

Algoritmo de balanceo de carga aplicado a enrutamiento multicast*

Yezid Donoso Meisel**, Iván Galera***

Grupo de Investigación en Redes de Computadores
Departamento de Ingeniería de Sistemas y Computación
Universidad del Norte, Barranquilla (Colombia)

Ramón Fabregat****

Grupo de Comunicaciones y Sistemas Distribuidos
Instituto de Informática y Aplicaciones
Universidad de Girona (España)

Resumen

En este artículo se proponen las bases para la creación del protocolo de enrutamiento multicast con características de ingeniería de tráfico: Pim-Dm/Sm-Te. Este protocolo consistiría en una serie de modificaciones realizadas al protocolo multicast Pim-Dm/Sm, para permitir que este protocolo realice balanceo de carga y sirva como fuente de información a protocolos de reserva de recursos.

Inicialmente se presentan definiciones y conceptos necesarios para la comprensión del funcionamiento del protocolo, tales como enrutamiento multicast en modo denso y enrutamiento multicast en modo esparcido. Se prosigue con la definición y explicación del funcionamiento de Pim-Dm/Sm, seguida de la propuesta y explicación de su funcionamiento. Por último, el lector tendrá acceso a resultados obtenidos a través de la herramienta de simulación, que confirman la utilidad del protocolo.

Palabras clave: Balanceo de carga, ingeniería de tráfico, multicast, optimización.

Abstract

In this paper the bases for the creation of the multicast routing protocol with traffic engineering characteristics (Pim-Dm/Sm-Te) is proposed. This protocol would consist of a series of modifications made to Pim-Dm/Sm multicast protocol to allow that this protocol makes load balancing and that this protocol serves like a source of information to resource reservation protocols. Initially, necessary definitions and

Fecha de recepción: 12 de julio de 2003
Fecha de aceptación: 5 de enero de 2004

* Este artículo forma parte de los resultados de la investigación «Aplicación de Ingeniería de tráfico en redes en Multicast».

** Ingeniero Sistemas y Minor en Gestión de Proyectos de Ingeniería, Universidad del Norte; Magíster en Ingeniería de sistemas y computación, Universidad de los Andes; D.E.A. y candidato a grado de Ph.D. en Redes Telemáticas, Universidad de Girona (España). Profesor del Departamento de Ingeniería de Sistemas, Universidad del Norte. ydonoso@uninorte.edu.co

*** Ingeniero Sistemas y Especialista en Redes de Computadores, Universidad del Norte.

**** Ingeniero en Informática, Universidad Autónoma de Barcelona; Ph.D. en Informática, Universidad de Girona (España). Profesor titular de esta misma universidad.

concepts for the understanding of the operation of the protocol will appear, such as routing multicast in dense mode and routing multicast in sparse mode. One continues with the definition and explanation of the operation of Pim-Dm/Sm, followed of the proposal and explanation of his operation. Finally, the reader will have access to results obtained through of the simulation tool, that confirm the utility of the protocol.

Key words: Load Balancing, Traffic engineering, multicast, optimization.

1. INTRODUCCIÓN

En la actualidad, uno de los objetivos más importantes de las ciencias de la computación, en el área de las comunicaciones, es la transmisión eficaz de voz y video, teniendo en cuenta ciertos requerimientos de retardo y ancho de banda, es decir, calidad de servicio.

Sin embargo, los protocolos de enrutamiento actuales no se ajustan correctamente a la estructura creada para este propósito, los cuales no son compatibles con MPLS y no soportan ingeniería de tráfico. Como consecuencia de lo anterior, es poca la calidad que en general puede asegurarse para este tipo de transmisión.

Para abordar este problema se hace necesario un proceso de adecuación de los protocolos actuales a estos nuevos requerimientos.

En este artículo se presenta una propuesta que se convierte en una herramienta para la solución de este problema: el protocolo de enrutamiento multicast PIM-DM/SM-TE. Este protocolo representa una mejora realizada al protocolo mutlicast PIM en sus versiones DM y SM. Se presenta inicialmente un marco teórico; posteriormente se continúa con la propuesta de los algoritmos de reserva de recursos y balanceo de carga a través de múltiples caminos, y finalmente se presentan los resultados que confirman la eficacia del protocolo.

2. PROTOCOLO DE ENRUTAMIENTO MULTICAST PIM-DM/SM

Llamamos multicast a una de las tres diferentes formas de direccionamiento IP (unicast, broadcast, multicast). Este direccionamiento se utiliza para enviar paquetes a un conjunto específico de receptores en una red, y que pertenecen a un mismo grupo multicast. El enrutamiento multicast puede ejecutarse en dos modos. El primero de ellos, modo denso, recibe su nombre debido a que se diseñó asumiendo que los receptores se encuentran distribuidos en forma densa en la red y se cuenta con amplio ancho de banda. Por el contrario, el

modo esparcido se diseñó para ambientes en los cuales los miembros del grupo se distribuyen en forma no densa (o esparcida) a través de la red. Este es el tipo de grupo que podríamos encontrar en una red MAN o WAN [8] [9].

PIM (*Protocol Independent Multicast*) es un protocolo de enrutamiento multicast, cuyo diseño no se basa en ningún otro protocolo unicast existente. Este protocolo no adquiere ni distribuye métricas y requiere de otro protocolo unicast capaz de descubrir la topología de la red.

Existen dos versiones de este protocolo, la versión en modo denso DM, y la versión en modo esparcido SM. En este artículo se utiliza la notación DM/SM para hacer referencia a los dos modos a la vez, y no debe confundirse con un tercer modo de operación inexistente [8] [9].

2.1. PIM-DM

PIM-DM asume inicialmente que todos los miembros de la red son receptores de tráfico multicast y transmite paquetes a toda la red. Cada enrutador que reciba dicho tráfico pero no lo requiera enviará un mensaje Prune a través de la interfaz por la cual lo recibió. Este proceso es conocido como poda del árbol.

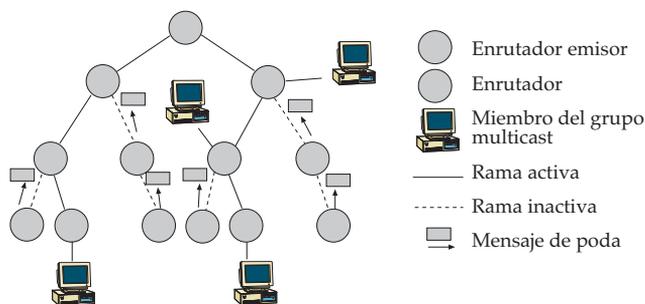


Figura 1. Poda en PIM-DM

Este estado de poda no es infinito en el tiempo. Cuando se cumple el tiempo de vida de este estado se elimina y el tráfico vuelve a transmitirse. Este proceso es conocido como ciclo de inundación y poda, y es común en los protocolos densos [8].

2.2. PIM-SM

PIM-SM difiere de PIM-DM principalmente en que sus usuarios deben informar al protocolo su necesidad de pertenecer a un grupo multicast. Además PIM-SM

utiliza un enrutador RP (*Rendezvous Point*), el cual recibe los paquetes multicast de la fuente y luego los reenvía a los destinos del grupo a través de un árbol de distribución compartido [9].

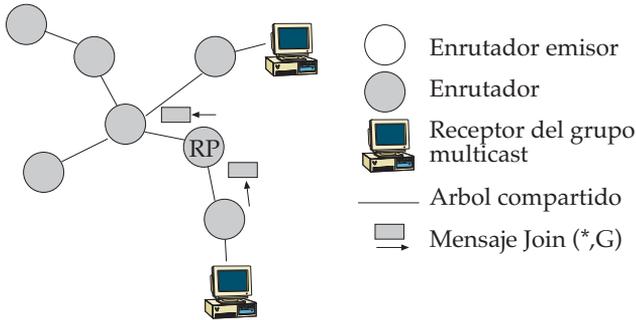


Figura 2. Unión de un nodo destino a un RP en PIM-SM

3. PIM-DM/SM-TE

Antes de presentar el funcionamiento del protocolo propuesto en este artículo PIM-DM/SM-TE se presentan los algoritmos que se diseñaron para soportar las características de ingeniería de tráfico, tales como reserva de ancho de banda y balanceo de carga a través de múltiples rutas.

3.1. Algoritmo de reserva de ancho de banda

Este algoritmo reserva ancho de banda de las rutas disponibles, en orden de calidad de las rutas, hasta ofrecer el ancho de banda necesario al flujo o hasta agotar los recursos.

Variables utilizadas en este algoritmo:

- $BW_{disp\ total}$ = Suma total de ancho de banda disponible de las rutas
- D_{total} = Tamaño del flujo
- $D_{asignado}$ = Tamaño asignado del flujo
- $\sum A_i$ = Suma de aportes de todas las rutas
- N = Número de rutas

Arreglos: El arreglo $A(i)$ debe encontrarse ordenado ascendentemente, y los demás arreglos deben ordenarse paralelamente a $A(i)$.

$BW(i)$ = Ancho de banda disponible para reservar de cada ruta

$D(i)$ = Ancho de banda asignado para el flujo en cada ruta
 $A(i)$ = Aporte de cada ruta

```

Algoritmo RAB;
 $D_{\text{asignado}} = 0$ 
Mq  $((D_{\text{total}} - D_{\text{asignado}}) > 0) \wedge (i \notin N)$ 
  Si  $((D_{\text{total}} - D_{\text{asignado}}) > BW(i))$ 
     $D(i) \leftarrow BW(i)$ 
  sino
     $D(i) \leftarrow D_{\text{total}} - D_{\text{asignado}}$ 
  fin si
   $D_{\text{asignado}} \leftarrow D_{\text{asignado}} + D(i)$ 
   $i \leftarrow i + 1$ 
fin mq
fin Algoritmo
  
```

Figura 3. Algoritmo de reserva de ancho de banda

3.2. Algoritmo de balanceo de carga

Este algoritmo envía el flujo por la mejor ruta disponible, hasta que la utilización de la misma alcanza un valor límite de la reserva realizada, cuyo valor de reserva es dado por el algoritmo presentado en la sección 3.1. Cuando se llega al valor de reserva, el flujo se envía por otras rutas en su orden de prioridad. Cuando estas rutas alcanzan también el valor límite, se envía un bloque de información en forma proporcional por cada una de las rutas teniendo en cuenta la capacidad de transmisión que posee cada uno. Al terminar de enviar este bloque se intenta enviar de nuevo, a través de la mejor de las rutas disponibles, sin superar el límite de utilización. Si no es posible, se continúa enviando proporcionalmente.

Variables:

N = Número de rutas
 P = Paquete a enviar
 TB = Tamaño del bloque que se enviará utilizando balanceo de carga
 Balanceo = Variable booleana que indica si ya se ha iniciado el proceso de balanceo de carga
 $\sum A_i$ = Suma de aportes de todas las rutas

Arreglos:

El arreglo A (i) debe encontrarse ordenado ascendentemente, y los demás arreglos deben ordenarse paralelamente a A (i)

BW (i) = Ancho de banda de cada ruta

L (i) = Capacidad disponible máxima de canal para utilización del flujo antes de iniciar balanceo de carga

CB (i) = Cantidad enviada por la ruta i utilizando balanceo de carga

A (i) = Aporte de cada ruta

Subrutinas:

Enviar (i, p): Es la función que envía el paquete p a través del canal i. No devuelve resultados.

Siguiente (): Esta función retorna el siguiente paquete que se va a enviar, o mantiene el algoritmo en espera hasta que sea creado.

Tam (p): Retorna el tamaño de un paquete p.

Util (i): Devuelve la cantidad de reserva de canal usada actualmente por el flujo en la ruta i.

```
Algoritmo BC;
p= Siguiente ( )
Mq (true)
  i ← 0
  haga
    i ← i + 1
  Hasta (i > N)/(Util (i) < L (i))
  Si (i ∈ N)
    Enviar (i, p)
    p= Siguiente ( )
  sino
    Para i = 1 hasta N
      CB (i) ← 0
    fin para
    i ← 1
    Mq (i ∈ N)
      Enviar (i, p)
      p= Siguiente ( )
      CB (i) = CB (i) + Tam (p)
      Si (CB (i) - TB ´ A (i) /SAi )
        i ← i + 1
      fin si
    fin mq
  fin si
fin mq
fin Algoritmo
```

Figura 4. Algoritmo de balanceo de carga

3.3. Funcionamiento del protocolo

3.3.1. Requerimientos

PIM-DM/SM-TE almacenará información de las diferentes rutas por las que es posible transmitir la información, pero no distribuirá métricas de toda la topología de la red. Por esta razón, necesitará un protocolo unicast con TE tal como OSPF-TE o IS-IS-TE que sea capaz de hacerlo.

3.3.2. Descubrimiento de rutas y proceso de Join

PIM-DM/SM-TE debe poder utilizar más de una ruta como alternativa para llevar a cabo el balanceo de carga y de esta forma evitar la congestión. A continuación se presenta el procedimiento que utilizaría el protocolo para conocer las posibles rutas a cada destino.

■ Pre-Join. Detección de rutas

El enrutador destino enviará un mensaje Pre-JoinRq unicast dirigido a la fuente en modo DM, o al RP en modo SM. El enrutador fuente responderá con un mensaje Pre-Join a través de cada una de sus interfaces. El mensaje Pre-Join almacenará la lista de todos los enrutadores que recorra antes de llegar al enrutador destino, la métrica unicast de la ruta y el menor ancho de banda disponible de los enlaces de la ruta. El mensaje no puede llegar dos veces al mismo enrutador intermedio. De ser así, los mensajes repetidos serán descartados por el nodo intermedio.

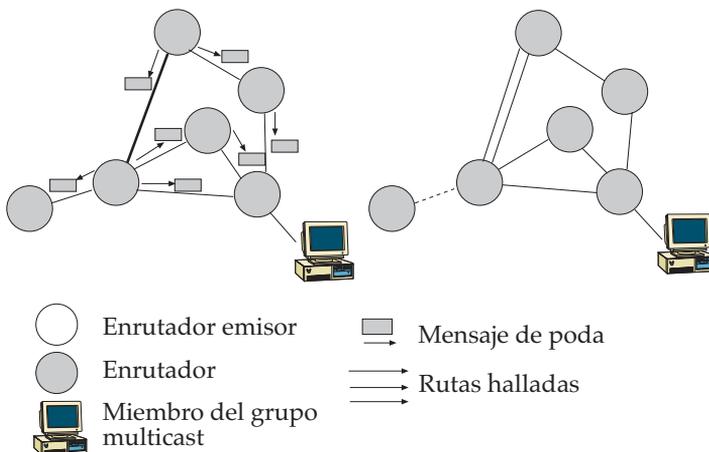


Figura 5. Pre-Join

Cuando el mensaje llegue al enrutador destino, contendrá una ruta sin ciclos y su métrica, desde el enrutador fuente al enrutador destino, es decir, un camino del árbol. Además, cada Pre-Join que llegue contendrá un camino diferente.

■ *Join. Creación de estados*

Una vez el enrutador destino recibe al menos un mensaje Pre-Join, envía un mensaje Join a través de la ruta que este mensaje contenía, en dirección del nodo origen en PIM-DM o al RP en PIM-SM. Este mensaje difiere del Join original de PIM-DM o PIM-SM en que debe contener los atributos de la ruta y la lista de enrutadores recibida del mensaje Pre-Join para poder recorrer exactamente el mismo camino.

Cada enrutador que recibe un Join almacena también las métricas TE de las rutas a la que pertenece y que fueron obtenidas del protocolo unicast a través del Pre-Join.

Cada vez que el enrutador destino reciba y acepte un mensaje Pre-Join comparará la ruta que contiene y la ruta que se encuentra en uso en ese instante. El enrutador destino debe enviar un mensaje Join a la mejor ruta y un JoinStandBy a la ruta alternativa. De lo anterior podemos deducir que se necesita crear un nuevo tipo de estado, el estado StandBy, además de los Forward y Prune ya existentes.

Los enrutadores en estado StandBy mantendrán la misma información que mantendrían si su estado fuese Forward, pero no reenviarán los paquetes que reciban, excepto paquetes LoadBalancing.

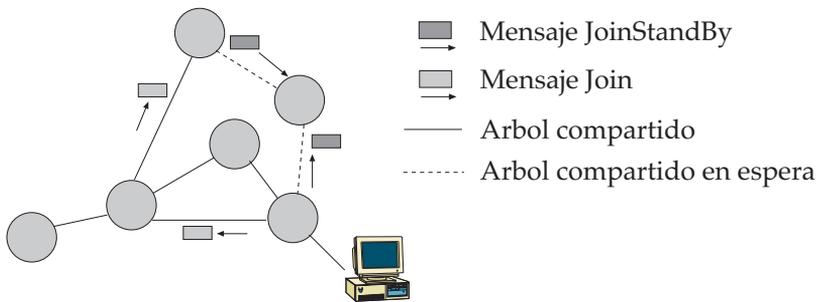


Figura 6. Cambio de ruta

Cuando un enrutador detecte que se ha alcanzado la cantidad límite de tráfico destinada para la ruta en uso, enviará un mensaje JoinStandByRq al destino, para indicar que, debido al tráfico, dicha ruta debe ser puesta en espera. En respuesta a la anterior acción, el enrutador destino enviará un mensaje JoinStandBy a la ruta en uso y un mensaje Join a la siguiente mejor ruta.

Cuando la congestión desaparezca, el enrutador que la detectó es responsable de enviar un mensaje PathON al enrutador destino, para que éste habilite la ruta puesta en estado StandBy anteriormente.

Cuando todas las cantidades límites de tráfico han sido alcanzadas por todas las rutas el enrutador destino envía un mensaje LoadBalancingON a la fuente o al RP.

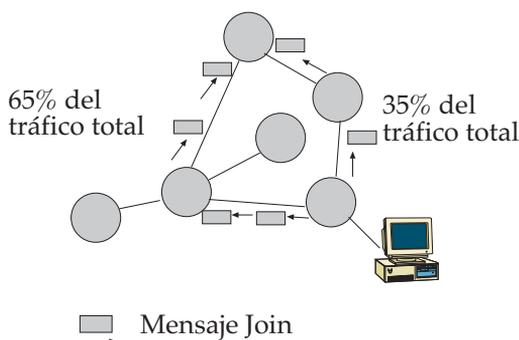


Figura 7. Reenvío proporcionado

Al usar este mecanismo, el nodo origen enviará el tráfico multicast encapsulado en mensajes LoadBalancing, que indicarán qué ruta debe seguir el mensaje, sin importar si la ruta se encuentra StandBy. La cantidad de tráfico por ruta depende de las métricas y ancho de banda de la ruta.

En el momento en que alguna ruta se encuentre de nuevo en condiciones normales de tráfico y el enrutador destino reciba un mensaje PathON de algún enrutador de la misma ruta, el primero enviará un mensaje Join a la ruta y un mensaje LoadBalancingOFF a la fuente, lo cual desactivará el proceso de reenvío proporcionado.

Con la información de ancho de banda y métricas de las rutas contenidas en los enrutadores PIM-DM/SM-TE, los protocolos CR-LDP y RSVP podrán ejecutar una reserva de ancho de banda óptima de acuerdo con los requerimientos del flujo y características de la red.

3.3.3. Diferencias entre PIM-DM-TE y PIM-SM-TE

Para PIM-DM, donde no existe Join, se propone la siguiente modificación. Tras la inundación inicial, los enrutadores que no necesiten tráfico multicast enviarán un mensaje Prune, versión TE. Los enrutadores que lo requieran enviarán un mensaje PreJoinRq en dirección a la fuente, y el proceso continuará tal como se describió anteriormente.

4. ANÁLISIS DE RESULTADOS

Para el protocolo PIM-DM/SM-TE se puede obtener una aproximación muy cercana a la realidad del comportamiento del mismo a través de la herramienta de la simulación Arena. A continuación se presentan resultados de la simulación.

La simulación se ejecutó con base la topología de la *National Science Foundation Network*, con los siguientes parámetros:

- Tamaño de buffer: 32KB= 262144b
- Tiempo de simulación: 3 segundos
- Intervalo de creación de paquetes: Distribución exponencial: 0.0007 seg
- Tamaño paquetes: Distribución triangular (64, 256, 1500) Bytes
- Capacidad canales: 256 kbps = 256000 bps
- Retardo procesamiento por enrutador: 0.25mseg = 0.00025 seg
- Retardo en los enlaces: Figura 8.

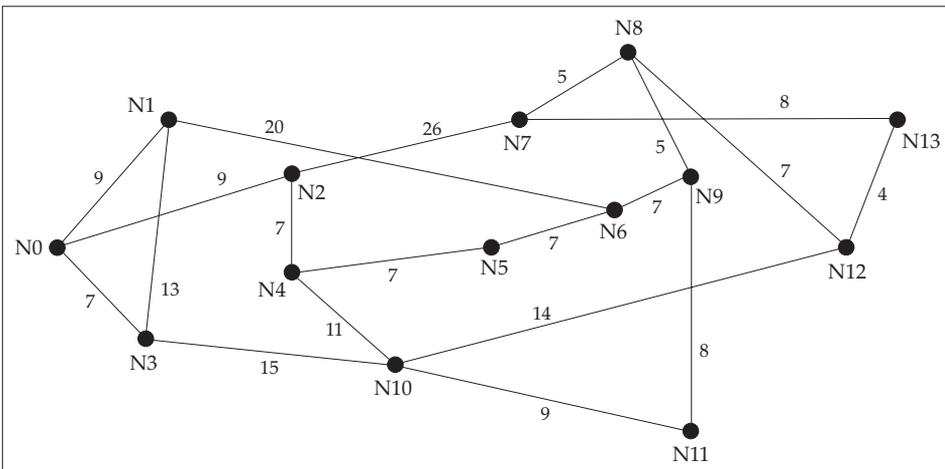


Figura 8. NSF Network

Retardo de paquetes

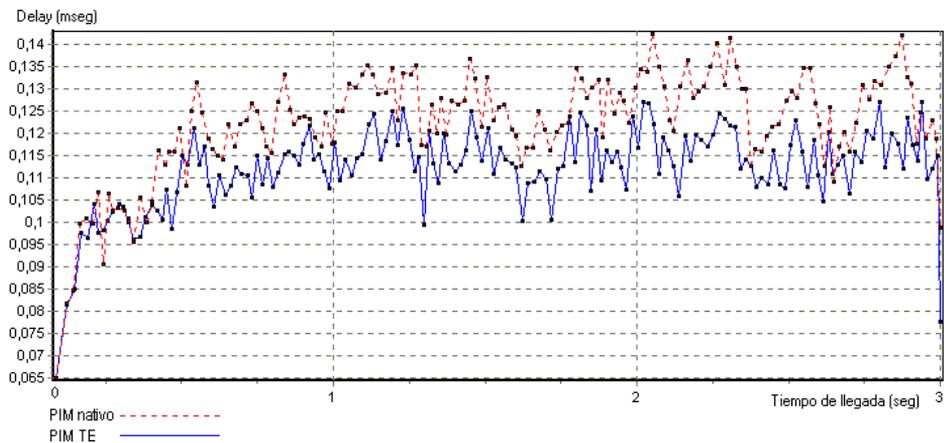


Figura 9. Retardo de paquetes en el tiempo

Retardo/ Protocolo	PIM-DM/SM TE	PIM-DM/SM nativo
Máximo	0.3207seg	0.2229 seg
Mínimo	0.3785 seg	0.0403 seg
Media	0.1125 seg	0.1224 seg

Cada punto de la gráfica corresponde al promedio del retardo de 80 paquetes. Los datos de la tabla fueron obtenidos utilizando todos los datos resultantes de la simulación. Como puede verse en la figura 9, el comportamiento del retardo en la red congestionada fue mejor en el protocolo propuesto en este artículo, PIM-DM/SM TE, que el funcionamiento tradicional PIM-DM/SM nativo.

Paquetes perdidos

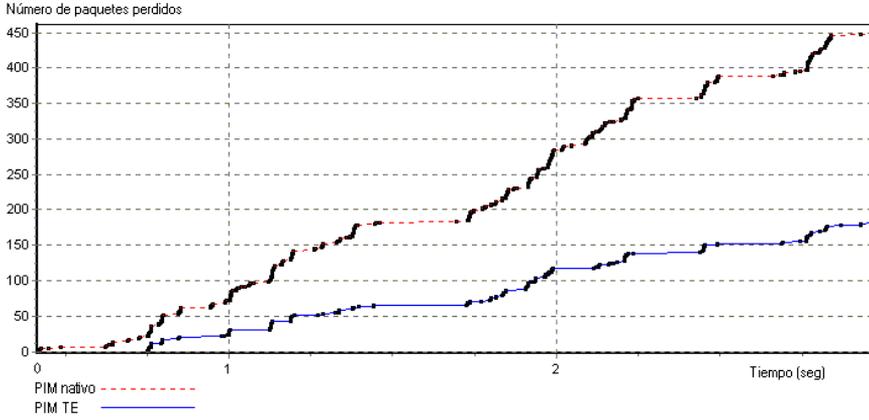


Figura 10. Paquetes perdidos en el tiempo

	PIM-DM/SM-TE	PIM-DM/SM nativo
Paquetes perdidos	185	458

Como puede apreciarse en la figura 10, la cantidad de paquetes perdidos sin el uso de ingeniería de tráfico fue mucho mayor (458 paquetes perdidos) que los perdidos con el uso de la misma ingeniería de tráfico (185).

Utilización de canales

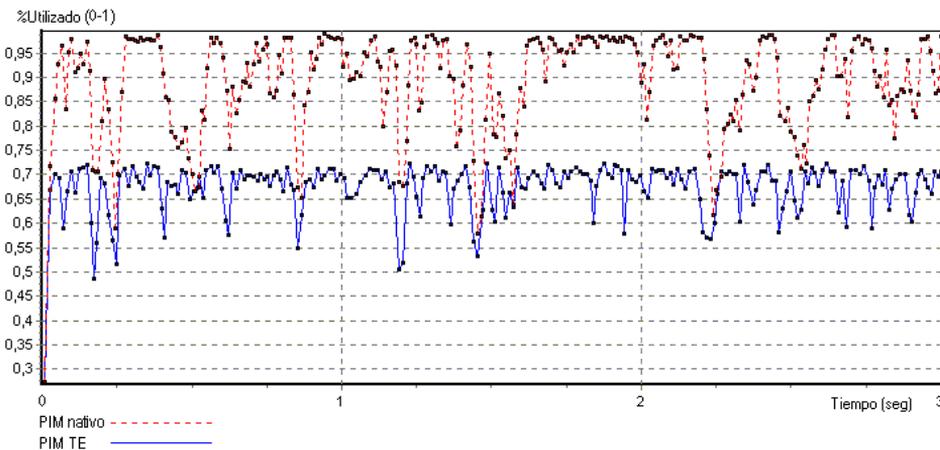


Figura 11. Utilización en el tiempo del canal 1-3

Utilización % / Protocolo	PIM-DM/SM-TE	PIM-DM/SM nativo
Máximo	74.9969%	100%
Mínimo	1.3438%	1.3438%
Media	67.3955%	89.5048%

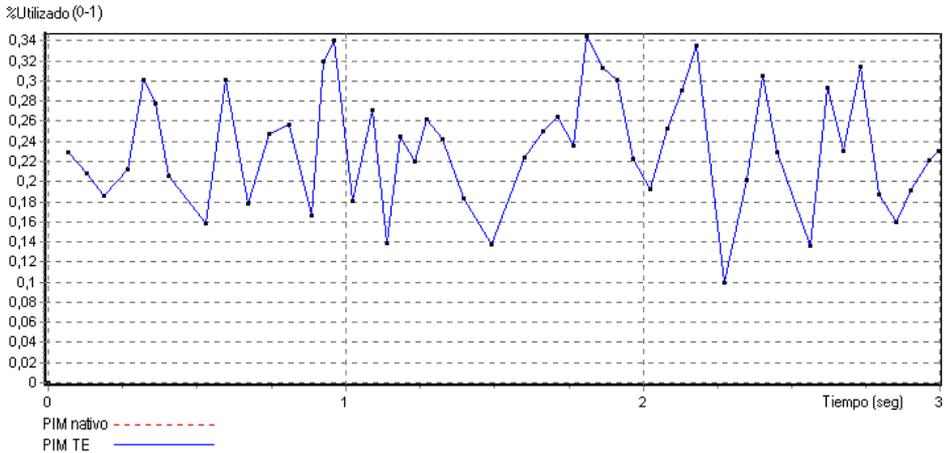


Figura 12. Utilización en el tiempo del canal 1-2

Utilización % / Protocolo	PIM-DM/SM-TE	PIM-DM/SM nativo
Máximo	42.2531%	0%
Mínimo	0%	0%
Media	23.4615%	0%

En las figuras 11 y 12 se observa una mejor utilización de los enlaces cuando se realiza el balanceo de carga, debido a que más canales son utilizados en comparación del esquema tradicional, y por lo tanto la distribución de la transmisión de la información se realiza entre más caminos, lo cual trae como consecuencia una mayor capacidad de transmisión sobre toda la topología de la red.

CONCLUSIONES

A partir de la investigación propuesta y de las pruebas realizadas se puede llegar a las siguientes conclusiones:

La implementación de los protocolos PIM-DM-TE y PIM-SM-TE sería ventajosa debido a la disminución del retardo y de la cantidad de paquetes perdidos en la red debido a congestión. Este protocolo podría verse limitado por la gran

cantidad de información que necesita almacenar, sin embargo, esto es poco en comparación con las mejoras que ofrece en cuanto a reserva y administración de recursos para multicast IP. La congestión en los canales sobreutilizados se reduce en gran medida y es posible utilizar canales que de otra forma permanecerían ociosos a pesar de su utilidad. Además, la información que almacena de las posibles rutas a los destinos pertenecientes a grupos multicast específicos, puede ser utilizada por MPLS y por protocolos de reserva de recursos como CR-LDP o RSVP-TE.

Estas ventajas del protocolo se verían reflejadas en un mejor servicio para el usuario final, menor retardo promedio en la recepción de datos, menor cantidad de pérdida de información, en general, mayor calidad en el servicio recibido.

Referencias

- [1] TANENBAUM, Andrew, *Redes de computadoras*.
- [2] SEMERIA, Chuck & MAUFER, Tom, *Introduction to IP Multicast Routing*.
- [3] <http://www.ietf.org>
- [4] OSI IS-IS Intra-domain Routing Protocol <RFC1142>
- [5] IS-IS extensions for Traffic Engineering <draft-ietf-isis-traffic-04.txt>
- [6] OSPF Versión 2 <RFC2328>
- [7] TE LSAs to extend OSPF for Traffic Engineering <draft-srisuresh-ospf-te-02.txt>
- [8] Protocol Independent Multicast - Dense Mode (PIM-DM): Protocol Specification (Revised) <draft-ietf-pim-dm-new-v2-02.txt>
- [9] Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol specification (Revised) <draft-ietf-pim-sm-v2-new-06.txt>
- [10] WANG, Z., «Internet QoS, Architectures and Mechanisms for Quality of Service». Morgan Kaufmann Publishers, 2001.
- [11] CIDON, I., ROM, R. & SHAVIT, «Analysis of multi-path routing», *IEEE / ACM Transactions on Networking*, vol. 7, pp. 885-896, Dec. 1999.
- [12] AMMAR, M.A., POLYZOS, G.C. & TRIPATHI, S.K., Network support for multipoint communications: Guest Editorial. *IEEE Journal on Selected Areas in Communications*, vol. 15, N° 3, April 1997.
- [13] RAO, N. & BATSELL, S. «QoS Routing via multiple Paths using Bandwidth Reservation». *Infocom 1998*.