

01126  
52



UNIVERSIDAD NACIONAL AUTÓNOMA  
DE MÉXICO

FACULTAD DE INGENIERÍA

PROPUESTA DE IMPLEMENTACIÓN DE LOS  
PROTOCOLOS DE ENRUTAMIENTO DE LA RED DE  
ALTO DESEMPEÑO DEL BACKBONE DE CUDI

TESIS CON  
FALLA DE ORIGEN

T E S I S

QUE PARA OBTENER EL TÍTULO DE:  
INGENIERO MECÁNICO ELECTRICISTA  
(ÁREA ELÉCTRICA - ELECTRÓNICA)  
P R E S E N T A :  
HANS LUDWIG REYES CHÁVEZ



DIRECTOR DE TESIS:  
ING. ALFREDO HERNÁNDEZ MENDOZA

MÉXICO, D.F. CIUDAD UNIVERSITARIA,

JUNIO 2003.

1



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

**TESIS CON  
FALLA DE  
ORIGEN**

# **PAGINACION DISCONTINUA**

### Agradecimientos

Al creador por haberme dado la vida y la oportunidad de estudiar esta carrera.

A mis padres por su apoyo desde su inicio hasta el término de esta carrera.

A la Universidad Nacional Autónoma de México y a los profesores de la Facultad que dejan parte de sus vidas en las aulas de clases y de los cuales he recibido la enseñanza de alta calidad y que me han permitido llegar hasta este momento.

A los ingenieros Alfredo Hernández y Gabriela Medina por la oportunidad de ser parte del equipo de RedUNAM.

Y a todos los familiares, amigos y compañeros de trabajo que es difícil enumerarlos a todos en este momento.

TESIS CON  
FALLA DE ORIGEN.

Hans Ludwig Reyes Chávez

<b>Índice</b>	i
<b>Introducción</b>	<b>1</b>
<b>Capítulo 1 Antecedentes</b>	<b>4</b>
1.1 Inicios de Internet y Evolución	5
1.2 El Internet de Hoy	7
1.2.1 Requerimientos de la NSFNET	8
1.2.2 Networks Access Points	9
1.2.3 El proyecto de Route Arbiter	11
1.2.4 Jerarquías de ISP	11
1.3 ¿Por que surge el proyecto de Internet2?	12
1.3.1 Membresías de Internet2	14
1.3.2 Estructura de participación con Abilene	14
1.3.3 Abilene conectores/GigaPop	15
1.3.4 Miembros Regulares "Universidades"	15
1.3.5 Corporaciones en Norteamérica	15
1.3.6 Corporaciones en Europa-Medio Oriente	18
1.3.7 Corporaciones en Asia-Pacífico	18
1.4 Internet 2 en México	19
1.4.1 Organización	20
1.4.2 Asociados Academicos	22
1.4.3 Asociados Institucionales	22
1.4.4 Afiliados	23
1.4.5 Aspectos técnicos	24
1.4.6 Backbone Actual	27
1.4.7 NOC de Internet2-México	28
1.4.7.1 Funciones del NOC-I2	29
<b>Capítulo 2 Conceptos Básicos</b>	<b>31</b>
2.1 Modelo OSI	32
2.1.1 El modelo OSI hoy en día	32
2.1.2 Las capas del modelo OSI	32
2.1.3 Proceso de encapsulamiento del modelo OSI	35
2.1.4 Beneficios de la utilización de capas	36
2.2 Descripción del protocolo IP	37
2.2.1 Posición de TCP/IP de acuerdo al modelo OSI	37
2.2.2 Formato del datagrama IP	38
2.2.3 Campos del datagrama IP	39
2.2.4 Diferencias entre TCP/IP y el modelo OSI	41
2.2.5 Direccionamiento IPv4	42
2.2.6 VSLM - Variable Length Subnet Masks	49
2.2.7 CIDR - Classless Inter-Domain Routing	51
2.3 Enrutamiento	53
2.3.1 Características de Diseño	53
2.3.2 Clasificación	55
2.3.2.1 Dinámico / Estático	56

# TESIS CON FALLA DE ORIGEN

## Índice

2.3.2.2 Single Path / Multi Path	56
2.3.2.3 Plano / Jerárquico	56
2.3.2.4 Distance Vector / Link State	57
2.3.2.4 Interdominio / Intradominio	57
2.4 Jerarquía en capas	57
2.4.1 Capa de Acceso	58
2.4.1 Capa Distribución	58
2.4.2 Capa de Core	58
2.5 Tipos de Core (Backbone)	59
2.5.1 Backbone Switcheados	59
2.5.2 Backbones Enrutados	60
<b>Capítulo 3 Protocolo de Enrutamiento OSPF</b>	<b>61</b>
3.1 Historia de OSPF	62
3.2 Descripción preliminar de OSPF	62
3.3 Operación de OPSF	63
3.4 Vecinos y Adyacencias	64
3.5 El protocolo de Hello	65
3.6 Soporte de OSPF a diferentes tipos de redes	67
3.6.1 Redes Punto a Punto	67
3.6.2 Redes Broadcast	67
3.6.3 Redes NBMA	68
3.6.4 Redes Punto multipunto	68
3.6.5 Virtual Links	69
3.7 El Router Designado y Backup del Designado	69
3.8 Estados de la interfaces de OSPF	71
3.9 Fases para el establecimiento de las adyacencias	73
3.10 Estados de los vecinos	74
3.11 Areas en OSPF	75
3.11.1 Tipos de traficos en relación a las áreas	77
3.11.2 Enrutamiento Jerárquico	77
3.11.3 Tipos de Routers en OSPF	77
3.12 Tipos de LSA	78
3.13 Tipos de Áreas en OSPF	80
3.13.1 Áreas Stub	81
3.13.2 Totally Stubby Áreas	82
3.13.3 Not-So_Stubby Áreas	82
3.14 Link state data base	83
3.15 La tabla de enrutamiento	84
3.15.1 Tipos de camino en OSPF	85
<b>Capítulo 4 Enrutamiento Interdominio y BGP</b>	<b>87</b>
4.1. Enrutamiento Interdominio	88
4.1.1 Historia del AS (Sistema Autónomo)	88
4.1.2 Definición de AS	88
4.1.3 Evolución del enrutamiento Interdominio	89
4.2 Protocolo de Enrutamiento BGP	89

4.2.1 Operación de BGP Descripción General de BGP-4	89
4.2.2 Tipos de mensajes de BGP	92
4.2.2.1. Mensaje de Open	92
4.2.2.2. Mensaje de Keepalive	93
4.2.2.3. Mensaje de Update	93
4.2.2.4. Mensaje Notification	94
4.2.3 Estados de BGP	94
4.2.3.1. Estado de Idle	94
4.2.3.2. Estado de Connect	95
4.2.3.3. Estado Active	95
4.2.3.4. Estado OpenSent	95
4.2.3.5. Estado OpenConfirm	96
4.2.3.6. Estado Established	96
4.2.4 Path Attributes	96
4.2.4.1. Categorías de los Path Attributes	97
4.2.4.2. El atributo de Origen (Origin)	98
4.2.4.3. El atributo de AS_PATH	98
4.2.4.4. El atributo NEXT-HOP	99
4.2.4.5. El atributo LOCAL_PREF.	100
4.2.4.6. El atributo MULTI-EXIT-DISC	101
4.2.4.7. Los atributos ATOMIC_AGGREGATE y AGGREGATOR	102
4.2.4.8. El atributo de community	104
4.2.4.9. Atributos ORIGINATOR_ID y CLUSTER_LIST	105
4.2.4.10. Administrative weight	106
4.2.4.11. Tipos de AS_PATH	106
4.2.5 Proceso de decisión de BGP	109
4.2.6 Penalizaciones a la rutas (Route-Dampening)	110
4.2.7 Sincronización	112
4.2.8 Implementación a gran escala de BGP	114
4.2.8.1. Comunidades	114
4.2.8.2. Route Reflectors	115
4.2.8.3. Confederaciones	120
4.2.9 ¿Quién Necesita BGP?	122
4.2.9.1. Single home	123
4.2.9.2. Multi-home	124
4.2.9.3. Peering / Transito	126
4.2.9.4. IXPs/NAPs	128
<b>Capítulo 5 Descripción de los protocolos Multicast</b>	<b>130</b>
5.1 Concepto de Multicast	131
5.1.1 Ventajas de Multicast	133
5.2 Tipos de Direcciones	133
5.2.1 Direcciones Unicast	133
5.2.2 Direcciones Broadcast	134
5.2.3 Direcciones Multicast	134



5.3 Direcciones de Multicast capa 2	130
5.3.1 Ethernet/FDDI	136
5.3.2 Token Ring	138
5.4 Soporte de Multicast en IP	139
5.4.1 Funcionamiento de IGMP	140
5.4.1.1 Ingresando a un grupo	142
5.4.1.2 Permanencia en un grupo	143
5.4.1.3 Abandonando un grupo	143
5.4.2 IGMPv2	144
5.4.3 Árboles de distribución de Multicast	146
5.4.3.1 Árboles de distribución de una fuente	146
5.4.3.2 Árboles de distribución compartida	147
5.4.3.3 Características de los árboles de distribución	148
5.5 Enrutamiento Multicast	148
5.5.1 ¿Qué es RPF?	150
5.5.2 Comprobación de RPF	151
5.5.3 Tipos de Protocolos Multicast	152
5.5.4 Topologías Sparse vs Dense	153
5.5.5 Diferencias entre el enrutamiento Unicast y Multicast	154
5.6 Protocolo de enrutamiento PIM-Sparse Mode	156
5.6.1 Comprobación de RPF en PIM-SM	156
5.6.2 PIM-SM adición al árbol compartido	156
5.6.3 Registro del Remitente	157
5.6.4 Mecanismos de PIM-SM	158
5.6.4.1 Descubrimiento de vecinos	159
5.6.4.2 Estado de PIM	160
5.6.4.3 Ingresos	161
5.6.4.4 Registros	162
5.6.4.5 STP-Switchover	164
5.6.4.6 Podas	167
5.6.4.7 Maquina de estados	170
5.7 MBGP extensiones para Multicast	171
5.7.1 Bases de información de enrutamiento	171
5.7.2 Mensaje de actualización de MBGP	172
5.7.2.1 Atributo MP_REACH_NLRI	173
5.7.2.2 Atributo MP_UNREACH_NLRI	174
5.7.2.3 Capacidad de Negociación MBGP	175
5.7.2.4 Información NLRI	176
5.7.3 Topologías congruentes Unicast-Multicast	178
5.7.4 Topologías incongruentes Unicast-Multicast	179
5.7.5 Conversión NLLI de Unicast a Multicast	179
5.8 MSDP	180
5.8.1 MSDP Peers	181
5.8.2 Mensajes MSDP	182
5.8.3 Comprobación RPF del Mensaje SA	183
5.8.4 MSDP Mesh-Groups	184
5.8.5 MSDP SA Caching	185

5.8.6 Ventajas del uso de MSDP	186
<b>Capítulo 6 El protocolo de la siguiente generación de Internet (IPv6)</b>	<b>187</b>
6.1 Metas de IPv6	188
6.2 Redes de IPv6	188
6.2.1 6bone	189
6.2.2 6REN	189
6.3 El paquete de Ipv6	189
6.3.1 Tamaño de las direcciones	189
6.3.2 Representación de las direcciones	190
6.3.3 Asignación del espacio de direcciones	191
6.3.4 Estructura de las direcciones	192
6.3.5 Formato de las direcciones	192
6.3.6 Formatos especiales de direcciones	195
6.3.6.1 Unspecified	195
6.3.6.2 Loopback	195
6.3.6.3 IPv4 contenido dentro de IPv6	195
6.4 Tipos de direcciones	196
6.4.1 Direcciones de Link-local	196
6.4.2 Direcciones de Site-local	197
6.4.3 Direcciones de Anycast	197
6.4.4 Direcciones de Multicast	199
6.4.5 Direcciones requeridas por los nodos	201
6.5 Encabezado de Ipv6	202
6.5.1 Formato del encabezado	202
6.6 Funcionalidades de Ipv6	204
6.6.1 ICMPv6	204
6.6.2 Descubrimiento de vecinos	205
6.7 Autoconfiguración	206
6.7.1 Stateless Autoconfiguration	206
6.7.2 Stateful Autoconfiguration	208
6.8 RIPv6	208
6.9 BGP-4 Multiprotocol extensiones para Ipv6	211
6.10 Mecanismos de Transición de IPv4 a IPv6	212
6.10.1 IPv6 a través de túneles de IPv4	212
6.10.2 Dual Stack	212
6.11 DNS para IPv6	213
<b>Capítulo 7 Propuesta de Diseño</b>	<b>214</b>
7.1 Propuesta de OSPF como protocolo de enrutamiento Intradominio	215
7.1.1 Diseño escalable de OSPF	216
7.1.2 Reglas en el diseño de las áreas	216
7.1.3 El diseño del direccionamiento	217
7.1.4 Características de OSPF que se deben considerar	218
7.1.5 Agregando redes al proceso de OSPF	219
7.1.6 Implementación en el backbone de CUDI	219
7.2 Propuesta de BGP como protocolo de enrutamiento Interdominio	220

7.2.1 Técnicas de Escalamiento	220
7.2.1.1 Soft reconfiguration	221
7.2.1.2 Peer-groups	222
7.2.2 Políticas de enrutamiento	223
7.2.3 Interacción entre BGP y el protocolo IGP	224
7.2.3.1 Funciones del IGP	224
7.2.3.2 Funciones de BGP	224
7.2.4 Recomendaciones en la implementación de EBGp:	225
7.2.4.1 Lo que nunca se debe de hacer:	226
7.2.4.2 Herencias de BGP que deben ser deshabilitadas:	226
7.2.4.3 Anuncios de redes que no deben ser recibidos:	226
7.2.5 Inyectando los anuncios de los prefijos al proceso de BGP	227
7.3 Topología de Internet 2	227
7.3.1 Topología en México	228
7.3.2 Función del Backbone	229
7.3.3 Función de los miembros Asociados	230
7.3.4 Función de los miembros Afiliados	232
7.4 Propuesta para el soporte de Multicast	232
7.4.1 DVMRP	232
7.4.2 PIM-DM	233
7.4.3 PIM-SM	233
7.4.4 MBGP extensiones para Multicast	235
7.4.5 MSDP	235
7.4.5.1 Redundancia al RP	236
7.4.6 Propuesta para el backbone	238
7.5 Propuesta para el soporte de Ipv6	240
7.5.1 Ripv6	240
7.5.2 MBGP extensiones para IPv6	241
<b>Conclusiones</b>	<b>242</b>
<b>Bibliografía</b>	<b>245</b>
<b>Glosario</b>	<b>248</b>
<b>Anexo A</b>	<b>262</b>

TESIS CON  
FALLA DE ORIGEN

# Introducción

# TESIS CON FALLA DE OMBLEN

Introducción

## Introducción

Uno de los objetivos de una universidad es el desarrollo tecnológico y la investigación, de ahí que la Universidad Nacional Autónoma de México se encuentre realizando diversas investigaciones en el desarrollo y la implementación de nuevas tecnologías de comunicación. El empleo cotidiano de los servicios de comunicación ofrecidos por la tecnología hace que se vuelvan imperceptibles, aún cuando se han convertido en una parte esencial en el desarrollo de nuestras vidas.

Hoy en día prácticamente en todas las universidades del país existen trabajos de investigación sobre nuevas tecnologías de telecomunicaciones. Dentro de estos trabajos la red mundial Internet no está ausente. En la mayoría de las universidades existen, investigadores, profesores y alumnos que están desarrollando nuevas aplicaciones enfocadas al uso de los recursos de la red de datos de sus universidades, de forma paralela al desarrollo de tecnologías, se encuentra la tarea de extender este desarrollo desde las universidades hacia el resto de las comunidades en México abarcando tanto a las empresas del sector público como el privado.

Como una respuesta de las universidades dentro del área de la tecnología, es la creación de un organismo como la Corporación Universitaria para el Desarrollo de Internet, conocida por sus siglas como CUDI. Uno de los principales objetivos de la Corporación Universitaria es promover la transferencia del desarrollo tecnológico que se da en las universidades hacia el resto del país. El objetivo general de esta tesis es elaborar una propuesta de implementación de los protocolos de enrutamiento para la red de CUDI que permitan dar soporte a las aplicaciones de Internet2 y adicionalmente dar un documento de referencia que coadyuve a las universidades que se encuentran conectadas a la red, de tal manera que una vez que conozcan el funcionamiento de la red dorsal, puedan implementar los protocolos en sus redes de forma compatible con el funcionamiento de la red de CUDI.

La estructura de esta tesis se encuentra dividida en siete capítulos, en los primeros seis se hace una descripción de los conceptos teóricos de los protocolos IPv4, OSPF, BGP, PIM-

SM, MBGP, MSDP e IPv6 y del funcionamiento de Internet e Internet2. El último capítulo se enfoca a proponer un esquema de configuración de la RedCUDI para un funcionamiento saludable y que no este limitado a la cantidad de universidades conectadas actualmente, así como la forma en que las universidades se comunicarán con la red. A continuación se describe brevemente el contenido de cada capítulo.

El capítulo 1 contiene una breve descripción de la historia de Internet desde sus inicios, su evolución, sus principales componentes y la creación del proyecto de colaboración llamado Internet2.

El capítulo 2 hace una descripción de los conceptos básicos del Protocolo de Internet (IP) y un gran número de conceptos relacionados con las redes de datos.

El capítulo 3 contiene una breve descripción del funcionamiento de las partes más importantes del protocolo de enrutamiento intradominio OSPF.

El capítulo 4 realiza una breve descripción de las partes más importantes del protocolo de enrutamiento interdominio BGP, también describe muchos conceptos relacionados con el enrutamiento entre diferentes sistemas autónomos.

El capítulo 5 elabora una descripción de los protocolos de Multicast PIM-SM, MBGP (extensiones para Multicast) y MSDP. También se describen muchos conceptos relacionados con el enrutamiento Multicast.

El capítulo 6 contiene una descripción de los protocolos de IPv6, se menciona también muchas de las ventajas y deficiencias que este protocolo pretende solucionar que se presentan con el actual protocolo IPv4, también se describen los protocolos de enrutamiento RIP y MBGP (extensiones para IPv6) utilizados para soportar al protocolo IPv6 en forma nativa.

El capítulo 7 es la propuesta de implementación de la red de CUDI siguiendo las mejores practicas (Best Practices) de funcionamiento al utilizar los diferentes protocolos de enrutamiento, estas practicas son una recopilación de recomendaciones hechas por los fabricantes, compañías de telecomunicaciones y por los grupos de trabajo de Internet2.

TESIS CON  
FALLA DE ORIGEN

TESIS CON  
FALLA DE ORIGEN

# Capítulo 1

## Antecedentes



## Capítulo I Antecedentes

Internet inició como un proyecto a finales de los 60s de ARPA (Advanced Research Projects Agency), ahora llamada DARPA del departamento de defensa de los EE.UU. DARPA experimento con la conexión de redes de computadoras dando acceso a múltiples universidades y compañías privadas involucradas.

### 1.1 Inicios de Internet y Evolución

En diciembre de 1969, la red experimental estuvo en línea con la conexión de cuatro nodos conectados a través de enlaces de 56kbps. Esta nueva tecnología proveía una alta confiabilidad y fue la base en la creación de dos redes militares similares. MILNET en EE.UU. y MINET en Europa, miles de hosts y usuarios conectaron sus redes privadas (de universidades y gobierno) a la ARPANET, creando entonces la "ARPA Internet". ARPANET tenía una *politica de uso aceptable* –conocida como AUP "Acceptable Use Policy", la cual prohibía el uso del Internet para fines comerciales. ARPANET fue declinada en 1989.

Para 1985, la ARPANET ya era ampliamente usada y se encontraba congestionada. En repuesta la NSF (National Science Foudation) dio inicio a la fase uno en el desarrollo de la NSFNET. La NSFNET fue compuesta de múltiples redes regionales y grupos de redes (como la NASA Science Network) conectada a un backbone de mayor capacidad que constituyo el core de toda la NSFNET.

En su forma más temprana, en 1986 la NSFNET fue creada con una arquitectura de tres niveles. La arquitectura conecto campus universitarios y organizaciones de investigación, las cuales se conectaban a redes regionales y a su vez a un backbone principal enlazado a seis centros de súper computó a nivel nacional. Los enlaces originales fueron de 56kbps y en 1988 fueron actualizados a enlaces T1s (1.544Mbps) como resultado de una solicitud de servicios de red más rápidos de la NSFNET, que fue concesionada a Merit Networks, con alianzas de MCI, IBM y el estado de Michigan. La NSFNET se formo con un backbone de enlaces T1s conectando un total de 13 sitios que incluyeron Merit. BARRNET, MIDnet,

Westnet, NorthWestNet, SEQUINET, SURAnet, NCAR (National Atmospheric Research), y cinco centros de supercomputo de la NSF.

En 1990, Merit, IBM, y MCI comenzaron una nueva organización conocida como Advanced Network and Services (ANS). Merit Network Internet Engineering Group desarrollo una base de datos con políticas de enrutamiento con consultas y administración de servicios para la NSFNET, mientras que ANS operó los routers del backbone y el "Centro de Operación de la Red" (Network Operation Center, NOC).

Para 1991, el tráfico de datos se incremento en gran medida, por lo cual se necesito de actualizar el backbone a enlaces T3 (45Mbps) como lo ilustra la figura 1-1.

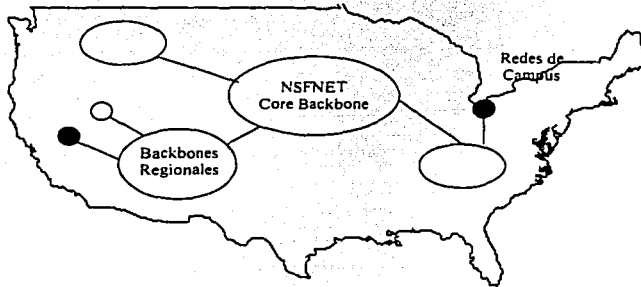


Figura 1-1 Red NSFNET basada en el ambiente Internet.

Todavía al inicio de los 90s, la NSFNET estaba reservada para aplicaciones de investigación y educación, y los backbones de las agencias del gobierno fueron reservados para una misión claramente definida. Pero surgieron nuevas presiones alimentadas por redes emergentes. Adicionalmente diferentes agencias necesitaban interconectarse con otras redes. Surgió entonces un interés comercial y de propósito general que fue requerido en el acceso a la red, que dio lugar a los Proveedores de Servicios de Internet (Internet Service Providers, ISP), como solución a esta demanda, redes en otros lugares fuera de EE.UU.

fueron desarrolladas en forma paralela, creando un interés internacional de interconexión. Debido a la variedad de las nuevas y existentes entidades cada una de ellas con diferentes metas, la complejidad de las conexiones e infraestructura crecieron.

Las redes de las agencias de gobierno se conectaron a través de un Federal Internet eXchange point (FIX) en ambas costas Este y Oeste. Al mismo tiempo alrededor del mundo particularmente en Europa y Asia se desarrollaron sustancialmente infraestructuras y conectividad. Sprint fue asignada por NSFNET para administrar las conexiones internacionales (Internacional Connections Manager, ICM) y proveer la conectividad entre los backbones de EE.UU., Europa y Asia. Finalmente la NSFNET fue declinada en Abril 1995.

## 1.2 El Internet de Hoy

La declinación de la NSFNET fue hecha en pasos bien definidos para asegurar la continúa conectividad de las instituciones y agencias de gobierno que estaban conectadas a través de redes regionales. La estructura del Internet de hoy se formo partiendo del core de la NSFNET hacia una arquitectura distribuida operada por múltiples proveedores comerciales como Sprint, MCI, BBN y otros conectados a través de redes de grandes capacidades conocidos como puntos de intercambio (exchange points). La figura 1-2 ilustra la forma general del Internet de hoy.

El Internet actual es una suma de proveedores de servicios ISPs (el termino ISP es usado cuando se refiere a alguien quien provee servicios a usuarios finales u otros proveedores) que tienen puntos de conexión sobre múltiples regiones llamados Puntos de Presencia (Point of Presence, POP). Esta colección de POPs y la forma en que están interconectados constituyen la red del ISP. Los clientes son conectados a los proveedores vía los POPs, los clientes de los proveedores incluso pueden ser otros ISPs. Los proveedores que tienen POPs a través de todo un país son llamados proveedores nacionales.

Los proveedores que cubren una región (Regional Providers) se conectan a otros proveedores en uno o múltiples puntos. Para permitir que los clientes de un proveedor se

comunicuen con los clientes de otro proveedor se crearon los NAPs (Network Access Point) como puntos de interconexión. El término NSP (Network Service Provider) es usualmente restringido a proveedores bien definidos que administran los NAPs, como Sprint, Ameritech, y MFS. El termino NSP también es usado con menor frecuencia para hacer referencia a un proveedor que esta conectado a todos los NAPs.

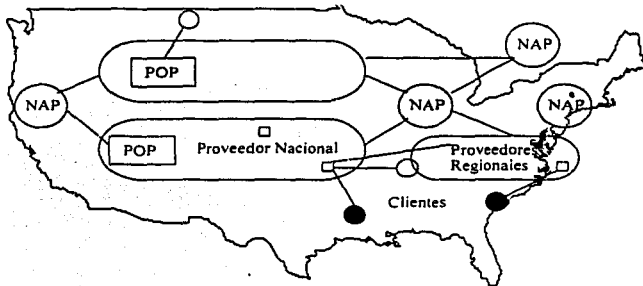


Figura 1-2 Estructura general del Internet actual.

### 1.2.1 Requerimientos de la NSFNET

Desde 1986 la NSFNET soportó las necesidades de las redes dedicadas a la investigación, también soportó las metas de alto rendimiento en cómputo y comunicaciones (High Performance Computing and Communications, HPCC), la cual promovió el liderazgo en investigación y ciencias. La NREN (National Research and Education Network), es una subdivisión del programa HPCC para la investigación y educación, también llamada *Giga-bit-per-second Networking* que tuvo lugar a mitad de los 90s. Todas estas necesidades además de la expiración de los acuerdos corporativos de la NSFNET en Abril de 1995 dieron lugar a lo que fue conocido como los requerimientos de la NSFNET.

El primer requisito fue establecido en 1987 y se refería a la actualización del backbone a enlaces T3 para finales de 1993. En 1992 la NSF busco el desarrollo de los siguientes requerimientos que pudieran alojar y promover el papel de los servicios comerciales de los

proveedores y que proporcionarán los cimientos a una nueva y robusta estructura de Internet. Al mismo tiempo la NSF dio un paso atrás en la operación de la red enfocándose en la investigación de características e iniciativas. Al final los resultados fueron presentados en Mayo de 1993.

El requerimiento final incluyó como propósito el desarrollo de cuatro proyectos por separado:

- La creación de una serie de NAPs que fueran los proveedores más grandes que realizarían la interconexión de sus redes para el intercambio de tráfico.
- La implementación de un RA (Route Arbiter), este proyecto facilitaría el intercambio de políticas y direccionamiento de los múltiples proveedores conectados a los NAPs.
- Encontrar un proveedor de un backbone con servicios de alta velocidad para redes educacionales y de gobierno, para este fin se creó la vBNS (Very high-speed Backbone Network Services).
- La transición entre las redes existentes y/o la reasignación de redes regionales para el soporte de conectividad interregional (IRC) por conexión a NSP y estos a los NAPs o directamente a conectados a los NAPs. Cualquier NSP seleccionado para este propósito debería de estar conectado cuando menos a tres NAPs.

### 1.2.2 Networks Access Points

El requisito para este proyecto fue una invitación hacia las compañías para implementar y administrar un número de NAPs donde la vBNS y otras redes podrían interconectarse. Estos NAPs debían de ayudar a las redes regionales, proveedores de servicios de red y a las comunidades de investigación y educación a conectarse e intercambiar tráfico unos con otros. También debían de proveer la interconexión de redes en un ambiente en el que no estuvieran sujetos al uso de las políticas de la NSF (esta política fue creada para restringir el uso del Internet para la investigación y la educación). Por lo tanto el uso general incluyendo el uso comercial, puede ser transportado también a través de los NAPs.

*¿Que es un NAP?*

Un NAP esta definido como una red o punto de intercambio de alta velocidad al cual pueden estar conectados un cierto número de routers con el propósito de intercambiar tráfico. Los NAPs deben de operar al menos a velocidades de 100Mbps y deben de poder ser actualizables a las velocidades que sean demandadas. Un NAP físicamente puede ser un simple switch de FDDI (100Mbps), ATM (155Mbps) o Ethernet (100/1000/10000 Mbps) pasando el tráfico de un proveedor de servicios a otro.

El concepto de NAP fue construido sobre la base de los conceptos de FIX (Federal Internet eXchange point) y CIX (Comercial Internet eXchange point). Los cuales fueron construidos alrededor de anillos de FDDI con redes conectadas a velocidades de 45Mbps o mayores.

Las compañías a las cuales la NFS les dio la concesión de los NAPs fueron:

- Sprint NAP – Pensauken, NJ
- PacBell NAP – San Francisco, CA
- Ameritech Advanced Data Service (AADS) NAP – Chicago, IL
- MFS Datanet (MAE-East) NAP – Washington. D.C.

El Backbone de la NFSNET fue conectado físicamente al Sprint NAP el 13 Septiembre de 1994, a mediados de Octubre de ese mismo año se conecto PacBell NAP, y finalmente Ameritech NAP en Enero de 1995. El backbone de la NFSNET se actualizo a FDDI gracias a un ofrecimiento de MFS en Marzo de 1995. NAPs adicionales alrededor del mundo se han creado para proveer la interconectividad hasta ahora requerida.

Las redes de interconexión a los NAPs deben ser redes de altas velocidades, inicialmente se utilizaron enlaces de 1.544Mbps (T1) pero estos se han actualizado a mayores velocidades.

TESIS CON  
FALLA DE ORIGEN

### 1.2.3 El proyecto de Route Arbiter

Otro de los proyectos requeridos por la NFS fue la creación de un Route Arbiter (RA), el cual proporcionaría un trato igualitario a las redes de los diferentes proveedores de servicio. El RA debe de proveer una base datos con la información de enrutamiento para proporcionar escalabilidad y administrabilidad a las redes.

Los múltiples proveedores conectados a los NAPs tenían que crear una forma escalable de realizar el intercambio de información, ya que cada proveedor debía de establecer una conexión con todos los otros proveedores. El proyecto de RA se creo para reducir las conexiones de todos contra todos (Full mesh peering). En lugar de que cada proveedor estableciera una sesión con cada uno de los otros proveedores, solo se establecería una conexión con un sistema central llamado el *Route-Server*. El route server mantendría una base de datos con toda la información necesaria para definir las políticas de enrutamiento requeridas por los proveedores.

### 1.2.4 Jerarquías de ISP

Actualmente a los ISP se les clasifica en diferentes niveles conocidos como Tiers, en donde el nivel más alto es el 1 y el último normalmente es el 3, pueden existir más niveles de Tiers realizando diferentes funciones dependiendo del nivel en el que estos se encuentren. El intercambio de información de enrutamiento se hace utilizando exclusivamente el protocolo BGP (Border Gateway Protocol).

#### *Tier 1*

Son ISPs con una cobertura internacional o nacional con redes de altas velocidades. Sirven de tránsito para la mayoría de los ISPs de niveles Tier 2 y Tier 3.

#### *Tier 2*

Tienen una menor cobertura o recursos de red que los Tier 1 y la mayoría de las ocasiones requieren de los servicios de los Tier 1.

*Tier 3 y otros*

Debido a su menor cobertura y/o infraestructura, este y los restantes niveles de Tiers requieren de los servicios de los Tiers 1 y 2.

La figura 1-3 muestra la forma en que se interconectan los diferentes niveles de Tiers.

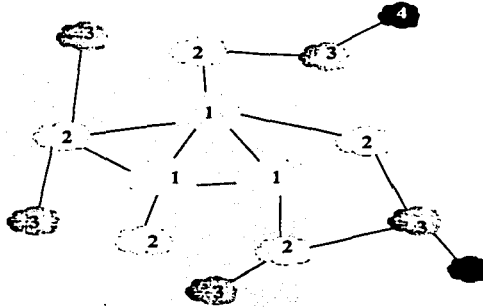


Figura 1-3 Niveles de Jerarquía de los Tiers.

**1.3 ¿Por que surge el proyecto de Internet2?**

Una vez que la red de la NSF llego a su termino, el Internet fue administrado prácticamente por compañías privadas, se vio entonces la necesidad de crear una red que promoviera el uso del Internet como en sus inicios. Una red orientada a la educación, la investigación de tecnologías avanzadas y grandes anchos de banda que sirvieran como laboratorio de pruebas para el desarrollo de nuevas aplicaciones y nuevas tecnologías que después fueran utilizadas en Internet. Fue entonces cuando surgió la "*iniciativa de colaboración entre las universidades y centros de investigación*", a la cual se le dio el nombre de Internet2. Uno de los requerimientos de la NSF fue la creación de una red que cumpliera este propósito como lo fue la creación de la vBNS. Siguiendo entonces esta iniciativa de colaboración se creó el proyecto que hoy se conoce como Internet2.

TESIS CON  
FALLA DE ORIGEN



Muchas veces se confunde al proyecto de Internet2 con las redes que se han desarrollado para soportar el proyecto.

De acuerdo a la página [www.internet2.edu](http://www.internet2.edu): Internet 2 es un consorcio dirigido por más de 190 universidades que trabajan en asociación con el gobierno y la industria para desarrollar e implementar aplicaciones y tecnologías avanzadas, acelerando la creación del Internet del mañana. Internet2 esta creando la sociedad entre la academia, industria y gobierno que se fomento en los inicios del Internet actual.

Los principales objetivos de Internet2 son:

- La creación de una red de alta capacidad para la comunidad de investigación.
- Permitir el desarrollo de aplicaciones revolucionarias.
- Asegurar la rápida transferencia de servicios de red y aplicaciones a todo el Internet.

Los miembros de Internet2 colaboran a través de grupos de trabajo e iniciativas sobre temas como:

- Aplicaciones Avanzadas.
- Middleware.
- Nuevas tecnologías de red.
- Infraestructuras Avanzadas de Red.
- Asociaciones y alianzas.
- Iniciativas.

TESIS CON  
FALLA DE ORIGEN

Una de las redes dorsales de Internet2 en EE.UU. es la red de Abilene. En algunas partes de EE.UU. se han creado redes regionales de forma similar a los inicios de Internet, las cuales se conectan a su vez a la red de Abilene, en ocasiones las universidades se pueden conectar directamente.

### 1.3.1 Membresías de Internet2

En el diagrama de la figura 1-4 se muestran las posibles membresías dentro de Internet2 en EE.UU. incluyendo miembros, no miembros y participantes de una avanzada red de backbone.

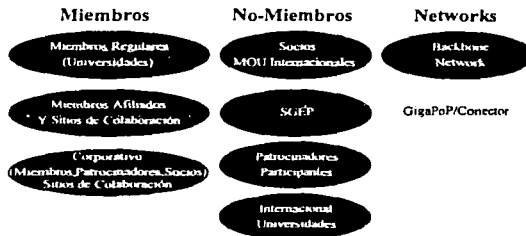


Figura 1-4 Membresías dentro de la red Internet2 en EE.UU.

### 1.3.2 Estructura de participación con Abilene

En el diagrama de la figura 1-5 se muestra la estructura de participación con la red de Abilene.

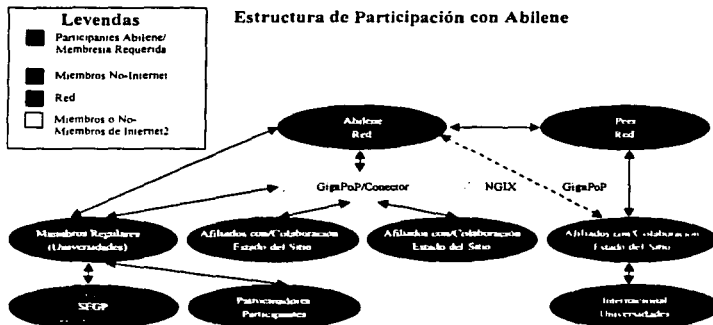


Figura 1-5 Estructura de Participación de la red Abilene en Internet2<sup>1</sup>.

<sup>1</sup> Diagrama tomado de la página: [www.internet2.edu/members/html/Internet2diagrams.html](http://www.internet2.edu/members/html/Internet2diagrams.html)

### 1.3.3 Abilene conectores/GigaPoP

- Los conectores de Abilene son centros de investigación e instituciones educativas que están conectadas directamente a la red de Abilene.
- Los conectores Abilene pueden ser GigaPoPs, universidades, miembros afiliados u otras redes regionales.

### 1.3.4 Miembros Regulares “Universidades”

Son universidades que se han unido al esfuerzo de Internet2 para desarrollar nuevas capacidades de red y aplicaciones avanzadas, necesarias para la investigación en este nuevo siglo. Las universidades miembro son responsables de los honorarios de Internet2 y de los costos de conexión con la red Abilene. Alrededor del mundo se han creado proyectos con iniciativas similares a estados unidos, a continuación se mencionan algunas de las más importantes alrededor del mundo y con las cuales la Corporación Universitaria para el Desarrollo de Internet (CUDI) tiene acuerdos<sup>2</sup>.

### 1.3.5 Corporaciones en Norteamérica

A continuación se mencionan algunos de las redes de Internet2 y corporaciones con que CUDI tiene algún acuerdo de colaboración en Norteamérica:

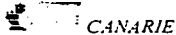


Abilene es un proyecto de la University Corporation for Advanced Internet Development (UCAID) desarrollado para soportar la iniciativa de Internet2. Abilene es una red troncal avanzada que conecta redes regionales a través de puntos de conexión, llamados GigaPoPs, para soportar el trabajo de desarrollo de aplicaciones avanzadas de Internet de las universidades afiliadas al proyecto de Internet2.

TESIS CON  
FALLA DE ORIGEN

<sup>2</sup> Fuente: [www.cudi.edu.mx](http://www.cudi.edu.mx)

El proyecto Abilene complementa otras redes de alto desempeño. Abilene proporciona conectividad a un avanzado backbone que soporta la demanda de los investigadores que incursionan en aplicaciones avanzadas de red en desarrollo por los miembros deUCAID. Abilene provee una red separada disponible para probar capacidades de red avanzadas anteriores a su introducción comercial. Estos servicios son esperados para incluir estándares de Calidad de Servicio (QoS) por sus siglas en ingles "Quality of Service". Multicasting, seguridad avanzada y protocolos de autentificación.



**TESIS CON  
FALLA DE ORIGEN**

CANARIE Inc. es la Organización Canadiense para el Desarrollo del Internet Avanzado. Fue establecida en 1993 y ha trabajado con el gobierno, la industria y las comunidades de investigación y educación para elevar la infraestructura de Internet en Canadá, el desarrollo de aplicaciones y su uso. Es una organización privada, sin fin de lucro soportada por la industria de Canadá, actualmente cuenta con 120 miembros y más de 500 proyectos.

CANARIE, Inc. continuará con el desarrollo de la red nacional óptica para Investigación y Desarrollo (R&D), CA\*Net 3, fundada en 1998 con presupuesto federal. Este proyecto es un intento para contribuir a la creación de una red de investigación sostenible para la comunidad académica de Canadá, y proveer de fabricantes de equipo, carriers y otros sectores de la tecnología de información y la comunicación.

● *STAR TAP*

STAR TAP (Puntos de Tránsito para Ciencia, Tecnología e Investigación) es una infraestructura persistente, fundada por la National Science Foundation Advanced Networking Infrastructure and Research division, la cual es parte de la Computer and Information Sciences and Engineering (CISE), que facilitará la interconexión a largo plazo y la interconectividad de redes avanzadas internacionales soportando aplicaciones, midiendo el desempeño, y evaluando tecnologías. El proyecto STAR TAP se ancla en el proyecto internacional vBNS.



A través de una asociación con MCI WorldCom, la NSF da un soporte fundamental a la investigación avanzada en Internet a través de la very High Performance Backbone Network Service (vBNS).

La red ha comenzado a operar con enlaces a 622 Mbps con la esperanza de llegar a 2.4 Gbps (2,400 Mbps) para finales de este año. Se espera que la red vBNS sea siempre capaz de soportar más información y de forma más rápida que las redes de telecomunicaciones comerciales actualmente disponibles.



La misión de UCAID, es la de facilitar y coordinar el desarrollo, despliegue, operación y transferencia de tecnología de redes basadas en aplicaciones y servicios avanzados para fomentar el liderazgo de los Estados Unidos en la investigación y educación superior, así como acelerar la disponibilidad de nuevos servicios y aplicaciones en Internet. Actualmente esta constituida por 159 universidades miembro.


● *NGI Next Generation*

La National Science Foundation (NSF) participa coordinando algunos esfuerzos separados que tienen como objetivo adentrarse al futuro de Internet mediante la utilización de medios de comunicación más rápidos, confiables y capaces de transmitir mayor información.

La NSF tiene un gran liderazgo en interconectividad de redes. La agencia creó la NSFNET en la década de 1980 —como un "backbone" de alta velocidad que utiliza la tecnología desarrollada por ARPAnet en los años setentas— para enlazar grupos de investigación con otros centros de supercómputo de los Estados Unidos. NSF involucró rápidamente a socios comerciales en el proceso de desarrollo con lo que ha dado comienzo a una importante nueva industria.

### 1.3.6 Corporaciones en Europa-Medio Oriente

A continuación se muestra un listado de corporaciones y países en Europa y Medio Oriente con quienes CUDI tiene algún entendimiento de colaboración<sup>3</sup>.

- |   |  |                  |
|---|--|------------------|
| • ARNES  | • GRNET                                | • HOCADU.net     |
| • BELNET  | • HEAnet                               | • POL34          |
| • CARTNE  | • <del>ARNET</del>                     | • RCTS           |
| • <del>ARNET</del>  | • INFN                                 | • REDIRIS        |
| • DANTE   | • Israel-IUCC <small>NET, INC.</small> | • <i>Renater</i> |
| • DFN-Verein  | • JISC                                 | • GARR           |
| • <i>Renater</i>  | • <del>ARNET</del>                     | • SWITCH         |
| • TEN-155   | • TEN-34                               | • TERENA         |

### 1.3.7 Corporaciones en Asia-Pacífico

A continuación se muestra un listado de corporaciones y países en Asia-Pacífico con quienes CUDI tiene algún entendimiento de colaboración<sup>4</sup>:

- |   |                                  |  |
|---|----------------------------------|--|
| • ACSys                                       | • Malasia                        | • Proyecto <del>ARNET</del>                                    |
| • APAN  | • Tailandia                      | • SingAREN   |
| • Australia                                   | • Taiwan                         | • TEMAN  |
| • China                                       | • Singapur                       | • CERNET <small>CHINA EDUCATION &amp; RESEARCH NETWORK</small> |
| • Corea                                       | • Japón                          | • Proyecto WIDE  |
| • Hong Kong                                   | • TANet: Taiwan Academic Network |  |
| • JAPAN ADVANCED INTERNET RESEARCH CONSORTIUM |                                  |  |

<sup>3</sup> Fuente: [www.cudi.edu.mx](http://www.cudi.edu.mx)

<sup>4</sup> Fuente: [www.cudi.edu.mx](http://www.cudi.edu.mx)

TESIS CON  
FALLA DE ORIGEN

#### 1.4 Internet2 en México

Previendo la importancia y las necesidades futuras para ejecutar bajo Internet aplicaciones cada vez más complejas, el Gobierno de México, la Comunidad Universitaria y la Sociedad de México toman la iniciativa de desarrollar y participar en el proyecto llamado Internet2. Este proyecto dotará a las nuevas y viejas aplicaciones de una mayor capacidad y velocidad para ejecutarse en las redes de datos mundiales. Esta participación permitirá el desarrollo de una nueva red de telecomunicaciones para crear nuevas generaciones de investigadores y de aplicaciones. Las principales acciones tomadas que dieron inicio a este proyecto son:

8 de Abril de 1999.- se oficializó en los Pinos la Constitución de la Corporación Universitaria para el Desarrollo de Internet (CUDI). El CUDI es una asociación civil, sin fines de lucro, que cuenta con miembros de los sectores académicos y empresariales, tanto públicos como privados. Tiene como propósito promover y coordinar el desarrollo de redes de telecomunicaciones y de cómputo, con capacidades avanzadas, enfocadas al desarrollo científico y educativo.

20 de Mayo de 1999.- en San Diego, California, representantes del CUDI firman 2 memorándums de entendimiento con 2 importantes corporaciones universitarias que promueven y coordinan la disponibilidad de redes avanzadas para aplicaciones de investigación y educación en los EE.UU., las cuales desde ese momento colaboran con el CUDI en el desarrollo de tecnologías y aplicaciones para el nuevo proyecto. Estas corporaciones son: UCAID (University Corporation for Advanced Internet Development) y CENIC (Education Network Initiatives in California).

20 de Mayo de 1999.- se firmó un convenio con TELMEX participando como Asociado Institucional.

En la tabla 1-1 se muestra la infraestructura institucional de cada una de las universidades que cuentan con una de las tecnologías más modernas de telecomunicaciones y cómputo, que utilizan tecnología de punta y medios de transmisión de alta velocidad; y que además son quienes dieron origen a esta organización:

	UNAM	IPN	ITESM	UAM	UDLA	UdeG	UANL	Total
Alumnos	269.000	164.000	80.000	45.000	6.500	158.000	110.000	832.500
Profesores	27.300	11.700	6.000	3.700	450	9.000	5.000	63.150
Planteles	36	63	3	3	1	88	72	293
Carreras	100	56	59	59	51	53	27	492
PCs	21.000	5.000	6.000	6.000	2.100	8.700	6.000	67.800

Tabla 1-1 Infraestructura Institucional en México.

Se ha pensado que para el desarrollo exitoso de Internet2 los participantes deben contar con:

Proyectos de aplicaciones avanzadas y con equipos de alta tecnología, una avanzada regulación y una moderna infraestructura de telecomunicaciones. La estrategia inicial de implantación de Internet2 en México se basa en la voluntad de las universidades (UNAM, IPN, UAM, ITESM, UDLA, UdeG Y UANL) de absorber, en proporción, el costo de instalar la red de alta velocidad necesaria que interconecte sus redes internas entre sí y con las universidades de alta velocidad en EE.UU. y Canadá.

#### 1.4.1 Organización

Como se mencionó en párrafos anteriores, el CUDI es el organismo que se encarga de manejar la red Internet2 en México, y su objetivo es establecer una infraestructura de telecomunicaciones entre las principales universidades del país, basada en medios de transmisión de alta velocidad para apoyar la investigación y la educación y permitir el desarrollo de aplicaciones para impulsar la nueva generación de Internet. Administrativamente el CUDI esta formado como lo muestra del esquema de la figura 1-6:

TESIS CON  
FALLA DE ORIGEN



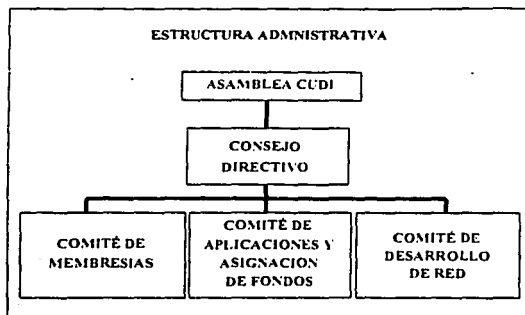


Figura 1-6 Estructura Administrativa de CUDI.

Los miembros de CUDI deberán estar comprometidos con el desarrollo, evolución, y utilización de facilidades de redes avanzadas y aplicaciones enfocados a la investigación y educación.

El CUDI puede recaudar fondos y equipos que contribuyan al desarrollo de la red y apoyar el desarrollo de aplicaciones. Además es el encargado de vigilar el manejo eficiente de los fondos recaudados y de la propia red.

En lo que respecta al equipo, la red se establece con: infraestructura de las empresas de telefonía de larga distancia (TELMEX y AVANTEL) y equipos terminales aportados por los fabricantes. la red tiene que cursar tráfico exclusivamente de carácter educativo o de investigación

El CUDI esta formado por 3 categorías de membresias:

- a) **Asociados Académicos.**- universidades con proyectos avanzados de educación e investigación y redes de alta velocidad.
- b) **Asociados Institucionales.**- empresas patrocinadoras.
- c) **Afiliados.**- universidades interesadas en el avance tecnológico sin infraestructura de telecomunicaciones de alta velocidad.

Para cada tipo de participante se establecen diferentes requisitos y derechos a los cuales se hacen responsables:

#### 1.4.2 Asociados Académicos

Requisitos:

- Instalar equipos de GigaPoP.
- Sufragar conectividad de banda ancha propia.
- Desarrollar y utilizar aplicaciones en educación e investigación.
- Participar en el consejo de administración.
- Aportar recursos humanos y financieros para el desarrollo de aplicaciones.
- Aportar cuotas requeridas.

Derechos:

- Participación en la asamblea de miembros.
- Utilizar la red de alta velocidad en aplicaciones aprobadas.
- Participación en el consejo directivo.
- Derecho de consulta a los reportes semestrales de avance de cada proyecto de investigación.
- Asistencia a reuniones de avance semestrales.
- Derecho de acceso al acervo de información de CUDI vía Internet.

#### 1.4.3 Asociados Institucionales

Requisitos:

- Aportar las cuotas establecidas en efectivo y especie.
- Aportar recursos para el desarrollo de aplicaciones específicas de su interés.

Derechos:

- Participación en la asamblea de miembros.
- Participación en el consejo directivo.
- Participación con investigadores en proyectos patrocinados.

- Derecho de consulta a los reportes semestrales de avance de cada proyecto de investigación.
- Asistencia a reuniones de avance semestrales.
- Derecho de acceso al acervo de información de CUDI vía Internet

#### 1.4.4 Afiliados

Las universidades que no cuenten con un GigaPoP propio podrán conectarse a la red dorsal como afiliados, pagando la infraestructura necesaria. También podrán participar como afiliados las personas morales del sector público, privado o social que deseen efectuar una aportación de menor cantidad de la que aportan los Asociados Institucionales.

Dentro de los Asociados Académicos se encuentran:

- Universidad Nacional Autónoma de México (UNAM).
- Instituto Politécnico Nacional (IPN).
- Universidad Autónoma Metropolitana (UAM).
- Instituto Tecnológico de Estudios Superiores de Monterrey (ITESM).
- Universidad Autónoma de Nuevo León (UANL).
- Universidad de Guadalajara (UdeG).
- Universidad de las Américas (Puebla) (UDLA).

Dentro de los Asociados Institucionales se encuentran:

- Consejo Nacional de Ciencia y Tecnología (CONACYT).
- Teléfonos de México (TELMEX).

Dentro de los Afiliados se encuentran:

- Instituto Tecnológico Autónomo de México (ITAM).
- Universidad Anáhuac del Sur.
- Universidad Autónoma de Chihuahua.
- Universidad Autónoma de Coahuila.
- Universidad Autónoma de Colima.
- Universidad Autónoma de Tamaulipas.



- Universidad Iberoamericana.
- Universidad tecnológica de México.
- Universidad del Valle de México.
- Centro de Investigación Científico y de Educación Superior de Ensenada (CICESE).

Debido a que la lista de miembros crece constantemente para obtener el listado completo se recomienda consultar la pagina de CUDI <http://www.cudi.edu.mx>.

#### 1.4.5 Aspectos técnicos

Se realiza un convenio importante con TELMEX, en el convenio se logra que TELMEX aporte el equipo principal para lograr la comunicación entre las instituciones del país y las estadounidenses. En él se especifica a TELMEX como: Asociado Institucional.

Actualmente TELMEX proporciona el backbone de la Red CUDI, incluyendo los enlaces y equipos entre los nodos de la red y sus conexiones con la red Internet2 en EE.UU. en 2 puntos, La red sólo se usará para aplicaciones de educación e investigación. El backbone planeado para la red de Internet2 en México se muestra en la figura 1-7.

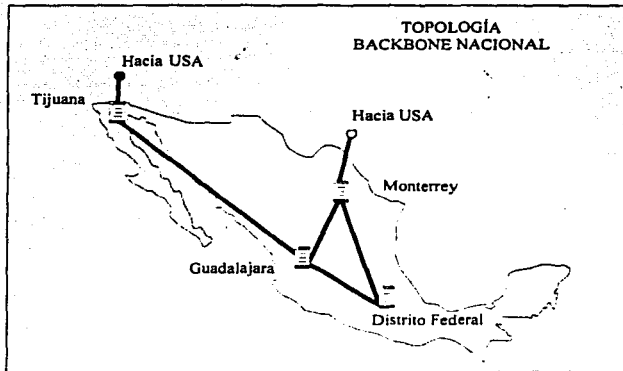


Figura 1-7 Backbone para la red Internet2 en México.

Sin embargo, en la primera parte de la implantación del proyecto en México, la red inicial tiene la siguiente estructura:

Los cuatro nodos localizados en las ciudades de México, Guadalajara, Monterrey y Tijuana estarán conectados entre sí por enlaces de capacidad de hasta STM-1 (155Mbps), hay equipos de alta capacidad de conmutación con tecnología IP/ATM y 2 conexiones de 155 Mbps hacia EE.UU., la construcción de la red se dará en 2 fases, la configuración de la fase inicial tendrá solo un punto de conexión con EE.UU. como se muestra en la figura 1-8.

TELEX no se hará responsable del equipo involucrado en la conexión a la red Internet2 que se encuentre físicamente en las instalaciones de los miembros del CUDI (asociados académicos y afiliados), La contribución de TELMEX no incluye el costo de los enlaces para conectar a los afiliados a la red.

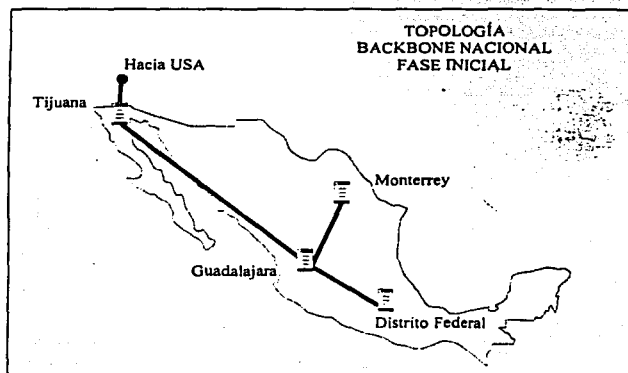


Figura 1-8 Topología del backbone nacional en su fase inicial.

TELMEX proporcionará un acceso E3 (155 Mbps) para conectar a cada uno de los siguientes asociados académicos hacia el backbone y cuya topología de acceso es ilustrada en la figura 1-9:

**TESIS CON  
FALLA DE ORIGEN**

1. UNAM en el campus de Ciudad Universitaria.
2. UAM en el campus de Azcapotzalco.
3. UdeG en su Administración General.
4. ITESM en el campus de Monterrey.
5. UANL en el campus de Ciudad Universitaria.
6. UDLA en el campus de Cholula, Puebla.

**TESIS CON FALLA DE ORIGEN**

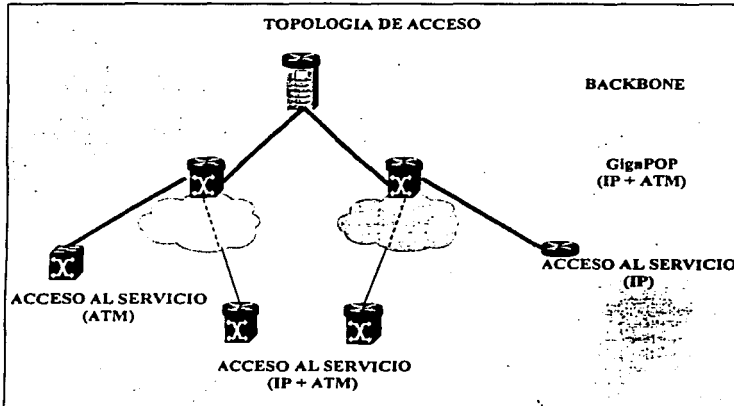


Figura 1-9 Topología de acceso a la red Internet2.

Respecto a la operación y administración de la red:

La operación de la red estará dividida en red física y red de servicios, TELMEX se hará cargo de la operación de la red física, esto es, atención a fallas, configuración, registro de uso, desempeño y mantenimiento de los enlaces de red e infraestructura del backbone.

El CUDI designará a la unidad responsable de la operación de los servicios de la red, esto es, atención a fallas, configuración, registro de uso, desempeño y seguridad de las aplicaciones que hagan uso de la red.

TELMEX apoyará al CUDI en la operación de los servicios en caso de contingencia y a solicitud de CUDI.

### 1.4.6 Backbone Actual

TESIS CON FALLA DE ORIGEN

El backbone actual se muestra en la figura 1-10, el cual es prácticamente el mismo que se definió en la fase inicial con la excepción del nodo de Cd. Juárez y la conexión a UTEP.

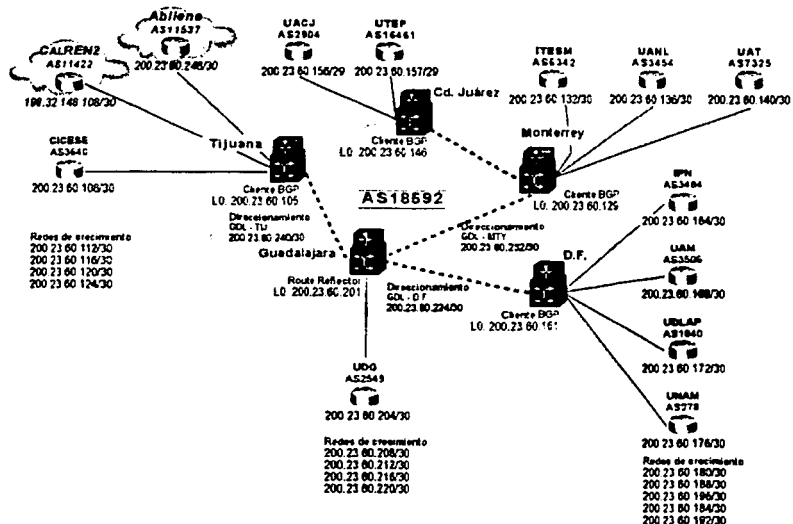


Figura 1-10 Backbone actual de la red de CUDI.

En últimas fechas se ha recibido la propuesta de incorporación de otro carrier "AVANTEL", con lo cual se extenderá la cobertura al doble del backbone actual, se agregará una conexión de alta velocidad a la red de VBNS y se conectarán directamente al backbone a otros ocho miembros asociados de CUDI.

#### 1.4.7 NOC de Internet2-México

Después de la creación del proyecto en México llamado Internet2. El CUDI tuvo la tarea de designar a una entidad como responsable de la operación de los servicios de la red, esto es, atención a fallas, configuración, registro de uso, desempeño y seguridad de las aplicaciones que hagan uso de la red, esta entidad es conocida como el Centro de Operación de la Red Internet2 (NOC-I2). La administración del NOC sería entonces delegada a una universidad de forma similar a la red de ABILENE y al contrario de la red VBNS en donde el NOC es administrado por una compañía privada, una de las principales razones para hacer esto es por el hecho de que al estar el NOC en una universidad, estaría formado por personas universitarias, siendo quienes conocen más las necesidades que tienen las redes de las universidades.

A la Universidad Nacional Autónoma de México se le ha delegado la administración del NOC-I2 debido a que ha sido pionera en la implementación de nuevas tecnologías de redes y en la administración de las mismas, así como en el desarrollo de Internet en México, lo que la hace contar con una moderna y robusta infraestructura de cómputo y telecomunicaciones, aunado a esto, la UNAM creó el primer Centro de Operación Académico, Mexicano. Por estas razones, la UNAM se estableció como el mejor candidato para albergar al Centro de Operación de Internet2.

Actualmente no se ha recibido la ratificación de la asignación del NOC-I2 a la UNAM, aún cuando esto no ha sucedido, el centro de operación incorpora de manera conjunta el uso de tecnología de punta y personal altamente capacitado.

El Centro de Operación de Internet2 actualmente se encuentra localizado en el campus de Ciudad Universitaria de la UNAM, México Distrito Federal, y es operado por personal de la Subdirección de Redes de la DGSCA-UNAM desde la formación del proyecto de CUDI.



#### 1.4.7.1 Funciones del NOC-I2

El Centro de Operación de la Red Internet2 (NOC-I2), es el encargado de mantener funcionando de manera eficiente la columna vertebral o "Backbone" de la red de CUDI. Con la finalidad de lograr este objetivo, el NOC-I2 cumple con las siguientes actividades.

##### *Soporte*

Entre las funciones para las que fue creado, se encuentra la de proporcionar apoyo a las instituciones que son miembro de esta red, así como a las que se van integrando; respecto a la forma en como deben de establecer la conexión desde sus redes hacia el backbone, respondiendo así a la problemática que se presenta con el incremento del número de nodos en la red.

##### *Monitoreo*

De igual manera es importante señalar que el NOC-I2 se encarga de evaluar el desempeño de la red, dentro de lo cual, el monitoreo es una de las partes fundamentales, que consiste en la recolección, extracción e interpretación de información relacionada con el comportamiento de la operación de la red. Este registro histórico permite que los administradores del NOC-I2 puedan determinar de manera más eficiente el comportamiento de la red Internet2 y con esto se pueda minimizar el impacto de fallas en la red o el deterioro del comportamiento de los protocolos.

##### *Generación de Estadísticas*

Con la finalidad de conocer el desempeño de los enlaces que conforman la red para planear su futuro crecimiento, se han implementado herramientas que permiten generar y visualizar de manera gráfica algunas de las variables más importantes involucradas en el funcionamiento de la red

##### *Administración de equipos de enrutamiento*

Una de las actividades más importantes que desempeña el Centro de Operación de Internet2, es la administración de los equipos de enrutamiento. Esta administración

involucra tareas de configuración, mantenimiento e implementación de nuevas tecnologías en el backbone, con la finalidad de impulsar la nueva generación de Internet.

*Seguimiento de Reportes (Troubleshooting)*

La atención de reportes de fallas es una actividad cotidiana. Una falla en la red, significa que algún dispositivo dentro de la misma no se desempeña como se espera. El NOC se encarga de minimizar el número de fallas y reducir el tiempo de resolución de las mismas.

Las actividades básicas a desarrollar durante el seguimiento de reportes son:

- Detección de la falla.
- Diagnóstico de la falla.
- Determinación de la falla.
- Resolución de la falla.



*Documentación Tecnológica*

Esta actividad representa un buen complemento en la formación profesional del staff del NOC-I2 y representa información técnica especializada que puede interesar a las personas relacionadas con el manejo de redes de datos.

TESIS CON  
FALLA DE ORIGEN

## Capítulo 2

# Conceptos Básicos

## Capítulo 2 Conceptos Básicos

### 2.1 Modelo OSI

OSI por sus siglas en inglés Open Systems Interconnections, es el modelo de referencia para comunicaciones. OSI es ampliamente definido como un conjunto de especificaciones de protocolos con muchas opciones para realizar la misma tarea: Los que participaron en la creación y desarrollo de OSI buscaron crear los protocolos que pudieran ser usados por todas las aplicaciones de redes. El gobierno de los EE.UU. fue aún más lejos al requerir el soporte de OSI en todas las computadoras que ellos comprarán (como lo fue hasta inicios de los 90s), este veredicto fue llamado GOSIP (Government OSI Profile), el cuál trataba de incentivar a los fabricantes para que escribieran código OSI en sus productos.

#### 2.1.1 El modelo OSI hoy en día

Todavía hoy en día algunos de los protocolos OSI permanecen alrededor del mundo, cuando el gobierno de los EE.UU., revoco la directiva GOSIP en Mayo de 1994, marcó el final de cualquier posible crecimiento e implementación duradera de OSI. Algunas de sus implementaciones actualmente es posible encontrarlas como legados del modelo OSI, por ejemplo muchos routers todavía utilizan el enrutamiento de redes OSI, incluso algunos dispositivos de ATM (Asynchronous Transfer Mode) utilizan direcciones de OSI NSAP (Network Service Access Point) para la señalización, de igual manera los equipos Digital DECnet Phase V usan muchas partes de OSI, incluyendo la capa de red, el direccionamiento y el enrutamiento de redes OSI. Actualmente el modelo OSI es usado prácticamente solo como modelo de referencia para la discusión de especificaciones de otros protocolos.

#### 2.1.2 Las capas del modelo OSI

El modelo OSI consiste de 7 capas, cada una de las cuáles pueden (y normalmente lo hacen) tener muchas subcapas. Los nombres de las capas del modelo OSI dan idea de la función que realizan.

Las capas superiores del modelo (Aplicación, Sesión, Presentación y Transporte) definen funciones enfocadas a las aplicaciones, siempre son implementadas en software y podemos

decir que solo se encuentran en los hosts, con la excepción de algunos switches que realizan algunas funcionalidades de capa 4 y superiores.

Las capas inferiores (Física, Enlace y Red) definen funciones enfocadas a la entrega de datos de extremo a extremo (*end-to-end delivery*). Las dos primeras son implementadas en hardware y la capa de red en software.

Cabe mencionar que cada capa en el equipo origen, se comunica con su capa adyacente en el equipo destino, esto lo hacen, agregando su propio encabezado (*header*) y opcionalmente pueden agregar un campo al final (*trailer*). Esto es necesario porque cada capa inserta cierta información, que sólo puede ser interpretada por la capa del mismo nivel en la máquina destino, figura 2-1. Existiendo entonces una comunicación vertical entre capas y horizontal entre capas del mismo nivel.

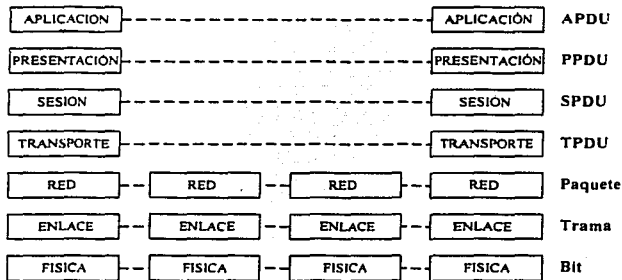


Figura 2-1 Comunicación entre capas adyacentes existente en la transferencia de datos entre el host destino y el host origen.

Las principales funcionalidades de cada capa son descritas a continuación:

*Aplicación (7):* Es la interfaz con el usuario, ejemplos de esta capa son las aplicaciones como telnet, FTP, WWW-browsers, etc.

*Presentación (6):* La principal función de esta capa es la de definir el formato de los datos, esto es, asegura que la información que se envía pueda ser comprendida por la capa de aplicación del sistema remoto. En esta capa se definen, por ejemplo: el tipo de código a utilizar ASCII o EBCDIC, si la información será compactada o incluso si se realizará algún cifrado a la información.

*Sesión (5):* Esta capa define el inicio, control y fin de la conversación (sesiones). Esto incluye el control y la administración de mensajes bidireccionales, haciendo que la aplicación sea notificada solo después de haberse completado una serie de mensajes. Ejemplo de un protocolo de esta capa es: NFS (*Network File System*) utilizado en los sistemas UNIX.

*Transporte (4):* Esta capa brinda la facilidad de utilizar protocolos para la detección y corrección de errores, también es la encargada de realizar el multiplexaje de los diferentes flujos de datos de las aplicaciones dentro del host. La reordenación de los datos es otra de sus funciones. Ejemplos de protocolos que pueden ser ubicados en esta capa son: TCP, UDP y SPX.

*Red (3):* Responsable de la entrega de los paquetes de datos de extremo a extremo. Para lograr esto, la capa de red define un direccionamiento lógico donde cada punto final debe ser identificado de forma única. Adicionalmente se define la forma en que deben de ser aprendidas las rutas para llegar a cierto destino. También define la forma en que deberán ser fragmentados los paquetes para poder ser enviados a través de los diferentes tipos de medios de transmisión.

*Enlace de Datos (2):* En esta capa se encuentran las especificaciones con relación a cierto tipo de enlace y métodos de acceso al medio. La capa de enlace define la entrega de datos a través de un enlace. Ejemplos de estos protocolos son la especificaciones de la IEEE 802.3 y 802.5 para redes LAN otros ejemplos son HDLC, PPP y Frame Relay usados comúnmente en redes WAN.

*Física (1)*: En esta capa la mayoría de las especificaciones provienen de organizaciones externas a las cuáles OSI hace referencia. Estas especificaciones son características físicas de los medios de transmisión: niveles de voltajes, corrientes, frecuencias, tipos de conectores codificación, etc.

### 2.1.3 Proceso de encapsulamiento del modelo OSI

El encapsulamiento es el proceso por el cuál los datos que se deben enviar a través de una red, se deben colocar en paquetes que se puedan administrar y rastrear, este proceso es mostrado en la figura 2-2.

Las tres capas superiores del modelo OSI (Aplicación, Presentación y Sesión) preparan los datos para su transmisión creando un formato común para la transmisión.

La capa de Transporte divide los datos en unidades de un tamaño que se pueda administrar, denominadas *segmentos*. También asigna números de secuencia a los segmentos para asegurarse de que los hosts receptores vuelvan a ensamblar los datos en el orden correcto.

La capa de Red encapsula el segmento creando un *paquete*. Le agrega una dirección de red destino y origen.

En la capa de Enlace de datos continúa el encapsulamiento del paquete, con la creación de una *trama*. A la trama se le agrega la dirección local (MAC) origen y destino. Luego, la capa de enlace de datos transmite los *bits* binarios de la trama a través de los medios de la capa física.

Cuando los datos se transmiten simplemente en una red de área local, se habla de las unidades de datos en términos de tramas, debido a que la dirección MAC es todo lo que se necesita para llegar desde el host origen hasta el host destino. Pero si se deben enviar los datos a otro host a través de una red interna o Internet, los paquetes se transforman en la unidad de datos a la que se hace referencia. Esto se debe a que la dirección de red del paquete contiene la dirección destino final del host al que se envían los datos.

Las tres capas inferiores (Red, Enlace de Datos y Física) del modelo OSI son las capas principales para el transporte de los datos a través de una red interna o de Internet. La

excepción principal a esto, es un dispositivo denominado gateway. Este es un dispositivo que ha sido diseñado para convertir los datos desde un formato, creado por las capas de aplicación, presentación y sesión, en otro formato. De modo que el gateway utiliza las siete capas del modelo OSI para hacer esto.

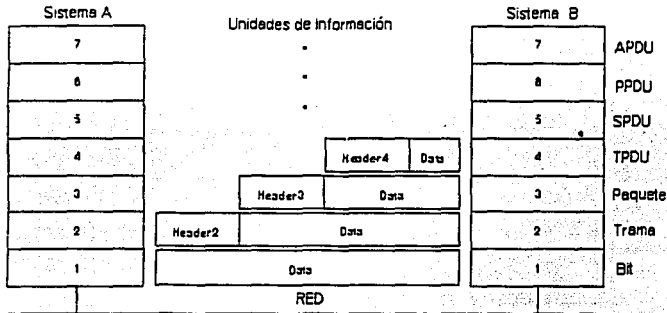


Figura 2-2 Proceso de encapsulamiento de los datos en el modelo OSI.

#### 2.1.4 Beneficios de la utilización de capas

Muchos beneficios se obtienen al dividir las funciones y tareas de las redes en pequeñas partes llamadas capas y definiendo las interfases de comunicación entre ellas. A continuación mencionaremos algunos de los principales beneficios de la fragmentación en el proceso de comunicación por capas son:

- Las personas pueden discutir y aprender más fácilmente los detalles acerca de las especificaciones de un protocolo.
- La normalización de las interfases entre las capas facilita la creación de una ingeniería modular.
- Se crea un mejor ambiente de interoperabilidad. Un vendedor puede escribir el software para las capas superiores y otro para las capas inferiores.
- Reduce la complejidad, permitiendo realizar cambios a los programas, lo cuál ayuda a una rápida evolución.



## 2.2 Descripción del protocolo IP

Es de suma importancia mencionar al protocolo IP, debido a la estrecha relación que tiene con el enrutamiento de paquetes que se lleva a cabo en redes de tecnología Internet. TCP/IP es un conjunto (Stack) de protocolos de comunicaciones de datos. Obtiene su nombre de dos de sus protocolos más importantes TCP (Transfer Control Protocol) e IP (Internet Protocol). TCP/IP fue adoptado en los años 80s como estándar para la red militar ARPANET.

Debido a la necesidad de establecer una comunicación global fue necesaria la creación de protocolos que pudieran comunicar equipos no importando la arquitectura, marcas, etc. Muchos protocolos se desarrollaron, sin embargo los protocolos que hicieron esto posible y con los cuáles esta funcionando la red internacional Internet, son la familia de protocolos de TCP/IP. Algunas de las principales características que hicieron posible que TCP/IP se convirtiera en el estándar para la red de alcance mundial Internet son:

- Es ideal para comunicar diferentes tipos de hardware y software, ya sea a través de Internet o en aplicaciones locales.
- Es independiente del nivel físico. TCP/IP integra diferentes tipos de redes como Ethernet, Token Ring, Frame Relay, etc.
- Provee de un esquema de direccionamiento común, capaz de identificar y establecer comunicación con cualquier otro dispositivo en la red, incluso en una red a nivel mundial.
- Estandariza los protocolos de alto nivel para proporcionar una interfaz que soporte cualquier tipo de aplicación de usuario.

### 2.2.1 Posición de TCP/IP de acuerdo al modelo OSI

A diferencia de OSI, TCP/IP se encuentra dividido en cuatro capas, en donde cada capa al igual que OSI, realiza funciones específicas. La capa más importante en el desarrollo de esta tesis, es la capa de Internet. El enrutamiento de paquetes toma lugar en esta capa con la interacción de protocolos como IP, ICMP, IGMP y protocolos de enrutamiento: OSPF, BGP, PIM, etc., como se ilustra en la figura 2-3.

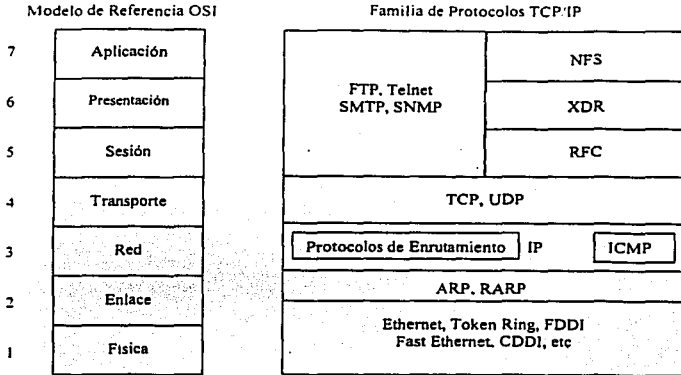


Figura 2-3 Comparación entre el modelo de referencia OSI y el modelo TCP/IP.

### 2.2.2 Formato del datagrama IP

El datagrama IP ilustrado en la figura 2-4 es la unidad básica de transferencia de datos entre el origen y el destino. viaja en el campo de datos de las tramas físicas de las distintas redes que va atravesando. Cada vez que un datagrama tiene que atravesar un router, el datagrama *saldrá* de la trama física de la red que abandona y se *acomodará* en el campo de datos de una trama física de la siguiente red. Este mecanismo permite que un mismo datagrama IP pueda atravesar redes distintas: enlaces punto a punto, redes ATM, redes Ethernet, redes Token Ring, etc. El propio datagrama IP también tiene un campo de datos en donde se transporta a los paquetes de las capas superiores.

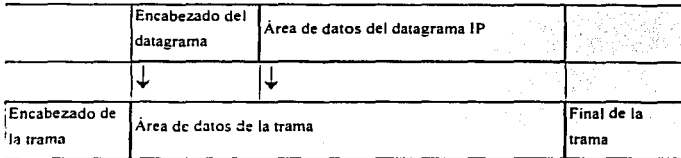


Figura 2-4 Datagrama IP.

**TESIS CON FALLA DE ORIGEN**

### 2.2.3 Campos del datagrama IP

La figura 2-5 muestra los campos del datagrama IP:

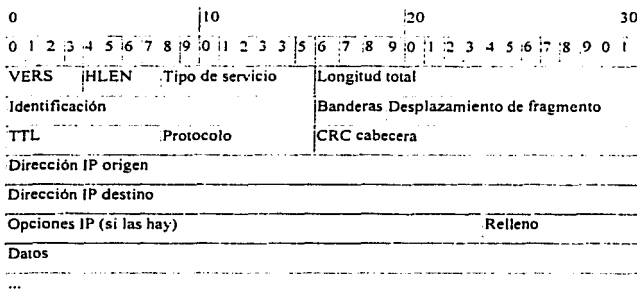


Figura 2-5 Campos de datagrama IP.

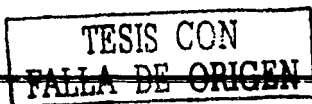
A continuación se hace una descripción de los campos que componen al datagrama IP:

**VERS** (4 bits). Indica la versión del protocolo IP que se utilizó para crear el datagrama. Actualmente se utiliza la versión 4 (IPv4) aunque ya se han definido la mayoría de las especificaciones de la siguiente versión, la versión 6 (IPv6).

**HLEN** (4 bits). Longitud de la cabecera expresada en múltiplos de 32 bits. El valor mínimo es 5, correspondiente a 160 bits = 20 bytes.

**Tipo de servicio (Type Of Service)**. Los 8 bits de este campo se dividen a su vez en:

**Prioridad** (3 bits). Un valor de 0 indica baja prioridad y un valor de 7 indica prioridad máxima.



Los siguientes tres bits indican la preferencia con la que se transmitirá el mensaje, es decir, son sugerencias a los routers que se encuentren a su paso, los cuáles pueden tenerlas o no en cuenta.

- *Bit D (Delay)*. Solicita retardos cortos (enviar rápido).
- *Bit T (Throughput)*. Solicita un alto rendimiento (enviar mucho en el menor tiempo posible).
- *Bit R (Reliability)*. Solicita que se minimice la probabilidad de que el datagrama se pierda o resulte dañado (enviar bien).
- Los siguientes dos bits no tienen uso.

*Longitud total* (16 bits). Indica la longitud total del datagrama expresada en bytes. Como el campo tiene 16 bits, la máxima longitud posible de un datagrama será de 65535 bytes.

*Identificación* (16 bits). Número de secuencia que junto a la dirección origen, dirección destino y el protocolo utilizado identifica de manera única un datagrama en toda la red. Si se trata de un datagrama fragmentado, llevará la misma identificación que el resto de los fragmentos.

*Banderas o indicadores* (3 bits). Sólo 2 bits de los 3 bits disponibles son utilizados actualmente. El bit de *Más Fragmentos (MF)* indica la existencia de más fragmentos que pertenecen al datagrama IP. Y el bit de *No Fragmentar (NF)* prohíbe la fragmentación del datagrama. Si este bit está activado y en una determinada red se requiere fragmentar el datagrama, éste no se podrá transmitir y se descartará.

*Desplazamiento de fragmentación* (13 bits). Indica el lugar en el cuál se insertará el fragmento actual dentro del datagrama completo, medido en unidades de 64 bits. Por esta razón los campos de datos de todos los fragmentos excepto el último tienen una longitud múltiplo de 64 bits. Si el paquete no está fragmentado, este campo tiene el valor de cero.

*Tiempo de vida o TTL* (8 bits). Especifica el número máximo de segundos que un datagrama puede existir en la red de redes. Cada vez que el datagrama atraviesa un router o

un host se resta 1 a este número. Cuando el valor del TTL llega a cero, el datagrama se descarta y se devuelve un mensaje ICMP de tipo "tiempo excedido" para informar al origen de la incidencia.

*Protocolo* (8 bits). Indica el protocolo utilizado en el campo de datos: 1 para ICMP, 2 para IGMP, 6 para TCP y 17 para UDP, etc.

*CRC cabecera* (16 bits). Contiene la suma de comprobación de errores sólo para la cabecera del datagrama. La verificación de errores de los datos corresponde a las capas superiores.

*Dirección origen* (32 bits). Contiene la dirección IP del origen.

*Dirección destino* (32 bits). Contiene la dirección IP del destino.

*Opciones IP*. Este campo no es obligatorio y especifica las distintas opciones solicitadas por el usuario que envía los datos (generalmente para pruebas de red y depuración).

*Relleno*. Si las opciones IP (en caso de existir) no ocupan un múltiplo de 32 bits, se completa con bits adicionales hasta alcanzar el siguiente múltiplo de 32 bits (recuérdese que la longitud de la cabecera tiene que ser múltiplo de 32 bits).

#### 2.2.4 Diferencias entre los modelos TCP/IP y OSI

- OSI define claramente las diferencias entre los servicios, las interfases, y los protocolos.

Servicio: la función que un nivel hace.

Interfaz: cómo se pueden acceder los servicios.

Protocolo: la implementación de los servicios.

TCP/IP no tiene una clara separación.

- Debido a que OSI fue definido antes de que los protocolos fueran implementados, los diseñadores no tenían mucha experiencia sobre el lugar en donde se debían

ubicar las funcionalidades, cabe mencionar que muchas de estas funcionalidades no se han ubicado. Por ejemplo, OSI originalmente no tiene ningún apoyo para broadcast.

- El modelo de TCP/IP fue definido después de los protocolos fueron definidos y por tal motivo estos se adecuan perfectamente. Esta adecuación no se tiene con otros stacks de protocolos.
- OSI no tuvo éxito debido a las siguientes razones:
  1. Mal momento de introducción: insuficiente tiempo entre las investigaciones y el desarrollo del mercado a gran escala para lograr la estandarización.
  2. Mala tecnología: OSI es complejo, es dominado por una mentalidad de telecomunicaciones sin pensar en computadores, carece de servicios sin conexión, etc.
  3. Malas implementaciones.
  4. Malas políticas: investigadores y programadores contra los ministerios de telecomunicación.
- Sin embargo, OSI es un buen modelo (no los protocolos). TCP/IP es un buen conjunto de protocolos, pero el modelo no es general.

### 2.2.5 Direccionamiento IPv4

Los equipos conectados a la red internacional de Internet con tecnología TCP/IP requieren de identificadores únicos, llamados direcciones IP. Sin embargo, cuando se tiene una red local TCP/IP sin conexión a Internet se puede asignar cualquier dirección IP válida, de lo contrario, si la red está conectada a la red internacional se deberá asignar una dirección de red IP única. Para garantizar esto existen organismos a los que se les ha delegado la asignación de direcciones IP, las cuáles reciben solicitudes de asignación de direcciones.

El formato de la dirección IP versión 4 está compuesta por 32 bits (4 bytes), representada mediante una notación decimal de la siguiente forma: W.X.Y.Z donde W, X, Y, Z pueden tomar el valor de 0 a 255. Un ejemplo sería: 132.247.1.253.

TESIS CON  
FALLA DE ORIGEN

La dirección IP esta constituida por dos partes: un identificador de red (NetID) y un identificador de equipo (HostID) Existen diferentes clases de redes IP: A, B, C y D mostradas en la tabla 2-1.

Clase	Formato (r=red, h=host)	Número de redes	Número de hosts por red	Rango de direcciones de redes	Máscara de subred
A	r.h.h.h	128	16.777.214	0.0.0.0-127.0.0.0	255.0.0.0
B	r.r.h.h	16.384	65.534	128.0.0.0-191.255.0.0	255.255.0.0
C	r.r.r.h	2.097.152	254	192.0.0.0-223.255.255.0	255.255.255.0
D	grupo	-	-	224.0.0.0-239.255.255.255	-
E	no válidas	-	-	240.0.0.0-255.255.255.255	-

Tabla 2-1 Clases de direcciones IPv4.

Otra forma de identificar a la clase de red a la que pertenece una dirección IP, es observando el valor de los primeros bits, la tabla 2-2 hace referencia a esto.

	0	1	2	3	4	8	16	24	31	
Clase A	0	Red				host				
Clase B	1	0	Red				host			
Clase C	1	1	0	red			host			
Clase D	1	1	1	0	grupo de multicast					
Clase E	1	1	1	1	(direcciones reservadas. no se pueden utilizar)					

Tabla 2-2 Clases de direcciones IP y sus rangos de valor posibles para las direcciones de red y de host.

TESIS CON  
FALLA DE ORIGEN

*Las direcciones de clase A*

Corresponden a redes que pueden direccionar hasta 16.777.214 máquinas cada una.

Las direcciones de red de clase A tienen siempre el primer bit con un valor de 0.

0 + Red (7 bits) + Máquina (24 bits)

Solo existen 124 direcciones de red de clase A.

*Las direcciones de clase B*

Las direcciones de red de clase B permiten direccionar 65.534 máquinas cada una.

Los dos primeros bits de una dirección de red de clase B son siempre 01.

01 + Red (14 bits) + Máquina (16 bits)

Existen 16.382 direcciones de red de clase B.

*Las direcciones de clase C*

Las direcciones de clase C permiten direccionar 254 máquinas.

Las direcciones de clase C empiezan con los bits 110.

110 + Red (21 bits) + Máquina (8 bits)

*Las direcciones de clase D*

Las direcciones de clase D son un grupo especial que se utiliza para dirigirse a grupos de máquinas (direcciones Multicast).

Las direcciones de clase D empiezan con los bits 1110.

*Las direcciones de clase E*

Estas direcciones son muy poco utilizadas.

Los cuatro primeros bits de una dirección clase E empiezan con los bits 1111.





*Direcciones IP especiales y reservadas*

No todas las direcciones comprendidas entre la 0.0.0.0 y la 223.255.255.255 son válidas para un host, algunas tiene significados especiales. Las principales direcciones especiales se resumen en la tabla 2-3. Su interpretación depende del host desde el que se utilicen.

Bits de red	Bits de host	Significado	Ejemplo
Todos 0		Mi propio host	0.0.0.0
Todos 0	Host	Host indicado dentro de mi red	0.0.0.10
Red	Todos 0	Red indicada	192.168.1.0
Todos 1		Difusión a mi red	255.255.255.255
red	Todos 1	Difusión a la red indicada	192.168.1.255
127	Cualquier valor valido de host	Loopback (mi propio host)	127.0.0.1

Tabla 2-3 Direcciones especiales IPv4.

Algunos rangos de direcciones de cada una de las clases de red han sido reservados y designados como rangos de direcciones "reservadas" o "privadas". Estas direcciones están reservadas para el uso de redes privadas y no son enrutadas en Internet. Son usadas normalmente por organizaciones con su propia *Intranet*<sup>1</sup>, pero incluso las redes pequeñas suelen encontrarlas útiles. Las direcciones de red reservadas se muestran en la tabla 2-4.

Clase	Rango de direcciones de redes reservadas
A	10.0.0.0 - 10.255.255.255
B	172.16.0.0 - 172.31.255.255
C	192.168.0.0 - 192.168.255.0

Tabla 2-4 Direcciones reservadas IPv4.

<sup>1</sup> *Intranet.*— Red privada que utiliza los protocolos TCP/IP. Puede tener salida a Internet o no. En el caso de tener salida a Internet, el direccionamiento IP permite que los hosts con direcciones IP privadas puedan salir a Internet pero impide el acceso a los hosts internos desde Internet. Dentro de una Intranet se pueden configurar todos los servicios típicos de Internet (web, correo, mensajería instantánea, etc.) mediante la instalación de los correspondientes servidores. La idea es que las intranets son como "Internets" en miniatura o lo que es lo mismo, Internet es una intranet pública gigantesca.

Si el tipo de red es una clase A o B, es posible obtener "subredes". Una subred es un rango de direcciones que forman parte de la red original. Por ejemplo, una red clase B se puede dividir en subredes, donde cada subred es del mismo número de direcciones que una red clase C, a este proceso se le conoce como subneteo. Para poder "subnetear" una red, es necesario utilizar una máscara de red que nos permita distinguir el NetID del HostID. La máscara de red es de 32 bits; y debe de estar formada por "unos" contiguos a partir de la izquierda para identificar al NetID y los bits restantes deben de ser "ceros" para identificar al HostID.

Una máscara de red permite dividir la parte del HostID en dos partes, por medio de la operación booleana AND bit por bit con lo que se obtiene:

La primera parte identifica al número de subred y la segunda parte identifica al host en esa subred.

*Dirección IP regular*

Identificador de Red	Identificador de Equipo
----------------------	-------------------------

*Dirección Subneteadas*

Identificador de Red	Identificador de Subred	Identificador de Equipo de la Subred
----------------------	-------------------------	--------------------------------------

Máscara de Red    11111111    11111111    11111111    00000000  
                          255.                    255.                    255.                    0

10000100 . 11111000 . 00000000 . 00000000 = 132.248.0.0 de manera binaria  
 11111111 . 11111111 . 11111111 . 00000000 = 255.255.255.0 de manera binaria

Para el mejor entendimiento del "subneteo" se ejemplifica un caso dentro de RedUNAM. La UNAM cuenta, entre otras redes, con una red clase B con dirección 132.248.0.0 y con máscara 255.255.255.0 que generan  $2^{16}$  (65536) direcciones para asignar a equipos. Debido a las necesidades de asignar direcciones a las diferentes dependencias de la UNAM se optó por dividir la red en 256 subredes con 256 hosts en cada subred, entonces la máscara resultante es 255.255.255.0, la tabla 2-5 contiene el resultado de este subneteo.

a) Formato Hexadecimal:

Red :	132.248.0.0		
Clase :	B		
Máscara Natural :	255.255.0.0		
Máscara Aplicada :	255.255.255.0	Máscara 24 bits	
No. de subredes :	256	Utilizables:	254
No. de hosts por subred :	256	Utilizables:	254

b) Formato Binario:

FORMATO BINARIO	DECIMAL	SIGNIFICADO
<b>Subred 0</b>		
11111111 . 11111111 . 00000000 . 00000000	132.248.0.0	NetID de la subred 0
11111111 . 11111111 . 00000000 . 00000000	132.248.0.1	Primera dirección
11111111 . 11111111 . 00000000 . 00000010	132.248.0.2	Segunda dirección
.	.	.
.	.	.
11111111 . 11111111 . 00000000 . 11111101	132.248.0.253	Penúltima dirección
11111111 . 11111111 . 00000000 . 11111110	132.248.0.254	Última dirección
11111111 . 11111111 . 00000000 . 11111111	132.248.0.255	Dirección broadcast subred
<b>Subred 1</b>		
11111111 . 11111111 . 00000001 . 00000000	132.248.1.0	NetID de la subred 1
11111111 . 11111111 . 00000001 . 00000000	132.248.1.1	Primera dirección
11111111 . 11111111 . 00000001 . 00000010	132.248.1.2	Segunda dirección
.	.	.
.	.	.
11111111 . 11111111 . 00000001 . 11111101	132.248.1.253	Penúltima dirección
11111111 . 11111111 . 00000001 . 11111110	132.248.1.254	Última dirección
11111111 . 11111111 . 00000001 . 11111111	132.248.1.255	Dirección broadcast subred 1
.	.	.
.	.	.

TESIS CON  
FALLA DE ORIGEN

De lo anterior resulta la siguiente tabla:

## TESIS CON FALLA DE ORIGEN

c)

No	ID RED	BROADCAST	RANGO			UTILIZABLE
1	132.248.0.0	132.248.0.255	132.248.0.1	-	132.248.0.254	NO
2	132.248.1.0	132.248.1.255	132.248.1.1	-	132.248.1.254	SI
3	132.248.2.0	132.248.2.255	132.248.2.1	-	132.248.2.254	SI
4	132.248.3.0	132.248.3.255	132.248.3.1	-	132.248.3.254	SI
5	132.248.4.0	132.248.4.255	132.248.4.1	-	132.248.4.254	SI
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
252	132.248.251.0	132.248.251.255	132.248.251.1	-	132.248.251.254	SI
253	132.248.252.0	132.248.252.255	132.248.252.1	-	132.248.252.254	SI
254	132.248.253.0	132.248.253.255	132.248.253.1	-	132.248.253.254	SI
255	132.248.254.0	132.248.254.255	132.248.254.1	-	132.248.254.254	SI

Tabla 2-5 (a), (b) y (c), Subneteo de una red clase C.

Existe una fórmula que ayuda a calcular el valor de la máscara y determinar el número de hosts y subredes que más convenga a las necesidades de cada red:

$N^{\circ}$  de hosts o  $N^{\circ}$  de subredes =  $2^n - 2$ , donde  $n = N^{\circ}$  de bits.

Como consecuencia del explosivo crecimiento de Internet, uno de los mayores problemas que enfrenta la comunidad de Internet es el agotamiento de direcciones IP; esto nos lleva a la implementación de nuevas estrategias en el manejo de direcciones IP: Variable Length Subnet Masks (VLSM) y Classless Inter-Domain Routing (CIDR). A continuación se describen estas estrategias.

### 2.2.6. VLSM (Variable Length Subnet Masks)

El término VLSM, se refiere a que una red puede ser configurada con diferentes máscaras de red. La idea de VLSM es ofrecer más flexibilidad para dividir —de acuerdo a las diferentes necesidades— la red en múltiples subredes utilizando diferentes máscaras de red para cada una de ellas y así tener un número adecuado de hosts en cada subred. Sin VLSM solo se puede aplicar una máscara a una subred.

Supóngase, por ejemplo, que se tiene la red clase C 192.214.11.0 y se necesita dividir esta red en tres subredes, con 100 hosts en una subred y las dos restantes con 50 hosts. Incluyendo las direcciones 0 y 255, teóricamente se tendrían disponibles 256 direcciones, que van desde la 192.214.11.0 hasta la 192.214.11.255. La división que se plantea no puede hacerse sin el empleo de VLSM.

Existen máscaras del tipo 255.255.255.X, que pueden ayudar a dividir la red clase C 192.214.11.0 en más subredes. La tabla 2-6 muestra algunos ejemplos de máscaras que pueden segmentar las 256 direcciones disponibles en más subredes.

252 (1111 1100)	64 subredes con 4 hosts cada una.
248 (1111 1000)	32 subredes con 8 hosts cada una.
240 (1111 0000)	16 subredes con 16 hosts cada una.
224 (1110 0000)	8 subredes con 32 hosts cada una.
192 (1100 0000)	4 subredes con 64 hosts cada una.
128 (1000 0000)	2 subredes con 128 hosts cada una.

Tabla 2-6 Posibles máscaras aplicadas a una red clase C.

Sin VLSM, se tendría que escoger el uso de una máscara 255.255.255.128 y dividir la red en dos subredes de 128 hosts cada una o usar la máscara 255.255.255.192 y dividir a la red en 4 subredes con 64 hosts, lo cual no cumple con los requerimientos.

Sin embargo utilizando múltiples máscaras se puede usar la máscara 255.255.255.128 para dividir la red en dos subredes con 128 hosts cada una y usar la máscara 255.255.255.192 para dividir a una subred de ambas en dos con 64 hosts cada una como se observa en la figura 2-6.

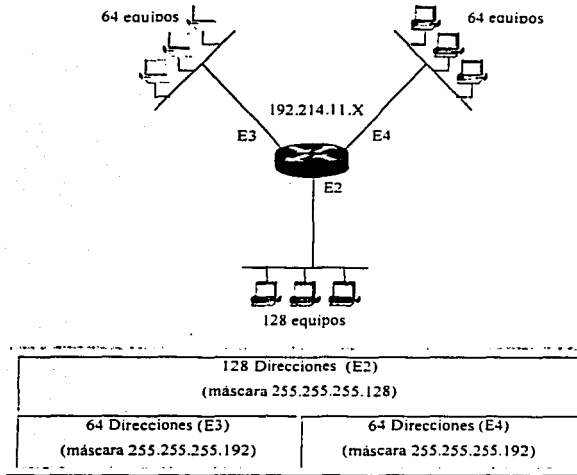


Figura 2-6 Subneteo de una red aplicando VLSM.

Sin embargo existen limitaciones, no todos los protocolos de enrutamiento pueden manejar VLSM. Por ejemplo RIP versión 1 e IGRP no pueden utilizar máscaras variables, pero protocolos como OSPF, EIGRP, ISIS y RIP versión 2 si soportan éste esquema.

TESIS CON  
FALLA DE ORIGEN

### 2.2.7 CIDR (Classless Inter-Domain Routing)

Recientemente, las tablas de enrutamiento IP en los routers de Internet han crecido en gran tamaño, provocando que éstos empiecen a saturarse tanto en procesamiento como en memoria (dos de los requerimientos más críticos en el enrutamiento). *"Se hicieron estudios de crecimiento los cuales indican que las tablas de enrutamiento se han duplicado en un lapso de 10 meses entre 1988 y 1991. Si no hubiese existido algún plan, las tablas de enrutamiento hubieran crecido aproximadamente a unas 80,000 rutas en 1995, sin embargo, en 1996 el tamaño de las tablas de enrutamiento fue de alrededor de 42,000 rutas"*<sup>2</sup>.

Este decremento en el crecimiento es atribuido a CIDR. CIDR ofrece una solución alternativa que pretende resolver el problema de direccionamiento IPv4. Este problema lleva consigo el incremento en las tablas de enrutamiento y el agotamiento de las direcciones de redes clase B.

En CIDR, una supernet está representada por un prefijo (dirección IP), junto con una "length"<sup>3</sup> (equivalente a la serie de bits con "unos" contiguos más significativos dentro de ésta dirección IP). A la representación *prefijo/length* se le llamará *agregado*. Por ejemplo la red 198.32.0.0, solía ser una red clase C ilegal, ahora es válida con la siguiente notación 198.32.0.0/16. La length "/16" indica que estamos utilizando 16 bits de máscara de red, empezando a contar desde la izquierda. La notación anterior es similar a tener 198.32.0.0 255.255.0.0.

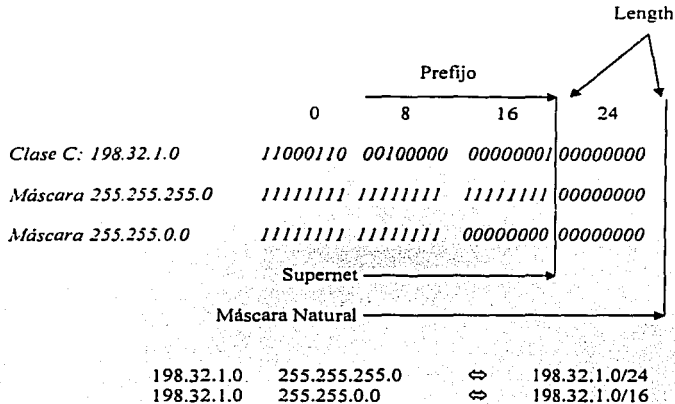
También en CIDR, una red es llamada supernet, cuando la máscara que se esté utilizando sea más pequeña a la máscara natural de esa red. Una red clase C 198.32.1.0, por ejemplo, tiene una máscara natural 255.255.255.0. La representación 198.32.0.0 255.255.0.0 puede denotarse como 198.32.0.0/16. (la cual es una máscara menor a la máscara natural de la clase C), por lo tanto es una supernet.

TESIS CON  
FALLA DE ORIGEN

<sup>2</sup> Dato obtenido del Libro Internet Routing Architectures, de Bassam Halabi.

<sup>3</sup> Notación decimal parecida a una máscara de red

Este esquema de direccionamiento es ilustrado a continuación:



Esta notación, nos permite agrupar a todas las redes específicas de la red 198.32.0.0 (como lo son la 198.32.1.0, 198.32.2.0 y así sucesivamente) en un solo anuncio llamado agregado.

Todas las redes que son parte de un bloque CIDR son llamadas prefijos "mas específicos" porque proporcionan más información acerca de la localización de la red. Los prefijos más específicos tienen una length más grande que el agregado como se observa:

198.213.0.0/16      Agregado con una length de 16 bits.

198.213.1.0/20     Prefijo más específico con length de 20 bits.

TESIS CON  
FALLA DE ORIGEN



### 2.3 Enrutamiento

Internet está formado por una serie de Sistemas Autónomos (AS)<sup>4</sup>, que definen políticas de administración y de enrutamiento de diferentes organizaciones. Un Sistema Autónomo ocupa Protocolos de Compuerta Interna (IGPs Interior Gateway Protocol), tales como RIP, OSPF, ISIS, etc. para enrutar información del mismo Sistema Autónomo. Estos a su vez se interconectan vía un Protocolo de Compuerta Externa (EGP Exterior Gateway Protocol), tales como EGP y BGP para intercambio de información entre Sistemas Autónomos.

Un router es un dispositivo de red de propósito específico, que a través de los protocolos de enrutamiento IGPs y EGPs avanza los paquetes de información de un lugar origen a un destino, determinando la mejor ruta. Este equipo actúa sobre la capa 3 (Red) del modelo OSI. Los routers construyen tablas de enrutamiento conteniendo información de los mejores caminos (paths), a todos los destinos que conocen.

Existen dos tipos de enrutamiento: El directo y el indirecto.

#### *Directo:*

Es la transmisión directa de un paquete de información de una máquina a otra, siempre y cuando ambas máquinas pertenezcan al mismo segmento lógico y por lo tanto físico.

#### *Indirecto:*

Se presenta cuando las máquinas que se quieren comunicar no están en la misma red (lógica y física), lo cual se traduce en la necesidad de utilizar un equipo adicional, que se encargue de comunicar la red origen al destino.

Un router se comunica con otro router, con el fin de intercambiar tablas de enrutamiento, información de control y diversos mensajes. Este intercambio de información se da gracias a los algoritmos de enrutamiento.

---

<sup>4</sup> Un Sistema Autónomo es un conjunto de redes bajo una administración común, que comparte políticas de enrutamiento estratégicas.

Los algoritmos de enrutamiento comúnmente emplean una tabla de enrutamiento en la que almacenan información referente a los posibles destinos y la forma en como llegar a ellos. Las tablas se generan a partir de dos procesos: Iniciación del proceso de enrutamiento e intercambio de tablas con otros routers.

### 2.3.1 Características de Diseño

Existen diferentes algoritmos de enrutamiento y cada uno de ellos fue diseñado con un propósito en particular, es por eso que tienen diferentes impacto sobre la red y sobre los recursos de los routers, cada uno de estos algoritmos utilizan una variedad de *métricas*<sup>5</sup>, que afectan directamente en el cálculo de la mejor ruta. Los algoritmos de enrutamiento generalmente tienen una o más de las siguientes características de diseño:

#### *Robustez*

La robustez nos indica que el algoritmo de enrutamiento debe de soportar una gran cantidad de rutas en las tablas de enrutamiento, además debe ser tolerante a malas implementaciones y a fallas de hardware.

#### *Estabilidad*

Esta característica de diseño nos indica que los anuncios de enrutamiento que se encuentran en nuestros equipos, deben de ser coherentes con lo que está aconteciendo en la red en todo momento.

#### *Flexibilidad*

Los algoritmos de enrutamiento deben adaptarse rápidamente a la gran variedad de circunstancias que ocurran en la red, es decir, los algoritmos deben de ser programados para adaptarse a los cambios en la red como son el ancho de banda, tamaño de colas, retardos en la red, y otras variables.

---

<sup>5</sup> Métrica. Es una medida de confiabilidad asignada a un anuncio de enrutamiento. Se calcula por medio de una evaluación de ciertas variables tales como: Anchos de Banda, Retardo, Costo, Número de Saltos, Confiabilidad y Carga. Que intervienen en la determinación de la ruta óptima para poder llegar a un destino y la forma de calcularse depende del algoritmo que se esté utilizando.

### *Simplicidad y Bajo Overhead*

El algoritmo de enrutamiento debe de ser lo más simple posible, en otras palabras, debe ofrecer funcionalidad y eficiencia con un mínimo de recursos. La eficiencia es particularmente importante cuando la implementación del algoritmo de enrutamiento se ejecuta en un router con recursos físicos limitados. Por ejemplo el router no debe de utilizar mucho ancho de banda para transportar sus anuncios hacia otro router.

### *Rápida Convergencia*

La convergencia es el proceso de conformidad por todos los routers sobre las rutas óptimas, es decir, cuando un anuncio de una red deja de ser anunciada por un router, se deben de mandar mensajes de actualización de rutas para recalcular rápidamente la ruta óptima, para poder llegar a ese destino por otra trayectoria. Si no existiera una convergencia rápida, se provocaría una inconsistencia en los anuncios de las redes. La convergencia se lleva a cabo una vez que los routers poseen la misma información de enrutamiento.

### *Optimización*

Si tenemos varias rutas en nuestra tabla de enrutamiento para poder llegar a un destino determinado, el protocolo de enrutamiento debe de ser lo suficientemente inteligente para poder seleccionar la mejor ruta.

## **2.3.2 Clasificación**

Los algoritmos de enrutamiento pueden ser clasificados en base a varios criterios tales como:

- Estáticos o Dinámicos
- Single-Path o Multipath
- Plano o Jerárquico
- Intradominio o Interdominio
- Link State o Distance Vector

### 2.3.2.1 Estáticos o Dinámicos.

En los algoritmos de enrutamiento estático el administrador de la red crea manualmente una tabla de enrutamiento. Dicha tabla no cambia a menos que el administrador lo haga. Los algoritmos que usan rutas estáticas son fáciles de diseñar y trabajan bien en ambientes en donde el tráfico de la red es relativamente predecible y el diseño de la red es relativamente simple.

Actualmente este tipo de algoritmos no son utilizados debido a que no son capaces de detectar los cambios que acontecen en la red.

Los algoritmos de enrutamiento dinámico van ajustando las rutas dinámicamente en tiempo real, gracias a que analizan los mensajes de actualización de rutas. Este tipo de algoritmos también permite la implementación de rutas estáticas cuando sean necesarias.

### 2.3.2.2 Single-Path o Multipath.

Existen protocolos de enrutamiento más sofisticados que aceptan múltiples rutas para un mismo destino. Estos algoritmos son llamados Multipath, y pueden balancear cargas a través de múltiples líneas. Los algoritmos Single-Path, por el contrario, solo aceptan una sola ruta por cada red.

### 2.3.2.3 Plano o Jerárquico.

En sistemas de enrutamiento plano, todos los routers son parejas de todos. En un sistema jerárquico, existen diferentes niveles, en donde el nivel más alto es el nivel de backbone, los routers se encuentran interconectados con otros routers de backbone y con algunos de los routers en el siguiente nivel. Los routers en los siguientes niveles se conectan con otros routers en este nivel creando un grupo de routers o área, también se conecta con uno o varios routers del backbone pero no con todos. Para la comunicación con otras áreas es necesario que la información primero sea enviada al nivel de backbone y después al área destino.

TESIS CON  
FALLA DE ORIGEN

#### 2.3.2.4 Link State o Distance Vector

Los algoritmos Link State (conocidos como algoritmos Shortest Path First), envían su información de enrutamiento a todos los nodos de la red. Sin embargo, cada router envía solo una porción de su tabla de enrutamiento, la cual describe el estado de sus interfaces. Los algoritmos Distance Vector (conocidos como algoritmos Bellman-Ford), envían toda la tabla de enrutamiento solamente a sus vecinos. Los algoritmos de tipo Link State son menos propensos a loops de enrutamiento, porque convergen más rápidamente que los algoritmos Distance Vector. Una de las desventajas los algoritmos Link State, es que su ejecución requiere del uso de más memoria y CPU.

#### 2.3.2.5 Intradominio o Interdominio

Algunos algoritmos de enrutamiento trabajan solo dentro de dominios (un dominio es un grupo de routers que hablan el mismo protocolo de enrutamiento, normalmente coincide con el Sistema Autónomo), otros trabajan dentro y entre dominios, es decir, los routers que trabajan con algoritmos Intradominio solo intercambian tablas de enrutamiento dentro de su Sistema Autónomo. en cambio los routers que utilizan algoritmos Interdominio pueden intercambiar sus tablas de enrutamiento dentro y fuera de su Sistema Autónomo.

#### 2.4 Jerarquía en capas

El éxito de una red se basa en permitir su crecimiento sin que se comiencen a presentar problemas de saturación, por esta razón se vuelve indispensable un diseño jerárquico que divida sus funciones en capas.

Si una red es diseñada en forma jerárquica, cada una de las capas subsecuentes actuará como un filtro, permitiendo que la red pueda crecer de una manera eficiente. De esta forma se logra mantener al "tráfico local" de manera local, y solo los datos e información acerca de recursos globales que necesiten viajar fuera del dominio inmediato sean quienes tengan que hacerlo, evitando así que el tráfico local consuma los recursos de la red de forma inadecuada. El número de capas que son necesarias depende del control de tráfico que se requiera. Para determinar el número de capas requeridas, se debe de identificar la función que tendrá cada una de ellas.

Cada capa en un diseño jerárquico es responsable de prevenir que tráfico innecesario sea enviado a las capas superiores. La meta es permitir que solo el tráfico relevante pueda atravesar la red y con esto reducir la carga en la red. Si la meta es claramente conocida, la red puede escalar más eficientemente. Un esquema que normalmente es suficiente, es un modelo jerárquico en tres capas, tales capas son:

- La capa de Acceso
- La capa de Distribución
- La capa de Core

#### 2.4.1 Capa de Acceso

La capa de Acceso es aquella en donde se conectan los dispositivos de orilla, los cuáles conectan a las redes de los usuarios, en donde son necesarias grandes densidades de puertos. Los dispositivos de capa 3 (como los routers) son los responsables de la entrada y salida de tráfico, asegurando que el tráfico local no se extienda por toda la red. En ocasiones dependiendo del tipo de usuario llegan a ser necesarias conexiones redundantes a la capa de distribución.

#### 2.4.2 Capa de Distribución

La capa de Distribución es responsable de la conexión al Core filtrando todas las actualizaciones y seleccionando en forma granular el acceso de los usuarios. Una de sus funciones es la seguridad. Se recomienda tener conexiones redundantes al Core, aunque en ocasiones llegan a ser necesarias conexiones redundantes hacia los usuarios o entre dispositivos en la misma capa.

#### 2.4.3 Capa de Core

La capa de Core tiene como función principal garantizar la conectividad a lo largo de toda la red. Su característica principal es la confiabilidad. Una falla a este nivel produce que grandes partes de la red queden incomunicados. Para asegurar una conectividad continua se debe diseñar con altos niveles de redundancia tantos como sean posibles (preferentemente en una topología de full-mesh o parcial-mesh), también se debe de eliminar toda la latencia que sea posible, debido a esto, no es recomendable que se tengan que tomar decisiones

complejas de enrutamiento como lo podrían ser los filtros. Por esto se recomienda que los filtros se implementen en las capas de Acceso y Distribución dejando al Core la responsabilidad de transportar los datos de la forma más rápida y confiable. En algunas ocasiones se implementan ciertas políticas de QoS (Calidad de Servicio).

La figura 2-7 es un diseño que muestra cada una de las capas y la forma en como se puede llegar a tener conexiones redundantes entre ellas o dentro de la misma capa.

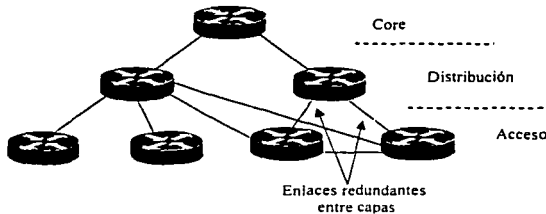


Figura 2-7 Jerarquía en las capas de una red.

## 2.5 Tipos de Backbone

Una característica muy importante en el diseño del backbone, es definir los tipos de servicios que se van a proporcionar; voz, datos y video, y que tecnología se utilizará para soportarlos. Podemos hacer una clasificación de los backbones de acuerdo al modelo OSI, si va a ser construido para brindar servicios de capa 2 se trata de un backbone "Switchado" y cuando se construye para soportar servicios de capa 3 se habla de un backbone "Enrutado".

### 2.5.1 Backbone Switchado

Este tipo de backbone se crea utilizando tecnologías como Frame Relay y/o ATM, el Core esta formado por switches que a su vez se encuentran rodeados por Routers. Algunas de las principales ventajas de este tipo de Backbone son; se puede tener un control más específico en cuanto asignación de anchos de banda y calidades de servicio (QoS), son tecnologías

orientas a circuitos. Sus principales desventajas son el enrutamiento y la solución de problemas complejos. la mayoría de sus características crecen en complejidad de forma exponencial. figura 2-8.

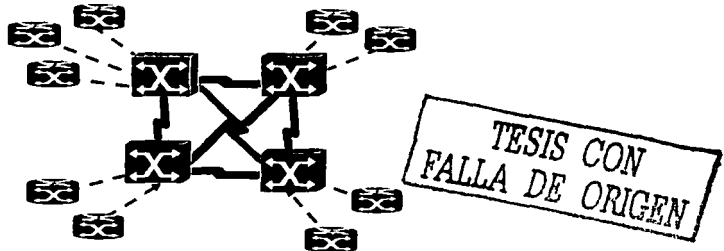


Figura 2-8 Backbone Switchheadado.

### 2.5.2 Backbone Enrutado

La infraestructura de este tipo de backbone esta formada por routers, las conexiones entre ellos son realizadas con enlaces PPP y/o HDLC. Las configuraciones de enrutamiento son más fáciles de realizar al igual que la solución de problemas, además también se obtiene un menor *overhead*<sup>6</sup>, ver figura 2-9.

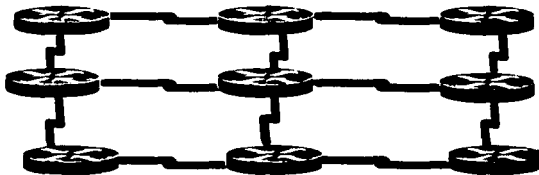


Figura 2-9 Backbone Enrutado.

<sup>6</sup> Es información de control que requieren los equipos de comunicaciones, pero que no es parte de la información de los usuarios. Es un parámetro que se debe de minimizar debido a que consume recursos de la red.



# Capítulo 3

## Protocolo de

# Enrutamiento OSPF

TESIS CON  
FALLA DE ORIGEN

---

## Capítulo 3 Protocolo de Enrutamiento OSPF

### 3.1 Historia de OSPF

OSPF (Open Shortest Path First) fue desarrollado por la IETF (Internet Engineering Task Force) como respuesta a la comunidad de Internet para introducir un nuevo protocolo de compuerta interna (IGP) de alta funcionalidad, que no fuera una implementación propietaria y que pudiera reemplazar al protocolo de enrutamiento RIP, debido a las deficiencias que presenta principalmente para escalar. Actualmente OSPF es el protocolo recomendado por la IETF para ser usado como protocolo intradominio. Los estudios comenzaron en el año de 1988, pero fue hasta el año de 1991 en que se formalizaron las especificaciones de este protocolo.

OSPF es un protocolo de estado de enlace, que utiliza el algoritmo de Dijkstra SPF (Shortest Path First) y que es abierto (Open) esto significa que no es propietario de alguna organización o compañía. OSPF envuelve un gran número de RFCs, de los cuales muchos fueron escritos por John Moy. La primer versión del protocolo fue especificada en el RFC 1131; esta versión nunca fue más allá de los laboratorios. La versión 2, la cual es la que se utiliza actualmente, tiene su primera especificación en el RFC 1247 y la más reciente en el RFC 2328.

### 3.2 Descripción preliminar de OSPF

Como todos los protocolos de estado de enlace, la mayor ventaja de OSPF sobre los protocolos vector distancia es su rápida convergencia, soporte para redes de gran tamaño y una baja susceptibilidad para conservar información errónea de enrutamiento. Otras características de OSPF son:

- El uso de áreas, la cuales reducen impacto del protocolo en la memoria y tiempo de CPU, contienen el flujo que debe seguir el tráfico de acuerdo a la información de enrutamiento y hace posible la construcción de una red de topología jerárquica.
- Un completo comportamiento Classless, eliminando los problemas presentados en las redes Classful, como son el uso de redes discontinuas.

- Soporte de rutas de tipo Classless, soporte para el uso de subredes de mascarar variables conocido como VLSM (Variable Length Subnet Mask) y Supernet para lograr una administración más eficiente del espacio de direcciones.
- Una métrica arbitraria sin dimensiones.
- Uso de un "costo" igual para lograr un balanceo de cargas más eficiente de enlaces múltiples paralelos.
- Usa un espacio reservado de direcciones de Multicast, para reducir el impacto producido en dispositivos que no necesitan escuchar los anuncios de OSPF.
- Soporte de Autenticación para hacer más seguro el enrutamiento.
- El uso de una etiqueta para poder realizar el seguimiento de rutas externas a OSPF.
- Soporte de autenticación al poder definir llaves de encriptación entre los routers que intercambian información.

### 3.3 Operación de OSPF

La operación de OSPF en una forma general se puede describir brevemente de la siguiente manera:

1. Todos los routers hablando OSPF envían paquetes de "Hello" fuera de sus interfaces. Si dos routers comparten un medio con ciertos parámetros específicos dentro de sus mensajes de Hello se convierte en vecinos.
2. Adyacencias (Adjacencies), las cuales incluso pueden ser conexiones virtuales punto a punto entre ciertos vecinos. OSPF define muchos diferentes tipos de routers. El establecimiento de las adyacencias es determinado por la forma de intercambio en los mensajes de Hello y por el tipo de redes sobre las cuales los mensajes son intercambiados.
3. Cada router anuncia el estado de sus interfaces por medio de mensajes de LSAs (*link State Advertisements*). Los LSAs informan sobre el estado en que se encuentran todos los enlaces (interfaces) del router. Estos enlaces pueden ser a redes aisladas (Stub), ósea que no contienen otros routers, hacia redes en otras áreas, o hacia redes externas (redes aprendidas por otro protocolo). Como OSPF tiene diferentes tipos de enlaces, OSPF define diferentes tipos de LSAs.

4. Cada router que recibe un LSA de un vecino, lo almacena en una base de datos conocida como *link-state database* y envía una copia del LSA a todos sus vecinos.
5. Por medio de una inundación (Flooding) de LSAs en un área, todos los routers construyen una base de datos idéntica.
6. Cuando las bases de datos están completas, cada router usa el algoritmo de SPF para calcular una gráfica libre de loops describiendo los caminos más cortos (menor costo) a cada uno de los caminos conocidos, con él como raíz. Esta gráfica es conocida como el árbol de SPF.
7. Cada router construye su tabla de enrutamiento de acuerdo al árbol de SPF.

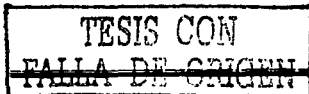
Una vez que la información de todos los enlaces ha sido enviada por medio de una inundación a todos los routers dentro de un área, es entonces cuando la información de todas las bases de datos se encuentran sincronizadas y finalmente los routers construyen su tabla de enrutamiento. OSPF es un protocolo llamado. Los mensajes de Hello son usados como mensajes de Keepalive, los LSAs son retransmitidos cada 30 minutos, si la topología se ha mantenido sin cambios.

Antes de que los LSAs puedan ser enviados, los routers de OSPF deben descubrir a sus vecinos y establecer adyacencias. Los vecinos deben ser almacenados dentro de una tabla de vecinos junto con la interfaz en la cual se encuentran localizados, además de información adicional relacionada con ellos.

### 3.4 Vecinos y Adyacencias

Se dice que dos routers son vecinos si tienen enlaces conectados a una misma red. Esto es, que se encuentran conectados al mismo segmento físico. El descubrimiento de vecinos se da por medio de los mensajes de Hello.

Las adyacencias se forman cuando vecinos han intercambiado información, tienen la misma base topológica y sus bases de datos se encuentran sincronizadas.



Para poder realizar un seguimiento de cada uno de los routers, es necesario que todos los routers tengan un identificador único (un router id), una forma de asignar este router ID es la selección de un dirección IP, la cual es un identificador único. Una práctica común es la utilización de interfases lógicas normalmente nombradas por los fabricantes como interfases de Loopback, a las cuales se les asigna una dirección IP para ser usadas como identificador del router, a estas interfases también es común asignarles el nombre de *Loopbacks de enrutamiento* por que su principal propósito es servir como identificador del router, una ventaja de utilizar este tipo de interfases es no tener una dependencia física, debido a esto, la interfaz siempre se encuentra en operación.

### 3.5 El protocolo de Hello

Las principales funciones de este protocolo son:

- Ser el mecanismo para el descubrimiento de vecinos.
- El anuncio de varios parámetros con los cuales dos routers deben estar de acuerdo para poder ser vecinos.
- Asegurar una comunicación bidireccional entre dos vecinos.
- La elección de un Router Designado (DR-Designated Router) y el respaldo del designado (Backup DR) en redes Broadcast y Nonbroadcast multiacceso.

Los routers que hablan OSPF envían periódicamente mensajes de Hello desde las interfases que tienen habilitado OSPF. Este periodo es conocido como *HelloInterval* y el tiempo puede ser configurado, un valor típico es de 10 segundos. Si un router no escucha un mensaje de Hello desde su vecino dentro de un cierto periodo, considerará que su vecino esta fuera de operación, este periodo de tiempo es conocido como *DeadInterval*, un valor común para este tiempo es cuatro veces el valor del HelloInterval.

Cada paquete de Hello contiene la siguiente información:

- El router ID del router que origino el mensaje.
- El área ID de la interfaz del router que origino el mensaje.
- La mascara de la interfaz que origino el mensaje.

TESIS CON  
FALLA DE ORIGEN

- El tipo de autenticación y la información de autenticación para la interfaz que lo origino.
- El HelloInterval de la interfaz originante.
- El DeadInterval de la interfaz originante.
- La Prioridad del router.
- El DR y el BDR.
- Cinco banderas de bits con un significado opcional.
- Los router IDs de los routers vecinos originantes. Esta lista solo contiene los routers desde los cuales los mensajes de Hello fueron avanzados en las interfases originantes dentro del último Router DeadInterval.

Cuando un router recibe un mensaje de Hello desde un vecino, este verificará el área ID, la autenticación, la mascara de red, el HelloInterval, el RouterDeadInterval y otros valores opcionales que se hayan configurado en las interfases. Si esto no se puede realizar entonces el paquete es desechado y no se establece una adyacencia.

Pero si todo concuerda, el mensaje de Hello es declarado válido. Si el ID del router originante ya se encuentra listado dentro de la tabla de vecinos de la interfaz donde se recibió el mensaje, el reloj del RouterDeadInterval es reiniciado. Si el router no se encuentra dentro de la tabla de vecinos, entonces es agregado a la tabla.

Cuando un router envía un mensaje de Hello incluye en el mensaje la siguiente información: los router ID de todos los vecinos listados en el enlace en el cual el paquete será transmitido. Si un router recibe un mensaje de Hello en el cual encuentra su propio Router ID, el router sabe que han sido establecidas dos vías de comunicación (Two-way).

Una vez que las dos vías de comunicación se han establecido, las adyacencias deben de establecerse. Pero como lo mencionamos anteriormente, no todos los vecinos pueden llegar a formar adyacencias. El tipo de red también determina el tipo de paquetes que OSPF debe transmitir.

### 3.6 Soporte de OSPF a diferentes tipos de redes

OSPF define cinco tipos diferentes de red:

- Redes tipo Punto a Punto (Point to Point).
- Redes tipo Broadcast.
- Redes tipo Nonbroadcast Multi-acceso (NBMA).
- Redes Tipo Punto Multi-punto.
- Enlaces Virtuales (Virtual Link).

TESIS CON  
 FALLA DE ORIGEN

#### 3.6.1 Redes Punto a Punto

Este tipo de redes se presenta cuando dos sistemas están directamente conectados, tanto para recibir como para transmitir información. Por ejemplo, las redes punto a punto se forman en enlaces seriales como son los E1/T1 o fracciones de estos, en donde solamente se encuentran conectados un par de routers, como se muestra en la figura 3-1. En estos enlaces los vecinos siempre forman adyacencias. La dirección destino utilizada por OSPF en los paquetes de este tipo de redes siempre debe ser la dirección de la clase D 224.0.0.5, conocida como "AllOSPF Routers".

En este tipo de redes no es necesario designar un DR y un BDR.

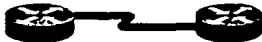


Figura 3-1 Red punto a punto.

#### 3.6.2 Redes Broadcast

Las redes Broadcast son redes multiacceso que pueden conectar a más de dos dispositivos, todos ellos son de tipo Broadcast, esto es, que pueden enviar mensajes que son recibidos y procesados por todos los dispositivos conectados.

Ejemplos de estas redes son Ethernet, Token Ring, FDDI, (figura 3-2) en estos tipos de redes es necesaria la elección de un DR y un BDR, los mensajes de Hello que son generados por el DR y BDR son enviados por medio de direcciones Multicast

"ALIOSPF Routers". La dirección MAC que se utiliza para llevar estos mensajes es 0100.5E00.0005. Todos los demás routers utilizan la dirección de la clase D "224.0.0.6" conocida como "AllDRouters" para realizar las actualizaciones o acuses de recibo (acknowledgement). La dirección MAC para estas tramas es 0100.5E00.0006.

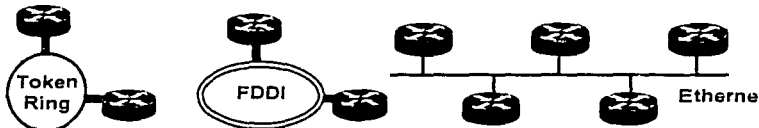


Figura 3-2 Redes Broadcast.

### 3.6.3 Redes NBMA

La figura 3-3 muestra este tipo de redes, las cuales tienen la capacidad de conectar a más de un router, pero no tienen la capacidad de soportar el envío de Broadcast, esto es, que un paquete enviado a un router puede que no sea recibido por los demás routers que se encuentren conectados. En estas redes también se realiza la selección de un DR y un BDR, y todos los paquetes son enviados por medio de direcciones Unicast. Ejemplos de estas redes son X.25, Frame-Relay y ATM.



Figura 3-3 Red NonBroadcast MultiAccess.

### 3.6.4 Redes Punto Multipunto

Estas redes son un caso especial de configuración de redes NBMA, en la cual, la red es tratada como un conjunto de enlaces punto a punto. La figura 3-4 hace referencia a este tipo de redes en las que no se realiza la elección de un DR y BDR, debido a que las redes son vistas como enlaces punto a punto, y en las que para el envío de paquetes se utiliza Multicast.



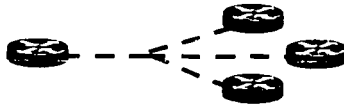


Figura 3-4 Red Punto Multipunto.

### 3.6.5 Virtual Links

Un virtual link es una conexión a una área remota que no tiene una conexión física al área 0. Sin embargo OSPF trata a esta área como si estuviera directamente conectada en un solo salto al área de backbone. También los virtual links se pueden utilizar para unir el área de backbone como se muestra en la figura 3-5.

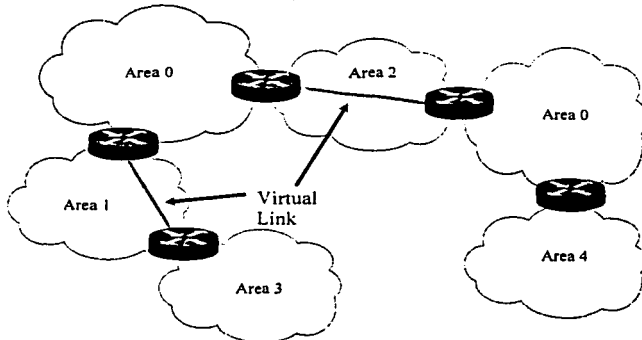


Figura 3-5 Virtual Links.

### 3.7 El Router Designado y Backup del Designado

Las redes multiacceso presentan dos problemas para OSPF con relación a la inundación de LSAs.

- La información de una adyacencia entre cada router conectado podría crear demasiados LSAs. Si  $n$  es el número de routers en una red multiacceso, existirían  $n(n-1)$  adyacencias. Cada router enviaría  $n-1$  LSAs, resultando en  $n^2$  LSAs originados por la red.

- La inundación en sí misma sería caótica. Un router enviaría un LSA a todos sus vecinos adyacentes y lo mismo harían los otros routers conectados en el mismo enlace, lo cual crearía muchas copias de un mismo mensaje de LSA en la misma red.

Para prevenir este problema se elige a un router como designado en las redes multiacceso. El DR tiene las siguientes funciones:

- Representar a la red multiacceso y estar conectado a ella.
- Administrar el proceso de inundación en la red multiacceso.

El concepto detrás del DR, es que las demás redes lo consideren como un pseudo-nodo o un router virtual, que representa a todos los routers conectados en la red multiacceso. Cada router en la red multiacceso forma una adyacencia con el DR (figura 3-6), el cual representa al pseudo nodo. Entonces solamente el DR enviará LSAs al resto de las redes. Hay que tener en cuenta que el DR podría ser un router conectado en una red multiacceso pero podría no ser el DR en otra red multiacceso a la cual también pudiera estar conectado. Esto es, que el DR es propietario a la interfaz del router y no al router entero.

Un problema que puede no estar muy lejos con el esquema del DR, se da cuando se presenta una falla en el router que fue elegido DR, entonces nuevas adyacencias deben ser establecidas y todos los routers en la red deben sincronizar nuevamente sus bases de datos con el nuevo DR. Mientras todo este mecanismo ocurre, la red no está disponible para el tránsito de paquetes.

Para prevenir este problema, se designa un router como respaldo (Backup) del Designado (Backup DR, BDR). Entonces todos los routers forman adyacencias tanto con el DR como con el BDR. El DR y BDR también forman adyacencias entre ellos como se ilustra en la Figura 3-6. Entonces si el DR falla, el BDR se convierte en el DR ya que este todavía tiene adyacencias con todos los routers de la red, minimizando así el tiempo de indisponibilidad de la red.

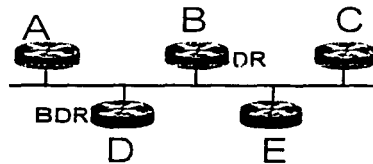


Figura 3-6 Router Designado y Backup del Router Designado.

### 3.8 Estados de las interfases de OSPF

Una interfaz de OSPF se habilita solo después de haber pasado a través de una serie de estados, antes de estar completamente operacional. Estos estados son: Down, Point-to-Point, Waiting, DR, Backup, DRother, y Loopback.

*Down:* Es el estado inicial de la interfaz. La interfaz no es operacional, todos los parámetros de las interfases son enviados a sus valores iniciales y ningún protocolo es enviado o transmitido en la interfaz.

*Point-to-Point:* Este estado es aplicable solo para interfases conectadas en redes punto a punto, multipunto y virtual link. Cuando una interfaz hace la transición a este estado, es completamente operacional y comenzará a enviar paquetes de Hello e intentará establecer adyacencias con sus vecinos al final del enlace.

*Waiting:* Este estado solo es aplicable a redes de tipo Broadcast o NBMA. Cuando una interfaz hace la transición a este estado, comenzará a enviar paquetes de Hello y definirá un cierto tiempo. El router deberá intentar identificar el DR y el BDR de la red mientras se encuentre en este estado.

*DR:* En este estado el router es el DR en la red a la que se encuentra conectado y establecerá adyacencias con los otros routers en la red multiacceso.

*Backup:* En este estado el router es el BDR de la red en la que se encuentra conectado y establecerá adyacencias con los otros routers conectados en la red multiacceso.

*DRother:* En este estado el router no es el DR, ni el BDR en la red a la que se encuentra conectado, y solo formará adyacencias con el DR y BDR aunque deberá seguir el comportamiento de todos los vecinos en la red.

*Loopback:* En este estado la interfaz se encuentra en "loopback" por software o hardware. Aunque los paquetes no pueden pasar a través de la interfaz, el anuncio se continúa realizando a través de los LSAs. La figura 3-7 muestra los estados de las interfases y los eventos que causan las transiciones. Estos eventos son descritos en la tabla 3-1.

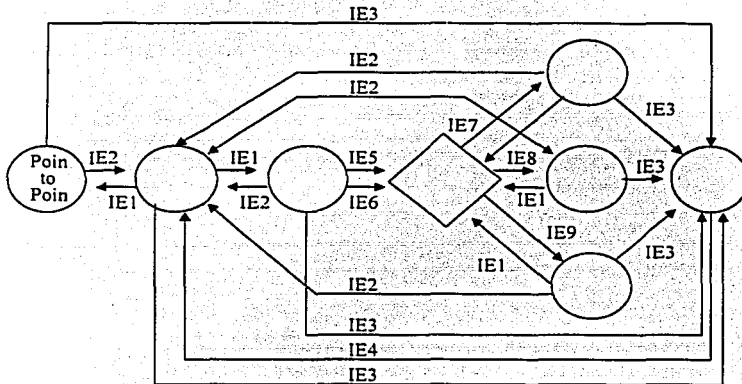


Figura 3-7 Máquina de Estado de las Interfases OSPF.

Evento	Descripción
IE1	Los protocolos de las capas inferiores indican que la interfaz de red es operacional.
IE2	Los protocolos de las capas inferiores indican que la interfaz de red no es operacional.
IE3	El administrador de red o las capas inferiores indican que la interfaz esta en loop-up*1.
IE4	El administrador de red o las capas inferiores indican que la interfaz esta en loop-down*1.
IE5	Un paquete de Hello fue recibido, originado por cualquiera de los vecinos, listandolo como el BDR o el vecino que lo origino se lista a si mismo como el DR y no indica un BDR.
IE6	El contador de wait ha expirado.
IE7	El router es elegido el DR.
IE8	El router es elegido el BDR.
IE9	El Router no ha sido elegido como DR o BDR para esa red.
IE10	<p>Un cambio ha ocurrido en un grupo de vecinos válidos en la red. Este cambio puede ser alguno de los siguientes:</p> <ul style="list-style-type: none"> <li>(1) El establecimiento de una comunicación bidireccional con un vecino.</li> <li>(2) Se ha perdido la comunicación bidireccional.</li> <li>(3) La recepción de un paquete Hello de un vecino, en el cual el vecino originante se lista como el DR o BDR</li> <li>(4) La recepción de un paquete Hello desde el DR en el cual él ya no es listado como el DR.</li> <li>(5) La recepción de un paquete desde el BDR en el cual él ya no es listado como el BDR.</li> <li>(6) La expiración del contador RouterDeadInterval sin haber recibido un Hello desde el DR o BDR o ambos.</li> </ul>

Tabla 3-1 Eventos que ocasionan los cambios de estados de las interfaces

### 3.9 Fases para el establecimiento de las adyacencias

A continuación se listan las fases para el establecimiento de las adyacencias:

1. Descubrimiento de vecinos.
2. Comunicación bidireccional. Esta comunicación es realizada cuando dos vecinos tienen los router IDs en sus paquetes de Hello.
3. Sincronización de las Bases de Datos. La descripción de la base de datos y los paquetes de actualización son intercambiados para asegurar que los vecinos tengan la

TESIS CON  
 FALLA DE ORIGEN

misma información en sus link-state database. El propósito de este proceso es que uno de los routers sea el maestro y el otro el esclavo. Como su nombre lo dice, el maestro controlará el intercambio de los paquetes descriptivos de la base de datos.

#### 4. Adyacencia completa (Full).

### 3.10 Estados de los vecinos

Los routers que se vuelven adyacentes con el DR y BDR tendrán exactamente la misma base de datos topológica. A continuación se mencionan los estados que deben de pasar los routers antes de obtener la adyacencia.

#### 1. DOWN

Ninguna información ha sido recibida por ningún router en ese segmento.

#### 1'. ATTEMPT

En redes multiacceso tipo non broadcast multiacces como Frame Relay y X.25, este estado indica que no se ha recibido información reciente de su vecino. Cuando pasa esta situación el router trata de contactar nuevamente a su vecino a través de paquetes Hello.

#### 2. INIT

La interfaz ha detectado un paquete Hello proveniente de un vecino, pero la comunicación bidireccional aún no se ha establecido.

#### 3. TWO-WAY

Existe una comunicación bidireccional con el vecino. El router se ha visto a él mismo en los paquetes Hello provenientes de su vecino. Al final de esta etapa se ha elegido al DR y BDR, los routers serán capaces de decidir con que routers harán adyacencias y con que routers no.

#### 4. EXSTART

Los routers tratan de establecer el *Initial Sequence Number* que va a ser utilizado para el intercambio de paquetes de información. Este número de secuencia asegura que los routers siempre obtengan la información más reciente.

### 5. EXCHANGE

Los routers describirán su base de datos topológica enviando Database Description Packets. En este punto, los paquetes deben de ser enviados a todas las interfaces del router por medio del proceso de inundación.

### 6. LOADING

En este paso los routers han finalizado el intercambio de información, han construido una lista de Link State Request y una lista de Link State Retransmission. Cualquier información incompleta o sin actualizar, se pondrá en la lista de peticiones "request". Cualquier actualización que es enviada se pondrá en la lista de retransmisiones "retransmission" hasta que llegue su acuse de recibo (acknowledge).

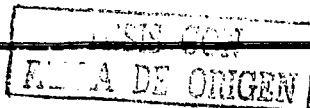
### 7. FULL

La adyacencia esta completa. Los routers vecinos están completamente adyacentes unos con otros.

OSPF siempre tendrá una adyacencia con un vecino que se encuentre conectado en una interfaz punto a punto. En estos casos no existen los conceptos de DR y BDR.

### 3.11 Áreas en OSPF

Debido a las múltiples bases de datos y a la complejidad de sus algoritmos, OSPF demanda una gran cantidad de memoria y una gran cantidad de tiempo de CPU del router. Cuando las redes crecen, estas demandas crecen de forma muy significativa que incluso pueden llegar a provocar que un router falle. Aún así, el mecanismo de inundación es mucho más eficiente que las actualizaciones periódicas de las tablas completas de RIP o IGRP, las cuales se pueden convertir en una carga inadmisibile a través de los enlaces de una red de gran tamaño. Contrariamente a la creencia popular de que el algoritmo de SPF es quien hace una intensiva demanda de tiempo de CPU, son los procesos relacionados a él quienes realizan la mayor demanda, tales como los procesos de inundación y de mantenimiento de la base de datos.



OSPF usa áreas para reducir estos efectos adversos. En el contexto de OSPF, un área es la agrupación lógica de routers de OSPF y sus enlaces, los cuales dividen un dominio de OSPF en subdominios. Los routers dentro de un área no tienen un conocimiento detallado de la topología fuera de su área. Por estas condiciones:

- Un router debe compartir una idéntica link state database con todos los routers dentro de su área y no con toda la red. El reducir el tamaño de la base de datos reduce el impacto de memoria en los routers.
- Una link state database más pequeña significa menos LSAs procesados y por lo tanto un menor impacto en el CPU.
- Ya que la link state database debe ser mantenida solo dentro del área, la mayoría de las inundaciones están limitadas solo al área.

Las áreas son identificadas por un identificador (*Area ID*) de 32 bits como lo muestra la figura 3-8, el área ID puede ser expresado de dos formas: como un número decimal o con números decimales separados por puntos (similar a una dirección IP).

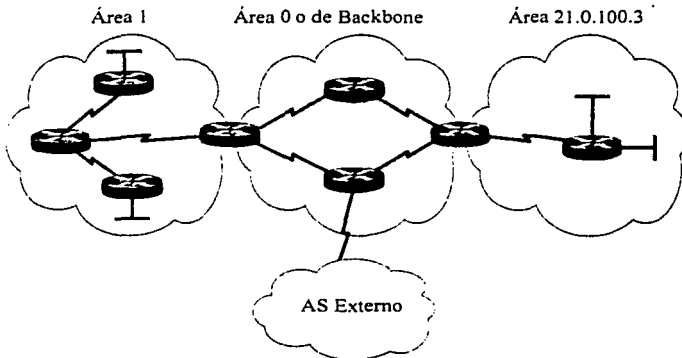


Figura 3-8 Notación para hacer referencia a un área.



### 3.11.1 Tipos de tráfico en relación a las áreas.

*Tráfico Intra-Área.* Es el tráfico de paquetes que son pasados por routers en la misma área

*Tráfico Inter-Área.* Es el tráfico de paquetes que pasan entre routers de diferentes áreas.

*Tráfico Externo.* Es el tráfico de paquetes que son pasados entre un router dentro del dominio de OSPF y un router dentro de otro sistema autónomo.

### 3.11.2 Enrutamiento Jerárquico

OSPF nos permite implementar un esquema modular y jerárquico de enrutamiento de información. El área ID 0 o 0.0.0.0 esta reservada para el backbone. El backbone es responsable de la sumarización topográfica de un área a cualquier otra área. Por esta razón, todo el tráfico inter-área debe pasar a través del backbone; las áreas no-backbone no pueden intercambiar tráfico directamente. Otro caso especial de área son las Áreas-Stub, antes de hablar de ellas, primero necesitamos definir los diferentes tipos de routers y los LSAs asociados a ellos.

### 3.11.3 Tipos de Routers en OSPF

Los routers al igual que el tráfico pueden ser clasificados de acuerdo a las áreas. Todos los routers de OSPF pueden ser uno de los cuatro tipos de routers, figura 3-9.

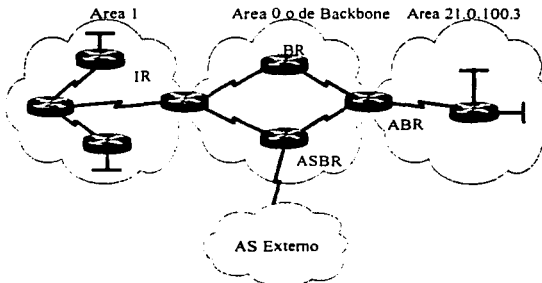


Figura 3-9 Tipos de Routers en OSPF.

TESIS CON  
FIRMA DE ORIGEN

*Router Interno (Internal Router) IR*

Todas las redes que se conectan a este tipo de router pertenecen a una misma área. Cada uno de estos routers mantiene una base de datos topológica única.

*Router de Borde de Área (Area Border Router). ABR.*

Este tipo de router se encuentra conectado a diferentes áreas. Estos routers mantienen una base de datos topológica por cada área a la que estén conectadas y otra para el área cero. Así mismo, este tipo de routers resume la información de la topología de las áreas que poseen para posteriormente redistribuirla al área cero. El área cero se encargará de distribuirla a las demás áreas.

*Router de Backbone (Backbone Router): BR.*

Es un router con una interfaz al backbone. Estos incluyen a los routers que poseen entre sus interfaces a más de un área, por ejemplo un ABR. Por el contrario, los routers de backbone no tienen que ser ABRs y los routers con todas sus interfaces conectadas al área cero son considerados IR.

*Router de frontera de Sistema Autónomo (AS Boundary Router): ASBR.*

Este router intercambia información de enrutamiento con otros routers de diferentes Sistemas Autónomos, otros protocolos de enrutamiento u otros procesos de OSPF para después anunciarlo en el proceso de OSPF propio. La ruta para llegar a cada ASBR debe ser conocida por todos los routers del Sistema Autónomo. Esta clasificación es completamente independiente de las anteriores por lo que pueden existir combinaciones.

### 3.12 Tipos de LSA

Debido a que en OSPF se definen múltiples tipos de routers, es necesario también definir múltiples tipos de LSAs. Por ejemplo un DR debe anunciar la red multi-acceso y todos los routers conectados en la red. Cada tipo de LSA describe un aspecto de OSPF. La tabla 3-2 lista los diferentes tipos de LSAs y el código que los identifica.

Código	Descripción
1	Router LSA
2	Network LSA
3	Network Summary LSA
4	ASBR LSA
5	AS external LSA
6	Group membership LSA
7	NSSA External LSA
8	External Attributes LSA
9	Opaque LSA (link-local scope)
10	Opaque LSA (area-local scope)
11	Opaque LSA (AS scope)

Tabla 3-2 Tipos de LSA.

*Router LSA:* son producidos por cada router. Es el LSA más fundamental, lista todos los enlaces que contiene el router, el estado del enlace y el costo de cada enlace. Estos LSAs son inundados dentro del área en la cual son originados

*Network LSA:* son producidos por el DR en cualquier tipo de red multi-acceso. El DR representa a toda la red multi-acceso y a todos los routers conectados a ella. El Network LSA lista a todos los router conectados incluyendo al DR, al igual que el Router LSA el Network LSA se difunde solamente dentro del área que lo origina.

*Network Summary LSA:* son originados por los ABRs y son enviados dentro de un área para anunciar destinos fuera del área. Esto es, los LSAs le dicen a los routers internos cuales son los destinos externos al área que pueden alcanzar a través de un cierto ABR. Además un ABR anuncia al área de Backbone a los destinos de las redes conectadas dentro de su área por medio de LSAs de tipo Network Summary LSA. Las rutas por default que son externas al área pero internas al sistema autónomo de OSPF también son anunciadas por medio de este tipo de LSA.

*ASBR Summary LSA:* también son originados por los ABRs. Los ASBR Summary LSA son idénticos a los Network Summary LSA, con la excepción de que el destino que anuncian es un ASBR.

*Autonomous System External LSA o External LSA:* son originados por los ASBRs y anuncian los destinos externos al sistema autónomo de OSPF, o una ruta por default externa al sistema autónomo de OSPF.

*Group Membership LSA:* son usados por una variante de OSPF conocida como Multicast OSPF (MOSPF). MOSPF realiza el enrutamiento de paquetes de una sola fuente a múltiples destinos, o grupos de miembros utilizando direcciones de clase D.

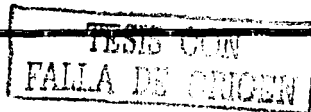
*NSSA external LSA:* son originados por los ASBRs dentro de áreas Not-so-stubby (NSSAs). Un NSSA external LSA en el formato del paquete es en su mayor parte idéntico a un AS external LSA, pero a diferencia de un AS External LSA los cuales son inundados en todo el sistema autónomo, los NSSA external LSA solamente son inundados dentro del área not-so-stubby que las generó.

*External Attributes LSA:* su propósito es una alternativa a usar Internal BGP (iBGP) para transportar la información de BGP a través de un dominio de OSPF.

*Opaque LSA:* son un tipo de LSA que consiste de un encabezado seguido por una aplicación específica. La información puede ser usada directamente por OSPF o indirectamente por otras aplicaciones para distribuir información a través del dominio de OSPF. Una aplicación de este tipo de LSA se encuentra en MPLS-TE.

### 3.13 Tipos de Áreas de OSPF

En OSPF existen diferentes tipos de áreas de acuerdo a los tipos de LSA que permite que se difundan a los routers que la conforman.



Un ASBR que ha aprendido destinos externos, los deben de anunciar por medio de inundaciones de LSAs de tipo AS External, a todo el sistema autónomo de OSPF. En muchos casos estos LSAs de tipo externo pueden ocupar un gran porcentaje de la base de datos de cualquier router.

### 3.13.1 Áreas Stub

Un área Stub es un área dentro de la cual los LSAs de tipo externo no pueden ser inundados. Y si un LSA de tipo 5 no es conocido dentro del área, los de tipo 4 son innecesarios; entonces estos tipos de LSA son también bloqueados. Los ASBR en las fronteras harán uso del LSA Network Summary para anunciar solo la ruta por default (destino 0.0.0.0) dentro del área. Cualquier destino que no se encuentre dentro de las rutas conocidas en el área, se encontrará dentro de la ruta por default. Como la ruta por default es transportada por LSAs de tipo 3, estos no se deberán anunciar fuera del área.

El desempeño de los routers dentro de áreas Stub puede ser mejorado, al reducir el tamaño de las bases de datos se reduce el uso de la memoria en un router. Por supuesto que la mejora es más marcada en redes con un gran número de LSAs de tipo 5; sin embargo existen cuatro restricciones para el uso de áreas Stub:

- Como cualquier otra área, todos los routers del área deben de tener la misma link state database. Para asegurar esta condición, todos los routers deberán utilizar la bandera (bit-E) con el valor de cero en los paquetes de Hello, con esto, los routers no aceptaran ningún paquete de Hello con el bit-E con un valor de uno, como resultado, las adyacencias no podrán ser establecidas con algún otro router que no este configurado como un Stub Router.
- Los Virtual Link no pueden ser configurados dentro de un área Stub o de tránsito.
- Un router dentro de un área Stub no puede ser un ASBR. Esta restricción se puede entender intuitivamente ya que los LSAs de tipo 5 que son producidos por un ASBR no pueden ser inyectados dentro de un área Stub.
- Un área Stub puede tener más de un área ABR. Pero debido a la existencia de la ruta por default, los routers no pueden seleccionar cual es el mejor gateway hacia el ASBR.

TESIS CON  
FALLA DE ORIGEN

### 3.13.2 Áreas Totally Stubby

En estas áreas la única ruta que es inyectada por el ABR es la ruta por default. Todas las rutas intra-área emplean la ruta por default para todos los destinos internos y externos al dominio de OSPF.

Las áreas Totally Stubby solamente usan la ruta por default para destinos tanto fuera del sistema autónomo de OSPF como para destinos externos al área. Para hacer esto el ABR de una área totally stub bloqueará tanto los LSA de tipo externo como los de sumarización de redes, con la excepción de los de tipo 3 que anuncian la ruta por default.

### 3.13.3 Áreas Not-So Stubby

Las Áreas de tipo NSSA permiten el anuncio de rutas dentro del sistema autónomo de OSPF, manteniendo las características de un área stub para el resto del sistema autónomo. Para hacer esto el ASBR dentro de una NSSA deberá anunciar las rutas externas por medio de mensajes de tipo 7. Estos LSA's NSSA External son inundados a través de toda la NSSA pero son bloqueados por los ABRs.

Los LSAs NSSA externos tienen una bandera conocida como el bit-P dentro de su encabezado. Los NSSA ASBRs tienen la posibilidad de modificar este bit. Si el NSSA ASBR recibe un LSA de tipo 7 con el bit-P en uno, este lo deberá de transformar en un LSA de tipo 5 e inundarlo a través de las otras. Si el bit-P es puesto en cero, entonces no se realiza la transformación y el LSA destino es un tipo 7 que no deberá ser anunciado fuera del NSSA.

La tabla 3-3 resume los LSAs permitidos en cada área.

Tipo de Área	1 y 2	3 y 4	5	7
Backbone (Área 0)	Si	Si	Si	No
Non-Backbone, Non-Stub	Si	Si	Si	No
Stub	Si	Si	No	No
Totally Stubby	Si	No	No	No
Not-so-Stubby	Si	Si	No	Si

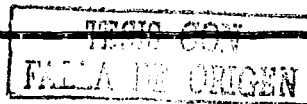
Tabla 3-3 LSAs permitidos en las áreas de OSPF.

### 3.14 Base de Datos Topológica

La base de datos topológica conocida también como base de datos Link State, describe la topología completa de una red, esto es: routers, segmentos de red y la forma en que estos se interconectan. Se representa a través de una tabla donde los vértices son routers y redes como se representa en la figura 3-10. Cuando existe una intersección en la tabla entre dos routers indica que estos se conectan a través de una interfaz física de red punto a punto (un enlace serial). Cuando existe una intersección de un router hacia una red, indica que el router tiene una interfaz asociada hacia a esa red. La intersección entre vértices puede tener diferentes tipos de valores de acuerdo a la tarea o función que la red o router tenga.

En el caso de que un router solamente transporte información, y que no tenga como destino u origen a algunas de sus redes (esto es, que sólo sea un router de tránsito), se representará en la tabla por medio de una intersección en ambos sentidos, es decir, tanto de entrada como de salida.

<sup>1</sup> Excepto para un tipo 3 de LSA del ABR, el anuncio de la ruta de default.



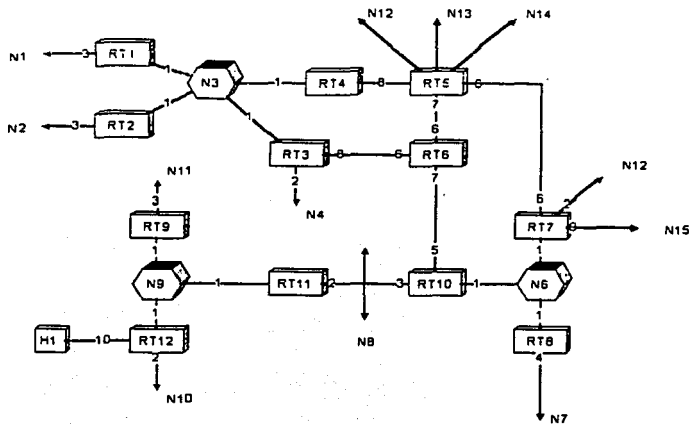


Figura 3-10 Base de datos topológica o base de datos Link State.

### 3.15 La tabla de enrutamiento

El algoritmo de Dijkstra es usado para calcular el Shortest Path Tree desde las link state database. El SPF se ejecuta una sola vez para construir las ramas de los árboles, las cuales son los enlaces de cada nodo (router) en el área. El algoritmo se ejecuta por segunda vez para agregar las hojas del árbol que serían las redes conectadas a cada router.

OSPF determina los caminos más cortos en base a una métrica arbitraria llamada costo, la cual es asignada a cada interfaz. El costo de una ruta es la suma de todos los costos de las interfaces de salida a un destino. El RFC-2328 no especifica valores para el costo a utilizar en las diferentes interfaces, una fórmula utilizada por los fabricantes es  $108/BW$ , donde BW es el ancho de banda de cada interfaz. La tabla 3-5 muestra los valores típicos utilizados.

TRIS CON  
FALLA DE ORIGEN



Tipo de Interfaz	Costo (10e8/BW)
FDDI, Fast Ethernet	1
HSSI (45M)	2
16M Token Ring	6
Ethernet	10
4M Token Ring	25
T1 (1.544M)	64
DS0(64kM)	1562

Tabla 3-5 Valores de costo default para las diferentes interfaces.

### 3.15.1 Tipos de caminos en OSPF

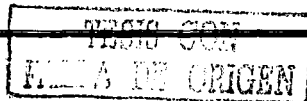
Los caminos se pueden clasificar en dos tipos de acuerdo en donde se encuentra el destino.

*Intra-area path:* son destinos que se encuentran en routers dentro de la misma área.

*Inter-area path:* son destinos que se encuentran en otra área pero dentro del mismo Sistema Autónomo de OSPF.

OSPF define dos tipos diferentes de caminos externos los cuales los denomina de tipo 1 y de tipo 2. a continuación se describe cada uno

*Type 1 external Path:* son destinos fuera del sistema autónomo de OSPF. Cuando una ruta es redistribuida dentro cualquier sistema autónomo, se le debe asignar una métrica que sea significativa al protocolo de enrutamiento del sistema autónomo. Dentro de OSPF, el ASBR es el responsable de la asignación de un costo a rutas externas. Las rutas externas que se definan como tipo 1 tendrán un costo que será el resultado de la suma del costo externo más el costo del camino a ASBR.



*Type 2 external Path:* son también destinos fuera del sistema autónomo de OSPF, pero no se les suma el costo para llegar al ASBR.

En la figura 3-11 se muestra un ejemplo con los dos tipos de destinos externos.

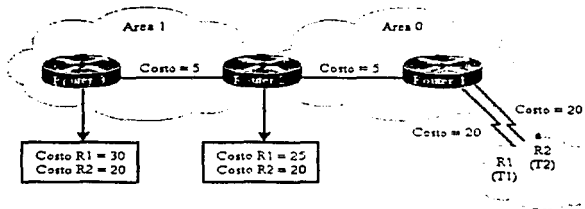


Figura 3.11 Destinos externos en OSPF.

# Capítulo 4

## Enrutamiento

### Interdominio y BGP

---

## Capítulo 4 Enrutamiento Interdominio y BGP

### 4.1. Enrutamiento Interdominio

En los inicios de los años 80s. los routers (gateways) que fueron utilizados en la ARPANET ejecutaban un protocolo vector distancia conocido como *-Gateway-to-Gateway Protocol (GGP)-*. Cada gateway conocía una ruta para alcanzar a cada una de las redes de la ARPANET. con una "distancia" medida en la cantidad de gateways por los cuales tenía que pasar.

Cuando la ARPANET creció, los ingenieros previeron muchos de los problemas que actualmente la mayoría de los administradores encuentran cuando las redes comienzan a crecer y cuando sus protocolos de enrutamiento no escalan adecuadamente.

#### 4.1.1 Historia de AS

Eric Rosen en el RFC 827 realizo una crónica de los problemas de escalabilidad y la solución que propuso fue que la ARPANET se fragmentara en pequeños grupos de redes conocidas como Sistemas Autónomos (*Autonomous System, AS*). Cada Sistema Autónomo sería libre de registrarse como él lo eligiera.

#### 4.1.2 Definición de AS

Sistema Autónomo (Autonomous System). Término utilizado para una colección de routers que están bajo la administración de una entidad y que interoperan utilizando un mismo protocolo de compuerta interior (IGP, Interior Gateway Protocol)<sup>1</sup>.

Nota. Entre los routers que pertenecen a diferentes sistemas autónomos debe de existir el acuerdo de utilizar un protocolo de compuerta exterior común a ambos, el cual permita la comunicación entre ambos.

---

<sup>1</sup> Definición tomada del glosario de términos de la American National Standard (<http://www.atis.org/tg2k/t1g2k.html>)

### 4.1.3 Evolución del enrutamiento Interdominio

Cada sistema Autónomo es libre de poder establecer el tipo de protocolo de enrutamiento que implementaran para su comunicación, conocido como Interior Gateways Protocol (IGP).

Por lo cual GGP es el primer protocolo (IGP) que se utilizo. Sin embargo el interés en un protocolo de enrutamiento más moderno y simple, motivó el desarrollo de Routing Information Protocol (RIP) en 1982, y otros como, RIPv2, OSPF, ISIS, etc.

Por otro lado, debido a que cada AS esta conectado con otro AS por medio de uno o más routers el RFC 827 proponía el uso de un protocolo de compuerta exterior para el intercambio de información entre ASs conocido como EGP (Exterior Gateway Protocol). Cuando se habla de este protocolo, se le hace referencia como un protocolo vector distancia, aunque realmente no es considerado como tal, ya que no existe un algoritmo con el cual se calculen mejores rutas. EGP es un lenguaje que los routers exteriores utilizan para calcular el alcance de las rutas.

### 4.2 Protocolo de Enrutamiento BGP

A EGP se le hicieron hechas muchas mejoras, finalmente un verdadero protocolo fue desarrollado para la comunicación interdominio. Su primera aparición fue en 1989 en el RFC 1105, esta primera versión fue actualizada un año después por medio del RFC 1163, la cual también fue actualizada al siguiente año con el RFC 1267, muchos se refieren a ellas como: BGP-1, BGP-2 y BGP-3 respectivamente.

#### 4.2.1 Operación de BGP descripción general de BGP-4

La versión actual de BGP es la número 4, fue introducida en 1995 en el RFC-1771, una de las principales diferencias con sus predecesoras es la característica de un comportamiento Classless que a diferencia de las anteriores tenían un comportamiento Classfull.

De una forma similar a EGP, BGP realiza una comunicación a través de una conexión basada en Unicast con cada uno de sus vecinos de BGP. dos routers que intercambian

información utilizando BGP también reciben el nombre de peers (parejas). Para incrementar la confiabilidad de la comunicación entre los peers, BGP utiliza TCP (puerto 179) como mecanismo para el intercambio de información. Este es el mecanismo con el que BGP realiza las actualizaciones, el cual es muy simple permitiendo a TCP realizar sus funciones de acuse de recibo (acknowledgments), retransmisiones y secuenciación. Como la comunicación de BGP se realiza sobre TCP, inicialmente se debe establecer una conexión punto a punto entre cada peer.

BGP es un protocolo de tipo vector distancia en el cual cada nodo confía en sus vecinos enviando sus rutas desde sus tablas de enrutamiento; los nodos hacen sus propios cálculos sobre las rutas que reciben y los resultados se envían a sus otros vecinos dentro y fuera de su AS. A diferencia de otros protocolos vector distancia como RIP donde el valor que utilizan para calcular la distancia de una ruta es el total de una suma de routers que tiene que cruzar para llegar a un destino, BGP utiliza una lista de números de AS por los cuales un paquete tendría que atravesar para alcanzar un destino. Debido a que esta lista describe completamente el camino que debe seguir un paquete, BGP es llamado protocolo de enrutamiento *path-vector*. La lista con los números de AS es llamada *AS\_PATH* y es uno de los muchos atributos de una trayectoria (Path attributes) asociado con una ruta, figura 4-1.

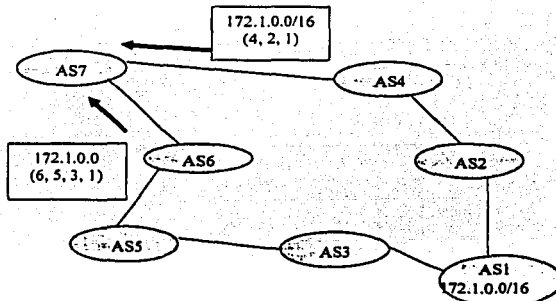


Figura 4-1 BGP determina el Loop-Free más corto Inter-AS Path desde una lista de números AS conocidos como el atributo *AS\_PATH*.

El atributo AS-PATH permite calificar a BGP como un verdadero protocolo de enrutamiento. debido a que con este atributo, BGP puede calcular los caminos más cortos y evitar loops de enrutamiento. Los caminos más cortos son fácilmente encontrados ya que solo basta encontrar el camino con el menor número de ASs. En la figura 4-1, el AS7 está recibiendo dos rutas hacia la red 172.1.0.0/16, uno de los caminos tiene que pasar a través de cuatro AS y el otro solo tres. El AS7 tiene entonces que seleccionar el más corto.

Los loops de enrutamiento son fácilmente detectados por medio del atributo AS\_PATH. Si un router recibe una actualización conteniendo en ella su AS dentro del AS\_PATH entiende que se ha encontrado un loop de enrutamiento y en tal caso ignora la actualización. En la figura 4-2 el AS7 ha anunciado una ruta al AS8, a su vez el AS8 anuncia esta ruta al AS9, el cual la anuncia al AS7, el AS7 encuentra su propio número de AS dentro del AS\_PATH y esto hace que la actualización no sea aceptada, con esto se evita un potencial loop de enrutamiento.

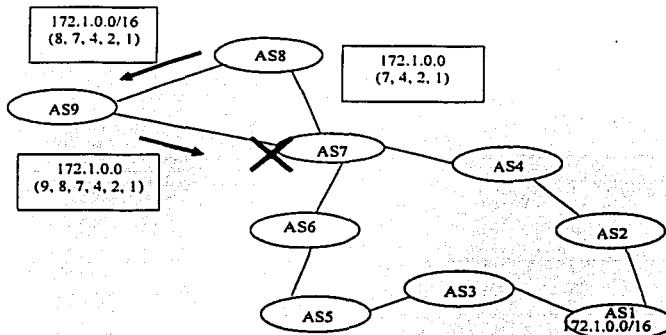


Figura 4-2 Si un router BGP ve su propio número de AS en el AS\_PATH de una ruta desde otro AS. Rechaza la actualización.

BGP no muestra una visión detallada de la topología dentro de cada AS porque BGP solo ve un árbol de Sistemas Autónomos, se puede decir que BGP tiene una visión más grande de Internet a diferencia del IGP que solo ve la topología dentro de su propio AS.

Los vecinos con los cuales habla un "speaker BGP"<sup>2</sup> puede ser cualquier router en su propio AS o en uno diferente. Si un vecino se encuentra en otro AS, entonces se dice que es un vecino de tipo externo (external BGP, EBGP). Si los vecinos se encuentran dentro del mismo AS, el vecino es de tipo interno (internal BGP, IBGP).

#### 4.2.2 Tipos de mensajes de BGP

Antes del establecimiento de una conexión de BGP, los dos vecinos deben de realizar el saludo de tres vías (three-way handshake) requerido por TCP y abrir una conexión de BGP a través del puerto 179. TCP proveerá las funciones de fragmentación, retransmisión, acuses de recibo, y de secuencia necesarias para una comunicación confiable, liberando a BGP de estas tareas. Todos los mensajes de BGP son Unicast y son enviados sobre la conexión de TCP.

BGP utiliza cuatro diferentes tipos de Mensajes:

- Open
- Keepalive
- Update
- Notification

##### 4.2.2.1 Mensaje Open

Después de establecer la sesión de TCP, ambos vecinos envían el mensaje de OPEN. Cada vecino usa este mensaje para identificarse y para enviar parámetros operacionales de BGP. El mensaje de OPEN incluye la siguiente información.

- El número de versión de BGP.- Especifica la versión (2,3 o 4) de BGP que se esta ejecutando en el router que origino el mensaje.

<sup>2</sup> Un Speaker BGP es un router que se comunica con otro a través del protocolo BGP.



- Autonomous System Number- Este parámetro indica el número de AS del router que origino el mensaje.
- Hold Time – Es el número de segundos que pueden pasar antes de que el router tenga que enviar una mensaje de Keepalive o de Update. El valor de Hold Time puede ser cualquiera incluso 0 o al menos 3 segundos, un valor común es de 180 segundos. Si el valor difiere entre los vecinos se selecciona el más pequeño de los dos.
- BGP identifier – Este valor es una dirección IP que identifica al router. Es común utilizar una interfaz lógica (interfases de loopbacks) y usar la dirección IP configurada en ella para evitar tener una dependencia del medio físico.
- Optional parameters – Este campo es usado para anunciar el soporte de alguna capacidad adicional como autenticación, multi-protocolo, y reenvío de rutas.

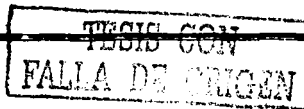
#### 4.2.2.2 Mensaje Keepalive

Si un Router acepta el parámetro especificado por el vecino en su mensaje de OPEN entonces responde con un keepalive. Los subsecuentes keepalives son enviados comúnmente cada 60 segundos o en un periodo que sea una tercera parte del Hold time acordado.

#### 4.2.2.3 Mensaje Update

El mensaje de Update anuncia las rutas factibles, el retiro de las rutas o ambas. El mensaje incluye la siguiente información:

- Network Layer Reachability Information (NLRI) – Esto es, uno o más prefijos que anunciar.
- Path Attributes – Contiene los atributos y las características NLRI a anunciar. Estos atributos proporcionan la información que le permitirá a BGP seleccionar la mejor ruta, detectar loops, determinar políticas de enrutamiento.
- Withdrawn Routes –Son rutas que han llegado a ser inalcanzables y han comenzado a ser retiradas del servicio.



Cabe mencionar que sin embargo, aunque múltiples prefijos pueden ser incluidos en el campo de NLRI, cada mensaje de Update solo describe a una sola ruta de BGP (ya que el path attribute describe solamente a un solo camino, pero el camino puede llegar a múltiples destinos). Esto nuevamente ayuda a hacer énfasis en que BGP tiene una visión más alta de una internetwork a diferencia de los IGP, quienes siempre llevan una sola ruta para un solo destino.

#### 4.2.2.4 Mensaje Notification

Este mensaje es enviado siempre que un error es detectado y siempre causa que BGP cierre la conexión de TCP. Un ejemplo del uso de este mensaje, se da cuando después de haber establecido la conexión de TCP, uno de los routers que solo soporta BGP-3 recibe un mensaje de OPEN desde su vecino especificando una versión 4, el router responde con un *mensaje de notification* indicando que esa versión no es soportada. La conexión se cierra y el vecino intentará reestablecer una conexión de BGP de la versión 3.

#### 4.2.3 Estados de BGP

Los estados de establecimiento y mantenimiento de una conexión de BGP pueden ser descritos por medio de una máquina de estados finitos. La figura la tabla 4-1 muestra la máquina de estados finitos de BGP y las entradas que pueden causar una transición a otro estado.

##### 4.2.3.1 Estado Idle

BGP siempre comienza en este estado, en el cual rechaza todas las peticiones de conexión. Cuando se presenta un "Start Event", el proceso de BGP inicializa todos los procesos: como el ConnectRetry timer, la conexión de TCP, espera una inicialización desde el vecino, y cambia al estado "Connect". El Start event es causado por: un operador al configurar un proceso de BGP, al reiniciar un proceso existente o por el router después de haberse reiniciado.

TRIS CON  
FALLA DE ORIGEN

#### 4.2.3.2 Estado Connect

En este estado, el proceso de BGP esta esperando que la conexión de TCP se establezca. Si el establecimiento es exitoso, el proceso de BGP limpia el ConnectRetry timer, completa la inicialización, envía un mensaje de OPEN al vecino, y realiza la transición al estado "OpenSent". Si el establecimiento de la conexión no fue exitoso, el proceso de BGP continua esperando otra conexión de inicialización de parte del vecino, reinicia el contador de ConnectRetry y hace una transición al estado Active.

Si el contador ConnectRetry expira en el estado Connect, el contador es reiniciado, se realiza otro intento para establecer una conexión TCP y el proceso se queda en Connect. Cualquier otra entrada causa una transición a Idle.

#### 4.2.3.3 Estado Active

En este estado, el proceso de BGP esta tratando de establecer una conexión de TCP con el vecino. Si la conexión de TCP se logra establecer, el proceso de BGP reinicializa el contador ConnectRetry, completa la inicialización, envía un mensaje de OPEN al vecino, y hace la transición al estado de OpenSent. Al Hold timer se le asigna un valor de 4 minutos. Si el contador de ConnectRetry expira mientras BGP se encuentra en el estado Active, el proceso regresa al estado de Connect y reinicia el contador ConnectRetry. Esto además inicializa una conexión de TCP con el vecino y continúa escuchando. Si el vecino esta tratando de establecer una sesión de TCP con una dirección IP no esperada, el contador ConnectRetry es reiniciado, la conexión es rechazada y el proceso se mantiene en el estado de Active. Cualquier otra entrada (con excepción de un Start event, la cual es ignorada en este estado) causa una transición al estado de Idle.

#### 4.2.3.4 Estado OpenSent

En este estado, un mensaje de Open ha sido enviado al vecino y se espera recibir un mensaje Open desde el vecino. Cuando un mensaje de Open es recibido, son revisados todos los campos. Si existe un error, un mensaje de Notification es enviado al vecino.

Si no hay error en el mensaje de Open recibido, se envía un mensaje de Keepalive y el contador de Keepalive es inicializado. El Hold timer es negociado, y se selecciona el más pequeño. Si el Hold timer es cero, los contadores de hold timer y keepalive no son iniciados. El tipo de conexión EBGP o IBGP, esta basada en el número de AS y el estado es cambiado a OpenConfirm.

Si una orden de desconexión es recibida por TCP, el proceso de BGP se cierra, reinicia el ConnectRetry, y comienza a escuchar esperando una nueva solicitud de conexión desde el vecino y una transición al estado de Active. Cualquier otro evento (excepto el de Start Event, el cual es ignorado) causa una transición a Idle.

#### 4.2.3.5 Estado OpenConfirm

En este estado, BGP esta esperando un mensaje de Keepalive o de Notification. Si se recibe un mensaje Keepalive, mueve su estado a Established. Si se recibe un mensaje de Notification o una orden de desconexión de TCP se realiza una transición al estado de Idle. Si el contador de Hold timer expira, un error es encontrado o un Stop event ocurre, entonces se envía un mensaje de notification al vecino y la conexión de BGP es cerrada y se cambia el estado a Idle.

#### 4.2.3.6 Estado Established

En este estado, el proceso de BGP esta completamente establecido y los vecinos pueden intercambiar mensajes de Update, Keepalive y Notification. Si se recibe un mensaje de Update o Keepalive, se reinicia el Hold timer (si el Hold timer no es cero). Si se recibe un mensaje de Notificación, se hace la transición al estado de Idle. Cualquier otro evento (con excepción de Star event, el cual es ignorado) causa que un mensaje de Notification sea enviado al vecino y ocurre una transición al estado de Idle.

#### 4.2.4 Path Attributes

Los Path Attributes son una característica de los anuncio de BGP, algunos paths attributes son de uso familiar como la dirección IP destino y el next-hop, estas características se pueden encontrar en otros protocolos de enrutamiento, otros path attributes, como el

Atomic-Aggregate son exclusivos de BGP. Además de la información básica de enrutamiento, los Path attributes le permiten a BGP definir y comunicar políticas de enrutamiento.

Categoría de los Path Attributes:

- Well-Know mandatory
- Well-know discretionary
- Optional transitive
- Optional nontransitive

#### 4.2.4.1 Categorías de los Path Attributes

Las categorías pueden ser vistas como dos subclases: Well-Know y Optional, y dentro de cada subclase existen otras dos subclases. Well-know, significa que deberá ser reconocido por cualquier implementación de BGP. Optional, significa que la implementación de BGP no requiere soportarla para funcionar.

Respecto a los atributos Well-know que son mandatory, significa que deben ser incluidos dentro de los mensajes de Update y los discretionary, significa que pueden ser o no incluidos en los mensajes de Update.

Si un atributo Optional es transitive, el proceso de BGP deberá aceptar el path en el cual esta contenido, incluso si el atributo no es soportado, este deberá ser enviado en el path a los otros peers.

Si un atributo Optional, es nontransitive, el proceso de BGP podrá ignorarlo en la actualización (Update) en la cual este incluido y no anunciarlo a sus otros peers.

En la tabla 4-1 se muestra una lista de los Path attributes más comunes:

Atributo	Clase
ORIGIN	Well-known mandatory
AS_PATH	Well-known mandatory
NEXT_HOP	Well-known mandatory
LOCAL_PREF	Well-known discretionary
ATOMIC_AGGREGATE	Well known discretionary
AGGREGATOR	Optional transitive
COMMUNITY	Optional transitive
MULTI_EXIT_DISC	Optional nontransitive
ORIGINATOR	Optional nontransitive
CLUSTER_LIST	Optional nontransitive

Tabla 4-1 Algunos atributos especificados en el RFC 1771.

#### 4.2.4.2 El atributo ORIGIN

El origen es de tipo Well-know mandatory. este atributo especifica el origen de la actualización de enrutamiento NLRI (Network Layer Reachability Infomation). El origen puede ser uno de los siguientes:

- IGP – Indica que el NLRI fue aprendido desde un protocolo interno al AS.
- EGP – Indica que el NLRI fue aprendido desde el protocolo EGP.
- Incomplete – El NLRI fue aprendido por alguna forma diferente a las anteriores. Incompleto no significa que la ruta sea invalida, solamente la forma en que fue generada. A las rutas que son redistribuidas a BGP se les asigna este atributo.

#### 4.2.4.3 El atributo AS\_PATH

El atributo AS\_PATH es de tipo Well-know mandatory que usa una secuencia de números de ASs, al cual describe el camino a través de los diferentes ASs (Inter-AS Path) hacia el destino. Cuando un BGP-speaker origina una ruta, anuncia el NLRI a sus peers agregando su AS, los cuales anuncian la ruta a sus peer externos. ellos a su vez agregan su AS antes del final (PreEnd). El resultado es que el AS\_PATH describe todos los sistemas autónomos

por los cuales paso el anuncio. empezando con el AS más reciente y terminando con el AS origen. ver figura 4-3.

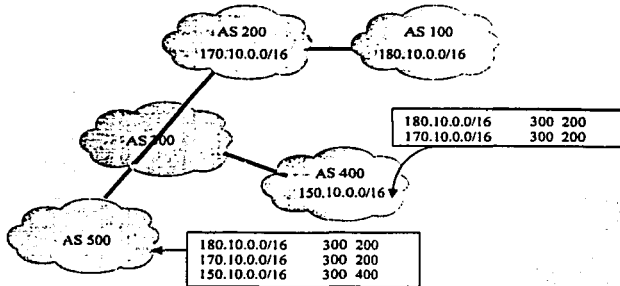


Figura 4-3 Los números de AS son agregados a lo largo de la trayectoria de una ruta.

#### 4.2.4.4 El atributo NEXT\_HOP

Como su nombre lo dice, este atributo de tipo Well-know mandatory describe la dirección IP del router NETX\_HOP en el camino hacia el destino. La dirección IP descrita por BGP no siempre es la dirección IP del router vecino.

Este atributo sigue las siguientes reglas:

- Si el router que realiza el anuncio y el router que lo recibe se encuentran en diferentes sistemas autónomos (peer externos), el NETX\_HOP es la dirección de la interfaz del router que los anuncia, figura 4-4.
- Si el router que lo anuncia y el router que recibe el anuncio se encuentra en el mismo sistema autónomo (peers internos), y el NRLI de la actualización (Update) se encuentra dentro del mismo AS, la dirección IP del NETX\_HOP es la dirección IP del vecino que anuncia la ruta, figura 4-5.
- Si el router que lo anuncia y el router que lo recibe son peer internos y el NRLI de la actualización hace referencia a un destino dentro del sistema autónomo. El NETX\_HOP es la dirección IP del peer externo desde el cual fue aprendida la ruta.

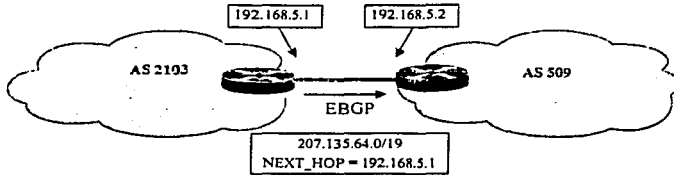


Figura 4-4 Si una actualización de BGP es anunciada vía EBGP, el atributo NEXT\_HOP es la dirección IP del peer externo.

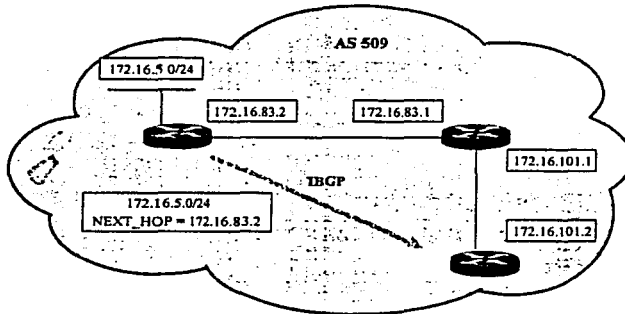


Figura 4-5 Si una actualización de BGP es anunciada vía IBGP, y el router destino esta en el mismo SA, el atributo NEXT\_HOP es la dirección IP del router origen.

#### 4.2.4.5 El atributo LOCAL\_PREF

LOCAL\_PREF es la abreviación de preferencia local. Este atributo es de tipo Well-know discretionary, el cual solo es anunciado a los peers internos, y no es enviado a los peers externos. Es usado para comunicar la preferencia local de una ruta en el AS a los routers que hablan iBGP. Si un BGP-speaker interno recibe múltiples rutas para el mismo destino compara el atributo de LOCAL\_PREF para seleccionar la mejor ruta. Este atributo es utilizado para modificar la forma en que el tráfico sale del AS hacia cierto destino, figura 4-6.



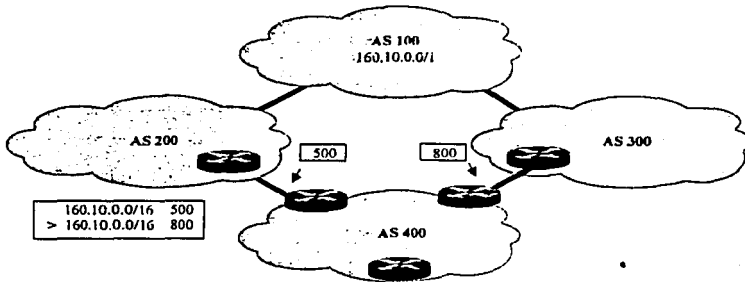


Figura 4-6 El atributo de LOCAL\_PREF comunica un grado de preferencia a los peers internos, prefiriendo el valor más alto.

#### 4.2.4.6 El atributo MULTI\_EXIT\_DISC

El atributo MULTI\_EXIT\_DISC es usado para influenciar el tráfico de regreso al AS, el MULTI\_EXIT\_DISC también abreviado como MED, es un atributo de tipo Optional nontransitive que es transportado en las actualizaciones de EBGP para permitir informarle a otro AS sobre su punto de regreso preferido, figura 4-7.

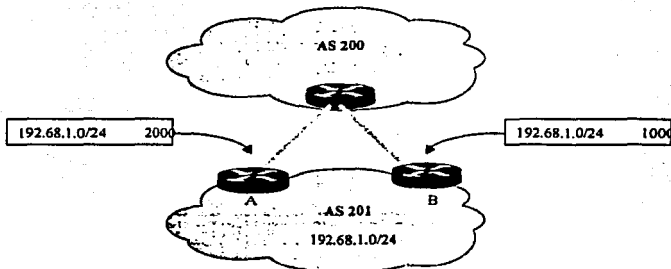


Figura 4-7 El atributo MULTI\_EXIT\_DISC hace que el tráfico de regreso al AS prefiera el enlace con el mayor MED, en este caso se prefiere el enlace hacia el router A.

#### 4.2.4.7 Atributos ATOMIC\_AGGREGATE y AGGREGATOR

Un router que esta hablando BGP puede propagar rutas superpuestas a otros speakers de BGP. Las rutas que se superponen no son rutas idénticas desde el punto de vista del destino. Cuando se realiza la selección de la mejor ruta, un router siempre seleccionará la ruta más específica, sin embargo, el BGP-speaker tiene varias opciones para seleccionar con las rutas que se superponen, estas opciones son:

- Anunciar ambas rutas, la más específica y la menos específica.
- Anunciar solo la ruta más específica.
- Anunciar solo la parte de la ruta que no se superpone.
- Anunciar las dos rutas y anunciar la agregada.
- Anunciar la ruta menos específica.
- No anunciar ninguna.

Cuando se realiza una sumarización, parte de la información de una ruta se pierde y el enrutamiento es realizado de una forma menos precisa. La agregación es realizada dentro de un router que habla BGP (BGP speaker). La información que se pierde son los detalles para llegar a un destino. La figura 4-8 ilustra la perdida de los detalles para llegar a un destino.

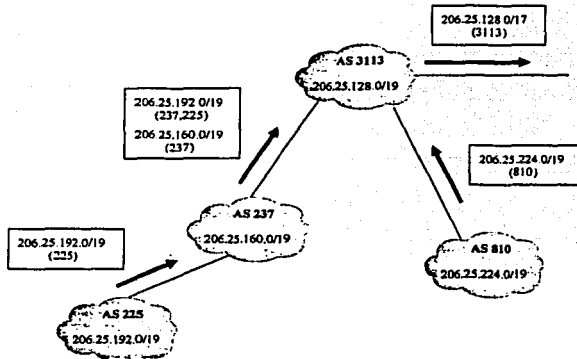


Figura 4-8 La agregación de rutas de BGP da como resultado la perdida de información de la trayectoria.

TESIS CON  
FALLA DE ORIGEN

El AS 3113 esta anunciando un agregado que representa una serie de sistemas autónomos, como él es quien origina el agregado, el anuncio solo incluye su sistema autónomo en el AS\_PATH. La información de los prefijos de los otros sistemas autónomos que son representados por el agregado se pierde.

ATOMIC\_AGGREGATE es un atributo Well-known discretionary usado para alertar a los routers que en el camino ha ocurrido una perdida de información. En cualquier momento un speaker de BGP puede sumarizar un grupo de rutas en un agregado menos específico y perder información como consecuencia de este agregado, el speaker de BGP debe estar ligado con el atributo de ATOMIC\_AGGREGATE al agregado de rutas. Cualquier router que recibe una ruta con un ATOMIC\_AGGREGATE no puede modificar ninguna parte de la información del NLRI de una ruta más específica, cuando realiza el anuncio a otros vecinos, se debe de mantener el valor del atributo de ATOMIC\_AGGREGATE. Cuando se coloca un atributo de ATOMIC\_AGGREGATE, el speaker de BGP también tiene la opción de colocar el atributo de AGGREGATOR. Este atributo opcional provee información acerca de lugar en de donde fue realizada la agregación al incluir el sistema autónomo y la dirección IP del router que origino la ruta de agregación, figura 4-9.

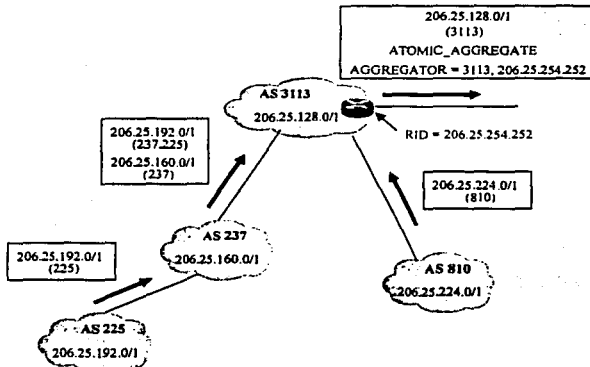


Figura 4-9 El atributo ATOMIC\_AGGREGATE indica que ha la pérdida de información de la trayectoria ha ocurrido, el atributo AGGREGATOR indica en donde ocurrió la agregación.

#### 4.2.4.8 El atributo de COMMUNITY

COMMUNITY es un atributo de tipo optional transitive que fue diseñado para simplificar las políticas de enrutamiento. originalmente era una implementación propietaria (de los routers CISCO) actualmente es un estándar definido en el RFC 1997. Las comunidades son un atributo que identifica a un destino como un miembro de alguna comunidad de destinos, las cuales comparten una o más propiedades. Por ejemplo, un proveedor de servicios puede asignar una comunidad para todas las rutas de sus clientes. El proveedor de servicios puede asignar ciertos valores de MED y LOCAL\_PREF en base a esta comunidad, en lugar de realizarlo en base a cada ruta.

La comunidad es un grupo de cuatro octetos, El RFC 1997 especifica que los primeros dos octetos sean el número de AS y los otros dos octetos sean administrativamente definidos, dándoles el siguiente formato AS:NN. Por ejemplo en la figura 4-10 las rutas del AS 300 tienen un identificador de comunidad con un valor de 1. El atributo de comunidad en el formato de AS:NN es 300:1 y en forma hexadecimal es representado como la concatenación de los dos números: 0x012C0001, donde 300=0x012C y 1=0x0001.

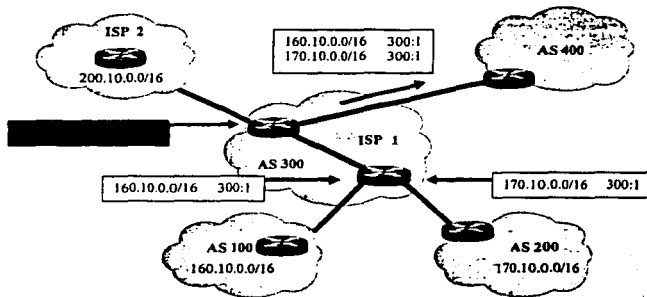


Figura 4-10 El atributo COMMUNITY identifica a un destino como miembro de una comunidad de destinos que comparten una o más propiedades comunes.

TESIS CON  
FALLA DE ORIGEN

Los valores de las comunidades van desde 0 (0x00000000) hasta 65535 (0x0000FFFF) y los valores desde 4294901760 (0xFFFF0000) al 4294967295 (0xFFFFFFFF) están reservados. Fuera de este rango se han definido una serie de comunidades de tipo "well-known" que realizan ciertas acciones sobre las rutas a las que se les asignan estas comunidades:

- INTERNET – La comunidad de Internet no tiene valor; todas las redes tienen este valor por default. Las rutas con esta comunidad son anunciadas sin restricciones.
- NO\_EXPORT (4294967041, o 0xFFFFFFFF01) – Las rutas recibidas con este valor no pueden ser anunciadas a los vecinos de EBGP, si se tratara de una confederación estas rutas no pueden ser anunciadas fuera de la confederación.
- NO\_ADVERTISE (4294967042, o 0xFFFFFFFF02) – Las rutas recibidas con este valor no serán anunciadas tanto a vecinos EBGP como a IBGP.
- LOCAL\_AS (4294967043, o 0xFFFFFFFF03) – El RFC llama a este atributo NO\_EXPORT\_SUBCONFED. Las rutas recibidas con este valor no pueden ser anunciadas a los vecinos de EBGP, incluyendo a los vecinos de EBGP dentro de una confederación.

#### 4.2.4.9 Atributos ORIGINATOR\_ID y CLUSTER\_LIST

Los atributos ORIGINATOR\_ID y CLUSTER\_LIST son al mismo tiempo de tipo optional y nontransitive, usados por los route reflectors, ambos atributos son empleados para evitar loops de enrutamiento. El ORIGINATOR\_ID es un número de 32 bits creado por el router que es configurado como route reflector. El valor es el Router\_ID del router que origino la ruta en el AS local. Si el router que origino el mensaje ve su propio Router\_ID en el ORIGINATOR\_ID de una ruta recibida, se da cuenta que existe un loop de enrutamiento y la ruta es ignorada.

El CLUSTER\_LIST es una secuencia de los clusters IDs por los cuales ha pasado la ruta. Si un route reflector ve su propio cluster ID dentro del CLUSTER\_LIST al recibir una ruta, sabe que ha ocurrido un loop de enrutamiento y la ruta es ignorada.

TESIS CON  
FALLA DE ORIGEN

#### 4.2.4.10 Administrative weight

El parámetro administrative weight (peso administrativo) es una implementación propietaria de varios fabricantes, la cual es local al router. Este valor nunca es comunicado a otros routers. El weight es un valor entre 0 y 65535 que puede ser asignado a una ruta; Las rutas con un mayor valor de weight son preferidas. Cuando se realiza la selección de una ruta en el proceso de decisión, se considera primero el weight sobre todos los demás atributos. Por default todas las rutas aprendidas tienen un valor de 0, y todas las rutas generadas por el router tienen un valor de 32768.

La figura 4-11 muestra el uso del parámetro weight para preferir una ruta sobre otras.

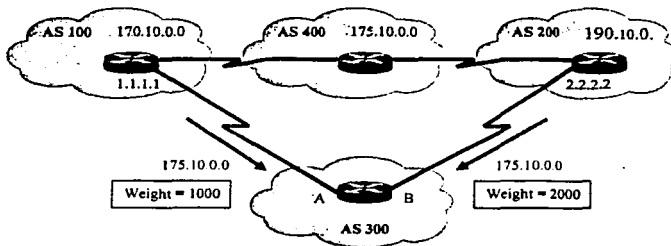


Figura 4-11 El AS 300 puede alcanzar a la red 175.10.0.0 por ambos enlaces, el parámetro Weight con un valor de 2000 hace que se prefiera el enlace B sobre el A para alcanzar esta red.

#### 4.2.4.11 Tipos de AS\_PATH

El atributo de AS\_PATH no es solo una lista ordenada de números de ASs que describe el camino hacia un destino. De hecho existen dos tipos de AS\_PATH:

- AS-SEQUENCE – Este es una lista ordenada de números de ASs a lo largo de un camino para llegar a un destino.
- AS\_SET – Es una lista desordenada de números de ASs a lo largo de un camino para llegar a un destino.

Uno de los principales objetivos del AS\_PATH es la prevención de loops de enrutamiento. Si un router que habla BGP ve su propio número de AS dentro de una actualización de un vecino externo, sabe que se ha presentado un loop de enrutamiento y por lo tanto ignora la actualización. Cuando se realiza una agregación, se pierde el detalle de la información y como resultado se incrementa la posibilidad de crear loops.

Por ejemplo en la figura 4-8, el AS 810 tiene una conexión alternativa a otro AS (ver la figura 4-12). Un agregado realizado por el AS 3113 es anunciado al AS 6571, y de éste otra vez al 810.

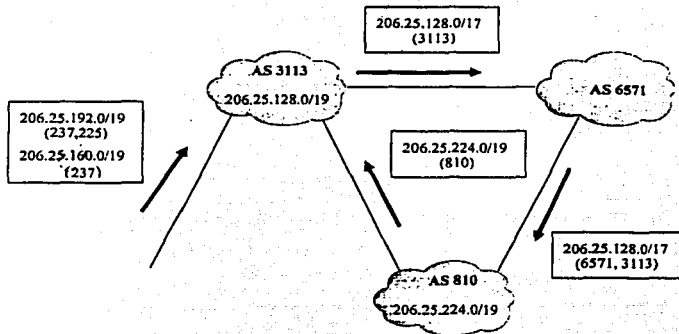


Figura 4-12 La pérdida del detalle de la trayectoria cuando se realiza la agregación, puede provocar loops de enrutamiento Inter-AS.

Como los números de AS atrás del punto de agregación no son incluidos en el AS\_PATH, el AS 810 no puede detectar un potencial loop de enrutamiento. Suponiendo que una red dentro del AS 810, por ejemplo la red 206.25.225.0/24, fallará. Los routers dentro del AS encontrarían en la agregación proveniente del AS 6571 una entrada válida para avanzar paquetes hacia ese camino ocurriendo entonces un loop de enrutamiento entre estos dos ASs.

Si se analiza un poco, para la prevención de un loop de enrutamiento no es necesario que los números de ASs sean incluidos con un orden en particular. Todo lo que se requiere es que el router que recibe la actualización pueda reconocer si su propio número de AS esta en la actualización. En esta parte es donde el AS\_SET comienza a ser útil.

Cuando un speaker de BGP crea un agregado a partir de un NLRI aprendido desde otro sistema autónomo, puede incluir todos los números de AS como un AS\_SET. Por ejemplo la figura 4-13 muestra la red de la figura 4-9 con un AS\_SET adicionado al agregado.

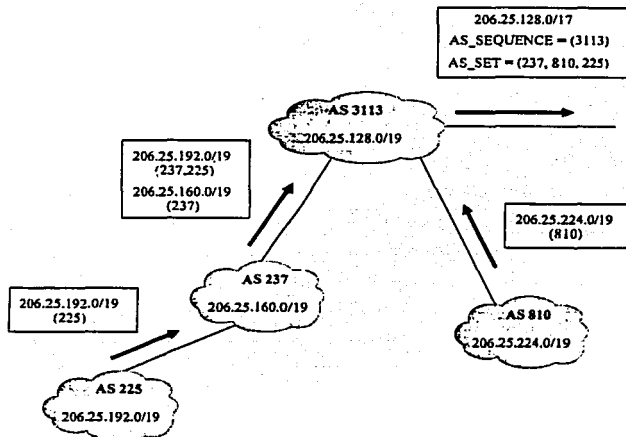


Figura 4-13 Incluyendo un AS\_SET en el AS\_PATH de una ruta agregada se restaura el loop y se evita la pérdida de información en la agregación.

El router que origina el agregado inicia un AS\_SEQUENCE. de esta manera los routers que los reciben pueden rastrear quien realizo el agregado. Aún así un AS\_SET es incluido para prevenir loops de enrutamiento. En la figura 4-13 también se puede ver porque el AS\_SET es una lista desordenada de números de AS. Detrás del AS3113 se encuentran las ramas de los caminos hacia los sistemas autónomos en los cuales residen las redes del agregado.



Cuando un AS\_SET es incluido en un AS\_PATH, el ATOMIC\_AGGREGATE no debe ser incluido con el agregado. El AS\_SET sirve entonces para notificar a los routers que una agregación ha ocurrido e incluye más información en comparación con el atributo ATOMIC\_AGGREGATE.

En la mayoría de las veces en la vida real, el AS\_SET sirve de compensación. También se entenderá que la sumarización sirve para proporcionar estabilidad, ya que si una red falla, la falla no es propagada más allá del punto de agregación, la sumarización también reduce en gran medida el tamaño de las tablas de enrutamiento de Internet.

#### 4.2.5 Proceso de decisión de BGP

La base de datos de enrutamiento (RIB) de BGP consiste de tres partes:

- *Adj-RIBs-In* - Es la parte en donde se encuentra la información de enrutamiento que ha sido aprendida de actualizaciones recibidas desde los vecinos. Todas las rutas en la Adj-RIBs-In se consideran como válidas.
- *Loc-RIB* - Contiene las rutas que el router ha seleccionado como válidas después de aplicar sus políticas de enrutamiento a la información recibida de la Adj-RIBs-In.
- *Adj-RIBs-Out* - Contiene las rutas que se anunciarán a los vecinos de BGP.

Múltiples atributos en las rutas de BGP pueden ser modificados o asignados para forzar que el tráfico siga ciertas políticas de enrutamiento dentro de un router, entre vecinos de IBGP o incluso en los Sistemas Autónomos adyacentes y más allá de estos. Por esto es importante conocer el proceso de selección de BGP cuando se presentan múltiples caminos hacia un destino. A continuación se describen los criterios de desempate:

1. Se prefiere la ruta con el mayor weight.
2. Si los weights son iguales, se prefiere la ruta con el mayor valor de LOCAL\_PREF.
3. Si los LOCAL\_PREF son iguales, se prefiere la ruta que fue originada localmente en el router, esto es, la ruta fue aprendida desde un IGP en el mismo router.

4. Si el LOCAL\_PREF es igual y la ruta no fue localmente originada, se prefiere la ruta con el AS\_PATH más corto.
5. Si la longitud del AS\_PATH es el mismo, se prefiere el camino con el origen más pequeño. IGP es menor que EGP, el cual es menor que Incompleto.
6. Si el origen es el mismo, se prefiere la ruta con el menor valor de MULTI\_EXIT\_DISC. Esta comparación es hecha solo si todas las rutas tienen los mismos números de AS.
7. Si el MED es el mismo, son preferidas las rutas de EBGp sobre las rutas de IBGP de la confederación y estas son preferidas sobre las de IBGP.
8. Si las rutas todavía son iguales, se prefiere la ruta con el camino más corto al NEXT-HOP de BGP. Esto es, se prefieren las rutas con una métrica más corta al next-hop de acuerdo con el IGP.
9. Si las rutas siguen siendo iguales, quiere decir que provienen del mismo sistema autónomo vecino, por lo cual si se tiene activada una opción de balanceo de tráfico ésta será realizada.
10. Si no se tiene activada alguna opción de balanceo es seleccionada la ruta con el menor BPG Router\_ID.

#### 4.2.6 Penalizaciones a las Rutas (Route Dampening)

Las oscilaciones de rutas (route flap) provocan la mayor parte de las inestabilidades en Internet y no importa de quien provengan. Una oscilación ocurre cuando una ruta es declarada como inválida e inmediatamente después se declara nuevamente válida. Cada vez que ocurre un cambio de estado, el cambio debe ser anunciado a través de todas las redes, y en cada router se deben realizar todos los cálculos necesarios, esto hace que se consuma tanto ancho de banda como tiempo de procesamiento.

Estos cambios son debidos principalmente a fallas en los enlaces o por cambios realizados por los administradores. Son cambios indeseados debido a que de existir un gran número de rutas cambiando de estado, pueden llegar a saturar a los routers al no poder procesar tal cantidad de rutas.

Como comentamos anteriormente, crear agregados ayuda a la estabilidad de Internet pero no contribuye a la estabilidad dentro de su propio AS.

La penalización de rutas es un método creado para detener las rutas inestables, esto no previene que un router acepte las rutas inestables, pero si previene de avanzarlas a los otros routers, en el RFC 2439 se encuentra especificada la forma de realizar las penalizaciones.

Un router usando penalizaciones asigna a cada ruta una figura dinámica de merito que refleja el grado de estabilidad de la ruta. Cuando una ruta oscila acumula penalidades. Esto también es llamado media-vida (Half-life). La penalidad es reducida a la mitad al final de cada half-life. Si la penalidad excede un umbral conocido como el limite de supresión (limit suppressed), la ruta es suprimida, esto es, la ruta se deja de anunciar. La ruta es suprimida hasta que el half-life se reduce a un umbral menor conocido como limite de rehusó (reuse limit). En este tiempo la ruta es anunciada nuevamente. Las penalidades también pueden ser manualmente eliminadas por los administradores.

Las penalidades podrían continuar acumulándose mientras se encuentran por arriba del limite de supresión, de tal modo que una ruta permanecería penalizada por periodos muy largos de tiempo. Sin embargo las rutas no pueden ser penalizadas más allá de un máximo limite de supresión (Maximum suppress limit). La figura 4-14 muestra gráficamente los valores típicos de las penalizaciones en los routers:

- Penalty – 1000 por flap.
- Suppress Limit – 2000.
- Reuse limit – 750.
- Half-life – 15 minutes.
- Maximum suppress time – 60 minutos, o 4 veces Half-life.

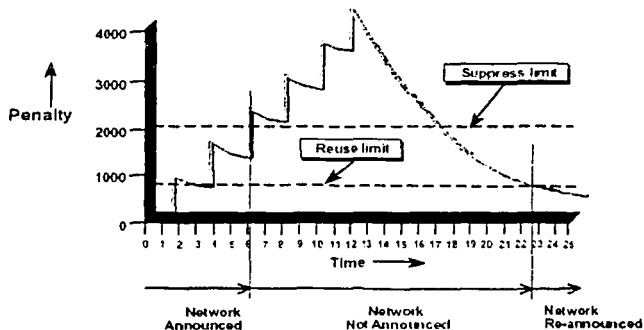


Figura 4-14 Valores típicos de las penalizaciones en los routers.

#### 4.2.7 Sincronización

Dentro de BGP existe el concepto de sincronización, el cual define la relación que debe existir entre el IBGP y el protocolo IGP. Una pregunta que muchas veces surge: ¿cuestiona porque es necesario tener un IGP cuando se podría utilizar a IBGP para que transporte los anuncios de enrutamiento a través del AS, una posible respuesta es que sería necesario que todos los routers tendrían que poseer sesiones de IBGP en Full-Mesh, para evitar loops de enrutamiento dentro del mismo AS y asegurar que todos los routers conozcan una ruta de BGP para avanzar los paquetes.

La regla de sincronización dice:

*Antes de que una ruta sea aprendida desde un vecino IBGP, deberá existir una entrada de ella en la tabla de enrutamiento del IGP, o para ser anunciada a un vecino de BGP deberá ser conocida primero por el IGP.*

Para demostrar por que es necesaria la sincronización, analicemos la figura 4-15. En este caso IBGP no es usado como IGP, para ello se está utilizando un IGP (OSPF, ISIS, etc.). Los routers A y B están conectados a dos diferentes sistemas autónomos y ambos se

anuncian las redes que aprenden de sus vecinos de EBGP por medio de una conexión de IBGP. La sesión de TCP pasa a través de los routers E y F.

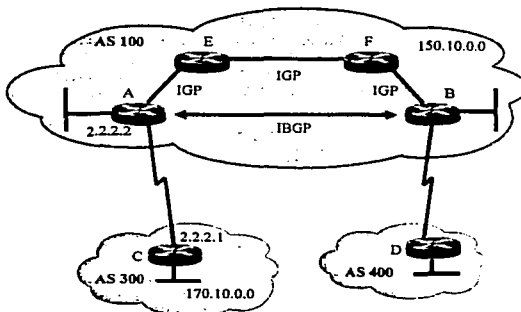


Figura 4-15 Regla de sincronización de BGP.

Suponiendo que A aprende la red 170.10.0.0 desde el AS 300 y anuncia la ruta al router B como si fuera el mismo (Next-Hop-Self) el router B anunciará la red al AS 400. El router en el AS 400, enviará entonces los paquetes con destino hacia el Router B (recordando que cuando se anuncia una ruta se hace una promesa de entrega).

Es en este momento cuando las paquetes van al router B, este hace una búsqueda para la red 170.10.0.0, determinando que es conocida a través de A, entonces busca por la dirección IP de A y se da cuenta que es a través del router F. Entonces los paquetes son avanzados a F, pero como F no conoce rutas externas (debido a que solo están intercambiando rutas dentro del mismo AS con un IGP) y al no conocer una ruta los paquetes son descartados (creando un hoyo negro), también es común que los routers E y F tuvieran una ruta por default hacia los routers A o B lo cual crearían un loop de enrutamiento en el caso de que la ruta por default de F fuera hacia B.

La sincronización evita que se presenten hoyos negros debidos a que el IGP no conoce de la suficiente información de enrutamiento para avanzar los paquetes.

TESIS CON  
FALLA DE ORIGEN

La sincronización hoy en día resulta un concepto anticuado para redes grandes que de forma implícita requieren realizar redistribuciones de BGP hacia el IGP para que el IGP tenga la suficiente información de enrutamiento.

En resumen para que IBGP funcione correctamente, se tienen dos opciones:

- Las rutas externas deben ser redistribuidas hacia el IGP para asegurar que el IGP se pueda sincronizar con el IBGP. La implementación de esta opción llega a ser inadecuada debido al gran tamaño en las tablas de enrutamiento en Internet/Internet2 que en la mayoría de los casos llegan a saturar a los routers.
- Los routers deben de poseer sesiones de IBGP en Full-Mesh y la sincronización debe ser deshabilitada. De esta forma todos los routers tienen el conocimiento de las rutas externas vía BGP, y teniendo deshabilitada la sincronización los routers anunciarán y aceptarán las rutas sin tener que informar primero al IGP.

En redes muy grandes el Full-Mesh llega hacer un problema para poder crecer, debido al gran número de sesiones de IBGP, para solucionar este problema se utilizan otras técnicas como son los route-reflectors y las confederaciones, que son descritas más adelante.

#### 4.2.8 Implementación de BGP a gran escala

##### 4.2.8.1 Comunidades

Las comunidades aplican a un grupo de rutas con las mismas políticas. Un router agrega una comunidad preconfigurada de un valor definido del atributo COMMUNITY, con lo cual la ruta se hace miembro de esa comunidad. Entonces los routers vecinos pueden aplicar sus políticas, como son; el filtrado o la modificación de otro atributo de BGP de acuerdo al valor del atributo COMMUNITY. Una ruta puede ser miembro de varias comunidades.

La tabla 4-3 muestra el uso de las comunidades que fue enviado a la lista de NANOG<sup>3</sup>:

<sup>3</sup> NANOG por sus siglas en inglés North American Network Operators' Group.

SHORT NAME	COMMUNITY	WHAT IT DOES
Local Pref = 80	701:80	set localpref 80
Local Pref = 120	701:120	set localpref 120
AS Path prepend 1	701:1	prepend 1x: 701 [cust-AS]
AS Path prepend 2	701:2	prepend 2x: 701 701 [cust-AS]
AS Path prepend 3	701:3	prepend 3x: 701 701 701 [cust-AS]
Cust but not peers	701:20	propagate to custs, not peers
keep cust routes in North America	701:30	send to custs & peers, but not 702, 703...
keep AS7046 in AS701	no-export	don't propagate beyond AS701
peers	701:666	don't propagate beyond this AS
peers	701:1030	don't propagate beyond this AS

Tabla 4-3 Ejemplo del uso de las comunidades.

El uso de las comunidades evita que los administradores tengan que realizar filtros demasiado extensos y difíciles de administrar que contienen todas las redes a las que se les desean definir políticas de enrutamiento. El valor de la comunidad se puede definir tanto en sesiones IBGP como en sesiones EBGP, lo cual permite a los vecinos externos poder definir políticas de enrutamiento fuera de su sistema autónomo, pudiendo de ésta forma definir caminos primarios y secundarios dentro y fuera de su sistema autónomo.

Una herramienta importante es tener una base de datos que almacene la información de enrutamiento de una ruta (número de AS, local\_pref, etc) para ello se crearon los IRR (Internet Routing Register), los cuales pueden almacenar toda esta información.

#### 4.2.8.2 Route Reflectors

Los Route Reflectors (Reflectores de Rutas) son útiles cuando dentro de un AS se tiene una gran cantidad de peers de IBGP. A menos que las rutas de EBGP sean redistribuidas dentro del IGP del AS, todos los IBGP deberán poseer sesiones en Full-Mesh. Entonces para  $n$  routers se deberán tener  $n(n-1)/2$  sesiones de IBGP dentro del mismo AS. En la figura 4-16 se muestra que para 8 routers serían necesarias 56 conexiones de IBGP.

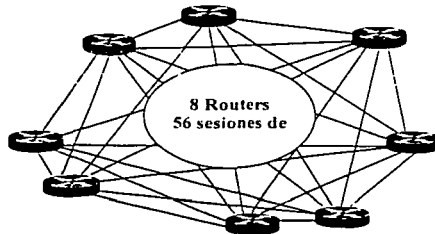


Figura 4-16 Sesiones de IBGP para 8 routers en Full-Mesh.

El uso de Route Reflectors es una alternativa al Full-Mesh. Un router es configurado como un Route Reflector (RR), y los otros routers con sesiones de IBGP, los cuales son conocidos como clientes del RR, pero preferentemente solo se realizan sesiones con el RR en lugar de realizarlas hacia el resto de los otros routers, el resultado es algo similar a la figura 4-17. Como resultado el número de sesiones requeridas son reducidas de  $n(n-1)/2$  a  $n-1$ . Un Route Reflector y sus clientes todos en conjunto son conocidos como cluster (grupo).

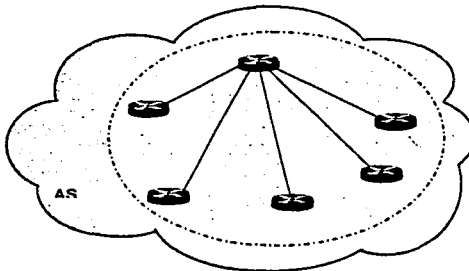


Figura 4-17 Route Reflector y sus clientes.

El Route Reflector flexibiliza la regla que especifica que los vecinos de IBGP no pueden anunciar las redes que ellos aprenden de otros vecinos de IBGP. En la red de la figura 4-18 el RR aprende redes de sus clientes. A diferencia de los otros routers de IBGP, el RR puede anunciar las rutas a sus otros clientes y no-clientes, en otras palabras las rutas de un



cliente son reflejadas por medio del RR a los otros clientes. Para evitar posibles loops de enrutamiento el RR no puede modificar los atributos de las rutas que recibe desde sus clientes.

Si un RR recibe múltiples rutas para un mismo destino, utilizará el proceso normal para la selección de la mejor ruta. El RFC1966 define las tres reglas que un RR debe utilizar para determinar cual ruta debe anunciar a quien, dependiendo de cómo fue aprendida:

- Si una ruta fue aprendida desde un IBGP no-cliente, solamente será reflejada a los clientes.
- Si una ruta fue aprendida desde un cliente, será reflejada a todos los clientes y no-clientes.
- Si la ruta fue aprendida desde un vecino EBGP, será reflejada a todos los clientes y no clientes.

La funcionalidad del RR solamente tiene que ser soportada dentro del router reflector. Desde la perspectiva de los clientes, ellos solamente se están conectando con otro router vía IBGP. Esta es una característica muy atractiva, ya que un router con una implementación básica de BGP pueden ser cliente en un cluster.

Un solo RR introduce un punto de falla para todo un sistema. Si el RR falla los clientes pierden toda la información, debido a que tienen una sola fuente de información, por lo tanto para poseer redundancia en un cluster se puede tener más de un RR como lo muestra la figura 4-19. Se recomienda que los clientes tengan conexiones físicas hacia ambos RR, de esta forma si un RR falla los clientes aún tienen la conexión con el otro RR, por lo tanto, los clientes no pierden su información.

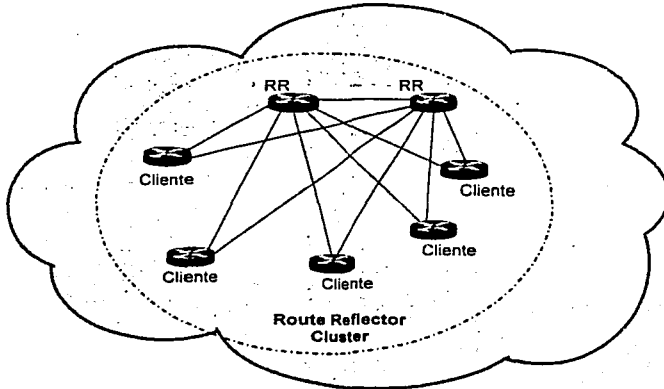


Figura 4-18 Redundancia empleando múltiples Route Reflectors.

La figura 4-19 muestra que un AS puede tener múltiples clusters, en donde cada cluster puede tener más de un RR para redundancia y para una mayor redundancia los clientes incluso pueden estar conectados entre ellos.

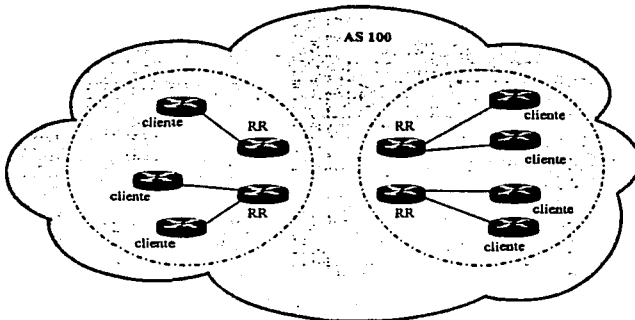


Figura 4-19. Múltiples clusters en un AS.

Debido a que los clientes no saben que lo son, un Route Reflector puede ser cliente de otro RR, como resultado los clusters de RR se pueden anidar tal y como lo muestra la figura 4-20.

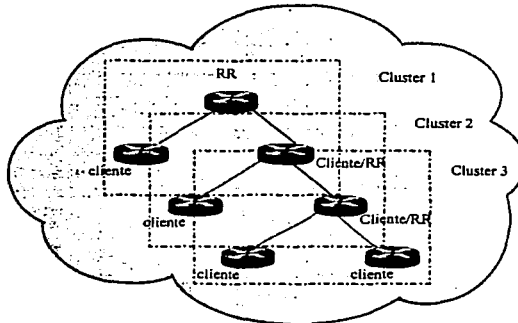


Figura 4-20 Clusters anidados.

Además, los clientes no pueden hacer peer con un router fuera de su mismo cluster, pero sí con otros routers que pertenezcan a su cluster. Incluso los clientes de un mismo cluster pueden estar en Full-Mesh. Configurando de esta forma al RR, no refleja las rutas de un cliente a otro, en lugar de eso el RR refleja solo las rutas de los no-clientes a los clientes y las de los clientes a los no-clientes, esto es ilustrado en la figura 4-21.

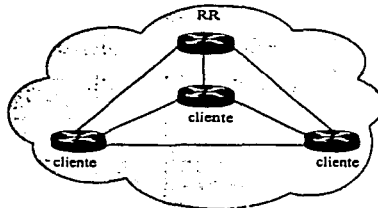


Figura 4-21 Cluster en Full-Mesh.

Recordando el concepto de sincronización de IBGP e IGP, la cual dice que BGP no avanzará una ruta que fue aprendida por un vecino interno, debido a que no se modifica el atributo AS-PATH dentro del AS y esto podría crear loops de enrutamiento. Si embargo teniendo en cuenta que un RR es un router de BGP en el cual esta regla fue flexibilizada para prevenir loops de enrutamiento, el RR usa dos atributos de una ruta de BGP: ORIGINATOR\_ID y CLUSTER\_LIST.

El ORIGINATOR\_ID es un atributo opcional, nontransitive que es creado por el RR. El ORIGINATOR\_ID es el Router\_ID del router que origino la ruta dentro del AS. si el router que origino el anuncio recibe una actualización con su propio Router\_ID, la actualización es ignorada.

Cada cluster dentro de un AS debe ser referido con un identificador único de cuatro octetos. Si el cluster contiene un solo RR el cluster ID es el Router\_ID del RR. Si el cluster contiene múltiples RRs, el cluster ID debe ser configurado manualmente.

El CLUSTER\_LIST de forma similar al AS\_PATH es un atributo opcional, nontransitive que permite rastrear los clusters por los cuales paso una actualización. Cuando un RR refleja una ruta a clientes y no-clientes, éste agrega al final su CLUSTER\_ID al CLUSTER\_LIST. Si el CLUSTER\_LIST esta vacío el RR crea uno. Entonces cuando un RR recibe una actualización, checa el CLUSTER\_LIST, si ve su propio cluster ID en la lista, se da cuenta de que ha ocurrido un loop de enrutamiento e ignora la actualización.

#### 4.2.8.3 Confederaciones

Las confederaciones surgen como otra forma de controlar el gran numero de sesiones de IBGP. Una confederación es un AS que se subdivide en grupos de subsistemas autónomos (sub-AS), conocidos como miembros del sistema autónomo, figura 4-22.

Los speakers de BGP dentro de una confederación hablan IBGP con sus vecinos de su mismo sub-AS y EBGP con los vecinos de otros sub-AS. A la confederación se le asigna una *confederation ID*, el cual es el número de AS que representará a toda la confederación y

también representará a todos los routers fuera de la confederación. Los vecinos Externos a la confederación no podrán ver la estructura interna de la confederación, solamente podrán ver un solo AS. En la figura 4-22, el AS 9184 es el confederation ID.

Anteriormente fueron descritos dos tipos de "AS\_PATH": AS\_SEQUENSE y AS\_SET. Las confederaciones agregan dos tipos más de AS\_PATH:

- AS\_CONFED\_SEQUENSE – Es una lista ordenada de AS para llegar a un cierto destino. Este es usado de la misma manera que el AS\_SEQUENSE, excepto que los números de AS en la lista pertenecen a los sub-sistemas autónomos usados localmente en la confederación.
- AS\_CONFED\_SET – Es una lista desordenada de números de AS para llegar a un destino. Este atributo es usado de la misma forma que el AS\_SET, excepto que los números de AS en la lista pertenecen a los sub-sistemas autónomos usados localmente en la confederación.

Debido a que el atributo de AS\_PATH es usado en las actualizaciones entre los diferentes subsistemas se deben de prevenir loops. Desde la perspectiva de un Router que habla BGP miembro de un sub-AS, todos los vecinos en otros sub-AS son vecinos externos.

Cuando una actualización se envía a un vecino externo a la confederación, la información del AS\_CONFED\_SEQUENSE y del AS\_CONFED\_SET es removida del atributo de AS\_PATH, y el confederation ID es agregado al final en el AS\_PATH. Esto es para que los vecinos externos a la confederación la vean como un solo AS en lugar de una colección de ASs. En la figura 4-22 se muestra una práctica que también es común, la cual consiste en utilizar el rango reservado para ASs privados para asignarlos a los sub-AS.

Cuando se realiza la selección de una ruta, el proceso de BGP se mantiene igual, con la adición de un paso: Las rutas EBGp externas a la confederación se prefieren sobre las rutas de EBGp de los sub-AS, las cuales son preferidas sobre las de IBGP. Otra diferencia entre las confederaciones y los sistemas autónomos normales es la forma en la que algunos

atributos son tratados. Atributos como NEXT-HOP y MED son anunciados sin cambios a los vecinos en otros sub-AS de la confederación.

A diferencia del uso del Router Reflector en el cual la funcionalidad solo debe ser soportada por el Router elegido como el reflector, todos los Router dentro de una confederación deben de soportar esta funcionalidad. Ya que deben de poder reconocer los tipos de AS\_PATH: AS\_CONFED\_SEQUENCE y AS\_CONFED\_SET.

En un sistema autónomo muy grande se pueden utilizar juntas la implementación de confederaciones y Route-reflector. Se pueden configurar uno o más clusters dentro de un sub-AS para tener un mejor control de los vecinos de IBGP.

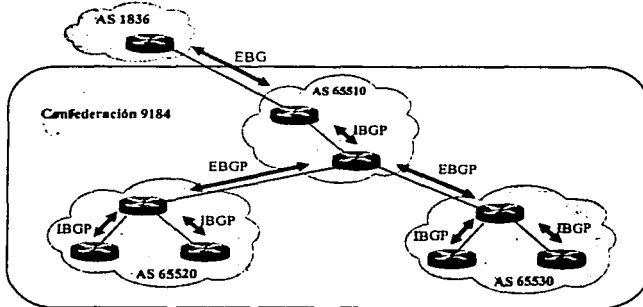


Figura 4-22 Confederaciones.

#### 4.2.9 ¿Quien necesita BGP?

No todas las organizaciones requieren de BGP para conectarse a Internet, otro concepto erróneo es que organizaciones muy grandes, como lo podría ser una empresa transnacional o internacional que se requiere dividir un varios dominios de enrutamiento emplee como solución la implementación de BGP para unir los dominios. Muchas veces no es necesario agregar tal nivel de complicación, la posible mejor solución es convertir los dominios en áreas de OSPF.

TECNOLOGÍA  
FALLA DE ORIGEN

#### 4.2.9.1 Single-homed

En la figura 4-23 se muestra una conexión típica de un suscriptor a Internet vía un proveedor de servicios (ISP), en el cual se tiene solo una conexión al ISP. BGP o cualquier otro protocolo de enrutamiento es innecesario en esta topología. Si el enlace falla, no se realizará ningún proceso, esto es por que no existe una ruta alterna. Entonces un protocolo de enrutamiento no ayuda en nada. En esta topología bastará con que el suscriptor agregue una ruta estática de default-route y la redistribuya al resto de la organización vía algún IGP.

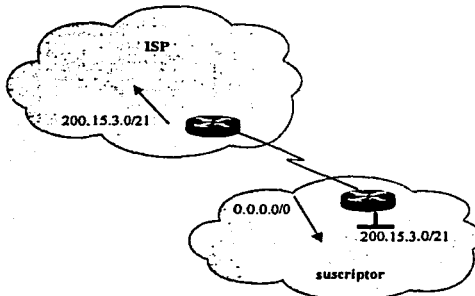


Figura 4-23 Topología de Single-homed.

De forma similar el ISP deberá agregar una ruta estática apuntando al rango asignado al suscriptor y redistribuirla al resto de su sistema autónomo vía un IGP. Lo ideal es que el rango del suscriptor sea parte del rango asignado al ISP y aun cuando no lo sea, el ISP deberá anunciar las redes del suscriptor con su número de AS como origen al resto del Internet.

#### 4.2.9.2 Multi-homed

En la figura 4-24 se muestra una topología más adecuada, cuando se requiere tener redundancia en el enlace a un mismo ISP. La forma en que se realiza la manipulación del tráfico de entrada (incoming) y salida (outgoing) depende de la forma en la que los enlaces serán usados. Por ejemplo un uso típico puede ser definir a un enlace como el principal y al otro como respaldo, en este caso no es necesario utilizar BGP, bastaría con que el ISP utilice un valor diferente en las métricas del IGP –dependiendo del IGP un valor mayor o menor- de tal forma se prefiera el enlace primario, esto sería para el tráfico de entrada al suscriptor y el suscriptor deberá hacer algo similar pero anunciando la ruta por default con diferente métrica, en la figura 4-24 se muestra un ejemplo con costos para OSPF.

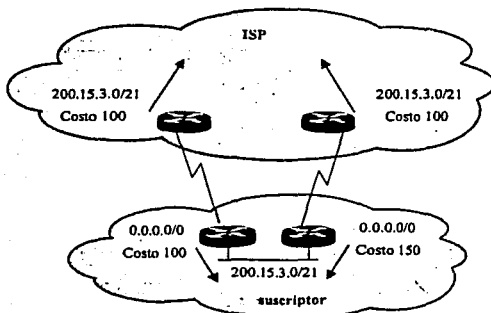


Figura 4-24 Multihomming a un solo Sistema Autónomo

En el caso de que la conexión sea a más de un proveedor (figura 4-25), lo cual proporciona una mayor redundancia al suscriptor debido a que si alguno de los ISPs tuviera una falla dentro de su sistema autónomo, todavía tendría su salida y regreso a través del otro ISP.



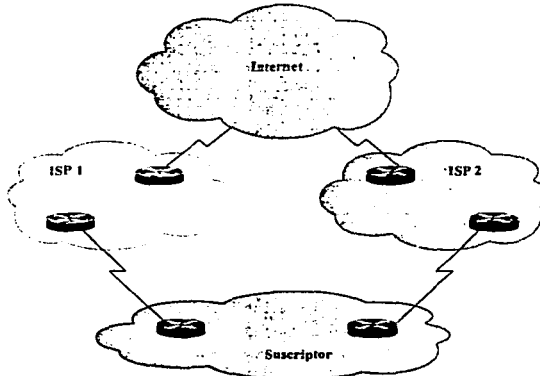


Figura 4-25 Multihoming a diferentes ISP.

En este caso el suscriptor puede ser una organización o un ISP pequeño que está conectado a otros ISPs. En este caso el suscriptor tiene que superar varios obstáculos, en caso de que no tuviera su propio espacio de direcciones y utilizará parte del rango de uno de los ISPs:

- El proveedor del espacio de direcciones debe ser persuadido de hacer un hoyo en su bloque de CIDR.
- El segundo proveedor debe ser persuadido de anunciar un bloque de direcciones que pertenece a otro proveedor.
- Ambos proveedores se deben de coordinar para realizar el anuncio del espacio de direcciones asignado al suscriptor.
- Si el espacio de direcciones es menor a un prefijo de 19 bits, algunos backbones en Internet podrían no aceptar el anuncio.

Los mejores candidatos a ser multihoming a varios ISP son organizaciones bastante grandes o ISPs de menor tamaño que normalmente tienen su propio bloque de direcciones y su propio número de AS.

En estos casos se requiere BGP, debido a que BGP es rico en atributos para poder modificar la forma en que el tráfico viaja a través de las múltiples conexiones que se posean.

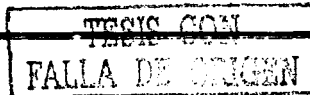
Regresando al ejemplo de la figura 4-25 el suscriptor puede utilizar BGP para definir a uno de ellos como el ISP primario y el otro como respaldo, para lo cual bastaría que el ISP primario anunciara la ruta por default y el otro ISP también tendría que anunciar la ruta de default pero con una menor preferencia además de los prefijos de su propio AS y de sus clientes. En algunas ocasiones en lugar de anunciar las rutas por default, se realiza el anuncio de todas las rutas que el ISP conoce (a veces todas las de Internet).

Normalmente los ISPs prefieren ser ellos mismos quienes modifiquen los anuncios, debido a que esto les proporciona el control y la posibilidad de cambiar el flujo del tráfico cuando ellos lo deseen y como lo deseen.

#### 4.2.9.3 Peering/Tránsito

Del ejemplo anterior surgen dos conceptos importantes para los ISPs: ¿cuando un AS es de tránsito? y ¿cuando solo se realiza intercambio de tráfico?, a este segundo concepto se le acostumbra llamar peering.

Cuando un AS propaga los anuncios de un ISP al resto del Internet, lo que hace es decirle al resto de las redes de Internet que pueden pasar a través de para llegar a ese ISP. Como se muestra en la figura 4-26 el ISP1 sirve de "tránsito" para el ISP2.



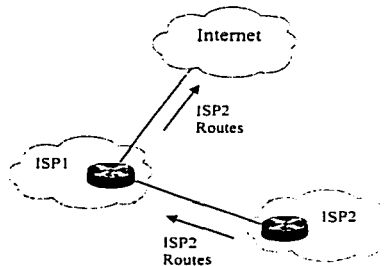


Figura 4-26 ISP de tránsito a Internet.

Con referencia a la conexión al resto de Internet, al ISP1 se le conoce como el ISP de Upstream y al ISP2 como Downstream.

Cuando solo se acuerda realizar el intercambio de tráfico entre dos ASs los anuncios son recibidos por ambos ISPs, pero los anuncios no son propagados a todo el Internet. Como se muestra en la figura 4-27 el ISP 1 y el ISP 2 solamente están realizando "peering" ya que ambos no propagan los anuncios del otro al resto de Internet.

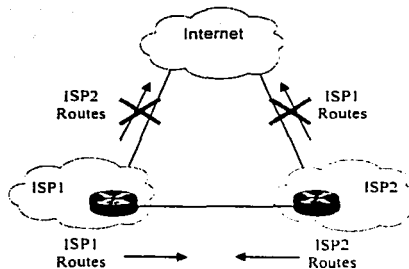


Figura 4-27 ISP en peering.

4.2.9.4 IXPs/NPAs

Es muy común que varios ISPs se unan para realizar el intercambio de tráfico en un cierto punto, esto se hace para poder ofrecer sus servicios a mas bajo costo sin tener que pagarle a otro ISP para comunicarse. Estos puntos son llamados puntos de intercambio -Exchange Point, IXP-. En la figura 4-28 se muestra la forma en como se realiza una conexión de este tipo.

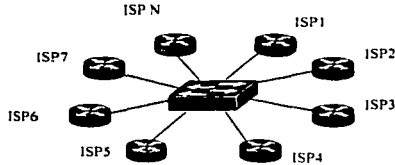


Figura 4-28 Punto de intercambio (IXP).

Como se mencionó en el capítulo 1 de este trabajo de tesis, después de la creación de la NFSNET, se crearon puntos de intercambio con la obligación de tener conexión a todas las redes. En la figura 4-29 se muestra el funcionamiento del Internet actual, en donde en la parte más alta se encuentran los NAPs después pueden existir varios ISPs antes de llegar a los usuarios de Internet.

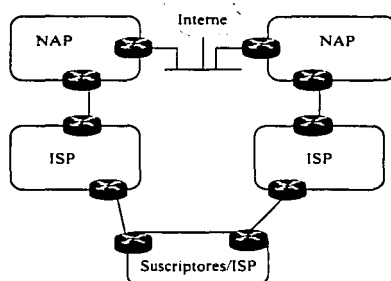
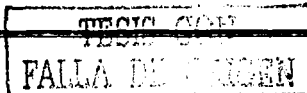
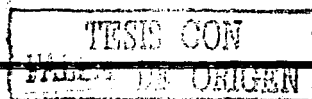


Figura 4-29 Estructura del Internet actual.



Uno de los riesgos del protocolo de enrutamiento BGP es que no conoce sobre anchos de banda, pudiendo ocurrir que el tráfico escoja los caminos menos adecuados y dada la gran cantidad de conexiones, es muy posible que se presenten caminos asimétricos (el camino de salida es diferente del de regreso) lo cual puede ocasionar efectos indeseados en muchas aplicaciones, al tener tiempos de respuesta diferentes, en la transmisión y recepción.



# Capítulo 5

## Descripción de los protocolos de Multicast

TESIS CON  
FALLA DE ORIGEN

---

## Capítulo 5 Descripción de los protocolos Multicast

### 5.1 Concepto de Multicast

Multicasting es el proceso de enviar datos a un grupo de receptores. Puede ser que sea discutido que Unicasting y Broadcasting son subconjuntos de Multicasting. En el caso de Unicasting, hay solamente un solo miembro del grupo; en el caso de Broadcasting, todos los posibles receptores son miembros del grupo.

Los tres requerimientos básicos para utilizar Multicast a través de una red enrutada son los siguientes:

- Debe de haber un conjunto de direcciones para los cuales los grupos de Multicast sean identificados.
- Debe de existir un mecanismo con el cual los hosts puedan unirse o abandonar grupos.
- Debe de haber un protocolo de enrutamiento que permita a los routers entregar eficientemente tráfico de Multicast a los miembros del grupo sin sobre-utilizar los recursos de la red.

Un grupo de Multicast es definido por su dirección IP Multicast; los grupos pueden ser permanentes o transitorios. *Permanente* se refiere al hecho de que el grupo siempre tiene asignada una misma dirección, y no que los miembros siempre están asignados al mismo grupo. De hecho, los hosts son libres de unirse o abandonar cualquier grupo. Los grupos *transitorios* son los grupos que no tienen una existencia permanente, una dirección sin reservar es asignada al grupo y es liberada cuando el grupo deja de existir.

La figura 5-1 ilustra la transmisión de un paquete realizando un comparativo entre Unicast y Broadcast.

TESIS CON  
FALLA DE ORIGEN

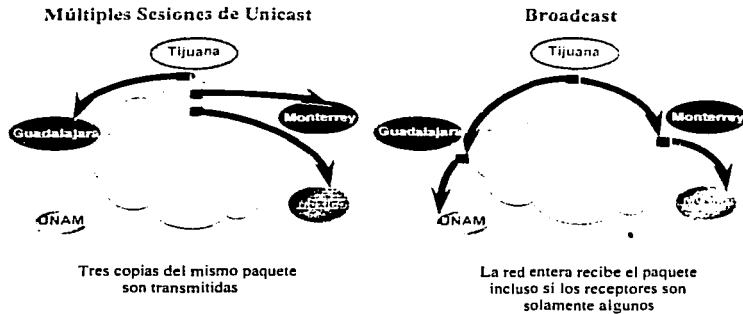


Figura 5-1 Dos técnicas multipunto ineficientes.

En la transmisión Multicast es posible enviar un solo paquete de datos a múltiples receptores, como lo muestra la figura 5-2.

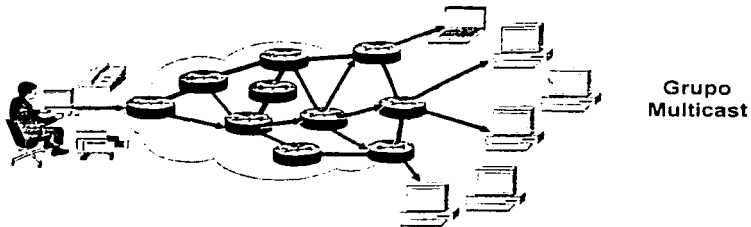


Figura 5-2 Multicast es la solución para enviar un paquete a varios destinatarios.



### 5.1.1 Ventajas de Multicast

La transmisión de Multicast ofrece muchas ventajas sobre la transmisión Unicast en ambientes uno-a-muchos (one-to many) o muchos-a-muchos (many-to-many).

- *Alta Eficiencia:* el ancho de banda disponible en la red se utiliza más eficientemente, puesto que los múltiples flujos de datos son reemplazados con una sola transmisión.
- *Funcionamiento Óptimo:* menos copias de datos requieren ser avanzados y procesados.
- *Aplicaciones Distribuidas:* permite el empleo de aplicaciones distribuidas:
  - Con la transmisión Unicast el nivel de tráfico y de clientes se incrementa en una tasa de 1:1
  - Con la transmisión de Multicast la tasa de incremento es muy reducida.
- *La transmisión de Multicast:* envía un solo paquete de Multicast direccionado a todos los destinatarios independientes.
- Introduce una nueva clase de direcciones IP. "Clase D" = (224.0.0.0 - 239.255.255.255).
- Eficiente comunicación y transmisión.

### 5.2 Tipos de direcciones

**TESIS CON  
FALLA DE ORIGEN**

#### 5.2.1 Direcciones Unicast

Unicast es la comunicación existente entre un solo emisor y un solo receptor sobre una red. El término existe en contradicción a Multicast, que es la comunicación un solo emisor y múltiples receptores, y la comunicación Anycast existente entre cualquier emisor y el receptor más cercano de un grupo en una red, un término más fácil, es la comunicación punto-a-punto que tiene un significado similar a Unicast.

Este tipo de direcciones hace referencia a un único host (interfaz) dentro de la subred.

- Un ejemplo de dirección IP Unicast es 192.168.100.9. Una dirección MAC Unicast es, por ejemplo, 80:C0:F6:A0:4A:B1.

### 5.2.2 Direcciones Broadcast

Broadcast es la comunicación entre un solo emisor y todos los receptores de una red. Con una dirección de este tipo se consigue direccionar a todos los hosts (interfases) dentro de una subred.

- Una dirección IP Broadcast es 192.168.100.255 y una dirección MAC Broadcast es: FF: FF: FF: FF: FF: FF.

### 5.2.3 Direcciones Multicast

Este tipo de direcciones identifica a un conjunto de interfases de la red, de manera que el paquete se envía a todos los miembros de un grupo.

El espacio de direccionamiento IP se distribuye en tres grupos o clases de direcciones, las direcciones de clase A, B y C. Hay una cuarta clase, la clase D, reservada para las direcciones Multicast. La clase D tiene reservado el rango de direcciones IPv4 entre la 224.0.0.0 y la 239.255.255.255.

Los 4 bits de mayor peso de la dirección IP permiten direccionar entre el valor 224 y el 239. Los 28 bits restantes de menor peso, están reservados para el identificador del grupo Multicast, tal y como se muestra en la figura 5-3.

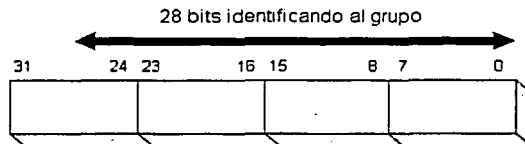


Figura 5-3 Formato de las direcciones IP Clase D.

TESIS CON  
FALLA DE ORIGEN

- Direcciones IP de Multicast
  - 224.0.0.0–239.255.255.255
  - Clase "D" Espacio de direcciones
    - Bits de mayor orden del primer octeto = "1110"
- Direcciones Reservadas en Enlaces Locales
  - 224.0.0.0–224.0.0.255
  - Transmitidas con TTL = 1
  - Ejemplos:
    - 224.0.0.1 Todos los sistemas en esta subred
    - 224.0.0.2 Todos los routers en esta subred
    - 224.0.0.4 DVMRP Routers
    - 224.0.0.5 OSPF Routers
    - 224.0.0.13 PIMv2 Routers
- Alcance administrativo de las direcciones
  - 239.0.0.0–239.255.255.255
  - Espacio de direcciones privado
    - Similares a las direcciones de Unicast del RFC1918.
    - No utilizadas para el tráfico global de Internet.
    - Uso de alcance limitado.
    - Las mismas direcciones pueden estar en uso en diferentes locaciones para diferentes sesiones de Multicast.
- GLOP-RFC 2770: asignamiento de direcciones estáticas de grupo.
  - Método temporal para satisfacer necesidades inmediatas.
  - Rango del grupo: 233.0.0.0 - 233.255.255.255.
    - Un número de AS es insertado en medio de dos octetos.
    - El octeto restante de orden inferior es usado para la asignación de grupo.

TESIS CON  
 FALLA DE ORIGEN

### 5.3 Direcciones Multicast de capa 2

Las direcciones Multicast IPv4 a nivel de red, deben asociarse con las direcciones MAC (capa de enlace) correspondientes al tipo de red con el que se esté trabajando. Si se estuviese trabajando con direcciones a nivel de red Unicast, se obtendría la dirección MAC asociada haciendo uso del protocolo ARP, en el caso de direcciones de red Multicast, no se puede usar ARP y habrá que obtener la dirección MAC asociada mediante un procedimiento diferente. Se han definido varios documentos RFC que especifican la forma de realizar esta asociación:

- Correspondencia de direcciones Multicast IPv4 a direcciones MAC Ethernet: RFC 1112.
- Correspondencia a redes FDDI: RFC 1390.
- Correspondencia a redes Token-Ring: RFC 1469.

#### 5.3.1 Ethernet/FDDI

Una tarjeta de interfaz de red (NIC) de un miembro de un grupo también debe conocer Multicast. Cuando un host se une a un grupo, la NIC determina una dirección MAC fiable. Para lograr esto, todas las NICs de Ethernet, Token Ring y FDDI usan la dirección reservada IEEE 802 0100.5E00.0000 para determinar una única dirección MAC de Multicast. Esto significa que el octavo bit de esta dirección es 1; ese bit, en el formato 802, es el bit Individual/Grupo (I/G). Cuando este bit tiene un valor igual a uno, indica que la dirección es una dirección de Multicast.

Las interfases Ethernet al igual que las FDDI asocian los 23 bits más bajos del grupo de una dirección IP a los 23 bits más bajos de la dirección MAC reservada para formar una dirección MAC de Multicast.

La figura 5-4 muestra un ejemplo. Aquí, la dirección IP 235.147.18.23 clase C es usada para crear la dirección MAC 0100.5e13.1217.

TESIS CON  
FALLA DE ORIGEN

<b>Dirección IP Multicast</b>						
Decimal:	235	147	18	23		
Hex:	EB	93	12	17		
Binario:	11101011	10010011	00010010	00010111		
<b>Dirección MAC base</b>						
01	00	5E	00	00	00	
00000001	00000000	01011110	00010011	00000000	00000000	
<b>Dirección MAC Multicast</b>						
00000001	00000000	01011110	00010011	00010010	00010111	
01	00	5E	13	12	17	

Figura 5-4 Mapeo de una dirección IP Multicast a una dirección de capa 2 Ethernet/FDDI.

Calculando el cociente del número total de direcciones clase D ( $2^{28}$ ) al número de posibles direcciones MAC bajo el prefijo reservado ( $2^{23}$ ) muestra que 32 diferentes direcciones IP de clase D pueden ser asociadas a una sola dirección MAC, figura 5-5.

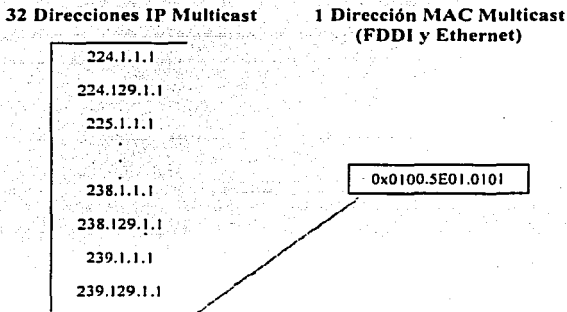


Figura 5-5 32 Direcciones IP Multicast pueden ser mapeadas a una misma dirección MAC Ethernet/FDDI.

TESIS CON  
FALLA DE ORIGEN

### 5.3.2 Token Ring

La Arquitectura de Referencia de las Redes Token Ring proporciona varios tipos de mecanismos de direccionamiento. Esto incluye el direccionamiento Unicast y el Multicast. Todas las direcciones de capa 3 IP Multicast se asocian a una sola *Dirección Funcional Token Ring*<sup>1</sup> o a la Dirección de Broadcast. En Token Ring hay un número limitado de direcciones funcionales (31) y por lo tanto varias funciones sin relación deben compartir la misma dirección funcional.

#### *Formato de la dirección MAC destino*

Actualmente existen dos métodos para direccionar tramas de Token Ring que transportan paquetes de Multicast:

- Usar la dirección de Broadcast FFFF.FFFF.FFFF para todas las tramas que transportan paquetes de Multicast.
- Usar una sola dirección funcional reservada, C0000.0004.0000.

#### *Dirección Funcional IP Multicast*

Debido a que hay una escasez de direcciones funcionales en Token Ring, todas las direcciones IP Multicast han sido asociadas a una dirección funcional de Token Ring. En forma canónica, esta dirección es 03-00-00-20-00-00. En forma no canónica es C0-00-00-04-00-00. Se debe observar que solamente hay 31 posibles direcciones funcionales (figura 5-6), puede haber otros protocolos que también asignen esta dirección funcional. Por lo tanto, el hecho de que se envíe una trama a la dirección funcional 03-00-00-20-00-00 no significa que es una trama de IP Multicast.



<sup>1</sup> Un subconjunto especial de grupo de direcciones es llamado "direcciones funcionales" y son indicadas por un bit en la dirección MAC destino.

**32 Direcciones IP Multicast**

224.0.1.0  
 224.0.1.1  
 225.0.1.2  
 .  
 239.239255.252  
 239.255.255.253  
 239.255.255.254  
 239.255.255.255

**1 Dirección Funcional  
 Token Ring**

0xFFFF.FFFF.FFFF  
 0  
 C0-00-00-04-00-00

**TESIS CON  
 FALLA DE ORIGEN**

Figura 5-6 En la asociación de direcciones capa 3 IP Multicast a direcciones de capa 2 Token Ring, 31 posibles direcciones IP pueden ser asociadas a una sola dirección Dirección Funcional Token Ring o a la Dirección de Broadcast.

**5.4 Soporte de Multicast en IP**

La transmisión Multicast ofrece una gran cantidad de servicios y beneficios que aumentan la funcionalidad de la red sobre la transmisión Unicast y Broadcast. En marzo de 1992 el Internet Engineering Task Force (IETF) transmitió el audio de su conferencia en San Diego y en ese momento inició la utilización de Multicast.

Las direcciones IP Multicast se suelen denominar 'grupo Multicast'. debido a que no están asignadas a un equipo en concreto de forma permanente, sino a un grupo determinado y de forma temporal.

Se han desarrollado una gran cantidad de herramientas para realizar el intercambio de aplicaciones multimedia utilizando Multicast: editores de texto (nt), pizarras electrónicas compartidas (wb), intercambio de hipertextos, sincronización de equipos (NTP Multicast) o (FTP Multicast) por citar algunos ejemplos. Algunas de estas aplicaciones experimentales, como son las de transmisión de audio y vídeo, se han convertido prácticamente en estándares de Multicast y son ampliamente utilizadas por un elevado número de usuarios

con cierta regularidad. Dentro de este grupo el "rat" y el "vat" son las más empleadas para la transmisión/recepción de audio, y el "vic" para la transmisión/recepción de vídeo. El ivs (INRIA Videoconference System), desarrollado por el INRIA (Institute National de Recherche en Informatique et en Automatique), permite la integración de los canales de audio y vídeo sobre una misma aplicación

### *EL MBONE.*

MBone (Multicast Backbone On Internet) existe desde 1992 como una red virtual para la experimentación del uso de IP Multicast en Internet. Esta red se ha empleado mayoritariamente para el estudio de herramientas de audio/vídeo conferencias multipunto, aunque en principio puede ser empleada para el intercambio de cualquier tipo de información multimedia. Su principal ventaja, o debiéramos decir característica, es la de proporcionar el intercambio de información de uno a muchos, pero sin los inconvenientes de tener que duplicar dicha información para cada uno de los receptores y en función del número de ellos.

Las misiones del trasbordador de la NASA, las reuniones periódicas de los grupos de trabajo del IETF, así como emisiones de radio por Multicast, son algunas de las sesiones disponibles con regularidad en MBone y seguidas por un gran número de usuarios en todo el mundo. Una utilidad fundamental en MBone, que permite la integración de todas las utilidades mencionadas anteriormente, es el SDR (Session Directory Tool), o directorio de sesiones MBone que nos permite conocer las sesiones que están activas en todo momento, conectarnos a cualquiera de ellas, o definir nuestra propia sesión MBone.

#### **5.4.1 Funcionamiento de IGMP**

Sin tener en cuenta cual de los varios protocolos de enrutamiento es usado en una red interna de Multicast, IGMP es siempre el "protocolo" empleado entre los hosts y los routers. Todos los hosts que quieran ingresar a grupos de Multicast, y todos los routers con interfases en subredes que contengan hosts Multicast, deben implementar IGMP.



El principal propósito de IGMP es permitir que los hosts que desean recibir tráfico de IP Multicast se comuniquen con los routers de la red local. Alternadamente, esto, permite a los routers con soporte de IP Multicast "Ingresar" a un grupo específico de Multicast y de esta manera, empezar a avanzar tráfico de Multicast hacia el segmento de red.

La especificación inicial para IGMP (v1) fue documentada en el RFC 1112, "*Host Extensions for IP Multicasting*". Desde entonces, muchos problemas y limitaciones han sido descubiertos dentro de IGMPv1. Esto condujo al desarrollo de la especificación IGMPv2, la cual fue ratificada como el RFC 2236 en noviembre de 1997.

Incluso, antes de que IGMPv2 fuera ratificado, los trabajos sobre la siguiente generación del protocolo IGMP, IGMPv3, ya habían comenzado. Sin embargo, la especificación IGMPv3 todavía se encuentra en la etapa de trabajo y aun no ha sido implementada por ningún vendedor.

#### *Queries de Membresía IGMP*

Las solicitudes de Membresía (*Membership Queries*) de IGMPv1 son enviados por el router a la dirección Multicast 224.0.0.1 "*All-Hosts*" para solicitar cuales son los grupos de Multicast que tienen receptores activos en la red local.

#### *Informes de Membresía de IGMP*

Los Informes de Membresía (*Membership Report*) de IGMPv1 son enviados por los hosts que desean recibir tráfico de un grupo específico de Multicast. Los Membership Reports son enviados con un TTL=1 a la dirección de Multicast del grupo del cuál los hosts desean recibir tráfico. Los hosts envían informes de forma asincrónica (cuando desean ingresar a un grupo) o en respuesta a los Membership Queries. En el último de los casos, la respuesta es usada para mantener al grupo, en un estado activo, así que el tráfico para el grupo continúa siendo avanzado hacia el segmento de red.

TESIS CON  
FALLA DE ORIGEN

*Informe de Supresión*

El informe de supresión es usado entre miembros del grupo, de modo que todos los miembros no tengan que responder a una solicitud. Esto ahorra tiempo de CPU y ancho de banda en todos los sistemas.

La regla en la membresía de Multicast es que mientras un miembro esta presente, el tráfico debe ser avanzado al grupo sobre este segmento. Por lo tanto, se requiere que solamente un miembro este presente para mantener el interés en un grupo dado, por lo cual el informe de supresión ayuda a mantener un uso eficiente de los recursos de la red.

*TTL*

Los paquetes de Membership Query y del Informe solamente tienen significado local, el TTL de estos paquetes siempre es fijado a 1. Así que, estos paquetes tampoco serán avanzados fuera de la subred.

5.4.1.1 Ingresando a un Grupo

TESIS CON  
FALLA DE ORIGEN

*Queries Generales*

Los queries generales son enviados a la dirección Multicast 224.0.0.1 "All-Hosts". Un miembro de cada grupo en el segmento responderá con un informe, ver figura 5-7.

Los queries generales son enviados periódicamente (el default es de 60 segundos).

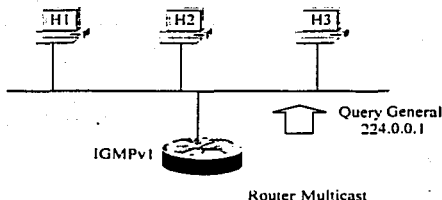


Figura 5-7 Envío periódico de Queries Generales a la dirección 224.0.0.1 para determinar a los miembros del grupo.

### *Querier IGMP*

En IGMPv1 no existe un proceso formal para la elección del *IGMP "Query Router"*. En su lugar, se deja el proceso de elección al protocolo de enrutamiento Multicast, diversos protocolos utilizan diferentes mecanismos. A menudo esto da lugar a múltiples "queriers" en una sola red multiacceso.

#### 5.4.1.2 Permanencia en un Grupo

Proceso de Respuesta a un Query

El router periódicamente envía Membership Queries IGMPv1 a la dirección de grupo 224.0.0.1 "*All-Hosts*".

Solamente un miembro de cada grupo responde con un informe a un query. Esto es para ahorrar ancho de banda en la subred y procesamiento en los hosts. Este proceso es llamado "*Response supresión*".

#### 5.4.1.3 Abandonando un Grupo

No hay un mecanismo especial de Abandono "*Leave*" definido en la Versión 1 de IGMP. En lugar de ello, en IGMPv1 los hosts abandonan un grupo "*pasivamente*" (*passively*) o "*tranquilamente*" (*quietly*) en cualquier momento sin enviar alguna notificación al router como lo muestra la figura 5-8.

No hay problema si hay múltiples miembros presentes en un grupo, debido a que el flujo Multicast es entregado a la subred. Sin embargo, cuando un miembro abandona el grupo y éste es el último miembro, existirá un periodo de tiempo en el que el router continuará avanzando innecesariamente el tráfico de Multicast después de que no haya miembros restantes dentro del segmento.

El tiempo de vida de un grupo expira, después de no haber obtenido respuesta a varios IGMP Queries Router. Esto es especialmente ineficiente si el número de grupos y/o tráfico de estos grupos es alto.

TESIS CON  
FALLA DE ORIGEN

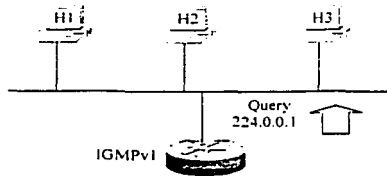


Figura 5-8 Abandonando un grupo.

#### 5.4.2 IGMPv2

Como resultado de algunas limitaciones en IGMPv1, se dio inicio a los trabajos sobre IGMPv2 en un intento por eliminar estas limitaciones. La mayoría de los cambios entre IGMPv1 e IGMPv2 son principalmente dirigidos a la solución de los problemas de la latencia de Ingreso y de Abandono, así como a las ambigüedades en la especificación del protocolo original.

A continuación se mencionan algunos de los cambios más significativos.

##### *Queries Específicos de Grupo*

Un query Específico de Grupo fue agregado en la versión 2 para solamente permitir al router hacer queries de membresía en un solo grupo, en lugar de hacerlo a todos los grupos. Esta es una forma mejorada de averiguar rápidamente si algún miembro permanece en un grupo sin necesidad de preguntar a todos los grupos por un informe.

La diferencia entre el query Específico de Grupo y el Query General es que un Query General es un envío Multicast a la dirección 224.0.0.1 "All-Hosts", mientras un query Específico de Grupo para un Grupo "G", es un envío Multicast a la dirección Multicast del Grupo "G".

TESIS CON  
FALLA DE ORIGEN

*Intervalo del Query*

- Los queries de membresía son enviados por default cada 60 segundos.
- Las diferencias importantes entre IGMPv1 e IGMPv2 son las siguientes:
  - IGMPv1 no tiene mensajes de Abandono de Grupo "*Leave Group*". significa que hay un periodo más largo de tiempo entre en el que el último host abandona un grupo y el tiempo en el que el router deja de avanzar el tráfico de grupo.
  - IGMPv1 no tiene un Query Especifico de Grupo "*Group-Specific Query*". Esto corresponde al hecho de que no hay mensaje de Abandono de Grupo "*Leave Group*".
  - IGMPv1 no especifica un Tiempo de Respuesta Máximo "*Max Response Time*" en sus mensajes Query. En lugar de esto, los hosts tienen un Tiempo Máximo de Respuesta fijo de 10 segundos
  - IGMPv1 no tiene un proceso de elección del DR. En su lugar, diferentes protocolos usan diferentes mecanismos de elección. bajo IGMPv1 es posible tener más de un DR en al subred.

En algunos casos, las implementaciones de IGMPv1 e IGMPv2 pueden existir en la misma subred:

- Algunos miembros del grupo pueden ejecutar IGMPv1 mientras otros ejecutan IGMPv2.
- Algunos miembros del grupo pueden ejecutar IGMPv2 mientras otros ejecutan IGMPv1.
- El router puede ejecutar IGMPv2 mientras algunos miembros del grupo ejecutan IGMPv1.
- Un router puede ejecutar IGMPv1 mientras otros routers en la subred ejecutan IGMPv2.

TESIS CON  
FALLA DE ORIGEN

### 5.4.3 Árboles de Distribución de Multicast

Definen la trayectoria por la que el tráfico fluye desde la fuente hasta los receptores.

#### 5.4.3.1 Árbol de Distribución de una Fuente

- La trayectoria más corta o árbol de distribución de una fuente, es un árbol mínimo con los costos más bajos de la fuente hacia todas las ramas del árbol.
- Cada SPT (Shortest Path Tree) es enrutado hacia la fuente. Esto significa que para cada fuente que envía tráfico a un grupo, hay un correspondiente SPT como lo muestra la figura 5-9.
- Los paquetes se avanzan sobre la Trayectoria más Corta del Árbol, de acuerdo a la Dirección Fuente desde la que los paquetes fueron originados y la Dirección de Grupo "G" a la que los paquetes están direccionados. Por está razón nos referimos al estado de avance (forwarding) en el Árbol de Trayectoria más Corta con la notación (S, G), donde:

"S" es la dirección IP de la fuente (source)

"G" es la dirección de grupo (group) Multicast

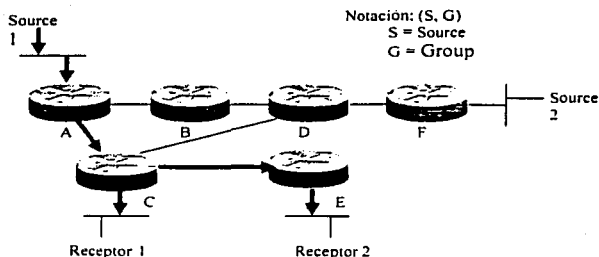


Figura 5-9 La trayectoria más corta "shortest path" entre la Fuente 1 "Source 1" y el Receptor 1 "Receiver 1" es a través de los Routers A y C, y la trayectoria más corta "shortest path" hacia el Receptor 2 "Receiver 2" es con un salto más vía el Router E.

5.4.3.2 Árbol de Distribución Compartida

- El Árbol de Distribución Compartida cuya raíz es un punto compartido en la red, es la trayectoria por la cuál fluyen los datos de Multicast para alcanzar a los receptores en la red. En PIM-SM, este punto compartido es llamado el Rendezvous Point (RP), tal como lo muestra la figura 5-10.
- El tráfico de Multicast es avanzado hacia el Árbol Compartido solo de acuerdo a la dirección de Grupo "G" a la que el paquete esta siendo direccionado, sin importar la dirección fuente. Por esta razón se hace referencia al estado de avance en el Árbol Compartido con la notación (\*, G)

Donde:

"\*" significa cualquier fuente  
 "G" es la dirección de grupo

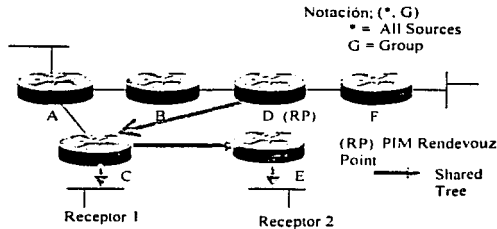


Figura 5-10 Rendezvous Point.

- Antes de que el tráfico se pueda enviar hacia abajo del Árbol Compartido "Shared Tree", de alguna manera debe ser enviado a la raíz del árbol.
- En PIM-SM, esto es realizado por el RP (Rendezvous Point) uniendo el Árbol de Trayectoria más Corta "Shortest Path Tree" a cada fuente, de modo que el tráfico pueda fluir al RP y de allí hacia abajo del árbol compartido. Para que el RP sea capaz de realizar esta función, debe ser notificado que una fuente se activará.
- En la figura 5-11, el RP ha sido informado que las Fuentes 1 y 2 están activas y subsecuentemente ha unido el SPT a las fuentes.

FALLA DE ORIGEN

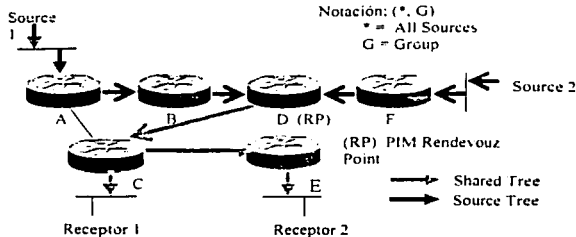


Figura 5-11 El RP ha sido notificado que las Fuentes 1 y 2 están siendo activadas y subsecuentemente unirá el SPT a estas fuentes.

### 5.4.3.3 Características de los Árboles de Distribución

#### *Trayectoria Más Corta o Árbol de Distribución Fuente*

Emplean más memoria pero obtienen trayectorias óptimas desde la fuente hacia todos los receptores; reduce al mínimo el retardo.

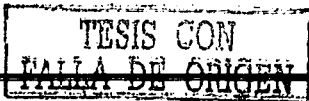
#### *Árboles Compartidos*

Emplean menos memoria pero obtienen trayectorias sub-óptimas desde la fuente hacia todos los receptores; pueden introducir retardo.

### 5.5 Enrutamiento Multicast

Los routers deben conocer el origen del paquete, en lugar del destino (opuesto a Unicast)

- la dirección IP origen denota una fuente conocida.
- la dirección IP destino denota a un grupo de receptores desconocido.





---

El enrutamiento Multicast emplea "Reverse Path Forwarding" (RPF)

- *Broadcast*: los paquetes inundan a todas las interfases salientes, excepto las interfases entrantes desde la fuente: inicialmente asumiendo que cada host en la red, es parte del grupo de Multicast.
- *Prune*: Elimina a las ramas del árbol que no poseen miembros del grupo de Multicast; corta la transmisión hacia LANs que no posean receptores interesados.
- *Selective Forwarding*: requiere integrar su propio protocolo de enrutamiento Unicast.

En el enrutamiento Multicast, la fuente envía tráfico a un grupo arbitrario de hosts que son representados por una dirección de grupo Multicast. El router Multicast debe determinar que dirección es upstream (hacia la fuente) y que dirección (o direcciones) son downstream. Si hay múltiples trayectorias downstream el router replicará el paquete y lo avanzará hacia las trayectorias downstream apropiadas, las cuales no son necesariamente todas las trayectorias. El concepto de avance de tráfico Multicast desde la fuente, en lugar que hacia el receptor, es llamado Reverse Path Forwarding (RPF)

Una terminología útil comúnmente empleada es upstream y downstream. Los paquetes de Multicast siempre deben fluir downstream de la fuente a los destinos, nunca upstream hacia la fuente. Para asegurar este funcionamiento, cada router de Multicast mantiene una tabla de avance de Multicast en la cuál el par de direcciones (fuente, grupo) o (S, G) son registradas. Los paquetes de una fuente y destino en particular para un grupo en particular siempre deben llegar en una interfaz upstream y ser avanzados en una o más interfases downstream. Por definición, una interfaz upstream esta más cerca de la fuente de cualquier interfaz downstream, como lo ilustra la figura 5-12:

TESIS CON  
FALLA DE ORIGEN

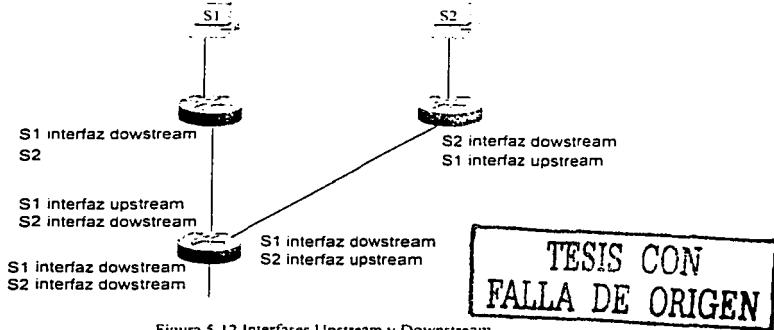


Figura 5-12 Interfases Upstream y Downstream.

### 5.5.1 ¿Que es RPF?

Un router avanzará un datagrama de Multicast, solamente si es recibido en una interfaz upstream de la fuente (siguiendo el árbol de distribución).

- Los routers avanzan paquetes de Multicast recibidos de una interfaz entrante al árbol de distribución que conduce a la fuente.
- Los routers verifican la dirección IP respecto a sus tablas de enrutamiento Multicast (Comprobación de RPF); asegurándose de que el datagrama de Multicast fue recibido en la interfaz entrante especificada.
- Observar que los cambios en la topología Unicast no necesariamente reflejarán inmediatamente un cambio en el RPF, esto depende de la frecuencia con que se realice la comprobación del RPF.
- Si se logra la comprobación del RPF, el datagrama es avanzado hacia todas las interfases salientes.
- Si la comprobación del RPF falla, el datagrama típicamente es desechado.
- Cuando un datagrama es avanzado, es enviado a cada interfaz en la lista de interfases salientes.
- Un paquete nunca es remitido hacia atrás de la interfaz RPF.

### 5.5.2 Comprobación de RPF

- La tabla de enrutamiento usada para Multicast es verificada respecto a la dirección "fuente" en el datagrama de Multicast.
- Si el datagrama llega en la interfaz especificada en la tabla de enrutamiento para la dirección fuente: entonces sucede la comprobación de RPF.
- De otra manera, la comprobación de RPF falla.
- La comprobación satisfactoria del RPF permite que el datagrama sea avanzado.
  - El datagrama es remitido hacia todas las interfaces salientes, pero no a la interfaz RPF en la que el datagrama fue recibido, ver figura 5-13.
- El fracaso de la comprobación del RPF causa que el datagrama sea tirado, como lo ilustra la figura 5-14.
- La fuente origen inunda la red con datos de Multicast
- Cada Router tiene indicada una interfaz entrante (interfaz RPF) en la cual los datos de Multicast se pueden recibir de una determinada fuente.
- Cada Router recibe datos de Multicast en una o más interfaces, pero realiza la comprobación del RPF para prevenir la expedición duplicada.

#### Comprobación RPF Exitosa

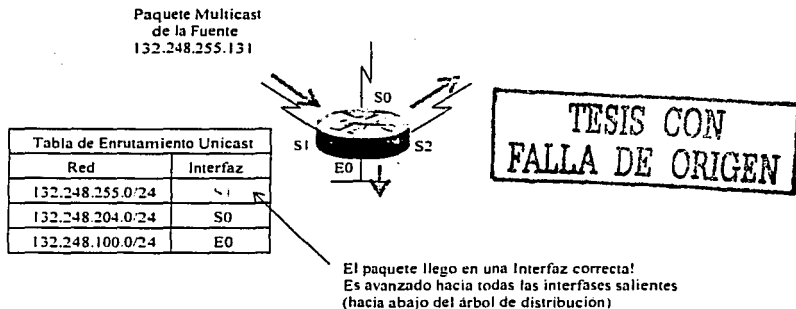


Figura 5-13 Comprobación RPF exitosa.

Comprobación RPF Fallida

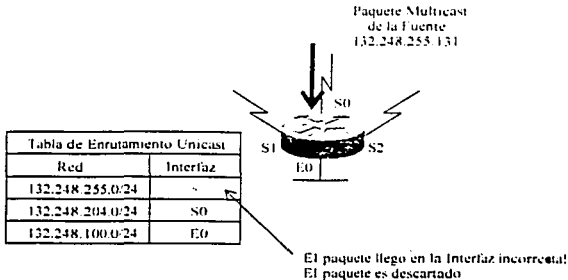


Figura 5-14 Comprobación RPF fallida.

5.5.3 Tipos de Protocolos Multicast

- Protocolos Dense Mode.
- Protocolos Sparse Mode.

Características protocolos Dense-Mode

- Emplean el Modelo "Empuje" (Push).
- Inundación de tráfico a través de la red.
- Podado del árbol multicast para recortar el tráfico hacia los lugares en los que no es deseado.
- Funcionamiento de Flood y Prune (típicamente cada 3 minutos).

Características protocolos Sparse-Mode

- Emplean el Modelo de "Jalar" (Pull).
- El tráfico es enviado solamente en donde es solicitado.
- Funcionamiento explícito de Ingreso.

Protocolos Dense Mode

- DVMRP - Distance Vector Multicast Routing Protocol.
- MOSPF - Multicast OSPF.
- PIM DM - Protocolo Independiente Multicasting (Dense Mode).

TESIS CON FALLA DE ORIGEN

*Protocolos Sparse Mode*

- PIM SM- Protocol Independent Multicasting (Sparse Mode).
- CBT - Core Based Trees.

**5.5.4 Topologías Sparse vs. Dense**

Una topología densamente poblada "Dense" es aquella en la cuál hay muchos miembros del grupo de Multicast en relación al número total de hosts en una red interna. Las topologías escasamente pobladas "Sparse" tienen pocos miembros de grupo en relación al número total de hosts. El termino sparse no significa que hay pocos hosts. Una topología sparse puede significar que hay 2.000 miembros de un grupo, por ejemplo, dispersos entre 100.000 hosts totales.

Ninguna proporción numérica específica diseña las topologías sparse y dense. Sin embargo, es seguro decir, que las topologías dense son usualmente encontradas en LANs switcheadas y ambientes de campus, y las topologías sparse usualmente involucran WAN's. Lo que es importante es que los protocolos de enrutamiento de Multicast son diseñados para trabajar mejor en una o la otra topología y son designados como protocolos "dense mode" o protocolos "sparse mode" como lo muestra la tabla 5-1.

Protocolo	Dense Mode	Sparse Mode
DVMRP	X	
MOSPF	X	
PIM-DM	X	
PIM-SM		X
CBT		X

Tabla 5-1 Protocolos de enrutamiento Multicast Dense Mode y Sparse Mode.

**TESIS CON  
FALLA DE ORIGEN**

### 5.5.5 Diferencias entre el enrutamiento Unicast y Multicast

- El enrutamiento Multicast funciona de forma inversa al enrutamiento Unicast.
- El enrutamiento Unicast está relacionado con el envío del paquete basado en el destino.
- El enrutamiento Multicast está relacionado con la procedencia del paquete, es decir, está basado en el origen.

La función del protocolo de enrutamiento Unicast es encontrar el camino más corto a un destino en particular. La interfaz por la que un paquete de Unicast es avanzado, es conocida como el próximo salto "next-hop", denominada interfaz downstream desde la perspectiva del protocolo de enrutamiento Unicast, esta interfaz es la más cercana al destino.

En contraste, la función del protocolo de enrutamiento Multicast es determinar la interfaz upstream (interfaz más cercana a la fuente). A causa de que los protocolos de enrutamiento Multicast se involucran con el camino más corto al origen, en lugar del camino más corto al destino, el procedimiento de avanzar paquetes de Multicast es conocido como "*Reverse Path Forwarding*".

La forma más fácil de que un protocolo de enrutamiento Multicast determine el camino más corto al origen es consultar la tabla de enrutamiento Unicast. Sin embargo, los paquetes de Multicast son avanzados basados en la información en una tabla de enrutamiento separada. La razón para esto, es que el router no solamente debe grabar la interfaz upstream para el origen de un particular par (S, G), sino también las interfaces asociadas con el grupo.

Así la función de un protocolo de enrutamiento Multicast es determinar las interfaces downstream reales asociadas con un par (S, G). Cuando todos los routers han determinado sus interfaces upstream y downstream para un particular origen y grupo, se dice que se ha establecido un árbol de Multicast, ver figura 5-15. La raíz del árbol es el router directamente conectado a la fuente, y las ramas conducen a todas las subredes en las cuales residen los miembros del grupo. Ninguna rama conduce a subredes "vacías" (subredes con

ningún miembro del grupo asociado). El avance de paquetes por interfasés que solamente conducen a miembros del grupo es llamado "Reverse Path Multicast" (RPM).

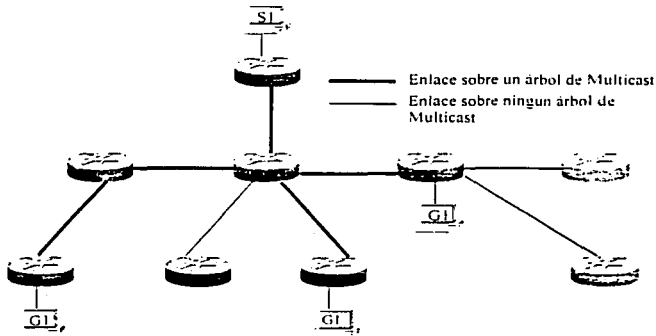


Figura 5-15 Los caminos principales desde la fuente de Multicast a todas las subredes en donde residen miembros del grupo, forman el árbol de Multicast.

Los árboles de Multicast solamente subsisten mientras existe la sesión de Multicast. Y debido a que los miembros pueden ingresar y abandonar a grupos desde el principio y hasta el fin de la sesión, la estructura es dinámica. La tercera función de un protocolo de enrutamiento Multicast es administrar el árbol, "insertando" (grafting) ramas conforme los miembros ingresen al grupo y "recortando" (pruning) ramas conforme los miembros abandonen el grupo.

TESIS CON  
FALLA DE ORIGEN

## 5.6 Protocolo de enrutamiento PIM-Sparse Mode

Protocolo independiente de Multicast (PIM) Sparse-mode (RFC 2362). PIM es un protocolo-independiente de enrutamiento IP y puede influenciar a cualquiera de los protocolos de enrutamiento Unicast que son usados para poblar la tabla de enrutamiento Unicast. PIM emplea esta información de enrutamiento Unicast para realizar la función de Multicast Forwarding. Aunque PIM es llamado protocolo de enrutamiento Multicast, realmente utiliza la tabla de enrutamiento Unicast para realizar la función de comprobación del RPF, en lugar de construir una tabla de enrutamiento Multicast completamente independiente. A diferencia de otros protocolos de enrutamiento, PIM no envía ni recibe actualizaciones de enrutamiento entre los routers.

### 5.6.1 Comprobación de RPF en PIM-SM

La comprobación del RPF depende del tipo de árbol de distribución Multicast.

Si el tráfico esta fluyendo hacia abajo del árbol de distribución compartido, el mecanismo de comprobación del RPF empleará la dirección IP del RP para realizar la comprobación RPF.

Si el tráfico esta fluyendo hacia abajo del árbol de distribución más cortó, el mecanismo de comprobación del RPF empleará la dirección IP de la Fuente para realizar la comprobación RPF.

### 5.6.2 PIM-SM uniendo los Árboles Compartidos

En la figura 5-16, hay un receptor activo, el cual ingreso al grupo Multicast "G". El router conoce la dirección IP del RP para el grupo "G", así que envía un par (\*, G) "Join" hacia el RP. Este par (\*, G) "Join" viaja salto-por-salto hacia el RP, de tal forma que construye una rama del Árbol Compartido que se extiende desde el RP hasta el router "ultimo salto" directamente conectado al receptor. En este punto, el tráfico del grupo "G" puede fluir hacia abajo del Árbol Compartido hasta llegar al receptor.

TESIS CON  
FALLA DE ORIGEN



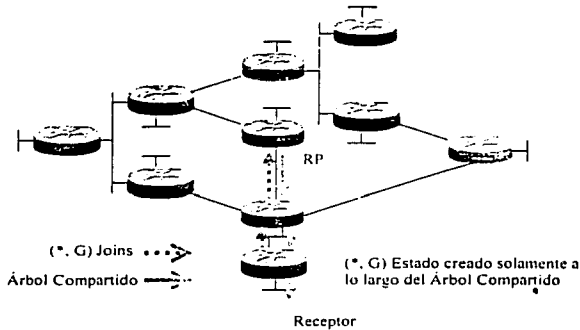


Figura 5-16 PIM-SM Shared Tree Joins.

### 5.6.3 Registro del Remitente

Tan pronto como una fuente activa para el grupo "G" envía un paquete al router al que se encuentra conectado, éste es responsable de registrar (registering) a esta fuente con el RP y solicitar que el RP construya un árbol hacia ella, ver figura 5-17.

El router fuente encapsula los datos Multicast en un mensaje especial de PIM-SM. Llamado "Registro" (Register) y los envía al RP en forma Unicast.

Cuando el RP recibe el mensaje de "Registro" hace dos cosas:

- Desencapsula el paquete de datos Multicast dentro del mensaje de Registro y lo envía hacia el árbol compartido.
- También el RP envía un (S, G) "Join" hacia la red fuente "S" para crear una rama de un (S, G) Árbol de Trayectoria más Corta. Esto resulta en un estado (S, G) que es creado en todos los routers a lo largo del SPT, incluyendo al RP.

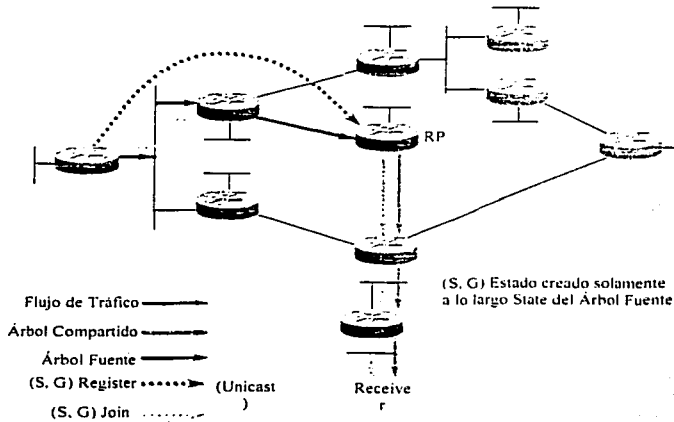


Figura 5-17 Proceso de Registro del Remitente

Tan pronto como es construido el SPT desde el router fuente hasta el RP, el tráfico de Multicast comienza a fluir nativamente desde la fuente "S" hasta el RP. Una vez que el RP empieza a recibir datos nativamente (es decir hacia abajo del SPT) desde la fuente "S" envía un mensaje "Register Stop" al router *first-hop* de la fuente para informarle que puede parar de enviar los mensajes Unicast de Registro "Register".

#### 5.6.4 Mecanismos de PIM-SM

PIM-SM emplea los siguientes mecanismos:

- PIM Descubrimiento de Vecinos (Neighbor Discovery)
- PIM Estados (State)
- PIM SM Ingresos (Joining)
- PIM SM Registros (Registering)
- PIM SM SPT-Switchover
- PIM SM Podas (Pruning)
- PIM SM Maquina de estados (State Maintenance)

TESIS CON  
 FALLA DE ORIGEN

#### 5.6.4.1 Descubrimiento de Vecinos

Mensajes PIM "Hello" son enviados periódicamente para descubrir la existencia de otros routers PIM en la red para elegir al router designado DR. En redes Multi-Acceso (tales como Ethernet) los mensajes PIM "Hello" son enviados en forma Multicast a la dirección de grupo 224.0.0.13 "All-PIM-Routers".

##### *Router Designado (DR)*

En redes Multi-Acceso, un Router Designado es elegido. en redes PIM SM, el DR es responsable de enviar mensajes "Join" al RP para los miembros en la red multiacceso y enviar mensajes "Register" a los RP para las fuentes en la red multiacceso. En PIM DM, el DR no tiene ningún significado. La excepción a esto es cuando se emplea IGMPv1. En este caso, el DR también funciona como el IGMP Querier para la red Multi-Acceso.

##### *Elección del Router Designado (DR)*

Para elegir al DR, cada nodo PIM en una red multiacceso examina el mensaje PIM "Hello" recibido de sus vecinos y compara la dirección IP de su interfaz con la dirección IP de sus vecinos PIM. El vecino PIM con la dirección IP más alta es elegido el DR, como lo muestra la figura 5-18.

Si después de algún periodo de tiempo (configurable), no se han recibido mensajes PIM "Hello" desde el DR elegido, nuevamente se ejecuta el mecanismo para la elección de un nuevo DR.

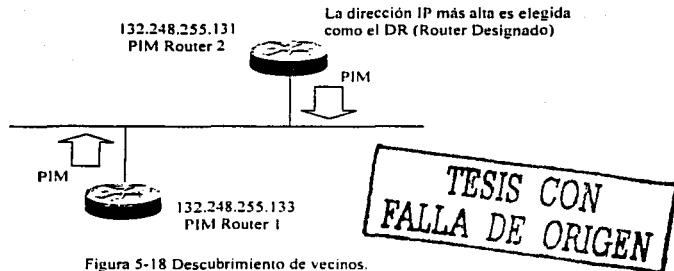


Figura 5-18 Descubrimiento de vecinos.

### *Encontrando al Rendezvous Point (RP)*

Un árbol compartido es enraizado en un router en alguna parte de la red de Multicast en lugar de la fuente. PIM-SM llama a este router "*Rendezvous Point*" (RP). Antes de que un árbol compartido pueda ser establecido, los routers que van a ingresar deben de saber como encontrar al RP. El router puede aprender la dirección del RP en tres formas:

- La dirección del RP puede ser estáticamente configurada en todos los routers.
- Un protocolo estándar y abierto de bootstrap puede ser usado para designar y anunciar el RP.
- El protocolo Auto-RP propietario de Cisco puede ser usado para designar y anunciar al RP.

### **5.6.4.2 Estado de PIM**

En general, el Estado Multicast "*Multicast State*" describe la forma en la que el router entiende al árbol de distribución Multicast en un punto en la red.

Sin embargo a decir verdad, "*Multicast State*" describe el estado de "*forwarding*" que el router emplea para remitir el tráfico Multicast.

### *Tabla de Enrutamiento Multicast (mroute)*

El "*state*" Multicast es almacenado en la tabla de enrutamiento Multicast (*mroute*). Las entradas en la tabla se componen de entradas (\*, G) y (S, G), las cuales contienen:

Información RPF que consiste de una interfaz entrante (o RPF) y la dirección IP del router vecino RPF en la dirección de la fuente.

Outgoing Interface List (OIL), la cual contiene una lista de interfaces hacia las cuales debe ser remitido el tráfico de Multicast. (El tráfico Multicast se debe recibir en la interfaz entrante antes de que sea remitido hacia fuera de las interfaces). Si el tráfico Multicast no se recibe en una interfaz entrante, simplemente es desechado.

TESIS CON  
FALLA DE ORIGEN

### 5.6.4.3 Ingresos

Cuando un router "last-hop" desea empezar a recibir tráfico de Multicast para el grupo "G", envía un mensaje de ingreso "(\*, G) Join" a su vecino PIM "up-stream" en la dirección del RP.

Mientras el mensaje "Join" es enviado en forma Multicast a la dirección de Multicast 224.0.0.2 "All-Routers", la dirección IP de los vecinos PIM "up-stream" es indicada en el cuerpo del mensaje PIM "Join". Esto permite que todos los routers PIM en una red Multi-Acceso estén enterados, pero solamente el vecino PIM "up-stream" indicado realizará el Join.

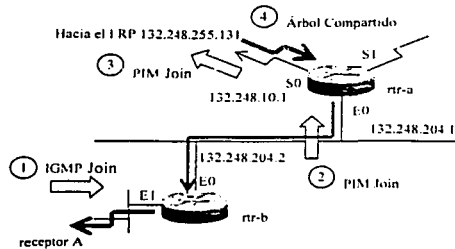
Cuando un router PIM recibe un mensaje de ingreso "(\*, G) Join" para el grupo "G" desde uno de sus vecinos PIM "down-stream", comprobará si existe algún "(\*, G) state" para el grupo "G" en su tabla de enrutamiento Multicast.

- Si ya existe un "(\*, G) state" para el grupo "G", entonces la interfaz en la que el mensaje de ingreso fue recibido es colocada en la "oolist" "(\*, G)" y un mensaje de ingreso "(\*, G)" será enviado hacia el RP.
- Si no existe un "(\*, G) state" para el grupo "G", se crea una entrada "(\*, G)", la interfaz desde la cual el mensaje de ingreso fue recibido es colocada en la "(\*, G) oolist" y un mensaje de ingreso "(\*, G)" es enviado hacia el RP.

El resultado final del mecanismo, es crear un "(\*, G) state" para todas las trayectorias desde el router last-hop al RP de modo que el tráfico de Multicast para el grupo "G" fluya hacia la parte baja del Árbol Compartido (RPT) hasta el router "last-hop".

Este mecanismo de ingreso es ilustrado en la figura 5-19.

TESIS CON  
FALLA DE ORIGEN



- ① El "receptor A" desea recibir tráfico del grupo "G". Envía un mensaje IGMP Join para "G".
- ② El "rtr-b" envía un mensaje (\*, G) Join hacia el RP.
- ③ El "rtr-a" envía un mensaje (\*, G) Join hacia el RP.
- ④ El Árbol Compartido es construido por todas las trayectorias hasta el RP.

Figura 5-19 Proceso de Ingreso a un grupo Multicast PIM-SM.

TESIS CON FALLA DE ORIGEN

#### 5.6.4.4 Registros

Todos los Remitentes No Necesariamente son Receptores y Viceversa. No es requisito que todas las fuentes sean receptores. En el caso de que un host sea solamente una fuente, para el host es permisible que simplemente empiece a enviar tráfico Multicast sin nunca ingresar al grupo vía IGMP.

*El Router "1st-hop" envía Mensajes "Register" al RP*

En PIM Sparse Mode, cuando un router "1st-hop" recibe el primer paquete de Multicast desde la fuente directamente conectada "S" para el grupo "G", crea un (\*, G) "state" y fija el bit "F" en la entrada (\*, G) para indicar que es una fuente directamente conectada "Source" y también configura la bandera "Registering" para indicar que está en el proceso de "Registering".

A continuación, el router "1st-hop" encapsula el paquete Multicast original en un mensaje PIM "Register" y lo envía en forma Unicast al RP. (Cualquier paquete subsecuente de Multicast recibido de la fuente directamente conectada "S" para el grupo "G", también será

encapsulado y enviado en forma Unicast al RP. Este proceso continuará hasta que un mensaje "Register-Stop" sea recibido desde el RP.)

*El RP recibe los mensajes "Register"*

Cuando un RP recibe un mensaje "Register", lo desencapsula. Si este paquete es para un grupo para el cual el RP tiene un (\*, G) "state", el RP hará lo siguiente:

- Envía el paquete original hacia todas las interfases salientes en la "oilist" de entrada (\*, G).
- Si no es así, el RP crea un (\*, G) "state" y envía un mensaje (\*, G) "Join" hacia la fuente para unir al Shortest-path Tree (SPT) con la fuente "S".

*El Router "1st-hop" recibe el mensaje de ingreso (S, G)*

Cuando el router "1st-hop" recibe el mensaje (\*, G) "Join" (enviado salto por salto desde el RP), lo procesa normalmente agregando la interfaz desde la cual el mensaje "Join" fue recibido para la "oilist" de la entrada (\*, G) existente. (Esta entrada fue creada originalmente cuando el router "1st-hop" recibió el primer paquete de Multicast proveniente de la fuente directamente conectada "S".) Esto completa la construcción del Shortest-Path Tree (SPT) desde la fuente "S" hasta el RP.

Ahora el router "1st-hop" comienza a enviar el tráfico de Multicast de la fuente "S" hacia la parte baja del recién construido Shortest-Path Tree (SPT) hasta el RP.

Nota: el tráfico (S, G) fluye temporalmente hasta el RP vía dos métodos: vía mensajes "Register" (hasta que un mensaje "Register-Stop" es recibido) y el "nativo" Shortest-Path Tree (SPT).

*El RP empieza a recibir el tráfico hacia la parte baja del (S, G) SPT.*

Tan pronto como el RP empieza a recibir tráfico (S, G) "nativamente" (es decir, cuando el mensaje "Register" no se ha encapsulado) hacia la parte baja del SPT, el RP configurará el bit "T" en la entrada (S, G) para denotar que el tráfico está fluyendo hacia el Shortest-Path Tree (SPT).

Cuando algún mensaje "Register" (S, G) sea recibido por el RP, él considerará que el bit "T" está configurado en la entrada (S, G) y responderá enviando un mensaje "Register-Stop" PIM (S, G) al router "1st-hop". Esto notificará al router "1st-hop" que el tráfico está siendo recibido "nativamente" en el SPT.

*El Router "1st-hop" recibe el mensaje "Register-Stop"*

Cuando el mensaje (S, G) "Register-Stop" es recibido por el router "1st-hop", limpia la bandera "Registering" en la entrada (S, G) y deja de encapsular en mensajes "Register" al tráfico (S, G). Ahora el tráfico solamente está fluyendo de abajo del SPT al RP.

**5.6.4.5 SPT-Switchover**

*Umbral-SPT*

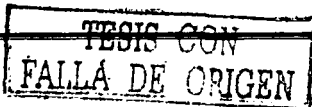
Para los routers "last-hop" (routers con miembros directamente conectados) PIM-SM tiene la capacidad de switchear el Árbol de Trayectoria Más Corta "Shortest-Path Tree" y puentear al RP si la tasa de tráfico sobrepasa el umbral configurado, llamado "SPT-Threshold".

El valor del "SPT-Threshold" indica a los routers que están directamente conectados a miembros de Multicast, el tiempo que deben de esperar para unir el SPT a la fuente, después de que el primer paquete llegue vía él (\*, G) árbol compartido.

En PIM Sparse Mode, los umbrales del SPT pueden ser configurados para controlar el switcheo al Shortest-Path Tree (SPT).

Los Umbrales-SPT son especificados en Kbps y pueden ser usados con listas de acceso para especificar los grupos a los que se les aplicaran.

El Umbral-SPT default es 0 Kbps. Esto significa que cualquiera y todas las fuentes son switcheadas inmediatamente al Shortest-Path Tree.





Si un Umbral-SPT es especificado para un grupo, las fuentes no serán switcheadas al Shortest-Path Tree (SPT) y permanecerán en el Árbol Compartido.

En la Figura 5-20, el router "last-hop" envía un mensaje (S, G) "Join" hacia la fuente para unir el SPT y puentear el RP.

Este mensaje (S, G) Join viaja salto-por-salto hasta el router first-hop (el router que esta directamente conectado a la fuente) de tal forma que crea otra rama del SPT.

Finalmente, mensajes especiales (S, G) RP-bit Prune son enviados hacia la parte alta del Árbol Compartido para recortar "prune" el tráfico (S, G) del Árbol Compartido. Si esta función no fuera realizada, el tráfico (S, G) continuaría fluyendo hacia la parte baja del Árbol Compartido, dando como resultado que paquetes duplicados (S, G) lleguen al receptor.

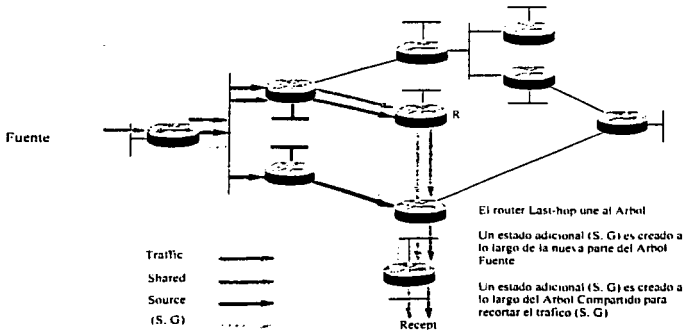


Figura 5-20 Intercambio del Árbol de Trayectoria Más Corta.

*Excediendo el Umbral*

Cuando el Umbral-SPT del Grupo es excedido en un router last-hop, el siguiente paquete recibido para el grupo hará que un mensaje (S, G) "Join" sea enviado hacia el origen del paquete. Esto construye un Shortest-Path Tree desde la fuente "S" hasta el router last-hop.

### Ventaja

Switcheando al Shortest-Path Tree (SPT). Usualmente la trayectoria más corta se utiliza para entregar el tráfico de Multicast. Dependiendo de la localización de la fuente con relación al RP, este switcheo al SPT puede reducir sustancialmente la latencia en la red.

### Desventaja

En redes con un gran número de remitentes, una cantidad creciente de estado latente debe mantenerse en los routers. En algunos casos, un umbral de infinito se puede emplear para forzar a ciertos grupos a permanecer en el Árbol Compartido cuando la latencia no es un problema.

### Umbral-SPT Myth

En este mecanismo, la tasa agregada total del tráfico de grupo que fluye hacia el *Shared Tree (RPT)* es calculado una vez por segundo. Si esta tasa agregada total es excedida, entonces el siguiente paquete de Grupo recibido causará que la fuente sea switchheada al Shortest-Path Tree (SPT).

### Mecanismo SPT-Switchover

Una vez cada segundo, la tasa agregada de tráfico es procesada y comparada con el Umbral-SPT. Si la tasa agregada total del tráfico de todos los grupos que fluye hacia el Shared Tree (RPT) excede el umbral, entonces la bandera "J" es configurada en la entrada (\*, G).

Debido a que cada paquete de Multicast es recibido en el Shared Tree, el bit "J" es revisado en la entrada (\*, G).

- Si la bandera "J" esta configurada, se crea una nueva entrada (\*, G) para el origen del paquete.
- Se envía un mensaje (S, G) "Join" hacia la fuente para ingresar al SPT.
- Se configura la bandera "J" en la entrada (S, G) para denotar que esta entrada fue creada como resultado del SPT-Threshold switchover.
- Se reinicia la bandera "J" en la entrada (\*, G).

TESIS CON  
FALLA DE ORIGEN

#### 5.6.4.6 Podas

El host localmente conectado envía un mensaje "Leave" de IGMP (o mensajes IGMP "state times out" en el router) para el grupo "G".

La interfaz es removida de la entrada (\*, G) y de todas las entradas (S, G) en la tabla de Enrutamiento Multicast.

- Si ahora la "Outgoing Interface list" (\*, G) no tiene efecto, entonces se envía un mensaje (\*, G) "Prune up" por el Shared Tree (RPT) hacia el RP.
- Cualquier entrada restante (S, G) es permitida para hacer que expire su tiempo y después sea borrada de la Tabla de Enrutamiento Multicast.

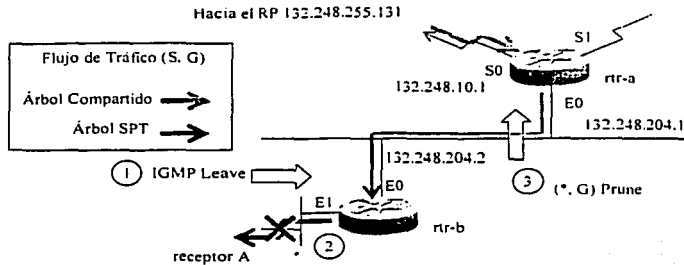
Cuando los routers de la parte alta del Shared Tree reciben el mensaje (\*, G) "Prune" (poda), remueven la interfaz en la cual el mensaje "Prune" fue recibido desde su (\*, G) "Outgoing interface list".

- Si como resultado de remover la interfaz, la (\*, G) "Outgoing Interface list" se convierte en inválida, entonces se envía un mensaje (\*, G) "Prune" por el Shared Tree (RPT) hacia el RP.
- Cualquier entrada (S, G) restante es permitida para hacer que expire su tiempo y después sea eliminada de la Tabla de Enrutamiento Multicast.

El proceso de "Pruning" para el caso de un Árbol Compartido es ilustrado en la figura 5.21.

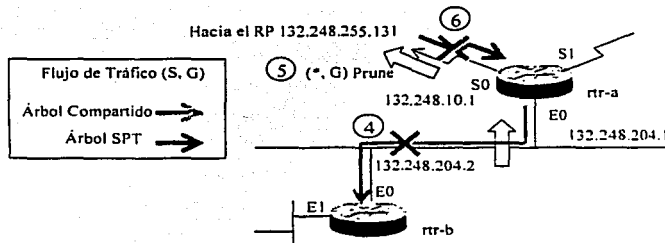
TESIS CON  
FALLA DE ORIGEN

a)



- 1 El "rtr-b" es un Router Leaf. El último host "receptor A", abandona el grupo "G".
- 2 El "rtr-b" elimina la entrada (\*, G) del E1 y también elimina algunas entradas (S, G) "oijist".
- 3 En el "rtr-b" la (S, G) "oijist" esta vacía, y envía un mensaje (\*, G) "Prune" hacia el RP.

b)



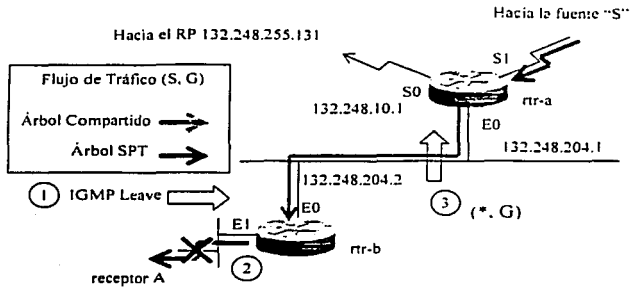
- 4 El "rtr-a" recibe el "Prune", elimina la (S, G) "oijist" de su interfaz E0.
- 5 La (S, G) "oijist" en el "rtr-a" esta vacía, y envía un mensaje (\*, G) "Prune" hacia el RP.
- 6 El "Pruning" continúa hacia el RP.

TESIS CON  
 FALLA DE ORIGEN

Figura 5-21 Proceso "Pruning" en el caso de un Árbol Compartido.

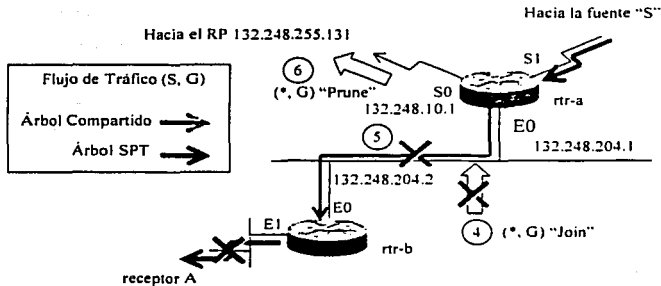
El proceso de "Pruning" (Poda) para el caso de un SPT es ilustrado en la Figura 5.22.

a)



- ① El "rtr-b" es un Router Leaf. El último host "receptor A", abandona el grupo "G".
- ② El "rtr-b" elimina la entrada (\*, G) de la E1 y cualquier entrada (S, G) "oolist".
- ③ En el "rtr-b" la (S, G) "oolist" esta vacía y envía un mensaje (\*, G) "Prune" hacia el RP.

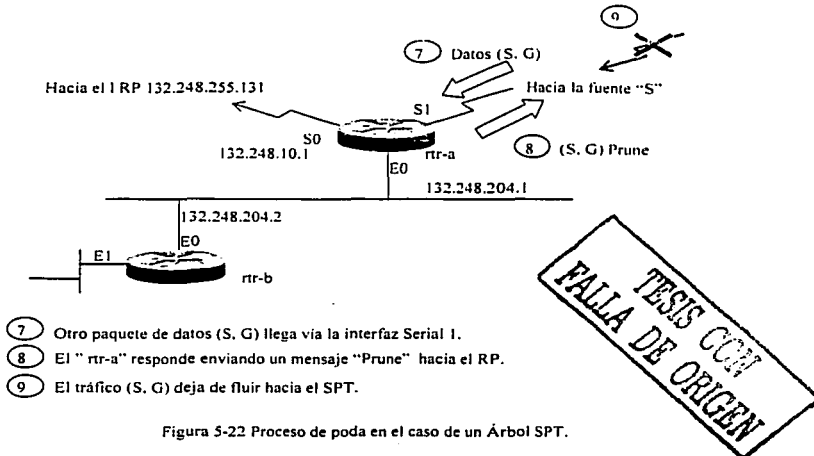
b)



- ④ El "rtr-b" deja de enviar los mensajes periódicos (S, G) "Join".
- ⑤ El "rtr-a" recibe el mensaje "Prune" y elimina la (\*, G) "oolist" de la interfaz E0
- ⑥ La (\*, G) "oolist" en el "rtr-a" esta vacía y envía un mensaje (\*, G) "Prune" hacia el RP.

TESIS CON  
 FALTA DE ORIGEN

c)



- 7) Otro paquete de datos (S, G) llega vía la interfaz Serial 1.
- 8) El "rtr-a" responde enviando un mensaje "Prune" hacia el RP.
- 9) El tráfico (S, G) deja de fluir hacia el SPT.

Figura 5-22 Proceso de poda en el caso de un Árbol SPT.

### 5.6.4.7 Máquina de estados

En PIM SM, la información del estado Join/Prune tiene un tiempo de expiración normal de 3 minutos. Si un mensaje Join/Prune no es recibido en forma periódica para refrescar esta información de estado, el estado expira y automáticamente es borrado. Por lo tanto, un router PIM envía mensajes Join/Prune en forma periódica a sus vecinos para mantener esta información de estado.

Cuando se recibe un mensaje Join desde un Vecino PIM, el contador de tiempo de la interfaz (en la Outgoing Interface List) en la cual se recibe el mensaje se reajusta a tres minutos. Si el contador de tiempo "expiración" para la interfaz llega a cero, la interfaz se elimina de la "Outgoing interface list". (Esto puede originar un mensaje "Prune" si la interfaz eliminada hace que la "Outgoing interface list" se convierta en invalida.)

Cuando se recibe un mensaje "Prune" en PIM Sparse Mode, la interfaz en la que se recibió el mensaje, normalmente es removida de la "Outgoing interface list". La excepción a esto, es el caso especial del mensaje (S, G) "RP -bit Prunes", el cual es usado para recortar el tráfico (S, G) del Árbol Compartido. En este caso, mensajes (S, G) "RP -bit Prunes" deben ser enviados periódicamente para mantener el estado "Prune" en el vecino PIM "upstream" hacia el RP.

Todas las entradas (S, G) tienen un contador de tiempo para expirar, el cual es reajustado a tres minutos por la llegada de un paquete (S, G) recibido vía el Shortest-Path Tree (SPT). Si la fuente deja de enviar paquetes, el contador de tiempo alcanza el valor de cero y la entrada (S, G) es borrada.

### 5.7 MBGP Extensiones para Multicast

El Multiprotocolo BGP (MBGP) esta definido en el RFC 2283. Define las extensiones para el protocolo existente BGP, para permitir transportar algo más que prefijos de rutas de IPv4. Algunos ejemplos de los nuevos tipos de información de enrutamiento son:

- Prefijos IPv4 para enrutamiento Unicast.
- Prefijos IPv4 para la comprobación del RPF Multicast.
- Prefijos IPv6 para enrutamiento Unicast.

Una idea común falsa es que MBGP es un reemplazo para PIM. Esto es incorrecto. MBGP no propaga *ninguna* información del estado Multicast, ni construye ninguna clase de árbol de distribución Multicast. MBGP *puede* distribuir los prefijos Unicast que se pueden utilizar para la comprobación RPF Multicast.

Debido a que MBGP es una extensión para el protocolo BGP existente, se aplican las mismas reglas básicas para la selección de la trayectoria, validación de trayectoria, etc.

#### 5.7.1 Bases de Información de Enrutamiento

Previamente, BGP mantenía solamente una sola Routing Information Base (RIB) para los prefijos Unicast de IPv4. En el caso de MBGP, para cada tipo de información de

enrutamiento que esta siendo intercambiada se debe de mantener una RIB. Esto implica que una Unicast RIB (U-RIB) y una Multicast RIB (M-RIB) pueden ser mantenidas de forma separada por MBGP.

*Unicast RIB (U-RIB)*

Esta RIB contiene los prefijos Unicast que previamente fueron usados por BGP para avanzar el tráfico Unicast de IPv4.

*Multicast RIB (M-RIB)*

Esta nueva RIB contiene el mismo tipo de prefijos que posee la U-RIB, excepto que los prefijos almacenados en la M-RIB son usados para la comprobación RPF del tráfico de Multicast que se esta recibiendo.

**5.7.2 Mensaje de Actualización de MBGP**

En la figura 5-23, se puede observar que el mensaje de actualización de MBGP es idéntico al mensaje de actualización de BGP, con la diferencia de que dos nuevos tipos de atributos han sido agregados. Estos dos atributos son:

- MP\_REACH\_NLRI
- MP\_UNREACH\_NLRI

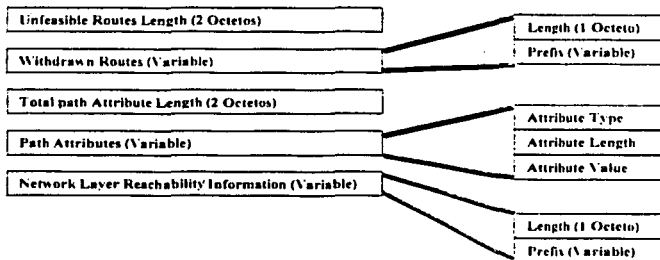


Figura 5-23 Formato del mensaje de actualización de MBGP.



### 5.7.2.1 Atributo MP\_REACH\_NLRI

Las principales características de este Nuevo atributo son los campos de "Address Family Identifier" y "Sub-Address Family Identifier" como lo ilustra la figura 5-24. Estos dos campos definen el tipo de información de enrutamiento que es transportada en el campo "NLRI" de este atributo.

- *Address Family Information (AFI)*

Este campo esta basado en las familias de direcciones definidas en el RFC1700.

AFI = 1 (IPv4)

AFI = 2 (IPv6)

- *Sub-Address Family Information (Sub-AFI)*

Este campo contiene la información adicional con respecto al tipo de información de enrutamiento que esta siendo intercambiada en el campo NLRI. A continuación se mencionan las definiciones actuales para los códigos secundarios Sub-AFI asociados con la Familia de Direcciones IPv4.

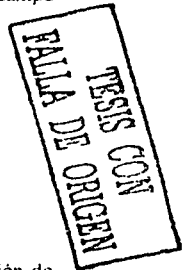
Sub-AFI = 1 (NLRI es empleado para el enrutamiento Unicast).

Sub-AFI = 2 (NLRI es usado para la comprobación RPF Multicast).

Sub-AFI = 3 (NLRI se emplea tanto como para el enrutamiento Unicast como para la comprobación RPF Multicast.)

La información de "Next-Hop Address" esta contenida en el campo siguiente a los campos "AFI" y "Sub-AFI".

Los siguientes campos al campo "Next-Hop Address" son cero o más campos "SNAP". Este campo contiene los atributos asociados con el campo "NLRI". (Para el AFI Ipv4, estos atributos son los mismos atributos de BGP). Finalmente, el campo "NLRI" contiene la información de la Longitud y del Prefijo de la ruta que esta siendo anunciada como alcanzable.



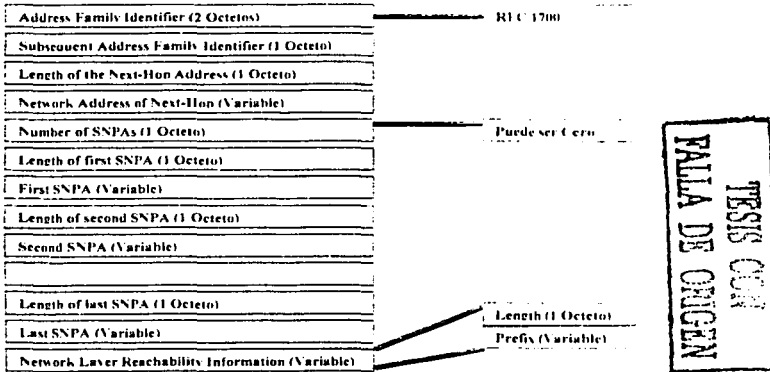


Figura 5-24 Campos del atributo MP\_REACH\_NLRI

### 5.7.2.2 Atributo MP\_UNREACH\_NLRI

Este nuevo atributo permite que las rutas irrealizables de los nuevos tipos de protocolos sean retiradas de la misma manera que lo hace el campo "Withdrawn Routes" usado en BGP.

En la figura 5-25 se puede observar que este atributo también transporta los campos "AFI" y "Sub-AFI" a lo largo de la Longitud y el Prefijo asociados de la ruta aislada.

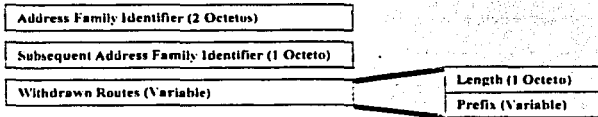


Figura 5-25 Formato de los campos del atributo MP\_UNREACH\_NLRI.

### 5.7.2.3 Capacidad de Negociación MBGP

MBGP ha ampliado el formato del Mensaje "Open" para incluir nuevos parámetros opcionales para la *Capacidad* de negociación.

MBGP enruta la negociación al conjunto de capacidades común más bajo usando estos campos opcionales de *Capacidad*.

Si dos "peers" de MBGP son incapaces de convenir sobre las Capacidades soportadas, la sesión de MBGP es terminada y un mensaje de error es mandado a consola.

Si los peers tienen configurado el mismo conjunto de Capacidades, entonces el NRLI de Unicast y de Multicast puede ser intercambiado a través de la sesión como lo muestra la figura 5-26.

Si no hay "match" entre las Capacidades, el *peering* no será levantado.



Figura 5-26 Sesión de BGP para el NRLI Unicast y Multicast

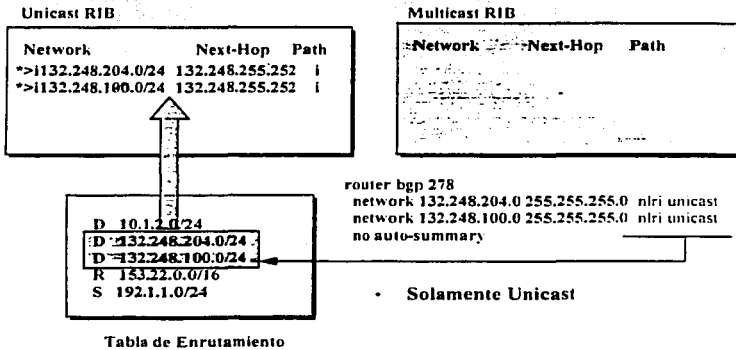
5.7.2.4 Información NLRI

*Controlando la Información en la RIB*

Al igual que el comando "neighbor", el comando "network" también acepta la nueva palabra clave "keyword" "nlri". Esto proporciona el control para la información RIB (U-RIB, M-RIB, o ambas) que es inyectada a la red.

Si se omite la cláusula "nlri", se asume la Unicast RIB (U-RIB).

En la figura 5-27. Solamente se emplea Unicast en él (nlri). Esto hace que las redes coincidan en la Tabla de Rutas local para ser inyectadas dentro de la MBGP Unicast RIB.



TESIS CON  
 FALLA DE ORIGEN

Figura 5-27 Empleando el comando NLRI para controlar la información de la RIB Unicast.

En la figura 5-28. Solamente se emplea Multicast en el (nlri). Esto hace que las redes coincidan en la Tabla de Rutas local para ser inyectadas dentro de la MBGP Multicast RIB.

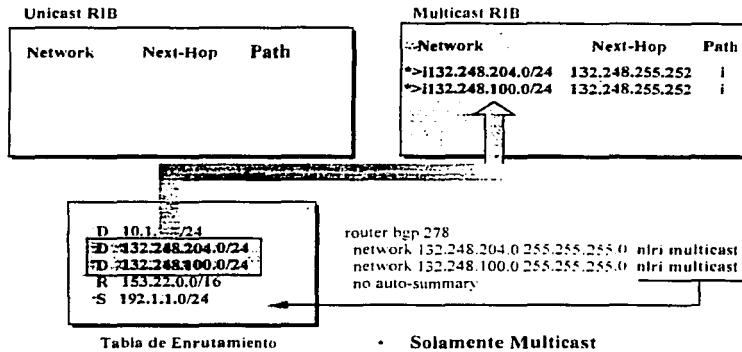
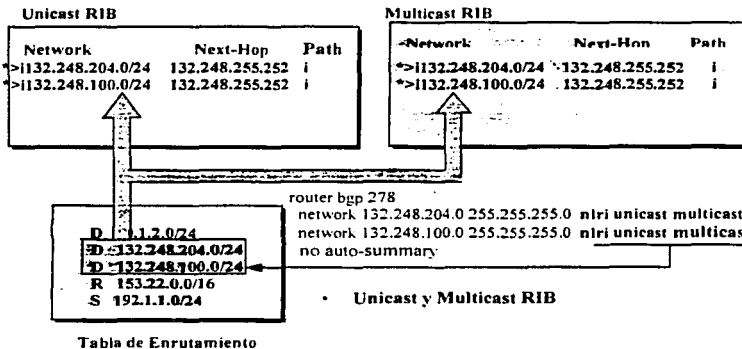


Figura 5-28 Empleando el comando NLRI para controlar la información de la RIB Multicast.

En la figura 5-29. Se emplea Unicast y Multicast en el (nlri). Esto hace que las redes coincidan en la Tabla de Rutas local para ser inyectadas dentro de las MBGP Unicast y Multicast RIBs.



**TESIS CON  
FALLA DE ORIGEN**

Figura 5-29 Empleando el comando NLRI para controlar la información de las RIBs Unicast y Multicast.

*Recibiendo Mensajes de Actualización*

El almacenamiento de la información del NRLI que se recibe, depende de los campos AFI/Sub-AFI en el atributo MP\_REACH\_NLRI.

*Contenido de los Mensajes de Actualización RIB*

La información que se recibe en los mensajes de Actualización se configura en la RIB (Unicast RIB, Multicast RIB o en ambas), dependiendo de los valores de los campos AFI/Sub-AFI.

- AFI/Sub-AFI = 1/1 (IPv4 / Unicast): solamente la U-RIB.
- AFI/Sub-AFI = 1/2 (IPv4 / Multicast): solamente la M-RIB
- AFI/Sub-AFI = 1/3 (IPv4 / Unicast-Multicast): ambas RIBs

**5.7.3 Topologías Congruentes Unicast-Multicast**

Cuando los flujos del tráfico Unicast y Multicast siguen la misma trayectoria, se dice que las dos topologías son congruentes, ver figura 5-30.

En la figura 5-30 ambas sentencias “neighbor” y “network” en el router de la izquierda contienen la cláusula “nrlri unicast-multicast”. Esto hace que el router negocie una sola sesión de BGP sobre la cual el BRLI Unicast-Multicast sea intercambiado.

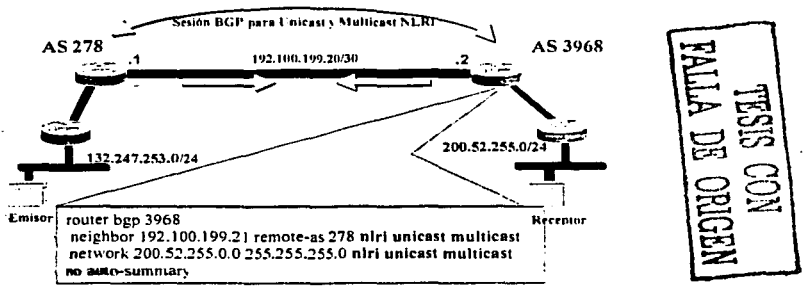


Figura 5-30 Topologías Congruentes.

### 5.7.4 Topologías Incongruentes Unicast-Multicast

En algunos casos es deseable tener una trayectoria para el flujo de tráfico Multicast y otra para Unicast. Cuando esto sucede, se dice que las topologías son "incongruentes".

En la figura 5-31, dos declaraciones "neighbor" separadas son usadas para que el router de la derecha pueda negociar y establecer dos sesiones de BGP con su router vecino de la izquierda.

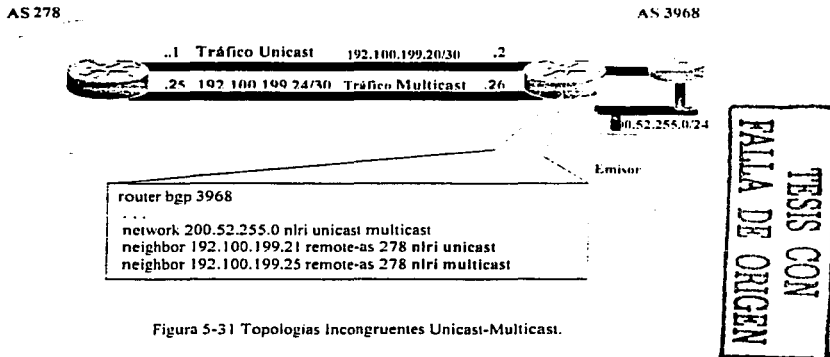


Figura 5-31 Topologías Incongruentes Unicast-Multicast.

### 5.7.5 Conversión NRLI de Unicast a Multicast

En algunos casos de transición, puede ser deseable convertir el Unicast NRLI en Multicast NRLI a favor de un AS que emplea una versión vieja de BGP y que además es incapaz de actualizar a MBGP, para poder recibir el tráfico Multicast.

- Debido a que es un AS BGP stub, para este AS es posible emplear un solo RPF usando una ruta estática o los prefijos Unicast NRLI recibidos desde el AS upstream.

### *Ventajas*

MBGP ofrece las siguientes ventajas:

- Una Internet puede soportar topologías incongruentes de Unicast y Multicast.
- Una Internet puede soportar topologías congruentes de Unicast y Multicast que tengan diferentes políticas (BGP filtering configurations).
- Todas las políticas posibles de BGP se pueden aplicar a MBGP.
- Todas las políticas posibles de enrutamiento de BGP se pueden aplicar a MBGP.
- Todos los comandos de BGP se pueden emplear en MBGP.

### *Restricciones*

- No se pueden conectar nubes de MBGP junto con nubes de BGP. Es decir, no se puede redistribuir rutas de MBGP dentro de BGP

### **5.8 MSDP**

Multicast Source Discovery Protocol (MSDP) es un mecanismo para conectar múltiples dominios PIM-SM. MSDP permite que las fuentes Multicast de un grupo conozcan a todos los rendezvous point(s) (RPs) en diferentes dominios. Cada dominio PIM-SM emplea su propio RP y no necesita depender de RPs en otros dominios. Un RP ejecuta MSDP sobre TCP para descubrir fuentes de Multicast en otros dominios.

Un RP en un dominio PIM-SM tiene una relación de parejas con routers en otros dominios que están ejecutando MSDP. La relación de parejas sucede sobre una conexión TCP, en donde principalmente una lista de fuentes enviadas a los grupos de Multicast es intercambiada. Las conexiones TCP entre los RPs son realizadas entre sistemas autónomos adyacentes. El RP receptor emplea la lista de fuentes para establecer una trayectoria hacia la fuente (source path).

El propósito de esta topología es descubrir de fuentes Multicast en otros dominios. Si las fuentes de Multicast son de interés para un dominio que posee receptores, los datos de Multicast son entregados sobre el árbol-fuente source-tree normal construido por el mecanismo de PIM-SM.

TESIS CON  
FALLA DE ORIGEN



MSDP también es empleado para anunciar las fuentes de un grupo. Estos anuncios deben originarse en el RP del dominio.

MSDP depende excesivamente de MBGP para el funcionamiento interdominio.

### 5.8.1 MSDP Peers

Al igual que BGP, MSDP establece relaciones con sus vecinos (neighbor relationships) de MSDP también conocidos como peers, ilustrado en la figura 5-32.

Los peers MSDP se conectan empleando el puerto 639 de TCP. El peer con la dirección IP más baja toma la función para abrir la conexión TCP. El peer con la dirección IP más alta espera en el estado "Listen" mientras el otro peer realiza la conexión:

- Los peers de MSDP envían mensajes "Keepalive" cada 60 segundos. La recepción de datos realiza la misma función que el "Keepalive" y también se encarga de mantener a la sesión activa.
- Si no hay mensajes "Keepalive" o no se han recibido datos en un periodo de tiempo de 75 segundos, la conexión de TCP se reinicia y nuevamente se reestablece.

Los dispositivos MSDP deben de ejecutar BGP:

- Este requerimiento, es debido al hecho de que el mensaje "Source Active" (SA) para el mecanismo de comprobación RPF, emplea información AS-PATH contenida en la MBGP M-RIB o en la U-RIB.
- Existen algunos casos especiales en donde se elimina la comprobación RPF para el mensaje SA que se recibe. Este caso se da, cuando solamente hay una sola conexión peer MSDP, o cuando se usan grupos de MSDP. En estos casos, no es necesario que los dispositivos MSDP ejecuten BGP.

TESIS CON  
FALLA DE ORIGEN

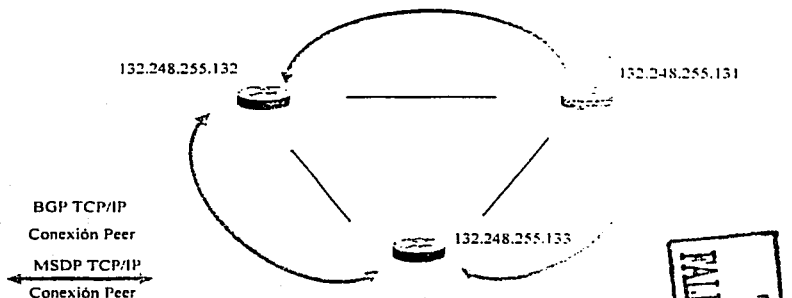


Figura 5-32 Peers MSDP.

### 5.8.2 Mensajes MSDP

Existen cuatro tipos de mensajes básicos en MSDP:

- Keepalives
- Source Active (SA)
- Source Active Request (SA-Req)
- Source Active Response (SA-Resp)

#### *Mensajes Source Active (SA)*

Estos mensajes son usados para anunciar fuentes activas existentes en un dominio. Además, los mensajes SA pueden contener el primer paquete de datos Multicast que fue enviado por la fuente. Transportar este primer paquete de datos en el primer mensaje SA ayuda a tratar el problema de desbordamiento de fuente con bajas tasas de transmisión como son los anuncios de SDR (Session Description Protocol).

Los mensajes SA contienen la dirección IP del RP que los origina, así como unos o más pares (S, G) que están siendo anunciados. Además, el mensaje SA puede contener un paquete de datos encapsulado.

TESIS CON  
 FALLA DE ORIGEN

*Mensajes Source Active Request (SA-Req)*

Estos mensajes son empleados para solicitar una lista de las fuentes activas para un grupo específico. Los mensajes son enviados un "Cache Server SA MSDP", el cual mantiene la lista de pares (S, G) activos.

La latencia de Ingreso se puede reducir empleando esta técnica para solicitar la lista de fuentes activas para un grupo, en lugar de tener que esperar hasta 60 segundos para que sean anunciadas por el RP que las origina.

*Mensajes Source Active Response (SA-Resp)*

Estos mensajes son enviados por el Cache Server SA MSDP en respuesta a un mensaje SA-Req.

Los mensajes SA-Resp contiene la dirección IP del RP que los origina, así como uno o más pares (S, G) de las fuentes activas en el dominio del RP.

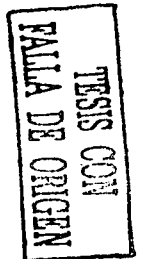
*Mensajes Keepalive*

Estos mensajes son enviados cada 60 segundos para mantener activa la sesión de MSDP. Si en 75 segundos no se han recibido mensajes Keepalive o SA, la sesión MSPD es reiniciada y nuevamente se vuelve a establecer.

**5.8.3 Comprobación RPF del Mensaje SA**

Los mensajes SA deben ser aceptados solamente desde el peer RPF MSDP que esta en la mejor trayectoria hacia el peer que lo origina. El mismo mensaje SA que se recibe desde otros peers debe ser ignorado debido a que si no es así, se pueden presentar loops.

Deterministicamente para seleccionar el peer RPF MSDP de un mensaje SA que se recibe, se debe de conocer la topología de MSDP. Sin embargo, MSDP no distribuye información de su topología. Esto significa que la topología de MSDP se debe deducir de alguna otra forma.



La solución, es utilizar los datos de enrutamiento (m)BGP como la mejor aproximación de la topología de MSDP para realizar el mecanismo de comprobación RPF del SA. Esto tiene las siguientes implicaciones:

- La topología de MSDP debe seguir la misma topología general que la topología del peer BGP.
- Esto significa que con un par de excepciones, generalmente un peer de MSDP también debe ser un peer de m(BGP).

Las reglas para realizar la comprobación RPF dependen del peering de BGP entre los peers de MSDP:

- Regla 1: Esta regla es aplicada cuando el peer de MSDP que envía, también es un peer i(m)BGP.
- Regla 2: Se aplica cuando el peer de MSDP que envía, también es un peer e(m)BGP.
- Regla 3: Es aplicada cuando el peer de MSDP que envía, no es un peer (m)BGP.

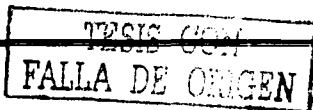
En los siguientes casos la comprobación RPF no es realizada:

- Si el peer MSDP que envía es único Peer MSDP.
- Si el peer MSDP que envía es un peer Mesh-Group.
- Si la dirección del peer MSDP que envía es la misma dirección que contiene el mensaje SA.

#### 5.8.4 MSDP Mesh-Groups

Un MSDP Mesh-Groups puede ser configurado en un grupo de peers MSDP que están completamente conectados en malla. En otras palabras, cada uno de los peers MSDP en el grupo tiene una conexión MSDP a cada uno de los peers MSDP en el grupo.

Cuando se configura un MSDP Mesh-Group entre un grupo de peers, se reduce la inundación de mensajes SA. Esto es, debido a que cuando un peer MSDP en el grupo recibe un mensaje SA de otro peer MSDP en el grupo, puede asumir que este mensaje fue enviado



a todos los otros peer MSDP en el grupo. Consecuentemente, no es necesario ni deseable que el peer de MSDP receptor inunde con mensajes SA a los otros peers MSDP del grupo.

Los MSDP Mesh-Groups también pueden ser usados para eliminar la necesidad de realizar la comprobación (M)BGP RPF de los mensajes que se reciben. Esto es debido a que los mensajes SA nunca son enviados a través de la inundación a los otros peers MSDP en el mesh-group. Consecuentemente, no es necesario realizar la comprobación RPF de los mensajes que se reciben.

### 5.8.5 MSDP SA Caching

Cuando se configura el SA caching, el router empezará a almacenar todos los pares (S, G) que se reciban en los mensajes SA. Esto reduce la latencia de ingreso mientras el RP mantiene una lista de todas las fuentes activas. Por lo tanto, cuando el primer receptor ingresa al grupo, el RP no tiene que esperar al siguiente mensaje SA durante 60 segundos, antes de enviar un mensaje (S, G) Join.

Debido a que los mensajes SA son anunciados periódicamente desde el cache (en lugar de enviarlos tan pronto como se reciban desde un vecino), la propagación de los mensajes SA a través de la red, es más eficiente. Esto evita sobre ejecución de entradas TCP en los peers de MSDP, las cuales dan lugar a inestabilidad en la sesión de MSDP.

#### *Ventajas de SA Caching*

- Reduce la Latencia de Ingreso.
- Es una valiosa herramienta para la resolución de fallas.
- El contenido del cache SA es una valiosa fuente de información de MSDP para la resolución de fallas.
- Ayuda a Prevenir las Tormentas de Mensajes SA.

TESIS CON  
FALLA DE ORIGEN

*Desventajas de SA Caching*

- Consumo de Memoria
- El impacto que se tiene por el hecho de almacenar los mensajes SA en los RPs, generalmente es muy pequeño.

**5.8.6 Ventajas del uso de MSDP**

- Separa el árbol de distribución Multicast compartido. Se puede hacer el árbol compartido local a su dominio. Sus miembros locales unen el árbol local, y el ingreso de mensajes para el árbol compartido nunca tendrá que dejar su dominio.
- Los dominios de PIM-SM solamente pueden confiar en sus propios RPs. Así la confianza decrece sobre RPs en otros dominios. Esto incrementa la seguridad debido a que podemos impedir que nuestras fuentes sean conocidas fueran de su dominio.
- Los dominios con solamente receptores pueden recibir datos sin asociaciones de grupo globalmente anunciados.
- La tabla de enrutamiento global de fuentes Multicast no es requerida, de esta manera se ahorra memoria.

TESIS CON  
FALLA DE ORIGEN

# **Capítulo 6**

## **El protocolo de la siguiente generación de Internet (IPv6)**

TESIS CON  
FALLA DE ORIGEN

---

## Capítulo 6 El protocolo de la siguiente generación de Internet (IPv6)

La versión 6 del protocolo IP esta diseñada para ser el protocolo de la siguiente generación de la red Internet, con esta se pretende solucionar todas las deficiencias que se han comenzado a presentar con la actual versión de IP (versión 4).

Mucha gente se pregunta por que versión 6 y no 5, ¿entonces que paso con la versión 5 del protocolo IP?. La versión 5 del protocolo IP en realidad si existió y es una propuesta de otro protocolo de Internet, la cual es conocida como *Internet Stream Protocol*, definido en el RFC-1190.

### 6.1 Metas de IPv6

El Internet ha tenido un gran y muy difundido éxito, soportando todo tipo de redes: compañías, universidades, centros de investigación, etc. Hoy en día pocas empresas carecen de página WEB y de un servicio de correo electrónico. El acceso a Internet es tan importante como el teléfono. Pero ciertos aspectos de IPv4 han comenzado a limitar el crecimiento global de Internet. Un espacio de direccionamiento de 32 bits, limita el número global de hosts e incluso limita el tamaño jerárquico de Internet.

Para el continuo éxito de Internet también se requiere de incrementar otros requerimientos como son la integridad de los datos, los mecanismos de autenticación y la confiabilidad.

Todas estas deficiencias son las que el protocolo IPv6 promete solucionar, sumadas a otras ventajas adicionales que se mostrarán en este capítulo.

### 6.2 Redes de IPv6

Actualmente existen muchas redes de IPv6, no solo redes experimentales, incluso existen muchos proveedores de servicio que han comenzado a ofrecer el uso comercial de IPv6. A continuación mencionaremos dos de las redes experimentales más importantes en el desarrollo e implementación de IPv6 a nivel mundial.

TESIS CON  
FALLA DE ORIGEN



### 6.2.1 6bone

La red 6bone es usada como un "laboratorio de pruebas" (Testbed) para evaluar las características del protocolo, implementaciones del protocolo, mecanismos de transición de IPv4 a IPv6 y procedimientos de operación.

6bone es mundialmente utilizada para probar redes o productos que utilicen IPv6 antes de ponerlos en producción. Esta red actualmente soporta a más de 260 organismos en 39 países.

### 6.2.2 6REN

6REN es otra de las redes para la investigación y educación, proporciona a las organizaciones una forma de realizar su transición de sus redes IPv4 a IPv6. 6REN es un iniciativa coordinada por voluntarios.

### 6.3 El paquete de IPv6

El paquete de IPv6 contiene direcciones que son jerárquicas y encabezados que permiten a los routers realizar un proceso de enrutamiento de los paquetes de manera más eficiente. El gran tamaño de las direcciones, permite a Internet tener varios niveles de jerarquía y algunos de los campos de IPv4 fueron removidos.

Cuando se hicieron los primeros desarrollos de Internet en 1970 el mundo del Internet era muy diferente. El Internet era usado solo para investigación y educación, un espacio de direcciones de 32 bits fue visto como algo que sería suficiente mientras durará la vida del protocolo IP. El éxito de Internet lo hizo parte integral de muchas de las operaciones que se realizan día a día desde empresas hasta hogares, requiriendo una gran cantidad de direcciones.

#### 6.3.1 Tamaño de las direcciones

Una de las primeras preguntas de IPv6 fue ¿cual debería ser el tamaño adecuado de las direcciones IP?, si era demasiado pequeño limitaría nuevamente el crecimiento y si era demasiado grande se podrían hacer imposibles de administrar y de soportar en los routers.

ISSUE CON  
FALLA DE ORIGEN

Una máscara variable incrementaría la complejidad y alentaría el procesamiento de los paquetes en los routers. Una de las propuestas fue usar direcciones de NSAP (Network Service Acces Point), las cuales varían entre 1 y 20 octetos. Otra, fue la utilización de direcciones de 64-bits. Aunque 64-bits parecían suficientes se agregaron más bits. Previendo la gran complejidad que podría resultar con el crecimiento de Internet y para permitir un futuro crecimiento jerárquico entonces fueron seleccionados 128 bits.

Teóricamente la cantidad de hosts que pueden existir con 128 bits son 340.282.366.920.938.463.463.374.607.431.768.211.456. Si la población de la tierra fuera de 10.000 millones habría entonces  $3.4 \times 10^{27}$  direcciones IP por persona.

### 6.3.2 Representación de las direcciones

Las direcciones de IPv6 tienen 128 bits de longitud, las cuales se dividen en ocho piezas de 16 bits, separadas por dos puntos ":" y cada una de ellas representada por cuatro dígitos hexadecimales. La tabla 6-1 muestra algunos ejemplos de direcciones IPv6.

Formato Expandido	Formato comprimido
0:0:0:0:0:13:1:68:3	::13:1:68:3
0:0:0:0:0:129:144:52:38	::129:144:52:38

Tabla 6-1 Representación de direcciones IPv6.

La representación de los prefijos es de forma semejante a IPv4 en notación de CIDR.

Ejemplos:

Dir-IPv6/long-prefijo

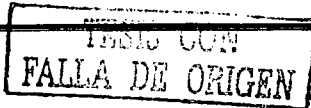
2001:200::/35

3FFE:600::/24

Donde:

Dir-IPv6: es cualquier dirección válida.

long-prefijo: es el tamaño de la máscara (número de bits).



### 6.3.3 Asignación del espacio de direcciones

IPv6 tiene una estructura más compleja a diferencia de IPv4, en donde antes de CIDR se habían definido una serie de rangos que se conocían como clases (A, B, C, D y E). De forma similar los bits de más alto orden, definen los diferentes tipos de direcciones en IPv6. El campo de longitud variable que abarca estos bits es llamado *format Prefix (FP)*. La tabla 6-2 muestra los rangos iniciales de estos prefijos.

Asignación	Prefijo (FP) (Binario)	Fracción del espacio de direcciones
Reservado	0000 0000	1/256
Sin Asignar	0000 0001	1/256
Reservado para NSAP	0000 001	1/128
Reservado para IPX	0000 010	1/128
Sin Asignar	0000 011	1/128
Sin Asignar	0000 1	1/128
Sin Asignar	0001	1/32
Direcciones globales de Unicast	001	1/8
Sin Asignar	010	1/8
Sin Asignar	011	1/8
Sin Asignar	100	1/8
Sin Asignar	101	1/8
Sin Asignar	110	1/8
Sin Asignar	1110	1/16
Sin Asignar	1111 0	1/32
Sin Asignar	1111 10	1/64
Sin Asignar	1111 110	1/128
Sin Asignar	1111 1110 0	1/512
Dirección de Link-local	1111 1110 10	1/1024
Dirección de Site-local	1111 1110 11	1/1024
Dirección de Multicast	1111 1111	1/1024

Tabla 6-2 Asignación del espacio de direcciones.

TESIS CON  
FALLA DE ORIGEN

Este bloque de direcciones esta asignado para poder realizar agregaciones globales de enrutamiento, tener direcciones de uso local y direcciones de Multicast, otra parte del direccionamiento esta reservada para implementaciones futuras de NSAP e IPX, otra porción esta reservada para realizar alguna función especial como es el caso de la dirección 0x00. Gran parte del espacio de direcciones aún esta sin ser asignado y podrá ser utilizado en un futuro para la expansión del que actualmente se esta usando o inclusive para nuevos usos.

#### 6.3.4 Estructura de las direcciones:

La estructura de las direcciones de IPv4 es usada tanto para realizar la asignación y el tipo de localización de esta. El avance de los paquetes esta basado en el concepto de "The longest best matched prefix" la mejor coincidencia del prefijo más largo (exacto), como actualmente se utiliza en IPv4. Una dirección de Unicast: puede ser un agregado global (global address), un enlace-local (local-link), un sitio-local (site-local) o de un formato especial de direcciones. Una interfaz de IPv6 puede tener múltiples direcciones al mismo tiempo Unicast, Multicast o Anycast.

#### 6.3.5 Formato de las direcciones

El agregado de direcciones globales deberá ser usado para conectarse al Internet público y para cualquier otro propósito que requiera de direcciones globales únicas y enrutables.

El agregado de las direcciones esta organizado en tres niveles jerárquicos: Público (public), Sitio (site) y a nivel de Interfaz (interface). La topología pública abarca a los proveedores de servicio (service providers) que ofrecen al Internet público servicios de tránsito, y los puntos de intercambio (exchange points).

El nivel más alto de esta topología es el público que normalmente es una zona en la cual todo el enrutamiento es explicito, es decir, una zona en la que no existe la ruta de default (default-zone-free, DZF) y por lo tanto todas las rutas deben de ser conocidas.

TESIS CON  
FALLA DE ORIGEN

La topología de Site es local a un sitio u organización que no ofrece servicios públicos de tránsito, aunque algunos servicios privados de tránsito pueden ser ofrecidos.

La figura 6-1 muestra el formato para un agregado global de Unicast

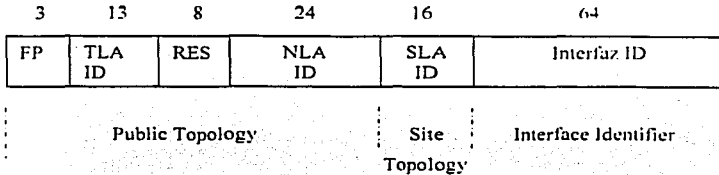


Figura 6-1 Formato de direcciones Unicast.

Los campos que definen el nivel público son FP, TLA, RES y NLA. SLA son a nivel de sitio y el campo Interfaz ID es a nivel de Interfaz. La porción de red de las direcciones son los primeros 64bits y la porción de los hosts son los últimos 64 bits.

Los campos de la dirección están definidos de la siguiente forma:

- FP es el formato del prefijo (001).
- TLA ID es el identificador de Top-level aggregation.
- RES es un campo reservado para uso futuro.
- NLA ID es el siguiente identificador del siguiente nivel.
- SLA ID es el identificador a nivel de sitio.

**TESIS CON FALLA DE ORIGEN**

*Identificador de Top-Level Aggregation (TLA ID)*

Estos identificadores (TLA ID) forman el nivel de enrutamiento más alto. Los routers en la default-zone-free deben de estar enrutando solo entradas de TLA ID y algunas entradas de su propia topología. El campo de TLA es de 13 Bits. Si más TLA fueran necesarios otro FP puede ser asignado, o el campo puede expandirse a la derecha dentro del campo reservado.

IANA le asigno inicialmente el bloque 0x0001, y esta a su vez lo subdividió para proveer bloques de TLA a los Register IP Regional (RIR). Los RIR asignaron Sub-TLA desde los bloques que les fueron asignados a los proveedores de IPv6.

*Reserved*

Este campo (RES) actualmente no es usado y todos los bits deben ser puestos a cero. El uso futuro que se espera tenga este campo es para incrementar los campos de TLA ID y NLA ID.

*Identificador de Next-Level Aggregation (NLA ID)*

Una organización a la cual se le asigno un TLA ID puede crear un direccionamiento jerárquico e identificar sus sitios usando el NLA ID. Las organizaciones que reciben la asignación de un bloque de direcciones desde un NLA son llamados NLA-registers.

*Identificador de Site-Level Aggregation (SLA ID)*

Un sitio es libre de crear tantos niveles como sean necesarios para crear una estructura jerárquica usando el SLA-ID. Esta puede ser un esquema plano sin subdivisiones o puede ser subdividido, dividiendo el SLA-ID en subredes, creando un esquema jerárquico.

*Identificador de Interfaz*

El identificador de interfaz es usado para enumerar a las interfases de un enlace específico, este es único en el enlace. Todas las direcciones con los bits de mayor orden dentro del rango 001 y 111, excluyendo a Multicast, deben tener un ID de interfaz de acuerdo al formato EUI. Actualmente solo los formatos especiales de direcciones asignados a NSAP e IPX, no están dentro de este rango.

La dirección MAC 0000:0C0A:2C51 se convierte en la dirección EUI-64 0200:0CFF:FE0A:2C51 por la inserción de FFFE después del identificador de la compañía y definiendo el valor del bit "universal/local".

TESIS CON  
FALLA DE ORIGEN

---

### 6.3.6 Formatos especiales de direcciones

Algunas de las direcciones usan un formato especial, todas ellas reconocidas por el FP reservado de 0x00. Las direcciones sin especificación (unspecified), las direcciones de loopback y las de "IPv4 contenidas en IPv6" son ejemplos de formatos especiales.

#### 6.3.6.1 Unspecified

Estas direcciones contienen solo ceros, 0:0:0:0:0:0:0 (::0). Estas nunca deben de ser asignadas a un nodo, actualmente representan la ausencia de una dirección. Uno de los usos para este tipo de direcciones, es emplearla como dirección fuente durante la inicialización de un host, antes de que se le haya asignado una dirección válida. Esta dirección nunca debe ser usada como dirección destino.

#### 6.3.6.2 Loopback

La dirección de loopback es la 0:0:0:0:0:0:1 (::1) y es análoga a la dirección IPv4 127.0.0.1. Un nodo puede usar esta dirección IP para enviarse paquetes a él mismo. Esta dirección no puede ser asignada a ninguna interfaz física.

#### 6.3.6.3 IPv4 contenido dentro de IPv6

Uno de los mecanismos de transición propuestos para ser usados durante la transición de IPv4 a IPv6 es el uso de túneles automáticos. En el cual los paquetes de IPv6 son automáticamente encapsulados en paquetes de IPv4 para ser transferidos sobre una red de IPv4. Este mecanismo requiere de un formato especial de paquetes de Unicast de IPv6. Los nodos que usan esta técnica son dual-stack -ejecutan tanto IPv4 como IPv6-. A los nodos dual-stack que soportan túneles automáticos se les debe de asignar una dirección IPv6, conteniendo a la dirección IPv4 en los últimos 32 bits. Todos los demás bits deben de ser ceros. El termino usado para este tipo de direcciones es "*IPv4-compatible con direcciones IPv6*" (también se dice que la dirección IPv6 contiene a la dirección IPv4) y siguen el siguiente formato.

::d.d.d.d

d.d.d.d es una dirección de normal de IPv4.

TESIS CON  
FALLA DE ORIGEN

Ejemplo:

::132.248.204.1

#### 6.4 Tipos de direcciones

Existen dos tipos de direcciones que tienen un significado local. Una dirección de Link-Local solo tiene significado en un enlace. El otro tipo son las direcciones de tipo Site-local, que tiene significado para los hosts del sitio en donde se encuentre. Estas direcciones globalmente no son únicas, ambas son únicas de acuerdo a su respectivo alcance.

##### 6.4.1 Direcciones Link-local

Estas direcciones pueden ser usadas por los nodos dentro de un enlace, para la autoconfiguración, el descubrimiento de vecinos, nodos sin capacidades de enrutamiento e incluso protocolos de enrutamiento pueden hacer uso de estas direcciones. Los routers no deben de avanzar paquetes con estas direcciones de Link-local como destino u origen. Sin embargo cualquier protocolo que desee asegurarse que los paquetes no serán enviados más allá del enlace local pueden hacer uso de una link-local. Una dirección de link-local esta definida por un FP = 111111010, seguido por 54 ceros y un ID de interfaz. Estas direcciones no contienen un TLA, NLA o SLA debido a que no contienen ninguna información de jerárquica. La figura 6-2 muestra el formato de una dirección de link-local.



Figura 6-2 Formato de una dirección Link-Local.

Cada nodo asigna a cada una de sus interfaces de IPv6 una dirección de link-local.. las cuales pueden ser configuradas automáticamente con autoconfiguración o pueden ser configuradas manualmente.

Ejemplos:

FE80::5ABC:01FF:FE0

FE80::0060:08FF:FEB1:7EA2

TESIS CON  
FALLA DE ORIGEN



**6.4.2 Direcciones de Site-local**

Un Site puede ser una organización o parte de ella. Un Site podría ser un cierta área topológica de algún lugar o puede ser multi-topológica que interconecte a varios lugares de alguna manera. Una red configurada con direcciones Site-local no deberá ser alcanzable desde organizaciones fuera de su organización. Los routers en la orilla de la organización (Site) deberán ser capaces de mantener el tráfico local a su organización y son también los responsables del control de la propagación de estas rutas. En adición a el FP de Site-local y del ID de interfaz estas direcciones tienen un identificador de subred, sin embargo no tienen los IDs de TLA o NLA. Estas direcciones deben de ser asignadas para ser usadas dentro de las organizaciones que no requieren de un prefijo global. El uso de estas direcciones es idéntico al de las direcciones privadas en IPv4 (RFC-1918).

Estas direcciones tienen definido el siguiente FP 1111111011 y es seguido por 38 ceros, un campo de subred de 16 bits el ID de la interfaz, como lo muestra la figura 6-3.

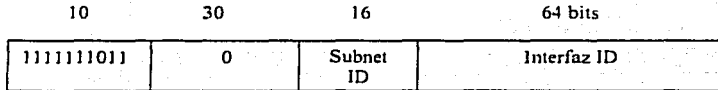


Figura 6-3 Direcciones Site-Local con el FP igual a 1111111011.

Ejemplos:

FECO::1:5ABC:1FF:FE01:1111

FECO::CAB:60:8FF:FEB1:7EA2

**6.4.3 Direcciones de Anycast**

El enrutamiento de tipo Anycast es un mecanismo de direccionamiento, mediante el cual múltiples interfaces, usualmente en diferentes nodos, tienen la misma dirección. El tráfico destinado a estas direcciones es enrutado al nodo más cercano. Existen muchos ejemplos de aplicación para este tipo de direcciones, uno de ellos es la definición del RP en Multicast. Las direcciones de Multicast son asignadas del espacio de direcciones de Unicast y por lo tanto no existe un FP especial para definir una dirección de Anycast. De hecho, las direcciones son tomadas del campo ID de la interfaz. La Subnet-Router de Anycast esta predefinida y todas las interfases de enrutamiento del enlace deben de tener asignada esta

dirección. Esta es una dirección Unicast que contiene solo ceros en la parte del identificador de interfaz. La figura 6-4 muestra el formato de la dirección de Subnet-Router.

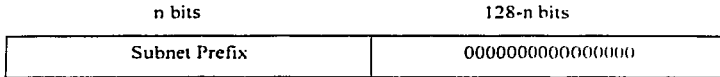


Figura 6-4 Formato de la dirección de Subnet-Router

De los 128 ID de interfaz, los más altos son reservados para ser asignados como direcciones de subred de Anycast. Una subred reservada de Anycast es una dirección de Anycast que puede estar disponible en cualquier subred de IPv6 sin importar el tipo de formato del prefijo. La figura 6-5 muestra la construcción de una dirección reservada de Anycast.

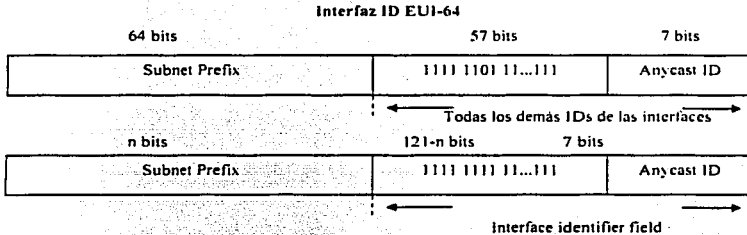


Figura 6-5 Construcción de una dirección reservada de Anycast.

La figura 6-5 muestra que el espacio reservado es tomado completamente desde el campo de interface ID. Una dirección reservada de Anycast es reservada para cada subred. Los 57 bits más significativos de la Interface Identifier son reservados para ser asignados como direcciones de Anycast. Los últimos 7 bits de la dirección identifican a una dirección específica de Anycast.

Como las direcciones de Anycast son *simbólicamente* indistinguibles de las direcciones de Unicast, una interfaz debe ser configurada en forma explícita para que pueda reconocer que su dirección, es una dirección Anycast.

TESIS CON  
 FALLA DE ORIGEN

#### 6.4.4 Direcciones de Multicast

Las direcciones de Multicast identifican un a grupo de interfaces, donde cada una de ellas puede contener también a otros grupos de Multicast. Las direcciones de Multicast pueden distinguirse de las de Unicast por que siempre comienza con 0xFF.

Las direcciones de Multicast en IPv6 pueden ser asignadas por algún organismo oficial de asignación (de tipo well-know address) o de transición para realizar algún propósito especial y por lo tanto asignada para un uso no-global. La asignación inicial de las direcciones de Multicast fue basada en las asignaciones que ya se habían realizado en IPv4. Entonces todas las asignaciones relevantes de IPv4 fueron convertidas a IPv6. En el RFC-2375 se encuentra un lista con todas las asignaciones de direcciones de Multicast de IPv6 que se han realizado actualmente.

Las direcciones de IPv6 también tienen definido un alcance. Estas direcciones tienen un campo que define el alcance que puede ser de alcance local al nodo, local al enlace (Link-local), local al Site (Site-local), o global. Las direcciones de Tránsito definidas con un cierto alcance tienen significado solo para los nodos dentro de ese alcance. Las mismas direcciones pueden ser definidas en diferentes lugares, o en diferentes redes que pueden tener un significado completamente diferente.

La figura 6-6 muestra el formato de una dirección de Multicast.



Figura 6-6. Formato de una dirección Multicast.

El primer octeto debe ser 11111111, para representar un dirección de Multicast.

*Flgs* es un grupo de bits. Los primeros tres bits están reservados y se les debe de asignar un valor de 0. El último bit indica si la dirección de Multicast esta una dirección

permanentemente asignada, conocida como de tránsito (transient). Un valor igual a 0 indica que la dirección de Multicast es bien conocida (well-know) y que alguna autoridad global de numeración la ha asignado.

El campo de *scop* es un valor de 4 bits usado para limitar el alcance de una dirección Multicast. La tabla 6-2 lista todos los posibles valores.

Valor	Descripción
0	Reservado
1	Alcance Node-local
2	Alcance Link-local
5	Alcance Site-local
8	Alcance local a la organización
E	Alcance global
F	Reservado



Tabla 6-2 Valores del campo Scop.

Cualquiera de los valores puede existir ya sea para una dirección Well-Known o transient.

El group ID identifica el grupo de Multicast, puede ser well-know o transient, con un cierto alcance.

La tabla 6-3 lista algunos de los grupos más comunes de Multicast con las direcciones y su alcance.

Dirección Multicast de IPv6 (Well-Known)	Dirección de Multicast IPv4 (Well-Known)	Grupo de Multicast
Alcance Node local		
FF01:0:0:0:0:0:1	224.0.0.1	All-nodes address

FF01:0:0:0:0:0:2	224.0.0.2	All-routers address
Alcance Link-local		
FF02:0:0:0:0:0:1	224.0.0.1	All-nodes address
FF02:0:0:0:0:0:2	224.0.0.2	All-routers address
FF02:0:0:0:0:0:5	224.0.0.5	OSPF-IGP
FF02:0:0:0:0:0:6	224.0.0.6	OSPF-IGP-Designated routers
FF02:0:0:0:0:0:9	224.0.0.9	RIP router
FF02:0:0:0:0:0:D	224.0.0.13	All PIM routers
Alcance Site-local		
FF05:0:0:0:0:0:2	224.0.0.2	All-routers address
Cualquier alcance valido		
FF0X:0:0:0:0:0:101	224:0:0:1	Network Time Protocol

Tabla 6-3 Grupos de Multicast "Well-Know".

#### 6.4.5 Direcciones requeridas por los nodos.

Los nodos requieren reconocer las múltiples direcciones que los identifican a ellos mismos. Los mecanismos de IPv6 requieren que los nodos reconozcan estas direcciones para trabajar correctamente.

Un host deberá poder reconocer las siguientes direcciones:

- Una dirección de local-link para cada interfaz.
- Todas las direcciones asignadas de Unicast .
- La dirección de loopback.
- La dirección de Multicast "All-nodes".
- La dirección de Multicast "solicited-nodes" para cada una de sus interfaces asignadas de Unicast y Anycast.
- La direcciones de Multicast de todos los otros grupos de los cuales es miembro el host.

TESIS CON  
FALLA DE ORIGEN

Además de todas las direcciones anteriores, un router también deberá reconocer las siguientes direcciones:

- La dirección de Anycast de "Subnet-router" para cada una de sus interfaces.
- Todas las demás direcciones de Anycast configuradas en el router.
- La dirección de Multicast "all-routers".
- Las direcciones de Multicast de todos los grupos de los que es miembro el router.

### 6.5 Encabezado de IPv6

Una de las principales metas de diseño de IPv6 es mejorar el encabezado con respecto al de IPv4, es más simple, más flexible, y más eficiente cuando se utiliza con diferentes opciones. Algunos de los encabezados de IPv4 fueron removidos y otros fueron renombrados. Las direcciones son cuatro veces más grandes, pero el encabezado es solo del doble de tamaño. La opción de codificación se modificó para hacer más eficiente al proceso y ofrecer una mayor flexibilidad en el tamaño y adición de opciones.

#### 6.5.1 Formato del encabezado

El Header (encabezado) es más simple. Este solo consta de ocho campos, incluyendo a las direcciones origen y destino. La figura 6-7 muestra el header de IPv6.

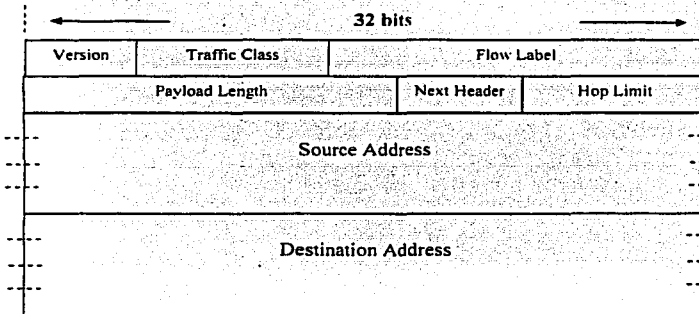


Figura 6-7 Formato del encabezado de IPv6.

Los campos del header de IPv6 son los siguientes:

- *Version* indica el tipo de versión (en este caso la versión 6)
- *Payload length* es el tamaño del paquete de IPv6, excluyendo al header. se mide en octetos. Los header de extensión se consideran parte de la carga útil (payload) del paquete y por lo tanto son incluidos dentro del tamaño del paquete.
- *Traffic flow* es utilizada por los nodos para indicar un trato especial a ciertos flujos de tráfico. Los nodos pueden etiquetar los flujos para solicitar a los routers de IPv6 una calidad de servicio (QoS) diferente a la de default.
- *Traffic Class* son bits que pueden ser utilizados por los nodos origen y/o routers intermedios para distinguir diferentes clases o prioridades de los paquetes IP. Estos bits pueden ser usados de la misma forma que en Type-of-Service de IPv4 y los bits de precedencia. Cuando se utiliza DiffServ se redefine el campo Traffic Class por DS. La definición del campo DS es la misma en IPv6 e IPv4.
- *Next Header* este es un campo que identifica el siguiente header después del de IPv6. El siguiente header puede ser cualquiera de los header de capa superior o un header de extensión de IPv6. Algunos ejemplos de Next header son:
  - Hop-by-Hop Options
  - Destination Options
  - Routing
  - Fragment
  - Authentication
  - Encapsulation Security Payload
  - Destination Options
  - Ospf para IPv6
- *Hop limit* se decrementa cada que pasa a través de un router/nodo. El paquete es descartado si el "hop limit" llega a cero. Su máximo valor es de 255
- *Direcciones destino / origen* son los campos de 128 bits que contienen las direcciones fuente y destino del paquete.



---

## 6.6 Funcionalidades de IPv6

A IPv6 se le han incluido desde el diseño un cierto número de funcionalidades, las cuales deben ser implementadas por cualquier nodo que soporte IPv6, incluyendo las siguientes:

- ICMPv6
- Neighbor discovery
- Stateless autoconfiguration
- Anycast
- Multicast
- MTU path discovery

Estos son funciones básicas de IPv6, las cuales mejoran las capacidades de IPv4. Otra de las ventajas de IPv6 es la habilidad de poder asignar múltiples direcciones a cualquier interfaz, facilitando el problema de reenumeración. No solo cualquier interfaz puede tener múltiples direcciones de IPv6 en múltiples prefijos, los nodos pueden comunicarse entre ellos directamente (sin necesidad de un router), sin importar los prefijos a los cuales pertenezcan, solo requieren pertenecer al mismo segmento físico.

### 6.6.1 ICMPv6

ICMPv6 es parte integral de IPv6 y cada nodo que lo implemente deberá cumplir completamente con ICMPv6. Es una versión modificada del ICMP de IPv4 con el soporte de reporte de errores más otras funcionalidades como el descubrimiento del MTU a través de un camino y el descubrimiento de vecinos.

El paquete de ICMPv6 sigue al header de IPv6 o a alguno de los header de extensión y es identificado con un valor de 58 bytes en el campo de Next Header (no es el mismo valor que se utiliza en IPv4). Los mensajes de información y error son identificados por el bit de más alto orden en el paquete de ICMPv6 dentro del campo type. Algunos ejemplos de mensajes de error de ICMP son:

- Destination Unreachable (Destino inalcanzable)
- Packet Too Big (Paquete demasiado grande)
- Time Exceeded (Tiempo de vida excedido)
- Parameter Problem (Problema en algún parámetro)

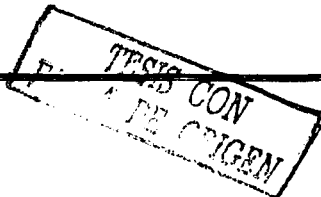


### 6.6.2 Descubrimiento de vecinos

El protocolo ND (Neighbor Discovery) resuelve muchos de los problemas que se presentan con los nodos de un enlace. Este protocolo proporciona las siguientes funcionalidades sin necesidad de un servidor (serverless) para poder realizar las configuraciones automática: router discovery (descubrimiento de routers), prefix discovery (descubrimiento del prefijo de red), la detección de un vecino inalcanzable, el MTU usado en el enlace, el siguiente salto (Next-Hop), detección de direcciones duplicadas. En IPv4 es necesario usar una combinación de múltiples protocolos para proporcionar estas funcionalidades (DHCP, ARP, ICMP router discovery).

ND usa ICMPv6 para realizar estas funcionalidades. ND utiliza cinco paquetes de ICMPv6 para proveer a los nodos de IPv6 de toda la información que ellos deben conocer antes de poder comunicarse con otros nodos.

- *Router Solitation (RS)* – Es enviado inmediatamente por los nodos usando Multicast cuando estos requieren enviar un "Router Advertisement", en vez de esperar por el siguiente anuncio. Durante la inicialización de un nodo este puede enviar un Router Solitation para de esta forma conocer los parámetros de configuración y de los routers existentes en el enlace.
- *Router Advertisement (RA)* - Son enviados periódicamente o enviados en repuesta a una solicitud. Los routers anuncian su presencia, con lo cual también proveen a los nodos de la información que ellos necesitan para configurarse.
- *Neighbor Solicitation (NS)* - Permite que los nodos conozcan las direcciones de la capa de enlace de datos de un vecino, o le permiten determinar si el vecino todavía esta presente por cierta dirección de la capa de enlace. Además permite a los nodos determinar si una dirección IPv6 esta duplicada en el enlace.
- *Neighbor Advertisement (NA)* – Es enviada en repuesta a un Neighbor Solitation. o sin ser solicitada si a un nodo se le cambia la dirección de la capa de enlace de datos.
- *Redirect* - Enviada por los routers para redireccionar el tráfico a un mejor salto siguiente.



## 6.7 Autoconfiguración

Debido a que la capacidad de administración de una red es crucial para el éxito de cualquier red, fue necesario que incluyeran procesos dentro del protocolo que ayudaran a realizarla. Las redes con configuraciones estáticas manualmente introducidas son difíciles de administrar cuando es necesario realizar cambios. Muchas herramientas se crearon para facilitar la administración en IPv4 como es DHCP para minimizar la cantidad de configuraciones estáticas, pero estas no son parte del protocolo. En IPv6 los nodos se pueden configurar automáticamente ya sea con DHCP o sin él, haciendo más fáciles los cambios de configuración en los hosts.

Los Router Advertisement son usados para decirles a los hosts como se deben de configurar, si es por medio de un proceso Statefull como de DHCP o de forma Stateless.

### 6.7.1 Stateless Autoconfiguración

Por medio de una combinación de lo que el nodo conoce (su identificador de Interfaz) y de lo que los routers conocen (el prefijo asignado al enlace), un nodo puede configurar su propia dirección IPv6, no es necesario utilizar un servidor para poder establecer una conectividad básica de IP. Esto solo funciona en interfaces que soporte Multicast.

Al momento que una interfaz se activa, un nodo genera una dirección de link-local para la interfaz. La dirección link-local es generada de la concatenación del identificador de interfaz con el prefijo ya conocido para direcciones locales (prefijo FE80). Los ceros de la derecha son reemplazados por el identificador de interfaz (normalmente la dirección MAC) formando una dirección de 128 bits. Los identificadores de interfaz normalmente son de 64 pero no siempre.

Si el identificador de interfaz es mayor a 118 de longitud, este no podrá ser concatenado con el FP para una dirección local. La autoconfiguración entonces fallará y el nodo deberá ser configurado manualmente.

TESIS CON  
FALLA DE ORIGEN

---

El nodo no puede asignar inmediatamente la dirección local a la interfaz. primero deberá determinar que esta dirección no se encuentra duplicada.

Cuando un nodo esta satisfecho de que su dirección no se encuentra duplicada. esta la asigna a su interfaz. En este punto, existe una conectividad básica a nivel de IP entre los host en un enlace sin ser necesario un router para comunicarlos.

Continuando con el proceso de autoconfiguración los hosts y no los routers. envían un mensaje de Multicast utilizando la dirección de "all-router" para encontrar los routers que se encuentra en el enlace. Todos los routers responden con un mensaje de "Router Advertisement". El RA les puede decir a los hosts que usen autoconfiguración stateful para configurar su dirección y para adquirir otra información que fuera necesaria. Los hosts usan la información del prefijo que les es dada para crear una dirección de Site-local. Para crear esta dirección. utilizan el FP de sitio que les es dado y lo concatenan con el ID de la interfaz. Un host no requiere de realizar un proceso de detección cuando se le asigna una dirección de sitio. Ya que en teoría esto fue verificado cuando se le asigno una dirección local. Esto significa entonces que su identificador de interfaz es único en el enlace. Como la dirección de sitio le es asignado un prefijo diferente al mismo identificador de interfaz la dirección de sitio es única. La dirección global es generada siguiendo el mismo proceso.

El RA además provee también de una lista de los prefijos que se encuentra en el enlace. Esta lista de prefijos y del tamaño de los mismo, son usadas por los hosts para construir sus listas de prefijos. La lista de prefijos es usada para determinar cuando un nodo esta en el enlace o fuera de él y por lo tanto hacer uso del default router para enviar el tráfico.

Un hosts puede también configurar el tamaño del MTU basado en la información del RA. La configuración de stateful es requerida para configurar otra información adicional como el DNS.

El proceso de autoconfiguración ocurre en cada interfaz de un host cuando esta es activada. Un nodo multihomed realizara un proceso de autoconfiguración para cada interfaz que

tenga de forma independiente. Una interfaz esta activada cuando realiza ó sucede alguno de los siguientes eventos:

- La inicialización de la interfaz se realiza al momento de iniciar el sistema.
- La interfaz se reactiva después de que se presenta una falla en la interfaz
- La interfaz es conectada al enlace por primera.
- La interfaz es habilitada después de ser administrativamente deshabilitada.

### 6.7.2 Stateful autoconfiguration

Stateful autoconfiguration puede ser usado en conjunto con stateless autoconfiguration. DHCP provee a IPv4 stateful autoconfiguration para IPv4. Una modificación de DHCP para IPv6 sea implementado para tomar ventaja de las numerosas características de IPv6, mejorando las capacidades de DHCP.

Se utiliza un servidor para la configuración de las direcciones asignadas y otra información como las direcciones de los servidores de DNS. Las direcciones asignadas son asociadas con un tiempo de vida (durante el cual son validas) con los prefijos usados en stateless autoconfiguration. Este servidor tiene la capacidad de pedir a todos los hosts que revaliden sus direcciones asignadas pudiéndose usar los valores de tiempos de vida para reenumerar redes.

### 6.8 RIPng

RIPng (ng de "next generation") esta basado en la RIP versión 2(RIP-2). Ninguno de los procedimientos operacionales, contadores, o mecanismos de estabilidad han sufrido cambios. RIPng es RIP-2, solo que se ha modificado para soportar las largas direcciones y múltiples direcciones que se pueden tener en cada interfaz de IPv6. El puerto de UDP es el 521. RIPng no puede soportar ambos protocolos IPv4 e IPv6 y por lo tanto no existe compatibilidad hacia atrás con RIP-2.

La figura 6-7 muestra el formato del mensaje de RIPng. La estructura es básicamente la misma que RIP-2.

Command (1)	Versión (1)	Deben ser Ceros (2)
Routing Table Entry 1 (20)		
Routing Table Entry N (20)		

Figura 6-7 Formato de Mensaje de RIPng.

Los campos del mensaje están definidos de la siguiente forma (con la longitud medida en bytes):

- **Command** se le asigna el valor de 1, para significar un petición, o 2 para una respuesta.
- **Version** actualmente 1.

El resto del mensaje contiene el listado de las entradas de la tabla de enrutamiento (route table entries, RTE): La figura 6-8 muestra el formato del RTE.

IPv6 Prefix (16)		
Route Tag (2)	Prefix length (1)	Metric (2)

Figura 6-8 Formato de las entradas de enrutamiento en RIPng.

Los campos en el formato del RTE son definidos de la siguiente forma:

- **IPv6 prefix** - Es la dirección del prefijo de IPv6 de 128 bits
- **Route Tag** - Este campo es idéntico al de RIP-2, el cual provee de un campo para la etiquetación de rutas externas o rutas que fueron redistribuidas desde otro proceso de RIPng
- **Prefix Length** - Especifica la parte significativa de la dirección del prefijo
- **Metric** - Este es igual al de RIP-2, es el número de salto para llegar a la red destino y es un valor entre 1 y 15.

TESIS CON  
FALLA DE ORIGEN

El número de rutas que puede contener un mensaje de actualización de RIPng depende del MTU del enlace, el número de octetos del encabezado que precede al mensaje de RIPng, el tamaño del header de RIPng el tamaño de los RTE.

Cada RIP-2 RTE contiene el campo de next-hop, el cual especifica a través de cual router se encuentra el mejor camino. Las direcciones IPv6 son tan largas que esto podría doblar el tamaño de los RTE. RIPng especifica un solo next-hop RTE que se aplica a todos los siguientes RTE hasta el final del mensaje o hasta encontrar la existencia de otro next-hop RTE. Cuando el valor 0:0:0:0:0:0:0 se encuentra en el campo de next-hop esto significa que el siguiente salto es quien origina el mensaje.

La dirección del next-hop debe ser la dirección de link-local del siguiente router. Si la dirección no es una dirección local, el receptor tratar el paquete como si la dirección del paquete fuera el valor 0:0:0:0:0:0:0.

Las actualizaciones periódicas y las respuestas de tipo triggered deben de mantenerse locales al enlace, estas no deberán atravesar un router. Ambas actualizaciones las periódicas y las triggered deberán tener la dirección local del router como dirección origen del anuncio y con un hop-limit de IPv6 de 255. El hop limit de 255 asegura que el anuncio no podrá a atravesar un router ya que este se decremента cada que pasa a través de un router. La dirección destino será la dirección de multicast "all-rip-routers" FF02::9.

IPv6 Next-Hop Address (16)		
Zeros (2)	Zeros (1)	0xFF (1)

Figura 8-8 Next-Hop RTE.

TESIS CON  
FALLA DE ORIGEN

---

### 6.9 BGP-4 Multiprotocol extensiones para IPv6

A BGP-4 se le han realizado modificaciones para poder soportar otros protocolos como IPX, Multicast e IPv6.

Tres piezas de información de BGP-4 son específicas de IPv4:

- El atributo de Next-hop
- El atributo de Aggregator
- El NLRI (Network Layer reachability information)

Dos nuevos atributos son definidos para el soporte de múltiples protocolos sobre BGP. El multiprotocol NLRI (MP-REACH-NLRI) y el multiprotocol-unreachable. Ambos opcionales y transitivos. Como su nombre lo sugiere, el atributo de multiprotocol-reachable NLRI describe la forma de alcanzar los destinos. El atributo contiene información acerca de la capa de red del protocolo a la cual la dirección pertenece y la dirección del next-hop a usar para avanzar los paquetes a sus destinos todo esto contenidos en una lista de prefijos. Cada mensaje de actualización de MP-REACH-NLRI contiene un next-hop y una lista de NLRI asociados a él. El NLRI es un par de elementos con el formato <Tamaño/Prefijo>, en el cual el Tamaño es la longitud del prefijo y Prefijo es una dirección IPv6 alcanzable.

El next-hop es la dirección a ser usada por speaker de BGP cuando se realice el avance de los paquetes. Las reglas por default para el next-hop son las mismas que IPv4:

- Si el router que realiza el anuncios y el router que recibe el anuncio están en diferentes sistemas autónomos (external peers), la dirección del NEXT-HOP será la dirección de la interfaz del router que realizo el anuncio.
- Si el router que realiza el anuncio y router que recibe el anuncio están en el mismo sistema autónomo (internal peers) y el NLRI hace referencia a un destino dentro del mismo sistema autónomo, la dirección del NEXT-HOP será la dirección del router que realizo el anuncio.
- Si el router que realiza el anuncio y el que lo reciben están en el mismo sistema autónomo y el NLRI de la actualización hace referencia a un AS diferente, el next-hop es la dirección del peer externo desde el cual la ruta fue aprendida.

## 6.10 Mecanismos de transición desde IPv4 a IPv6

Un nuevo protocolo de enrutamiento no puede ser implantado sin una clara metodología de transición. La facilidad de los procedimientos de transición, es un gran factor para que un nuevo protocolo pueda ser implementado. Es necesario que IPv6 interopere con IPv4. los nodos de IPv6 necesitan poder comunicarse con los nodos de IPv4. Los nodos de IPv6 al menos inicialmente necesitan comunicarse con los de IPv4 y esto todavía no está definido. El grupo de trabajo NGTRANS de la IETF a desarrollado un número de diferentes metodologías para facilitar la transición.

### 6.10.1 IPv6 a través de túneles de IPv4

La compatibilidad con IPv4 es posible de varias formas. Un nodo puede estar ejecutando las dos implementaciones al mismo tiempo. Esto hace posible comunicarse usando ambos. Otra forma es que un nodo encapsule los paquetes de IPv6 dentro de paquetes de IPv4 creando lo que se conoce como un túnel sobre una red de IPv4, permitiendo al nodo de IPv6 comunicarse con otros nodos de IPv6. Existen dos mecanismos para crear estos túneles para IPv6:

- Túneles automáticos
- Túneles configurados manualmente

En el formato IPv4 compatible con IPv6 está definido que los primeros 96 bits de la dirección IPv6 sean ceros los restantes 32 sean compuestos a partir de la dirección IPv4. por ejemplo ::172.0.1.1 es una dirección compatible con IPv4. Un nodo configurado con dirección compatible con IPv4 usa túneles automáticos.

### 6.10.2 Dual Stack

Una forma de que un nodo implemente IPv6 y mantenga la compatibilidad con IPv4 es la implementación de ambos. Un nodo que implementa ambos stacks es llamado nodo IPv6/IPv4. Un nodo IPv6/IPv4 puede comunicarse con nodos de IPv6 usando paquetes IPv6 y con nodos IPv4 usando paquetes IPv4.

TESIS CON  
FALLA DE ORIGEN



En un nodos IPv6/IPv4 se deben de configurar ambas dirección es IPv6 e IPv4. Las direcciones pueden estar relacionadas o no. Las direcciones compatibles con IPv4 pueden ser vistas como una sola dirección.

Las direcciones pueden ser configuradas de varias formas:

- Las direcciones IPv6 pueden ser configuradas de forma automática usando stateless o stateful. La dirección puede ser compatible con IPv4 o una sola dirección IPv6.
- Se puede utilizar cualquier mecanismo para configurar un nodo de IPv4.
- Se puede configurar una dirección compatible con IPv4 usando cualquier método de configuración de IPv4.

Pero un nodo con ambos protocolos debe de tener algún mecanismo para determinar que direcciones debe utilizar. Este mecanismo es provisto por el DNS.

#### 6.11 DNS para IPv6

Un nuevo tipo de registro de DNS se a definido para IPv6, el registro AAAA. Este nuevo tipo de registro provee el mapeo de direcciones IPv6. Un resolver de DNS deberá ser capas de manejar ambos registros, los registros A de IPv4 y los AAAA de IPv6. Cuando un nodo realiza una petición de una dirección, un registro A o AAAA deberá ser devuelto. El tipo de dirección determina el protocolo a utilizar. Si un registro A es regresado, este regresara una dirección IPv4 y el protocolo a utilizar será IPv4, de forma similar sucede con el registro AAAA que regresa una dirección IPv6.

Cuando se realiza una asignación compatible con IPv4 el DNS ambos registros deben de ser definidos, el A y AAAA. El resolver tiene entonces la opción de regresar ambas dirección es o solo una de ellas.

TESIS CON  
FALLA DE ORIGEN

# Capítulo 7

## Propuesta de Diseño.

TESIS CON  
FALLA DE ORIGEN

## Capítulo 7 Propuesta de Diseño

En el diseño de una red es de suma importancia tener en cuenta las características y el comportamiento de los protocolos que serán implementados, debido a que el comportamiento de muchos de los protocolos puede llegar a afectar de forma negativa el comportamiento de toda la red cuando son implementados en gran escala. De ahí que cuando se realiza el diseño de una red sea muy importante tomar en cuenta estos factores, de tal manera que una adecuada implementación de los protocolos, permita que los administradores puedan identificar un problema y en caso de ser necesario logren aislar fácilmente la parte de la red en donde se presenta la falla, en esos momentos cobra mayor importancia contar con esquemas de redundancia que permitan seguir ofreciendo los servicios de red al resto de los usuarios.

En este capítulo analizaremos varias de las mejores prácticas o recomendaciones a seguir referentes a cada uno de los protocolos propuestos para ser implementados. En el momento de empezar el diseño es muy importante definir los alcances del proyecto, debido a que de ello dependerá la utilización de cierta característica de algún protocolo, por ejemplo para una buena implementación de OSPF es importante delimitar un esquema de diseño que defina cuantas áreas deberán ser utilizadas y cuantos routers podrá contener cada una de ellas.

En la red CUDI quienes se conectan a la red son los asociados y afiliados, en este capítulo usaremos la palabra cliente para referirnos de forma indistinta a alguno de ellos.

### 7.1 Propuesta de OSPF como protocolo de enrutamiento Intradominio

En el capítulo 3 se realizó una descripción de la mayor parte de las características del protocolo de enrutamiento, pero no se mencionaron los problemas que tales características pueden presentar en la operación de la red cuando son implementadas, y las limitantes que se pueden presentar debido a las capacidades de los equipos de comunicación.

TESIS CON  
FALLA DE ORIGEN

### 7.1.1 Diseño escalable de OSPF

Para lograr que OSPF pueda soportar una red con una gran cantidad de routers, la red deberá ser jerárquica. Este punto se logra dividiendo al dominio de enrutamiento en áreas. OSPF es muy robusto en la variedad de áreas que se pueden utilizar, parte del objetivo de realizar esta división mediante la implementación de áreas, es poder realizar sumalizaciones de rutas entre ellas. Esto evita que los routers tengan que conocer los detalles de toda la infraestructura de la red. Para poder realizar la sumarización en un solo anuncio a un grupo de redes es necesario que las redes sean continuas.

En conclusión, para tener un diseño de red escalable es necesario tener:

- Una Jerarquía en la red en basada en áreas.
- Uso de áreas Stub.
- Empleo de un direccionamiento continuo.
- Sumarización de rutas.

### 7.1.2 Reglas en el diseño de las áreas

Cuando una red es dividida en áreas, es importante definir el área de mayor importancia, la cual siempre debe existir aun cuando el número de routers sea muy pequeño y no haya justificación para dividir a la red en áreas. Todas las áreas deben de tener conexión al área de backbone (área 0), los virtual-links solo deben de ser utilizados de forma temporal, la razón se debe a que cuando existen problemas en el área de tránsito, los routers en el área que se encuentra detrás del virtual-link no tienen forma de conocer lo que esta ocurriendo en el área de tránsito, debido a que el virtual-link no es una conexión física sino lógica al área cero, el área de backbone tampoco es la excepción y también debe de ser continua.

En resumen las reglas en el uso de áreas son:

- Debe de existir al menos el área cero.
- Todas las otras áreas deben de tener conexión física con el área cero.
- El área de backbone debe de ser continua.
- Evitar el uso de virtual-links.

TESIS CON  
FALLA DE ORIGEN

### 7.1.3 El diseño del direccionamiento

El diseño del direccionamiento es uno de los primeros puntos que se debe definir, para realizar un buen diseño, se deben de tener en mente los siguientes puntos:

- El objetivo debe ser el de mantener una link-state database.
- Crear un direccionamiento jerárquico que coincida con la topología.
- Tener bloques de direcciones separados para la infraestructura y para los usuarios (y/o clientes).
- Examinar la topología física y determinar si es Full Mesh o en estrella.
- Tratar de emplear áreas de tipo Stub tanto como sea posible.
  - Ayudan a reducir el overhead y el número de LSAs
- Impulsar la creación de un backbone.
  - Tratar de reducir el full-mesh y promover un diseño jerárquico.
- Emplear un solo proceso de SPF por área. una sola inundación de LSAs por área.
  - Tener cuidado de evitar sobrecarga a los ABR.
- La implementación de diferentes tipos de áreas resulta en diferentes tipos de inundaciones: áreas normales. Stub. NSSA. Totally stubby tal implementación ayuda a que los cambios en la red solamente sean propagados a las áreas que necesitan conocer tales cambios.
- Implementar redundancia.
  - Es recomendable que el diseño permita implementar dos enlaces hacia cada área.
  - Evitar demasiada redundancia.
    - Dos enlaces del backbone hacia un área Stub deben ser iguales, de lo contrario el enrutamiento se vuelve poco eficiente.
    - Con una mala sumarización, la redundancia excesiva en el área de backbone puede llegar a afectar su rapidez en la convergencia.

#### 7.1.4 Características de OSPF que se deben considerar

Algunas de las características a considerar cuando se realiza la configuración en los routers son:

El comando *"logging neighbour chances"* sirve en la detección de problemas. los administradores puede detectar que se ha presentado un problema en la red. Cuando un router al enviar un mensaje de logging detecta que ha perdido la comunicación con un vecino.

El comando de *"reference cost"* hoy en día es muy útil. a causa de que inicialmente el ancho de banda fue el valor que se utilizo para definir el costo en los enlaces. tomando el valor de 100Mbps como el valor de los enlaces de mayor ancho banda. actualmente existen enlaces con mayores velocidades. por lo cual. para tener un mejor funcionamiento es necesario mover esta referencia con la ayuda de este comando. de esta manera se evita la necesidad de modificarlo en cada enlace del router.

El comando *"Router ID"* sirve para definir el ID que el router utilizará para el proceso de OSPF. en caso de realizar la configuración manualmente. el router utilizará la dirección de alguna de sus interfaces, prefiriendo utilizar la de mayor ancho de banda. esto puede generar un problema debido a que si esta interfaz deja de estar en operación será necesario reiniciar todo el proceso de OSPF sin importar que todas las demás interfaces estén funcionando correctamente. Una práctica más adecuada es la utilización de interfaces de loopback. las cuales puede ser utilizadas por todos los protocolos y por otros procesos que realice el router con la ventaja de no tener una dependencia del medio físico y por lo tanto siempre estará operacional.

El comando *"process clear/start"* sirve para reiniciar el proceso de OSPF cuando por alguna razón es necesario.

En resumen, las características que se deben de considerar cuando se realiza la configuración son:

- OSPF "logging neighbour chances".
- OSPF "reference cost".
- OSPF "router ID".
  - Uso de interfaces de loopback para agregar estabilidad. (commando "Router Id").
- OSPF "process clear/restart".

### 7.1.5 Agregando redes al proceso de OSPF

Es muy importante definir cuales y la forma en que las rutas serán inyectadas al proceso de OSPF. A continuación se listan algunas recomendaciones que nos permitirán realizar este proceso de la mejor manera posible:

- OSPF solo debe transportar las rutas de la infraestructura de red.
- En caso de realizar redistribuciones hacia OSPF se deben de utilizar rutas externas de tipo 1.
- Emplear una sentencia "network" por cada enlace.
  - Cada interfaz necesita de un comando "network". Las interfaces que no requieren realizar el envío de paquetes de Hello necesitan del uso del comando "Passive-Interfase".

### 7.1.6 Implementación en el backbone de CUDI

Desde los inicios del diseño, se sugirió emplear el protocolo de enrutamiento OSPF como protocolo IGP en virtud de que es un protocolo estándar ampliamente recomendado por la IETF capaz de soportar redes de gran tamaño.

Actualmente el backbone de CUDI se encuentra en un proceso de crecimiento, por lo tanto, no contiene un gran número de nodos y enlaces, por esta razón no se hizo una división en áreas, de ahí que todos los equipos de backbone se encuentren en una sola área, el área de backbone (área 0), es importante señalar que su implementación final deberá permitir el crecimiento tanto en número de nodos como de enlaces. Para los routers ID requeridos por

OSPF y BGP se utilizan interfaces de loopback, el resto de su implementación sigue las recomendaciones antes mencionadas.

En el anexo A se encuentra una plantilla de configuración que se utiliza en los routers del Backbone, la cual por políticas de seguridad de CUDI, en esta tesis no es posible mostrar la configuración completa de los routers del backbone. Una buena referencia sobre la configuración son las páginas de Internet, por ejemplo la página de Rob Thomas (<http://www.cymru.com/index.html>) en las que se ofrecen plantillas de configuración para los routers, las cuales recomiendan deshabilitar muchos servicios que actualmente son innecesarios o que simplemente pueden llegar a causar un problema de seguridad.

## 7.2 Propuesta de BGP como protocolo de enrutamiento Interdominio

En el capítulo 4 se hizo la descripción del funcionamiento del protocolo de enrutamiento BGP y de sus atributos, también se comentaron brevemente algunas de las técnicas que se pueden utilizar cuando su implementación es realizada, en este capítulo profundizaremos en las "mejores prácticas" (Best-practices), que conjuntamente permiten una implementación de BGP a gran escala.

### 7.2.1 Técnicas de Escalamiento

Las siguientes partes del protocolo de BGP fueron descritas en el capítulo cuatro, su empleo ayuda a realizar una implementación adecuada:

- Route flap dampening
- Communities
- Route Reflectors

El uso de estas partes en conjunto con algunas facilidades que poseen los equipos de comunicaciones, las cuales aunque no son parte del protocolo, ayudan a tener una mejor administración y una rápida solución de los problemas de la red, permiten que los archivos de configuración sean más fáciles de leer y menos susceptibles a errores, a continuación listamos dos de las facilidades más notables que comúnmente son mayormente usadas en redes grandes:

TESIS CON  
FALLA DE ORIGEN



- Soft reconfiguration
- Peer groups

### 7.2.1.1 Soft reconfiguration

También conocida como reconfiguración dinámica. Recordando un poco la forma en la que BGP trabaja, antes de que BGP propague la información de enrutamiento, establece una sesión de TCP por medio de la cual se realizará el intercambio, una vez que el router establece la sesión y recibe el anuncio de una ruta, antes de integrarla a la tabla global de enrutamiento, el anuncio debe de ser validado antes de ser aceptado; existen muchas razones por las cuales puede que un anuncio no sea aceptado para su ingreso a la tabla de enrutamiento, una de estas razones podría ser que no se desea recibir el anuncio de esa ruta, para la cual el administrador colocó un filtro impidiendo que la actualización sea válida y por lo tanto sea desechada. Teniendo en mente que BGP solo intercambia sus tablas completas de enrutamiento al inicio de la sesión y posteriormente solo se envían mensajes que solamente contienen los cambios que han ocurrido, ¿qué sucede si después de cierto tiempo el administrador remueve el filtro para permitir que el anuncio sea válido?: simplemente esto no ocurrirá por que la actualización ya fue enviada y no se enviará ninguna actualización hasta que ocurra algún cambio.

Una forma de forzar a que los routers nuevamente se envíen las actualizaciones: es reiniciando la sesión de BGP con el vecino (también llamada hard-reset). Si las tablas de enrutamiento son muy grandes, el reestablecimiento completo de la sesión puede tomar mucho tiempo considerando que esto no solo ocurriría en los routers vecinos, sino también en todos los demás routers debido a que todos deben de propagar los anuncios.

El problema se origina debido a que los routers no almacenan los prefijos que son denegados por los filtros, por lo tanto un cambio en las políticas (configuradas en filtros) de entrada requieren que se realice un hard-reset.

Para solucionar este problema algunos fabricantes han creado la funcionalidad de "soft-reconfiguration", en la cual sus routers almacenan todas las actualizaciones de BGP en una

tabla de entrada, de tal forma que cuando se realiza algún cambio en las políticas de entrada no es necesario reiniciar la sesión, basta con consultar la tabla de entrada.

Esta funcionalidad trabaja con cualquier implementación estándar de BGP, debido a que solo es una funcionalidad local al router, pudiendo trabajar con otros routers que no la soporten, no es necesario tener una opción de soft-reconfiguración de salida ya que cuando una ruta cambia de inalcanzable a alcanzable se debe de realizar la notificación a los vecinos incluso si este cambio es debido por la aplicación de filtros.

### 7.2.1.2 Peer-groups

Otra funcionalidad de configuración incluida por varios fabricantes es de poder crear Peer-groups:

Sin el uso de Peer-groups:

- Los vecinos de IBGP reciben muchas veces las mismas actualizaciones (si recordamos que normalmente se encuentran conectados en full-mesh).
- Cuando se tienen muchos vecinos de IBGP la configuración llega a ser difícil de construir y mantener libre de errores.
- El router desperdicia mucho tiempo realizando el cálculo de actualizaciones repetidas

Se recomienda el uso de Peer-groups cuando:

- Cuando se tienen las mismas políticas de salida.
- Cuando se tienen que generar las mismas actualizaciones.

Normalmente cuando los vecinos de IBGP se encuentran en full-mesh requieren las mismas políticas de salida, incluso considerando que en muchas ocasiones se tienen las mismas políticas con vecinos de EBGP como puede ser el caso de un ISP.



Ventajas del uso de Peer-groups:

- Permite crear más fácilmente la configuración de full-mesh.
- La configuración es menos propensa a errores.
- Hace que la configuración sea más fácil de leer y de entender.
- Reduce la carga de procesamiento de CPU del router debido a que solamente realiza los cálculos para todo el Peer-group.
- Los miembros del Peer-group pueden tener diferentes políticas de enrutamiento.
- También se puede usar en vecinos EBGp.

### 7.2.2 Políticas de enrutamiento

Una parte importante para garantizar el correcto funcionamiento de cualquier organización es la definición de las políticas de enrutamiento, es decir, definir la forma en que fluye el tráfico, delimitar si un sistema autónomo es o no un sistema de tránsito, evitar que debido a una mala configuración accidental o mal intencionada un sistema autónomo ocasione que otro sistema sea usado como un sistema de tránsito cuando en realidad solo se está haciendo peering, todo esto se logra modificando la información y las características con las que el router la conoce. En el caso de BGP los valores de los atributos de las rutas ayudan a identificar estas políticas.

Las políticas de enrutamiento en BGP se puede implementar basándose en:

- AS-PATH, prefijos y comunidades.
- Aceptando o rechazando rutas.
- Modificando los atributos de BGP para afectar la selección de las rutas.

Las principales herramientas que se tienen y que pueden ser empleadas en la configuración de los routers son:

- Listas de acceso ACL/Prefix-list.
- Filtros de AS Filter-list.
- Router-Maps y comunidades.

TESIS CON  
FALLA DE ORIGEN

### 7.2.3 Interacción entre BGP y el protocolo IGP

Cuando se utiliza BGP una parte esencial referente a la estabilidad de la red: es la interacción que tiene con el protocolo IGP que se este utilizado debido a que sus funciones son muy distintas y una incorrecta implementación de alguno de ellos puede afectar la operación del otro y a su vez la operación de la red de toda la organización.

#### 7.2.3.1 Funciones del protocolo IGP

Todos los protocolos de compuerta interna (IGP) son diseñados para tener tiempos de convergencia muy bajos, propagan sus anuncios por medio de mensajes de Broadcast o multicast y generalmente solo tienen una visión que no va más allá de su organización o su sistema autónomo.

En resumen las funciones de los IGP's son:

- Deben de ser empleados para transportar los prefijos de la infraestructura.
- No deben de ser empleados para transportar los prefijos de los clientes.
- Diseñarlos de tal forma que se minimice el número de prefijos.

#### 7.2.3.2 Funciones de BGP

Una de las principales características de BGP es poseer una visión global de Internet a diferencia de los protocolos de compuerta interna (IGP's). otra característica de BGP es que no tiene un métrica para la selección de las mejores rutas, en lugar de esto, es rico en atributos, los cuales pueden ser modificados para alterar los caminos que siguen los paquetes hacia un destino dado.

A continuación mencionaremos algunas de las mejores prácticas en la implementación de BGP, es decir, ¿que hacer, como hacerlo?, y que no se debe de hacer. Básicamente son recomendaciones que se deben seguir cuando se implementa BGP para evitar muchos de los problemas que pueden presentarse en el momento en el que la infraestructura crece, o simplemente para garantizar un mejor desempeño de los equipos de comunicación, una explicación detallada del porque de cada una de las recomendaciones queda fuera del

TESIS CON  
FALLA DE ORIGEN

alcance de esta tesis y prefiero sugerir que se consulten las referencias bibliográficas para obtener información mucho más detallada.

Cada uno de los componentes de BGP realiza una función específica y por lo tanto no se deben de utilizar a este protocolo para realizar funciones para las cuales no fue diseñado o que sencillamente podrían ocasionar que los equipos de comunicación fallen. a continuación se listan las partes de BGP y las funciones que primordialmente deben de realizar. también se hace una recomendación sobre los esquemas que no se deben de efectuar.

#### *Recomendaciones en la implementación de IBGP:*

- Utilizarlo para transportar los prefijos de Internet a través del backbone.
- Emplearlo para transportar los prefijos de los clientes a Internet.
- Utilizar interfases de loopback en los routers para establecer las sesiones de IBGP. también se recomienda que sean el Router\_ID de BGP. Estas direcciones pueden tener una mascara de 32 bits.
- Usar Peer-groups.
- Usar passwords en la sesiones de IBGP.
- Si el IGP no contiene a las redes de la zona de demarcación (DMZ) usar la opción de configuración "next-hop-self".

#### **7.2.4 Recomendaciones en la implementación de EBG:**

- Usarlo para intercambiar prefijos con otros AS.
- Utilizarlo para implementar políticas de enrutamiento.
- Limitar el número máximo de prefijos que se pueden aceptar desde los vecinos de EBG.
- No ejecutar un IGP con los vecinos de EBG.
- Los vecinos deben de estar directamente conectados. no se debe utilizar EBG multi-hop.

TESIS CON  
FALLA DE ORIGEN

**7.2.4.1 Lo que nunca se debe de hacer:**

- Redistribuir los prefijos aprendidos desde BGP a un IGP.
- Redistribuir los prefijos de un IGP a BGP.
- Usar un IGP para transportar los prefijos de los clientes, con ellos se debe hablar BGP o enrutamiento estático.
- Usar a IBGP para transportar los prefijos de la infraestructura.

**7.2.4.2 Herencias de BGP que se deben deshabilitar:**

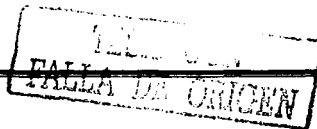
- Sincronización.- Para redes grandes es recomendable tener deshabilitada la sincronización, para lo cual se debe tener la certeza de que todos los vecinos de IBGP están conectados en full-mesh.
- Auto-Sumarización.- Aun cuando BGP es un protocolo de tipo Classless, en equipos de muchos fabricantes se tiene activo por default un comportamiento Classfull para la compatibilidad con equipos viejos, lo cual actualmente no es adecuado debido a que en Internet/Internet2 los anuncios son de tipo Classless, por lo tanto esta opción se debe deshabilitar.

**7.2.4.3 Anuncios de redes que no se deben recibir:**

En las políticas de enrutamiento se deben definir los prefijos que se estarán recibiendo desde los vecinos y delimitar hasta donde serán propagados para cada uno, adicionalmente siempre es recomendable tomar la precaución de evitar recibir los anuncios de ciertas redes que podrían llegar a afectar la operación de la red, como por ejemplo recibir el anuncio de redes de uso privadas, multicast, rutas por default, etc.

En resumen: el filtro de entrada aplicado en la sesión de BGP en cada vecino de EBGP debe evitar que se reciba el anuncio de los prefijos de:

- Las redes del RFC1918.
- La ruta por default o cualquier segmento de esta, en caso de haber acordado no recibir el anuncio.
- El anuncio del bloque de direcciones que tienen como origen el AS local.



- El anuncio de prefijos con una máscara de longitud demasiado grande. Para Internet: máscaras mayores a 24 bits y en el caso de Internet2 mayores a 27 bits.

### 7.2.5 Inyectando los anuncios de los prefijos al proceso de BGP

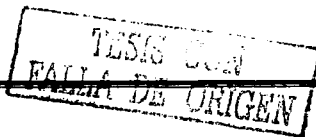
EL proceso de introducir los anuncios de los prefijos al proceso de BGP – también conocido como inyección de los prefijos a BGP- se puede realizar de tres formas: por redistribución de rutas estáticas, de forma semi-automática (uso del comando *Network*) y por redistribución de un IGP a BGP. El empleo de la última forma no es muy adecuada en redes grandes, muchos ISPs prefieren la redistribución de rutas estáticas debido a que suele ser más sencilla, aunque requiere se requiere la creación de un filtro para asegurar que solamente se redistribuirán los prefijos que se deseen, la forma semi-automática se realiza con una relativa facilidad de configuración mediante el comando *Network*, se dice que es semi-automática por que el router no incluirá el anuncio del prefijo hasta que este anuncio exista en el IGP o en una ruta estática.

La redistribución de rutas estáticas y la forma semi-automática son las más ampliamente recomendadas para su implementación en redes grandes siguiendo las siguientes recomendaciones:

- Usar a IBGP en lugar de emplear un IGP para transportar los prefijos de clientes.
- Utilizar una ruta estática que apunte hacia la interfaz física del cliente y definirla como permanente.
- La ruta estática debe abarcar todo el espacio de direcciones del cliente y de igual forma el prefijo se debe inyectar en BGP.

### 7.3 Topología de Internet 2

De forma similar a Internet, la topología de Internet 2 es jerárquica. En EE.UU. existen dos redes dorsales de Internet2, Abilene y vBNS, las cuales forman la parte más alta de la topología, en el siguiente nivel se encuentran las redes regionales y las conexiones a otros países y en el último nivel se encuentran las universidades y los centros de investigación, en la figura 7-1 se muestra la topología de Internet2 en EE.UU.



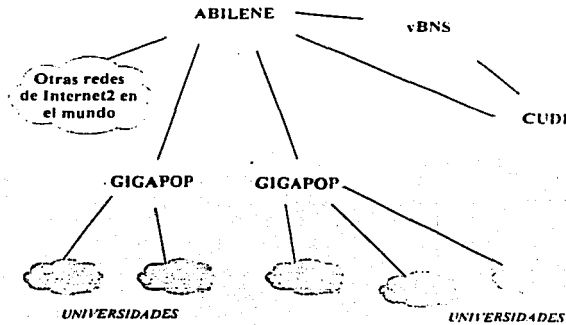


Figura 7-1 Topología de Internet2 en EE.UU.

### 7.3.1 Topología en México

En México la topología de la red también es jerárquica, en nuestro caso existe la red dorsal de CUDI, esta red representa el nivel más alto y esta formada con enlaces que poseen velocidades de transmisión de STM-1 (155Mbps), en el segundo nivel se encuentran los asociados académicos e institucionales y las conexiones internacionales, los asociados se conectan con enlaces de jerarquía de transmisión de E3 (34Mbps), el siguiente nivel esta formado por los afiliados académicos que se conectan a los asociados con enlaces mínimos de jerarquía E1 (2Mbps), en la figura 7-2 se ilustra la topología de la red de Internet2 en México, adjuntamente en la figura se puede apreciar que la red dorsal de CUDI no tiene conexión a Internet, pero los asociados y afiliados si pueden tener conexión tanto a Internet como a Internet2.

TESIS CON  
FALLA DE ORIGEN



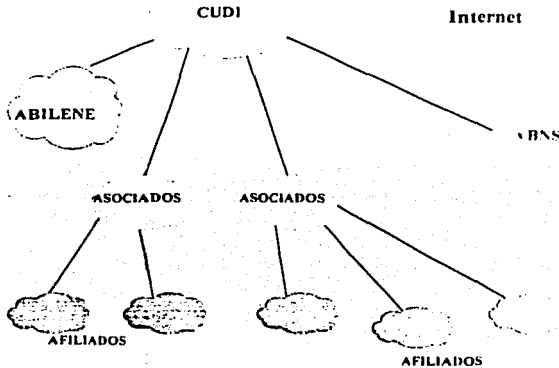


Figura 7-2 Topología de México Backbone, Asociados, Afiliados y conexiones Internacionales.

Una confusión que se presenta frecuentemente; es creer que para la red de Internet2 se requiere utilizar rangos de direcciones IP diferentes a los utilizados en Internet, esta idea es falsa debido a que las direcciones empleadas tienen que ser asignadas por algún organismo que administre las direcciones IP (como LACNIC, RIPE y en México NIC-México). lo que puede variar, es que los miembros pueden realizar el anuncio de una o varias partes de este bloque sin estar obligados a realizar el anuncio completo, pudiendo realizar el anuncio de prefijos con una longitud de máscara más grande que las usadas en Internet, teniendo como límite 27 bits de longitud, esto mismo sucede con el número de AS el cual debe ser asignado por alguno de los organismos anteriormente mencionados.

### 7.3.2 Función del Backbone

La función del backbone es servir de tránsito para el tráfico de Internet2 entre los diferentes asociados, afiliados y otras redes de Internet2 en el mundo por medio de enlaces internacionales, esto se ilustra en la figura 7-3. También se muestran los caminos que deben seguir los tráficos de Internet e Internet2.



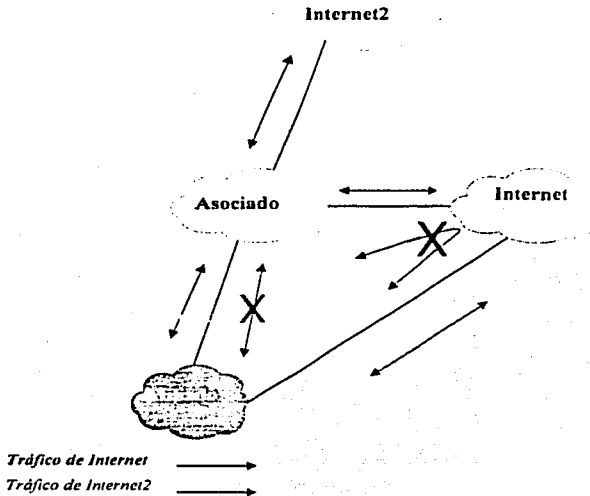


Figura 7-4 Tránsito del Internet e Internet2 entre Asociados y Afiliados.

La función de los asociados no se limita a conectar a los afiliados, el objetivo es tratar de crear redes regionales con otros asociados y afiliados, con la finalidad de tener acceso a Internet/Internet2 con menores costos y con una infraestructura propia sobre la cual tengan el completo control de los recursos de la red.

En este nivel se recomienda que los asociados intercambien la información de enrutamiento con los afiliados por medio del protocolo BGP, si esto no es posible por alguna dificultad técnica se aconseja emplear rutas estáticas, también de igual forma se sugiere tratar de seguir las recomendaciones antes mencionadas para la implementación de BGP.

### 7.3.4 Función de los miembros Afiliados

Los afiliados son los usuarios que se encuentran en el nivel más bajo, esto no impide que puedan conectar a otros afiliados, aunque su mayor aportación al proyecto son las aplicaciones de sus investigadores y usuarios, que en este caso son los maestros y alumnos que sumados con sus contra partes en los asociados y universidades en EE.UU, y otras partes del mundo completan la red de Internet.

En el anexo A existe una plantilla de configuración para una router del backbone, en esta plantilla se encuentran las líneas de configuración para vecinos IBGP y EBGP, la cual se sugiere consultar para realizar la implementación de BGP.

### 7.4 Propuesta para el soporte de Multicast

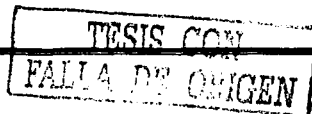
Para el soporte de multicast existen muchos protocolos de los cuales al igual que los protocolos de Unicast no todos son adecuados para redes de gran tamaño o para soportar un enrutamiento de multicast entre diferentes sistemas autónomos (enrutamiento interdominio).

Antes de tomar la decisión sobre el protocolo de enrutamiento multicast que utilizaremos, haremos una pequeña revisión de los protocolos comúnmente utilizados y trataremos de ubicarlos en los ambientes en donde su uso es más adecuado.

#### 7.4.1 DVMRP

Es uno de los protocolos de multicast con el que la gente está más familiarizado desde los inicios del enrutamiento multicast, es soportado en la mayoría de los fabricantes de equipos de comunicaciones. Utilizando túneles, DVMRP fue ampliamente utilizado en la red Mbone por muchos años.

Este protocolo es una versión modificada del protocolo RIP, por tanto, ha heredado muchas de las deficiencias que RIP tiene para escalar, debido a las actualizaciones periódicas y la utilización de envenenamiento de rutas, aunque en este caso no de redes sino de parejas (S, G), en redes grandes la estabilidad es otro problema en este protocolo.



Muchas de las características heredadas de RIP y otras desventajas, por ejemplo: tener dos protocolos de enrutamiento –uno para Unicast y otro para multicast– hacen que la implementación de DVMRP no sea apropiada para redes grandes o entre sistemas autónomos, su uso queda reducido a pequeñas redes LAN.

#### 7.4.2 PIM-DM

El protocolo de enrutamiento multicast PIM-DM es muy eficiente cuando se implementa en pequeñas redes piloto, también posee un comportamiento muy eficiente en los mecanismos de “flood” y “prune”, aunque si bien es cierto, este protocolo puede causar problemas en ciertas topologías de red. PIM-DM crea pares de estado (S, G) en cada router. Aunque es un protocolo fácil de configurar –Tan fácil como dos comandos– y que sus mecanismos de “flood” y “prune” son muy simples, las principales desventajas de este protocolo de enrutamiento multicast son:

- Poser un mecanismo para imponer su autoridad “assert” que resulta ser complejo.
- Mezcla los planos de control y de datos, lo cual puede dar lugar a comportamientos topológicos no determinísticos.
- No soporta árboles de distribución compartida.

En conclusión la implementación de PIM-DM es recomendable en redes pequeñas o en maquetas de laboratorio donde lo más importante es probar el funcionamiento de cierta aplicación de multicast y no el comportamiento del enrutamiento multicast.

#### 7.4.3 PIM-SM

PIM en modo esparcido (Sparse) es muy eficiente, debido a que posee un modelo de solicitud explícito que permite el ingreso de los hosts a los grupos de multicast, el tráfico de multicast solamente fluye en las partes de la red en donde realmente se demanda. A diferencia de PIM-DM, este protocolo realiza la separación entre los planos de control y de datos, lo cual permite que se tenga un comportamiento topológico determinístico. También resulta ser un protocolo más escalable a diferencia del modo denso, debido a que trabaja de igual forma en redes densa y/o escasamente pobladas de grupos de multicast, este protocolo

en comparación con PIM-DM soporta los árboles de distribución compartida y árboles de distribución fuente.

Debido a que PIM-SM emplea un modelo de "Explicit Join", el tráfico de multicast es forzado a circular solamente por las partes de la red en donde realmente se desea. Por lo tanto, PIM-SM no muestra ineficiencias de los mecanismos de "flood" y "prune" que presentan los protocolos como PIM-DM y DVMRP. Como resultado de esto, PIM-SM es más apropiado para redes de IP multicast que poseen miembros potenciales en el extremo de los enlaces WAN.

También podemos concluir que PIM-SM es la mejor opción de los protocolos de enrutamiento multicast tanto en redes intradominio como para la mayoría de la redes de multicast de propósito general. Las posibles excepciones son las redes de propósito especial que están diseñadas para ejecutar aplicaciones muy específicas bajo el completo control de los administradores de red.

Una vez que se ha tomado la decisión de emplear a PIM-SM como protocolo de enrutamiento intradominio multicast, aún faltan por agregar varias piezas que nos permitan realizar un enrutamiento entre diferentes sistemas autónomos. En el funcionamiento de PIM-SM se presentan dos problemas, el primero es que PIM-SM requiere de un enrutamiento unicast que debe ser compatible con el protocolo de enrutamiento que se tenga en los otros sistemas autónomos, el segundo problema es que se requiere que mediante algún mecanismo el RP local conozca sobre la existencia de las fuentes de multicast en los otros sistemas autónomos.

#### 7.4.4 MBGP extensiones para Multicast

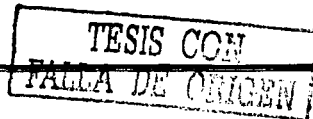
Para solucionar el primer problema, se debe de utilizar MBGP con las extensiones para el soporte de multicast. con la implementación de MBGP no solo se soluciona el problema de enrutamiento multicast. también se logra una solución muy robusta. esto es debido a que MBGP hereda todos los atributos de BGP. lo cual permite que se tengan diferentes políticas de enrutamiento para el enrutamiento unicast y multicast. igualmente permite la creación de topologías incongruentes. es decir. los caminos de multicast y unicast pueden ser diferentes. en resumen las ventajas de utilizar al protocolo MBGP son:

- Una red puede soportar topologías incongruentes de unicast y multicast
- Una red puede soportar topologías congruentes de unicast y multicast que poseen diferentes políticas (configuraciones de filtrado de BGP).
- Una red puede transportar información de enrutamiento de múltiples direcciones de familias de protocolos de la capa de enlace (por ejemplo. IPv4 o VPNv4).
- Una red que es compatible con routers que soportan las extensiones del multiprotocolo puede interoperar con routers que no las soporten.
- Todas las capacidades de las políticas de enrutamiento de BGP pueden ser aplicadas al multiprotocolo BGP (MBGP).
- Todos los comandos de configuración para BGP se pueden emplear en MBGP.

#### 7.4.5 MSDP

Para solucionar el segundo problema relacionado con la necesidad de que el RP conozca las fuentes de multicast en los sistemas autónomos. una primera idea seria tener un RP centralizado para todos los sistemas autónomos lo cual no es muy adecuado. debido a que el problema que ahora se presentaría. sería elegir en cual de todos los sistemas autónomos estaría ubicado. una vez solucionado este problema. ahora se presentarían problemas de operación. debido a que cada sistema autónomo no tendría la posibilidad de definir sus propias políticas.

Una mejor solución es que cada sistema autónomo tenga su propio RP en el cual puedan definir sus propias políticas de enrutamiento. aún cuando esta solución es la más viable. su implementación requiere de una herramienta que realice el intercambio de información



sobre las fuentes de multicast, es en esta parte en donde se requiere del empleo del protocolo MSDP. este protocolo permite que cada sistema autónomo pueda tener su propio RP y permite a cada sistema autónomo definir sus propias políticas sobre las fuentes y grupos en las que se desea intercambiar tráfico multicast.

#### 7.4.5.1 Redundancia al RP

Un último punto se refiere a que en toda red de producción son muy importantes los esquemas de redundancia, en esta parte PIM-SM tiene el problema de que el RP se puede convertir en un punto de falla de toda la red de multicast. La definición del RP se puede hacer de 3 formas: Definiéndolo estáticamente la dirección del RP en cada router, haciendo uso de un BSR y la última forma es por auto-RP.

Las dos últimas formas para definir al RP proporcionan un cierto nivel de redundancia, debido a que se pueden definir otros routers como respaldos de un router principal. Es importante mencionar que aunque se pueden tener routers de respaldo, todas las tareas se siguen concentrando en un solo router. La implementación de esta solución generalmente es adecuada desde redes pequeñas hasta redes con un tamaño mediano y que no poseen demasiadas conexiones a otros sistemas autónomos, para redes muy grandes es más apropiado definir varios RPs dentro del mismo sistema autónomo empleando direcciones de anycast, para evitar tener problemas de incongruencia en las fuentes multicast se deben de unir por medio de MSDP, para la definición de los RP se puede usar cualquiera de los métodos automáticos BSR o auto-RP.

La razón de utilizar direcciones de anycast es hacer creer a los routers que solo existe un RP, esto es porque en PIM-SM se tiene un solo RP, de esta forma si alguno de los RPs falla los routers se comunican con el otro RP creyendo que aún se trata del mismo router, este router que también es RP conoce toda la información de las fuentes, debido a que estuvo intercambiando esta información vía MSDP. Además, teniendo múltiples RPs se puede balancear la carga de trabajo al distribuir la cantidad de routers que atiende cada uno de los RPs, es importante señalar que puede haber más de dos RPs.



De igual forma que para los otros protocolos existen muchas recomendaciones que permiten implementar una solución completa, también existen recomendaciones para el protocolo MSDP, solo mencionaremos algunas de las más importantes sin profundizar a detalle en cada una de ellas, pues bien, esto cae fuera de los alcances de esta tesis.

Una de las primeras prácticas es tener preferentemente un solo protocolo de enrutamiento multicast para evitar realizar redistribuciones entre los protocolos y también permita asegurarnos de que el tráfico se mantenga en el modo en que se haya definido.

Cuando se realice la activación de multicast, se debe de tratar de hacerlo en la mayor parte de la red, al menos en todo el backbone, si es que se requiere cruzar.

Otra recomendación es activar multicast en todos los switches de capa 2, si no se hace esto, el switch en capa 2 avanzará todos los paquetes de multicast como si se tratarán de paquetes de broadcast, esto tendría como consecuencia un efecto negativo en el desempeño de la red.

En la sesión de MSDP se tiene que evitar que se registren direcciones de multicast del tipo bien conocidas "well known" las cuales tienen funciones específicas y son utilizadas por otros protocolos, como es el caso de OSPF que realiza el envío de LSAs empleando direcciones de multicast, existen direcciones que son de uso local, las cuales también se deben filtrar, otro tipo de direcciones son las de uso privado como las que define el RFC-1918 y por último el rango utilizado por otros protocolos de multicast como PIM-SM.

Se debe definir un filtro conocido como "filtro de frontera" el cual debe de evitar que los mensajes de los protocolos en modo denso o protocolos como OSPF se difundan más haya de las fronteras del sistema autónomo.

#### 7.4.6 Propuesta para el backbone

La propuesta para el backbone es solamente utilizar el protocolo PIM-SM para el registro de las fuentes de multicast. La definición del RP se realizará por auto-RP utilizando direcciones anycast para obtener redundancia de varios RP. En todos los enlaces del backbone se activará PIM-SM con un filtro de frontera en los enlaces a otros sistemas autónomos, como los muestra la figura 7-5.

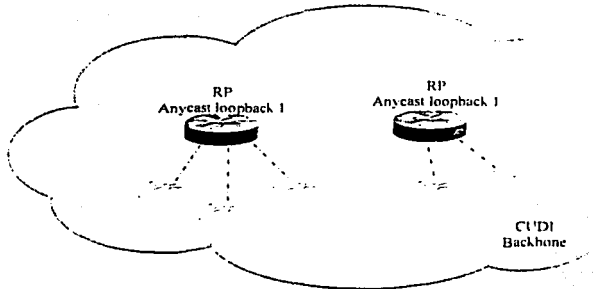


Figura 7-5 PIM-SM con direcciones de Anycast en las loopback para los RP.

En MBGP se siguen las mismas políticas de enrutamiento que se tienen para el caso del enrutamiento unicast, sirviendo de tránsito propagando los anuncios de los asociados y afiliados al resto de las redes de Internet2, como se muestra en la figura 7-6 los flujos que deben seguir los tráficos de multicast de Internet e Internet2.

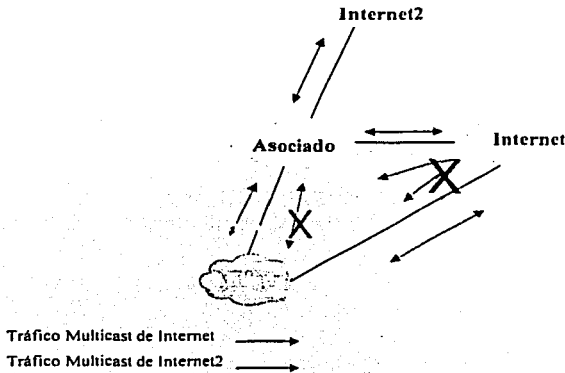


Figura 7-6 Flujo de los tráficos de Multicast de Internet e Internet2.

Cada uno de los asociados y afiliados que estén interesados en recibir o enviar tráfico multicast deberán definir su propio RP por el método que más les convenga y establecer una sesión de MSDP, los asociados con el backbone y los afiliados con los asociados. estas sesiones junto con las sesiones de MSDP deberán de contener los filtros recomendados de frontera. este escenario se ilustra en la figura 7-4.

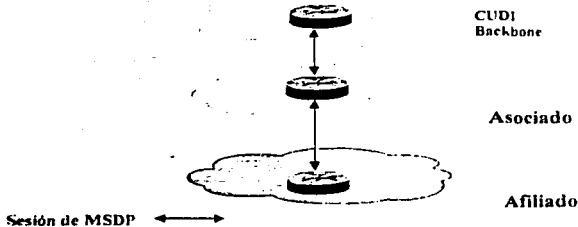


Figura 7-7 Relación en las sesiones de MSDP entre CUDI-asociados-afiliados.

### 7.5 Propuesta para el soporte de IPv6

En el capítulo 6 se mencionaron las principales características del protocolo IPv6, pero al igual que en los otros protocolos de enrutamiento unicast y multicast, en este capítulo mencionaremos la forma en que la red debe brindar el soporte de este protocolo, desafortunadamente aún faltan muchos protocolos por desarrollarse y que además puedan ser soportados por los equipos de telecomunicaciones, ya que por ejemplo, actualmente solo se cuenta con el protocolo RIP como IGP, aunque se están haciendo esfuerzos para desarrollar una versión de OSPF e ISIS para IPv6, de forma similar todavía no existe un protocolo para el intercambio de información entre ASs, por tal motivo, se hace uso de "MBGP extensiones para IPv6" para transportar los anuncios de los prefijos de IPv6 entre diferentes sistemas autónomos.

Actualmente todos los enlaces del backbone soportan tanto IPv4 como a IPv6, esto es, todo el backbone de CUDI esta soportando nativamente IPv6, es decir, sin túneles, incluso en las interfaces de loopbacks se tiene activo direccionamiento de IPv6 con mascararas de red con una longitud de 64 bits, la cual actualmente es la máxima longitud de mascara permitida, las direcciones que se están utilizando fueron delegadas por la UNAM a CUDI.

#### 7.5.1 RIPv6

Debido a que actualmente otros protocolos de enrutamiento de tipo link-state no son soportados por los routers del backbone, se ha tenido que recurrir al empleo de RIPv6 como protocolo de enrutamiento IGP.

El protocolo de enrutamiento se encuentra habilitado en todos los routers del backbone, actualmente la UNAM y el ITESM cuentan con enlaces nativos de IPv6, aunque también por parte de ABILENE se tiene la intención de poseer un enlace nativo en un futuro próximo.

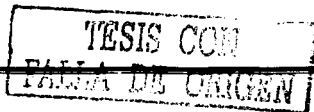
La configuración es muy sencilla ya que no existen muchas opciones de configuración, además de que la mayor parte de los valores se están dejando con sus opciones de default.

#### 7.4.2 MBGP extensiones para IPv6

Como ya se menciono anteriormente, para el intercambio de información no existe un protocolo de compuerta exterior, por lo cual se recurre al empleo de MBGP. El empleo de MBGP se realiza con las extensiones que le permiten transportar el anuncio de prefijos de otros protocolos de enrutamiento; en este caso IPv6, esta implementación no es la más adecuada, debido a que se depende de tener implementado a IPv4 en el backbone, quedando entonces todos los routers en una configuración conocida como dual-stack, se espera que muy pronto se tenga un protocolo de enrutamiento de compuerta exterior, aunque existen propuestas para que no se desarrolle ningún protocolo de este tipo, esto es debido a que IPv6 tiene una estructura jerárquica lo cual evitaría tener que utilizar un protocolo similar a BGP.

Para el enrutamiento interdomino, el backbone, los asociados y afiliados siguen las mismas políticas de tránsito de IPv4 a redes de IPv6 en Internet2 y se trata de evitar el tráfico comercial de IPv6 de Internet hacia las redes de Internet2.

En lo posible se trata de evitar el uso de túneles de IPv6 en IPv4 tanto con los asociados y afiliados como con el resto de las redes de Internet2, prefiriéndose los enlaces nativos de IPv6.



# Conclusiones

TESIS CON  
FALLA DE ORIGEN

Esta tesis cumplió con los objetivos de elaborar una propuesta de implementación de los protocolos de enrutamiento al incluir el soporte completo de todas las aplicaciones basadas en el protocolo IP. en Multicast y en la siguiente versión del protocolo IP (IPv6). De igual forma se cumplió con el segundo objetivo, que planteaba la creación de un documento que servirá de guía a las universidades y a cualquier otra entidad que se desee integrar al proyecto de Internet2.

La implementación de la propuesta elaborada permite que actualmente se ejecuten aplicaciones que en la red Internet no son posibles. hoy en día es posible ejecutar aplicaciones que emplean las características que ofrecen los protocolos propuestos que forman parte de la siguiente generación de Internet. Las ventajas que se obtuvieron como resultado de esta implementación son:

En el protocolo IP hoy en día existen muchas aplicaciones basadas en la actual versión que en el Internet comercial no son posibles de ejecutar. debido a las calidades de servicio y anchos de banda requeridos. algunos ejemplos de estas aplicaciones que actualmente ya se están ejecutando en la red Internet2 son: Voz y video sobre IP. Telemedicina. Teleimersión. modelado y simulación en tiempo real. bases de datos distribuidas. manipulación remota de telescopios. etc. Además de que también se da el soporte a las aplicaciones actuales del Internet comercial.

La propuesta de implementación de Multicast permite que cualquiera de los miembros de CUDI pueda enviar y/o recibir transmisiones de Multicast que anteriormente solo podían tener un alcance a nivel del campus universitario. Actualmente el alcance de las transmisiones de Multicast se extiende a cualquier universidad o institución conectada a la red Internet2 en México y en cualquier parte del Mundo. Algunas de las aplicaciones que actualmente se emplean en la red Internet2 son: transmisiones de audio y video de tipo broadcasting. access-grid. VoIP. H323. etc.

Se logro el soporte de la nueva generación del protocolo de Internet (IPv6). la importancia de que la propuesta soporte esta versión del protocolo IP. cobra importancia debido a que se

espera que en unos cuantos años reemplace a la actual versión y aunque faltan todavía muchos protocolos y aplicaciones por desarrollarse, México es uno de los países pioneros en la implementación de IPv6.

Finalmente todo el trabajo planteado en esta propuesta de implementación actualmente ya se encuentra en operación, soportando tanto a los protocolos de enrutamiento, así como a la forma en la que su implementación fue planteada, resultando ser coherente con las redes contraparte que integran el proyecto mundial llamado Internet2, con lo cual se crea una red nacional de tecnología avanzada, dedicada a la educación e investigación científica que inicialmente impulsen el desarrollo tecnológico en las universidades y organizaciones miembros de CUDI y posteriormente en el resto del país.



# Bibliografía

---

TESIS CON  
FALLA DE ORIGEN

Bibliografía

Internet Routing Architectures

Bassam Halabi

Cisco Press

CCIE Profesional Develoment

Routing TCP/IP

Volumen 1

Jeff Doyle

Cisco Press

CCIE Profesional Develoment

Routing TCP/IP

Volumen 2

Jeff Doyle

Cisco Press

CCNP Routing

Exam Certification guide

Gough

Cisco Press

Implementing IPv6

Mark A. Miller

M&T Book

IP Multicasting

"The complete guide to interactive

Corporate Networks"

Dave Kosiur

Wiley

TCP/IP illustrated

Volumen 1

Stevens

Adison Wesley

IP Multicasting

Concepts and applications

Marcus Goncalves

Mc Graw Hill

Developing IP Multicast Networks

Volumen 1

Beau Williamson

Cisco Press



OSPF Network Design Solution  
Thomas M.  
Cisco Press

Manual del curso de certificación  
"Building Scalable Cisco Networks"  
Cisco learning product

Manual del curso para ISP  
"ISP/IXP Workshop"

Links en Internet con referencia a la Tesis:

El fabricante de los equipos  
<http://www.cisco.com>

Paginas de Internet2  
<http://www.internet2.edu>  
<http://www.edu1.edu.mx>  
<http://www.internet2.unam.mx>  
<http://www.cenic.org>

Paginas de centros de operación  
<http://www.abilene.iu.edu>  
<http://www.vbns.net>  
<http://www.calren2.net>  
<http://www.noe-internet2.unam.mx>

Paginas de IPv6  
<http://www.6bone.net>  
<http://ipv6.internet2.edu>  
<http://www.ipv6.unam.mx>  
<http://www.abilene.iu.edu/doc/IPv6-cookbook.html>  
<http://www.m6bone.net>

Paginas de Multicast  
<http://www.abilene.iu.edu/mccook.html>  
<http://ftpeng.cisco.com/ftp/multicast/index.html>  
<http://www.internet2.edu/multicast>  
<http://www.nmsl.cs.ucsb.edu/mantra/routers/routers.html>

Otros sitios de Interés  
<http://www.nanog.com>  
<http://www.ietf.org>  
<http://www.atis.org/ig2k/1e2k.html>  
<http://www.ansi.org>  
<http://www.ieee.org>

# Glosario

---

TESIS CON  
FALLA DE ORIGEN

**ABR (Area Border Router).** Router de área fronteriza. Router ubicado en la frontera de una o más áreas OSPF que las conecta al backbone de la red.

**Acknowledgment.** Notificación enviada desde un dispositivo de red a otro, para confirmar que ha ocurrido algún evento (por ejemplo, la recepción de un mensaje).

**Actualización de Enrutamiento (routing update).** Es un mensaje enviado desde un router para indicar cuales son las rutas alcanzables de la red e información asociada con los costos. Las actualizaciones de enrutamiento se envían habitualmente a intervalos regulares y después de un cambio de topología en la red.

**Adyacencia (adjacency).** Es la relación que se establece entre routers vecinos seleccionados y nodos terminales, con el propósito de intercambiar información de enrutamiento. La adyacencia se basa en el uso de un segmento común.

**Algoritmo de Enrutamiento.** Mecanismo que determina la mejor ruta para enviar tráfico desde el origen hasta un destino en particular.

**Ancho de Banda (Bandwith).** Es la tasa máxima de transmisión en un medio físico determinado o en un protocolo dado.

**ANSI (American National Standards Institute).** Instituto nacional Americano de Estándares. Formado por voluntarios estadounidenses compuesto por miembros corporativos, de gobierno, y otras dependencias, que coordina las actividades relacionadas con los estándares.

**Anycast.** Son direcciones con las cuales se identifica a un grupo de hosts, pero solo uno de host, el más cercano es el que atiende las peticiones que se hacen al grupo.

**Área.** Es un conjunto lógico de segmentos de red (OSPF) y sus dispositivos conectados.

**Área non-stub.** Es un área de OSPF con grandes recursos que transporta una ruta predeterminada, rutas estáticas, rutas entre áreas, rutas intra-áreas y rutas externas. Son las únicas áreas de OSPF que pueden tener virtual-links configuradas a través de ellas y son las únicas áreas que pueden contener un ASBR.

**Área stub.** Área de OSPF que transporta una ruta predeterminada, rutas intra-área y rutas Inter.-área, pero no transportan rutas externas.

**ARP (Address Resolution Protocol).** Protocolo de Resolución de Direcciones. Protocolo de Internet que se usa para traducir una dirección IP a una dirección MAC. Está definido en el RFC 826.

**ARPANET (Advanced Research Projects Agency NET).** Red de la Agencia de Proyectos de Investigación Avanzada. Importante red de conmutación de paquetes establecida en 1969.

**ASBR (Autonomous System Border Router).** Router Fronterizo de Sistema Autónomo. El ABR se localiza entre un sistema autónomo OSPF y una red que no emplea OSPF. Los ASBRs ejecutan tanto el protocolo de enrutamiento OSPF como cualquier otro protocolo de enrutamiento. Los ASBRs residen en un área stub de OSPF.

**ASCII (American Standard Code for Information Interchange).** Código Estándar Americano para el Intercambio de Información. Es un código de 8 bits para la representación de caracteres (7 bits más el bit de paridad).

**ATM (Asynchronous Transfer Mode).** Modo de Transferencia Asíncrono. Estándar internacional para la conmutación de celdas, en el que se transportan varios tipos de servicios (voz, video y datos) por medio de celdas de longitud fija (53 bytes).

**Backbone.** Parte de la red responsable de conectar a todos los demás segmentos de red en una organización o empresa.

**BGP (Border Gateway Protocol).** Protocolo de Puerta de Enlace Fronteriza. Es un protocolo de enrutamiento entre dominios que reemplaza a EGP.

**BPDU (Bridge Packet Data Unit).** Unidad de Datos de Protocolo de Puente. Es un paquete hello del Spanning Tree Protocol.

**Broadcast.** Es un paquete de datos que se envía a todos los nodos de la red.

**CIDR (Classes InterDomain Routing).** Es un método enrutamiento donde no se toma en cuenta la definición de las clases de redes IP.

**CISCO.** Compañía fabricante de equipos de comunicación: switches, router, firewalls, etc.

**CIX (Comercial Internet eXchange point).** Punto de intercambio de tráfico comercial de Internet.

**Core.** Parte de la red donde cruza la mayor parte del tráfico de una red que normalmente coincide con el backbone.

**Costo.** Es un valor arbitrario, el cual típicamente se basa en el conteo de saltos, ancho de banda, y otras medidas, que es asignado a un enlace.

**CUDI.** Es la Corporación Universitaria para el Desarrollo de Internet en México.

**DARPA.** Defense Advanced Research Projects Agency

**Datagrama.** Es la unidad de datos de la capa de red.

**DCE.** Equipo de Comunicación de Datos. O equipo para la terminación de circuitos de datos. Son los dispositivos y conexiones de una red de comunicaciones que forman el extremo de red de la interfaz de usuario a red.

**Dirección IP.** Es una dirección de 32 bits asignada a los hosts que utilizan el protocolo TCP/IP.

**Dirección MAC.** Es la dirección estándar de la capa de enlace de datos que se requiere para cada dispositivo que se conecta a una LAN. Tiene una longitud de 6 bytes y esta controlada por la IEEE.

**DiffServ.** Metodología para la asignación de diferentes calidades de servicio basada en IP.

**DHCP Dinamic Host Configuration Protocol,** Protocolo de configuración dinámica de direcciones IP.

**Distancia Administrativa.** Es una medida de confiabilidad asignada a los anuncios de rutas que depende del protocolo de enrutamiento por el cual la ruta fue aprendida.

**DNS.** Sistema de Nombre de Dominios. Es utilizado en Internet para traducir los nombres de los nodos de red en direcciones IP.

**Downstream.** Termino utilizado para referenciar en una estructura jerárquica a los ISP que se encuentran en la parte inferior de esta.

**DTE.** Equipo Terminal de Datos. Dispositivo al extremo del usuario de una interfaz de usuario de red. que sirve como fuente de datos, destino o ambos. El DTE se conecta a una red a través de un dispositivo DCE.

**DVMRP (Distance Vector Multicast Routing Protocol).** Protocolo de Enrutamiento Multicast Basado en Vector Distancia. Protocolo de puerta de enlace de red. basado en gran parte en RIP.

**E1.** Esquema de transmisión digital de área amplia. utilizado principalmente en Europa para transportar datos a una velocidad de 2.048 Mbps.

**E3.** Esquema de transmisión digital de área amplia. utilizado principalmente en Europa para transportar datos a una velocidad de 34.368 Mbps.

**EBCDIC (Extended Binary Coded Decimal Interchange Code).** Código de Intercambio Decimal Codificado en Binario Extendido. Uno de varios grupos de caracteres codificados desarrollado por IBM. que constan de caracteres codificados de 8 bits.

**EBGP (Exterior BGP).** parte de protocolo de enrutamiento BGP que se emplea para comunicarse con vecinos externos al sistema autónomo.

**EGP (Exterior Gateway Protocol).** Protocolo de Puerta de Enlace Exterior. Protocolo de Internet para el intercambio de información de enrutamiento entre sistemas autónomos. Documentado en el RFC 904.

**EIGRP (Enhanced Interior Gateway Routing Protocol).** Protocolo de Enrutamiento de Puerta de Enlace Interior Mejorado. Es una versión avanzada de IGRP, desarrollada por CISCO. Presenta propiedades superiores de convergencia y eficiencia de operación.

**Encabezado (Header).** Información de control que se antepone en los datos cuando se van a encapsular para su transmisión a través de la red.

**Enrutamiento.** Proceso de búsqueda de una trayectoria hacia un host destino. **Estándar.** Es un conjunto de reglas y procedimientos ampliamente utilizados u oficialmente especificados.

**Ethernet.** Especificación de redes LAN. Las redes Ethernet emplean el método de acceso CSMA/CD y corren con una velocidad de 10 Mbps.

**FastEthernet.** Cualquier de las especificaciones Ethernet a 100 Mbps. Basada en la especificación IEEE 802.3.

**FDDI (Fiber Distributed Data Interface).** Interfase de Datos Distribuida por Fibra. Estándar LAN definido por la ANSI X3T9.5 que especifica una red de Token Ring a 100 Mbps que utiliza Fibra Óptica.

**FIX (Federal Internet eXchange point).** Punto de Intercambio de tráfico IP entre organismos Federales.

**Frame Relay.** Estándar Industrial, protocolo conmutado de la capa de enlace de datos que maneja circuitos virtuales múltiples utilizando encapsulamiento HDLC entre los dispositivos conectados.

**FTP (File Transference Protocol).** Protocolo de Transferencia de Archivos. Protocolo de aplicación, parte del stack de protocolos de TCP/IP, que se utiliza para la transferencia de archivos entre los nodos de la red. Definido en el RFC 959.

**Full mesh peering.** Arreglo topológico en el cual todos sus miembros tienen conexiones físicas o virtuales con todos los demás miembros.



**Gateway** (Puerta de Enlace). En la comunidad IP, es un termino con el que se hace referencia a un dispositivo de enrutamiento.

**GGP (Gateway-to-Gateway Protocol)**, Protocolo de Puerta de Enlace a Puerta Red Enlace. Protocolo MILNET que especifica la forma en que los routers principales deben de intercambiar la información de enrutamiento y alcance. Este protocolo utiliza un algoritmo distribuido de trayectoria más corta.

**Gigabit Ethernet**. Cualquiera de sus especificaciones a 1000Mbps

**GigaPop**. Sitio de comunicación con capacidades iguales o superiores a 1000Mbps en la tasa de transmisión de los equipos de comunicación.

**HDLC (High Data Link Protocol)**. Control de Enlace de Datos de Alto Nivel. Protocolo sincrónico de la capa de enlace orientado a bit desarrollado por la ISO.

**Host**. Dispositivo terminal alojado en una red. Este termino normalmente siempre implica un sistema de computo.

**IBGP Interior BGP**. parte de protocolo de enrutamiento BGP que se emplea para comunicarse con vecinos internos al sistema autónomo

**ICMP (Internet Control Message Protocol)**. Protocolo de Mensajes de Control de Internet. Protocolo de Internet de la capa de red que reporta errores y proporciona información relevante referente al procesamiento de paquetes.

**ICMPv6**. Protocolo de Mensajes de Control de Internet para IPv6. Es una versión modificada del ICMP para IPv6.

**IEEE (Institute of Electrical and Electronics Engineers)**. Instituto de Ingenieros en Electrónica y Electricidad. Organización profesional cuyas actividades incluyen el desarrollo de los estándares de comunicación y de redes.

**IEEE 802.3**. Protocolo IEEE LAN que especifica una implantación de la subcapa LLC de la capa de enlace de datos. Utiliza el método de acceso al medio CSMA/CD a una gran variedad de velocidades sobre diferentes medios de transmisión.

**IEEE 802.5**. Protocolo IEEE LAN que especifica una aplicación de la capa física y de la subcapa MAC de la capa de enlace de datos. Este estándar utiliza un método de acceso de token circulante a 4 o 16 Mbps. similar a Token Ring.

**IETF (Internet Engineering Task Force).** Fuerza de Trabajo de Ingeniería en Internet. Es una fuerza de trabajo que consiste en más de 80 grupos de trabajo responsables del desarrollo de estándares para Internet.

**IGMP (Internet Group Message Protocol).** Protocolo de membresía de Grupos de Internet. Este protocolo es utilizado por los hosts IP para reportar membresías de grupos de Multicast a un router de Multicast adyacente.

**IGP (Interior Gateway Protocol).** Protocolo de Puerta de Enlace Interior. Protocolo de Internet que se utiliza para el intercambio de información dentro de un sistema autónomo.

**IGRP (Interior Gateway Routing Protocol).** Protocolo de Enrutamiento de Puerta de Enlace Interior. IGP desarrollado por CISCO para resolver los problemas asociados con el enrutamiento de redes heterogéneas de gran tamaño.

**Interfaz.** Es una conexión de red.

**Interfaz de Loopback.** Interfaz lógica que se puede configurar en los equipos de comunicación como routers y switches las cuales no tienen dependencia de un medio físico.

**Internet2.** Es un consorcio dirigido por más de 190 universidades que trabajan en asociación con el gobierno y la industria para desarrollar e implementar aplicaciones y tecnologías avanzadas, acelerando la creación del Internet del mañana.

**Intranet.** Red privada que utiliza los protocolos TCP/IP.

**IP (Internet Protocol).** Protocolo de Internet. Protocolo de la capa de red en el conjunto de protocolos de TCP/IP que ofrece un servicio de red sin conexión. El protocolo IP proporciona características de direccionamiento, especificación del tipo de servicio, fragmentación y reensamblado y seguridad. Documentado en el RFC 791.

**IPv6.** Es la siguiente generación del Protocolo de Internet (IP).

**IS-IS (Intermediate System-to-Intermediate System).** Sistema Intermedio a Sistema Intermedio. Protocolo de enrutamiento jerárquico de OSI basado en el estado de enlaces y en el enrutamiento DECnet Fase V, donde los ISs (routers) intercambian información de enrutamiento basada en una sola medida para determinar la topología de la red.

**ISO (International Organization for Standardization).** Organización Internacional para la Estandarización. Organización internacional responsable de una amplia gama de estándares, incluyendo los pertinentes a las redes.

**ISP (Internet Service Provider).** Proveedor de Servicios de Internet.

**ITU-T.** Unión Internacional de Telecomunicaciones, sector de Estandarización de las telecomunicaciones. Organismo internacional que desarrolla estándares para las diferentes tecnologías de telecomunicaciones a nivel mundial.

**LAN (Local Area Network).** Red de Área Local. Red de datos de alta velocidad y baja tasa de errores, que cubre un área geográfica relativamente pequeña. Las LANs conectan estaciones de trabajo, periféricos, terminales y otros dispositivos en un solo edificio o área geográfica limitada.

**LSA (Link-state Advertisement).** Paquetes usados por los protocolos de estado de enlace que contienen la información acerca de vecinos y el costo de los caminos.

**Loops de enrutamiento.** Ocurren cuando los router o switches tiene información incorrecta acerca de la red y sucede que los paquetes en vez de ser enviados a sus destino son enviados a ellos mismo por una creencia errónea de los otros routers.

**MAC (Medium Access Control).** Control de Acceso a Medios. Es la subcapa inferior de las dos subcapas de la capa de enlace de datos definida por el IEEE. La subcapa MAC maneja el acceso al medio.

**Mascara de Red.** Combinación de 32 bits para indicar los bits de una dirección IP que se están utilizando para las direcciones de la subred. Conocida simplemente como mascara.

**MCI.** Compañía Norteamérica de Telecomunicaciones de Voz, datos y Video responsable de la administración de vBNS.

**Métrica.** Es una medida de confiabilidad asignada a un anuncio de enrutamiento. Se calcula por medio de una evaluación de ciertas variables tales como: Anchos de Banda, Retardo, Costo, Número de Saltos, Confiabilidad y Carga.

**Middleware.** Es un término general para cualquier programa que sirve para actuar como intermediario entre programas separados ya existentes

**MILNET.** Red militar. Porción no clasificada de la red de datos de la defensa nacional de los EE.UU.

**MINET.** Red militar europea similar a la red MILNET.

**MOSPF.** Protocolo de enrutamiento Multicast OSPF. Algoritmo de enrutamiento multicast entre dominios, empleado en las redes OSPF. Se aplican extensiones al protocolo base de unicast OSPF para soportar el enrutamiento IP multicast.

**MPLS-TE.** Multi Protocolo Label Switching – Traffic Engineering. Ingeniería de tráfico basada en etiquetas.

**MTU.** Maximun Trasnmission Unit, unidad de transferencia máxima en un enlace. medida en bits.

**Multicast.** Envío de paquetes de uno (o varios) a muchos.

**NAP (Network Access Point).** Un NAP esta definido como una red o punto de intercambio de alta velocidad al cual pueden estar conectados un cierto número de routers con el propósito de intercambiar tráfico.

**NIC (Network Interface Card).** Tarjeta de interfase de red. Es una tarjeta que proporciona capacidades de comunicación en la red hacia y desde un sistema.

**NREN (National Research and Education Network).** Red Nacional de Educación e Investigación. Es un componente del Programa HPCC (Comunicaciones y Computación de Alto Desempeño) diseñado para asegurar el liderazgo técnico de Estados Unidos en redes.

**NSF (National Science Foudation).** Fundación Nacional de Ciencias. Es una agencia del gobierno de Estados Unidos que financia la investigación científica en dicho país. La ahora desaparecida NSFNET fue financiada por la NSF.

**NSAP (Network Service Access Point).** Punto de Acceso al Servicio de Red. Son direcciones de red especificadas por la ISO. Un NSAP es el punto en el que el Servicio de Red de OSI se hace disponible a una entidad de la capa de transporte.

**NSFNET.** Red de la Fundación Nacional de la Ciencia. Es una gran red que fue controlada por la NSF y proporciono servicios de red para ayudar a la educación y la investigación en Estados Unidos de 1986 a 1995. Esta red actualmente ya no esta en servicio.

**NSP (Network Service Provider).** También es usado con menor frecuencia para hacer referencia a un proveedor que esta conectado a todos los NAP.

**OSI (Open System Interconnection).** Modelo de arquitectura de red. desarrollado por la ISO e ITU-T. y un modelo de referencia de interconexión de sistemas abiertos. El modelo consta de siete capas (Aplicación, Presentación, Sesión, Transporte, Red, Enlace de datos y Física) , cada una de las cuales especifica funciones particulares de red como direccionamiento, control de flujo, control de errores, encapsulamiento y transferencia confiable de mensajes.

**OSPF (Open Shortest Path First).** Algoritmo de enrutamiento jerárquico IGP basado en el estado de enlaces, propuesto como sucesor del protocolo de enrutamiento RIP en la comunidad de Internet. Entre las características de OSPF están el enrutamiento de menor costo, el enrutamiento multirayectoria y el balanceo de carga. OSPF se derivó de la versión antigua del protocolo IS-IS.

**Overhead.** Es información de control que requieren los equipos de comunicaciones, pero que no es parte de la información de los usuarios. Es un parámetro que se debe de minimizar debido a que consume recursos de la red.

**Paquete.** Es la unidad de datos de la capa de red.

**Payload.** Es la porción de una trama que contiene información (datos) de las capas superiores.

**PDU (Packet Data Unit).** Unidad de Datos de Protocolo. Término de la OSI para los paquetes.

**Peer.** Par de routers que han establecido una sesión de BGP o MSDP.

**PIM (Protocol Independent Multicast).** Protocolo de enrutamiento Multicast independiente del protocolo de enrutamiento Unicast, puede ser en modo denso o en modo disperso.

**POP (Point of Presence).** Es el punto físico de acceso a un Proveedor de Servicios de Internet.

**PPP (Point to Point Protocol).** Protocolo Punto a Punto. Sucesor del protocolo SLIP que ofrece conexiones entre un router y host a red sobre circuitos síncronos y asíncronos.

**Protocolo.** Descripción formal de un conjunto de reglas que rigen el modo en el que intercambian información los dispositivos de red.

**Query.** Petición o solicitud.

**QoS (Quality of Service).** Calidad de Servicio. Es la habilidad de administrar los recursos de la red para proporcionar diferentes servicios a diferentes aplicaciones.

**RA (Router Arbiter).** El RA debe de proveer una base de datos con la información de enrutamiento para proporcionar escalabilidad y administrabilidad a las redes.

**RARP (Reverse Address Resolution Protocol).** Protocolo de Resolución Inversa de Direcciones. Es el protocolo en el stack de TCP/IP que proporciona un método para encontrar direcciones IP con base en direcciones MAC.

**Red de Datos.** Red de comunicación basada en el envío de paquetes.

**Redistribución.** Permite la distribución de la información de enrutamiento, descubierta a través de un protocolo de enrutamiento, en los mensajes de actualización de otro protocolo de enrutamiento. Comúnmente llamado redistribución de rutas.

**RFC (Request for Comments).** Solicitud de comentarios. Es una serie de documentos que se utiliza como la manera principal para comunicar información con respecto a Internet. Algunos RFCs están designados por el IAB como estándares de Internet. La mayoría de los RFCs documentan especificaciones de protocolos.

**RIP (Routing Information Protocol).** Protocolo de Información de Enrutamiento. Es el IGP más común en Internet. El protocolo RIP utiliza el conteo de saltos como una medida de enrutamiento.

**RIPv2.** Es la segunda versión del protocolo RIP la cual tiene un comportamiento de CIDR y puede utilizar Multicast entre sus principales diferencias con la versión anterior.

**RIPv6.** Es la versión modificada del protocolo RIP para soportar el anuncio de redes de IPv6.

**RIR (Register IP Regional).** Organismo responsable del registro de las direcciones IP.

**Router.** Dispositivo de la capa de red que utiliza una o más medidas para determinar la mejor trayectoria a lo largo de la cual se avanzará el tráfico de la red. Los routers avanzan información de una red a otra con base a la información de la capa de red.

**Salto (hop).** Término que describe el tránsito de un paquete de datos entre dos nodos. Por ejemplo entre dos routers.

**Segmentos.** Unidad de datos en la capa de Transporte. Sección de una red que se encuentra delimitada por bridges, switches o routers.

**Sistema Autónomo (AS).** Es un conjunto de redes bajo una administración común, que comparte políticas de enrutamiento estratégicas.

**SMTP (Simple Mail Transport Protocol).** Protocolo Simple de Transferencia de Correo. Es un protocolo de Internet que proporciona servicios de correo electrónico.

**SNMP (Simple Network Management Protocol).** Protocolo Simple de Administración de Red. Protocolo de administración de red utilizado casi exclusivamente por las redes TCP/IP. Este protocolo representa un medio para supervisar y controlar los dispositivos de una red.

**SPF (Shortest Path First).** Algoritmo Abierto Primero la Trayectoria más Corta. Es un algoritmo de enrutamiento que itera según la longitud de la trayectoria. para determinar un árbol de recubrimiento de trayectoria más corta. Comúnmente se utiliza con algoritmos de enrutamiento basados en estado de enlaces (link-state).

**SPT (Shortest Path Tree).** Árbol de distribución de tráfico Multicast que utiliza la trayectoria más corta a la fuente de Multicast.

**SPX (Secuence Packet eXchange).** Protocolo de capa cuatro utilizado en redes Novell.

**Stack.** Conjunto de protocolos de comunicación que operan juntos y, como grupo, tiene que ver con la comunicación en alguna o en todas las capas del modelo de referencia OSI. No todas los stacks de protocolos cubren las siete capas del modelo OSI, y con frecuencia un solos protocolo en snack puede cubrir varias capas en una sola, tal es el caso de TCP/IP.

**STM-1 (Synchronous Transfer Module).** Modulo de Trasferencia Sincrono nivel 1. Es uno de los diferentes formatos de SDH que especifica la estructura de trama para las líneas con una velocidad de transmisión de 155 Mbps.

**STP (Spanning Tree Protocol).** Es un protocolo de la capa de enlace en redes Ethernet que utiliza el algoritmo de árbol de recubrimiento y evita ciclos en una topología de red.

**Subred.** En redes IP, una red que comparte una dirección de subred particular. Segmento de red que proporciona una estructura de enrutamiento jerárquico.

**Switch.** Dispositivo de red que filtra, direcciona y avanza tramas con base en la dirección de destino de cada trama. El switch opera a nivel de la capa de enlace de datos del modelo OSI.

**T1.** Instalación de Transporte en un red WAN. Transmite datos formateados en DS-1 a una velocidad de 1.544 Mbps , es el estándar americano.

**T3.** Instalación de Transporte en un red WAN. Transmite datos formateados en DS-3 a una velocidad de 44.736 Mbps.

**Tabla de Enrutamiento.** Es una tabla que se encuentra almacenada en los routers o en algún otro dispositivo de red que mantiene un registro de las rutas hacia los destinos de red particulares y las métricas asociadas con esas rutas.

**TCP (Transmission Control Protocol).** Protocolo de Control de la Transmisión. Protocolo orientado a conexión que pertenece a la capa de transporte y que ofrece una transmisión confiable de datos duplex total.

**TCP/IP (Transmission Control Protocol/Internet Protocol).** Protocolo de Control de la Transmisión/Protocolo Internet. Es el nombre común que se utiliza para nombrar al conjunto de protocolos desarrollado por el Departamento de Defensa de Estados Unidos en los años 70 para soportar redes con cobertura mundial.

**Tier.** Nivel de jerarquía de un ISP.

**Token Ring.** Es una red LAN con protocolo de acceso de token circulante desarrollada y soportada por IBM. La red Token Ring corre a 4 o 16 Mbps sobre una topología en anillo. Es similar al estándar 802.5.

**Topología.** Es el arreglo físico de los nodos y el medio de transmisión dentro de una estructura de red.

**ToS (Type of Service).** Indicación de cómo requiere un protocolo de las capas superiores que un protocolo de las capas inferiores trate sus mensajes.

**Trama.** Unidad de datos de la capa de enlace del modelo OSI.

**Trailer.** Campo al final.

**TTL (Time to Live).** Tiempo de Vida. Es un campo dentro del encabezado IP que indica el periodo dentro del cual se considera válido un paquete. Medido en saltos.

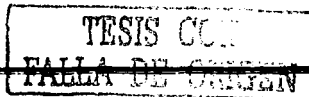
**UDP (User Datagram Protocol).** Protocolo de Datagrama de Usuario. Es el protocolo de la capa de transporte no orientado a la conexión, en el stack de protocolos de TCP/IP. UDP es un protocolo simple que intercambia datagramas sin reconocimiento o entregas garantizadas, y por ello requiere que el procesamiento de errores y la retransmisión sean manejados por otros protocolos. Especificado en el RFC 768.

**Unicast.** Envío de paquetes de uno a uno.

**Upstream.** Término utilizado para referenciar en una estructura jerárquica a los ISP que se encuentran en la parte superior de esta.

**vBNS (Very high-speed Backbone Network Services).** Red formada con enlaces de alta velocidad creada para reemplazar a la NSFNET.

**VLSM (Variable Length Subnet Mask).** Máscara de Subred de Longitud Variable. Es la habilidad de especificar una máscara de subred diferente para el mismo número de red en diferentes subredes. Ayuda a optimizar el espacio disponible para las direcciones IP.





**WAN (Wide Area Network).** Red de área amplia. Es una red de conmutación de datos que da servicio a usuarios localizados en una amplia área geográfica y son redes de menor velocidad en comparación con las redes LAN.

**X.25.** Estándar de la ITU-T que define la forma en como se mantienen las conexiones entre DTE y DCE para el acceso a terminales remotas. El protocolo X.25 especifica a LAPB, a un protocolo de 1 capa de enlace de datos, y a PLP, un protocolo de la capa de red.

TESIS CON  
FALLA DE ORIGEN

# Anexo A

TESIS CON  
FALLA DE ORIGEN

## Anexo A

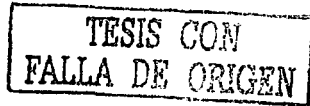
Este anexo contiene solamente las instrucciones para configurar los diferentes protocolos mencionados durante el desarrollo de la tesis y no pretende ser una guía completa de la configuración de un router para esto sugerimos revisar las páginas mencionadas en la bibliografía.

Plantilla de configuración para una router del backbone:

Configuración de OSPF

- Habilitar el proceso de OSPF.
- Habilitar el log del proceso de OSPF.
- Modificar la referencia del costo a el enlace de mayor ancho de banda.
- Poner en modo pasivo los enlaces a los vecinos de otros A.S.
- Agregar las redes de la infraestructura al proceso de OSPF.
- Definir el router ID.

```
router ospf 1
log-adjacency-changes
auto-cost reference-bandwidth 1000
passive-interface FastEthernet0/0
passive-interface ATM1/0.2
passive-interface POS1/0/0
passive-interface POS1/0/2
network 192.168.0.0 0.0.0.255 area 0
network 200.23.60.1 0.0.0.0 area 0
router-id 198.168.0.1
```



Configuración para las sesiones de IBGP

- Habilitar el proceso de BGP
- Definir los peer-groups
- Agregar una descripción
- Definir el vecino como un tipo interno. Los vecinos interno se definen al especificar el mismo número que el usado en el proceso de BGP.
- Establecer las sesiones usando interfaces de loopback
- Usar el comando de "Next-hop-self"
- Enviar las comunidades a todos los vecinos
- Especificar siempre la versión 4 de BGP.
- Usar passwords en cada sesión de IBGP
- Habilitar el uso de "soft-reconfiguration inbound"

```
router bgp 65001
neighbor VecinoIBGP peer-group
neighbor VecinoIBGP Description Vecinos IBGP
neighbor VecinoIBGP remote-as 65001
neighbor VecinoIBGP update-source loopback0
neighbor VecinoIBGP next-hop-self
```

```
neighbor VecinoIBGP send-community
neighbor VecinoIBGP version 4
neighbor VecinoIBGP password 7 04095E09
neighbor VecinoIBGP soft-reconfiguration inbound
neighbor 192.168.1.1 peer-group VecinoIBGP
neighbor 192.168.1.5 peer-group VecinoIBGP
```

#### Configuración para las sesiones de EBGp

- Definir el vecino con su número de AS
- Agregar una descripción del vecino
- Remover los sistemas autónomos privados
- Definir los filtros, tan extensivos como sean posible
- Asignar el password previamente acordado a la sesión de EBGp
- Limitar el número máximo de anuncios recibidos en cada sesión
- Habilitar el uso de "soft-reconfiguration inbound"

```
router bgp 65001
neighbor 192.168.2.1 remote-as 65002
neighbor 192.168.2.1 description SESION BGP CON EL VECINO X
neighbor 192.168.2.1 remove-private-AS
neighbor 192.168.2.1 prefix-list FiltroPrefix in
neighbor 192.168.2.1 soft-reconfiguration inbound
neighbor 192.168.2.1 maximum-prefix (# de prefijos)
neighbor 192.168.2.1 soft-reconfiguration inbound
```

**TESIS CON  
FALLA DE ORIGEN**

#### Configuración global al proceso de BGP

- Habilitar dampening
- Habilitar el log del cambio de estado de los vecinos
- Habilitar MED en modo determinístico
- Deshabilitar la sincronización
- Deshabilitar la sumarización automática de rutas
- Definir el router ID
- Agregar al proceso de BGP el bloque de direcciones del AS
- Agregar al proceso los bloques de direcciones propias de los clientes que no hablan BGP.

```
Router bgp 65001
bgp dampening route-map FLAPSPenalizacion
bgp log-neighbor-changes
bgp always-compare-med
bgp deterministic-med
no synchronization
no auto-summary
bgp router-id 192.168.0.1
network 192.168.0.0 mask 255.255.0.0
network 172.16.1.0 mask 255.255.255.0
```

network 172.16.2.0 mask 255.255.255.0

### Configuración para Multicast

- Habilitar globalmente el enrutamiento Multicast.
- Habilitar PIM-SM en las interfaces, definiendo el filtro de frontera en cada enlace hacia otros AS.
- Definir el RP del sistema autónomo.
- Habilitar la sesión de MBGP para el soporte de los prefijos de Multicast.
- Habilitar la sesión de MSDP.

*!Habilitando el enrutamiento Multicast*

*ip multicast-routing*

!

*!configuración de PIM-SM a nivel de Interfaz*

*interface ATM0/0.1 point-to-point*

*ip address 192.168.1.2 255.255.255.252*

*!Modo de operación de la interfaz*

*ip pim sparse-mode*

*ip multicast boundary multicast-boundary*

*!Filtro de Frontera*

*ip access-list standard multicast-boundary*

*deny 224.0.1.39*

*deny 224.0.1.40*

*deny 239.0.0.0 0.255.255.255*

*permit any*

*!Definición del RP*

*ip pim send-rp-announce Loopback0 scope 32*

*ip pim send-rp-discovery Loopback0 scope 32*

!

*!Configuración de MBGP*

*neighbor 192.168.1.1 remote-as 65001*

*address-family ipv4 unicast*

*neighbor 192.168.1.1 activate*

*neighbor 192.168.1.4 remote-as 65002*

*address-family ipv4 multicast*

*neighbor 192.168.1.4 activate*

!

*!Configuración de MSDP*

*ip msdp peer 192.168.1.1 connect-source Loopback0*

*ip msdp sa-filter out 192.168.1.1 list 111*

!

*!Filtro recomendado para MSDP (ACL 111)*

*access-list 111 deny ip any host 224.0.2.2*

*access-list 111 deny ip any host 224.0.1.3*

*access-list 111 deny ip any host 224.0.1.24*

*access-list 111 deny ip any host 224.0.1.22*

TESIS CON  
FALLA DE ORIGEN

```

access-list 111 deny ip any host 224.0.1.2
access-list 111 deny ip any host 224.0.1.35
access-list 111 deny ip any host 224.0.1.60
access-list 111 deny ip any host 224.0.1.39
access-list 111 deny ip any host 224.0.1.40
access-list 111 deny ip any 239.0.0.0 0.255.255.255
access-list 111 deny ip 10.0.0.0 0.255.255.255 any
access-list 111 deny ip 127.0.0.0 0.255.255.255 any
access-list 111 deny ip 172.16.0.0 0.15.255.255 any
access-list 111 deny ip 192.168.0.0 0.0.255.255 any
access-list 111 permit ip any any

```

#### Configuración para soporte de IPv6

- Habilitar el enrutamiento de IPv6.
- Configurar IPv6 en las interfaces
- Habilitar el proceso de RIPv6
- Habilitar la sesión de MBGP extensiones de IPv6

!Habilitando el enrutamiento de IPv6

```
ipv6 unicast-routing
```

!

!Habilitando IPv6 en las interfaces

```
interface Loopback0
```

```
ipv6 address 2001:448:3:55::1/64
```

```
ipv6 enable
```

!Agregando al proceso de RIPv6

```
ipv6 rip CUDI enable
```

!

```
interface ATM1/0.1 point-to-point
```

```
ipv6 address 2001:448:3:65::2/64
```

```
ipv6 rip CUDI enable
```

!Habilitando el proceso de RIPv6

```
ipv6 router rip CUDI
```

!

```
router bgp 65001
```

!Habilitando MBGP extensiones para IPv6

```
address-family ipv6
```

!Vecino IBGP

```
neighbor 2001:448:3:66::1 remote-as 65001
```

```
neighbor 2001:448:3:66::1 description IBGP_IPv6
```

```
neighbor 2001:448:3:66::1 activate
```

```
neighbor 2001:448:3:66::1 soft-reconfiguration inbound
```

```
neighbor 2001:448:3:66::1 filter-list 10 in
```

!Vecino EBGP

```
neighbor 2001:448:3:3:135::1 remote-as 65002
```

```
neighbor 2001:448:3:135::1 description IBGP_IPv6
```

TESIS CON  
FALLA DE ORIGEN

*neighbor 2001:448:3:135::1 activate*  
*neighbor 2001:448:3:135::1 soft-reconfiguration inbound*  
*neighbor 2001:448:3:135::1 filter-list 12 in*  
*!Bloque de direcciones del AS 65001*  
*network 2001:448:3:55::1/64*  
*aggregate-address 2001:448:3::/48 summary-only*  
*exit-address-family*

TESIS CON  
FALLA DE ORIGEN