UNIVERSIDAD NACIONAL DE INGENIERÍA

FACULTAD DE INGENIERÍA ELÉCTRICA Y ELECTRÓNICA

ARQUITECTURA DE ENRUTAMIENTO EN INTERNET

INFORME DE SUFICIENCIA

PARA OPTAR EL TÍTULO PROFESIONAL DE:

INGENIERO ELECTRÓNICO

PRESENTADO POR:

WILMER JESÚS DÍAZ OLIVET

PROMOCIÓN
1999 - II

LIMA – PERÚ
2005

ARQUITECTURA DE ENRUTAMIENTO EN INTERNET

DEDICO ESTE TRABAJO A:
MIS PADRES, POR SU APOYO
INCONDICIONAL EN MI CARRERA.

SUMARIO

El objetivo principal de este informe es comprender cómo funciona el enrutamiento en Internet el cual esta basado en el protocolo BGP utilizado para el intercambio de información de encaminamiento entre diferentes sistemas autónomos, los requerimientos necesarios para establecer una sesión BGP e IBGP, anunciar redes, manipulación de atributos, filtrar prefijos, crear mapas de rutas, confederaciones, reflectores de rutas, el NAP Perú, etc.

Enfocándolo al diseño y aplicaciones prácticas (casos de Telefónica del Perú), proporcionando soluciones reales a problemas de conectividad con los ISP. Aprender cómo integrar una red en la Internet global y construir sistemas autónomos de gran tamaño, a diseñar redes seguras y estables aplicando las políticas de enrutamiento requeridas enmarcándolas en las reglas que se manejan en Internet.

ÍNDICE

CAPÍTULO II

FUNDAMENTOS DE LOS PROTOCOLOS DE ENRUTAMIETO

CAPÍTULO III

PROTOCOLO BGP-4

CAPÍTULO IV

FUNCIONAMIENTO Y MANEJO DE ATRIBUTOS BGP

CAPÍTULO V

MANEJO DE ATRIBUTOS BGP

CAPÍTULO VI

CONTROL DE ENRUTAMIENTO DENTRO DE SISTEMAS AUTÓNOMOS

CAPÍTULO VII

CASOS PRÁCTICOS, ENRUTAMIENTO EN EL PERÚ

CAPÍTULO VIII

IMPLEMENTACIÓN DE UN ROUTER EN LINUX DESARROLLADO POR EL GRADUANDO

PRÓLOGO

Internet es una gran red de computadoras capaces de intercambiar información gracias a que utilizan un protocolo común de comunicaciones conocido como TCP/IP.

Esta red está formada por miles de redes de diversos tamaños interconectadas entre sí, y que están distribuidas por todo el mundo. A través de esta red de redes es posible intercambiar datos, voz, sonido e imágenes, lo cual crea un mundo virtual en el que las distancias entre los usuarios se acortan y la información fluye a todos los terminales conectados a la red.

En este contexto, hay que distinguir dos mercados: el mercado de infraestructura de acceso al ISP (*Internet Service Provider*); y el mercado del servicio de acceso a Internet. Esta distinción es desde el punto de vista del usuario de Internet, el cual debe usar algún medio para "llegar" al ISP, quien le proporciona el servicio de acceso a Internet propiamente dicho.

Dado que lo que transmiten son datos (paquetes TCP/IP) en formato digital, el acceso puede proveerlo cualquier empresa operadora de una red con capacidad de transmisión digital de datos.

CAPÍTULO I

INTRODUCCIÓN A LA EVOLUCIÓN DE INTERNET.

1.1 INTRODUCCIÓN.

El mercado del servicio de acceso a Internet está constituido principalmente por empresas que poseen una conexión con la red internacional (proveedores de Internet en Estados Unidos) y/o local de Internet (p.e. NAP local u otro ISP local) y ofrecen a los usuarios finales sus servicios de acceso a los contenidos de Internet. Estas empresas son las denominadas ISPs. Estos ISPs poseen dos necesidades de conexión: con la red Internet y con los usuarios que desean acceder a Internet.

Por otro lado, en el mercado peruano también existen pequeñas empresas o personas individuales que brindan el servicio de cabinas públicas de Internet, mediante las cuales los usuarios, después del pago respectivo, puede acceder a Internet. Las cabinas públicas de acceso a Internet constituyen principalmente un medio de acceso de los usuarios residenciales.

1.2 TÉCNICAS DE DIRECCIONAMIENTO Y ASIGNACIÓN IP.

Las estrategias de direccionamiento son de importancia directa y fundamental para la arquitectura de enrutamiento de cualquier red. Una de las funciones básicas de la arquitectura de enrutamiento y de los routers es alojar direcciones para todo el tráfico que dirigen. Con el explosivo crecimiento de Internet y del número de direcciones y la evolución de las nuevas estrategias de direccionamiento, se han presentado nuevos desafíos para las arquitecturas de enrutamiento. Una comprensión de la historia y los fundamentos del direccionamiento IP jugará sin duda un papel clave para que pueda asimilar rápidamente los conceptos de los protocolos de enrutamiento.

1.2.1 Direccionamiento IP básico.

Una dirección IP es un valor único de 4 octetos (32 bits) expresado en notación decimal con puntos de la forma W.X.Y.Z, donde los puntos se utilizan para separar cada uno de los 4 octetos de la dirección (por ejemplo, 10.0.0.1). El campo dirección de 32 bits consta de dos partes: un número de red o enlace (que representa la parte de red de la dirección) y un número de host (que identifica un host en el segmento de red).

Los límites de la red y el host se definían tradicionalmente basándose en la clase de la dirección IP, con cinco clases definidas (tres de las cuales se utilizan para direccionamiento de unidifusión): A, B, C, D y E.

Este esquema de direccionamiento basado en clases a menudo se conoce como modelo con clase. Las diferentes clases se prestan a diferentes configuraciones de red, dependiendo de la proporción deseada entre redes y hosts

1.2.2 VLSM.

El término **máscara de subred de longitud variable** (VLSM) hace referencia al hecho de que una red puede configurarse con diferentes máscaras. La idea básica tras las VLSM es ofrecer más flexibilidad al dividir una red en múltiples subredes, a la vez que se optimiza la asignación de cantidades variables de espacio en host entre las subredes. Sin VLSM, sólo puede aplicarse una máscara de subred a toda una red. Esto restringiría el número de hosts dado el número de subredes requeridas. Si selecciona la máscara de modo que tenga suficientes subredes, quizá no podría asignar suficientes números de host en cada subred. Lo mismo es cierto para los hosts; una máscara que permite suficientes hosts podría no proporcionar suficiente espacio de subred. VLSM proporciona la capacidad de asignar subredes con cantidades variables de hosts, permitiendo al administrador de red utilizar mejor el espacio de direcciones.

Supongamos, por ejemplo, que le asignan una red de Clase C 192.214.11.0 y que necesita dividirla en tres subredes. Una subred requerirá 100 números de host, y las otras dos requerirán 50 números de host cada una. Ignorando los limites de los dos extremos de red (0, número de red, y 255, dirección de difusión directa), teóricamente tendrá 256 números de host disponibles, desde 192.214.11.0 hasta 192.214.11.255. Como veremos, la deseada división en subredes no puede llevarse a cabo sin VLSM.

Para determinar las opciones de subred disponibles asociadas con la red 192.214.11. primero necesitará identificar la máscara de red, que en caso de esta tradicional red de Clase C está representada por 25.5.255.255.0 (todo unos en los primeros tres octetos). Se puede utilizar un puñado de máscaras de subred de la forma

255.255.255.X para dividir la red de Clase C 192.214.11.0 en más subredes. Una máscara debe tener un número continuo de bits 1, comenzando por el bit situado más a la izquierda, mientras que los bits restantes deben ser 0.

1.2.3 Agotamiento del espacio de direcciones IP.

La creciente demanda de direcciones IP ha supuesto una severa molestia en el modelo con clase. La mayoría de las empresas que solicitan direcciones de Clase B han determinado que una dirección de Clase B satisfacería mejor sus necesidades por el equilibro entre el número de redes y el número de hosts que ofrece. Una dirección de Clase A suele ser excesiva, con más de 16 millones de hosts, y una de Clase C tiene muy pocos hosts por red. Hacia 1991, comenzaba a ser obvio que el consumo de Clase B no estaba disminuyendo y era necesario tornar medidas para evitar su agotamiento.

Algunas de estas medidas consistían en asignaciones creativas de direcciones IP y en promover el uso de direcciones privadas en, empresas que no tenían conectividad global a Internet. Otras medidas dieron como resultado el inicio de grupos de trabajo y oficinas de dirección corno el grupo de trabajo de Enrutamiento y direccionamiento (ROAD) y la oficina de dirección de la próxima generación de IP (IPng). En 1992, el grupo de trabajo ROAD propuso el uso de enrutamiento entre dominios sin clase (CIDR) como alternativa al direccionamiento IP con clase. Al mismo tiempo, la oficina de dirección IPng estaba trabajando en el desarrollo de un nuevo esquema de direccionamiento IP mejorado que utilizara IP, versión 6 (IPv6), que con el tiempo solucionaría los problemas de uso con los que se ha encontrado el direccionamiento lPv4.

Las medidas para solucionar el agotamiento del direccionamiento IP pueden agruparse en las cuatro categorías siguientes:

- Asignación creativa de direcciones IP.
- Enrutamiento entre dominios sin clase (CIDR).
- Direccionamiento IP privado y Traducción de Direcciones de Red (NAT).
- IP versión 6 (IPv6).

Además de la preocupación por el agotamiento, la creciente demanda de direcciones IP generó la necesidad de convertir el proceso de asignación del direccionamiento IP desde un registro central. Originalmente, la IANA y el Registro en Internet (IR) tenían un control completo de la asignación de direcciones. Las direcciones IP eran asignadas a las organizaciones secuencialmente sin considerar los factores geográficos y de cómo y dónde una organización se conectara a Internet. Este método tenía el efecto de crear huecos en el espacio de direcciones IP (segregando números individuales o pequeños de direcciones IP y eliminando rangos de números grandes, contiguos).

Era necesario un enfoque diferente, en el que se otorgasen rangos contiguos de direcciones a administradores diferentes (como los proveedores de servicios), y dichos proveedores de servicio asignasen por turno direcciones a los clientes de su propio espacio. En general, este método de asignación de direcciones pronostica un método de distribución de direcciones IP más controlado y jerárquico. Es de algún modo similar al enfoque empleado por el esquema de asignación de la red telefónica, donde los códigos de área están asociados a regiones (redes de proveedores de servicios), los prefijos con subconjuntos de dichas regiones (dientes de los proveedores de servicios) y el resto con clientes individuales (hosts).

1.2.4 CIDR.

En los últimos años, las tablas del enrutamiento IP global han crecido de tal forma que los routers comenzarán a saturarse debido a la potencia de procesamiento y la asignación de memoria. Las estadísticas y las proyecciones medias de crecimiento sugieren que las tablas de enrutamiento doblaron su tamaño cada diez meses entre 1991 y 1995, y han crecido significativamente desde 1998. La Figura siguiente ilustra este crecimiento.

Sin ningún plan de acción, la tabla de enrutamiento habría crecido hasta unas 80.000 rutas en 1995. Sin embargó, los datos reales a principios de 2000 mostraban que el tamaño de la tabla de enrutamiento era aproximadamente de 76.000 rutas. Esta reducción en el crecimiento se debe al esquema de asignación de direcciones IP y a CIDR

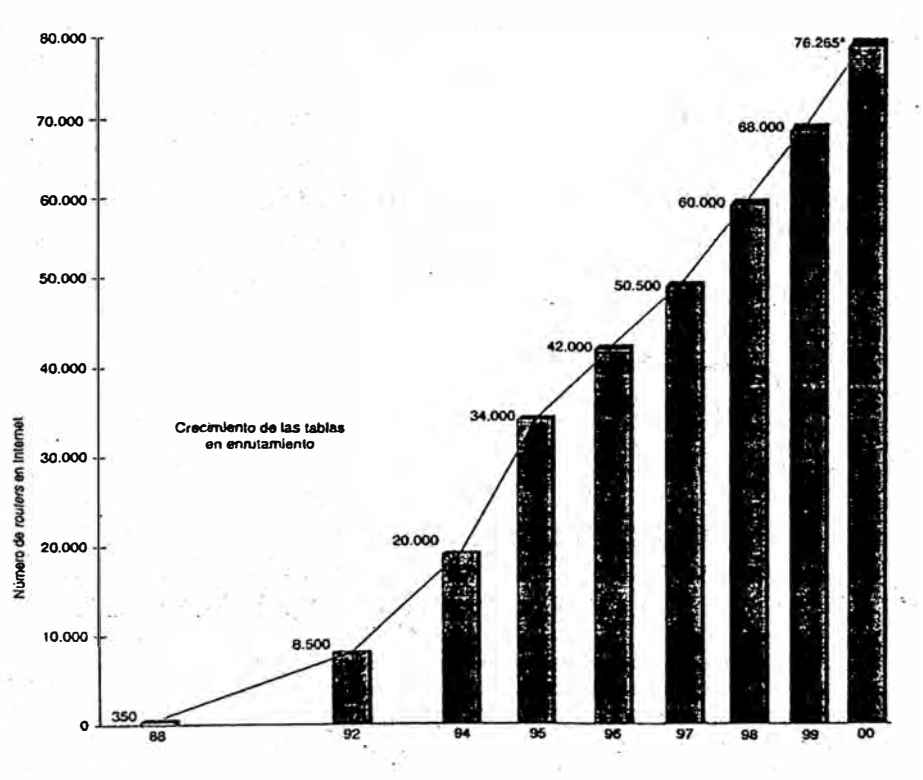


Figura 1.1 Crecimiento de las tablas de enrutamieto.

CIDR fue un paso evolutivo más allá de las tradicionales direcciones IP con clase; a saber, redes clase A, B y C. Con CIDR, una red IP se representa mediante un prefijo, que es la dirección IP de una red, seguido de una barra, y por ultimo una indicación de números de bits contiguos más a la izquierda correspondientes a la mascara de red asociada con dicha dirección de red.

Una red se denomina superred cuando el límite de la mascara de prefijo contiene menos bits que una mascara natural de red. Esta notación proporciona un mecanismo para reunir fácilmente todas las rutas más especificas (ejem. 198.32.0.0/16 contiene a 198.32.0.0, 198.32.1.0, 198.32.2.0 ... etc.) en una publicación denominada **agregada**.

1.3 ORGANIZACIONES IMPORTANTES.

1.3.1 IANA.

(Internet Assigned Numbers Authority). Es la autoridad de Asignación de Números en Internet. Se trata de la entidad que gestiona la asignación de direcciones IP en Internet. Incorporada a ICANN en 1999 El IANA es el organismo de la ISOC (Internet Society http://info.isoc.org/) de la administración de las direcciones Internet (direcciones IP) así como de la creación de nuevos dominios (DNS) (Actualmente se encuentra en estudio la creación de nuevos dominios como inc, co etc). La IANA delega la asignación de dominios ya creados a la InterNIC.

El IANA delega parte de su responsabilidad a un Internet Registry , el cual actúa como una central que almacena información de Internet y que proporciona alojamiento central a las redes y sistemas autónomos de usuarios. La IR (Internet Registry) se encarga asimismo de proveer mantenimiento central al DNS (Domain Name System), base de datos que apunta a la distribución subsidiada de servidores de DNS distribuidos a lo largo de Internet. Esta base de datos tiene como función enlazar a un host/cliente y a el nombre de una network/red, con sus direcciones en Internet. Esta es la razón por la cual este organismo cumple una función crítica en la operatividad a gran escala de los protocolos TCP/IP, incluido el servicio de e-mail.

1.3.2 ICANN.

Es una corporación sin fines de lucro establecida recientemente, con sede en California, para coordinar la asignación de nombres de dominios, de direcciones IP y el desarrollo de protocolos en Internet. Estas funciones habían venido por muchos años realizándose a través de IANA (Internet Assigned Numbers Authority), con financiamiento del gobierno de Estados Unidos a través de diversos contratos, pero durante el año pasado se comenzó un proceso de traspaso de estas responsabilidades al sector privado. A través de dos documentos (llamados el Green Paper y el White Paper), el gobierno de Estados Unidos dio inicio a un proceso mediante el cual pidió a la propia comunidad que se organizara para constituir ICANN y tomar el control de las funciones realizadas por IANA.

Así, se han definido tres organizaciones de apoyo: la DNSO (Domain Names Support Organization), la ASO (Address Support Organization) y la PSO (Protocol

Support Organization). Uno de los problemas más importantes que ICANN abordó en Santiago fue cómo resolver los conflictos por nombres de dominios que se sucitan en los dominios llamados genéricos , es decir, .com, .net y .org.

1.3.3 LACNIC.

El Registro de Direcciones de Internet para América Latina y el Caribe (LACNIC), es la organización que administra el espacio de direcciones IP, Números de Sistemas Autónomos (ASN), Resolución Inversa y otros recursos para la región de América Latina y el Caribe (LAC) en nombre de la comunidad Internet.

Tiene como objetivos representar y promover los puntos de vista de la comunidad de la región así como contribuir al desarrollo y crecimiento de Internet en la misma, además de promover oportunidades educacionales y políticas públicas relativas a la Internet.

Entre sus pautas LACNIC pretende ofrecer un servicio neutral, participativo, democrático y no lucrativo de calidad con un directorio elegido por sus miembros. En este contexto, la membresía se atendrá a reglas específicas, de acuerdo a la ubicación geográfica y al carácter de ISPs. Los miembros se regirán por un sistema de votación según el tamaño del espacio de direcciones que administren.

Las políticas y los procedimientos se basarán en las RFCs relacionadas con la administración de IP, números de sistemas autónomos y Resolución Inversa. Así mismo se implementarán permanentemente mecanismos de participación y discusión

tendientes a la actualización y modificación de las políticas. LACNIC es una organización sin fines de lucro, basada en membresía y establecida jurídicamente en el Uruguay.

Objetivos:

- Proveer servicios de registro de direcciones IP, ASN, Resolución Inversa y sus recursos asociados, con el propósito de permitir y facilitar las comunicaciones a través de redes informáticas.
- Representar y promover los puntos de vista e intereses de la región ante organismos internacionales, en el área de su competencia.
- Colaborar en el crecimiento de Internet en Latinoamérica y el Caribe.
- Asistir a la comunidad latinoamericana y caribeña en el desarrollo de procedimientos, mecanismos y estándares para la asignación eficiente de recursos de Internet.
- Promover oportunidades educacionales a sus miembros en áreas técnicas y políticas de su competencia.
- Proponer y desarrollar las políticas públicas en el área de su competencia.

1.4 SERVICIOS Y CARACTERÍSTICAS DE LOS ISP.

Un Proveedor de Servicio de Internet (ISP, Internet Service Provider), es una empresa que ofrece a sus usuarios conexión a la red mundial Internet y su gama de servicio relacionado, como correo electrónico y navegación grafica, entre otros

Un ISP tiene acceso a Internet por un canal dedicado usando una conexión permanente y un conjunto de equipos configurados para ofrecer los múltiples servicios a sus clientes

Existen varias formas de conexión al proveedor de Internet desde el punto de vista del usuario. Una de ellas es por línea telefónica, la más común; que consiste en una comunicación no permanente la cual requiere marcar el numero telefónico del ISP La segunda es una conexión dedicada(24 horas al día), en la que se establece una dirección de protocolo de comunicación IP (Protocolo de Internet) fija en la computadora del usuario, la cual ofrece varias ventajas a la industria y al comercio con respecto a la dirección temporal, esto significa disponer de una dirección permanente donde se pueda ubicar a este usuario o empresa en Internet

En ambos casos, el ISP valida al usuario a través de un nombre de identificación y una palabra clave para verificar su acceso a la red y se le asigna una dirección IP temporal para el caso de conexión telefónica. El proveedor de Internet posee un rango de direcciones IP, y a su vez entrega estas direcciones a los usuario ya sea temporal o permanente dependiendo del tipo de enlace.

1.5 Registros de enrutamiento en Internet.

Debido a los riesgos que conlleva la comisión de un error en la configuración de los anuncios BGP que se originan en un sistema autónomo y con la intención de paliar la propagación de dichos errores a través de la Internet es política habitual en los

grandes ISPs la configuración de listas de filtrado de prefijos en las sesiones BGP que tienen con peers y clientes.

De esta forma si uno de sus peers BGP anuncia un prefijo que no le corresponde este anuncio será filtrado por dichas políticas y no será propagado por el resto de la Internet. Inicialmente la generación de estas listas de prefijos se hacía de forma manual, mediante comunicación entre los NOCs de las partes afectadas que se comunicaban los cambios en las políticas de enrutamiento y las altas y bajas de prefijos.

Esta política llego a ser inviable con el crecimiento de la Internet ya que hay sesiones de peering en las que se manejan más de 20000 prefijos, lo que hace su mantenimiento manual prácticamente imposible.

Del deseo de automatizar este tipo de tareas nace la idea de los Routing Registries, que son bases de datos en las que se almacena la información sobre políticas de routing necesaria para poder generar automáticamente estas listas de filtrado de prefijos.

1.5.1 ¿Qué es un Routing Registry?

Un routing registry es una base de datos en la que se almacena, entre otras cosas, información sobre los prefijos que se tiene intención de anunciar en una sesión BGP, de forma que en caso de que se cometa un error este no se propague por la Internet.

Las principales entidades de Internet están concentrando los esfuerzos en unificar distintos modelos de Routing Registry que han ido apareciendo hasta la fecha y han creado el Internet Routing Registry (http://www.irr.net). Uno de estos esfuerzos ha sido el de la utilización del RPSL (Routing Policy Specification Language), publicado como RFC 2622 por el IETF. Este lenguaje permite definir la información contenida en la base de datos del registro de rutas de una forma estándar y así poder utilizar herramientas normalizadas para acceder a dicha información.

CAPÍTULO II

FUNDAMENTOS DE LOS PROTOCOLOS DE ENRUTAMIETO.

2.1 PROTOCOLOS DE ENRUTAMIENTO.

Internet es un conglomerado de sistemas autónomos que define la autoridad administrativa y las políticas de enrutamiento de distintas organizaciones. Los sistemas autónomos están compuestos por routers que ejecutan Protocolos de gateway interior (IGP), como el Protocolo de información de enrutamiento (RIP), eI Protocolo de enrutamiento de gateway interior mejorado (EIGRP), Primero la ruta libre más corta (OSPF) y el Sistema intermedio sistema intermedio (IS-IS), en el interior de sus limites y se interconectan mediante un Protocolo de gateway exterior (EGP). EI protocolo EGP estándar de hecho en Internet actualmente es el Protocolo de gateway fronterizo, versión 4 (BGP-4), definido en la RFC 17711.

2.1.1 Visión general de los routers y el enrutamiento.

Los routers son dispositivos que dirigen el trafico entre los hosts. Construyen tablas de enrutamiento que contienen información sobre las mejores rutas a todos los

destinos a los que se pode llegar. Los pasos para el enrutamiento básico son los siguientes:

Paso 1. Los routers ejecutan programas denominados protocolos de enrutamiento tanto para transmitir como para recibir información hacia y desde otros routers de la red.

Paso 2. Los routers utilizan esta información para rellenar las tablas de enrutamientos que están asociadas con cada protocolo de enrutamiento en particular.

Paso 3. Los routers analizan las tablas de enrutamiento desde diferentes protocolos (si se esta ejecutando mas de un protocolo de enrutamiento) y seleccionan la(s) mejor(es) ruta(s) para cada destino.

Paso 4. Los routers asocian con dicho destino la dirección de la capa de enlace de datos del dispositivo de próximo salto y la interfaz local de salida que sera utilizada cuando se reenvíen paquetes al destino. Tenga en cuenta que el dispositivo de próximo salto podría ser otro router, o inc1uso el host de destino.

Paso 5. La información de reenvío del dispositivo de próximo salto (la dirección de la capa de enlace de datos mas la interfaz de salida) esta situada en la tabla de reenvíos del router.

Paso 6. Cuando un router recibe un paquete, examina su cabecera para determinar la dirección de destino.

Paso 7. El router consulta la tabla de reenvío para obtener la interfaz de salida y la dirección del próximo salto para alcanzar el destino.

Paso 8. El router ejecuta cualquier función adicional requerida (como reducir el TTL IP o manipular las configuraciones del TOS IP) Y entonces reenvía el paquete hacia el dispositivo apropiado.

Paso 9. Esto continua hasta que se alcanza el host de destino. Este comportamiento refleja el paradigma del enrutamiento salto a salto que se utiliza generalmente en redes de switching de paquetes.

Los EGP, como por ejenplo BGP, fueron introducidos porque los IGP no escalaban bien en redes que iban más allá del nivel corporativo, con miles de nodos y cientos de miles de rutas. Nunca se pretendió utilizar los IGP para este propósito.

2.2 SISTEMAS AUTÓNOMOS.

Un sistema autónomo (SA) es un conjunto de routers que tiene una única política de enrutamiento, que se ejecuta bajo una única administración técnica, y que habitualmente utiliza un único IGP (el SA podría ser también un conjunto de IGP trabajando juntos para proporcionar enrutamiento interior). Para el mundo exterior, el SA es visto como una única entidad. Cada SA tiene un numero identificador, que se Ie asigna mediante un Registro de Internet, 0 un proveedor de servicios en el caso de SA privados. La información de enrutamiento entre varios SA se intercambia mediante un protocolo de gateway exterior como BGP-4, según se muestra en la Figura.

Lo que hemos ganado al dividir el mundo en administraciones es la capacidad de tener una gran red (en el sentido de que Internet podría haber sido una gran red OSPF 0 IS-IS) dividida en redes mas pequeñas y manipulables. Dichas redes, representadas como SA, pueden implementar su propio conjunto de reglas y políticas que distinguirán unívocamente sus redes y los servicios asociados ofrecidos de otras redes. Cada SA puede ahora ejecutar su propio conjunto de IGP, independientemente

de los IGP de otros SA.

Las siguientes secciones explicaran configuraciones de red potenciales con redes de conexión única, redes con múltiples conexiones no de transito, y redes de tránsito con múltiples conexiones.

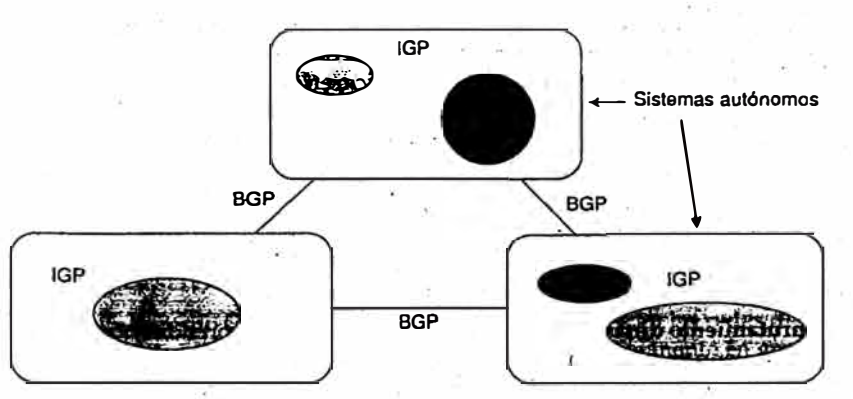


Figura 2.1 Sistemas autónomos.

2.2.1 SA de conexión única.

Se considera que un SA es de conexión única cuando alcanza las redes exteriores a su dominio a través de un único punto de salida. La Figura ilustra un SA de conexión única.

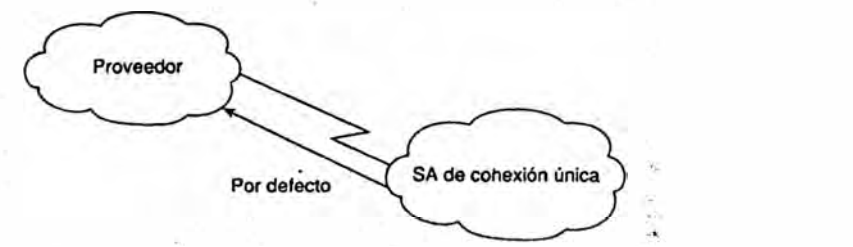


Figura 2.2 Sistemas autónomos de conexión única.

Un SA de conexión única no tiene que aprender realmente las rutas de Internet de su proveedor. Dado que solo hay una salida, todo el trafico puede ir por defecto al proveedor. Cuando se usa esta configuración, el proveedor puede utilizar diferentes métodos para publicar las rutas del cliente a otras redes.

Una posibilidad para el proveedor es enumerar las subredes del cliente como entradas estáticas en su router. EI proveedor publicara entonces dichas entradas a Internet a través de BGP. Este método escalaría muy bien si las rutas del cliente pudieran representarse mediante un conjunto pequeño de rutas agregadas. Cuando el cliente tiene demasiadas subredes no contiguas, enumerar todas esas subredes por medio de rutas estáticas resulta ineficaz.

Alternativamente, el proveedor puede emplear IGP para publicar las redes del cliente. Puede utilizarse un IGP entre el cliente y el proveedor para que el cliente publique sus rutas. Esto tiene todos los beneficios del enrutamiento dinámico donde la información de red y los cambios son enviados dinámicamente al proveedor. Sin embargo, esto es muy poco común, principalmente porque no escala muy bien debido a que la inestabilidad del enlace del cliente puede dar lugar a inestabilidades IGP.

El tercer método par el cual el ISP puede aprender y publicar las rutas del cliente consiste en utilizar BGP entre el cliente y el proveedor. En la situación de un SA de conexión única, es difícil obtener el numero de un SA registrado de un IRR porque las políticas de enrutamiento del cliente son una extensión de las políticas de un único proveedor.

En su lugar, el proveedor puede otorgar al cliente un numero de SA de la colección de sistemas autónomos (65412-65535), asumiendo que las políticas de enrutamiento

del proveedor han establecido soporte para usar espacio privado de los SA con los clientes, como se describe en la RFC 227010.

Nota: La RFC 1930 proporciona un conjunto de normativas para la creación, selección y registro de números de sistemas autónomos.

2.2.2 SA de múltiples conexiones sin tránsito.

Un SA es de múltiples conexiones si tiene mas de un punto de salida hacia el exterior. Un SA puede tener múltiples conexiones hacia un único proveedor o hacia varios proveedores. Un SA sin tránsito no permite tráfico de tránsito a través de el. El trafico de tránsito es cualquier tráfico que tenga origen y destino fuera del SA. La Figura 4.5 ilustra un SA (AS1) sin tránsito y con múltiples conexiones a dos proveedores, ISPI e ISP2.

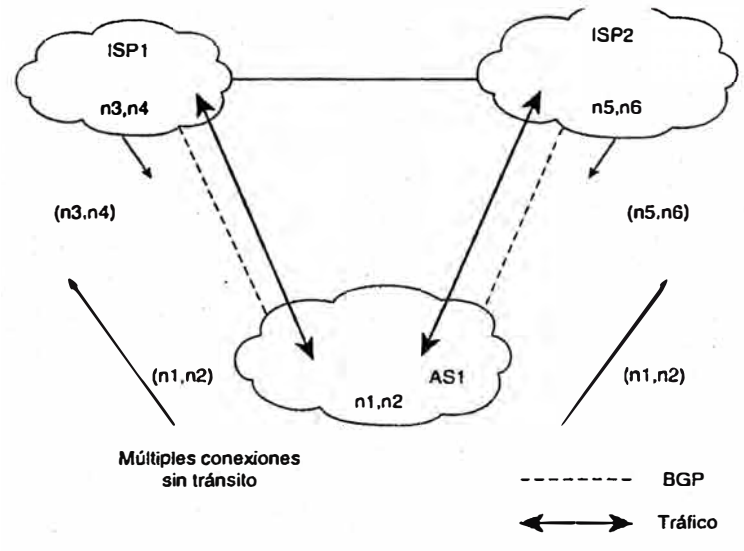


Figura 2.3 Sistemas autónomos de múltiples conexiones.

Un SA sin tránsito solo publicara sus propias rutas y no propagara rutas que haya aprendido de otros SA. Esto asegura que el tráfico hacia cualquier destino que no pertenezca al SA no será dirigido hacia el SA. En la Figura 4.5, AS1 aprende las rutas n3 y n4 por medio de ISPI y las rutas n5 y n6 por medio de ISP2. ASI publica solo sus rutas locales (n1,n2). No pasara hacia ISP2 las rutas que aprendió de ISP1 o hacia ISP1 las rutas que aprendió de ISP2. De esta forma, AS1 no se abre al tráfico del exterior, como ISP1 intentando alcanzar n5 0 n6 e ISP2 intentando llegar a n3 y n4 a través de AS1. Por supuesto, ISPI 0 ISP2 pueden obligar a que su trafico se dirija hacia ASI por medio de enrutamiento predeterminado o de enrutamiento estático. Como precaución contra esto, AS1 solo podría filtrar cualquier tráfico entrante hacia el con un destino que no pertenezca a AS1.

2.2.4 SA de tránsito con múltiples conexiones.

Un SA de tránsito con múltiples conexiones tiene mas de una conexión con el exterior y todavía puede ser utilizado para el tráfico de tránsito por otros SA.

El tráfico de tránsito (relativo al SA con múltiples conexiones) es cualquier tráfico que tenga un origen y un destino que no pertenezca al SA local.

Aunque BGP-4 es un protocolo de gateway exterior también puede utilizarse dentro de un SA como un conducto para intercambiar actualizaciones BGP. Las conexiones BGP entre routers dentro de un sistema autónomo son denominadas BGP internos (IBGP); mientras que las conexiones BGP entre routers en sistemas autónomos Separados son denominadas BGP externos (EBGP). Los routers que están utilizando IBGP se denominan routers de transito cuando transportan el tráfico de tránsito que va a través del SA.

Un SA de tránsito publicara a un SA las rutas que haya aprendido de otro SA. De esta forma, el SA de tránsito se abrirá al tráfico que no le pertenezca. Es aconsejable que los SA de tránsito de múltiples conexiones utilicen BGP-4 para sus conexiones a otros SA y para proteger sus routers internos sin transito de las rutas de Intemet. No todos los routers dentro de un dominio necesitan ejecutar BGP; los routers internos sin transito pueden ejecutar un enrutamiento predeterminado hacia los routers BGP, lo que alivia el número de rutas que los routers internos sin tránsito deben transportar. Sin embargo, en la mayoría de las grandes redes de los proveedores de servicios todos los routers transportan habitualmente un conjunto lleno de rutas BGP internamente.

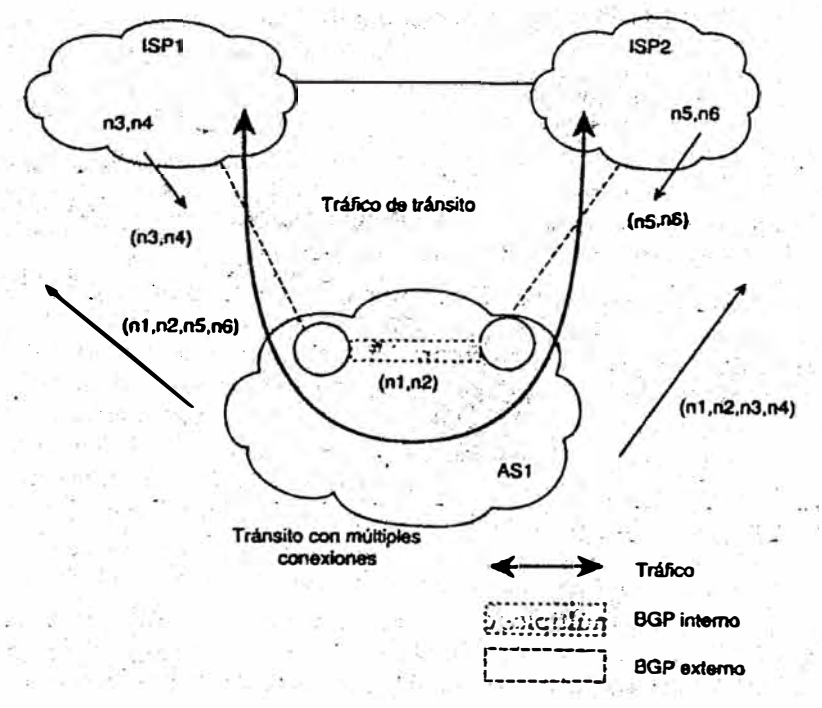


Figura 2.4 Sistema autónomo de tránsito a dos ISP.

La Figura ilustra un sistema autónomo de tránsito con múltiples conexiones, AS1, conectado a dos proveedores diferentes, ISP1 e ISP2. AS1 aprende las rutas n3, n4, n5 y n6 de ISP1 e ISP2, y publica por turnos todo lo que aprende, incluyendo sus rutas locales, hacia ISP1 e ISP2, En este caso, ISP1 podría utilizar AS1 como un SA de tránsito para llegar alas redes n5 y n6, e ISP2 podría utilizar AS1 para alcanzar las redes n3y n4.

CAPÍTULO III

PROTOCOLO BGP-4.

El Protocolo de gateway fronterizo (BGP) ha pasado por diversas fases y mejoras desde su versión original, BGP-l, en 1989. La distribución de BGP-4 comenzó en 1993. Es la primera versión de BGP que administra la agregación (enrutamiento entre dominios sin clase [CIDR]) y las superredes.

BGP no impone restricciones sobre la topología de red subyacente. Asume que el enrutamiento dentro de un sistema autónomo se hace mediante un protocolo de enrutamiento intra-sistema autónomo (Protocolo de gateway interior [IGP]) Para lo que significa dentro de una entidad, e inter significa entre entidades. BGP construye un grafico de sistemas autónomos basados en la información intercambiada entre los routers BGP. Este entorno gráfico se denomina en ocasiones árbol. En lo que concierne a BGP, Internet es un grafico de SA, con cada SA identificado por un número de SA único. Las conexiones entre dos SA juntos forman una ruta de acceso, y el conjunto de información de rutas de acceso forma una ruta para llegar a

un destino especifico. BGP utiliza la información de ruta de acceso asociada con un destino dado para asegurar el enrutamiento entre dominios libre de bucles.

3.1 COMO TRABAJA BGP.

BGP es un protocolo por vector de ruta utilizado para transportar información de enrutamiento entre sistemas autónomos. El termino Vector de ruta viene del hecho de que la información de enrutamiento de BGP transporta una secuencia de números de SA que identifica la ruta de SA que un prefijo de red ha seguido. La información de ruta de acceso asociada con el prefijo se utiliza para activar la prevención de bucles.

BGP utiliza TCP como su protocolo de transporte (puerto 179). Esto asegura que toda la seguridad del transporte (como la retransmisión) queda al cuidado de TCP y no necesita ser implementada en BGP, simplificando así la complejidad asociada con la fiabilidad de diseño en el propio protocolo.

Los routers que ejecutan un proceso de enrutamiento BGP a menudo se conocen como portavoces BGP. Dos portavoces BGP que forman una conexión TCP entre ambos con el propósito de intercambiar información de enrutamiento se denominan vecinos o iguales. La Figura ilustra esta relación. Los routers de iguales intercambian mensajes abiertos para determinar los parámetros de conexión. Estos mensajes son utilizados para comunicar valores como el número de versión del portavoz BGP.

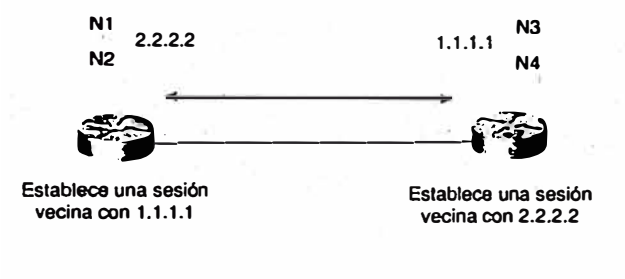


Figura 3.1 Establecimiento de una sesión BGP.

BGP también proporciona un mecanismo para cerrar elegantemente una conexión con un igual, En otras palabras, en caso de desacuerdo entre iguales, sea resultante de configuración, incompatibilidad, intervención de operador u otras circunstancias, se envía un mensaje de error NOTIFICATION, y la conexión al igual no se establece o se corta si ya estaba establecida. EI beneficio de este mecanismo es que ambos iguales comprenden que la conexión no puede ser establecida o mantenida, por lo que no se desperdician recursos que de otra forma serán requeridos para mantener o reintentar establecer la conexión a ciegas. El mecanismo de cierre elegante simplemente asegura que todos los mensajes pendientes principalmente mensajes de error NOTIFICA TION, sean entregados antes de que se cierre la sesión TCP.

Inicialmente, cuando se establece una sesión BGP entre un conjunto de portavoces BGP, todas las rutas BGP candidatas son intercambiadas, como se muestra en la Figura Después de haber establecido la sesión y se haya producido el intercambio de la ruta inicial, solo se envían las actualizaciones incrementales como cambios en la información de la red. El enfoque de actualización incremental ha mostrado una mejora enorme en el gasto de CPU y asignación de ancho de banda en comparación con las periódicas actualizaciones completas utilizadas por los anteriores protocolos,

como EGP.

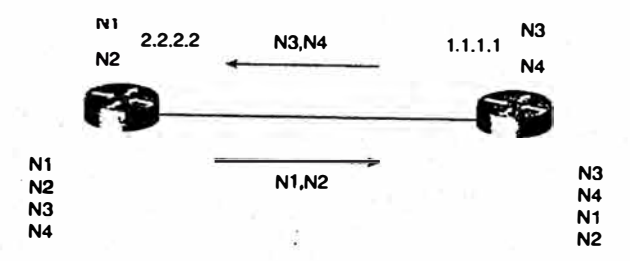


Figura 3.2 Proceso de intercambio de rutas mediante BGP.

Las rutas se publican entre un par de routers BGP en los mensajes UPDATE. El mensaje UPDATE contiene, entre otras cosas, una lista de pares <longitud, prefijo> indicando los destinos que pueden ser alcanzados a través de un portavoz BGP. EI mensaje UPDATE también contiene los atributos de la ruta de acceso, que incluye información como el grado de preferencia para una ruta en particular y la lista de SA por los que la ruta ha pasado.

En el caso de que una ruta se vuelva inalcanzable, un portavoz BGP informa a sus vecinos dando de baja la ruta incorrecta, las rutas retiradas son parte del mensaje UPDATE. Esas rutas ya no estarán disponibles para el uso. Si la información asociada con una ruta ha cambiado, o se ha seleccionado una nueva ruta de acceso para el mismo prefijo, no es necesaria una retirada; es suficiente publicar una sustitución de ruta.

Los mensajes KEEPALIVE son enviados periódicamente entre los vecinos BGP para asegurar que la conexión se mantiene viva, Los paquetes KEEPALIVE (de 19 bytes cada uno) no deberían causar ninguna variación en la CPU del router o el

ancho de banda del enlace, porque consume una cantidad mínima de ancho de banda (un paquete instantáneo de 152 bits cada sesenta segundos, o unos 2,5 bps por cada igual cada sesenta segundos).

BGP guarda el número de versión de la tabla para estar al corriente de la instancia actual de la tabla de enrutamiento de BGP, Si la tabla cambia, BGP incrementa el numero de versión de la tabla. Una versión de tabla que se incrementa rápidamente es normalmente una indicación de inestabilidad en la red (aunque esto es bastante común en grandes proveedores de servicio de Internet) por esto, la inestabilidad introducida por las redes conectadas a Internet en cualquier lugar del mundo supondrán un incremento del numero de versión de tabla en cada portavoz BGP que tenga una vista completa de las tablas de enrutamiento en Internet. Route flap dampening y otras medidas han sido diseñadas para minimizar los efectos de esta inestabilidad.

3.2 FORMATO DE LA CABECERA DE MESAJE BGP.

EI formato de cabecera de mensaje BGP es un campo Marcador de 16 bytes, seguido de un campo Longitud de 2 bytes y un campo Tipo de 1 byte. La Figura ilustra el formato básico de la cabecera de mensaje BGP. Dependiendo del tipo de mensaje, podría haber o no una porción de datos a continuación de la cabecera los mensajes KEEP ALIVE por ejemplo, constan solo de la cabecera de mensaje, sin datos a continuación.

El campo Marcador, de 16 bytes, se utiliza para autentificar mensajes BGP entrantes o para detectar perdidas de sincronización entre dos iguales BGP. El campo Marcador puede tener uno de dos formatos:

Si el tipo del mensaje es OPEN, o si el mensaje OPEN no tiene información de autenticación, el campo Marcador debe ser todo 1.

De otro modo el campo Marcador se computaría en función de la parte del mecanismo de autenticación utilizado y el uso de la opción de firma TCP MD5 de este marcador.

El campo Longitud, de 2 bytes, se utiliza para indicar la longitud total del mensaje BGP, incluyendo la cabecera. El mensaje BGP mas pequeño no es menor de 19 bytes (16+2+1) ni mayor de 4.096 bytes. EI campo Tipo de 1 byte, indica el tipo de mensaje, con las siguientes posibilidades:

. OPEN.

. UPDATE.

. NOTIFICATION.

. KEEP ALIVE.

Las siguientes secciones examinan mas detalladamente el propósito y el formato de cada uno de los cuatro tipos de mensajes.

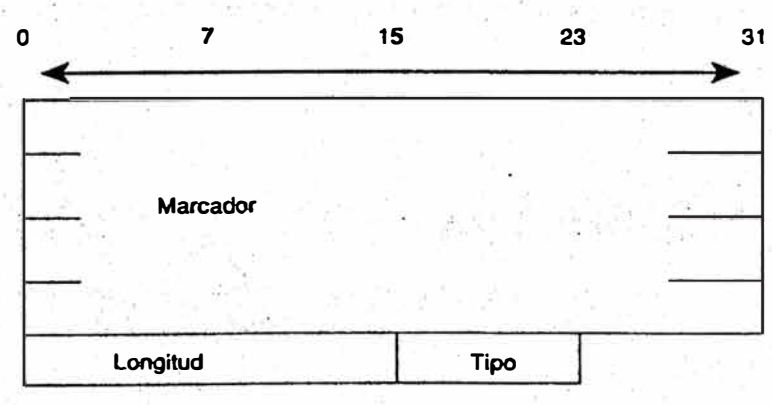


Figura 3.3 Formato de la cabecera de un mensaje BGP.

3.3 NEGOCIACIÓN DE VECINO BGP.

Uno de los pasos básicos del protocolo BGP es establecer sesiones entre iguales BGP. Sin una finalización con éxito de este paso, no se producirá el intercambio de actualizaciones. La negociación vecina se basa en la finalización satisfactoria de una conexión de transporte TCP, el procesamiento con éxito del mensaje OPEN y la detección periódica de los mensajes, UPDATE o KEEPALIVE.

Formato del **mensaje OPEN**

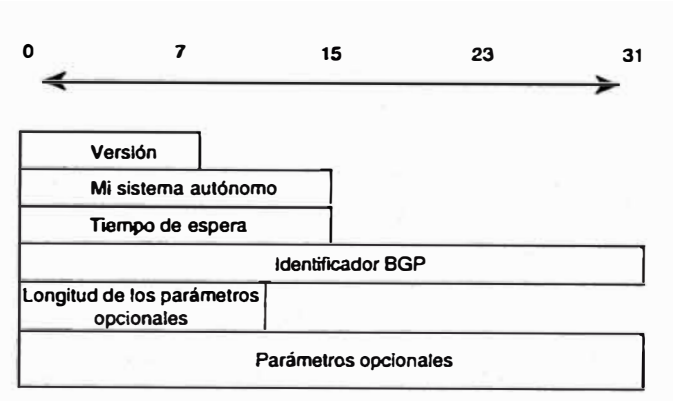


Figura 3.4 Formato del mensaje OPEN.

Las siguientes descripciones resumen cada uno de los campos del mensaje OPEN:

Versión. Un entero sin signo de 1 byte que indica la versión del mensaje BGP, como BGP-3 o BGP-4. Durante la negociación vecina, los iguales BGP se ponen de acuerdo sobre el número de versión BGP. Los iguales BGP intentan negociar la versión superior común que soporten ambos. Reinician la sesión BGP y renegocian hasta que una versión soportada en común es determinada por ellos. **Mi sistema autónomo**. Un campo de 2 bytes que indica el número de SA del portavoz BGP.

Temporizador de espera. EI temporizador de espera es un entero sin signo de 2

bytes que indica la máxima cantidad de tiempo en segundos que puede transcurrir entre la recepción de mensajes KEEPALIVE o UPDATE sucesivos. El temporizador de espera es un contador que se incrementa desde 0 al valor del tiempo de espera. La recepción de un mensaje KEEPALIVE 0 UPDATE provoca que el temporizador de espera se reinicie a O. Si se excediera el tiempo de espera de un vecino determinado, el vecino se considerara extinto.

Identificador BGP. Un entero de 4 bytes que indica el valor del ID BGP del emisor, normalmente esto es igual al ID del router (RID), que se calcula como la dirección IP más alta del router.

Longitud de los parámetros opcionales. Es un entero sin signo de 1 byte que indica la longitud total en bytes del campo Parámetros opcionales. Una longitud de 0 indica que no hay parámetros opcionales.

Parámetros opcionales. Este es un campo de longitud variable que indica una lista de parámetros opcionales utilizados en la sesión BGP de negociación vecina. Este campo se representa mediante la tripleta <Tipo de parámetro, Longitud del para metro, Valor del para metro> con longitudes de 1 byte, 1 byte y longitud variable, respectivamente. Un ejemplo de parámetros opcionales es el parámetro de autenticación de información (tipo 1), que se utiliza para autentificar la sesión con un igual BGP.

3.4 PERSPECTIVA DE LA MÁQUINA DE ESTADO FINITO.

La negociación vecina de BGP transcurre por diferentes etapas antes de que la conexión este completamente establecida. En la Figura se ilustra una máquina

simplificada de estado finito que destaca los grandes eventos del proceso con una indicación de mensajes (OPEN, KEEPALIVE, NOTIFICATION) enviada al igual en la transición de un estado a otro.

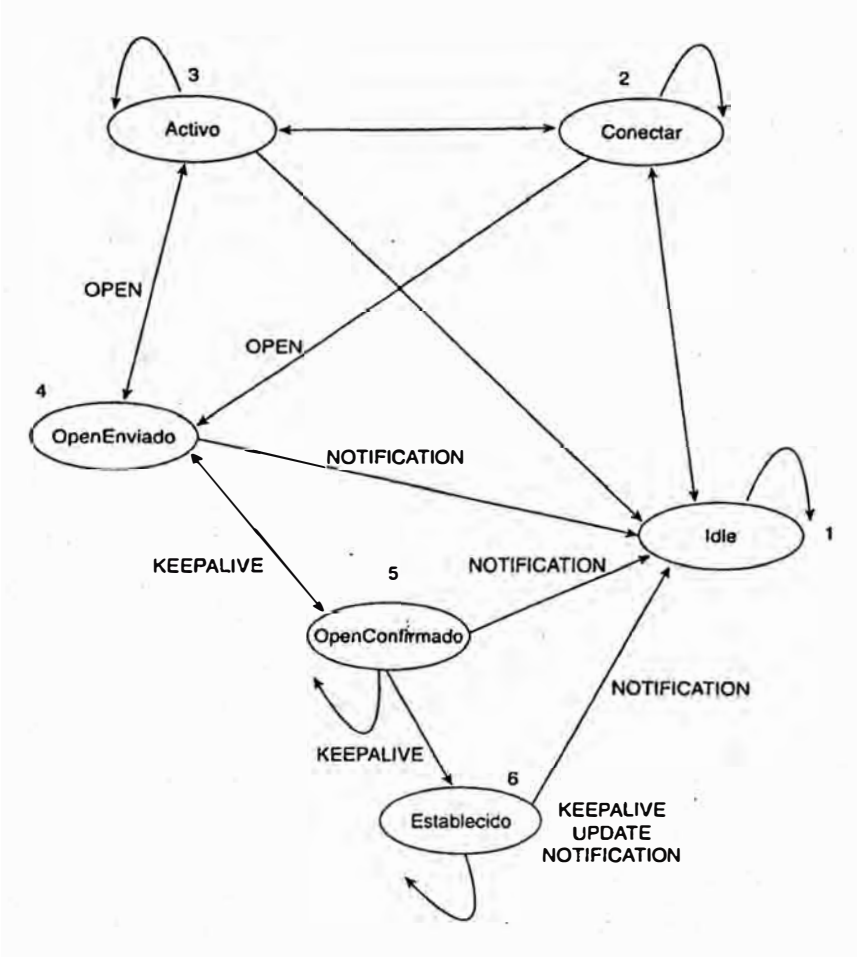


Figura 3.5 Máquina simplificada de estado finito.

1. Inactivo: Esta es la primera etapa de la conexión. BGP esta esperando un evento Inicio, que es iniciado por un operador del sistema BGP. Un administrador que establezca una sesión BGP a través de la configuración del router o reiniciando una sesión ya existente, normalmente produce un evento Inicio. Tras el evento Inicio, BGP inicializa sus recursos; reinicia un temporizador Reintento Conexión, inicia una conexión de transporte TCP, y comienza a escuchar esperando una conexión que

puede iniciarla un igual remoto. BGP entonces realiza la transición a un estado Conectar. En caso de error, BGP vuelve al estado Inactivo.

2. Conectar: BGP esta esperando a que se complete la conexión del protocolo de transporte. Si la conexión de transporte TCP tiene éxito, las transiciones de estado hacia OpenEnviado (esto es donde se envía el mensaje OPEN). Si la conexión de transporte no tiene éxito, el estado pasa a Activo. Si el temporizador Reintento Conexión expira, el estado se mantienen en la etapa Conectar, se reinicia el temporizador y se inicia una conexión de transporte. En caso de cualquier otro evento (iniciado por el sistema o por el operador), el estado vuelve a Inactivo.

3. Activo: BGP intenta adquirir un igual iniciando una conexión de protocolo de transporte. Si se establece la conexión de transporte, pasa a OpenEnviado (se envía un mensaje OPEN). Si el temporizador ReintentoConexión expira, BGP reinicia el contador ReintentoConexion y vuelve al estado Conectar. Además, BGP continua esperando escuchar una conexión que podría ser iniciada por otro igual. EI estado podría volver a Inactivo en caso de otros eventos, como un evento Stop iniciado por el sistema o el operador.

En general, un estado vecino que oscila entre Conectar y Activo indica que algo no va bien en la conexión de transporte TCP. Podría ser debido a muchas retransmisiones TCP o a la incapacidad de un vecino para alcanzar la dirección IP de su igual.

4. OpenEnviado: BGP esta esperando un mensaje OPEN de su igual. Se comprueba que el mensaje OPEN es correcto. En caso de error, como un numero de versión erróneo o un SA inaceptable, el sistema envía un mensaje de error NOTIFICATION

y vuelve a Inactivo. Si no hay errores, BGP comienza a enviar mensajes KEEPALIVE y reinicia el temporizador KEEPALIVE. En esta etapa, el tiempo de espera es negociado, y se toma el valor más pequeño. En caso de que el tiempo negociado sea 0, el temporizador de espera y el de KEEP ALIVE no se reinician. En el estado OpenEnviado, BGP reconoce, comparando su numero de SA con el número de SA de su igual, si el igual pertenece al mismo SA (BGP Interno) o a un SA diferente (BGP externo) Cuando se detecta una desconexión de transporte TCP, el estado cae de nuevo al estado Activo. Para cualquier otro error, como una expiración del temporizador de espera, BGP envía un mensaje NOTIFICATION con el código de error correspondiente y vuelve al estado Inactivo. También, en respuestas a un evento de parada iniciado por el sistema o el operador, el estado retrocede al estado Inactivo.

5. OpenConfirmado: BGP espera un mensaje KEEPALIVE. Si se recibe un KEEPALIVE, el estado va a Establecido, y la negociación vecina se completa. Si el sistema recibe un mensaje KEEPALIVE, reinicia el temporizador de espera (suponiendo que el tiempo de espera negociado no es 0). Si se recibe un mensaje NOTIFICATION, el estado retrocede al estado Inactivo. EI sistema envía mensajes KEEPALIVE periódicos ala velocidad establecida por el temporizador KEEPALIVE. En caso de cualquier notificación de desconexión de transporte o en respuesta a cualquier evento de parada (iniciado por el sistema o el operador), el estado retrocede a Inactivo. En respuesta a cualquier otro evento, el sistema envía un mensaje NOTIFICATION con un código de error FSM (Maquina de estado finito) y vuelve al estado Inactivo.

6. Establecido: Este es el estado final de la negociación vecina. En esta etapa, BGP

comienza a intercambiar paquetes UPDATE con sus iguales. Suponiendo que es distinto de cero, el temporizador de espera se reinicia al recibir un mensaje UPDATE o KEEPALIVE. Si el sistema recibe cualquier mensaje NOTIFICATION (si ha ocurrido un error), el estado retrocede a Inactivo. Los mensajes UPDATE son comprobados en busca de errores, tales como atributos perdidos, duplicados, y demás. Si se encuentran errores, se envía un mensaje NOTIFICATION hacia el igual, y el estado retrocede a Inactivo. Si el temporizador de espera expira, o se recibe una notificación de desconexión del protocolo de transporte, o se recibe un evento Stop, o en respuesta a cualquier otro evento, el sistema retrocede al estado Inactivo.

3.4.1 Mensaje NOTIFICATION.

Del examen precedente de la Maquina de estado finito, debería ser aparente que existen muchas oportunidades entre los distintos estados para detectar los errores. Un mensaje NOTIFICATION se envía siempre donde quiera que se detecta un error. Después de eso, se cierra la conexión con el igual. Los administradores de redes necesitan evaluar esos mensajes NOTIFICATION para determinar la naturaleza especifica de los errores que surgen en el protocolo de enrutamiento. La Figura ilustra el formato general del mensaje.

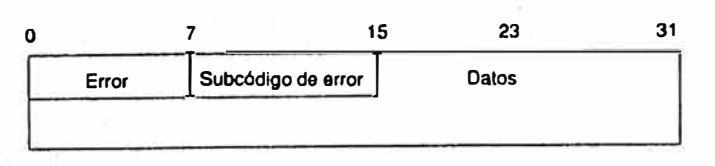


Figura 3.6 Formato del mensaje NOTIFICACIÓN.

EI mensaje NOTIFICATION esta compuesto del Código de error (1 byte), el subcódigo de error (1 byte) y el campo Datos (variable).

EI código de error indica el tipo de la notificación, y el subcódigo de error proporciona información mas especifica sobre la naturaleza del error. El campo Datos contiene datos relevantes para el error, como una mala cabecera, un numero de SA erróneo e información por el estilo. La Tabla siguiente enumera los posibles errores y sus subcódigos.

Código de error	Subcódigo de error
1. Error en cabecera de mensaje	1. Conexión no sincronizada. 2. Longitud de mensaje incorrecta. 3. Tipo de mensaje incorrecto.
2. Error en el mensaje OPEN	1. Número de versión no soportado. 2. Igual AS erróneo. 3. Identificador BGP incorrecto. 4. Parámetro opcional no soportado. 5. Fallo de autenticación. 6. Temporizador de espera inaceptable. 7. Capacidad no soportada.
3. Error en el mensaje UPDATE	1. Lista de atributos mal formada. 2. Atributo conocido no reconocido. 3. Atributo conocido perdido. 4. Error en los *flags* del atributo. 5. Error en la longitud del atributo. 6. Atributo de origen erróneo. 7. Bucle de enrutamiento en el SA. 8. Atributo NEXT_HOP erróneo. 9. Error en atributo opcional. 10. Campo de red erróneo. 11. AS_PATH mal formado.
4. Temporizador de espera expirado	N/A.
5. Error de la máquina de estado finito (para errores detectados por la FSM)	N/A.
6. Cesar (para errores fatales además de los ya indicados)	N/A.

Tabla 3.1 Códigos de errores de NOTIFICACIÓN.

3.4.2 Mensaje KEEPALIVE.

Los mensajes KEEPALIVE son mensajes periódicos intercambiados entre iguales para determinar si estos son accesibles. Como se dijo anteriormente, el tiempo de espera es el periodo de tiempo máxima que puede transcurrir entre la recepción de mensajes KEEPALIVE o UPDATE sucesivos. Los mensajes KEEPALIVE se envían a una velocidad que asegure que el tiempo de espera no expirará (La sesión se considera viva). La velocidad recomendada para KEEPALIVE es un tercio del valor del temporizador de espera. Si el valor de este ultimo es 0, los mensajes KEEPALIVE periódicos no se envían. Como se menciono previamente. El mensaje KEEPALIVE es una cabecera de mensaje BGP de 19 bytes sin datos a continuación, o puede ser eliminado durante un intervalo si se envía un mensaje UPDATE.

3.4.3 Mensaje UPDATE e información de enrutamiento

En el centro del protocolo BGP esta el concepto de actualizaciones de enrutamiento, que contienen toda la información necesaria que utiliza BGP para construir una imagen de la red libre de bucles. Los bloques básicos de un mensaje UPDATE son los siguientes:

- Información de accesibilidad de la capa de red (NLRI).
- Atributos de la ruta de acceso.
- Rutas no factibles.

La Figura siguiente ilustra esos componentes en el contexto del formato de un mensaje UPDATE.

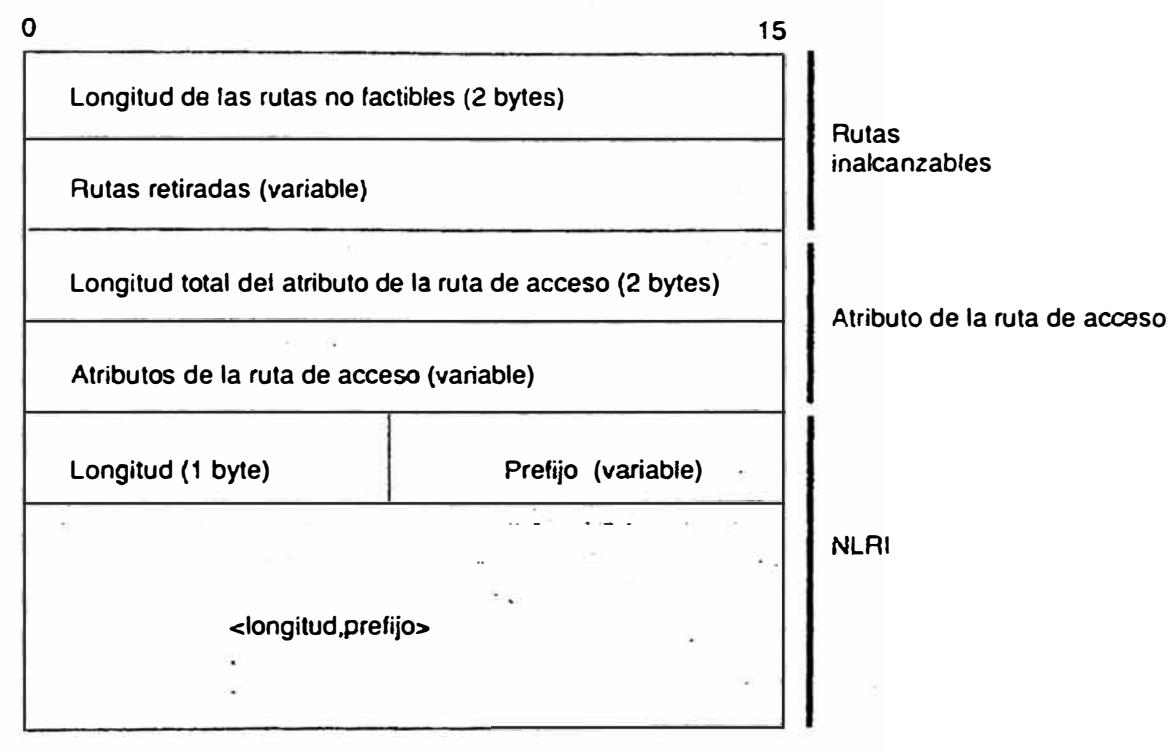


Figura 3.7 Formato de un mensaje UPDATE.

3.5 Opción de firma MD5.

La opción de firma MD5 para TCP ,definida en la RFC 2385, se utiliza para ayudar a BGP a protegerse de segmentos TCP spoofed y particularmente, reinicios de TCP.

Esta opción proporciona un mecanismo para que TCP transporte un boletín de mensajes en cada segmento TCP, donde el boletín utiliza la información conocida solo para los puntos extremos de la conexión y actúa como una firma para el segmento.

Aplicar el algoritmo MD5 a los siguientes elementos, en el orden indicado, produce el boletín creado para un segmento dado:

- Pseudocabecera TCP, por este orden: dirección IP origen, dirección IP destino, numero de protocolo rellenado con ceros y longitud de segmento.
- Cabecera TCP, excluyendo opciones y suponiendo una paridad e cero.
- Segmento de datos TCP.
- Clave o contraseña especificada independientemente, conocida tanto para el emisor TCP como para el receptor.

Cuando TCP recibe un segmento firmado, el receptor debe validarlo utilizando su clave local para calcular su propio boletín y comparar el valor con el boletín recibido, si la comparación da como resultado valores desiguales, deberá descartarse el segmento y no producir ninguna respuesta hacia el emisor.

La opción MD5 aparece en todos los segmentos y siempre tiene 16 bytes de longitud (recordemos los 16 bytes disponibles en la cabecera de mensaje BGP que estaban reservados para este propósito).

CAPÍTULO IV

FUNCIONAMIENTO Y MANEJO DE ATRIBUTOS BGP.

4.1 Construcción de sesiones EBGP e IBGP.

Una conexión vecina (también denominada conexión entre iguales) entre dos routers puede ser establecida dentro del mismo SA, en cuyo caso BGP se denomina BGP interno (IBGP). Asimismo, una conexión de iguales entre routers de SA diferentes se conoce como BGP externo (EBGP). La Figura contrasta dichos entornos.

En el establecimiento de la sesión vecina y durante la negociación de intercambio de mensajes OPEN, los routers de iguales comparan números de SA y determinan si son iguales del mismo SA o de SA diferentes. La diferencia entre EBGP e IBGP se manifiesta en como cada igual procesa las actualizaciones de enrutamiento que llegan del otro igual y en la forma en que distintos atributos BGP son transportados sobre conexiones externas en comparación con las conexiones internas.

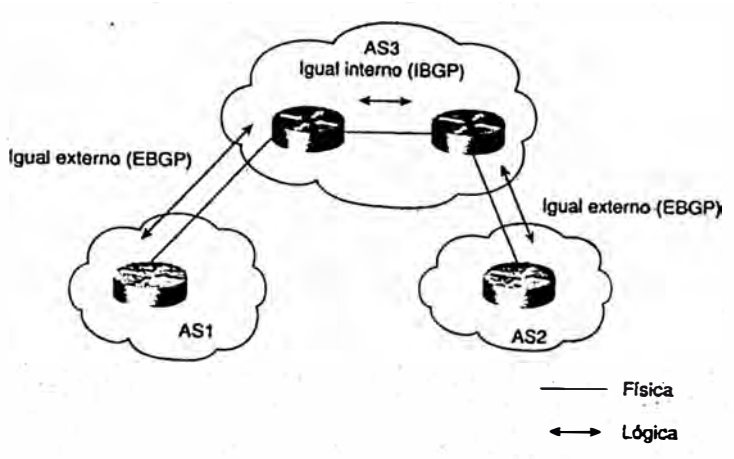


Figura 4.1 Sesiones EBGP e IBGP.

El proceso de la negociación vecina es básicamente el mismo para vecinos internos y externos hasta donde se construye la conexión TCP en el nivel de transporte. Es esencial tener conectividad IP entre los dos vecinos para que se establezca la sesión de transporte.

La conectividad IP debe lograrse a través de un protocolo diferente a BGP, ejemplo:

"Los vecinos pueden llegar unos a otros a través de algún Protocolo de gateway interior (IGP), se establece la sesión BGP, y se intercambian los mensajes BGP. La conexión IGP se cae por alguna razón, pero la sesión TCP de BGP sigue activa porque los vecinos todavía pueden llegar de unos a otros vía BGP. Con el tiempo, la sesión caerá porque la sesión BGP no puede depender de BGP en si mismo para la conectividad vecina; el substrato subyacente proporciona accesibilidad NEXT_HOP. Otro ejemplo es si una ruta mas especifica que la utilizada para establecer la conexión es aprendida a través de BGP".

En las sesiones peering IBGP es más común que un Protocolo de gateway interior

(lGP) o una ruta estática puedan ser configurados para lograr conectividad IP. En esencia, un paquete ping, que contiene una dirección lP origen (la dirección lP de un igual BGP) y una dirección lP destino (la dirección IP del segundo igual), debe tener éxito para que se inicie la sesión de transporte. Generalmente, para las sesiones BGP externas, una ruta a través de una interfaz conectada directamente establece la accesibilidad IP.

4.1.1 Conexiones físicas frente a lógicas.

Los vecinos BGP externos tienen como restricción que deben estar físicamente conectados, adyacentes unos a otros. BGP ignora cualquier mensaje UPDATE de su igual BGP externo si el igual no esta conectado físicamente, a menos que se especifique otra cosa. Sin embargo, surgen algunas situaciones en las que los vecinos externos pueden no estar en el mismo segmento físico. Tales vecinos están conectados lógicamente (a muchos saltos lP de distancia) pero no físicamente conectados. Un ejemplo seria ejecutar BGP entre vecinos externos a través de routers no BGP. En esta situación, Cisco (y muchos otros fabricantes) ofrecen una opción adicional para no evadir esta restricción. BGP requeriría algo de configuración adicional para indicar que su igual externo no esta conectado físicamente.

Una sesión BGP formada entre iguales BGP externos que no están físicamente conectados se denomina EBGP multisalto. En la Figura, RT2 no puede ejecutar BGP, pero RT1 y RT3 si . De esta forma, los vecinos externos RT1 y RT3 están conectados lógicamente y un igual con otro vía EBGP multisalto (observe, sin embargo, que RT2 debe aprender de alguna forma la información de enrutamiento

apropiada para evitar potenciales bucles de reenvío o paquetes de agujeros negros).

Por otro lado, los vecinos dentro del mismo sistema autónomo (vecinos internos) no tienen restricciones sobre si el igual esta conectado físicamente o separado por muchos saltos IP. Mientras haya conectividad IP entre los dos vecinos, BGP no requiere configuración adicional. En la Figura, RT1 Y RT4 están conectados lógica, pero no físicamente. Dado que ambos están en el mismo SA, no necesitan configuración adicional para ejecutar IBGP.

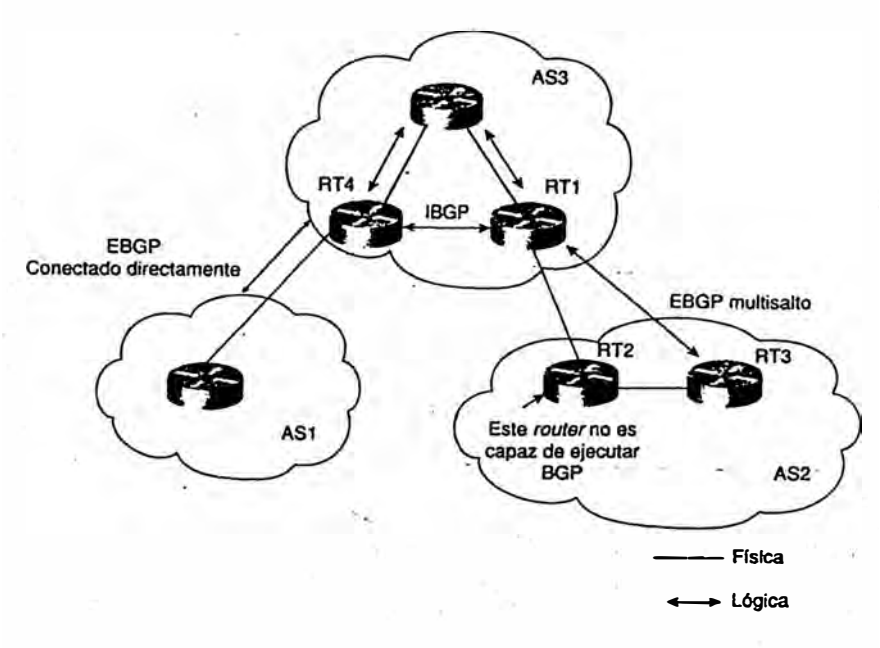


Figura 4.2 Conexiones físicas frente a lógicas.

4.1.2 Como obtener una dirección IP.

La dirección IP del vecino podría ser la dirección de cualquiera de las interfaces de los routers, tales como Ethernet, Token Ring o serie. Tener en cuenta que la estabilidad de la conexión vecina depende de la estabilidad de la dirección lP que elija.

5i la dirección IP pertenece a una tarjeta, Ethernet que tiene algún problema de

hardware y que se esta cayendo cada pocos minutos, la conexión vecina y la estabilidad del sistema de enrutamiento sufrirán las consecuencias, se proporciona la capacidad de configurar una interfaz virtual, denominada interfaz del bucle de prueba (loopback), que se supone está activa todo el tiempo. Vinculando la conexión vecina BGP con una interfaz de bucle de pruebas asegurara que la sesión BGP no es necesaria en cualquier interfaz de hardware que pueda ser problemática.

No es necesario añadir interfaces de bucle de prueba en cada situación (realmente requiere mas configuración), Si los vecinos BGP externos están conectados directamente y las direcciones IP del segmento conectado directamente se utilizan para la negociación por vecindad, una dirección de bucle de prueba no añade ningún valor. Si el enlace físico entre dos iguales es problemático, la sesión se romperá con o sin bucle de prueba.

4.2 Sincronización dentro de un SA

Por definición, el comportamiento predeterminado de BGP requiere que deba ser sincronizado con el IGP antes de que BGP pueda publicar las rutas de transito hacia SA externos. Es importante que su SA sea coherente con las rutas que publica para evitar trafico de agujeros negros innecesario. Por ejemplo, si un portavoz IBGP fuese a publicar una ruta a un igual externo antes de que todos los routers dentro de su SA la hubiesen aprendido a través del IGP, su SA podría recibir tráfico hacia destinos para los que algunos de los routers podrían no tener aun la información sobre como llegar.

Cuando un router recibe una actualización sobre un destino desde un igual IBGP, el router intenta verificar la accesibilidad interna para dicho destino antes de publicarla a otros iguales EBGP. El router lo hace, primero, comprobando el prefijo de destino

antes de ver si existe una ruta hacia el router de próximo salto, y segundo, para ver si-existe un prefijo de destino en el IGP. Esta comprobación de router indica si los routers no BGP pueden desviar trafico a ese destino. Suponiendo que IGP reconozca dicho destino, el router lo publica a los otros iguales EBGP. De otro modo, el router trata el prefijo de destino como si no estuviera sincronizado con el IGP y no lo publica.

Considere la situación ilustrada en la Figura ISP1 e ISP2 utilizan ISP3 como SA de transito, ISP3 tiene multiples routers en su SA y esta ejecutando BGP solo en los router fronterizos (incluso aunque RTB y RTD estén transportando tráfico de transito, ISP3 no ha configurado BGP en esos routers). ISP3 esta ejecutando un Protocolo de gateway interior dentro del SA para conectividad interna.

Suponga que ISP1 esta publicando la ruta 192.213.1.0/24 a ISP3. Dado que RT A Y RTC están ejecutando IBGP, RTA propaga la ruta hacia RTC. Observe que otros routers además de RTA Y RTC no están ejecutando BGP y no tienen conocimiento mas allá de la existencia de la ruta 192.213.1.0/24.

En la situación ilustrada en la Figura, si RTC publica la ruta a ISP2, el tráfico hacia el destino 192.213.1.0/24 comenzara a fluir hacia RTC. RTC ejecutara una búsqueda en su tabla de enrutamiento IP y dirigirá el trafico hacia RTB. RTB, no teniendo visibilidad de las rutas BGP, cortara el tráfico porque no tiene conocimiento del destino. El tráfico se corta porque BGP e IGP no están sincronizados.

La regia de BGP dice que un router BGP no debería publicar a vecinos externos destinos aprendidos de vecinos IBGP a menos que dichos destinos se conozcan a través de un IGP. Esto se conoce como sincronización. Si un router conoce dichos destinos mediante un IGP, supone que la ruta ya ha sido propagada dentro del SA y

que la accesibilidad interna esta asegurada.

La consecuencia de insertar rutas BGP dentro de un IGP es costosa. Redistribuir las rutas desde BGP hacia el IGP dará como resultado una sobrecarga mayor de los routers internos, principalmente desde una perspectiva de escalabilidad en IGP, porque, los IGP no están diseñados para administrar tantas rutas. Además, transportar todas las rutas externas dentro de un SA no es necesario. EI enrutamiento puede ser fácilmente llevado a cabo teniendo routers internos no BGP por defecto hacia uno de los routers BGP. Por supuesto, esto supondrá un enrutamiento no optimo porque no garantiza que se utilizara el camino más corto para cada ruta, pero este coste es mínimo comparado con el de mantener miles de rutas dentro de un SA. Por supuesto, administrar rutas por defecto en una situación como esta puede ser extremadamente complejo y puede dar lugar a bucles de enrutamiento.

Sin embargo, la mayoría de las implementaciones BGP ofrecen una opción software que permite al operador de la red desactivar la sincronización. Como sospecharemos, configurando el subcomando BGP **no synchronization** de Cisco ordenara a BGP ignorar el requisito de sincronización y permitir publicar rutas aprendidas vía IBGP, sin tener en cuenta la existencia. de una ruta IGP. En la practica, la mayoría de las situaciones permiten que la sincronización se desactive de forma segura en los routers fronterizos, suponiendo que todos los routers de transito del SA están ejecutando una malla IBGP. En esta situación, la accesibilidad interna esta garantizada porque una ruta que es aprendida vía EBGP en cualquier router fronterizo será automáticamente transmitida vía BGP a todos los routers de transito.

Dicho esto, la configuración más común en las redes conectadas a Internet es

desactivar la sincronización BGP y confiar en una malla de routers IBGP. La idea de insertar decenas de miles de rutas en un IGP es bastante aterradora.

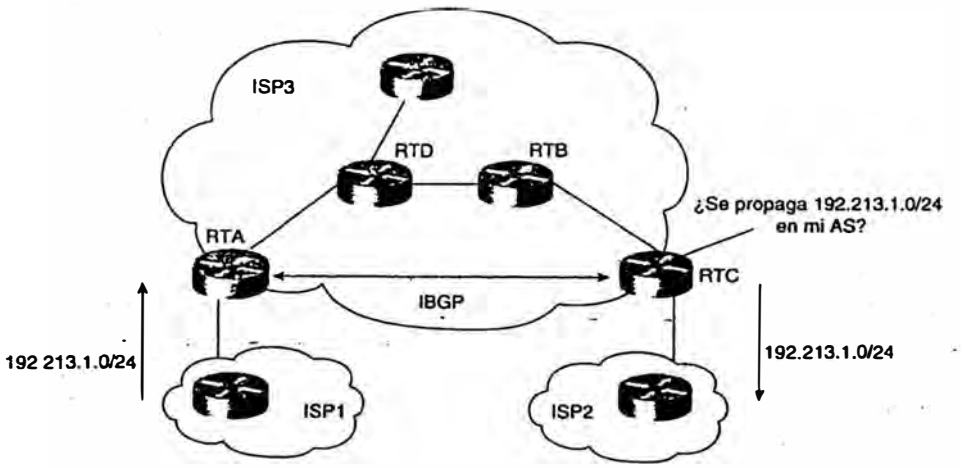


Figura 4.3 Sincronización dentro de un SA.

4.3 Solapamiento de protocolos, puertas traseras

Con diferentes IGP y EGP trabajando juntos para lograr el enrutamiento, las rutas pueden ser aprendidas mediante distintos protocolos; elegir un protocolo sobre otro afecta a cómo fluye el tráfico. Por ejemplo, si el tráfico sigue una ruta RIP, podría cruzarse con un enlace, mientras que si sigue una ruta BGP externa, podría acabar en otro enlace. Los enlaces de puerta trasera ofrecen una ruta IGP alternativa que puede ser utilizada en lugar de la ruta BGP externa. Las rutas IGP que pueden ser alcanzadas por el enlace de puerta trasera se denominan rutas de puerta trasera. Con la existencia de tales rutas alternativas, se necesita un mecanismo que otorgue preferencia a un protocolo sobre otro, se tiene un parámetro de preferencia llamado **distancia administrativa** de un protocolo. Cuanto más baja sea la distancia

administrativa de un protocolo de enrutamiento, más alta será la preferencia del protocolo.

Debería destacarse que la distancia administrativa es un parámetro que es relativo solo para el router configurado localmente y no es conocido por, ni comunicado a, otros routers del SA. Por ello, si intenta modificar la distancia administrativa de un router del SA, es altamente recomendable que modifique la distancia administrativa en todos los routers del SA para garantizar una decisión de enrutamiento consistente. La Tabla enumera las distancias de acuerdo con la implementación de Cisco.

Protocolo	**Distancia**
Conectado directamente	0
Estático	1
EBGP	20
EIGRP (intemo)	90
IGRP	100
OSPF	110
ISIS	115
RIP	120
EGP	140
EIGRP (externo)	170
IBGP	200
BGP local	200
Desconocido	255

En el ejemplo se ilustra el uso de rutas de puerta trasera. En la figura, AS1 esta

recibiendo actualizaciones sobre NetA de dos fuentes diferentes. AS1 esta recibiendo rutas vía EBGP en el enlace hacia AS3 y mediante el enlace de puerta trasera. ejecutando RIP entre AS1 y AS2. De acuerdo con la Tabla, el router dará automáticamente una distancia de 20 ala ruta EBGP y una distancia de 120 a la ruta RIP. En AS1, los routers que aprenden la ruta vía EBGP (routers fronterizos del SA) instalaran la distancia mas baja de la tabla de enrutamiento. Por tanto, el trafico hacia NetA seguirá la ruta BGP indirecta vía AS3 y luego AS2, en lugar de la ruta RIP directa vía AS2. "

Se proporciona una forma de obligar alas rutas IGP a tomar precedencia sobre las rutas EBGP. El concepto es simple. Las rutas especificas de EBGP pueden ser etiquetadas como rutas de puerta trasera, lo que establece que la distancia de dichas rutas sea la misma que la distancia de la ruta "BGP local" (por defecto, es 200). De acuerdo con la Tabla , esta distancia es superior a la de cualquier ruta IGP aprendida, y se preferirá la ruta IGP de puerta trasera.

Alternativamente, como se explico previamente, otra opción es utilizar el subcomando distance de BGP para alterar la distancia administrativa de todos los prefijos BGP aprendidos por el router.

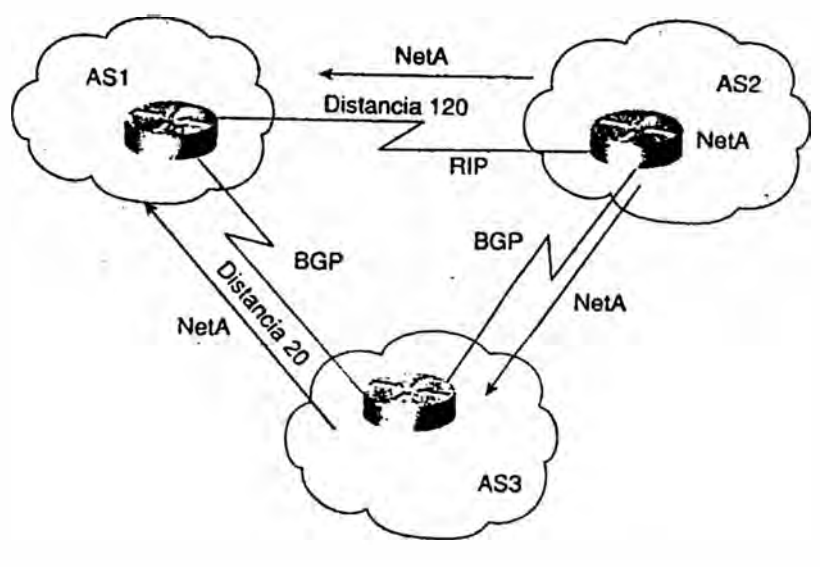


Figura 4.4 Variación de la distancia administrativa.

4.4 RESUMEN DEL PROCESO DE DECISIÓN DE BGP.

BGP basa su proceso de decisión en el valor de los atributos. Cuando se encuentra con muchas rutas de la misma longitud de prefijo hacia el mismo destino, BGP elige la mejor ruta para el tráfico de enrutamiento hacia el destino. El siguiente proceso resume como BGP elige la mejor ruta:.

1. Si el próximo salto es inaccesible, la ruta se ignora (debido a esto es importante tener una ruta IGP al siguiente próximo salto).
2. Prefiere el camino con mayor peso .
3. Si los pesos son iguales, prefiere la ruta con el mayor valor de preferencia local.
4. Si no hay rutas originadas localmente y la preferencia local es la misma, prefiere la ruta con el AS_PATH mas corto.
5. Si la longitud de AS_PATH es la misma, prefiere la ruta con el tipo de

origen mas bajo (donde IGP es mas bajo que EGP, y EGP es mas bajo que INCOMPLETE).

6. Si el tipo de origen es el mismo, prefiere la ruta con el valor MED mas bajo si las rutas fueron recibidas del mismo SA (o si bgp always-compare-med esta habilitado).
7. Si las rutas tienen el mismo valor MED, prefiere las rutas EBGP alas rutas IBGP.
8. Si todos los escenarios precedentes son idénticos, prefiere la ruta que puede ser alcanzada mediante el vecino IGP mas próximo (es decir, tomar la ruta interna mas corta dentro del SA para llegar al destino).
9. Si la ruta interna es la misma, prefiere la ruta que proceda del vecino con ROUTER ID más bajo.

CAPÍTULO V

MANEJO DE ATRIBUTOS BGP.

Los atributos BGP son un conjunto de parámetros utilizados para llevar la cuenta de la información específica de ruta como la información de ruta de acceso, grado de preferencia de una ruta, el valor NEX_HOP de una ruta y la información de agregación. Estos parámetros se utilizan en el filtrado de BGP y en el proceso de decisión de ruta. Cada mensaje UPDATE tiene una secuencia de longitud variable de atributos de ruta de acceso. Un atributo de ruta de acceso es una tripleta <tipo de atributo, longitud de atributo, valor de atributo>. El tipo de atributo es un campo de 2 bytes que consta de un flag de atributo de 1 byte y un código de tipo de atributo de 1 byte. La Figura ilustra la forma general del campo tipo Atributos de ruta de acceso.

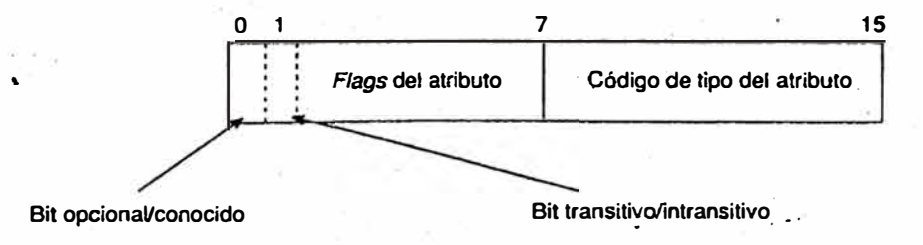


Figura 5.1 Formato general de un atributo BGP.

La siguiente tabla muestra los atributos y la documentación RFC respectiva.

Número de atributo	Nombre de atributo	Categoría/código de tipo	RFC/boceto Internet relacionado
1	ORIGIN	Obligatorio conocido, Código de tipo 1	RFC 1771
2	AS_PATH	Obligatorio conocido, Código de tipo 2	RFC 1771
3	NEXT_HOP	Obligatorio conocido, Código de tipo 3	RFC 1771
4	MULTI_EXIT_DISC	Intransitivo opcional, Código de tipo 4	RFC 1771
5	LOCAL_PREF	Discrecional conocido, Código de tipo 5	RFC 1771
6	ATOMIC_AGGREGATE	Discrecional conocido, Código de tipo 6	RFC 1771
7	AGGREGATOR	Transitivo opcional, Código de tipo 7	RFC 1771
8	COMMUNITY	Transitivo opcional, Código de tipo 8	RFC 1997[1]
9	ORIGINATOR_ID	Intransitivo opcional, Código de tipo 9	RFC 1966[2]
10	Lista de grupos	Intransitivo opcional, Código de tipo 10	RFC 1966
11	DPA	Atributo del Punto de Destino para BGP	Boceto Internet expirado.
12	Publicador	Servidor de Ruta BGP/IDRP	RFC 1863[3]
13	RCID_PATH/CLUSTER_ID	BGP/IDRP Servidor de Ruta	RFC 1863
14	Multiprotocolo alcanzable NLRI	Intransitivo opcional, Código de tipo 14	RFC 2283[4]
15	Multiprotocolo no alcanzable NLRI	Intransitivo opcional, Código de tipo 15	RFC 2283
16	Comunidades extendidas		draft-ramachandra-bgp-ext-communities-00.txt, "trabajo en progreso".
256		Reservado para el desarrollo	

Tabla 5.1 Atributos y documentación RFC respectiva

5.1 ATRIBUTOS DE RUTAS BGP.

Los atributos BGP son un conjunto de parámetros que describen las características de un prefijo (ruta). El proceso de decisión de BGP empareja esos atributos con el prefijo que describen, compara todas las rutas disponibles para llegara un destino dado, y luego selecciona las mejores rutas que serán utilizadas para llegar a ese

destino. Recordemos que los atributos son parte de cada paquete UPDATE de BGP y describen la información de ruta del prefijo asociado. Las siguientes secciones abordan esos atributos y como pueden ser manipulados para afectar al comportamiento del enrutamiento.

5.1.1 El Atributo ORIGIN.

El atributo ORIGIN es un atributo conocido obligatorio (código de tipo 1) que indica el origen de la actualización de enrutamiento con respecto al sistema autónomo que la originó. BGP considera tres tipos de orígenes:

-IGP. La Información de accesibilidad de la capa de red (NLRI) es interna para el SA que la origina.

-EGP. La Información de accesibilidad de la capa de red es aprendida a través del Protocolo de gateway exterior (EGP).

-INCOMPLETE. La Información de accesibilidad de la capa de red es aprendida por otros medios.

BGP considera el atributo ORIGIN en su proceso de toma de decisión para establecer un ranking de preferencia entre múltiples rutas. Específicamente, BGP prefiere la ruta con el tipo de origen mas bajo, donde el IGP es mas bajo que EGP y EGP es mas bajo que INCOMPLETE.

5.1.2 El atributo AS_PATH.

El atributo AS_PATH es un atributo conocido obligatorio (código de tipo 2) que contiene una secuencia de números de sistemas autónomos que representan el camino por el que atraviesa una ruta. Internamente para un SA, las rutas pasadas

entre porta voces BGP dejan la información AS_PATH intacta; sın embargo, cuando se envían rutas a iguales BGP externos, el SA que origina la ruta añade su propio numero de SA. De allí en adelante, cada SA que recibe la ruta y la transmite hacia otros iguales EBGP añadirá su número de SA hacia la lista. EI mecanismo conocido como prepending es el acto de añadir el numero de SA al comienzo de la lista. La lista final representa todos los números de SA por los que una ruta ha atravesado. EI numero de SA del SA que origino la ruta queda al final de la lista (justo antes del código ORIGIN). Este tipo de lista AS_PATH se conoce como AS_SEQUENCE, porque todos los números de SA están ordenados secuencialmente.

BGP utiliza el atributo AS_PATH como parte de las actualizaciones de enrutamiento (paquete UPDATE) para asegurar una topología libre de bucles en Internet.

La Figura ilustra el atributo AS_PATH, la ruta 172.16.10.0/24, originada en AS1 y pasada a AS2, luego a AS3 y a AS4, y de vuelta a AS1. Observe como cada SA que pasa la ruta a los otros iguales externos añade su propio numero de SA al principio de la lista. Cuando la ruta vuelve a AS1, el router BGP fronterizo de AS1 se da cuenta de que esa ruta ya ha pasado por su SA (el numero 1 de SA aparece en la lista) y no aceptara la ruta.

La información de AS PATH es uno de los atributos que BGP utiliza para determinar la mejor ruta a tomar para llegar a un destino. Comparando dos o más rutas diferentes, dado que todos los atributos de prioridad superior son iguales, siempre se prefiere un AS_PATH mas corto sobre uno más largo.

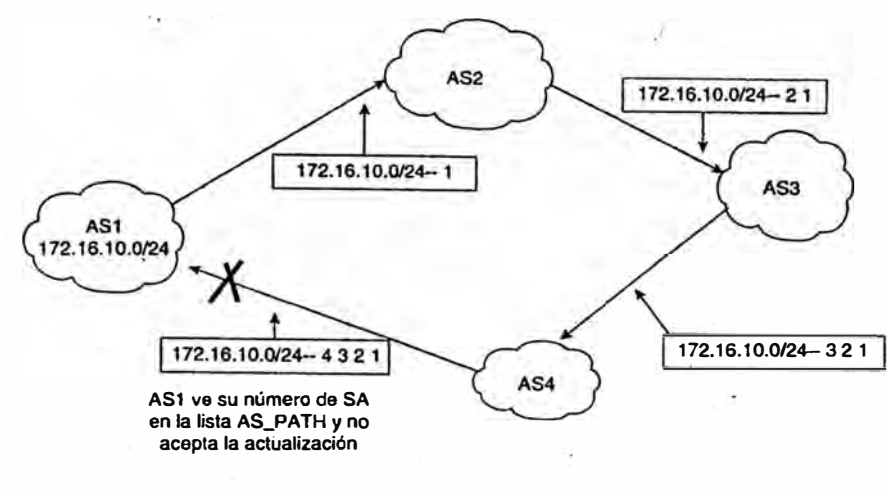


Figura 5.2 Actualización de los AS_PATH.

5.1.3 EI atributo NEXT HOP.

EI atributo NEXT_HOP es un atributo conocido obligatorio (código de tipo 3). Varia ligeramente cuando se utiliza en el contexto de un IGP, donde el próximo salto para llegar a un destino es la dirección IP de la interfaz del router conectado que anunció la ruta.

EI concepto de próximo salto con BGP esta ligeramente mas elaborado. Toma una de las cuatro formas siguientes:

- Para sesiones EBGP, el próximo salto es la dirección IP del vecino que anunció la ruta.
- Para las sesiones IBGP, para rutas originadas dentro del SA, el próximo salto es la dirección IP del vecino que anunció la ruta.
- Para las rutas insertadas en el SA vía EBGP; el próximo salto aprendido de EBGP es transportado sin alteraciones hacia IBGP. EI próximo salto es la dirección IP del vecino EBGP del que la ruta fue aprendida.

- Cuando la ruta se publica en un medio multiacceso (como Ethernet, Frame Relay, y demás), el próximo salto normalmente es la dirección IP de la interfaz del router conectado al medio que originó la ruta.

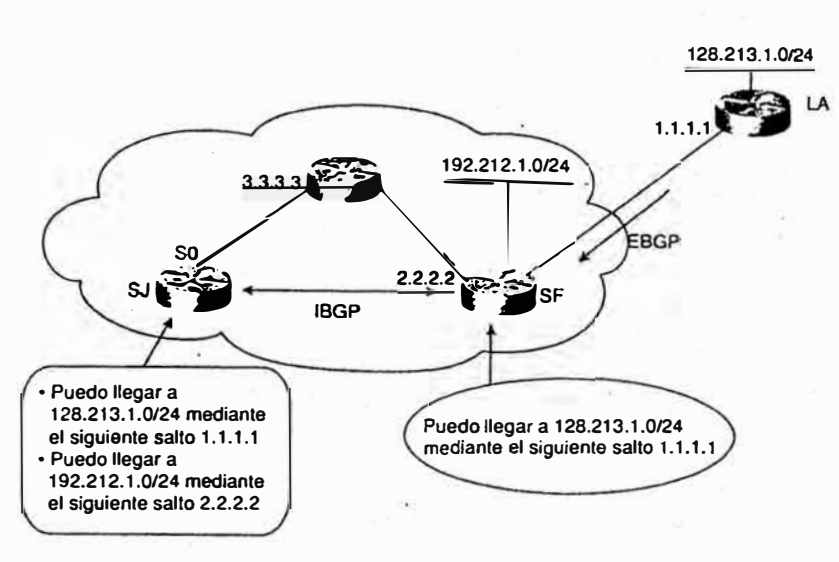


Figura 5.3 Manejo del atributo NEXT_HOP.

Como podemos ver en este ejemplo, el próximo salto no es necesariamente accesible a través de una conexión directa. Por ejemplo, el próximo salto de SJ para 128.213.1.0/24 es 1.1.1.1. De esta forma, el comportamiento del próximo salto ordena una búsqueda IP recursiva de un router para saber dónde enviar el paquete. Para llegar al próximo salto 1.1.1.1, el router de 5J mirara recursivamente en su tabla de enrutamiento IGP para ver si, y cómo, se puede llegar a 1.1.1.1. Esta búsqueda recursiva continua hasta que el router asocia el destino 1.1.1.1 con una interfaz saliente. EI mismo comportamiento recursivo es ejecutado para llegar al próximo salto 2.2.2.2. Si un salto no puede ser alcanzado, BGP considera la ruta inaccesible.

5.1.4 EI atributo MULTL_EXIT_DISC (MED).

EI atributo Discriminador multisalida de BGP (MULTI_EXIT_DISC 0 MED) es un atributo opcional (código de tipo 4). Es un indicio para los vecinos externos sobre la ruta preferida en un SA que tiene múltiples puntos de entrada. EI valor MED también es conocido como la métrica externa de una ruta. Un valor mas bajo de MED es preferible sobre un valor mas alto. A diferencia de LOCAL_PREF, el atributo MED se intercambia entre los SA, pero un atributo MED que es recibido por un SA no abandona el SA. Cuando una actualización entra en el SA con un cierto valor de MED, dicho valor se utiliza para tomar decisiones dentro del SA. Cuando BGP pasa la actualización de enrutamiento a otro SA, el valor de MED se reinicia a 0 (a menos que el MED saliente este asignado explícitamente a un valor especifico). Cuando la ruta es originada por el SA en si mismo, la practica más común es que el valor MED siga la métrica IGP interna de la ruta. Esto es útil cuando un cliente tiene múltiples conexiones al mismo proveedor.

En el ejemplo ilustrado en la Figura el MED muestra como un SA puede influir en la decisión saliente de otro SA. En la Figura, ANET e YNET intentan influir en el trafico saliente de XNET mediante el envío de diferentes valores de MED.

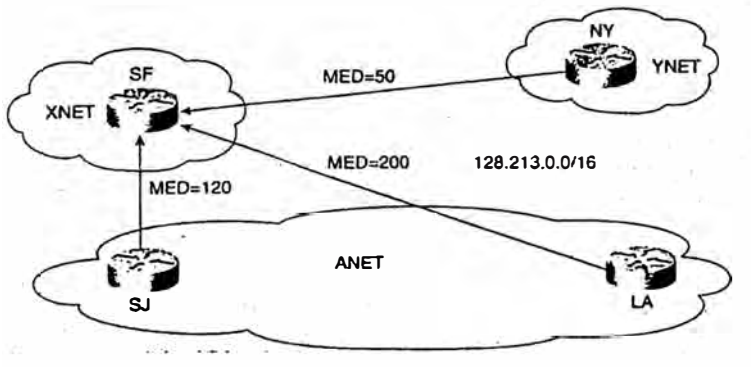


Figura 5.4 Influencia del atributo MED.

XNET esta recibiendo actualizaciones de enrutamiento sobre 128.213.0.0/16 desde tres orígenes diferentes: SJ (MED 120), LA (MED 200) y NY (MED 50). SF comparara los dos valores de métrica que vienen de ANET (mismo AS) y preferirá el router de SJ porque esta publicando una métrica mas baja (120). Cuando el comando bgp always-compare-med se utilice en el router de SF, comparara la métrica 120 con la métrica 50 procedente de NY y preferirá NY para llegar a 128.213.0.0/16.

5.1.5 EI atributo LOCAL PREFERENCE.

EI atributo de preferencia local (LOCAL_PREF) es un atributo conocido y discrecional (código de tipo 5). El atributo de preferencia local es un grado de preferencia dado a una ruta para compararla con otras rutas hacia el mismo destino. Un valor de preferencia local superior indica que la ruta es mas preferida. La preferencia local, como su propio nombre indica, es local al sistema autónomo y se intercambia solo entre iguales IBGP. Un SA conectado vía BGP a otros SA obtendrá actualizaciones de enrutamiento sobre el mismo destino de diferentes SA. La preferencia local normalmente se utiliza para establecer el punto de salida de un SA para llegar a un cierto destino. Dado que este atributo es comunicado dentro de los routers BGP dentro del SA, todos los routers BGP tendrá una vista. común de como salir del SA.

Consideremos el entorno ilustrado en la Figura. Suponga que la empresa ANET ha adquirido conexiones a Internet mediante dos proveedores de servicio, XNET e YNET. ANET esta conectada a YNET mediante un enlace T3 primario y a XNET mediante un enlace T1 de respaldo.

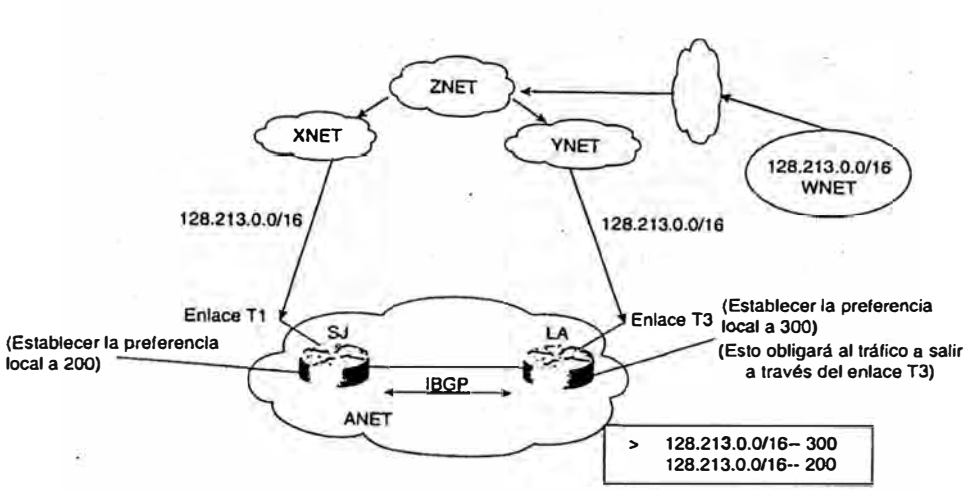


Figura 5.5 Influencia del atributo LOCAL_ PREFERENCE en el tráfico.

Es importante para ANET decidir que ruta tomara el tráfico saliente. Por supuesto, ANET prefiere utilizar el enlace T3 vía YNET en un funcionamiento normal porque es un enlace de alta velocidad. Aquí es donde la preferencia local entra en juego: EI router de LA asigna una preferencia local de 300 a las rutas recibidas de YNET. EI router de SJ designa un valor mas bajo de, digamos, 200 alas rutas recibidas de XNET. Dado que los routers tanto de LA como de SJ intercambian actualizaciones vía IBGP, ambos están de acuerdo en que el punto de salida del SA será vía YNET debido a la preferencia local superior.

5.1.6 EI atributo COMMUNITY.

En el contexto de BGP, una comunidad es un grupo de destinos que comparten algún tipo de propiedad común. Una comunidad no esta restringida a una red o a un sistema autónomo; no tiene limites físicos. Un ejemplo es un grupo de redes que pertenecen alas comunidades educacionales o gubernamentales. Esas redes pueden pertenecer a

cualquier sistema autónomo. Las comunidades son utilizadas para simplificar las políticas de enrutamiento mediante la identificación de rutas basadas en una propiedad lógica en lugar de en un prefijo IP o en un número de SA. Un portavoz BGP puede utilizar este atributo en conjunción con otros para controlar las rutas que se aceptarán, preferirán y transmitirán a otros vecinos BGP.

El atributo COMMUNITY (código de tipo 8) es un atributo opcional transitivo. Es de longitud variable y consta de un conjunto de valores de 4 bytes. Existen comunidades reservadas estas comunidades son conocidas; es decir, tienen un significado global, ejemplos de comunidades conocidas:

NO_EXPORT: Una ruta transportando el valor de esta comunidad no debería ser publicada a iguales fuera de un SA.

NO_ADVERTISE : Una ruta transportando este valor de comunidad, cuando se recibe, no debería ser publicada a ningún igual BGP.

A pesar de los atributos de comunidad conocidos, los atributos de comunidad privada pueden ser definidos para usos especiales. También pueden los definidos en la RFC 19981, que describe un mecanismo por el que las comunidades pueden ser utilizadas para manipular la selección de ruta BGP en las redes delos proveedores de servicio.

Una practica común es utilizar los dos primeros bytes del atributo de comunidad para el número de SA y los dos últimos bytes para definir un valor en relación a dicho SA. Por ejemplo, un proveedor (AS256) que desea definir una comunidad privada llamada mis rutas-de-igual podría utilizar la comunidad **256:1** representada en notación decimal. EI 256 indica que este proveedor particular ha definido la comunidad. El 1 tiene un significado especial para el proveedor. En este caso, es mis-rutas-de-igual.

La Figura muestra un uso simple del atributo COMMUNITY. XNET esta enviando hacia YNET las rutas X e Y con un atributo de comunidad NO_EXPORT, y la ruta Z sin modificación. EI router BGP de YNET propagara sólo la ruta Z hacia ZNET. Las rutas X e Y no serán propagadas debido al atributo de comunidad NO EXPORT.

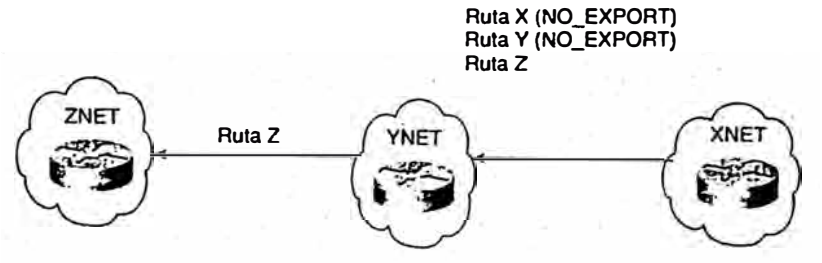


Figura 5.6 Influencia del atributo COMMUNITY.

5.2 GRUPOS DE IGUALES.

Un grupo de iguales BGP es un grupo de vecinos BGP que comparten las mismas políticas de actualización. En lugar de definir las mismas políticas para cada vecino individual, se define un nombre para el grupo de iguales y se le asignan políticas. Por ejemplo, un administrador que establece políticas para su iguales BGP establecerá probablemente las mismas políticas para la mayoría de sus iguales, definiéndolos de este modo como un grupo de iguales.

No solo los grupos de iguales ahorran al operador la configuración repetitiva de cada igual BGP, sino que además le ahorran al propio router BGP el esfuerzo de analizar las políticas secuencialmente para cada vecino.

5.3 AGREGACIÓN BGP-4

Una de las principales mejoras de BGP-4 sobre versiones anteriores de BGP es la capacidad de administrar CIDR y las superredes como un medio de controlar el crecimiento de las tablas de enrutamiento IP y el agotamiento del espacio de direcciones IP.

La agregación se aplica a rutas que existen en la tabla de enrutamiento de BGP. Esto contrasta con el comando network, que se aplica a las rutas que existen en la tabla de enrutamiento IP. La agregación puede ser ejecutada si al menos una ruta mas especifica del agregado existe en la tabla de enrutamiento BGP.

Se tiene una gran variedad de formas para manipular agregados a fin de asegurarse que se satisfacen todas las necesidades en Internet.

5.3.1 Solo agregación, suprimiendo las rutas más especificas

Este escenario ilustra un caso en el que un agregado es publicado y todas sus rutas especificas suprimidas. Esto se realiza normalmente cuando las rutas mas especificas no ofrecen ningún beneficio adicional, tal como tomar mejores decisiones en el reenvío del tráfico.

La Figura ilustra una situación en la que todas las actualizaciones de enrutamiento son agrupadas en un único agregado. Suponga que AS100 tiene el rango de subred 172.16.0.0/24 a 172.16.15.0/24. Esto incluye 172.16.0.X, 172.16.1.X y demás. La lista de prefijos específicos puede resumirse en el rango 172.16.0.0/20. El agregado 172.16.0.0/20 es enviado, y todos los prefijos más específicos son suprimidos.

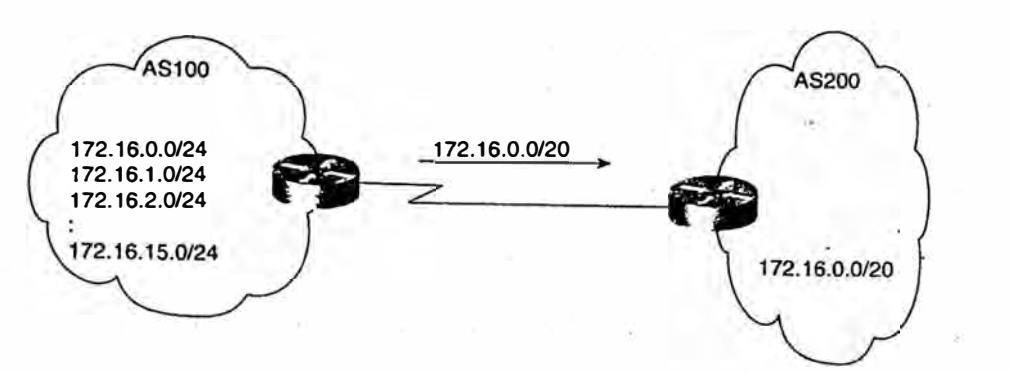


Figura 5.7 Agregación sin atributos específicos.

5.3.2 Agregación de más rutas mas especificas.

Existe un cierto número de situaciones en las que un SA enviara fuera un agregado, así como sus rutas mas especificas. Esto normalmente ocurre en situaciones en las que el cliente tiene múltiples conexiones con un mismo proveedor. El proveedor utilizara las rutas mas especificas para tomar mejores decisiones cuando envíe tráfico hacia el cliente. AI mismo tiempo, el proveedor puede propagar el agregado solo hacia el NAP para minimizar el número de rutas propagadas hacia Internet.

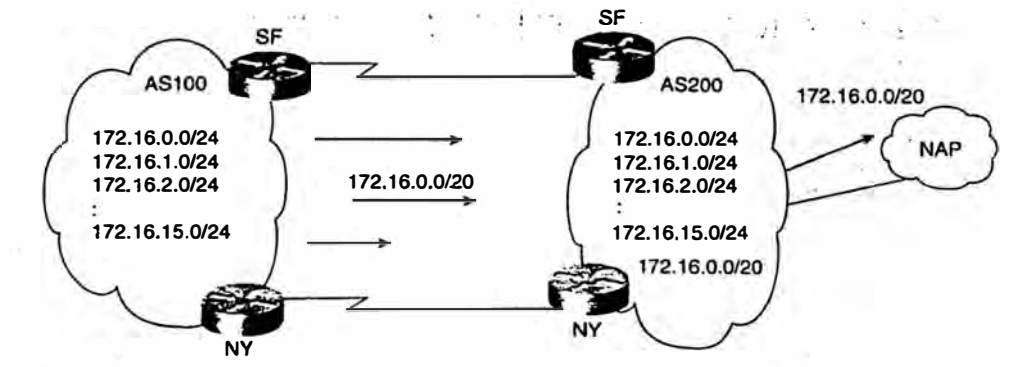


Figura 5.8 Agregación con atributos específicos.

CAPÍTULO VI

CONTROL DE ENRUTAMIENTO DENTRO DE SISTEMAS AUTÓNOMOS.

6.1 Control de los sistemas autónomos de gran potencia.

Los sistemas autónomos compuestos por cientos de nodos enrutados pueden suponer un serio problema de manipulación del enrutamiento para los administradores. Cada uno de los proveedores de servicios y clientes tiene su propio conjunto de problemas cuando se trata de lidiar con las redes grandes. En el lado del proveedor de servicio la mayoría de routers ejecutan el Protocolo de gateway fronterizo (BGP). Debido a la norma de BGP que afirma que un portavoz IBGP (Protocolo de gateway fronterizo interior) no puede publicar una ruta aprendida de otro portavoz IBGP a un tercer portavoz IBGP, la malla BGP puede crecer rápidamente mas allá del control del proveedor. Sin embargo, en el lado del cliente la mayoría de los routers ejecutan el Protocolo de gateway interior (IGP), que también podría crecer mas allá del control del cliente.

A continuación explicaremos los métodos y las técnicas que podemos utilizar para controlar mejor el despliegue de BGP e IGP dentro de los sistemas autónomos grandes.

6.2 Reflectores de ruta.

En las redes de algunos ISP, la malla BGP interior puede crecer bastante (más de 100 sesiones BGP interiores por router), lo que sugiere la implementación de algún mecanismo de peering el concepto de reflector de ruta1 esta basado en la idea de especificar un router de concentración para que actúe como un punto focal de las sesiones BGP interiores.

Muchos de estos routers BGP pueden conectarse como iguales con un servidor central (el reflector de ruta), y después los reflectores conectarse como iguales entre sí. Aunque las normas de BGP afirman que las rutas aprendidas a través de un portavoz IBGP no pueden publicarse a otro portavoz IBGP, la reflexión de ruta permite a los servidores reflectores de ruta "reflectar" rutas como se explicara posteriormente.

Sólo se recomiendan los reflectores de ruta par los SA que tienen una malla BGP interior grande. El concepto de reflector de ruta introduce el consumo del proceso en el servidor reflector de ruta y, si se configuró incorrectamente, podrían provocarse bucles en el enrutamiento e inestabilidad en el mismo. Como resultado, no se recomiendan los reflectores de ruta para cada topología.

La reflexión de ruta supone algunas ventajas para los servidores reflectores de ruta y sus dientes. Por ejemplo, la implementación de un servidor reflector de ruta podría optimizarse para simplificar la copia de los mensajes UPDATE cuando se envían a múltiples iguales, en lugar de generar mensajes únicos para cada igual. Además, los dientes normalmente solo se conectan como iguales con el servidor reflector de ruta local, disminuyendo significativamente por ello el número de sesiones que deben mantener.

6.2.1 Iguales internos sin reflectores de ruta.

Sin los reflectores de ruta, los porta voces BGP en un SA tendrán una malla lógica. En la Figura , RT A, RTB Y RTC forman una malla lógica BGP interior.

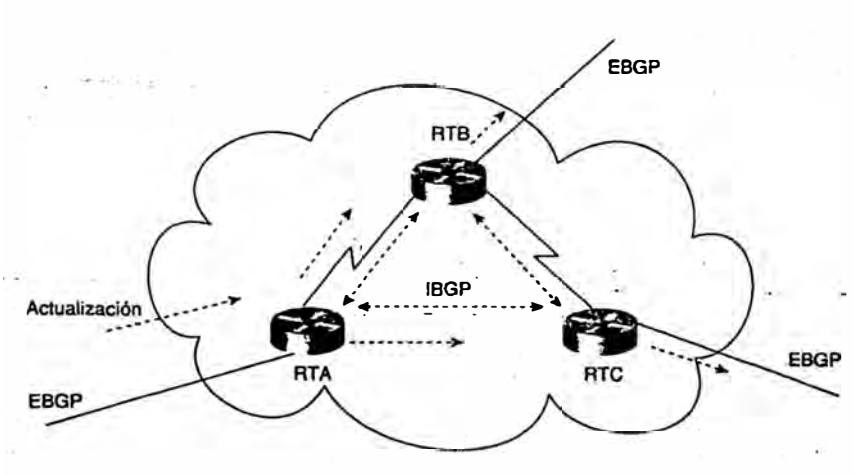


Figura 6.1 Sesiones BGP sin reflectores de rutas.

Cada router actúa como un igual BGP con los otros dos routers. RTA y RTB están conectados físicamente, como RTB y RTC. No existe ninguna conexión física entre RTA y RTC.

Cuando RTA recibe una actualización de un igual externo, la remite a sus dos iguales interiores, RTB y RTC, aunque no existe una conexión física entre RTA y RTC, RTA pasa la actualización a RTC a través de la sesión de peering BGP. RTB y RTC, a su vez, pasan la actualización a sus iguales externos.

EI mensaje UPDATE que RTB recibe de RT A no se vuelve a publicar a RTC, puesto que este es un igual interior, y el mensaje UPDATE que RTB recibió procedía de un igual interior (RT A). Sin una sesión BGP interna entre RT A y RTC, RTC no obtendría nunca la actualización; por tanto, es necesaria la malla IBGP.

6.2.2 Iguales internos con reflectores de ruta.

EI reflector de ruta actúa como punto de concentración para otros routers referidos como clientes. Los clientes se conectan como iguales con el reflector de ruta e intercambian información de enrutamiento con el. A su vez, el reflector de ruta pasa (0 refleja) la información entre los clientes y a los otros iguales IBGP y EBGP

En la Figura RTB esta configurado como un reflector de ruta con dos clientes, RTA y RTC, RTA recibe una actualización de un par externo y lo dirige a RTB. RTB refleja la actualización del cliente RTA al cliente RTC En esta configuración, no debería configurarse una sesión peering entre RTA Y RTC, puesto que el reflector de ruta esta propagando la información del protocolo BGP de RT A a RTC.

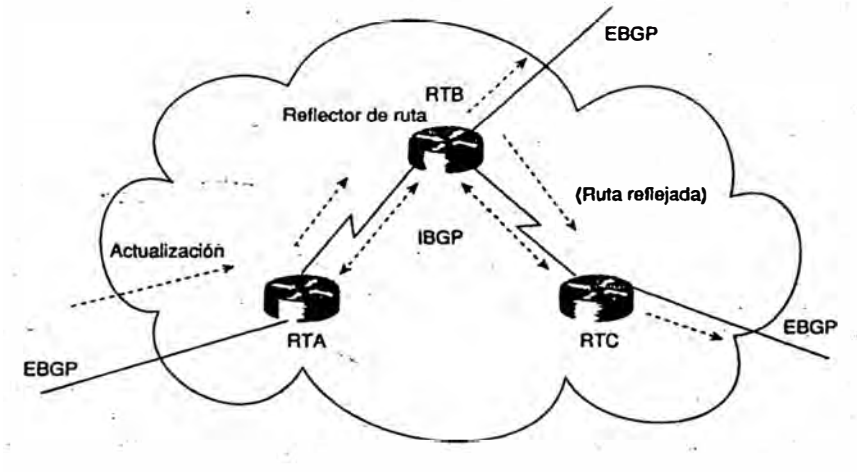


Figura 6.2 Sesiones BGP con reflectores de rutas.

En un SA donde el administrador tendría que construir un número sustancial de sesiones BGP entre los routers, el concepto de reflector de ruta proporciona una solución muy útil y escalable al problema.

6.3 CONFEDERACIONES.

Una confederación es otro modo de tratar con la explosión de una malla IBGP dentro de un SA. Igual que con la reflexión de ruta, las confederaciones sólo son recomendables para los casos en los que el peering IBGP involucra un gran numero de sesiones peering IBGP por router.

Las confederaciones BGP están basadas en el concepto de que un SA puede romperse en múltiples subSA. Dentro de cada subSA, se aplican todas las reglas de IBGP. Todos los routers BGP dentro del subSA, por ejemplo, deben estar completamente en malla. Como cada subSA tiene un numero de SA diferente, el BGP externo debe ejecutarse entre ellos. Aunque EBGP se usa entre los subSA, el enrutamiento dentro de la confederación se comporta como el enrutamiento IBGP dentro de un solo SA. En otras palabras, el próximo salto, MED, y la información de preferencia local se preserva al cruzar los limites del subSA. Para el mundo exterior, una confederación se asemeja a un solo SA.

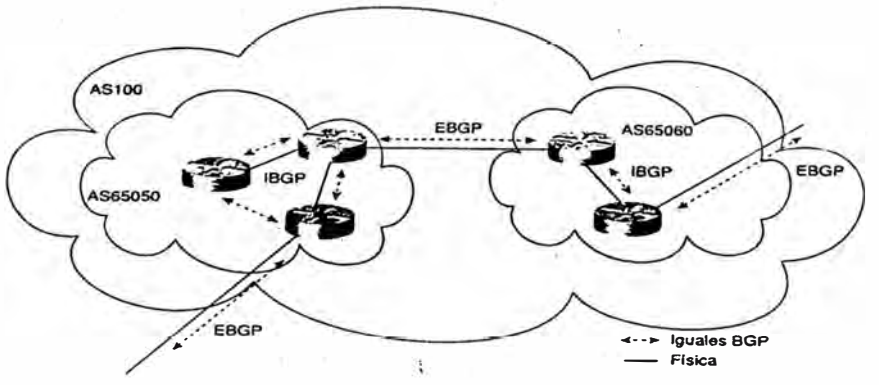


Figura 6.3 Divición de un AS mediante Confederaciones.

En la Figura, AS100 esta dividida en dos subSA: AS65050 y AS65060. EI SA,

como un todo, es ahora una gran confederación, identificada por un solo número de confederación, 100. Todos los subSA están protegidos del mundo exterior, y pueden recibir cualesquier números de SA. Los números podrían escogerse del rango de SA privados (64512 a 65534, como se especifica en la RFC 1930) para no usar ningún número de SA formal.

Como se menciono antes, dentro del subSA se usa una malla IBGP. EBGP se utiliza entre los subSA, así como entre la misma confederación y fuera de los SA. Las confederaciones pueden detectar fácilmente bucles de enrutamiento dentro del SA completo, porque EBGP se ejecuta entre los subSA. La lista de rutas del SA es un mecanismo para evitar los bucles usado para detectar actualizaciones de enrutamiento dejando un subSA e intentando reintroducir el mismo subSA. Una actualización de enrutamiento que intenta reintroducir un subSA que lo origino será detectado porque el subSA vera su propio número de subSA listado en la ruta de SA de actualización.

6.4 Inestabilidades de rutas en Internet y penalizaciones.

Uno de los modos en que los administradores impulsan sus redes hasta el limite es dejándolas crecer de tal modo que el protocolo IGP será difícil de manipular. Sea el protocolo IGP tan obsoleto como la versión 1 de RIP o tan avanzado como Primero la ruta libre más corta (OSPF) y Sistema intermedio-sistema intermedio (I5-IS), surgirá el tema de la escalabilidad. Un modo escalable de administrar la expansión de IGP es segmentar el SA en múltiples regiones, ejecutando cada una de ellas un solo y distinto protocolo IGP. Las regiones individuales, a su vez, deben conectarse a través de BGP.

Establecer y mantener la estabilidad de la ruta dentro y entre redes es crucial para asegurar una conexión fiable a Internet. Varios defectos de diseño y problemas pueden contribuir a desestabilizar las conexiones a Internet.

6.4.1 Inestabilidades de las rutas en Internet.

EI síntoma principal de la inestabilidad de las rutas es la desaparición de una ruta que existió previamente en la tabla de enrutamiento. Esta ruta podría desaparecer y reaparecer intermitentemente, una condición denominada a veces fluctuación. Lo que ocurre a nivel del protocolo de enrutamiento es que BGP envía una actualización de enrutamiento y luego la retira rápidamente. Un router que recibe mensajes UPDATE 0 WITHDRAWN debe propagar dichos mensajes a sus iguales. Estos mensajes son visibles para todas las redes del Protocolo de gateway fronterizo (BGP) conectadas a la Internet global. Si este comportamiento continua en cascada, el rendimiento del enrutamiento sufre.

He aquí algunos factores que afectan a la inestabilidad de las rutas en Internet:

- Inestabilidad qel Protocolo de gateway Interior (IGP).
- Hardware defectuoso.
- Problemas de software.
- Potencia de CPU insuficiente.
- Memoria insuficiente.
- Actualizaciones de red y mantenimiento de rutina.
- Error humano.
- Congesti6n del enlace.

6.4.2 Inestabilidad IGP.

Insertar dinámicamente IGP en BGP puede causar una fluctuación innecesaria de la ruta.

Los problemas que puedan ocurrir dentro de un dominio pueden traducirse en problemas fuera del dominio, la inserción estática del enrutamiento en BGP puede aliviar este problema.

La agregación de rutas en los routers fronterizos o del núcleo pueden reducir también los desagradables efectos laterales potenciales asociados con la inserción de IGP en BGP. Con la agregación, muchas entradas de ruta se introducen en BGP como un resumen agregado. La inestabilidad de una única ruta en cualquier elemento sencillo del agregado no afecta a la estabilidad del agregado en sí mismo.

Todavía, algunos diseñadores de red son obligados a confiar en el enrutamiento dinámico por razones aun validas:

. Las implementaciones de BGP solo pueden administrar un numero fijo de entradas de red para que sean publicadas estáticamente. El numero de rutas estáticas permitido varia de fabricante a fabricante. Cualquiera que sea el límite las redes que quieran traspasar este limite exigen que los administradores introduzcan IGP en BGP.

. Algunos administradores no están cómodos con el hecho de que las redes que se publican estáticamente puedan volverse inalcanzables por el router que las publica. Esto es comprensible, especialmente en los casos donde las rutas se publican desde diferentes puntos del SA. Publicar una ruta que no es accesible puede crear agujeros negros.

6.4.3 Hardware defectuoso.

Interfaces defectuosas, sistemas defectuosos o líneas defectuosas pueden afectar a la estabilidad de la ruta. Una interfaz que esta disponible intermitentemente puede hacer que la información de enrutamiento transite. Los fallos de hardware están, a un cierto grado, mas allá del control de los usuarios del servicio. La redundancia del sistema y el enlace son herramientas importantes para reducir la perdida de conectividad debida a fallos, pero cuando se produce un fallo físico, el enrutamiento se interrumpe, y cualquier interrupción inicia algún tipo de efecto en cascada por la ruta de enrutamiento

6.4.4 Errores del software.

Los errores del software (bugs) pueden provocar fallos en el sistema e inestabilidad de la red. Los equipos de desarrollo del protocolo de enrutamiento hacen todo lo que pueden para detectar dichos problemas antes de que el software llegue a los clientes. Sin embargo, es casi imposible prever todas las situaciones que podrían darse en las redes reales. Los administradores deberían experimentar con nuevo software o nuevas características en laboratorios de pruebas y en partes de bajo impacto de sus redes con el fin de obtener algún nivel de confianza antes de que el software se distribuya en un entorno de producción.

6.4.5 Potencia de CPU insuficiente.

Cuantas más actualizaciones de enrutamiento y sesiones de peering administre el router, mas potencia de CPU se requiere. Pensemos en el router como un pequeño veiculo 4 x 4, Y pensemos en el enrutamiento y el coste del trafico como en la carga que transporta. Nos sorprendería si el veículo tuviera problemas mientras transporta

una carga de 20 toneladas? Seleccionar el sistema correcto con la potencia de CPU correcta es muy importante para satisfacer sus necesidades particulares de enrutamiento.

En las etapas iniciales de la construcción de tablas BGP después de que se hayan establecido las sesiones BGP, un procesador de sistema puede pasar mas del 90% de su tiempo procesando actualizaciones. Cuando los enlaces se vuelven inestables y se sobrecargan, el router debería finalizar en una condición de lentitud la CPU esta demasiado ocupada gestionando actualizaciones, lo que causa que se caigan las sesiones BGP, lo que por su lado desencadena mas inestabilidad.

6.4.6 Memoria insuficiente.

Además de la memoria que necesita un router para ejecutar su propio sistema operativo, un router debe almacenar tablas de enrutamiento, bases de datos y otros paquetes de software para permitir el funcionamiento. Un router que alcanza su limite de memoria puede dejar de funcionar lo que provocaría la perdida de todas las rutas que conoce o publica.

En términos BGP, una entrada de enrutamiento consta de la entrada en la tabla de reenvío IF y cualquier información correspondiente esta disponible en la tabla de enrutamiento BGP. Actualmente las tablas de enrutamiento en Internet incluyen mas de 75.000 rutas, y este número se incrementa cada mes. Los sistemas que toman rutas completas de Internet de uno o mas proveedores apenas se mantienen (si es que se mantienen) con 32 MB de memoria (para almacenar BGP y otra información de enrutamiento). La mayoría de los proveedores han actualizado sus sistemas a 96, 128, e incluso 256 MB de memoria para la tabla de enrutamiento. La memoria

insuficiente por si misma a menudo produce inestabilidad, porque cuando en un router escasea la memoria, a menudo no se pueden recuperar los fragmentos de memoria, lo que se convierte en una fuente permanente (hasta que se reinicia) de fluctuaciones de la ruta.

Actualizaciones de red y mantenimiento de rutina, las redes son dinámicas. La mejora del rendimiento, la consolidación del sitio y la expansión del soporte requieren cambios y adaptaciones. Los cambios pueden incluir actualizaciones a nuevas versiones de software o hardware; adición de mas enlaces de más ancho de banda, o la reconfiguración de la composición de una red.

CAPITULO VII

CASOS PRACTICOS, ENRUTAMIENTO EN EL PERU.

7.1 Requisitos de conexión a Backbone de operadores Internacionales.

Las entidades que quieran conectarse a un operador Internacional han de cumplir los requisitos administrativos y técnicos que a continuación se detallan:

- Disponer de una conexión a Internet propia, es decir, que el proveedor no obtenga el acceso a Internet únicamente de su conexión al Punto Neutro.
- Los proveedores conectados al Punto Neutro han de figurar registrados en algún Registro Regional de Internet (IRR) de la IANA y estar al corriente del pago de las cuotas que les corresponda.
- Los proveedores habrán de disponer de su propio número oficial de Sistema Autónomo.
- Cada proveedor habrá de establecer acuerdos, como mínimo, de intercambio de tráfico con la totalidad de miembros presentes en el Punto Neutro para sus redes.
- Cada proveedor conectado al Punto Neutro ha de hacer públicos los datos de las personas de contacto técnico (para problemas operativos) y administrativo (para el establecimiento y seguimiento de los acuerdos de intercambio de tráfico) al resto

de miembros del Punto Neutro. Es obligatorio que todas las personas de contacto dispongan de la dirección de correo electrónico operativa.

- Cada proveedor ha de registrar previamente en el registro de enrutamiento del IRR de la IANA correspondiente, cualquier ruta y sistema autónomo que vaya a anunciar mediante el Punto Neutro, de acuerdo con los procedimientos que dispongan los IRR.
- Ningún proveedor podrá generar transiciones de enrutamiento ("routing flaps") innecesarias ni anunciar rutas específicas que no sean estrictamente imprescindibles.
- Todo proveedor habrá de establecer los filtros oportunos de enrutamiento para asegurarse de que únicamente anuncia las rutas y sistemas autónomos por él conectados, de acuerdo con la información registrada al IRR.
- El medio de acceso de todo proveedor será permanente y estará operativo las 24 horas del día, los 365 días al año (descontando averías y causas de fuerza mayor).
- El protocolo de enrutamiento a utilizar entre proveedores es el BGP-4, optimizando al máximo las capacidades de agregación permitidas por el CIDR.
- Los proveedores conectados no pueden llevar a cabo acciones que sean ilegales o que vayan en detrimento del uso del Punto Neutro de Internet por parte de otros proveedores. Estas acciones podrán ser motivo de expulsión.
- Aquellos proveedores que experimenten un estado de congestión grave, por exceso de tráfico en sus líneas de acceso al Punto Neutro, habrán de ampliar la capacidad de estas líneas en el plazo de 60 días. El estado de congestión grave será predeterminado por un procedimiento de auditoría aprobado por la Comisión Técnica

7.2 POLÍTICA MULTI- HOMED.

NIC tiene una política "multi-homed" la cual permite a organizaciones multi-homed que han utilizado eficientemente un bloque /21 le sea asignado un bloque /20. Una organización es muti-homed si recibe una conectividad de tiempo completo de más de un ISP y el cual tiene uno o más prefijos de ruteo anunciados al menos por dos de sus ISP. Este requisito es debido a restricciones técnicas y de implementación en el sistema de ruteo de Internet.

Con el fin de recibir una asignación de NIC las organizaciones Multi-homed deberán.

1. Demostrar la eficiente utilización de un mínimo (contiguos o no contiguos) /21 (ocho /24).

2. Proveer información de la utilización para un /29 y más corta y longitudes de prefijos usando SWIP para ser registrado en la BD WHOIS. Si en el futuro se requiere de espacio adicional la informacion de la BD de WHOIS deberá estar diponible al momento de la solicitud. Bloques mas pequeños que un /29 pueden ser documentados usando el siguiente formato.

Ciudad	Direcciones IP asignadas	Numero de puertos	Numero de clientes dial-up
Ciudad	Direciones IP asignadas	Numero de hosts internos	Propósito
Direcciones IP asignadas		Lista URLs para websites	

Tabla 7.1 Asigación de IPs mediante multi-homed

3. Proveer información detallada mostrando como el /20 será utilizado dentro de tres meses.

4. Deberán estar de acuerdo en renumerar el bloque /21 dentro de un plazo de 18 meses y regresar el espacio a su proveedor original. Este punto es indispensable para obtener el bloque /20 que se solicita. El bloque /20 asignado deberá ser usado para renumerar el bloque /21 asignado previamente

5. Planes de subneteo por al menos un año, incluyendo mascaras de subred y números de hosts sobre cada subred. El uso de VLSM es requerido (ver template).

6. Una descripción de la topología de la red.

7. Una descripción de los planes de ruteo de la red, incluyendo los protocolos de ruteo a ser usado también como cualquier limitación existente.

Requerimientos para el espacio de direcciones solicitado

La tasa de utilización es un factor clave en justificar. La tasa de utilización es el porcentaje de direcciones que la organización utilizará en un espacio de tiempo determinado. El establecido de acuerdo al RFC 2050 y adoptado por NIC es

25% de la tasa de utilización inmediata.

50% de la tasa de utilización en un año.

Una tasa de utilización más grande puede ser requerida basado en requerimientos individuales. Si la organización solicitante no cumple con esos parámetros se le retirarán las direcciones negociando un tiempo razonable para su renumeración

7.3 SLA.

EL modelo de Acuerdo de Nivel de Servicios (Service Level Agreement, SLA) consiste en un contrato en el que se estipulan los niveles de un servicio en función de

una serie de parámetros objetivos, establecidos de mutuo acuerdo entre ambas partes, así, refleja contractualmente el nivel operativo de funcionamiento, penalizaciones por caída de servicio, limitación de responsabilidad por no servicio, etc.

Este modelo no ha de estar relacionado necesariamente con la contratación de servicios a terceras partes, sino que puede implantarse a nivel interno, transformando una determinada unidad de negocio en centro de servicios que provea a la propia compañía.

En esta parte del contrato se describe y obliga a un nivel específico de calidad en el suministro.

Los principales puntos a cubrir deben ser:

- Tipo de servicio.
- Soporte a clientes y asistencia.
- Provisiones para seguridad y datos.
- Garantías del sistema y tiempos de respuesta.
- Disponibilidad del sistema.
- Conectividad.
- Multas por caída del sistema.

Estos puntos son importantísimos a la hora de formalizar de forma contractual una operación.

7.3.1 Implatanción de acuerdos de nivel de servicio con proveedores.

Para implantar con éxito un SLA han de tenerse en cuenta una serie de factores claves, de los que va a depender en gran medida la obtención de los resultados deseados:

- **Aspectos críticos.**

Los aspectos más críticos, son la definición de procedimientos estándares y los mecanismos de evaluación y seguimiento.

- **En la implatación de un SLA se deben seguir una serie de puntos:**

1. Definición de Objetivos: mejora de la eficacia, reducción de costes, formalización de la relación

2. Identificar expectativas: qué es lo que espera la organización de este acuerdo

3. Adecuada planificación temporal

4. Optimización/rediseño de procesos (revisar los procesos si el SLA no asegura ningún cambio o como mínimo formalizarlos).

- **Errores más frecuentes en la implantación.**

-Definir niveles de servicio inalcanzables

-Regulación excesiva

-Error en la definición de prioridades

-Complejidad técnica

¿Por qué necesito un SLA?

Un contrato que claramente esboza los derechos y las obligaciones de las partes reduce significativamente el ámbito de desacuerdos que se generan en el curso de las relaciones de negocios de éstas. Usted sabe en qué se está involucrando, cómo y cuándo.

7.3.2 Que tipo de información debe contener un SLA.

El SLA debe incluir medidas, y siempre que esté disponible y sea práctico, rangos del sector y criterios de cálculo. Para un BSP, las tres medidas principales de éxito ingresos (sostenidos y crecientes), rentabilidad y satisfacción de cliente son la razón para tener SLAs que cubren también la calidad.

El SLA debe incluir:

• Propósito del SLA	• Descripción del Servicio	• Duración del Servicio	• Asuntos Legales tales como Garantías
• Términos de Pago	• Condiciones de Terminación	• Cronograma de Instalación	• Fecha de iniciación de servicio, indemnizaciones, limitación de responsabilidad, etc.

Tabla 7.2 Términos incluidos en un contrato SLA.

¿Afectará el SLA la adición de nuevos usuarios o aplicaciones?

Ya que la duración de un SLA típico es uno a tres años, es poco probable que el ámbito de un servicio al cliente pudiese cambiar durante el período del contrato a través de la adición de nuevas aplicaciones, usuarios o ambos. Con el fin de evita malos entendidos, es importante que el SLA intente anticipar el impacto de servicio de tales cambios. Por ejemplo, el SLA puede enunciar que niveles especificados de servicio se emplean solamente para las aplicaciones y el número de usuarios delimitados en el SLA. Alternativamente, el SLA puede manifestar que el BSP garantizará el nivel de servicio descrito en el SLA aún si el cliente adiciona un

número especificado de nuevas aplicaciones o usuarios. Aún más, en otros casos, el SLA puede obligar al BSP y al cliente a discutir cualquier cambio a la configuración inicial y modificar el SLA, si fuese necesario. Cualquiera de estos planteamientos representa una solución efectiva, con tal que ambos, el BSP y cliente, entiendan claramente las condiciones bajo las cuales se aplica el SLA concurrente.

7.3.3 Ejemplo de los principales parámetros medibles

Los principales parámetros medibles y las características del servicio para el caso de Telefónica se describen a continuación:

Telefónica le garantiza un servicio de primera calidad gracias al riguroso control y gestión de Su red 24x7 desde su Centro de Gestión de Red Internacional, uno de los más modernos y seguros del mundo.

El servicio asegura la calidad que sus clientes demandan, y para garantizárselo ofrecen reembolsos según las condiciones especificadas en el SLA acordado para el caso de incumplimiento de los niveles comprometidos de disponibilidad, retardos y pérdida de paquetes. También ponen a disposición una URL pública en la que podrá comprobar las prestaciones de su Red en todo momento (URL: http://www.telefonica-wholesale.com).

A continuación se muestra los parámetros principales a tener en cuenta en un acuerdo SLA.

Niveles del servicio:
--Disponibilidad: hasta el 100% según SLA
-- MTTR < 4 horas
--Pérdidas de paquetes < 0.1%
--Retardos (Round Trip delay):
Intra Europa: < 50 ms
Europa-USA: < 80 ms
Latam-USA: < 120 ms

Tabla 7.3 Niveles de servicio en SLA.

Acceso	
Velocidad (1)	Hasta 2.5 Gbps
Tecnología de red	Cisco 12000 DWDM, Juniper
Peering privado extenso (Tier 1)	Sí
Autonomous System (2)	AS12956
Características del Servicio	
Overbooking	No
Disponibilidad garantizada, con penalizaciones por SLA (3)	Hasta 100%
Round trip delay garantizado	
Intra Europa	50 ms
Europa-USA	80 ms
Latam-USA	120 ms
Packet loss	<0.1%
Tarificación	Tarifa plana o por uso
Capacidad mínima comprometida	Sí
Notificación a clientes en caso de avería	<15 min
Opciones de servicio	
Routing multicast	Sí
Routing BGP4	Sí
Reserva direcciones IP	IPv4 e IPv6
DNS secundario	Sí
Gestión de servicio	
Help Desk	24x7
Estadísticas e Informes de Servicio On-Line	Tiempo Real
SLA	1 mes
Periodo de contrato	Mínimo 1 año

(1) Velocidades de acceso mayor disponibles bajo pedido
(2) Información de full routing accesible vía www.ripe.net
(3) Depende de la ciudad a la que se conecta el cliente

Tabla 7.4 Ejemplo de parámetros definidos por Telefónica.

7.4 NAP (Network Access Point).

Un NAP es definido como una red de alta velocidad a la cual se pueden conectar varios proveedores de acceso a Internet con el propósito de intercambiar tráfico, permitiendo por ejemplo la interconexión de redes regionales. Opera por lo general a velocidades iguales o superiores a 100Mbps, en algunos casos conformadas por switches FDDI o ATM (155Mbps), pasando tráfico de un proveedor a otro. Los proveedores de Internet que se conectan al NAP deben respetar las políticas definidas por los otros proveedores conectados a este. Otro nombre con el que se conocen los NAP es IXP, Internet Exchange Point.

7.4.1 Quien Administra los NAP y cuales son sus funciones.

La organización que administra cada NAP es usualmente elegida por los proveedores de servicio que se conectan a este. Las funciones de este Administrador del NAP son:

- Establecer y mantener el NAP, conectándolo a las redes que corresponda.
- Establecer políticas y cuotas a los proveedores que quieren conectarse al NAP.
- Proponer la ubicación del NAP, de acuerdo a las condiciones geográficas de la región.
- Proponer y establecer procedimientos para trabajar con el personal de otros NAPs, el Route Arbiter, y los proveedores de servicio que se conectan al NAP, con el fin de resolver problemas, soportar la conectividad de extremo a extremo y calidad de servicio para los usuarios de la red.
- Desarrollar estándares de confiabilidad y seguridad para el NAP, así como procedimientos para asegurar que esos estándares se cumplan.
- Especificar y proveer estadísticas y contabilidad del NAP.

7.4.2 Configuraciones físicas actuales de un NAP

La configuración física de, un NAP de hoy es una mezcla de switches FDDI, ATM y Ethernet (Ethernet, Fast Ethernet y Gigabit Ethernet). Los métodos de acceso varían desde FDDI y Gigabit Ethernet hasta DS3, OC3 y ATM OC12. La Figura muestra una posible configuración basada, en algunos NAP contemporáneos. Normalmente, el proveedor de servicios administra los routers colocados en los servicios del NAP mientras que el administrador define las configuraciones, las políticas y las tarifas.

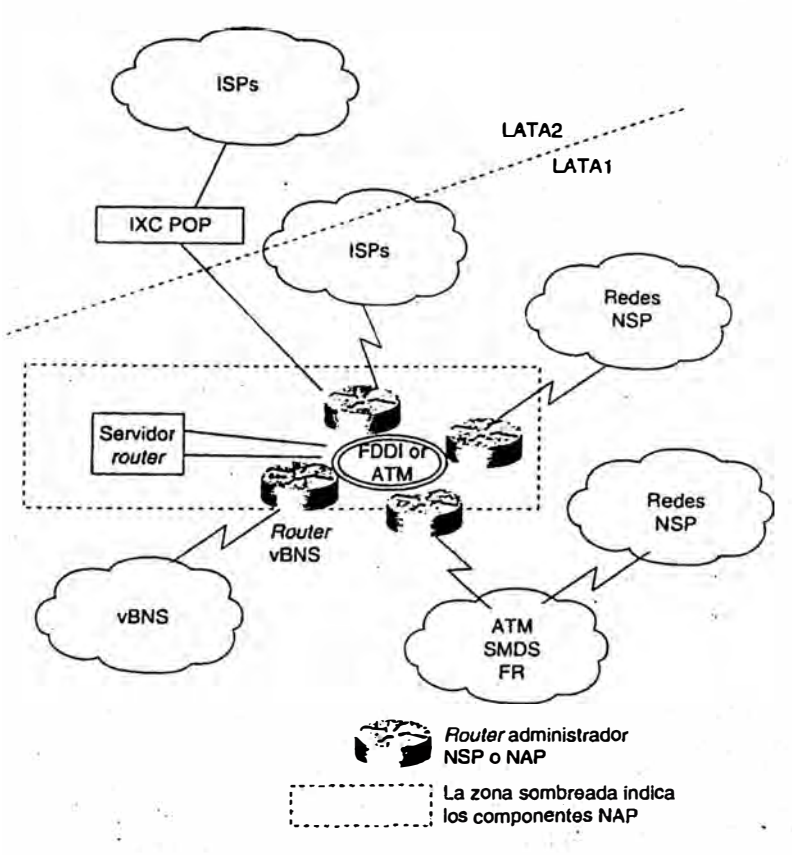


Figura 7.1 Componentes de un NAP.

7.4.3 Una alternativa a los NAP : Interconexíones directas.

Mientras Internet continúa creciendo, la enorme cantidad de tráfico intercambiado entre grandes redes está haciendo que muchos NAP no puedan soportarlo. A menudo, los problemas de capacidad de los NAP se traducen en pérdidas de datos e inestabilidad. Además, grandes redes privadas, y a veces los ISP, son reacios a confiar en administradores de NAP de terceros, aparentemente menos interesados en resolver los problemas que afecten al servicio y en proporcionar capacidad adicional. Por dichas razones, durante los últimos años ha evolucionado una alternativa a los NAP para la interconexión de proveedores de servicios: las interconexiones directas.

La idea que se esconde tras ellas es simple. Mediante el suministro de enlaces directos entre redes y evitando totalmente los NAP, los proveedores de servicios pueden reducir los tiempos de aprovisionamiento, incrementar la fiabilidad y escalar considerablemente la capacidad de interconexión. El ancho de banda del enlace y la situación de las interconexiones directas, se negocian normalmente de forma bilateral sobre una base de igual a igual. Las interconexiones directas no son perseguidas normalmente entre dos redes hasta que una o ambas partes implicadas se dan cuenta de los incentivos económicos asociados con el hecho de evitar los NAP.

Las interconexiones directas no sólo proporcionan ancho de banda adicional entre las redes interconectadas, sino que también alivian la congestión y liberan ancho de banda en los NAP, mejorando consecuentemente el rendimiento de transferencia y la ejecución. Asimismo, puesto que los controladores del mercado normalmente dan como resultado grandes topologías de red que se reflejan fielmente unas a otras, el parecido entre las topologías de red y los requisitos de interconexión permite proporcionar a las interconexiones directas una mejor distribución geográfica para el

intercambio de datos que los NAP. Las interconexiones directas pueden proporcionar una arquitectura que regionalizará óptimamente el intercambio de tráfico entre redes, incrementando de ese modo el rendimiento de transferencia en la red mientras decrece la latencia entre un conjunto dado de hosts.

Los proveedores regionales más pequeños y los proveedores de servicios más nuevos, probablemente no estarán en posición de llegar a acuerdos de interconexión con grandes proveedores por un par de razones:

- Los costes asociados a proveedores ya existentes que mantienen grandes cantidades de infraestructura con el fin de acomodar interconexiones directas.
- El incremento de tarifas asociado con el número de facilidades de circuito requerido por los LEC (Local Exchange Carriers) y los IXC (IntereXchange Carriers).

Afortunadamente, la mayoría de los grandes proveedores siguen manteniendo una presencia en los NAP, utilizando sus conexiones para intercambiar tráfico con redes que no pueden justificar todavía los costes de interconectarse directamente.

7.4.4 El NAP Peru.

Un logro para el desarrollo de Internet en el Perú se plasmó el viernes 25 de agosto con la firma del acuerdo para la creación del NAP Perú. Las cinco empresas de telecomunicaciones más importantes del país, BellSouth, COMSAT Perú, FirstCom, Infoductos y Telecomunicaciones del Perú (RCP) y Telefónica del Perú SAA, hacen posible la realización de este proyecto largamente ambicionado, no sólo por los proveedores de servicios Internet, sino también por la mayoría de empresas y usuarios que conocen su importancia.

El NAP (Network Access Point) es un punto de acceso y conexión de redes que permitirá a las distintas empresas locales proveedoras de servicios Internet intercambiar el tráfico local sin necesidad de utilizar los enlaces internacionales. Hasta ahora, antes de la creación del NAP, un correo electrónico de un cliente del proveedor A, destinado a un cliente del proveedor B, ambos proveedores en el Perú, era enviado a algún punto de conexión de redes, ubicado usualmente en los Estados Unidos. Así, por ejemplo, un correo entre Miraflores y La Victoria podía recorrer un camino vía Nueva York. Esto implicaba el uso innecesario de enlaces internacionales, además de causar demoras en el servicio en desmedro de su calidad.

A partir de la puesta en marcha del NAP peruano, el tráfico local de Internet entre diferentes empresas proveedoras no pasará por el enlace internacional y será manejado localmente, lo que acelerará las comunicaciones internas y el acceso a los recursos de los servidores locales (páginas Web, bases de datos, etc.).

Los usuarios peruanos, serán los principales beneficiados con la creación del NAP Perú, ya que permitirá una importante mejora y mayor rapidez en el servicio.

El Perú, tras la firma de este importante acuerdo, se convierte en el quinto país de América Latina en contar con este sistema, además de Colombia, Argentina, Chile y Brasil.

El NAP no afectará la libre competencia entre las empresas comprometidas, por lo cual será administrado por un ente totalmente neutral sin fines de lucro, cuyo único objetivo será impulsar el desarrollo del sector telecomunicaciones.

Actualmente existen algunos problemas entre los proveedores integrantes del NAP, y esto radica en que Telefonica tiene el 70% del total de los contenido en paginas a

nivel nacional, en donde todo los demás proveedores quieren conectarse a la información que tiene Telefonica. Entonces aquí empieza el detalle, mientras que telefónica viene gastando por ejemplo 100 en equipos para el NAP, los demás integrantes del NAP vienen gastando solamente 20 y tienen capacidad instalada ociosa por montones. Esto conlleva a que todos los integrantes del NAP le pidan solamente y exclusivamente a Telefonica que amplíe capacidades, a lo que Telefonica replica que no es equitativo de que se tenga que invertir en esas proporciones y los demás inviertan poco.

Actualmente la conexión entre Telefonica y el NAP es de 12 Mbp.

Actualmente las empresas que conforman el NAP Peru son: AT&T, RCP, BellSouth, Diveo, IMSAT, COMSAT y Telefonica, entre las cuales e maneja sesiones BGP en Full Mesh (todos contra todos) y se usan equipamientos donados por Cisco System, y se cuenta con un ancho de banda simétrico que suma 42 Mbps (42 Mbps de entrada y 42Mbps de salida).

7.5 PROYECTO ROUTING ARBITER.

Este proyecto fue creado, con el fin de ofrecer un tratamiento equitativo a los diversos proveedores de servicios de red en cuanto a administración del enrutamiento. El RA estipula una base de datos común de información sobre enrutamiento para promover la estabilidad y la administrabilidad de las redes..

El hecho de que múltiples proveedores se conectaran a un NAP, creó un problema de escalabilidad porque cada proveedor tenía que conectarse por igual con todos los demás para **intercambiar** información de enrutamiento y política. El proyecto RA fue desarrollado para reducir los requisitos de una malla de iguales entre todos los

proveedores. En lugar de conectarse entre sí mediante iguales, los proveedores pueden conectarse con un sistema central llamado servidor de ruta. El servidor de ruta mantendrá una base de datos con toda la información necesaria para que los proveedores establezcan sus políticas de enrutamiento. La Figura muestra la conectividad física y el peering lógico entre un servidor de ruta y varios proveedores de servicios.

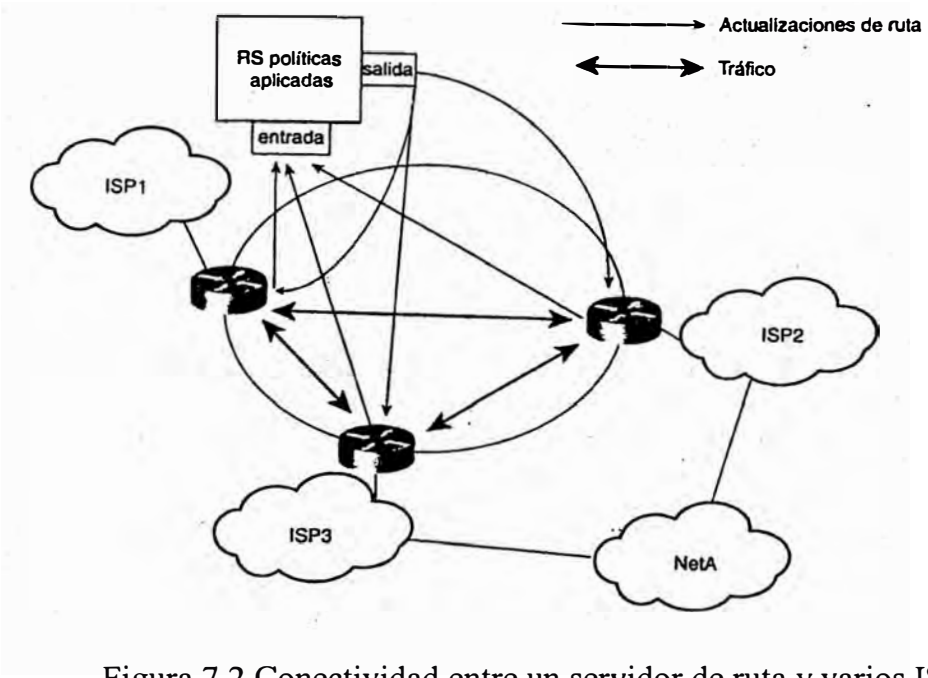


Figura 7.2 Conectividad entre un servidor de ruta y varios ISP.

Las siguientes son las principales tareas del RA:

- Promocionar la estabilidad y administrabilídad del enrutamiento en Internet. El servidor de ruta desempeña gran par . te de esta tarea reduciendo el número de iguales BGP que un router está obligado a mantener y aplicando la política antes de pasar información de enrutamiento hacia el igual, aliviando

así los recursos de los procesos requeridos por el router para filtrar la información de enrutamiento.

- Establecer y mantener bases de datos de topologías de red por métodos como intercambiar información de enrutamiento con los sistemas autónomos adjuntos (AS), y actualizar dinámicamente información de enrutamiento desde estos mismos sistemas, utilizando Protocolos de gateway_ exterior (EGP) estándar, como, por ejemplo, BGP.
- Proponer y establecer procedimientos para trabajar con personal de administración del NAP, el proveedor y redes adjuntas regionales y de otros tipos, para resolver problemas y para soportar conectividad de extremo a extremo y QoS para usuarios de la red.
- Desarrollar tecnologías de enrutamiento avanzado como el tipo de servicio y prece**dencia de enrutamiento, multidifusión,** ancho **de banda sobre petición** y servicios de asignación de ancho de banda en cooperación con la comunidad global de Internet.
- Mantener estrategias de enrutamiento simplificadas, como el enrutamiento por defecto para redes conectadas.
- Promocionar el funcionamiento y la administración distribuida de Internet.

El servicio RA constaba de dos proyectos:

7.5.1 El servidor de ruta (RS).

El RS puede ser tan simple como una estación de trabajo Sun desplegada en cada NAP. El servidor de ruta sólo intercambia información de enrutamiento con los

routers del proveedor de servicios conectado al NAP. Los requisitos de la política de enrutarmiento individual (RIPE 181) para cada proveedor se mantienen. El servidor de ruta, por sí mismo, no reenvía paquetes ni ejecuta ninguna función de switching entre proveedores de servicios.

El servidor facilita la interconexión entre los ISP recopilando las reglas y políticas de enrutamiento predefinidas de cada ISP y redistribuyendo dicha información de enrutamiento a cada ISP. Este proceso salva a cada router de tener que conectarse por igual con cada uno de los otros ruiters, reduciendo así el número de iguales de (n - 1) a 1, donde n es el número de routers.

En esta configuración, los routers de los distintos proveedores se concentran en intercambiar el tráfico entre ellos y en hacer un filtrado y una aplicación de la política relativamente pequeños.

7.5.2 Base de datos Routing Arbiter (RADB).

Esta es una de las diversas bases de datos de enrutamiento conocida colectivamente como Registro de enrutamiento en Internet (IRR). La política de enrutamiento en la RADB se expresa utilizando la sintaxis RIPE-181, desarrollada por el Centro de coordinación de red RIPE (RCC). La RADB fue desarrollada en modo dual con la Base de datos de políticas de enrutamiento.

7.6 Descripción general del servicio Infovia Plus brindado por TdP.

InfoVía Plus Básico es un servicio que permite a sus clientes (proveedores, empresas o instituciones) ofrecer a los usuarios finales acceso a través de conexiones conmutadas a la Red IP. A diferencia de otros servicios, no es necesario que el

proveedor disponga de una conexión por la que canalice el tráfico de sus usuarios.

Las principales características del servicio son las siguientes:

- Acceso al servicio desde cualquier punto con cobertura de la Red IP.
- Proporciona acceso conmutado bidireccional a la Red IP a usuarios finales, vía RTC o RDSI y protocolo PPP.
- El acceso es abierto, es decir, permite la comunicación con otros usuarios de la Red IP y el acceso a los servicios de valor añadido que se prestan desde ella (en el caso de que se hubiera contratado alguno).
- El servicio en esta versión se presta en dos modalidades:

  Modalidad de autenticación en red, en la que el proveedor proporciona a la Red IP los datos necesarios para que ésta, en su nombre, permita o deniegue los accesos de los usuarios de ese proveedor.

  Modalidad de autenticación delegada, en la que el proveedor mantiene su propia base de datos de usuarios, y a través de un servidor de autenticación conectado a la Red IP a través de un router, permite o deniega los accesos de sus usuarios.

Los usuarios accederán al servicio utilizando equipos (generalmente ordenadores personales) con las interfaces de acceso adecuadas y que soporten los protocolos PPP e IP. Estas interfaces (módems, adaptadores, tarjetas) no forman parte del servicio, y son responsabilidad del usuario.

El usuario para conectarse de forma conmutada a su proveedor deberá marcar un número de teléfono (140100).

Además del acceso directo de usuarios individuales, también es posible el acceso de

una red de área local a través de routers que dispongan de las interfaces de acceso adecuadas. Puesto que el servicio se limita a ofrecer la capacidad de acceso y se asigna al usuario una sola dirección IP, será necesario que el cliente por sus propios medios (y bajo su responsabilidad) habilite algún elemento que haga funciones de proxy o traductor de direcciones (NAT), si se desea el acceso simultáneo de varios usuarios.

Será posible realizar la agregación de varios canales de acceso

utilizando el protocolo MLPPP (MultiLink PPP).

En el servicio se consideran tres calidades de servicio: Oro, Plata y Bronce. Estas calidades de servicio se caracterizan por la diferente prioridad con que la información es encaminada dentro de la Red IP. La información de los usuarios que acceden con calidad Oro tiene mayor prioridad en su transporte por la Red IP que la de los usuarios que acceden con calidad Plata que a su vez tiene mayor prioridad que la de los usuarios que acceden con calidad Bronce.

Cada usuario de un proveedor tiene asignado un identificador (de la forma nombre@proveeedor) y una clave secreta, que se utiliza para autorizar y autenticar el acceso a la red.

Una vez conectados, los usuarios disfrutarán de un acceso abierto a Internet, pudiendo navegar libremente por ella y acceder a los servicios que allí se prestan.

El proveedor dispondrá de un límite al número de conexiones simultáneas de usuarios suyos que pueden establecerse. La Red IP impedirá la conexión de más usuarios cuando se alcance este tope.

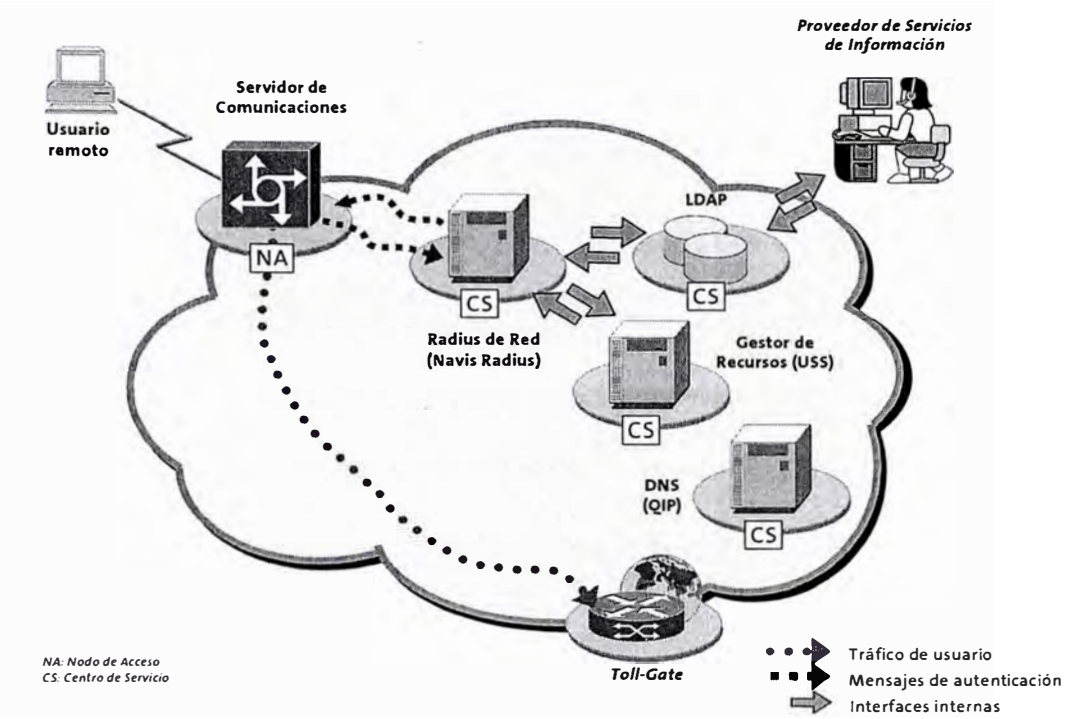


Figura 7.3 Modalidad de autenticación en red.

7.6.1 Software de usuario.

El acceso a la Red IP puede realizarse desde cualquier equipo que cumpla los requisitos exigidos, que son esencialmente la capacidad de conexión a la red de acceso, y el soporte de IP sobre PPP. Prácticamente todos los routers y equipos especializados en interconexión de redes disponen de estas capacidades.

La mayoría de los usuarios accederán desde sus ordenadores personales. Los sistemas operativos más comunes, como Windows 9x/NT o MacOS incluyen de serie el soporte de hardware de conexión a redes (módems, tarjetas de comunicaciones) y los protocolos de comunicaciones adecuados.

7.6.2 Proveedores de la modalidad de autenticación en red.

Los proveedores son los clientes de Telefónica que contratan el servicio para ofrecérselo a sus usuarios. Transfieren a la Red IP los datos necesarios para que sea la propia red la que realice la totalidad de las funciones de autorización en su nombre, siguiendo sus directrices. El mantenimiento de estos datos se realiza a través del procedimiento de altas, bajas y modificaciones masivas, descrito más adelante.

Los proveedores no necesitan contratar una conexión permanente con la Red IP hacia la que se canalice el tráfico de sus usuarios. Tampoco es necesario que dispongan de bloques propios de direcciones IP.

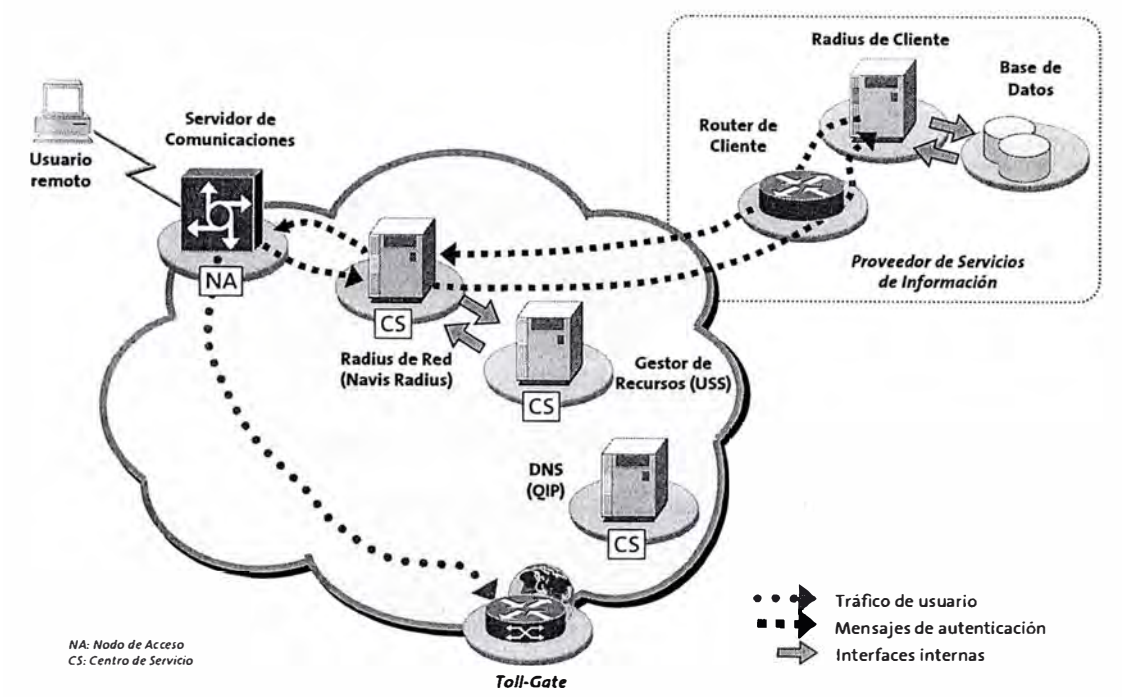
El mnemónico es indicado en el identificador de usuario cuando éste intenta conectarse (identificador: usuario@realm). Cuando la Red IP atiende una petición de conexión de un usuario determina a qué proveedor desea conectarse a partir del realm.

7.6.3 Proveedores de la modalidad de autenticación delegada.

Los proveedores son los clientes de Telefónica Data que contratan el servicio para ofrecérselo a sus usuarios. Los proveedores de la modalidad delegada realizan ellos mismos las funciones de autenticación y autorización de los accesos y gestión de la base de datos de usuarios, delegando en la Red IP otros aspectos del servicio, como la asignación de direcciones IP. Los proveedores son los responsables de permitir o denegar el acceso a la Red IP de los usuarios.

Para realizar la autorización de sus usuarios, estos proveedores deberán disponer de un servidor Radius accesible desde la Red IP a través de una conexión permanente, como se describe más adelante.

Los proveedores no necesitan canalizar a través de esta conexión permanente con la Red IP el tráfico de sus usuarios. Tampoco es necesario que dispongan de bloques propios de direcciones IP.



Fgura 7.4 Modalidad de autenticación delegada.

El servidor Radius de la Red IP es el responsable de la autorización de los accesos de los usuarios. En el caso de la autenticación delegada, es responsable de contactar con el servidor Radius del cliente y hacerle llegar la información necesaria para que pueda realizar la autorización. Para ello debe comportarse como proxy Radius. El servidor Radius de la Red IP realiza la asignación del número pool de direcciones IP a los usuarios.

7.7 Descripción del servicio Infovia Plus Directo brindado por TdP.

El servicio InfoVía Plus Directo es un servicio de acceso conmutado a redes privadas. El servicio permite el acceso de usuarios remotos a las redes privadas del cliente a través de túneles de nivel 2 establecidos entre un servidor de comunicaciones del nodo de acceso y un servidor de túneles (ST) en las dependencias del cliente. La utilización de estos túneles permite ofrecer sobre una red pública un servicio de acceso a redes privadas.

Por tanto, como principales características del servicio, cabe destacar:

- Acceso conmutado seguro a redes privadas virtuales a través de túneles dinámicos L2TP.
- Las funciones de autenticación de usuarios y asignación de direcciones IP las realiza el cliente.
- Direccionamiento independiente de red. La utilización de túneles permite al cliente el uso de cualquier tipo de direccionamiento, público o privado.

Desde el punto de vista de implementación técnica, el servicio InfoVía Plus Directo se apoya en el servicio Uno IP Básico como servicio básico de conectividad a la red. Es decir, el cliente debe al menos contratar un servicio Uno IP Básico para que sus usuarios remotos lleguen a su red a través de éste, así como para realizar la autenticación de los accesos.

- Servidores de túneles

  Los servidores de túneles (ST) son unos equipos situados en las dependencias del cliente que se encargan de terminar los túneles establecidos desde los servidores de terminales de la Red IP, y extraer

las comunicaciones de los usuarios.

- Servidor Radius de proveedor

  El proveedor tendrá en sus dependencias un servidor Radius que autorizará o denegará los accesos de sus usuarios, utilizando la información contenida en la base de datos de usuarios. También es el encargado de la asignación de direcciones IP.

- Base de datos de usuarios del proveedor

  El proveedor mantendrá una base de datos asociada al servidor Radius, en la que introducirá la información necesaria para la autenticación de sus usuarios. Esta base de datos puede ser desde un fichero plano hasta un sistema separado de gestión de bases de datos.

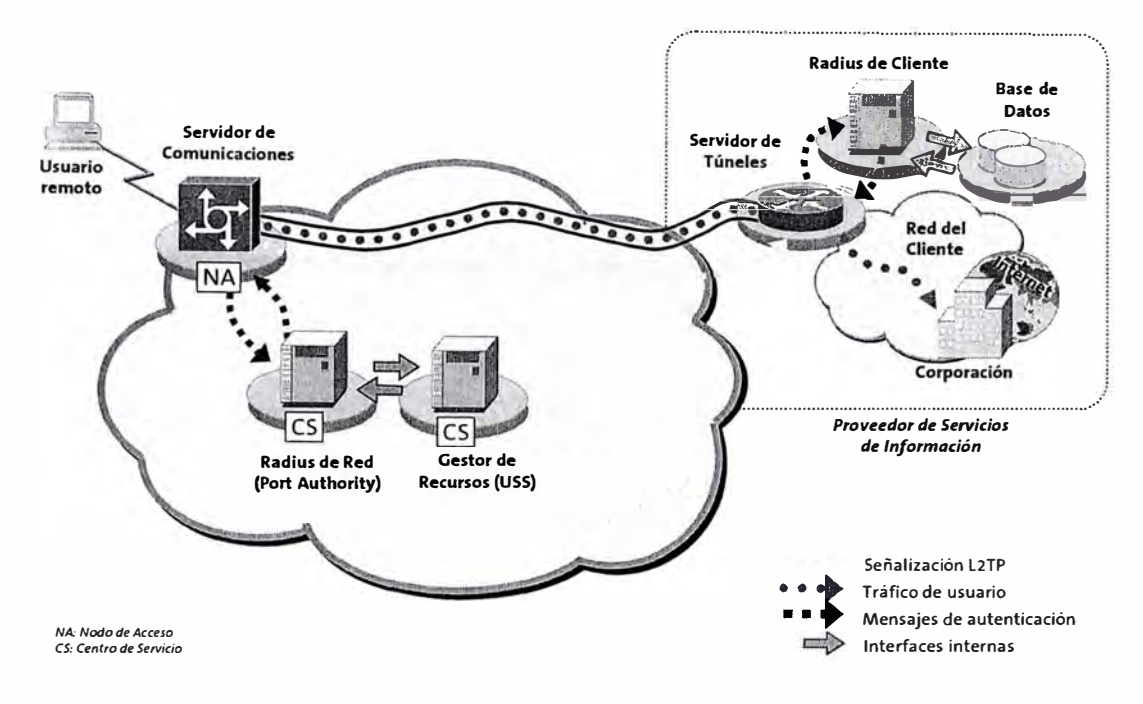


Figura 7.5 Servicio Infovia Plus Directo brindado por TdP

7.7.1 Usuarios del servicio.

El usuario para conectarse de forma conmutada a su proveedor deberá marcar el número de teléfono (140100). Además del acceso directo de usuarios individuales, también es posible el acceso de una red de área local a través de routers que dispongan de las interfaces de acceso adecuadas. Puesto que el servicio se limita a ofrecer la capacidad de acceso, será necesario que el cliente por sus propios medios (y bajo su responsabilidad) habilite algún elemento que haga funciones de router (y quizás de proxy o NAT), si se desea el acceso simultáneo de varios usuarios.

Cada usuario de un proveedor tiene asignado un identificador (de la forma nombre@proveeedor) y una clave secreta, el proveedor dispondrá de un límite al número de conexiones simultáneas de usuarios suyos que pueden establecerse. La

Red IP impedirá la conexión de más usuarios cuando se alcance este tope.

El mecanismo de túneles que se utiliza en el servicio requiere una negociación PPP inicial del equipo del usuario con el servidor de comunicaciones, en la que el acceso es autorizado por el servidor Radius de la Red IP. Si esta fase concluye con éxito, se produce un reinicio del PPP y una renegociación con el servidor de túneles, en la que el acceso es autorizado por el servidor Radius del proveedor. Por esta razón es necesario que el equipamiento del cliente soporte correctamente esta doble negociación.

7.7.2 Conexión del proveedor a la Red IP.

El proveedor estará conectado a la Red IP a través de una conexión que permita llegar desde cualquiera de los nodos de acceso hasta su servidor de túneles. Puesto que el protocolo de túneles empleado (L2TP) usa UDP/IP como transporte, el único requisito que necesita esta conexión es que ofrezca conectividad IP con la Red.

7.8 Descripcion del enrutamiento de los servicios, brindados por TdP.

Los diferentes servicios de acceso a Internet brindados por Telefonica del Peru hacen uso de una plataforma común para el acceso a Internet, cuenta para ello con routers los cuales ejecutan diferentes protocolos de enrutamiento (Rutas estáticas ,OSPF , BGP), teniendo para esto diferentes modelos y marcas de enrutadores (CISCO , LUCENT, UNISPHERE) que interactúan entre si, en forma sincronizada para dar acceso a Internet a los miles de usuarios de la Red IP y de la red ADSL.

7.8.1 Red IP (Lucent).

En el caso de la **Red IP**, los diferentes servicios que brinda esta red (anteriormente descritos) redistribuyen sus redes publicas y privadas por medio de el protocolo dinámico OSPF que ejecutan los routers (BSTDX) siendo este un IGP estable con la característica de tener una rápida convergencia.

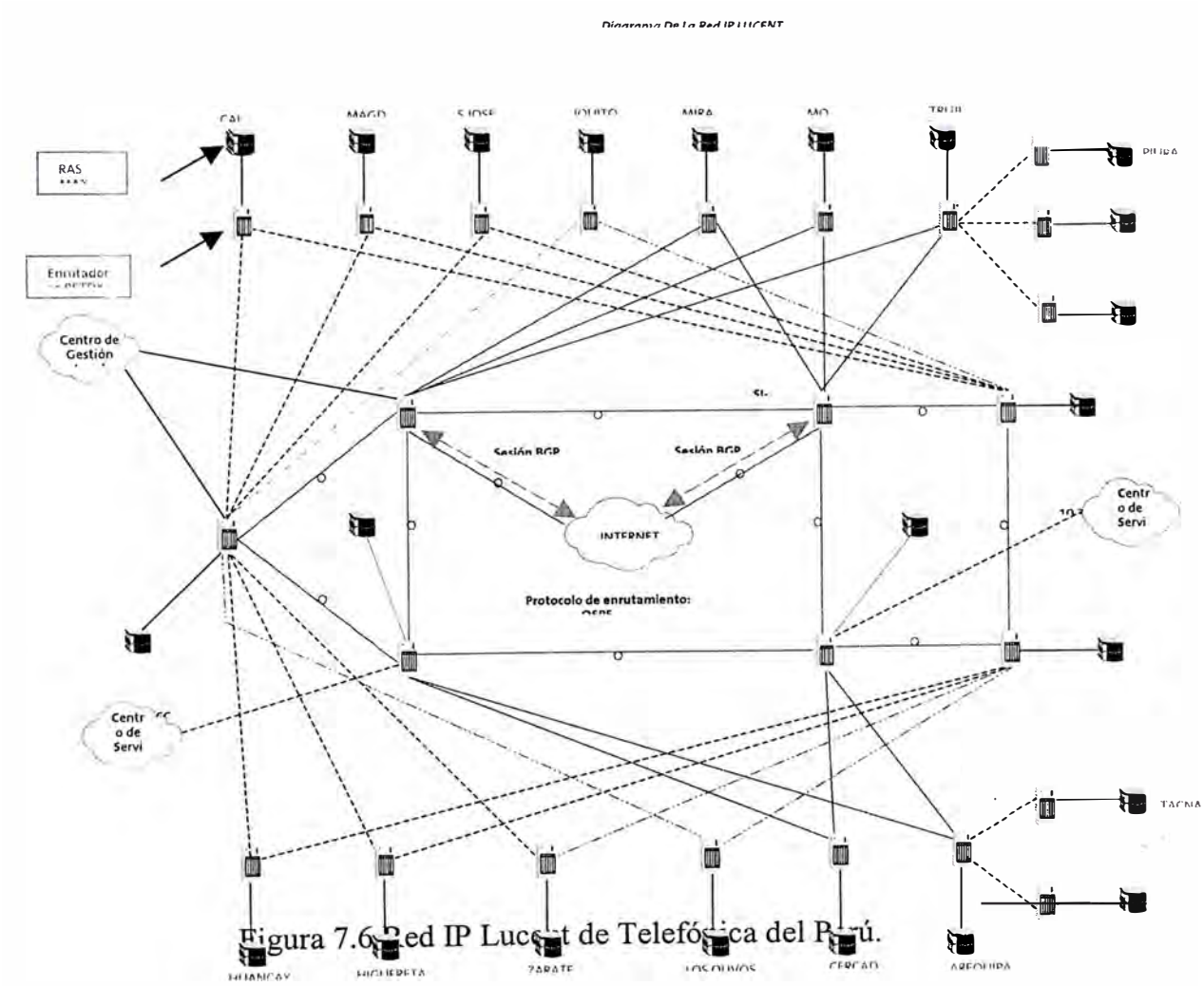


Figura 7.6 Red IP Lucent de Telefónica del Perú.

En el gráfico se observa los diferentes puntos de presencia de la **Red IP**, los cuales están formados por un servidor RAS MAX-TNT y un enrutador BSTDX, estos enrutadores ejecutan el protocolo de enrutamiento OSPF , de tal manera que las rutas

estáticas configuradas en los BSTDX para alcanzar a los usuarios dedicados o a los usuarios conmutados cuya sesión PPP termina en los MAX-TNTs serán redistribuidas en OSPF , y por lo tanto alcanzable desde cualquier punto de la red.
El acceso a Internet se realiza a través de los BSTDX1 ubicados en los locales de Washington y San Isidro, estos ejecutan el protocolo de enrutamiento BGP, en la Red IP se ha configurado un Sistema Autónomo privado que es el 64514 y se ha establecido una sesión EBGP con los enrutadores ERX de Unisphere ubicados también los locales de Washington y San Isidro los cuales tienen un Sistema Autónomo privado que es el 64513 como se observa en el primer grafico.

Por medio de esta sesión BGP se anuncian las direcciones de la Red IP hacia Internet, esto mediante una redistribución de las rutas aprendidas por OSPF hacia BGP de tal manera que ayudándonos de un mapa de ruta se anuncien solo las direcciones publicas y filtradas las direcciones privadas.

7.8.2 Red ADSL.

En el caso de la **red ADSL** Los usuarios acceden por medio de los DSLAMS y la red ATM formado por los BPX, para ello cada usuario tiene configurado un PVC a travez del cual levanta una sesión PPP hasta el agregador ERX de Unisphere como se observa en el grafico siguiente.

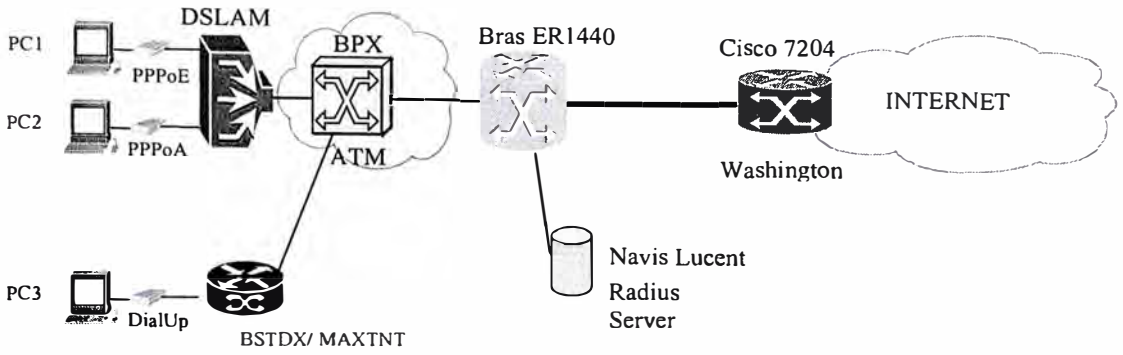


Figura 7.7 Red ADSL de Telefónica del Perú.

Actualmente la red ADSL consta de seis nodos de agregación, los cuales ejecutan el protocolo OSPF como se observa en la grafica.

Los agregadores estan configurados de tal manera que por cada sesión PPP levantada por el usuario se inscribe una ruta tipo access-internal y esta es redistribuida en el protocolo OSPF de tal manera que todos los elementos de la red conozcan como alcanzar a los diferentes usuarios.

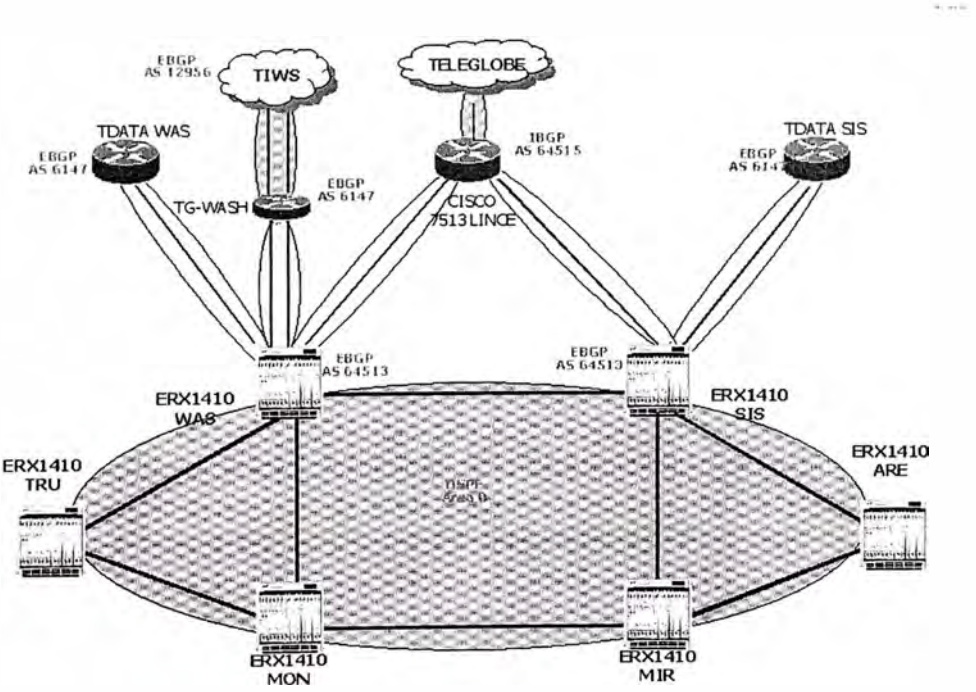


Figura 7.8 Topología del enrutamieto de Telefónica.

Como se observa en la gráfica anterior los ERX usados para agregación están ejecutando el protocolo OSPF de tal manera que tanto las rutas estáticas como las rutas originadas por el usuario son redistribuidas, y dentro de estas rutas se encuentran direcciones privadas y publicas.

Para anunciar nuestras redes a Internet se han establecido varias sesiones BGP básicamente por el tema de la redundancia, esto solo lo realizan los agregadores ubicados en los locales de Washington y San Isidro , para este fin se cuenta con un Sistema Autónomo privado el cual es el 64514.

Por medio de esta sesión BGP son anunciadas los pooles de la red ADSL mediante una redistribución de OSPF a BGP, para la cual se cuenta con filtros, mapas de ruta, que nos permiten evitar que se redistribuyan redes privadas.

Además para el anuncio de las redes, se ha considerado el tema de la agregación.

7.8.3 Enrutamiento de Tráfico Nacional e Internacional.

Para el tema de la salida a Internet se tienen enlaces Internacionales con nuestros Carriers o proveedores de acceso a Internet, a los cuales les anunciamos en forma sumarizada todas las pooles o redes utilizadas por los usuarios, esto debido a que en Internet se enruta como minimo redes de tipo claces C, teniendo mucho cuidado en que las redes no sean inestables o sea que estén por algún motivo apareciendo y desapareciendo de las tablas de ruta ya que ello ocasionaria lo que se conoce como flaping y esto traeria como consecuencia que nuestras redes sean penalizadas por los diferentes carriers de Internet.

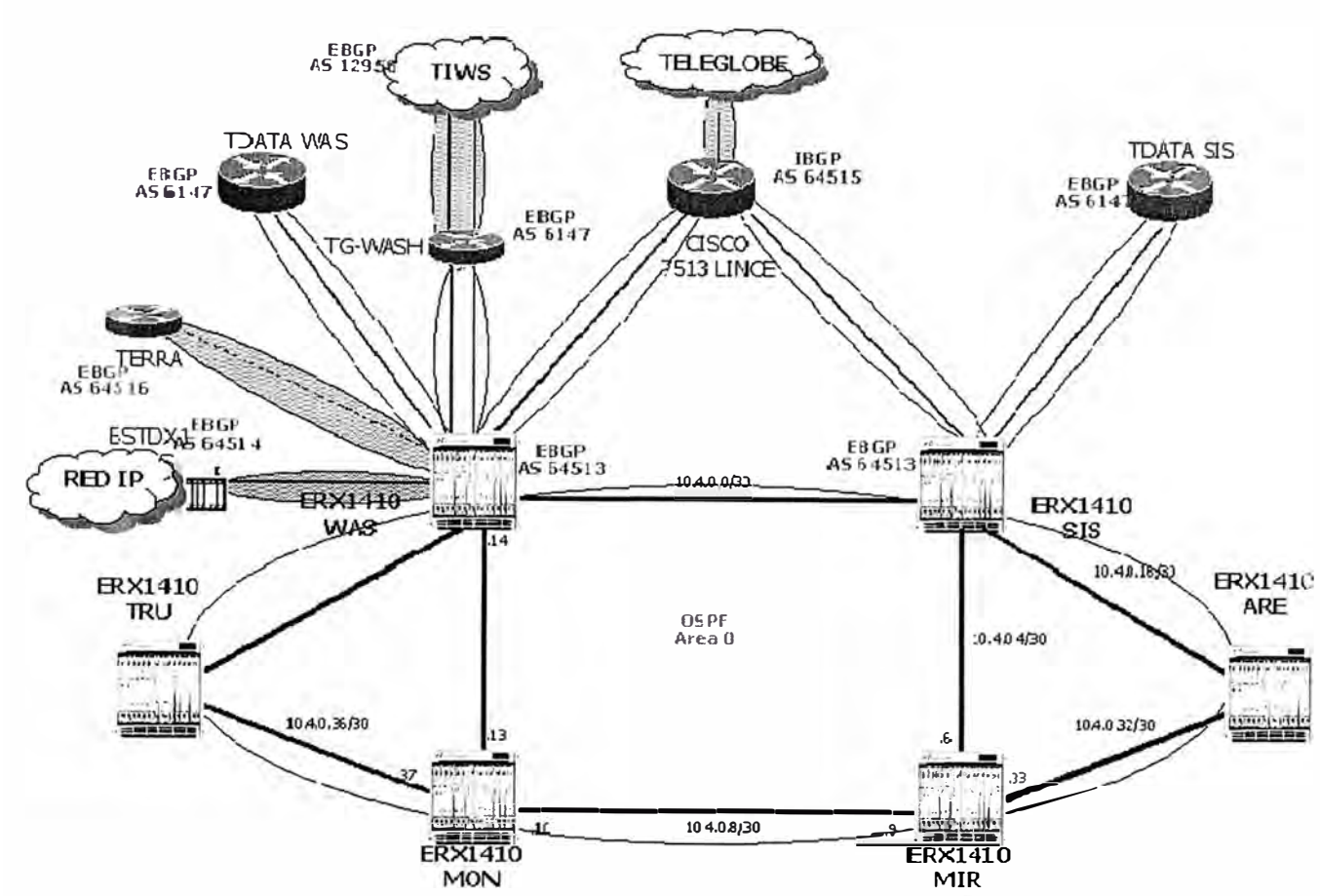


Figura 7.9 Enrutamiento de Tráfico Nacional e Internacional

Para el caso de nuestro tráfico de salida, esto lo realizamos por medio de una ruta por default a pesar de que estamos recibiendo de nuestros proveedores todo el full-ruting de Internet (aproximadamente 100,000 rutas) , esto lo realizamos por medio de listas de acceso aplicado a la sesión BGP en el sentido entrante a la RED.

Se tienen sesiones BGP establecidas con los equipos de la red IP/MPLS de TData esto debido a que TData tiene conexión directa al NAP del Perú , de tal manera por intermedio de dicha sesión recibimos todas las redes Nacionales esto para cuando nuestros usuarios tanto de la red IP como de la Red ADSL deseen acceder a paginas WEB nacionales sea por intermedio de estos enlaces y no haga uso de los enlaces

Internacionales, evitando saltos no necesarios por routers de Internet para que finalmente retorne nuevamente al Perú, ocasionando demasiado retardo y y consumo de ancho de banda nuestros enlaces Internacionales los cuales son bastante cotosos.

Esto se aplica también al trafico entrante ya que por intermedio de la sesión BGP establecida con TData nuestros pooles son anunciados hacia el NAP del Perú de tal manera que todos los proveedores nacionales de acceso a Internet (ATT , Milicom,...) saben que para alcanzar a cualquiera de nuestros usuarios deben hacer uso de estos enlaces

CAPÍTULO VIII

IMPLEMENTACIÓN DE UN ROUTER EN LINUX DESARROLLADO POR EL GRADUANDO.

8.1 Implementación de un router y servidor de rutas ZEBRA en Linux.

Zebra es un paquete de software de encaminamiento que proporciona encaminamiento basado en servicios de TCP/IP con protocolos de encaminamiento que soportan RIPv1, RIPv2, RIPng, OSPFv2, OSPFv3, BGP-4 y BGP-4+. Zebra también soporta el comportamiento especial de BGP Route Reflector y Route Server. Además de los protocolos de encaminamiento tradicionales basados en IPv4, Zebra también soporta protocolos de encaminamiento basados en IPv6

Zebra utiliza una arquitectura de software avanzada para proporcionar una gran calidad, con un motor multi servidor de encaminamiento. Zebra tiene un interfaz de usuario interactivo para cada protocolo de routing y soporta comandos de cliente en sus interfaces. Debido a su diseño es posible añadir nuevos demonios de protocolos

fácilmente a Zebra. Zebra se puede también utilizar como librería para un programa cliente de interfaz de usuario.

¿Qué es Zebra?

Un sistema con Zebra instalado actúa como router dedicado. Con Zebra, una máquina intercambia información de routing con otros routers utilizando protocolos de routing. Zebra utiliza esa información para actualizar el núcleo de las tablas de routing de forma que la información correcta esté en el lugar correcto. Zebra permite la configuración dinámica y es posible ver la información de la tabla de routing desde el interfaz de terminal de Zebra.

Añadiendo soporte al protocolo de routing, Zebra puede configurar las banderas (flags) de las interfaces, direcciones de los interfaces, rutas estáticas y muchas más cosas. Si se utiliza en una red pequeña o en una conexión xDSL, la configuración del software Zebra es muy sencilla. Lo único que hay que pensar es en levantar los interfaces e introducir unos pocos comandos sobre rutas estáticas y/o rutas por defecto. Si en cambio estamos utilizando una red más grande, o la estructura de la red cambia frecuentemente, entonces utilizaremos la ventaja que nos ofrece Zebra sobre los protocolos de routing dinámicos, soportando protocolos como **RIP**, **OSPF**, o **BGP**.

Tradicionalmente, la configuración de un router basado en UNIX se realizaba mediante los comandos **ifconfig** y los comandos del tipo **route**. El estado de las tablas se podía mostrar mediante la utilidad **netstat**. Estos comandos solamente se podían utilizar trabajando como root. Zebra, sin embargo tiene otro método de administración. En Zebra existen dos modos de usuario. Uno es el modo normal y el

otro es el modo de enable (habilitado). El usuario de modo normal únicamente puede ver el estado del sistema, sin embargo el usuario de modo enable puede cambiar la configuración del sistema, Esta cuenta independiente de UNIX puede ser de gran ayuda para el administrador del router. Actualmente, Zebra soporte los protocolos de unicast más comunes. Los protocolos de routing Multicast como **BGMP**, **PIM-SM**, **PIM-DM** serán soportados en Zebra 2.0.

El soporte de **MPLS** está siendo programado actualmente. En el futuro, control de filtros TCP/IP, control de calidades de servicio QoS, la configuración de diffserv será añadida a Zebra. El objetivo de Zebra es conseguir un software de routing productivo de calidad y gratuito.

8.1.1 Arquitectura del Sistema

El software tradicional de routing esta compuesto por un programa o proceso único que proporciona todas las funcionalidades de los protocolos de routing. Zebra sin embargo tiene una visión distinta. Está compuesto por una colección de varios demonios que trabajan juntos para construir una tabla. Hay vario demonios de routing específicos que se ejecutan junto con e zebra, el kernel gestor del routing.

El demonio **ripd** maneja el protocolo RIP, mientras que el demonio **ospfd** controla el protocolo OSPFv2. **bgpd** soporta el protocolo BGP-4. Para cambiar la tabla de routing del kernel y la redistribución de rutas entre distintos protocolos de routing tenemos la table de routing del kernel controlada por el demonio **zebra**. Es sencillo añadir nuevos demonios de protocolos de routing el sistema global de routing sin afectar a otro software. Para ello hay sólo es necesario ejecutar los demonios asociados a los protocolos de routing a utilizar. Realizando esta operación, el usuario

puede ejecutar un determinado demonio y enviar reportes a la consola central de routing.

No es necesario ejecutar esos demonios en la misma máquina. Es posible ejecutar varias instancias del mismo demonio de routing en la misma máquina. Esta arquitectura crea nuevas posibilidades para el sistema de routing.

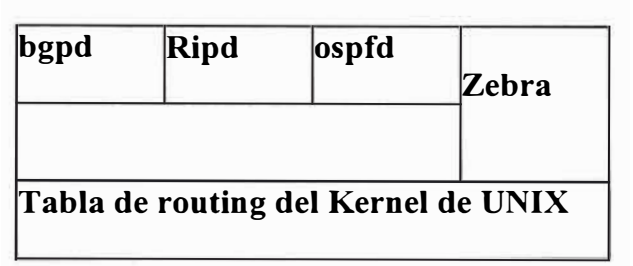


Figura 8.1 Arquitectura del software Zebra.

Arquitectura del Sistema Zebra

La arquitectura multiproceso nos permite un sistema más fácilmente extensible y gestionado y por supuesto nos permite un sistema totalmente modular, a la vez que nos permite varios ficheros de configuración e interfaz de terminal.

Ya que cada demonio tiene su propio fichero de configuración e interfaz de terminal, cuando se quiere configurar una ruta estática, esta se configura en el fichero de configuración **zebra**. Cuando se configura una red BGP hay que hacerlo en el fichero de configuración **bgpd**, esto es bastante fastidioso. Para solucionar este problema existe un interfaz shell integrado llamado **vsh**. **vsh** conecta cada demonio funcionando como un proxy para la entrada del usuario.

Plataformas Soportadas

GNU/Zebra ha sido probado en:

- GNU/Linux 2.0.37
- GNU/Linux 2.2.x
- GNU/Linux 2.3.x
- FreeBSD 2.2.8
- FreeBSD 3.x
- FreeBSD 4.x
- NetBSD 1.4
- OpenBSD 2.5
- Solaris 2.6
- Solaris 7

8.1.2 Servidor de Rutas

En un nodo de intercambio de Internet, muchos ISPs (Proveedores de Servicio de Internet), se encuentran conectados unos con otros a través de conexiones EBGP (BGP externo). Normalmente estas conexiones EBGP se hacen formando un mallado completo. Del mismo modo que en la formación de un mallado IBGP, este método tiene un problema de escalabilidad.

Este problema de escalabilidad es bien conocido, el Servidor de Rutas es un método para solventar este problema, el encaminador BGP de cada ISP establece sesión sólo con el Servidor de Rutas, este a su vez, sirve la información BGP a todos los demás

encaminadotes BGP. Aplicando este método el número de conexiones BGP se reduce de C=(n*(n-1)/2) a C=(n).

A diferencia de los encaminadores BGP normales, el Servidor de Rutas debe tener muchas tablas de rutas distintas para manejar cada una de las diferentes políticas de enrutamiento de cada portavoz BGP. Llamamos a cada una de estas tablas una vista (**view**) diferente. **bgpd** puede trabajar como un encaminador BGP normal, como un Servidor de Rutas o como los dos a la vez.

8.2 Implementación del Graduando.

Luego de investigar y realizar los pasos que anteriormente se detallan se logro instalar el software Zebra 0.93b en una PC. Linux Red Hat 7.3 , como se muestra en el siguiente grafico.

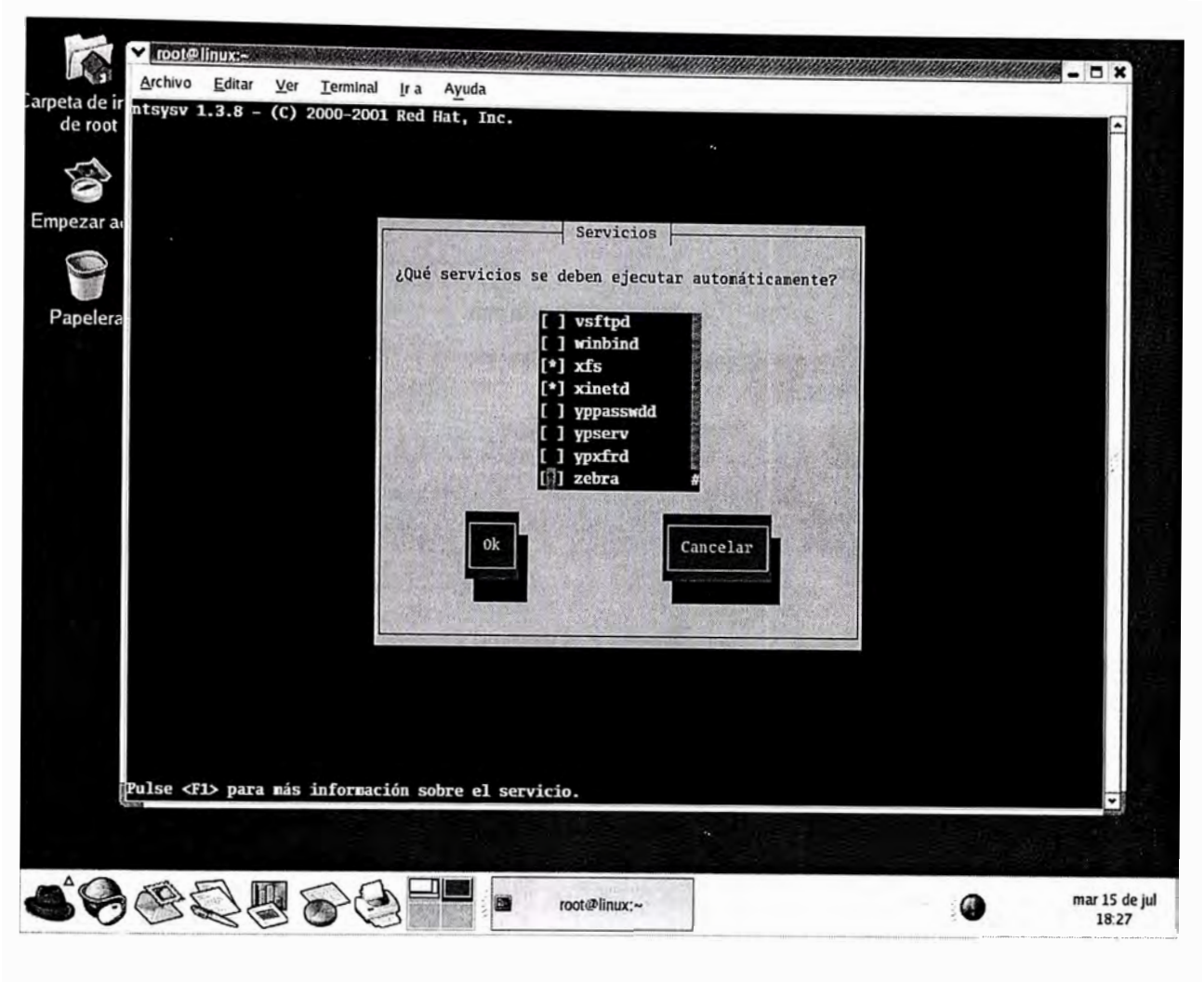


Figura 8.2 Captura de pantalla de Linux con servicio Zebra.

En el gráfico se obtiene de la captura de la pantalla de la PC Linux, se observa que el servicio zebra esta disponible.

Se implemento un escenario de pruebas, en la que se establece una sesión BGP entre la PC Linux y un router ERX de Unisphere, como se muestra en el siguiente dibujo.

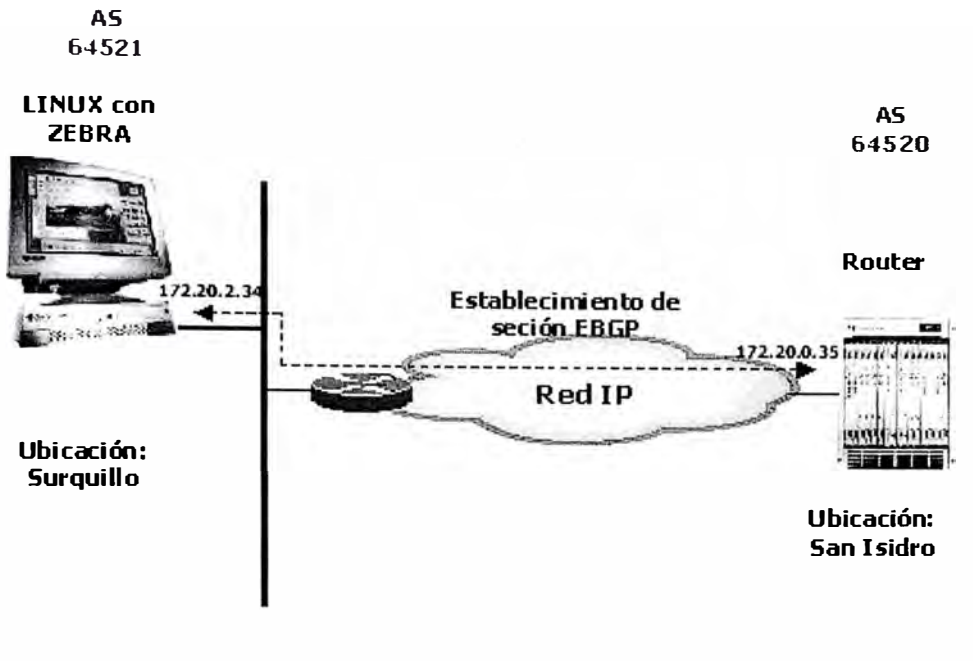


Figura 8.3 Escenario de pruebas con router Linux.

Se realizo la configuración adecuada de los equipos para dicho fin, a continuación se muestra la configuración realizada en el software Zebra en la PC Linux.

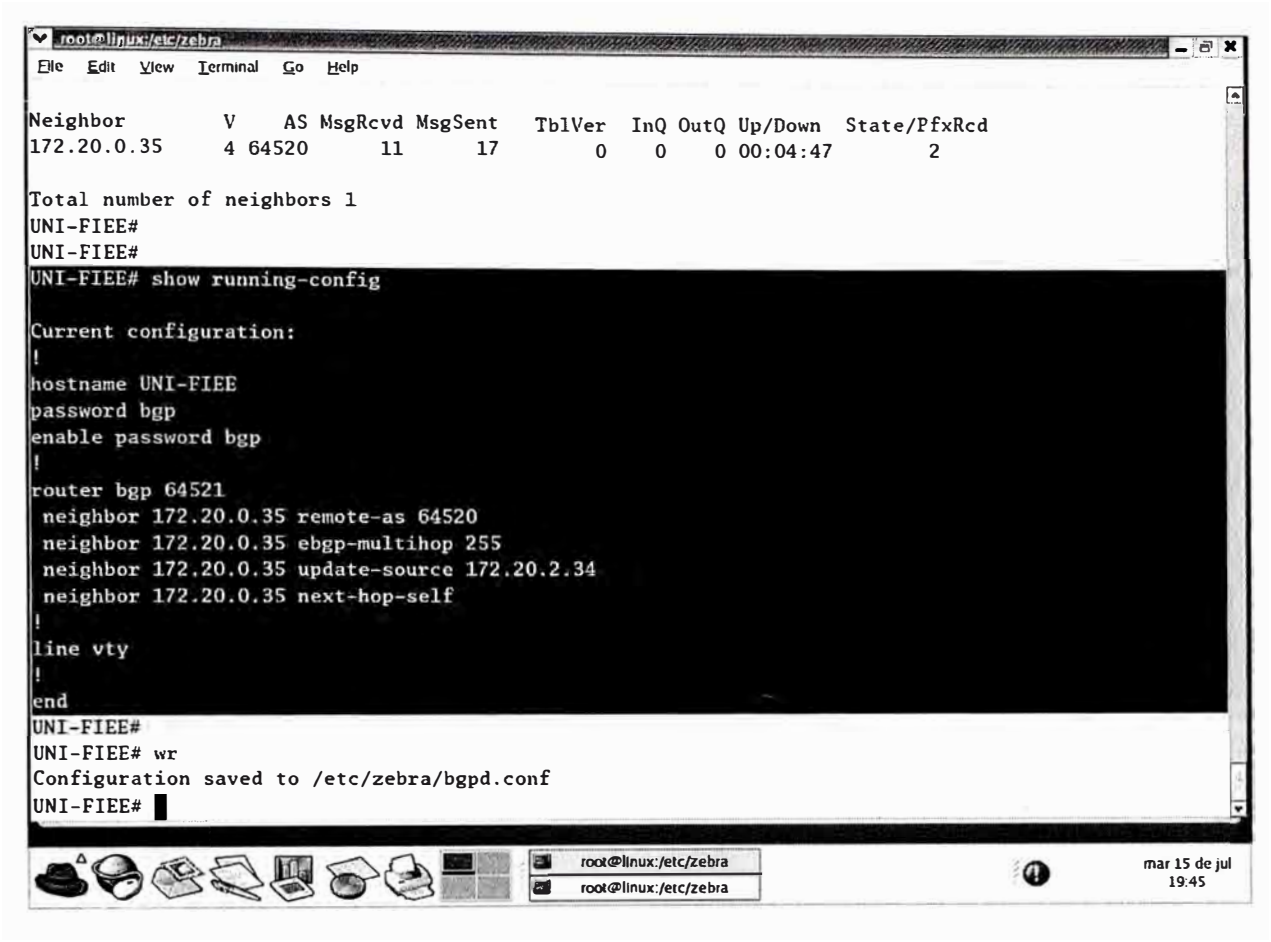


Figura 8.4 Captura de pantalla de la configuración de router Linux.

Se observa que el hostname es UNI-FIEE.

A continuación se muestra la configuración del router ERX.

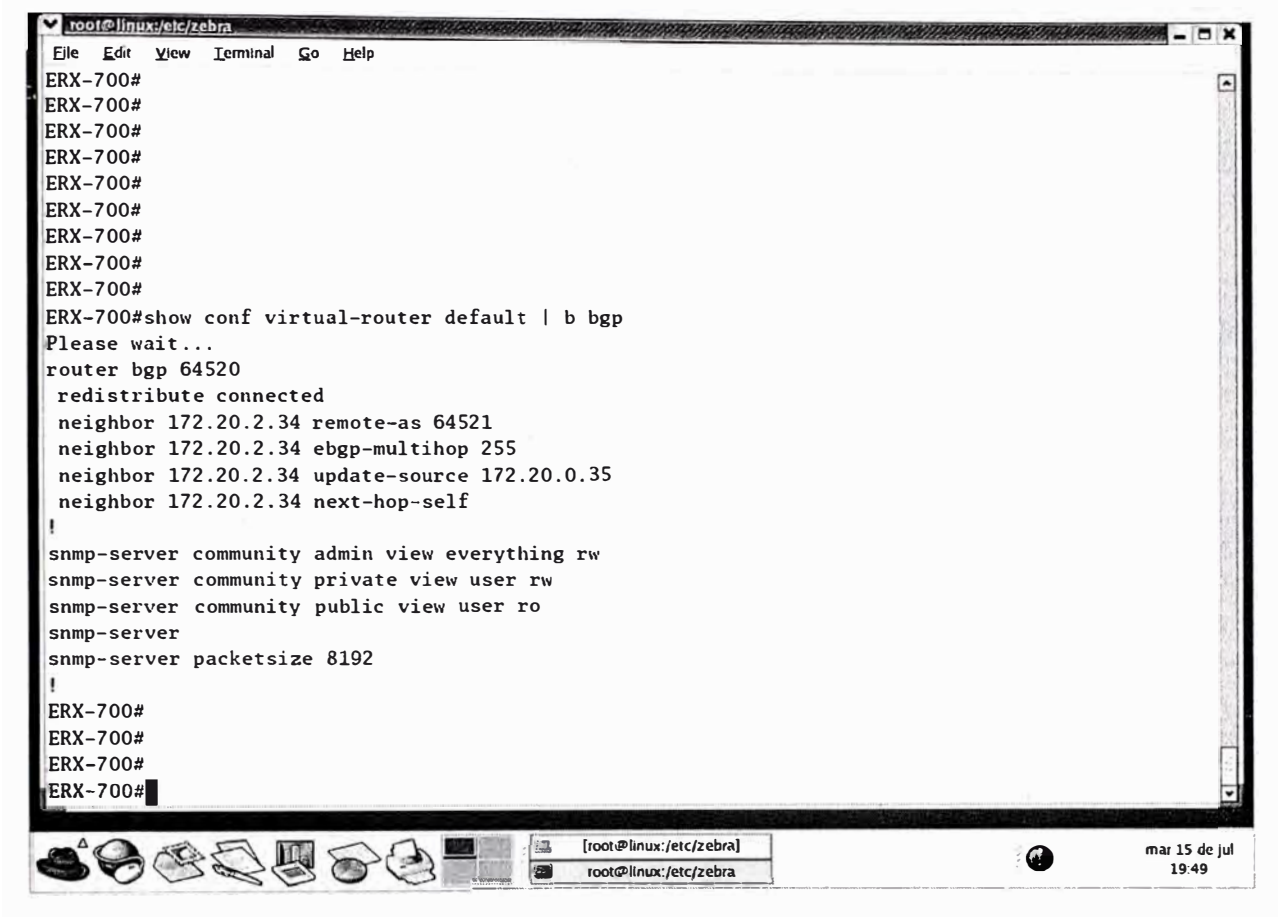


Figura 8.5 Captura de pantalla de la configuración de router ERX-700.

En la grafica siguiente se observa que la sesión BGP se ha establecido y que se están reciviendo 2 redes, asi también que se ha variado el hostname a TITULACION-WILMER

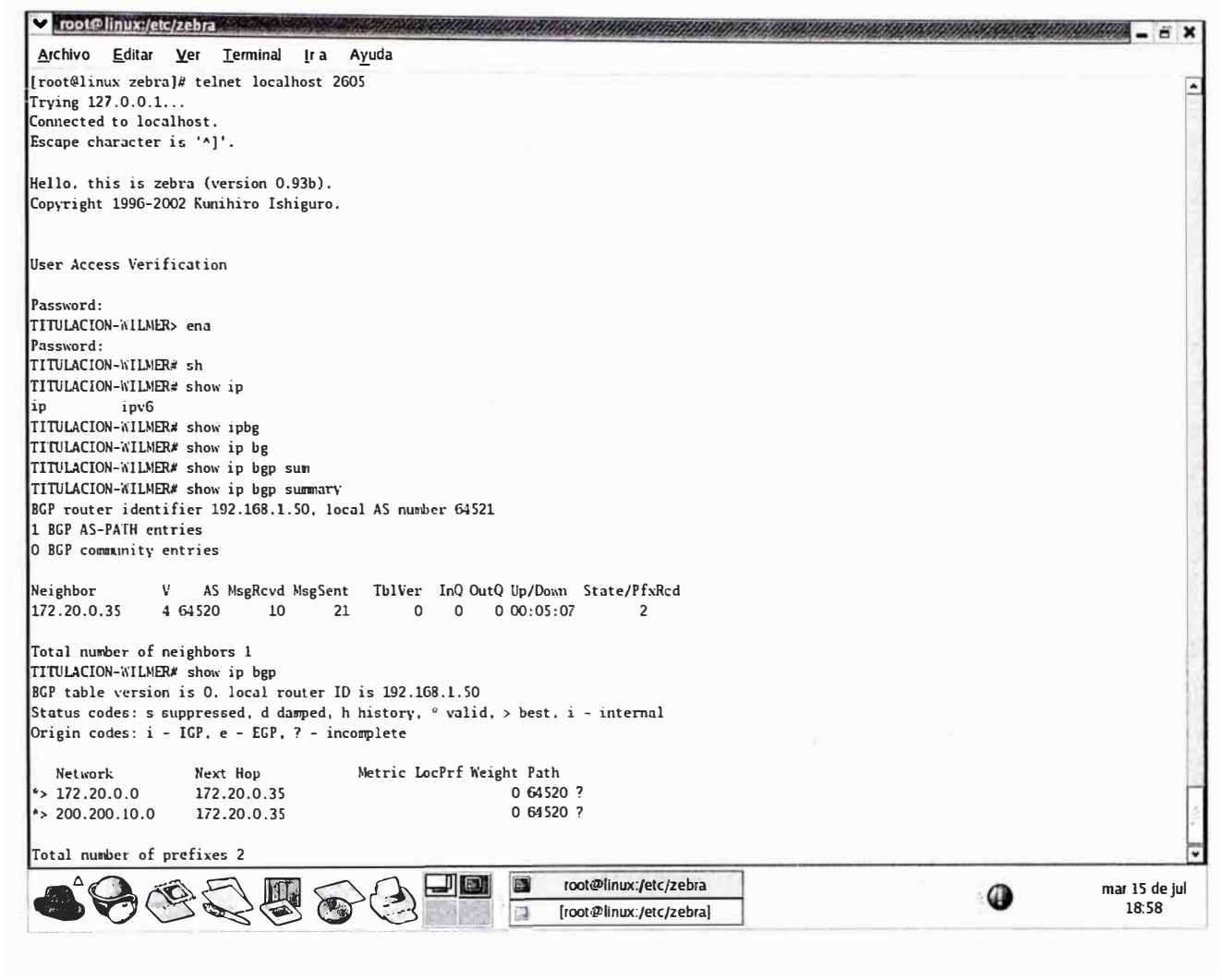


Figura 8.6 Captura de pantalla de la sesión BGP establecida.

CONCLUSIONES

1. La complejidad de los problemas de enrutamiento están estrechamente relacionadas con el crecimiento y la evolución de la Internet, por eso antes de investigar más profundamente los protocolos de enrutamiento, es importanmu y útil tener una perspectiva general sobre temas como el direccionamieto IP y el agotamiento del espacio de direcciones en Internet, organizaciones importantes, proyectos como el NAP, Los arbitros de rutas, Registros de Internet etc.
2. Es importante comprender los servicios básicos y las características de un proveedor de acceso a Internet (ISP), y los parámetros y causas que afectan la conexión , el precio no debería ser el principal factor sobre el que se basa las decisiones de elegir un ISP, sino más bien factores como la redundancia, diseño del backbone , estabilidad del sistema, los cuellos de botella etc.

3. Los comportamientos del enrutamiento en Internet se ven afectados por los comportamientos de los protocolos de enrutamiento y el tráfico dado sobre una infraestructura física ya establecida, el buen diseño de la infraestructura y su mantenimiento son factores primordiales para un enrutamiento eficiente en Internet.
4. El protocolo BGP, ha pasado por diversas fases y mejoras desde su versión original BGP-1, actualmente la versión BGP-4 posee muchas funcionalidades y atributos que nos permiten influir en el proceso de decisión de BGP, de tal manera que podamos decidir el curso de nuestro tráfico IP, la redundancia, simetría y equilibrio de carga.
5. En la actualidad es muy importante que el tráfico IP que fluye desde y hacia Internet cumpla ciertas características, de calidad como llegar a un destino con el menor número de saltos, que los retardos sean mínimos, que no exista pérdidas de paquetes, en caso de falla de nuestros proveedores de acceso a Internet esto se solucione en el menor tiempo posible, todo esto se estipula en los contratos con acuerdos de nivel de servicio conocidos como SLA.
6. El presente informe consta de una aplicación practica que consiste en la implementación de un router en Linux, usando el software libre Zebra , el cual es muy útil ya que puede ser usado en empresas que no manejen tan alto tráfico y con fines didacticos , esto principalmente por el tema de los elevados costos que implica equipos de marcas conocidas.

ANEXO A

RELACIÓN DE ACRÓNIMOS

ACRÓNIMOS

ATM. (Asynchronous Transfer Mode). Modo de Transferencia Asíncrono.

BIOS. (Basic Input/Output System). Sistema Básico de Entrada/Salida.

CLI. (Command Line Interface). Interfaz de Línea de Comandos.

DHCP. (Dynamic Host Configuration Protocol). Protocolo de Configuración Dinámica de Servidor.

DNS. (Domain Name System). Sistema de Nombres de Dominio.

DOS. (Disk Operating System). Sistema Operativo de Disco.

FDDI. (Fiber Distributed-Data Interface). Interfaz de Datos por Distribución de Fibra.

FTP. (File Transfer Protocol). Protocolo de Transferencia de Archivos.

HDLC. (High-Level Data Link Control). Control de Enlace de Datos de Alto Nivel.

HTML. (Hypertext Markup Language). Lenguaje que Señala Hipertexto.

HTTP. (Hypertext Transfer Protocol). Protocolo de Transferencia Hipertexto.

IEEE. (Institute of Electrical and Electronics Engineers). Instituto de Ingenieros Eléctricos y Electrónicos.

IP. (Internet Protocol). Protocolo de Internet.

ISDN. (Integrated Services Digital Network). Red Digital de Servicios Integrados.

LAN. (Local Area Network). Red de Área Local

NetBEUI. (NetBIOS Extended User Interface). Interfaz Extendida de Usuario

NetBIOS. (Network Basic Input/Output System). Sistema de Red Básico de Entrada/Salida.

NFS. (Network File System). Sistema de Archivos de Red.

NIC. (Network Interface Card). Tarjeta de Interfaz de Red.

NTFS. (NT File System). Sistema de Archivos de NT.

PC. (Personal Computer). Computadora Personal.

PPP. (Point-to-Point Protocol). Protocolo Punto a Punto.

SCSI. (Small Computer System Interface). Interfaz de Sistema de Computadoras Pequeñas.

SMTP. (Simple Mail Transfer Protocol). Protocolo de Transferencia de Correo Simple.

SOHO. (Small Office Home Office). Oficina en Casa/Oficina Pequeña.

TCP/IP. (Transmission Control Protocol/Internet Protocol). Protocolo de Control de Transmisión/Protocolo de Internet.

URL. (Uniform Resource Locator). Localizador Uniforme de Recursos.

WAN. (Wide Area Network). Red de Área Amplia.

ANEXO B

RELACIÓN DE FIGURAS Y TABLAS

FIGURAS

TABLAS

BIBLIOGRAFÍA

[1] Curso de Capacitación-Telematic, "Curso avanzado de BGP", Telematic, Abril- 2002

[2] Jordi Tarrigas, "Difusión de redes dentro de Internet", Editorial Lmdata, Agosto-2000

[3] Sam Halabi y Day McPherson, "Capabilities Advertisement with BGP-4, 2da Edición", Ciscopress, Julio-1999.

[4] Lacnic , "Implementado BGP", LACNIC 2002

[5] RFC 1771 – A Border Gateway Protocol 4 (BGP –4)

[6] RFC 2842 – Capabilities Advertisement whith BGP-4

[7] RFC 1700 – Assigned Numbers

[8] http://www.irr.net

[9] http://www.isi.edu

[10] http://www.ietf.org