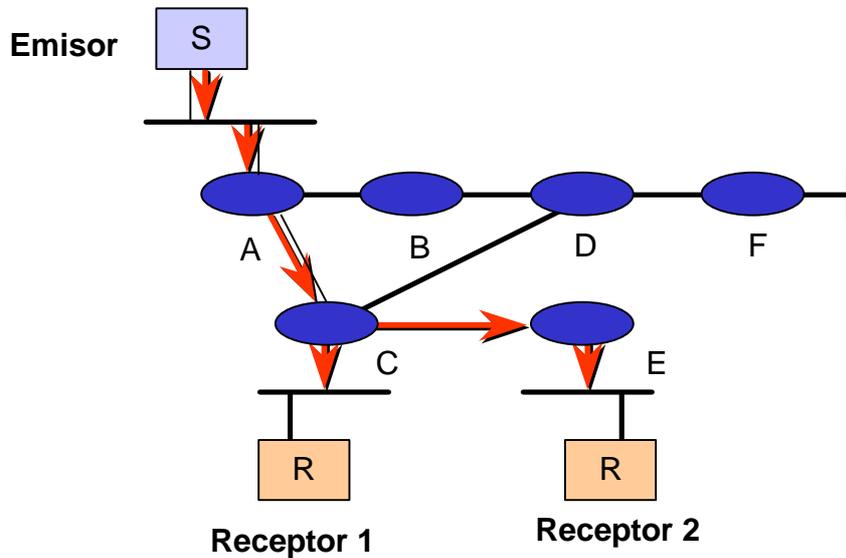


Figure 12.17 Árbol con raíz en el emisor

Sus características más sobresalientes son las siguientes:

- Es necesario que los routers dispongan de más memoria, del orden del producto del número de grupos por emisores.
- El camino desde el emisor es óptimo, lo que supone que se minimiza el retraso

De acuerdo con estas características resulta adecuada para situaciones en las cuales hay pocos emisores y muchos receptores, tales como aplicaciones de difusión de Radio

El árbol de distribución con raíz compartida se construye un árbol para todos los nodos emisores de datagramas multicast en el que se comparten las ramas en las cuales se encuentran los nodos receptores. Esto supone que todas las ramas del árbol serán compartidas a partir de un punto raíz del árbol que no está situado en ninguno de los nodos emisores.

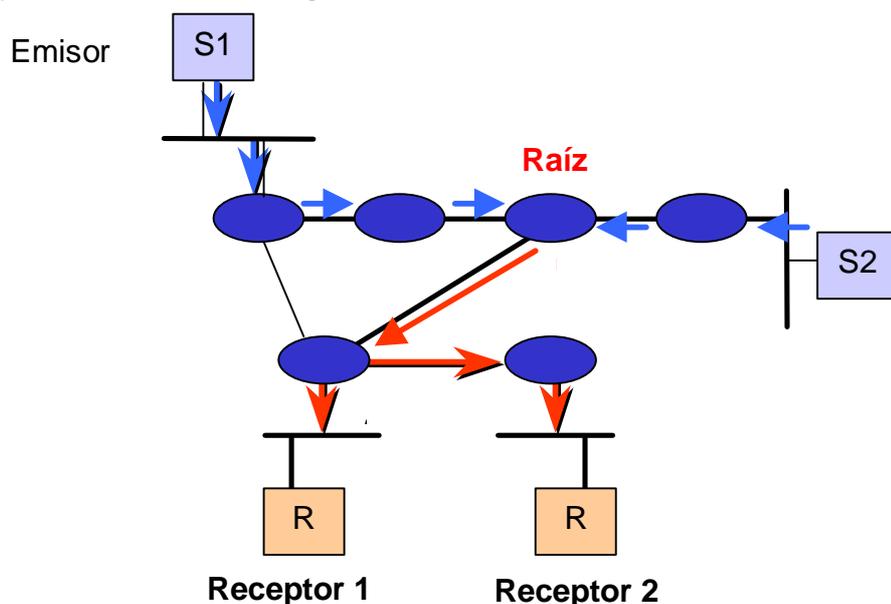


Figure 12.17 Árbol con raíz compartida

Sus características son las siguientes:

- La memoria necesaria en los routers es menor, del orden del número de grupos.
- Los caminos desde el emisor al receptor no son óptimos en todos los casos, por lo que pueden introducirse retrasos adicionales, generalmente aquellos que se producen desde el emisor a la raíz del árbol.
- Pueden producirse transferencias duplicadas por la duplicación de un camino desde el emisor a la raíz y de la raíz a los receptores.

De acuerdo con estas características resulta apropiado para situaciones en las que hay muchos emisores con poco ancho de banda y entornos en los que la mayor parte del árbol compartido es idéntico al árbol de distribución con raíz en el emisor.

El segundo criterio de clasificación tiene relación con la Política de Distribución de Datos, atendiendo al modo en que emisores y receptores se “encuentran”:

- Difundir los paquetes iniciales desde cada emisor y podar los segmentos sin ningún miembro: DVMRP, PIM-DM. También conocidos como de Tipo NACK. En este caso hay mucho tráfico inútil cuando hay pocos receptores en un área muy amplia, el motivo es que los primeros mensajes se distribuyen por toda el área hasta saber en qué segmentos hay nodos pertenecientes al grupo y en cuáles no y por lo tanto deben ser podados del árbol.
- Difundir las notificaciones de pertenencia desde cada receptor de toda la red: MOSPF
- Especificar un “punto de encuentro” al que los emisores envían sus paquetes iniciales, y los receptores se adhieren; precisa de una correspondencia entre direcciones de grupo multicast y el “punto de encuentro” . Ejemplos: CBT, PIM-SM. También se conocen como de Tipo ACK, y en ellos sólo se produce el envío de datagramas a los segmentos en los que los nodos explícitamente se han adherido al grupo. El rendimiento de este tipo es evidentemente superior al anterior, pero introduce una importante latencia por la propagación de las uniones y un coste elevado cuando hay muchos receptores como consecuencia de los mensajes de unión.

Una tercera forma de clasificación de los Protocolos de Routing Multicast IP sería la siguiente:

- Protocolos de modo denso. Suponen una pertenencia densa a grupos, es decir, que el número de nodos unidos a un grupo es elevado. Se basan en árboles de distribución con raíz en el emisor y son de tipo NACK. Los ejemplos más característicos son: DVMRP (Distance-Vector Multicast Routing Protocol), MOSPF (Extensiones Multicast a “Open Shortest-Path First Protocol”) y PIM-DM (Protocol-Independent Multicast, Dense Mode).
- Protocolos de modo disperso. Suponen pertenencia dispersa a grupos. Se basan en árboles compartidos y son de tipo ACK. Los ejemplos más representativos en este caso son: PIM-SM (Protocol-Independent Multicast, Sparse Mode) y CBT (Core-Based Trees).

12.7.2.- DVMRP: Distance Vector Multicast Routing Protocol

El protocolo Distance Vector Multicast Routing Protocol (DVMRP) fue definido originalmente en el RFC 1075, basándose en el protocolo RIP, con la diferencia de que RIP envía los datagramas unicast basándose en la información del siguiente salto hacia el destino, mientras que DVMRP construye árboles de distribución basándose en la información del salto previo hacia el emisor. La versión inicial de este algoritmo de routing vector-distancia construye los árboles basándose en el algoritmo TRPB. Las versiones posteriores de DVMRP, hasta la última 3.8, fueron mejoradas para utilizar RPM y son habitualmente utilizadas en los routers MBONE. La estandarización de la última versión de DVMRP está siendo llevada a cabo por el grupo de trabajo Inter-Domain Multicast Routing (IDMR) de la Internet Engineering Task Force (IETF).

Interfaces físicos y túneles

Los puertos de un router DVMRP pueden ser bien una interfaz física a una subred directamente conectada o una interfaz túnel otra isla multicast. Todos los interfaces se configuran con una métrica que especifica el coste para el puerto dado y un umbral TTL que limita el ámbito de la transmisión multicast. Además, cada interfaz túnel debe ser explícitamente configurada con dos parámetros adicionales, las direcciones IP de la interfaz del router local y de la interfaz del router remoto.

Initial TTL	Scope
0	Restricted to the same host
1	Restricted to the same subnetwork
32	Restricted to the same site
64	Restricted to the same region
128	Restricted to the same continent
255	Unrestricted in scope

Tabla 12.18 Valores de control de ámbito TTL

Un router multicast encaminará datagramas multicast a través de una interfaz sólo si el campo TTL de la cabecera IP es mayor que el umbral TTL asignado a la interfaz. La tabla 12.1 muestra los valores convencionales de TTL que se utilizan para restringir el ámbito de un multicast IP.

Por ejemplo, un datagrama multicast con un valor de TTL menor de 32 está restringido a la red de emisión y no sería encaminado a través de una interfaz a otros lugares de la misma región.

Operación básica

DVMRP implementa el algoritmo RPM, y por lo tanto, el primer datagrama de cualquier (emisor, grupo) es enviado a través de toda la red, siempre que el valor de TTL en el datagrama y el umbral de la interfaz del router lo permita. El datagrama inicial es enviado a todos los routers hojas, que transmiten mensajes de poda de vuelta hacia el emisor si no hay miembros del grupo en las subredes hojas directamente conectados a ellos. Los mensajes de poda hacen que se eliminen del árbol las ramas que no conducen a miembros del grupo, creando así un árbol de camino más corto específico para el emisor que contiene miembros del grupo en todas sus hojas. Después de un periodo de tiempo, las ramas podadas crecen de nuevo y el siguiente datagrama para el par (emisor, grupo) es enviado a través de toda la red, produciendo un nuevo conjunto de mensajes de poda.

DVMRP implementa un mecanismo para incorporar ("graft") rápidamente una rama previamente podada de un árbol de distribución. Si un router que ha enviado previamente un mensaje de poda para un (emisor, grupo) descubre nuevos miembros del grupo en una red hoja, envía un mensaje de "graft" a los routers del salto anterior del grupo. Cuando un router en el camino ascendente recibe un mensaje "graft", cancela el mensaje de poda recibido con anterioridad. Los mensajes "Graft" pueden producirse en cascada hacia el emisor, permitiendo restaurar como parte del árbol de distribución ramas previamente podadas.

Funciones de un router DVMRP

Cuando hay más de un router DVMRP en una subred, el Router Dominante es responsable de la transmisión periódica de mensajes IGMP "Host Membership Query". En la inicialización, un router DVMRP se considera Router Dominante para la subred hasta que recibe un mensaje "Host Membership Query" de un router vecino con una dirección IP menor.

Para evitar datagramas duplicados cuando hay más de un router DVMRP en una subred, se elige Router Dominante para la misma. En la figura 12.13, el Router C está más alejado ("downstream") del emisor y puede recibir potencialmente datagramas de la subred de emisión desde el Router A o el Router B. Si la métrica del Router A hacia la subred de origen de emisión es menor que la métrica del Router B, el Router A es dominante respecto del Router B para este emisor. Esto significa que el Router A encaminará tráfico desde la subred de emisión y el Router B descartará el tráfico procedente de dicha subred.

Sin embargo, si la métrica del Router A es igual a la de B, el router con la dirección IP más pequeña en su interfaz descendente (enlace hijo) será el Router Dominante.

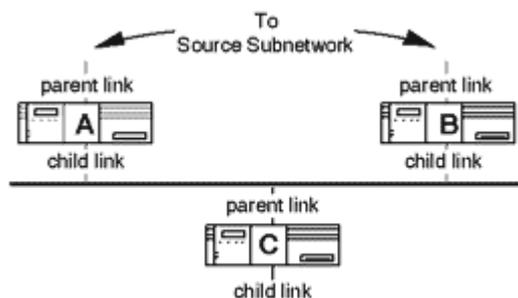


Figure 12.19: Router Dominante DVMRP

Tabla de routing DVMRP

Dado que DVMRP se desarrolló para encaminar datagramas multicast y no tráfico unicast, un router puede tener que mantener varios procesos de routing, uno para el envío de tráfico unicast y otro para el envío de tráfico multicast. El proceso DVMRP intercambia periódicamente mensajes de actualización de la tabla de encaminamiento con los vecinos con capacidad multicast. Estas actualizaciones son independientes de las generadas por cualquier protocolo IGP que soporte el routing unicast.

DVMRP se basa en la recepción de actualizaciones "poison reverse" para la detección de router hojas. Esta técnica precisa que un vecino que está por debajo notifique "infinito" para una subred de emisión al router en el salto previo del camino más corto hacia dicha subred de emisión. Si un router anterior ("upstream") no recibe una actualización "poison reverse" para la subred de emisión en la interfaz descendente, supondrá que la subred descendente es una hoja y elimina el puerto descendente de su lista de puertos de reenvío.

La figura 12.20 presenta un ejemplo de tabla de encaminamiento de un router DVMRP. A diferencia de las tablas de encaminamiento unicast, ésta contiene Subredes de Emisión y Saltos Previos en lugar de Subredes de Destino y Sigüientes Saltos. La tabla representa el árbol de expansión de camino más corto con raíz en el emisor para cada subred participante en el árbol RPB. La tabla DVMRP no considera la pertenencia de grupos o los mensajes de poda recibidos.

<u>Source Subnet</u>	<u>Subnet Mask</u>	<u>From Gateway</u>	<u>Metric</u>	<u>Status</u>	<u>TTL</u>	<u>InPort</u>	<u>OutPorts</u>
128.1.0.0	255.255.0.0	128.7.5.2	3	Up	200	1	2,3
128.2.0.0	255.255.0.0	128.7.5.2	5	Up	150	2	1
128.3.0.0	255.255.0.0	128.6.3.1	2	Up	150	2	1,3
128.4.0.0	255.255.0.0	128.6.3.1	4	Up	200	1	2

Figura 12.20 Tabla de routing DVMRP

Los elementos claves de la tabla de routing DVMRP son los siguientes:

- Source Subnet - Una subred que contiene un nodo que emite datagramas multicast
- Subnet Mask - Máscara de subred asignada a la subred de emisión.
- From-Gateway – El router anterior en el camino hacia la subred de emisión
- TTL – El valor time-to-live se utiliza para la gestión de la tabla e indica el número de segundos antes de que una entrada sea eliminada de la tabla.

Tabla de encaminamiento DVMRP

Como la tabla de routing DVMRP no contiene información sobre la pertenencia a grupos, el proceso DVMRP construye una tabla de encaminamiento, basada en una combinación de información procedente de la tabla de routing multicast, los grupos conocidos y los mensajes de poda recibidos. La tabla de encaminamiento representa el conocimiento de router local del árbol de distribución de camino más corto con raíz en el emisor para cada (emisor, grupo), es decir el árbol RPM.

<u>Source Subnet</u>	<u>Multicast Group</u>	<u>TTL</u>	<u>InPort</u>	<u>OutPorts</u>
128.1.0.0	224.1.1.1	200	1 Pr	2p 3p
	224.2.2.2	100	1	2p 3
	224.3.3.3	250	1	2
128.2.0.0	224.1.1.1	150	2	2p 3

Figure 12.21 Tabla de encaminamiento DVMRP

La Tabla de encaminamiento DVMRP que se muestra en la Figura 12.21, incluye los siguientes elementos:

- Source Subnet – La subred que contiene un nodo que emite datagramas multicast dirigidos a los grupos especificados.
- Multicast Group – La dirección IP de clase D a la que están dirigidos los datagramas multicast. Hay que tener en cuenta que una subred de emisión puede contener emisores para diferentes grupos multicast.
- InPort – El puerto “padre” para el par (emisor, grupo). Un valor "Pr" es esta columna indica que se ha enviado un mensaje de poda al router en sentido ascendente.
- OutPorts – Los puertos “hijo” por los que se envían datagramas multicast para el par (emisor, grupo). Un valor "p" en esta columna indica que el router ha recibido un mensaje de poda desde el router en sentido descendente.

DVMRP Jerárquico

El rápido crecimiento de MBONE ha producido un aumento en la capacidad de trabajo de los routers. La actual versión de DVMRP trata el MBONE como un solo y plano dominio de encaminamiento, donde cada router debe mantener información de encaminamiento detallada a cada subred del MBONE. A medida que continúa creciendo el número de subredes, el tamaño de las tablas de encaminamiento y de los mensajes de actualización periódicos continuarán creciendo hasta que la capacidad de memoria de los routers se agote y MBONE se colapse.

Para evitar esta amenaza se ha desarrollado una versión jerárquica de DVMRP, donde MBONE se divide en varios dominios de routing individuales. Cada dominio de routing ejecuta su propia instancia de un protocolo multicast. Otro protocolo, u otra instancia del mismo protocolo se utiliza para el routing entre los dominios individuales. El routing jerárquico reduce la demanda de recursos de los routers porque cada uno sólo precisa conocer los detalles explícitos de los destinos dentro de su propio dominio, pero nada de la estructura topológica detallada de cualquier otro de los dominios. El protocolo que opera entre los dominios individuales mantiene información sobre la interconexión de los dominios, pero no acerca de la topología interna de cada dominio.

Además de reducir la cantidad de información de routing, pueden conseguirse otros beneficios con la versión jerárquica de DVMRP:

- Los efectos del fallo de un enlace o un router se limitan a aquellos routers que operan dentro de dicho dominio. Además, los efectos de un cambio en la topología de interconexión se limitan sólo a los routers inter-dominio. Estas mejoras son especialmente importantes cuando el uso de un protocolo vector distancia puede producir tiempos de convergencia muy largos.
- El problema de la cuenta hasta infinito asociado con los protocolos vector distancia pone límites al diámetro máximo de la topología MBONE, mientras que el routing jerárquico limita este diámetro a un solo dominio y no al conjunto de MBONE.

Arquitectura jerárquica

El DVMRP jerárquico propone la creación de regiones de no intersección donde cada región tenga un identificador único. Los routers interiores de una región ejecutan cualquier protocolo de routing multicast (DVMRP, MOSPF, PIM, o CBT) como protocolo de "Nivel 1" (L1). Cada región debe tener al menos un router periférico que es responsable de proporcionar conectividad inter-regional. Los routers periféricos ejecutan DVMRP como protocolo de "Nivel 2" (L2) para enviar el tráfico entre regiones, como se muestra en la Figura 12.22.

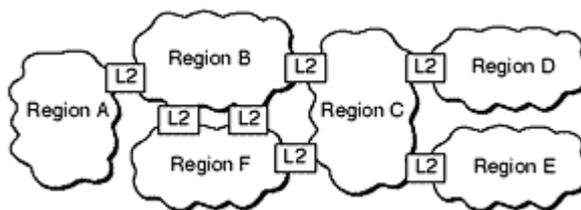


Figure 12.22 DVMRP jerárquico

Los routers L2 intercambian información de routing en forma de Identificadores de Regiones lugar de las direcciones de subred individuales contenidas dentro de cada región. Con DVMRP como protocolo L2, se construye un árbol de distribución multicast inter-regional basándose en el par (region_ID, grupo) en vez del par habitual (emisor, grupo).

Cuando un paquete multicast se origina dentro de una región, se envía de acuerdo al protocolo L1 a todas las subredes que contienen miembros del grupo. Además, el datagrama es enviada a cada uno de los routers periféricos (L2) configurados en la región de emisión. Los routers L2 etiquetan ("tag") el paquete con el Region-Id y lo colocan en una cabecera para su envío a otras regiones. Cuando el paquete llega a la región remota, los Routers L1 eliminan la cabecera antes de enviarlo a los miembros del grupo.

12.5.2.- MOSPF: Multicast Extensions to OSPF

Las extensiones multicast a OSPF (MOSPF) están definidas en RFC 1584 y se sitúan por encima de la versión 2 del protocolo Open Shortest Path First (OSPF) (RFC 1583). MOSPF utiliza la información de pertenencia a grupos, obtenida mediante IGMP y con la ayuda de la base de datos de OSPF construye árboles de distribución multicast. Estos árboles son los árboles de caminos más cortos construidos (bajo demanda) para cada pareja (emisor, grupo). Aunque MOSPF no soporta túneles puede coexistir e interoperar con routers no-MOSPF.

MOSPF soporta jerarquía de encaminamiento. Todos los nodos de la red son divididos en Sistemas Autónomos (AS) y cada AS es dividido a su vez en Areas. Vamos a estudiar a continuación cómo MOSPF lleva a cabo el routing multicast en estos tres niveles.

Routing Intra-Area

Routing Intra-Area describe el algoritmo de routing básico utilizado por MOSPF y que se ejecuta dentro de un área OSPF y soporta el envío multicast cuando el emisor y los miembros del grupo destino residen en el mismo área OSPF, o cuando el Sistema Autónomo completo es un área OSPF.

Base de Datos de Grupo Local

Del mismo modo que DVMRP, los routers MOSPF utilizan IGMP para monitorizar la pertenencia a grupos multicast en las subredes a las que están directamente conectados. Los routers MOSPF deben implementar una "base de datos de grupos locales" que mantenga una lista de los miembros de grupos directamente conectados y determine la responsabilidad del router local en el envío de datagramas multicast a dichos miembros de grupos.

En una subred determinada, la transmisión de Host Membership Queries IGMP se lleva a cabo sólo por el Router Designado (DR); además, la responsabilidad de escuchar los Host Membership Reports IGMP corresponde también al DR y al Router Designado de Backup (BDR). Esto supone que en un entorno mixto que contiene routers MOSPF y OSPF, debe elegirse un Router MOSPF como DR de la subred si deben generarse Queries IGMP. Esto se consigue asignando a todos los Routers no MOSPF un valor de RouterPriority de 0 para evitar que se conviertan en DR o BDR.

El DR es responsable de comunicar información de pertenencia de grupos a todos los demás routers en el área OSPF difundiendo LSAs Group-Membership. El DR origina un LSA Group-Membership para cada grupo multicast que tiene una o más entradas en la base de datos de grupos locales del DR. Del mismo modo que Router-LSAs y Network-LSAs, se difunden Group Membership-LSAs sólo a través un área. Esto garantiza que todos los datagramas multicast son enviados a la subred especificada para su distribución a los miembros del grupo locales.

Arbol del Camino más corto

El árbol del camino más corto describe el camino a seguir por un datagrama multicast al viajar por Internet desde la subred de emisión hasta cada uno de los miembros individuales del grupo. El árbol del camino más corto para cada par (emisor, grupo) se construye "bajo demanda" cuando un router recibe el primer datagrama multicast para dicho par.

Cuando llega el datagrama inicial, se ubica la subred de destino en la base de estado de enlace MOSPF. La base de estado de enlace MOSPF es implemente la base de estado de enlace OSPF con el añadido de los LSAs Group-Membership. Basándose en las Router-LSAs y Network-LSAs de la base de estado de enlace MOSPF, se construye un árbol de camino más corto con raíz en el emisor utilizando el algoritmo de Dijkstra's. Después de construir el árbol, se utilizan LSAs Group-Membership para podar las ramas que no conducen a subredes con miembros individuales de un grupo. El resultado del cálculo de Dijkstra es una poda del árbol de camino más corto con raíz en el emisor del datagrama.

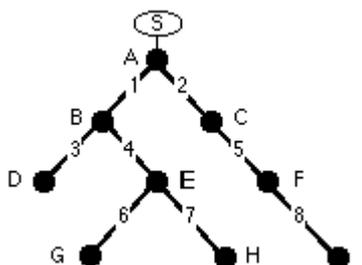


Figure 12.23 Shortest Path Tree for (S, G)

Para enviar un datagrama multicast hacia los miembros situados en la parte baja del árbol, cada router debe determinar su posición en el árbol de distribución de camino más corto. La figura 12.17 muestra el árbol de distribución de camino más corto para un par (emisor, grupo), en concreto el emisor S; El nodo más arriba del router E es el router B y en éste hay dos interfaces descendentes: un que le conecta a la Subred 6 y otro que le conecta a la subred 7.

El algoritmo de routing básico MOSPF tiene las siguientes propiedades:

- Para un datagrama multicast, todos los routers dentro de un área OSPF calculan el mismo árbol de distribución de camino más corto, garantizando que si existen varios caminos de igual costo, todos los routers acuerdan un mismo único camino a través del área. A diferencia del unicast OSPF, MOSPF no soporta el concepto de camino múltiple de igual costo.
- Las Bases de Datos sincronizadas que contienen LSAs Group-Membership permiten a un router MOSPF llevar a cabo de modo efectivo el cálculo de RPM. A diferencia de DVMRP, esto significa que el primer datagrama transmitido a un grupo no tiene que ser enviado a todos los routers del área.

La construcción "bajo demanda" del árbol de distribución de camino más corto tiene la ventaja de distribuir los cálculos en el tiempo, con un menor impacto en los routers participantes.

Caché de encaminamiento

Cada Router MOSPF toma sus decisiones de encaminamiento basándose en el contenido de la caché de encaminamiento, que es construida a partir del encaminamiento árbol de distribución de camino más corto para cada par (emisor, grupo) y la base de datos del grupo local del router.

Después de que el router descubre su posición en el árbol de distribución de camino más corto, se crea una entrada en la caché de encaminamiento con el par (emisor, grupo), el nodo ascendente y todas las interfaces descendentes. En este punto, se descarta el árbol de camino más corto liberando los recursos asociados con la creación del árbol. A partir de este punto, la caché de encaminamiento se emplea para encaminar todos los datagramas siguientes para el par (emisor, grupo).

<u>Destination</u>	<u>Source</u>	<u>Upstream</u>	<u>Downstream</u>	<u>TTL</u>
224.1.1.1	128.1.0.2	!1	!2 !3	5
224.1.1.1	128.4.1.2	!1	!2 !3	2
224.1.1.1	128.5.2.2	!1	!2 !3	3
224.2.2.2	128.2.0.3	!2	!1	7

Figure 12.24 Caché de encaminamiento MOSPF

La figura 12.24 muestra un ejemplo de caché de encaminamiento de un router MOSPF. Los elementos incluidos en ella son los siguientes:

- Destination – la dirección de grupo de destino a la que se encaminan los datagramas que casan.
- Source – La red de emisión del datagrama. Cada par Destino/emisor identifica una entrada de caché separada.
- Upstream – La interfaz desde el que debe recibirse un datagrama que case.
- Downstream – Los interfaces por los que un datagrama que case deberían ser encaminado para alcanzar los miembros del grupo de destino.
- TTL – El número mínimo de saltos que debe dar un datagrama para alcanzar los miembros del grupo multicast. Esto permite al router descartar datagramas que no tienen la posibilidad de alcanzar un miembro del grupo de destino.

La información de la caché de encaminamiento ni expira ni es refrescada periódicamente, y se mantiene mientras haya recursos disponibles en el sistema (i.e., memoria) o hasta el siguiente cambio de topología. En general, los contenidos de la caché cambian cuando:

- Cambia la topología de la red OSPF, forzando a que se recalculen todos los árboles de camino más corto.
- Hay un cambio en los LSAs Group-Membership indicando que la distribución de los miembros de un grupo ha cambiado.

Combinación de routers MOSPF y OSPF

Los routers MOSPF pueden combinarse con routers no-multicast OSPF. Esto permite un despliegue gradual de MOSPF y la experimentación con routing multicast a escala limitada. Cuando MOSPF y los routers no-multicast OSPF se mezclan dentro de un Sistema Autónomo, todos los routers interoperarán en el envío de datagramas unicast.

Es importante hacer notar que se precisa un router MOSPF para eliminar todos los routers no-multicast OSPF cuando se construye el árbol de distribución de camino más corto con raíz en el emisor. Un router MOSPF puede determinar con facilidad la capacidad multicast de otro basándose en el valor del bit multicast (MC-bit) del campo Options de cada notificación de estado de enlace de un router. La omisión de routers no-multicast puede crear numerosos problemas potenciales cuando se encamina tráfico multicast:

- Los datagramas multicast pueden enviarse a través de rutas subóptimas puesto que el camino más corto entre dos puntos puede precisar atravesar un router no-multicast OSPF.
- Incluso aunque haya conectividad unicast con un destino puede no haber conectividad multicast. Por ejemplo, la red puede estar dividida con respecto a la conectividad multicast dado que el único camino entre dos puntos atraviesa un router no-multicast OSPF.
- El envío de datagramas multicast y unicast entre dos puntos puede seguir caminos completamente diferentes a través de la red. Esto puede crear algunos problemas de routing difíciles de depurar.

- El Router Designado para una red multi-acceso debe ser un Router MOSPF. Si un router no-multicast OSPF se elige como DR, la subred no será seleccionada para enviar datagramas multicast puesto que un DR no-multicast no puede generar LSAs Group-Membership para su subred.

Inter-Area Routing

El routing Inter-area incluye el caso en que el emisor del datagrama y alguno de los miembros del grupo de destino se encuentren en áreas OSPF diferentes. Debe tenerse en cuenta que el encaminamiento de datagramas multicast continúa siendo determinado por la caché de encaminamiento que está aún construida a partir de la base de datos de grupos local y el árbol de camino más corto del datagrama. Las principales diferencias están relacionadas con el modo en que la información de pertenencia a un grupo se propaga y el modo en que se construye el árbol de camino más corto inter-area.

Encaminadores multicast Inter-Area

En MOSPF, un subconjunto de Routers periféricos de área (ABRs) operan como multicast "encaminadores multicast inter-area". Un encaminador multicast inter-area es responsable del encaminamiento de la información de pertenencia a un grupo y de difundir los datagramas entre áreas. Los parámetros de configuración determinan si un ABR determinado opera también o no como encaminador multicast inter-area.

Los encaminadores multicast inter-area resumen la información de pertenencia a grupos de sus áreas conectadas hacia el backbone originando nuevas LSAs Group-Membership en el área backbone. La sumariación de la pertenencia a grupos en MOSPF es asimétrica, lo que significa que esta información de áreas no-backbone se difunde hacia el backbone, pero el backbone no notifica la información de pertenencia a grupos del backbone o la información aprendida desde otras áreas no-backbone hacia ningún área no-backbone.

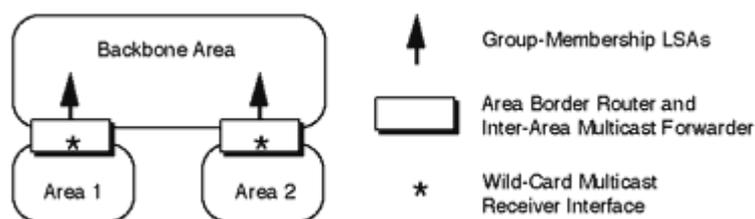


Figure 12.25 Arquitectura de routing Inter-Area

Para hacer posible el encaminamiento de tráfico multicast entre áreas, MOSPF introduce el concepto de un "receptor multicast comodín", que es un router que recibe todo el tráfico multicast generado en un área, independientemente de la pertenencia a grupos multicast. En áreas no-backbone, todos los encaminadores multicast inter-area operan de este modo. Esto garantiza que todo el tráfico multicast originado en un área no-backbone se envía a su encaminador multicast inter-area, y a continuación si es necesario al área backbone. Dado que el backbone tiene conocimiento de la pertenencia a grupos para todas las áreas, el datagrama puede ser encaminado hacia los miembros de grupo que se encuentran en el backbone y otras áreas no-backbone. El área backbone no precisa de receptores multicast comodín multicast porque los routers del área backbone tienen conocimiento completo de la información de pertenencia a grupos del sistema OSPF completo.

Árbol de Camino más corto Inter-Area

En el caso de routing multicast inter-area, a menudo resulta imposible construir un árbol de distribución de camino más corto completo. Se crean árboles incompletos porque no se distribuye entre áreas OSPF la información topológica detallada ni la información de pertenencia a grupos completa. Para superar estas limitaciones se realizan estimaciones topológicas mediante el uso de receptores comodín y LSAs Summary-Links OSPF.

Hay dos casos que resulta necesario considerar cuando se construye el árbol de distribución de camino más corto inter-área. El primero es que la subred de origen está situada en el mismo área que el router que efectúa el cálculo. El segundo se produce cuando la subred de emisión se localiza en un área diferente de aquella en que se encuentra el router que efectúa el cálculo.

Si el emisor de un datagrama multicast se encuentra en el mismo área que el router que efectúa el cálculo, el proceso de poda debe ser cuidadoso para asegurar que todas las ramas que conducen a otras áreas son podadas y eliminadas del árbol. Sólo se podan aquellas ramas que no contienen miembros de grupo ni receptores multicast comodín". Las ramas que contienen receptores multicast comodín" deben mantenerse, porque los routers locales no saben si hay miembros del grupo que residen en otras áreas.

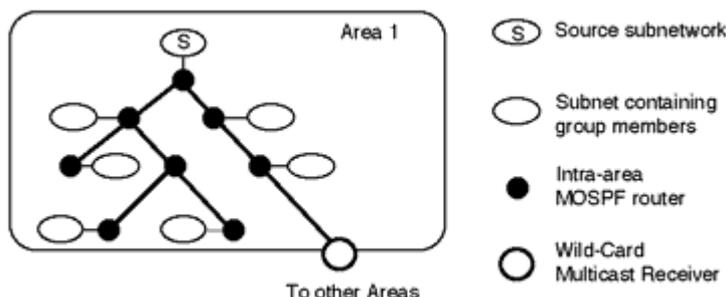


Figure 12.26 Árbol de camino más corto en el mismo Área

Si el emisor del datagrama multicast se encuentra en un área diferente de la del router que efectúa el cálculo, los detalles que describen la topología local que rodea a la estación emisora son desconocidos. Sin embargo, esta información puede ser estimada utilizando los LSAs Summary-Links de la subred de emisión. En este caso, la base del árbol comienza con ramas que conectan directamente a la subred de emisión con cada uno de los encaminadores multicast inter-área de las áreas locales. Los encaminadores multicast inter-área deben estar incluidos en el árbol, puesto que cualquier datagrama multicast originado fuera del área local entrará en el área por medio de un los encaminador multicast inter-área.

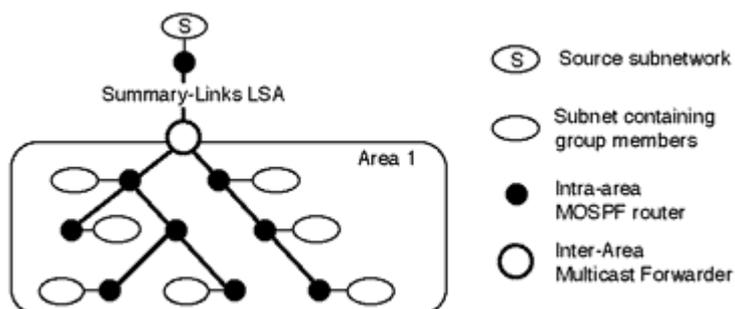


Figure 12.27 Árbol de camino más corto en un Área Remota

Dado que cada encaminador multicast inter-área es también un ABR, debe mantener una base de estado de enlace separada para cada área conectada. Esto significa que cada uno de ellos debe calcular un árbol de distribución separado para cada una de las áreas conectadas. Después de calculados los árboles individuales, son fundidos en una entrada caché de encaminamiento para el par (emisor, grupo) y se descartan los árboles individuales.

Routing Inter-AS

Los sistemas multicast Inter-Sistemas Autónomos contempla la situación en que el emisor del datagrama y al menos uno de los miembros del grupo de destino se encuentran en Sistemas Autónomos distintos. Para permitir el routing entre Sistemas Autónomos, se configuran algunos Routers Periféricos del Sistema Autónomo (ASBRs) como "encaminadores multicast inter-AS". MOSPF supone que cada encaminador multicast inter-AS ejecuta un protocolo de routing multicast inter-AS (como DVMRP), que encamina los datagramas multicast de modo (RPF). Cada encaminador

multicast inter-AS opera como un receptor multicast comodín” en cada una de sus áreas conectadas. Esto garantiza que cada encaminador multicast inter-AS permanece en todos los árboles de camino más corto y recibe todos los datagramas multicast, independientemente de la pertenencia a los grupos.

Es necesario considerar tres casos al describir la construcción de un árbol de distribución de camino más corto inter-AS. El primero se produce cuando la subred emisora se encuentra en el mismo área que el router que efectúa el cálculo. El segundo caso, cuando la subred de emisión se encuentra en un área diferente que el router que efectúa el cálculo y finalmente, cuando la subred de emisión está situada en un AS diferente que el router que efectúa el cálculo.

Los primeros dos casos son similares a los ejemplos inter-area descritos anteriormente. La única mejora es que los encaminadores multicast inter-AS deben ser incluidos también los árboles de distribución de camino más corto podados. Las ramas que contienen encaminadores multicast inter-AS deben ser retenidas, puesto que los routers locales no saben si hay miembros de un grupo residiendo en otros Sistemas Autónomos. Cuando llega un datagrama multicast a un encaminador multicast inter-AS, es responsabilidad del ASBR determinar si debería encaminarse el datagrama fuera del Sistema Autónomo local. La Figura 12.28 muestra un ejemplo de árbol de distribución de camino más corto inter-AS cuando la subred de emisión reside en el mismo área que el router que efectúa el cálculo.

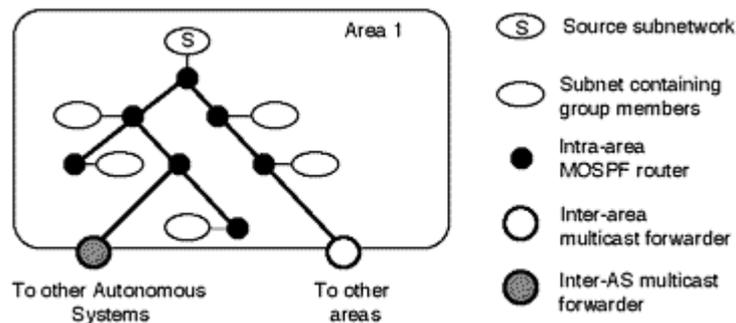


Figure 12.28 Árbol de distribución de camino más corto inter-AS en el mismo Área

Si el emisor del datagrama multicast reside en un Sistema Autónomo diferente que el router que efectúa el cálculo, los detalles que describen la topología local alrededor de la estación emisora son desconocidos. Sin embargo, esta información puede estimarse utilizando los AS-External Links que describen la subred de emisión. En este caso, la base del árbol comienza con ramas directamente conectadas a la subred de emisión y a cada uno de los encaminadores multicast inter-AS.

La Figura 12.29 muestra un ejemplo de árbol de distribución de camino más corto inter-AS cuando el encaminador multicast inter-AS reside en el mismo área que el router que realiza el cálculo, la topología alrededor del emisor se aproxima combinando los Summary-ASBR Link con los AS-External Link.

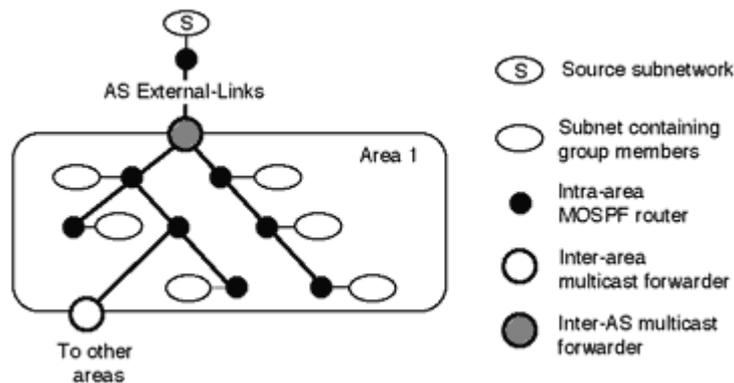


Figure 12.29 Árbol de distribución de camino más corto inter-AS en diferentes AS

Para finalizar, es importante resaltar que los AS External Links no son importados en las áreas Stub. Si el emisor está ubicado fuera del área stub, la topología en las proximidades del emisor se estima por los Default Summary Links originados por los encaminadores multicast intra-área de las áreas stub en lugar de por los AS-External Links.

12.7.1.- PIM: Protocol-Independent Multicast

Los protocolos PIM (Protocol Independent Multicast) han sido desarrollados por el grupo de trabajo Inter-Domain Multicast Routing (IDMR) de la IETF. IDMR se propuso desarrollar un conjunto de protocolos de routing multicast que independientes de cualquier protocolo de routing unicast pudieran proporcionar routing multicast escalable a nivel de Internet. Desde luego, PIM precisa la existencia de un protocolo de routing unicast para actualizar la tabla de información de encaminamiento y adaptarse a los cambios de topología.

La mayoría de los protocolos multicast propuestos funcionan bien si los miembros de los grupos están densamente distribuidos y el ancho de banda no es un problema. Sin embargo, el hecho de que DVMRP inunde periódicamente la red y de que MOSPF envíe información de pertenencia a grupos a través de los enlaces, hace que estos protocolos no resulten eficientes en casos donde los miembros de los grupos se distribuyen de forma dispersa en distintas regiones de la red y donde el ancho de banda no es inagotable.

Para solucionar estos problemas, PIM contempla dos protocolos: PIM - Modo Denso (PIM-DM) que es más eficiente cuando los miembros están distribuidos densamente y PIM - Modo Disperso (PIM-SM) que tiene mejor rendimiento en los casos en que los miembros están dispersamente distribuidos. Aunque estos dos algoritmos pertenecen a PIM y comparten mensajes de control similares, son esencialmente protocolos distintos. El modo denso se refiere a un protocolo diseñado para operar en un entorno en el que los miembros del grupo están agrupados densamente y con un ancho de banda importante. El modo no denso se refiere a un protocolo optimizado para entornos donde los miembros de los grupos están distribuidos en muchas regiones de Internet y el ancho de banda disponible no es necesariamente muy grande. Es importante tener en cuenta que el modo-disperso no implica que el número de miembros sea pequeño sino que estos se encuentran dispersos a través de Internet.

PIM-DM (Protocol-Independent Multicast - Dense Mode)

Aunque la arquitectura PIM fue concebida por la necesidad de proporcionar árboles de distribución de modo disperso escalables, también define un protocolo nuevo en modo-denso en lugar de los protocolos ya existentes como DVMRP y MOSPF. Se considera que PIM-DM se desplegará en entornos ricos en recursos, tales como un campus LAN donde la pertenencia a un grupo es relativamente densa y el ancho de banda disponible es importante.

PIM-DM es muy similar a DVMRP y utiliza el algoritmo RPM para formar los árboles de distribución. Si embargo, hay diferencias sustanciales entre estos dos protocolos. Aunque PIM-DM precisa la existencia de un protocolo de routing unicast para encontrar rutas hacia el nodo emisor, PIM-DM es independiente de los mecanismos utilizados por cualquier protocolo de routing unicast. DVMRP emplea intercambio de mensajes de tipo RIP para construir su tabla de encaminamiento unicast, y MOSPF se basa en la Base de Datos de estado de enlace OSPF, lo que constituye su principal diferencia con PIM-DM.

La otra diferencia entre PIM-DM y DVMRP es que PIM-DM envía datagramas multicast por todas las interfaces descendentes hasta que recibe un mensaje de poda, mientras que DVMRP envía tráfico multicast a los nodos hijo en el árbol de expansión. Por lo tanto, resulta obvio que PIM-DM necesita gestionar mensajes duplicados. Sin embargo, este método es elegido para eliminar dependencias del protocolo de routing y evitar la sobrecarga originada por el cálculo de las interfaces hijos de cada router. Para aquellos casos en los que aparecen miembros de un grupo en una rama podada del árbol de distribución, PIM-DM, al igual que DVMRP, utiliza mensajes "graft" para volver a unir al árbol de distribución ramas previamente podadas. Finalmente, en PIM-DM, el procesamiento de mensajes de control y el encaminamiento de paquetes de datos están integrados con la operación de PIM-SM operación de modo que un solo router puede ejecutar diferentes modos para diferentes grupos.

PIM-SM (Protocol-Independent Multicast - Sparse Mode)

PIM Sparse Mode (PIM-SM) ha sido desarrollado como un protocolo de routing multicast que proporcione una comunicación eficiente entre miembros de un grupo distribuidos de modo disperso, por lo demás, el tipo de grupos más común en redes extensas. Sus diseñadores creyeron que una situación en la que varios nodos desean participar en una conferencia multicast no justifica la difusión del tráfico multicast en toda la red, y se temían que los protocolos de routing multicast existentes experimentarían problemas de escalado si se producían varios miles de pequeñas conferencias simultáneamente, creando gran cantidad de tráfico agregado que podría, potencialmente, saturar las conexiones de Internet. Para eliminar estos problemas potenciales de escalado, PIM-SM ha sido diseñado para limitar el tráfico multicast de modo que sólo aquellos routers interesados en recibir el tráfico dirigido a un grupo particular lo "vean".

PIM-SM está definido en RFC 2117, y presenta dos diferencias fundamentales con los protocolos de modo denso (DVMRP, MOSPF, and PIM-DM). Los routers que emplean el protocolo PIM-SM Protocol necesitan anunciar explícitamente su deseo de recibir datagramas multicast de grupos multicast, mientras que los protocolos de modo denso asumen que todos los routers necesitan recibir datagramas multicast a menos que explícitamente envíen un mensaje de poda. La segunda diferencia es el concepto de "core" o "rendezvous point" (RP) que ha sido empleado en el protocolo PIM-SM.

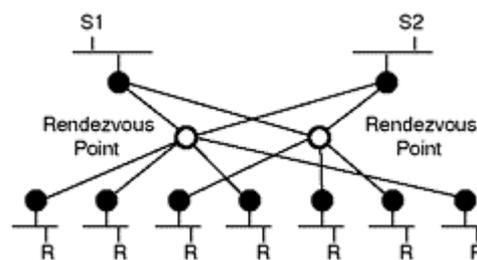


Figure 12.30 Rendezvous Points

PIM-SM tiene un enfoque similar al árbol Core-Based Tree (CBT) en el sentido de que emplea el concepto de rendezvous point (RP) donde los receptores "se encuentran" con los emisores. El iniciador de cada grupo multicast selecciona un RP y un pequeño conjunto ordenado de RPs alternativos, conocido como la lista-RP. Para cada grupo multicast hay un solo RP activo. Cada receptor que desea unirse a un grupo multicast contacta con su router directamente conectado, que por turno se une al árbol de distribución multicast enviando un mensaje explícito de adhesión al RP primario del grupo. El emisor utiliza el RP para anunciar su presencia y encontrar un camino a los miembros que se han unido al grupo. Este modelo precisa que los routers en modo disperso mantengan algún estado (i.e., la lista-RP) antes de la llegada de los datagramas. Por otro lado, los protocolos de routing multicast en modo disperso son dirigidos por los datos, puesto que no definen ningún estado para un grupo multicast hasta que llega el primer paquete de datos.

Un nodo directamente conectado se une a un grupo

Cuando hay más de un router PIM conectado a una LAN multiacceso, el router con la dirección IP más elevada es seleccionado para operar como Router Designado (DR) para la LAN. El DR es responsable de la transmisión de los mensajes IGMP Host Query, de enviar mensajes Join/Prune hacia el RP, y de mantener el status del RP activo para los emisores locales hacia grupos multicast.

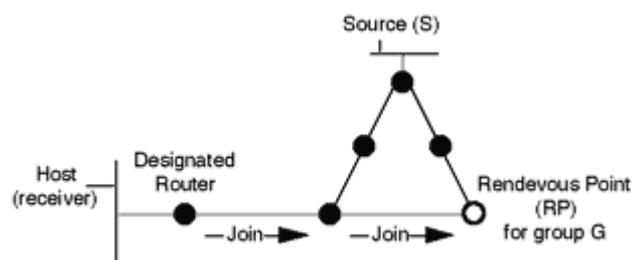


Figure 12.31 Unión de un nodo a un grupo multicast

Para facilitar la diferenciación entre grupos DM y SM, una parte del espacio de direcciones multicast de Clase D ha sido reservado para su uso por grupos SM. Cuando el DR recibe un mensaje IGMP Report para un nuevo grupo, el DR determina si el grupo está basado en RP examinando la dirección de grupo. Si la dirección indica un grupo SM, el DR lleva a cabo una búsqueda en la lista-RP asociada al grupo para determinar el RP primario de dicho grupo. La especificación describe un procedimiento para la elección del RP primario y el uso de RPs alternativos si el RP primario no está accesible.

Después de llevar a cabo la búsqueda, el DR crea una caché de encaminamiento multicast para el par (*, grupo) y transmite un mensaje unicast PIM-Join al RP primario; La notación (*, grupo) indica (cualquier fuente, grupo). Los routers intermedios envían el mensaje unicast PIM-Join y crean una caché de encaminamiento para el par (*, grupo). Los routers intermedios crean dicha entrada caché para saber cómo encaminar el tráfico dirigido al par (*, grupo) descendentemente hacia el DR que originó el mensaje PIM-Join.

Un emisor directamente conectado envía datagramas a un grupo

Cuando un emisor envía un datagrama multicast a cierto grupo, su DR debe encaminar el datagrama al RP primario para su distribución posterior a través del árbol de distribución del grupo. El DR de dicho emisor encapsula el primer datagrama en un paquete "PIM-SM-Register" y lo envía en modo unicast al RP de dicho grupo como un datagrama unicast. El mensaje PIM-SM-Register informa al RP de un nuevo emisor, lo que activa al RP para transmitir un mensaje PIM-Join al DR del emisor. Este intercambio se ha ilustrado en la Figura 12.32.

Mientras este mensaje es encaminado, los routers que se encuentran entre el DR del emisor y el RP mantienen estado de los mensajes PIM-Join recibidos, añadiendo una nueva entrada en sus tablas de encaminamiento multicast para la nueva pareja (emisor, grupo), de este modo sabrán cómo enviar los siguientes datagramas multicast no encapsulados desde la subred de emisión al RP. Obviamente, el RP será responsable de enviar estos datagramas multicast a los miembros del grupo. Hasta que estas entradas no hayan sido añadidas en las tablas de todos los routers intermedios, todos los datagramas multicast serán enviados como mensajes unicast encapsulados.

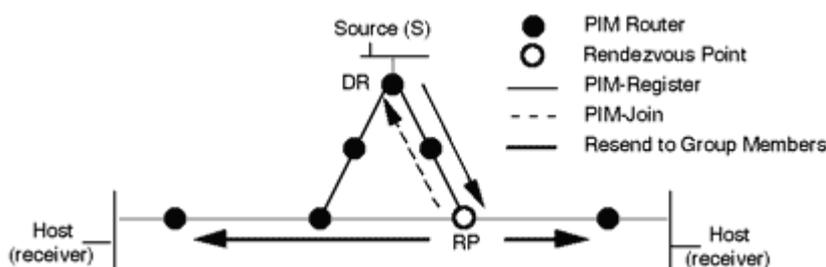


Figure 12.32 Envío desde un emisor a un grupo multicast

El DR de emisor cesa de encapsular los paquetes de datos en PIM-SM-Registers cuando recibe un mensaje Join/Prune desde el RP. En este punto, el tráfico de datos es enviado por el DR en su formato multicast nativo al RP. Cuando el RP recibe paquetes multicast desde el emisor, reenvía los datagramas por el árbol compartido-RP a todos los miembros del grupo en sentido descendente.

Arbol RP-Compartido o Arbol de camino más corto (SPT)

Aunque el envío de datagramas multicast a través de un árbol RP compartido es suficiente, si el número de participantes (o de datagramas siendo transmitidos a través de este árbol compartido) aumenta, el uso del mismo árbol compartido puede no ser muy deseable, puesto que los árboles RP compartidos no optimizan los caminos de envío a través de la red. PIM-SM proporciona un método para utilizar árboles con el camino más corto para algunos o todos los receptores. Los routers PIM-SM pueden seguir utilizando árboles RP, pero tienen la opción de utilizar árboles de camino más corto con raíz en el emisor creados por los receptores conectados. El árbol de camino más corto permite al miembro de un grupo reducir el retraso entre él y un emisor particular.

Un router PIM con receptores locales tiene la opción de conmutar al árbol de camino más corto al emisor tan pronto como comienza a recibir datagramas desde la estación emisora. El cambio puede ser producido cuando los datos de una estación emisora exceden de un umbral predeterminado. El DR del receptor local efectúa este cambio enviando un mensaje Join hacia el emisor activo. Al mismo tiempo, los mecanismos del protocolo garantizan que se transmite un mensaje de poda para el mismo emisor hacia el RP activo. Alternativamente, el DR puede ser configurado para continuar utilizando el árbol basado en RP y nunca conmutar al árbol de camino más corto al emisor. La figura 12.33 muestra tanto el Arbol RP-compartido como el Arbol de Camino más Corto hacia el emisor S.

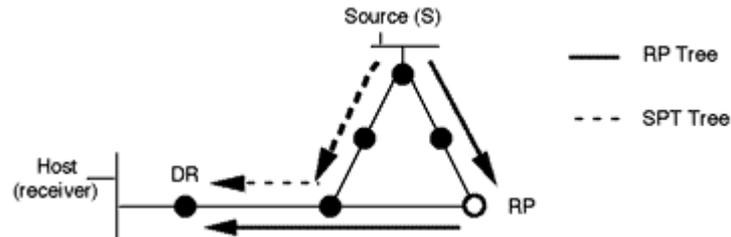


Figure 12.33: Arbol RP-compartido y Arbol de Camino más Corto

12.- Multicast.....	1
12.1.- Introducción.....	1
12.2.- Modelo de Servicio IP Multicast (RFC-1112)	3
12.2.1.- Direcciones IP Multicast.....	3
12.2.2.- Traducción de direcciones IP de Clase D a direcciones MAC.....	4
12.3.- IGMP	5
12.4.- Mbone.....	7
12.5.- Routing IP Multicast.....	8
12.6.- Algoritmos de routing multicast	9
12.6.1.- Inundación.....	9
12.6.2.- Árboles de expansión	9
12.6.3.- Reverse Path Broadcasting (RPB).....	10
12.6.4.- Truncated Reverse Path Broadcasting (TRPB)	11
12.6.5.- Reverse Path Multicasting (RPM)	11
12.6.6.- Árboles de Steiner (ST)	12
12.6.7.- Core-Based Trees (CBT).....	13
12.7.- Protocolos de routing multicast	13
12.7.1.- Clasificación de los protocolos de Routing Multicast	13
12.7.2.- DVMRP: Distance Vector Multicast Routing Protocol	15
12.7.1.- PIM: Protocol-Independent Multicast.....	25