UNIVERSIDAD NACIONAL DE INGENIERÍA

FACULTAD DE INGENIERÍA ELÉCTRICA Y ELECTRÓNICA

"INTERCONEXIÓN DE SISTEMAS AUTÓNOMOS EN REDES IP"

INFORME DE SUFICIENCIA

PARA OPTAR EL TÍTULO PROFESIONAL DE:

INGENIERO ELECTRÓNICO

PRESENTADO POR:

FERNANDO JUAN TORRES QUINTANA

PROMOCIÓN
1992-2

LIMA – PERÚ
2002

Dedicado a mis padres en retribución a su esfuerzo y comprensión, gracias al Señor.

INTERCONEXIÓN DE SISTEMAS AUTÓNOMOS EN REDES IP

SUMARIO

El presente estudio está enfocado a la interconexión de Sistemas Autónomos en redes TCP/IP; que se inicia con una breve introducción en el prólogo. En el capítulo I se define el concepto de Sistemas Autónomos enunciando sus principales características y el ámbito en el cual se aplican. En el capítulo II se estudia los protocolos de comunicación de rutas que se emplean al interior del Sistema Autónomo, comenzando desde el protocolo IP; para luego ocuparnos de las características generales de los protocolos de vector-distancia y estado de enlace; pasando después al estudio de dos de los más importantes protocolos utilizados, los protocolos RIP y OSPF.

El capítulo III se ocupa de los protocolos que se emplean en la interconexión entre los Sistemas Autónomos reseñando los orígenes de estos protocolos, en el estudio del protocolo EGP, el cual es el antecesor de todos; luego se pasa al estudio del protocolo BGP; protocolo que se ha convertido en el estándar de facto en la industria. En el capítulo IV se analiza la interconexión desarrollada en el Nap-Perú; este es el caso de estudio típico de interconexión de Sistemas Autónomos.

Por último se tienen las conclusiones alcanzadas; los anexos que detallan los aspectos considerados necesarios de consulta rápida para la presente obra, relacionados con los importantes conceptos de subredes, numeración IP y eventos del protocolo BGP; y finalmente se detalla la bibliografía.

ÍNDICE

PRÓLOGO

El presente estudio tiene como objetivo desarrollar y plantear soluciones al problema al cual se enfrentan los ingenieros de redes de datos en la interconexión de Sistemas Autónomos dentro del universo de redes TCP/IP. Con el fin de evitar confusión y con el objeto de mantener uniformidad con la literatura relativa a los temas abordados, se mantiene el nombre en inglés de algunos parámetros técnicos propios de cada protocolo.

Se comienza a enfocar el problema de la interconexión apoyándonos en el modelo de redes propuesto por OSI para dividir el problema de interconexión en capas o niveles. Se trata en primer lugar la interconexión propuesta a nivel físico, mostrándose las diversas soluciones disponibles en nuestro medio y que se pueden utilizar actualmente. Es oportuno mencionar que el énfasis del presente estudio esta puesto sobre la Interconexión a nivel IP, sin embargo para tener una idea cabal de la Interconexión en su conjunto, es necesario la revisión del aspecto físico y de enlace de datos involucrados. Luego se avanza al estudio de la Interconexión a nivel del Protocolo IP, desarrollándose los mecanismos que hacen posible que redes de

diversas tecnologías puedan comunicarse entre sí, envolviendo todas las diferencias existentes en las subcapas inferiores de los estratos físico y de enlace de datos, haciendo estas diferencias; intrascendentes, a nivel del estrato de red implementado por el Protocolo IP. En este punto del desarrollo, se define el concepto de Sistemas Autónomos en redes IP; se abarca el conjunto de protocolos de enrutamiento que hacen posible el uso de mecanismos mas elaborados; los cuales a su vez, permiten el empleo de técnicas de automatización en el control y administración de la dinámica de enrutamiento IP. Luego de definir las diferentes clases de protocolos existentes; se estudia con mas detalle los protocolos RIP y OSPF, los cuales son, cada uno de ellos, representativo de su clase. En el capítulo III, se estudian los protocolos externos de interconexión de Sistemas Autónomos EGP y BGP, EGP en su condición de ser el primer protocolo usado y antecesor a BGP, y BGP como el estándar actual en uso. Se tratan luego, los modelos de interconexión que es posible utilizar en el entorno de BGP.

Llegado a este punto del estudio, se esta en capacidad de abordar el entorno de interconexión desarrollado en el Nap-Perú. Se trata en primer lugar los antecedentes y tratativas desplegados durante varios años; proceso que dio como resultado la formación del grupo de trabajo del Nap-Perú; del cual el autor participa en calidad de miembro del Comité Técnico. Este grupo de trabajo fue encargado de afrontar los problemas técnicos; los de tipo administrativo; así como los de tipo normativo y legal. Se estudia a continuación los aspectos relativos a la interconexión a nivel de las tres capas del modelo OSI; terminando luego con el tratamiento del modelo de interconexión aplicado en el entorno BGP-4. Los siguientes puntos tratan las conclusiones y la bibliografía a recurrir para un estudio mas detallado de los

diversos temas involucrados en este texto. Los anexos contienen, entre otros puntos, un resumen referido a las clases de redes y subredes IP y conceptos relativos a CIDR y VLSM; muy importantes para entender el funcionamiento actual de los mecanismos encaminadores de paquetes IP.

Finalmente; es necesaria a estas alturas del presente escrito; la mención correspondiente a todo el grupo de profesionales que formaron el grupo interdisciplinario; como parte del Comité Técnico del Nap-Perú y otros que no siendo miembros del mismo, participaron del grupo de trabajo en algún momento, interviniendo en la ejecución del proyecto. Todos ellos cumplieron meritoriamente las labores encomendadas; suma de esfuerzos que tuvo como resultado el éxito esperado; hoy el Nap-Perú es una realidad.

CAPITULO I

SISTEMAS AUTÓNOMOS EN REDES IP

La Interconexión de Redes TCP/IP se puede analizar desde dos puntos de vista claramente diferenciados: la interconexión física de redes que involucra tanto el nivel físico propiamente dicho, así como el nivel de enlace de datos, y el estrato correspondiente al nivel de Red del Modelo OSI. Es sobre la base de la división anterior propuesta que se estudia la Interconexión de Redes TCP/IP.

1.1 Sistemas Autónomos en Redes IP

El Internet es una Red de redes; es decir; un gran conjunto de redes interconectadas entre sí. Cuando el número de estas redes fue en aumento; quedo claro que el método manual empleado en un comienzo, para definir la accesibilidad individual de cada red, se haría inmanejable e impráctico, aún sin tener presente los cambios que necesariamente deben hacerse a nivel de rutas, como consecuencia de los cambios que se hacen a nivel físico.

Para automatizar y reducir la complejidad de la administración de rutas; se crearon los protocolos de encaminamiento o enrutamiento. Punto fundamental en el diseño de estos protocolos, entre otros, son los dominios de enrutamiento. Dentro de

estos dominios se aplican políticas de rutas homogéneas y sirven para delimitar el ámbito de aplicación de las mismas. Sin embargo esto no significa que la entidad o institución que administra estos dominios sea una sola; puede darse el caso que dos o más administraciones o subunidades administrativas que controlen las políticas de rutas de sus respectivas redes; acuerden políticas comunes; las cuales se aplican de manera uniforme a sus respectivas redes.

Estos dominios de enrutamiento son inflexibles en cuanto a sus políticas, una vez acordadas se cumplen inexorablemente, caso contrario, simplemente, no sería posible la comunicación interna ni externa entre ellos. En estos dominios solo se utiliza en la amplia mayoría de casos un único protocolo de enrutamiento para las redes involucradas.

Nuevamente; a medida que creció de forma exponencial, el número de redes individuales así como los Dominios de Rutas; factores tales como la rigidez de funcionamiento y la necesidad de individualizar grupos de redes autónomas, cobraron mayor importancia; es debido a estos factores; que se hace indispensable descentralizar la administración de las redes en base a las divisiones de orden administrativo, de una manera natural y lógica. Es así como se populariza el término y cobran relevancia los Sistemas Autónomos dentro de las redes TCP/IP.

Sistemas Autónomos, es el nombre que toman las redes o grupos de redes que están manejadas por una sola entidad administrativa, no importando la naturaleza de la misma. En estos Sistemas Autónomos; representados en la figura 1.1; las administraciones correspondientes se encargan de diseñar y aplicar las políticas internas de cada uno de ellos. En otras palabras; estas administraciones; son las encargadas de definir; entre otras cosas; los protocolos de enrutamiento, políticas de

seguridad, topología, administración de números IP, etc., a emplearse, con total independencia.

Estos Sistema Autónomos se identifican en el ámbito de los mismos por un número ASN (Autonomous System Number), el cual es asignado de manera única, por la Autoridad de Registro que corresponda. Para el caso de Internet, esta asignación es única; es decir, que no se asigna el mismo número AS a administraciones distintas. Para el caso de redes privadas también se debe hacer asignaciones únicas, tan igual como se hace con los números IP. En términos generales, para todo conjunto de redes TCP/IP, sin conexión alguna con otro conjunto, se debe aplicar la regla de asignación única.

1.2 Interconexión Física

Las tecnologías disponibles actualmente permiten la interconexión en el estrato físico; sin embargo esto solo es posible en redes de características físicas iguales. En el caso de redes de área local, la tecnología Ethernet es la predominante; otras clases de redes en el ámbito LAN son Token Ring, FDDI o ATM. Si la distancia entre redes es corta; es decir, están dentro del rango de acción de estas redes, la conexión puede hacerse con puentes, repetidores, conmutadores o utilizar redes de mayor capacidad y alcance como FDDI, si fuera posible. En caso de que las distancias esten fuera del rango de acción de estas; se impone el uso de redes WAN de gran alcance. Las tecnologías disponibles de redes de área extensa WAN permiten en nuestros días interconectar, en la práctica, cualesquiera dos lugares del planeta o del espacio cercano a la Tierra; asimismo, la capacidad de estas va en aumento conforme el avance científico y tecnológico. La figura 1.2 muestra ejemplos de redes interconectadas.

Las tecnologías de interconexión fisica se pueden agrupar en 2 grandes campos: Comunicaciones Digitales Dedicadas y Comunicaciones Digitales Conmutadas. Para el primer caso tenemos por ejemplo los circuitos TDM (Time Division Multiplexing) cuyo ancho de banda se encuentra disponible únicamente para los puntos que interconectan; es decir, es un medio no compartido. La fracción mínima en esta tecnología es de 64kbps y múltiplos de 64 kbps, disponibles en jerarquías de T1 o E1 según sea el caso y en múltiplos de estos a su vez para capacidades mayores. Genéricamente se les conoce como conexiones punto a punto.

Para el caso de Comunicaciones Digitales Conmutadas, se tiene disponible ISDN, Frame Relay, ATM, SMDS SONET. Para el caso de ISDN están disponibles a través de puertos BRI y PRI accesos de 128 kbps y 2048 kbps respectivamente. Para el caso de Frame Relay y ATM existen capacidades mayores como T1, E1, T3, E3, STM-1 y múltiplos. Los circuitos SMDS y SONET son de gran capacidad, iguales o mayores a STM-1. Los medios de transporte pueden ser cables de cobre, radioenlaces, para el caso de baja y media capacidad y enlaces de radio y fibra óptica para el caso de alta capacidad de transporte, prefiriéndose esta última por su alta eficiencia y confiabilidad.

1.3 Interconexión IP

La interconexión IP es posible debido a su naturaleza de Arquitectura abierta, la cual; por definición; no esta restringida a un solo medio de transmisión fisica o tecnología de capa 1 o 2; conforme al Modelo de Referencia OSI; por el contrario; deja abierta la posibilidad de interoperar con cualquier otra tecnología de transmisión creada o por crearse, siendo necesaria para ello la implementación de las interfaces de las capas 1 y 2 que correspondan a fin de incorporar las tecnologías y protocolos

de las otras capas adicionales del Modelo de Referencia OSI; figuras 1.3 y 1.4. Una vez que ya tenemos implementada e interconectada la red a nivel físico, veamos como se implementa la interconexión a nivel TCP/IP.

Cada tecnología de red implementa un método de identificación única a nivel físico, que no son compatibles; necesariamente; con otras clases de redes. El protocolo ARP (Address Resolution Protocol) es el protocolo especificado para asociar la dirección física de cada nodo de la red; con la dirección IP correspondiente. Es el puente de tránsito entre el estrato IP y cualquier otro correspondiente al nivel dos de OSI. El formato de este protocolo no presenta una trama de encabezado fijo y su uso se restringe a redes directamente conectadas; formándose una tabla ARP que muestra la equivalencia entre un número IP y el número físico que le corresponde de manera biunívoca. Para el caso de las redes que no se encuentra directamente conectadas, es necesario el uso de encaminadores o enrutadores. Estos pueden ser una computadora de propósito general con un programa de enrutamiento, o un equipo dedicado, tal como lo es un ruteador. La función principal de estos equipos de enrutamiento es identificar la ruta requerida con la ruta que tienen registrada en su tabla de rutas; para de esta manera poder alcanzar la dirección de destino final de cada paquete IP que le es entregado.

Estos equipos mantienen una tabla de rutas con interfaces asociadas a cada ruta especifica de tal forma, que pueden encontrarse en estas tablas la ruta adecuada para alcanzar la red a la cual pertenece el número IP de destino; además de la tabla ARP que contiene el número IP asociado al número de estrato dos de OSI. Pero en el caso de que la ruta que se necesita no se encuentre especificada en su respectiva tabla de rutas, se designa una ruta genérica, denominada ruta por omisión.

1.3.1 Modelo de Decisión de Rutas IP

Para llevar a cabo la tarea de transportar el paquete a su destino; el algoritmo de encaminamiento IP aplica las decisiones, que veremos a continuación, en el proceso de entrega de datagramas, en el orden siguiente:

- Extrae la porción de red de la dirección IP destino.
- Si la red destino esta directamente conectada, revisa la tabla ARP para ubicar la dirección física y hacer la entrega directamente.
- Si la red destino corresponde a un anfitrión, la envía al salto correspondiente.
- Si la red destino corresponde a una red que no está directamente conectada, la envía al salto correspondiente.
- Si la red destino no corresponde a ninguna ruta de la tabla, la envía al salto correspondiente a redes no especificadas.
- Si la red destino no corresponde a ninguna ruta de la tabla de rutas, entonces elimina el paquete IP.

La figura 1.5 representa el diagrama de flujo de este proceso de toma de decisión. En este proceso, si alguna de las decisiones no es verdadera; entonces; se prosigue con la siguiente, hasta llegar eliminar el paquete, en el último paso de decisión. Cuando las rutas para alcanzar las redes destino no son numerosas ni tampoco las modificaciones que ocurren por eventuales fallas de comunicación o cambios en las redes; es posible administrar estas tablas de rutas manualmente de manera muy eficiente. Pero; supongamos que el número de redes aumenta rápidamente así como los cambios que se deban realizar; entonces; como es fácilmente deducible, la administración manual de estas tablas se convertirá tarde o temprano en un método muy engorroso e ineficiente, lo cual se reflejará en que las tablas de rutas no estén

debidamente actualizadas, con la consiguiente pérdida de paquetes. Es así como surge la necesidad de automatizar este proceso dando origen a los Protocolos de Encaminamiento.

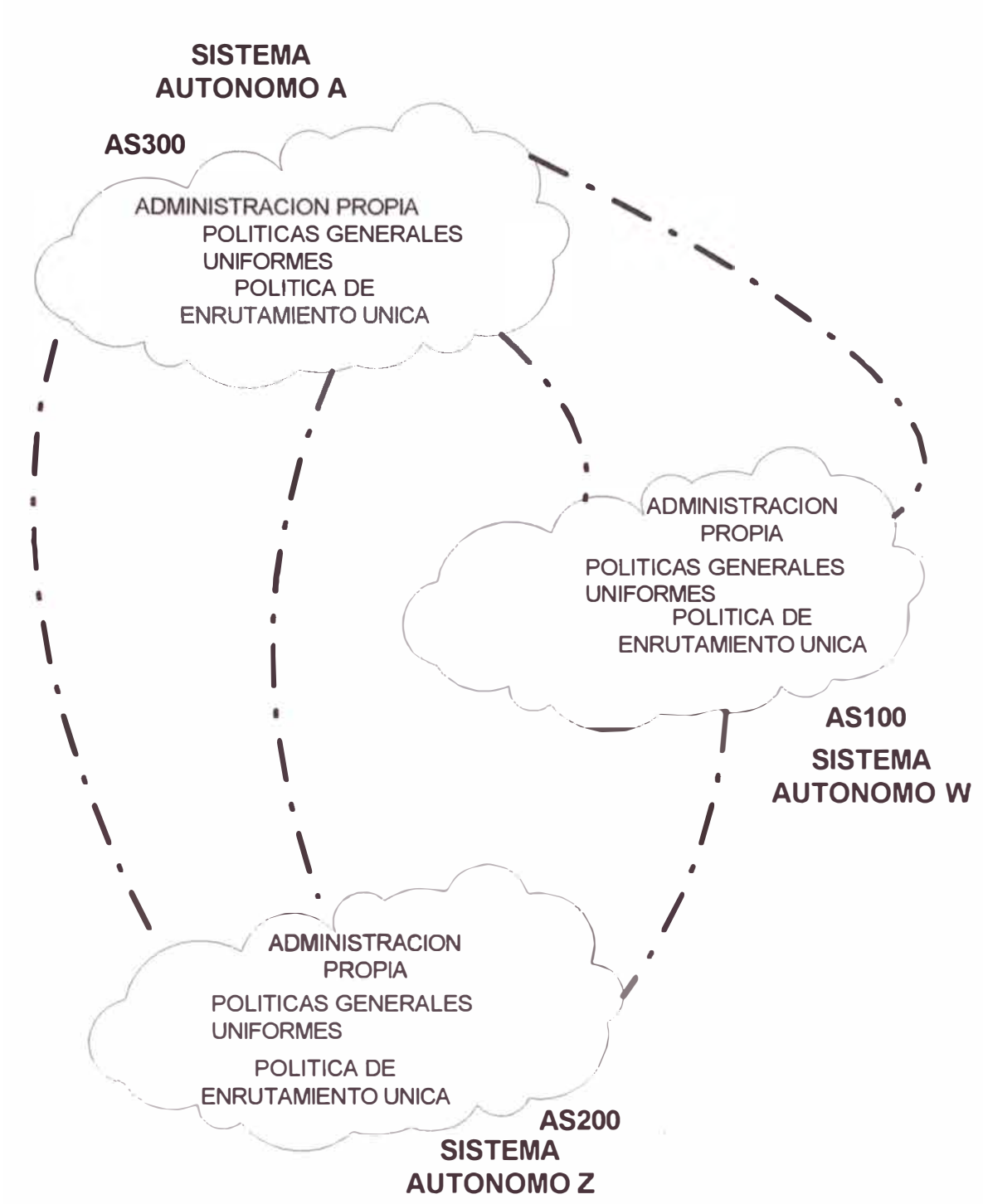


FIG 1.1 Sistemas Autónomos

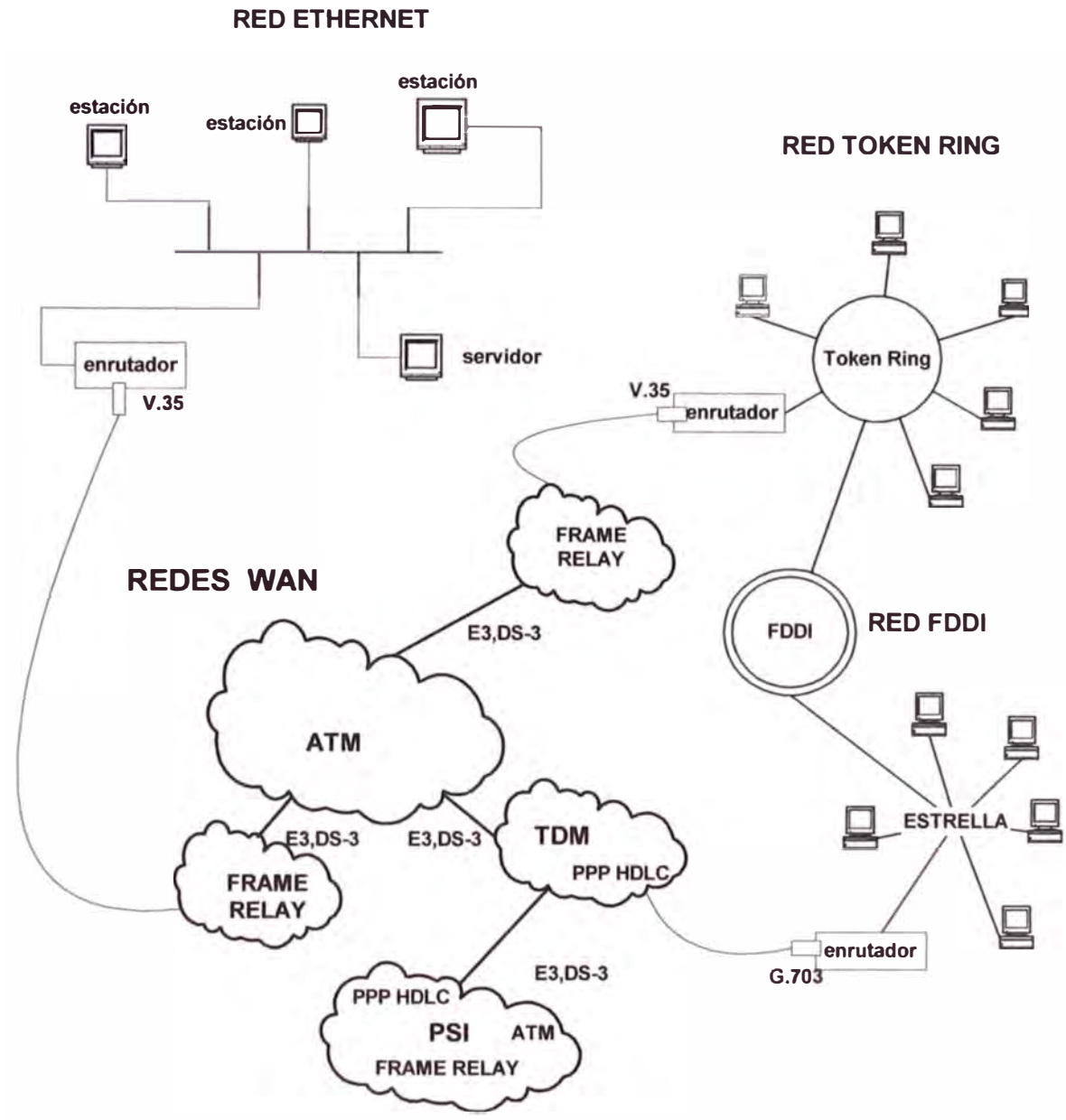
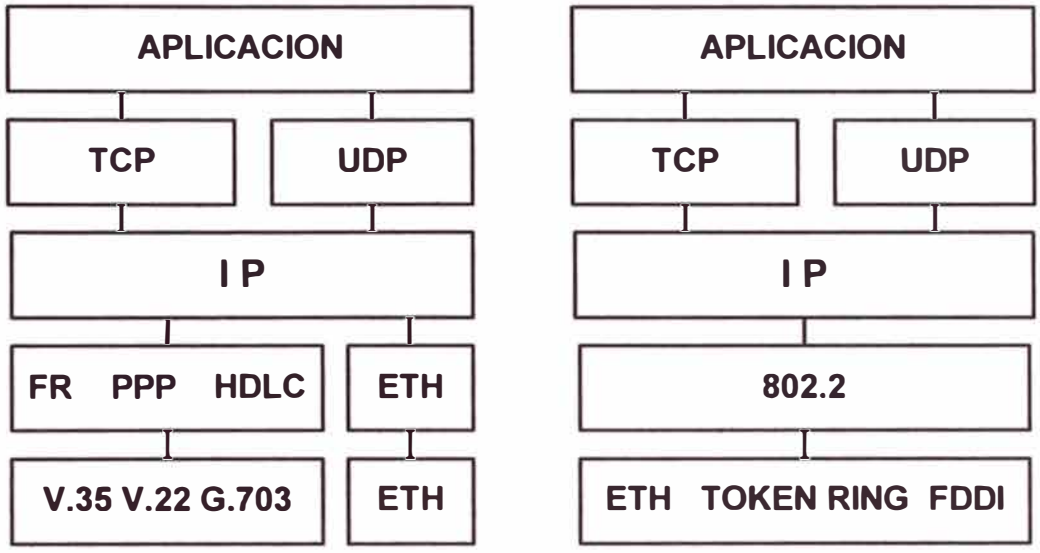


FIG 1.2 Interconexión de Redes



Modelo WAN **Modelo LAN**

FIG 1.3 Modelo de Interconexión por Capas

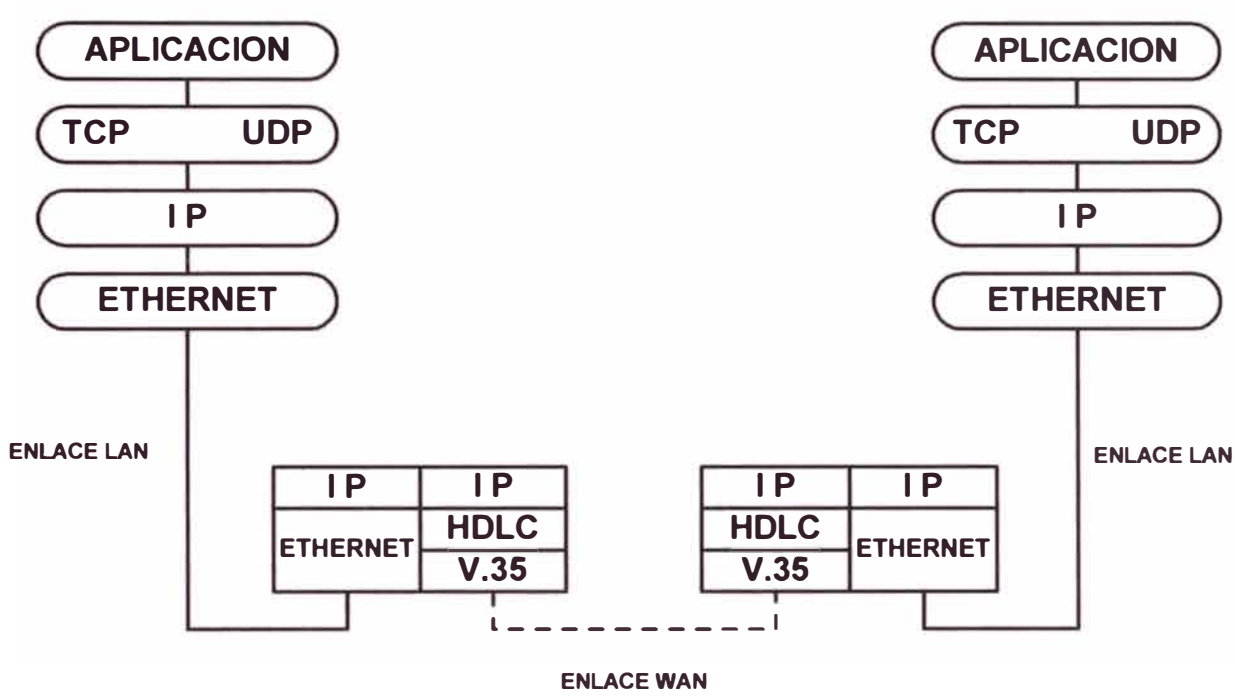


FIG 1.4 Interconexión IP

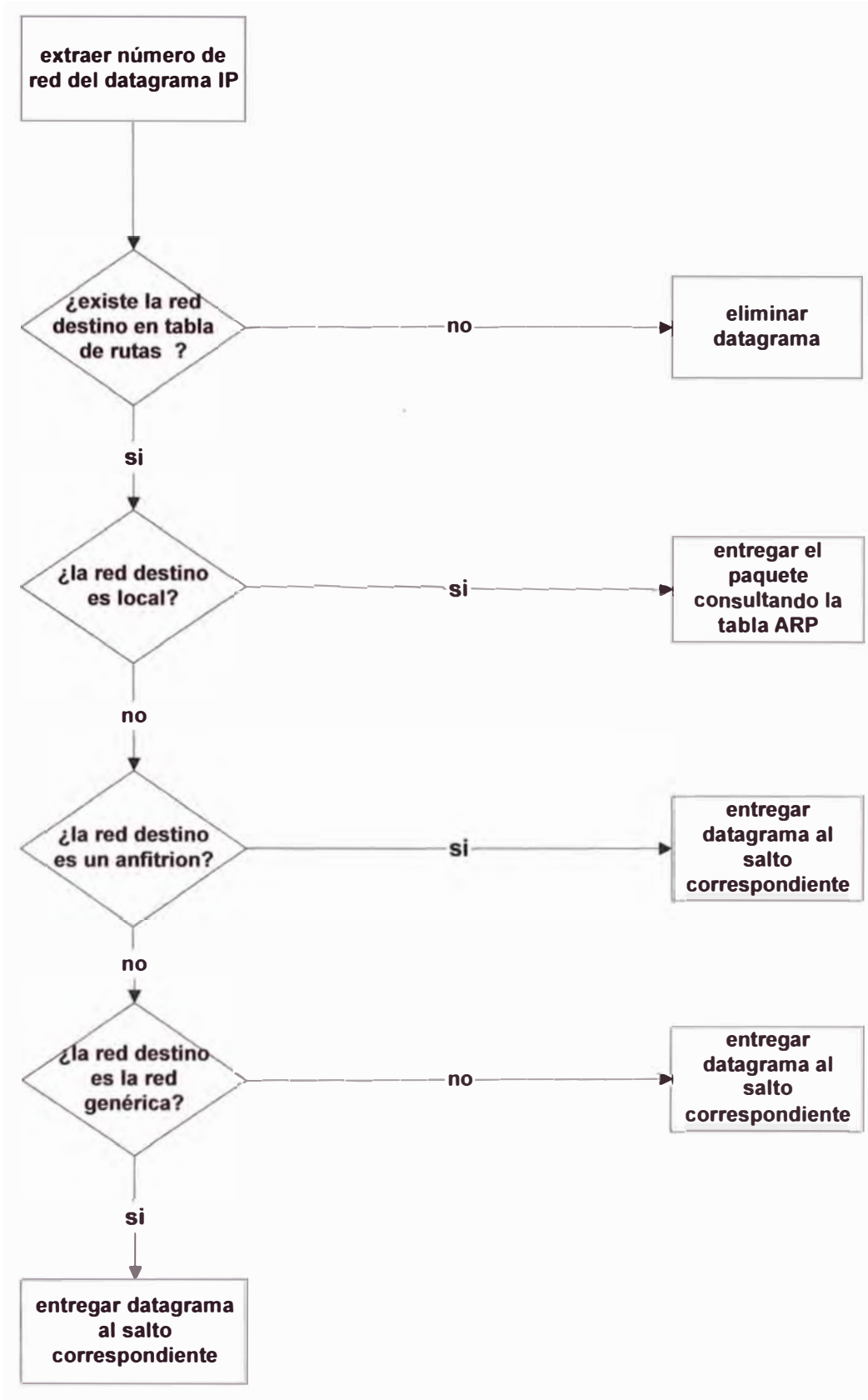


Fig 1.5 Algoritmo IP

CAPITULO II

PROTOCOLOS DE ENCAMINAMIENTO INTERNO DE SISTEMAS AUTÓNOMOS

2.1 Protocolos de Encaminamiento

Los protocolos de encaminamiento o ruteo como también se les conoce, surgen de la necesidad de hacer eficiente y manejable el constante crecimiento en el número de redes que se conectaban a otras redes, cada vez de manera mas acentuada, así como las modificaciones que inevitablemente se sucedían, a nivel fisico. La aparición de estos nuevos protocolos trajo como consecuencia la aplicación de nuevos conceptos como rutas dinámicas y rutas estáticas, asociándose las primeras con los protocolos de enrutamiento y las segundas con las rutas que se configuraban de manera manual en los ruteadores. Los Protocolos RIP y OSPF son ejemplo de estos protocolos de rutas.

2.1.1 Métricas

Existen algunas características muy importantes comunes a los protocolos de encaminamiento, las cuales se refieren con el nombre genérico de métricas. Estos parámetros son los que se consideran en su diseño, funcionamiento, performance y análisis; los cuales se listan a continuación:

- Retardo
- Ancho de banda
- Carga
- Número de saltos
- Costo
- Unidad de Transferencia Máxima

Retardo

Es el tiempo requerido por un paquete para llegar desde la dirección origen hasta la dirección destino a través de toda la Red. Este retardo depende de muchos factores como por ejemplo el ancho de banda disponible, la congestión de los enlaces, la capacidad de procesamiento de los ruteadores de transporte; así como la distancia física entre los puntos de comunicación.

Ancho de Banda

Es la capacidad de transporte disponible en cada enlace físico. Se refiere al valor máximo disponible de transmisión de un enlace. Por ejemplo en un enlace dedicado tenemos capacidades de 64Kbps, y en Fastethernet tenemos 100 Mbps.

Carga

Es el porcentaje de utilización de un recurso de la red que se usa para transmitir los paquetes. Puede ser la carga de la CPU, la carga de una Interface, la tasa de paquetes procesado por segundo. A mayor carga mayor retardo, sin embargo la relación no es necesariamente lineal.

Número de Saltos

Es el número de enlaces de interconexión que un paquete atraviesa hasta llegar a su dirección destino. A mayor número de saltos mas retardo.

Costo

Se refiere al costo financiero de usar un enlace. Es decir para una empresa puede ser más costoso usar una clase de línea de conexión alquilada, en lugar de otra; ya sea por el tipo de tecnología, por tener que alquilar el equipo que sea necesario, o por algún otro concepto.

Unidad de Transferencia Máxima

Es el tamaño de paquete máximo que puede contener una trama a través de un enlace en particular. Por ejemplo la UTM en Ethernet es 1.5 KB, en FDDI es 4 KB.

Estas han sido las métricas más comunes que se usan con relación al enrutamiento.

Se pueden separar los distintos protocolos de enrutamiento en dos grandes grupos; estos se revisan en la siguiente sección.

2.2 Clases de Protocolos.

Existen 2 clases de protocolos, los protocolos Vector-Distancia y los Protocolos de Estado de Enlace. La figura 2.1 resume las principales características de ambas clases.

2.2.1 Protocolos Vector –Distancia

Son aquellos protocolos basados en el algoritmo conocido como Bellman-Ford o Bellman-Fulkerson, algoritmo computacional que permite ubicar la distancia más corta entre 2 puntos. Se le llama Vector-Distancia por el hecho de que la información que es transmitida entre ruteadores contiene la distancia para llegar a una red; así como; el vector dirección, incluyendo el costo en que se incurre para alcanzar el destino a través de este vector. Este fue uno de los primeros algoritmo de rutas que se uso en los inicios de Internet. Este protocolo permite enviar toda la tabla

de rutas que cada ruteador contiene, a todos los ruteadores con los cuales se encuentra directamente conectado. Su punto débil, es el tiempo de convergencia mas alto; además de ser susceptible de generar circuitos cerrados de rutas; sin embargo; tiene la ventaja de ser más sencillo de implementar. Ejemplo de este es el protocolo, es el protocolo RIP (Routing Information Protocol).

2.2.2 Protocolo de Estado de Enlace.

Esta basado en el algoritmo conocido como SPF (Shortest Path First) o algoritmo Dijkstra. Estos protocolos envían periódicamente información actualizada a todos los demás ruteadores; pero; describiendo únicamente el estado de los enlaces propios de cada ruteador y no la tabla de rutas completa. Con esta información, cada ruteador es capaz de elaborar su propia tabla de rutas. Esta clase de protocolos tiene una convergencia rápida; sin embargo; requieren de mayor potencia de procesamiento en CPU y en memoria para operar. En OSPF; por ejemplo cada enrutador calcula de manera independiente su tabla de rutas, haciendo uso de la información del estado de los enlaces que todos los otros ruteadores le envían periódicamente. Cada uno de los encaminadores posee una base de datos donde se almacena la topología completa de toda la red, data mediante la cual calcula su propia tabla de rutas. Cada objeto de esta base de datos contiene identificadores de interfaces, número de enlace, métrica así como el estado del enlace.

2.3 Protocolo RIP

Protocolo de Información de Enrutamiento RIP

El protocolo RIP (Routing Information Protocol) fue desarrollado en base a estudios realizados en el Centro de Investigaciones de la empresa Xerox en EUA, sobre los protocolos de rutas XNS y PUP de Xerox. La popularidad que logró RIP

con el paso del tiempo; se debió en parte a que la Universidad de Berkeley en California EUA, distribuyó el software que implementaba RIP a través de sus sistemas UNIX-BSD, los cuales se volvieron muy populares en ambientes académicos y de investigación. Es oportuno mencionar que el desarrollo de redes IP, Internet y UNIX, están estrechamente ligados desde la década de los 70.

El RIP, que se diseñó originalmente para las redes LAN, se usa también en redes de áreas amplias y existe en dos versiones RIP-1 y RIP-2. Utiliza el protocolo UDP, puerto 520 como plataforma de comunicación;. Se distribuyó como parte de la serie de protocolos de Internet en la distribución de software de la Universidad Berkeley y se constituyó como uno de los primeros protocolos de rutas que por su simplicidad y disponibilidad logró un uso extendido en intranets privadas en los inicios de Internet.

RIP está clasificado como un protocolo de enrutamiento de tipo vector-distancia. RIP construye su tabla de rutas basando su métrica en el número de saltos intermedios que existen en el camino hacia el destino final. RIP utiliza el algoritmo conocido como el algoritmo Ford-Fulkerson o Bellman-Ford, los cuales fueron quienes estudiaron el algoritmo en el cual se cimenta el funcionamiento de RIP. A continuación una breve descripción del funcionamiento de RIP.

2.3.1 Funcionamiento del Protocolo RIP

RIP comunica las direcciones de red y la distancia (número de saltos) y emplea un contador de saltos para calcular el costo de cada ruta; además, asigna un valor máximo de 16 para indicar que una dirección no se puede alcanzar; es decir, 16 representa el valor de *"infinitos"* saltos para indicar que una red es inaccesible.

En la mayoría de los sistemas, un ruteador que usa RIP envía información de rutas cada 30 segundos. Si un ruteador no recibe una actualización por parte de otro

ruteador durante 180 segundos, éste marca como fuera de uso, la ruta de la cual no recibió ninguna actualización. Después de 240 segundos y si todavía no recibe ninguna actualización, el ruteador modifica todos los registros de su tabla de ruta con las *"nuevas"* modificaciones. Pasaremos a explicar con un ejemplo de una interconexión de 3 ruteadores, el funcionamiento del protocolo RIP.

2.3.2 Propagación de Enrutamiento RIP

Se tiene 3 ruteadores que operan bajo RIP, tal como se describe en la figura 2.2. Supongamos que el ruteador A anuncia la red 192.168.25.0 a sus vecinos B y C.

En el primer paso, con una métrica de 1 porque el numero de saltos es 1.

En el segundo paso, B y C reciben este anuncio y le suman 1 a la métrica pues la distancia desde cada uno de ellos a la red aumenta en un salto, y lo registran en su tabla de rutas.

En el tercer paso, B y C intercambian rutas entre ellos, y también con el ruteador A agregando 1 a la métrica, antes de registrar en la tabla de rutas verifican si existe la ruta y cual es la métrica. Dado que es mayor que la existente, entonces no se registra esta nueva ruta. La información que intercambian es de la forma salto-red, es decir para llegar a la red se tienen tantos saltos, que por analogía se asocia a un par de información de la forma vector-distancia. Veamos ahora un caso distinto; la fig. 2.3 representa una red compuesta de cinco segmentos de red nombradas desde la red 1 a la 5. El centro de referencia es el ruteador D (RD). Este propaga información acerca de la red 1 a partir de su interface. Los pares ordenados de la forma (R1, 1) es una indicación abreviada que se usa para informar que la dirección Red1 se encuentra a 1 salto a través del ruteador D; es decir, la métrica en saltos con respecto a una dirección; es decir, el término (vector,distancia).

RC envía su mensaje (R1,2) a la red 3, donde es recibido por RA y RB. Entonces RB inscribe en su tabla de rutas (R1,3), o en otras palabras, que la Red 1 se encuentra a 3 saltos de distancia.

Dado que todos los ruteadores se envían comunicaciones unos a otros, la accesibilidad hacia todas las redes puede calcularse. Como ejemplo, la figura 2.4 también muestra la tabla de rutas guardada por RB.

El campo destino en la tabla, es la dirección de la red de destino, tal como lo percibe RB, en este caso puede ser una dirección IP (RIP también se puede usar con otros protocolos como IPX-SPX). El siguiente salto (próximo nodo) en la tabla, es la dirección del nodo que tiene que recibir el tráfico dirigido a la red asociada. Este registro es 0 en la tabla si esta conectado directamente al ruteador. La columna Métrica expresa cuantos saltos hay entre el emisor del mensaje hacia la dirección de destino. La columna Directo o Remoto es ya sea D (directamente conectada) o R (red remota, no está directamente conectada). La columna RIP o Local es ya sea L (ruta aprendida porque es local y descubierta por ARP, DHCP, etc., o R (ruta aprendida mediante mensajes RIP) La columna Interface identifica el puerto fisico en R3 en el que se hizo la detección de la ruta.

Propagando los Cambios de Rutas

Puesto que RIP actualiza sus vecinos y éstos hacen lo propio con los suyos; y así sucesivamente a través de todo el dominio, a las actualizaciones les toma tiempo propagarse por toda la inter-red. Un ruteador envía toda su tabla de rutas, cada 30 segundos. Por lo tanto, los nodos al final de este anuncio de rutas; también llamado difusión, no descubrirán el cambio, hasta que todos los nodos que intervienen hayan terminado de enviar todas sus propias rutas.

Hay que recordar que la tabla de rutas tiene un registro llamado próximo nodo (o próximo salto). Cuando se recibe el mensaje, RIP chequea la dirección de envío del datagrama IP con el registro del próximo nodo en la tabla. Si concuerdan, se verifica la métrica del mensaje con la métrica de la tabla. Es fácil observar; que los cambios del RIP pueden tomar un tiempo prolongado hasta conseguir la convergencia de rutas; es decir; hasta que la tabla de rutas sea actualizada con todos los cambios producidos en la inter-red.

Unidifusión y Cancelación De Actualizaciones

RIP ha sido diseñado como un protocolo de difusión, pero es posible enviar mensajes RIP para nodos sin difusión, lo que puede ser útil cuando se conecta con otro ruteador en un enlace de punto a punto.

Asimismo, es posible inhabilitar el uso del RIP en interfaces especificadas, de esta forma el administrador de la red evita que los mensajes RIP sean generados en una o varias interfaces específicas.

2.3.3 Contenido de Mensajes RIP

RIP utiliza UDP de manera que sus mensajes se encapsulan en segmentos UDP y se dirigen a un puerto conocido (puerto 520). Dos versiones de RIP se encuentran disponibles: RIP-1 y RIP-2. La ilustración 2.5 y 2.6 muestra los formatos para ambas versiones. El campo de comandos puede contener los valores del 1 al 6, 1 y 2 son los únicos valores especificados Un código de comando igual a 1 identifica un mensaje de solicitud y un valor igual a 2 identifica una respuesta. Para ambas versiones, el campo de versión es 1 ó 2.

Los mensajes tienen una ligera diferencia. Las dos versiones contienen una familia de direcciones, lo cual se codifica como 2 para las direcciones IP. Lo que

sigue; se trata de una dirección IP y su métrica (contador de saltos). Estos campos de información pueden repetirse hasta 25 veces. Esta limitación mantiene el mensaje RIP por debajo de los 512 bytes (4 bytes fijos comando, versión, 2 bytes de todos los ceros + 20 bytes x 25 repeticiones posibles = 504 bytes).

En el caso de RIP-2, los campos que no se usan en RIP-1 se codifican de la siguiente manera:

- Dominio de ruta: Identificador de la ruta asociada con el mensaje. En UNIX, este campo es un proceso ID. Al emplear el dominio de ruta, una máquina puede manejar RIP múltiples y concurrentes.
- Etiqueta de ruta: Si es que RIP se emplea para dar soporte a EGP (exterior gateway protocol), este campo contiene, entonces, el número de Sistema Autónomo involucrado.
- Máscara de red: Asociada con la dirección IP en el mensaje
- Dirección IP del siguiente salto: la dirección IP, que indica dónde debe enviarse el paquete de datos.

2.3.4 Autenticación del Protocolo RIP-2

La versión RIP-2 proporciona autenticación, la versión 1 no. Cada paquete RIP se certifica con el receptor; para que esto suceda; la interface que corresponda debe configurarse para dar este soporte. Es posible proveer autenticación MD5, aunque los ruteadores puedan proporcionar otras opciones. El paquete RIP-2 para la autenticación es mostrado en la ilustración 2.7. El campo de la familia de direcciones se fija en 0xFFFF para la autenticación del paquete. El campo del tipo de autenticación se fija en 2, para un procedimiento de autenticación de texto plano y en 3, para el procedimiento MD5 (en texto plano no esta codificado).

Los bytes de información de la autenticación contienen el identificador que es un número clave y pueden haber múltiples números en una serie clave. Asimismo, estos bytes pueden contener campos que definan la duración de la clave o de las claves. Cada identificador clave del paquete se encuentra asociado con una clave almacenada localmente. El identificador clave y una interface asociada con el mensaje identifican al algoritmo de autenticación, así por ejemplo la clave de autenticación MD5 específica que se emplea para la operación de autenticación. Las versiones RIP-1 y RIP-2 pueden operar solas o combinándose entre ellas.

2.3.5 Problemas de la Convergencia

Las actualizaciones RIP que hasta el momento hemos descrito, son bastantes simples. Sin embargo, esta simplicidad puede ser motivo de problemas. Un problema es la posibilidad de enviar tráfico por medio de una ruta inexacta; el otro es la posibilidad de una actualización de la ruta a la que tome mucho tiempo lograr una convergencia, tiempo en el cual el Dominio de rutas es inestable y el tráfico pasa de manera deficiente e incorrecta, dando lugar a pérdida de paquetes en el "*ciberespacio*".

Para el caso de que se generen lazos cerrados, el problema continuará hasta que la métrica llegue a 16, momento en el cual los ruteadores volverán a evaluar la situación y a converger.

Supongamos que estamos ante el caso en que no se pueda acceder a una red por un problema fisico; y se genera un lazo cerrado, durante el conteo de saltos de esta red inaccesible, las comunicaciones se siguen transmitiendo entre los ruteadores, hasta que el valor máximo de la métrica que se alcance sea 16. Es este momento, en realidad en que no se puede llegar a la red destino con métrica 16. El proceso se

torna bastante deficiente si existe gran número de ruteadores; conduce además a una congestión de enlaces y a la posibilidad de tráfico perdido, inclusive de los mismos mensajes RIP.

2.3.6 Medidas Para Acelerar la Convergencia

La mayoría de las implementaciones RIP tienen medidas para contrarrestar las causas de problemas de convergencia. Una forma es, simplemente, evitar enviar propagación cada 30 segundos y solo enviar actualizaciones, cuando se tiene una modificación de ruta que actualizar, luego de producido el cambio. Por supuesto que esta actualización inmediata no resuelve el problema, pero si agiliza el tiempo para lograr una convergencia.

El Método Horizonte Dividido (Split Horizon)

Otro método que ayuda a reducir el tiempo de convergencia, se denomina, *"split horizon"*. La idea detrás de esta operación se sustenta en un concepto lógico de sentido común que propone que un ruteador no debería propagar nuevamente las direcciones a través de la interface desde la cual se recibió la misma dirección; pues naturalmente el equipo que las envió ya las conoce.

Esta postura es efectiva en la mayoría de las situaciones, pero no eliminará los problemas. Por ejemplo, si la red física es una topología que presenta lazos cerrados, el problema del conteo hasta infinito persiste (métrica 16).

La Variación *"Split Horizon with Poison Reverse"*

Una variación del *"split horizon"* es el *"split horizon with poison reverse"*. En lugar de no enviar las informaciones acerca de su interface 1, éstos envían la información; pero, con métrica 16. Si sólo dos ruteadores están comprendidos, *"split horizon with poison reverse"* elimina los bucles. Puede que no sea deseable emplear

un *"split horizon"* en enlaces en serie (enlaces sin difusión), como X.25, frame relay y el modo de transferencia asíncrono. Si es necesario, los ruteadores pueden configurarse para deshabilitar *"split horizon"*.

Ajustes del Tiempo de Actualización

La mayoría de los ruteadores que tienen múltiples puntos de destino pueden configurarse para variar la frecuencia de las actualizaciones de ruta RIP y otros parámetros. Estos parámetros de configuración RIP se encuentran disponibles:

- Intervalo de tiempo entre las actualizaciones (*update*): El valor por defecto para las actualizaciones es de 30 segundos.
- Intervalo de tiempo después del cual la ruta se declara no válida (*invalid*): Este intervalo debería ser el triple del valor del *update*, lo que significa que una ruta se convierte en invalida si la ruta no es "revitalizada" con las actualizaciones. Por consiguiente, esta ruta inaccesible tiene participación en el holddown. También puede utilizarse para el envío de paquetes. El valor por defecto para el invalid es de 180 segundos.
- Intervalo de tiempo durante el cual se suprime la información sobre mejores rutas (holddown): Este valor debería ser el triple del valor del update. Cuando el *holddown* caduca, se aceptan las rutas comunicadas por otras fuentes y la ruta ya no es inaccesible. El valor por defecto del holddown es de 180 segundos.
- Cantidad de tiempo en el que un registro de ruta debe permanecer en la tabla de ruta antes de que se le retire de éste (flush).

  Este intervalo de tiempo no debe ser mayor a la suma de los intervalos de tiempo invalid más el holddown. El valor por defecto del flush es de 240 segundos.

2.3.7 Filtros

A fin de fijar un control sobre la forma cómo se propaga la información de ruta a través del ruteador y desde las interfaces, puede ser deseable contar con un filtro de paquete de ruta. El ruteador puede configurarse para proporcionar los siguientes filtros RIP:

- Evita las actualizaciones de la ruta mediante una interface: Esta operación evita que otros ruteadores en una LAN sepan de manera dinámica acerca de las rutas.
- Controla la comunicación de rutas en actualizaciones de ruta: Esta operación permite que el administrador de la red evite que las rutas sean comunicadas en las actualizaciones del RIP.
- Controla el proceso de actualización de rutas: Suponiendo que las rutas sean comunicadas, esta operación no permitirá que la ruta "descubierta" sea procesada.
- Filtración de fuentes de información de ruta: Puede darse el caso de que se encuentre una ruta más confiable por medio de otro protocolo de ruta; es decir, algunos protocolos de ruta producen informaciones más exactas que otros. Es posible configurar un ruteador para priorizar la información de ruta proveniente de diversas fuentes. Esta función permite que el administrador de la red seleccione un valor (una distancia de administración) para una interface, direccionando así el ruteador para seleccionar la ruta desde el protocolo de ruta que tenga una menor distancia de administración

2.4 Protocolo OSPF

Protocolo Abierto Primero el Camino Mas Corto OSPF

El protocolo OSPF (Open Short Path First) es de la clase IGP y opera en Sistemas

Autónomos. Basa su funcionamiento en el algoritmo SPF o Dijkstra que fue desarrollado en la década del 50, su eficiencia es mayor que la de RIP en cuestiones relativas a uso de ancho de banda; sobre todo si el número de ruteadores que forman la red es elevado. El protocolo está especialmente diseñado para una red interna e incluye capacidades de direccionamiento de subred y de tipo de servicio (TOS).

El OSPF basa su decisión para descubrir una ruta en las direcciones y en la métrica de estado de enlace. El OSPF es un protocolo adaptable que se ajusta a los problemas típicos en una red (una falla en el nodo o en los enlaces, por ejemplo) y proporciona períodos menores de convergencia para estabilizar las tablas de ruta. Asimismo ha sido diseñado para evitar lazos cerrados de tráfico,

El OSPF a diferencia de RIP, usa directamente el protocolo IP; se encuentra encapsulado en el campo de datos del datagrama IP. El identificador del OSPF sobre el protocolo IP es 89. La figura 2.8 nos muestra un Sistema Autónomo con OSPF.

2.4.1 Funciones de un Ruteador en el Entorno OSPF

El OSPF permite que un ruteador asuma diversos papeles en un dominio con OSPF. Este puede actuar como un ruteador designado para un Sistema Autónomo, como un ruteador designado para un área dentro de un Sistema Autónomo y como un ruteador designado para una red a la que se le anexan otros ruteadores.

Dentro de estos dominios, el ruteador puede enviar y recibir diferentes tipos de LSA (Link State Advertisement). Algunos de éstos son una especie de "hola" entre los ruteadores, tales como un paquete "hello".

Otros LSA contienen información acerca de la base de datos de un nodo, otros son paquetes de actualización.

2.4.2 Características Principales del OSPF

Dada la importancia de la seguridad en las operaciones de redes de comunicación, OSPF provee estos mecanismos. La seguridad se aplica a los mensajes de OSPF haciendo uso de técnicas tales como MD5.

OSPF define una mejor ruta por medio de la métrica, lo que puede traducirse en capacidad, distancia, retardo, seguridad, etc. El OSPF no pone límites al costo total de la ruta. El valor de la métrica de por sí puede oscilar entre 1 a 65,535. En la práctica, las implementaciones comerciales del OSPF calculan la métrica basada en el ancho de banda del enlace en una interface del ruteador.

Una de las principales preocupaciones de cualquier nuevo diseño de protocolo de ruta es su escalabilidad; es decir, su capacidad para soportar el incremento en el número de redes y proveer la información de enrutamiento entre ellos de forma eficiente. OSPF desarrollo el esquema de jerarquías para facilitar la escalabilidad en las redes.

Existe el problema de las rutas múltiples para un destino. Los antiguos protocolos no resuelven adecuadamente esta situación. OSPF, brinda soluciones a estas operaciones críticas. Este proporciona herramientas para que el administrador elija entre múltiples rutas y se puede obtener además una eficaz distribución de la carga entre los distintos enlaces disponibles hacia un mismo destino. Los enlaces múltiples a un mismo destino eran poco comunes; sin embargo; en la actualidad los requerimientos de elevada disponibilidad los hacen necesarios.

Grafos de Red

El OSPF trabaja en base al concepto de grafos de red. Los grafos contienen valores entre dos puntos y las interfaces entre dos ruteadores (sus interfaces de

enlace). Los valores representan el valor de la ruta más corta ponderada, teniendo al mismo ruteador como punto de referencia. En consecuencia, el árbol del camino más corto, a cualquier punto en una red interna se determina por el ruteador que realiza el cálculo, el cual sólo revela el siguiente salto hacia el destino. La base de datos de estado de enlace, empleado en estos cálculos, se deriva de la información obtenida por las comunicaciones periódicas enviadas por los ruteadores a sus vecinos a lo largo del dominio de ruta.

La información se centra en la topología de la red o de las redes por medio de un grafo dirigido. Los ruteadores y las redes forman los nodos del grafo. Esta información puede difundirse periódicamente a todos los ruteadores del Sistema Autónomo. Un ruteador OSFP calcula la ruta más corta hacia los otros ruteadores en el dominio de ruta partiendo de la base de datos de estado de enlace. Si los cálculos muestran que las dos rutas tienen el mismo valor ponderado, el OSPF distribuirá el tráfico en esas rutas proporcionalmente. En teoría, las métricas de costo individual pueden calcularse para ambas direcciones de un enlace. Sin embargo, la mayoría de las implementaciones simplemente emplean el mismo valor para ambas direcciones del enlace.

Determinación del Enrutamiento

En principio, el OSPF sustenta sus decisiones en cuanto a rutas en dos campos en el datagrama IP: La dirección IP de destino y el tipo de servicio (TOS). El campo del tipo de servicio contiene diversos valores.

Los valores de campo que el OSPF puede utilizar para una ruta son: El retardo, el caudal, la fiabilidad y el costo, entre otros Mientras que el IP y el OSPF proporcionan reglas sobre cómo estos campos codifican, así como definiciones

generales sobre su significado, ningún protocolo norma cómo se traducen dichos valores a las operaciones de red.

Desde su implementación; en la práctica, el valor TOS no fue muy utilizado y terminó en desuso. En la actualidad, se usa otros medios en lugar de TOS, para asignar métricas a los enlaces y a las redes.

2.4.3 OSPF en Diversos Tipos de Redes

El OSPF opera en redes de difusión y de no-difusión. Ejemplos de redes de difusión son Ethernet, Token Ring y FDDI. Ejemplos de redes de no-difusión son X.25, Frame Relay y el modo de transferencia asíncrono ATM. El OSPF también opera en enlaces de punto a punto en los que no existen conjuntos de redes. Ejemplos de enlaces punto a punto son aquéllos que operan sobre PPP y HDLC.

OSPF brinda soporte a los enlaces en demanda tales como:

Las líneas de discado, las conexiones ISDN y las operaciones X.25 de llamada virtual conmutadas, el Frame Relay y el ATM crean un entorno en demanda que puede ser utilizado bajo OSPF. La idea base, es que el OSPF suprima algunos de los tráficos del paquete de comunicación y de contacto entre los ruteadores que se conectan en el enlace de solicitud, lo que permite que el enlace de solicitud esté inactivo; es decir, la capa del enlace de datos no este operando.

Cuando el enlace se activa, el OSPF reasume su comportamiento habitual y enviando los paquetes *"hello"* y la información de estado de enlace, de tal manera que se completa el intercambio de información necesario.

2.4.4 Operaciones Básicas del OSPF

Las operaciones OSPF varían dependiendo del tipo de red en la que operan. Su forma de proceder es ligeramente diferente si es que está en una red de punto a

punto, en una red de difusión o en una red de difusión múltiple. Pasaremos a revisar las operaciones del OSPF comunes a todos los tipos de redes.

El OSPF implementa un protocolo *"hello"*, el cual es un protocolo de contacto que permite que los ruteadores se reconozcan, que conozcan la información de intercambio y que, posteriormente, los ruteadores vecinos produzcan una trama de prueba para cerciorarse de que el enlace o el ruteador está en funcionamiento.

Después de que las operaciones *"hello"* se hayan completado, se considera que los ruteadores son adyacentes.

Como siguiente paso, los ruteadores intercambian información que describe el dominio de enrutamiento. A esta información se le denomina una descripción de base de datos y se la coloca en mensajes de información sobre el estado del enlace (LSA). Las descripciones de la base de datos; que envía cada nodo; no son toda la base de datos de estado de enlace, pero contienen suficiente información para el ruteador de recepción, que le permite saber si su base de datos de estado del enlace está conforme con las bases de datos de sus pares. Si todo está en orden, el vecino se definirá como completamente adyacente.

De aquí en adelante, *"hello"* periódicos se emiten para hacer que los puntos sigan percatándose el uno del otro. Igualmente, los LSA que el ruteador originó deben enviarse a sus puntos de destino cada cierto tiempo, sólo para asegurarse de que todas las bases de datos de estado de enlace estén sincronizadas.

Propagación de Información

La comunicación de la información se hace dentro del dominio de ruta, "inundando" de paquetes LSA a todos los ruteadores. A primera vista, podría parecer que el desbordamiento crearía una gran cantidad de carga de tráfico en el

dominio de ruta. De hecho, el desbordamiento puede dar como resultado que un nodo reciba más de una copia de un paquete, pero el OSPF establece diversas reglas para disminuir los excesos de copias de paquetes OSPF que se propagan en el dominio de ruta.

Algunas reglas para reducir el tráfico son las siguientes: Primero, un LSA sólo puede actualizarse (como máximo) cada 5 segundos. Segundo, a un ruteador no se le permite aceptar un LSA si la copia de la base de datos es de una antigüedad menor a la de un segundo; de manera que si un ruteador se configura de manera equivocada, y si su interface sube y baja sucesivamente o, dicho de otro modo, experimenta problemas, sus vecinos no procesarán el torrente de tráfico, tampoco enviarán estos LSA a otros ruteadores.

2.4.5 Las Áreas del OSPF

Las empresas con sistemas grandes pueden operar con muchas redes, ruteadores y servidores. Con la finalidad de administrar este gran número de componentes de redes de comunicación, es probable que muchos paquetes LSA tengan que ser intercambiados entre los ruteadores a fin de determinar cómo transmitir el tráfico dentro del Sistema Autónomo entre la computadora remitente y la receptora.

El administrador de la red debe evaluar cuánto tráfico tiene que enviarse entre los ruteadores, ya que el tráfico de ruta puede afectar el caudal de data del usuario, puesto que es práctica común, transmitir una gran cantidad de paquetes de informaciones de enrutamiento a todos los nodos en el dominio de ruta. Al mismo tiempo que se tiene que tomar medidas para reducir la cantidad de tráfico duplicado que un nodo recibe, no es raro que la red tenga sus nodos de ruta conectados de una forma que se puedan generar bucles, lo que quiere decir que es posible que la

información de ruta se reciba más de una vez. El OSPF permite contención de paquetes y provee control anti-bucles de tráfico.

El OSPF divide el Sistema Autónomo en partes más pequeñas (figura 2.9) llamadas áreas. Esta división reduce la cantidad de tráfico de ruta que se envía a través del Sistema Autónomo, ya que las áreas se aíslan unas de otras. El establecimiento de Areas disminuye la cantidad de información que un ruteador debe guardar; en lugar de tener información acerca de todo el Sistema Autónomo. Asimismo, esto implica que el costo de la información transmitida para mantener las tablas de ruta del OSPF se ve sustancialmente reducida.

Un ruteador designado, asume la responsabilidad de informar a los ruteadores del área acerca de las otras rutas, redes y ordenadores que residen en el Sistema Autónomo.

Contención de Paquetes

Como consecuencia de esta disposición, los ruteadores dentro del área no se preocupan por los detalles de todo el Sistema Autónomo (las otras áreas) y obtienen información a partir de un ruteador determinado, llamado ruteador designado suponiendo una red de difusión, como Ethernet, el OSPF emplea una difusión múltiple para restringir el procesamiento del paquete LSA en los nodos que no requieran examinar todos los paquetes En el caso en que la red sea de no-difusión (una red conmutada, como el Frame Relay), el OSPF emplea un procedimiento de filtración del paquete para disminuir el número de paquetes de rutas que se intercambian entre los ruteadores del área.

Áreas De Tope

El OSPF también brinda soporte al concepto de áreas de tope que es un área a la

cual no se envía la información sobre rutas externas. En vez de esto, el ruteador de borde del Area genera una ruta por defecto para los destinos fuera del Area Tope.

Enlaces Virtuales y Columnas Vertebrales

El OSPF requiere que todas las áreas se conecten juntas a través de un área de columna vertebral o de núcleo (denominada área 0). Por lo general, una columna vertebral se considera como un conjunto de nodos contiguos y sus enlaces de interconexión permiten que las redes secundarias (o áreas) se comuniquen con los demás por medio de la columna vertebral. La idea de una columna vertebral obliga a una topología simple en el que la información de ruta del área pasa por la columna vertebral para llegar a otra área.

Sin embargo, esta disposición podría ser muy restringida, ya que implica que las áreas anexadas a la columna vertebral deban ser áreas de tope. Para flexibilizar este comportamiento, el OSPF respalda el empleo de los enlaces virtuales; en los cuales los paquetes de ruta pueden enviarse por medio de un área a la otra; y sin tener que atravesar la columna vertebral. Los enlaces virtuales se movilizan entre los ruteadores y permiten que los paquetes de LSA pasen como por un túnel por las áreas. La ilustración 2.10 muestra un ejemplo de la columna vertebral y la disposición de los enlaces virtuales. Los enlaces virtuales permiten la distribución de la información de ruta a través de las áreas, bajo el concepto denominado "tunelización".

Las redes conectadas fuera de los Sistemas Autónomos del OSPF no son miembros de este Sistema Autónomo. Los ruteadores descubren estas redes mediante un protocolo EGP externo (como el BGP) luego comunican estas redes dentro del Sistema Autónomo, mediante LSA externos.

Designación de un Ruteador para una Red

El dividir un Sistema Autónomo entre áreas más manejables y más pequeñas, es importante en OSPF, pues reduce el tráfico relativo a enrutamiento, elevando la eficiencia para un Sistema Autónomo que se enfrenta con la interconexión de un gran conjunto de redes, ruteadores y servidores. A modo de ilustración, consideremos las redes 1 y 3 y el ruteador 1 de la figura 2.11. El ruteador 1 se conecta con dos redes, la red 1 y la red 3, una Ethernet y una Token Ring, respectivamente.

El OSPF requiere que un ruteador se asigne para una red. De este modo, los servidores y demás dispositivos no tienen que preocuparse por la información de rutas. En este ejemplo, el ruteador 1 asume la responsabilidad de realizar esta tarea para la red 1 y 3. De hecho, el papel del ruteador designado para esta red es el más indicado, ya que el ruteador 1 es el equipo que conecta a esta red. La designación del ruteador se logra por el empleo de paquetes "hello". En una subred con ruteadores múltiples se configura a cada ruteador con una prioridad para el parámetro de ruteador designado.

Este parámetro ayuda a determinar el ruteador designado para la subred. Si el ruteador designado falla, un ruteador suplente toma el control. Las redes de difusión, tales como Ethernet, permiten que todos los nodos de la red reciban un mensaje, con la dirección MAC codificadas todas con 1. El Ethernet también brinda soporte a la difusión múltiple, que permite que una misma trama Ethernet sea recibida por un subconjunto de nodos en la subred. El uso de un ruteador designado bajo OSPF quiere decir que los demás ruteadores no intercambian todos su mensajes entre sí, salvo al inicio en que se intercambia paquetes *"hello"*, y centraliza la responsabilidad de mantener la base de datos actualizada en el ruteador designado,

R2 y R3 mantienen su información actualizada y las bases de datos sincronizadas con R1, el ruteador designado.

Esta disposición es mucho más eficiente, pero significa que el sistema es más vulnerable a fallas; si se presentan problemas en R1 se afectará toda la red.

El OSFP emplea otra dirección de IP de difusión múltiple para identificar los ruteadores designados, incluyendo el ruteador de copia de seguridad. Esta es la dirección 224.0.0.6.

2.4.6 Los Paquetes OSPF

La ilustración 2.12 muestra la cabecera del paquete OSPF de 20 octetos. Cada paquete LSA del OSPF está anexado a esta cabecera. A continuación se hace una descripción general.

El campo Antigüedad, muestra el número de segundos a partir del origen del LSA. Comúnmente, éste va de 0 a 30 minutos y si la antigüedad excede a los 30 minutos, el ruteador de origen reenvía el LSA y fija este campo en 0 (también un incremento del número de la secuencia de LS).

El campo Opciones se emplea para indicar que el LSA debería procesarse de forma especial. Se han definido diversas opciones.

El campo Tipo, identifica el tipo de LSA. Actualmente, se han definido varios tipos de LSA. Las reglas para el uso de estos LSA son diversas y variadas.

El campo ID, se emplea para distinguir cada LSA del mismo tipo LS que se origina a través de un ruteador específico. En la práctica, por lo general contiene la información de una dirección. Por ejemplo, en una conexión de punto a punto, este campo es el ID del ruteador vecino, que por lo general es una dirección IP.

El campo Ruteador de Anuncio, contiene el ID del ruteador de origen.

El ruteador original del LSA aumenta el campo Número de Secuencia, en el momento que este ruteador desee actualizar el LSA. Así, un número de secuencia más largo en un LSA indica que es más reciente que un LSA con un número de secuencia menor.

La suma de verificación se aplica al receptor para revisar una cabecera de LSA, en caso de que haya sido alterada, así como para revisar las informaciones, par detectar errores. Asimismo, la suma de verificación se almacena en todas las bases de datos de estado de enlace de los ruteadores a fin de:

(a) Determinar si dos LSA con el mismo número de secuencia son, idénticos. (También se emplea el campo de la antigüedad para esta revisión),

(b) Determinar periódicamente si el hardware y el software del ruteador están alterando el registro del LSA en la base de datos.

El campo Longitud define la longitud de la cabecera y de los contenidos del LSA. Los contenidos de los campos del paquete OSPF que siguen a la cabecera varían, dependiendo del tipo de paquete LSA, del papel asignado al ruteador, si es que éste está, por ejemplo, en un enlace punto a punto o en una subred compartida.

La cabecera del LSA de 20 bytes precede a los campos de comunicación de la información. Estos campos se inician con el tipo de ruteador e identifica los ruteadores comunes, los ruteadores de límite, etc.

El campo Número de Enlaces especifica cuántos enlaces del ruteador de origen se vienen reportando.

El valor ID del Enlace varía. Para una conexión de punto a punto, el campo contiene el ID del ruteador del vecino. El campo Informaciones del Enlace también depende del tipo de comunicación de información. En el ejemplo, éste contiene el

número de interface del enlace del ruteador original, ifIndex que es una variable de la MIB de OSPF.

El campo Tipo de Enlace identifica el tipo de enlace, tales como uno de punto a punto, una red de tope, etc.

El campo Métrica del Tipo de Servicio no se emplea en Internet y se le omite de las especificaciones OSPF.

El ifIndex es el objeto de registro ID (OID) en hoja la del grupo de objeto de las interfaces en el MIB OSPF.

Los dos enlaces se configuran con una métrica de 5 para un enlace y 6 para el otro. Los campos LSA reflejan esta configuración y asocian la métrica con el valor ifIndex, así como con el enlace ID.

El enlace ID para una comunicación de información punto a punto es el identificador del nodo par.

2.4.7 Parámetros de Configuración OSPF

Cada ruteador que está configurado para OSPF puede adaptarse para operar con un conjunto de parámetros diferente. Estos son los siguientes:

- Costo OSPF: La métrica en la interface que especifica el costo del envío de un paquete en la interface.
- Intervalo de retransmisión: El intervalo de retransmisión indica el número de segundos para la retransmisión de paquetes LSA entre los nodos adyacentes a la interface.
- Retraso de transmisión: Cálculo del retraso (en segundos) para enviar una actualización de LSA en la interface.
- Prioridad: El número de prioridad para el ruteador, empleado para seleccionar el

ruteador designado para la red.

- El intervalo *"hello"*: El tiempo que transcurre, en segundos, entre el envío de los paquetes *"hello"*.
- El intervalo muerto: El tiempo, en segundos, que un vecino no ha enviado un "hello", antes de que se declare que el vecino está fuera de servicio.

  Clave de autenticación: La contraseña que los vecinos emplean entre ellos.
- La autenticación MD5: Establece el uso del MD5 para el proceso de autenticación

VECTOR DISTANCIA	ESTADO DE ENLACE
MINIMO NUMERO DE SALTOS	MINIMO VALOR DE METRICA
PUBLICA TABLA DE RUTAS	PUBLICA ESTADO DE ENLACE
CALCULO DE RUTA DISTRIBUIDO	CALCULO DE RUTA NO DISTRIBUIDO
CONVERGENCIA LENTA	CONVERGENCIA RAPIDA

FIG 2.1 Comparación de Protocolos

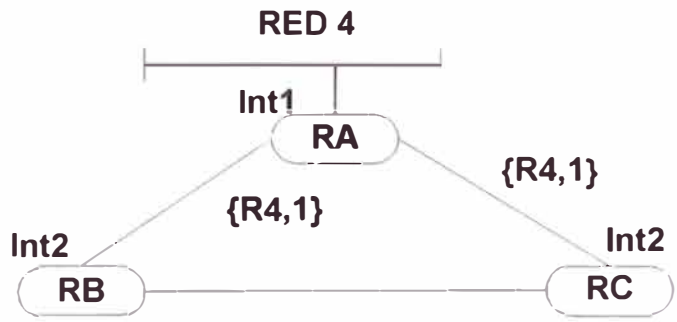


FIGURA A

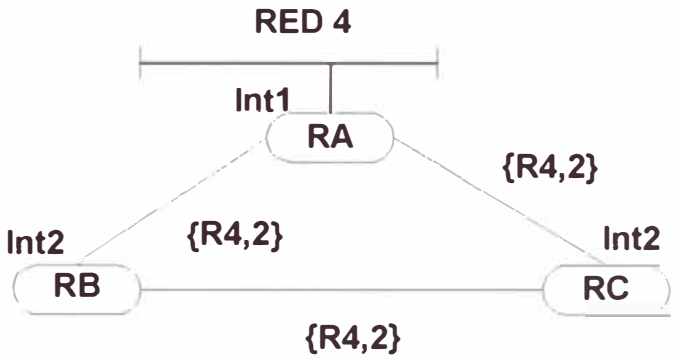


FIGURA B

RUTEADOR	SALTO SIGUIENTE	METRICA	CONEXION	RUTA	INTERFACE
R A	-	1	DIRECTA	LOCAL	INT 1
R B	RA	2	REMOTA	RIP	INT 2
R C	RA	2	REMOTA	RIP	INT 2

FIGURA C RUTAS PARA LA RED 4

FIG 2.2 Rutas en Red Bajo RIP

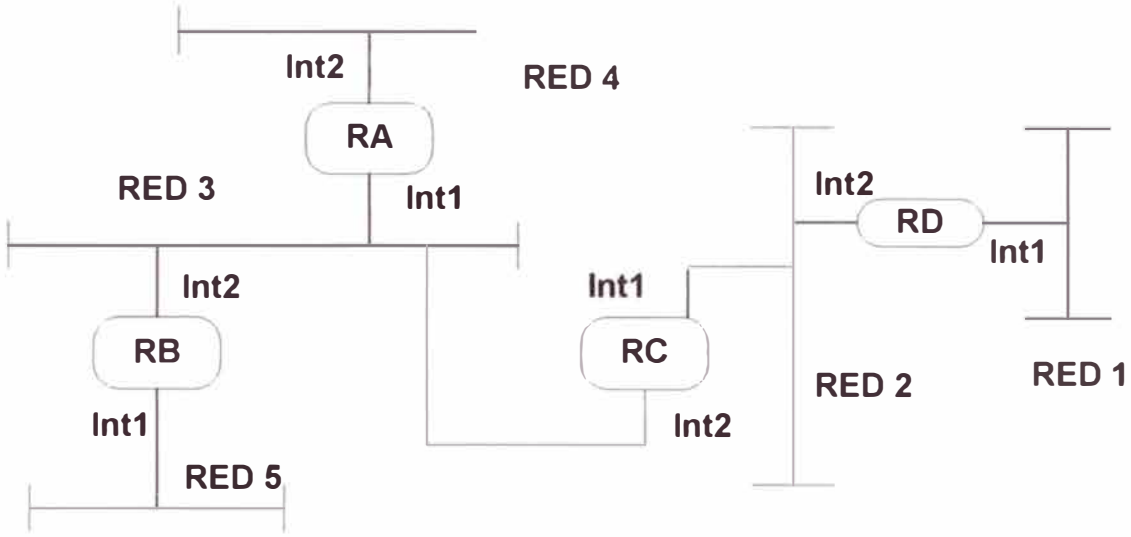


Fig 2.3 Aplicación RIP

RED DESTINO	SALTO SIGUIENTE	METRICA	CONEXION	RUTA	INTERFACE
RED 5	-	1	DIRECTA	LOCAL	INT 1
RED 4	RA	2	REMOTA	RIP	INT 2
RED 3	-	1	DIRECTA	LOCAL	INT 2
RED 2	RC	2	REMOTA	RIP	INT 2
RED 1	RC	3	REMOTA	RIP	INT 2

Fig 2.4 Tabla de Rutas en RB

Fig 2.5 Trama RIP-1

FIG 2.6 Trama RIP-2

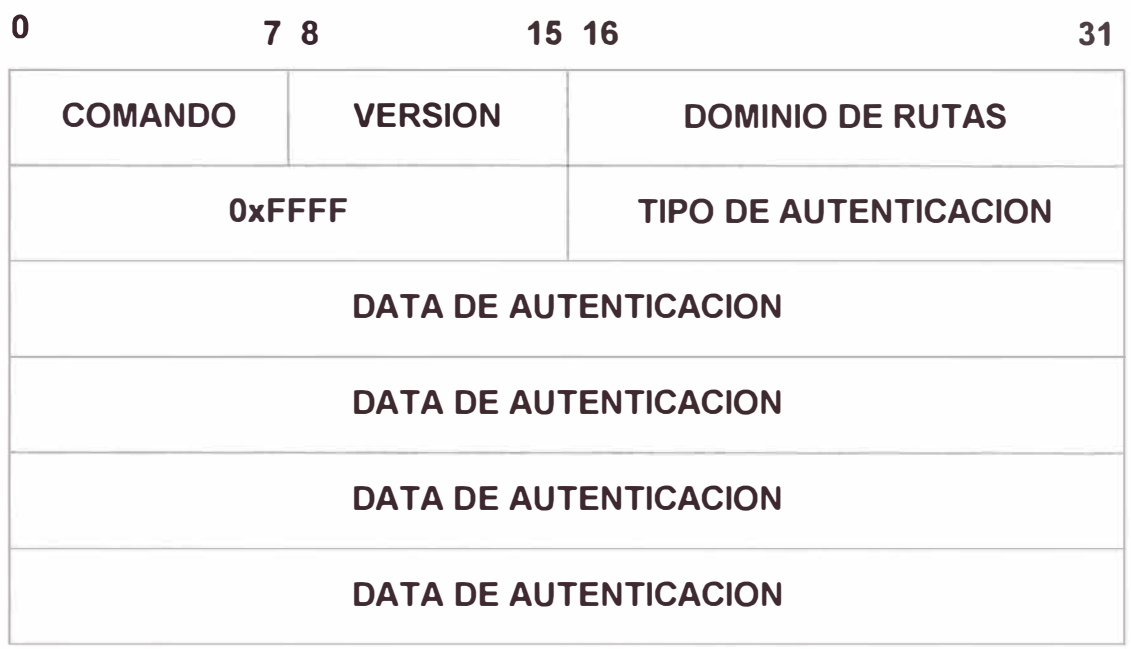


FIG 2.7 Trama de Autenticación RIP-2

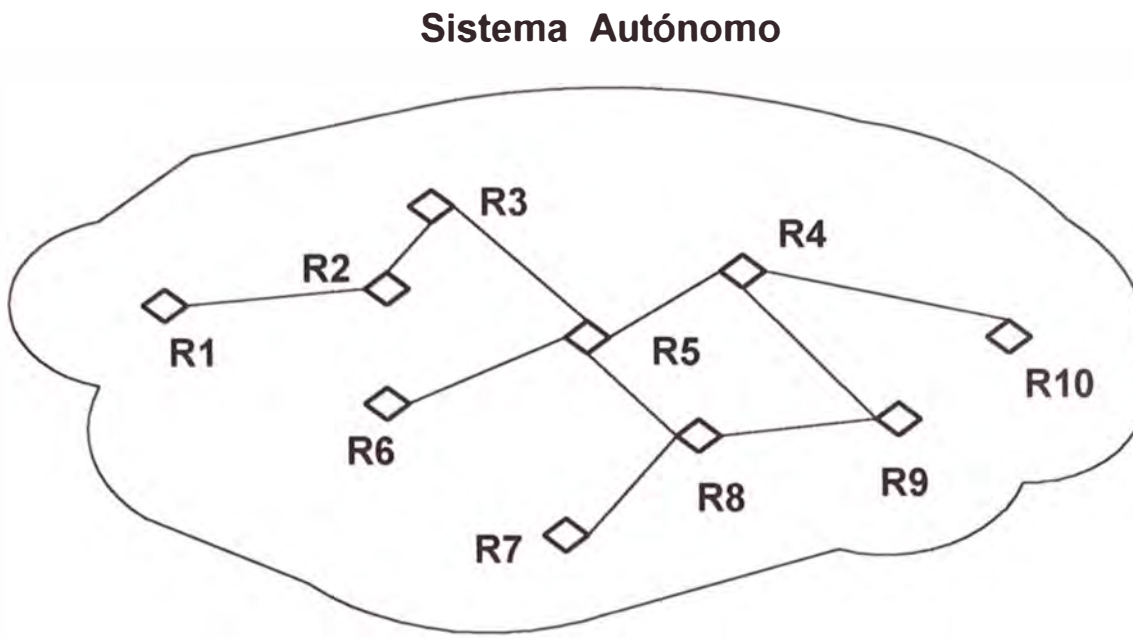


Fig 2.8 SA en OSPF

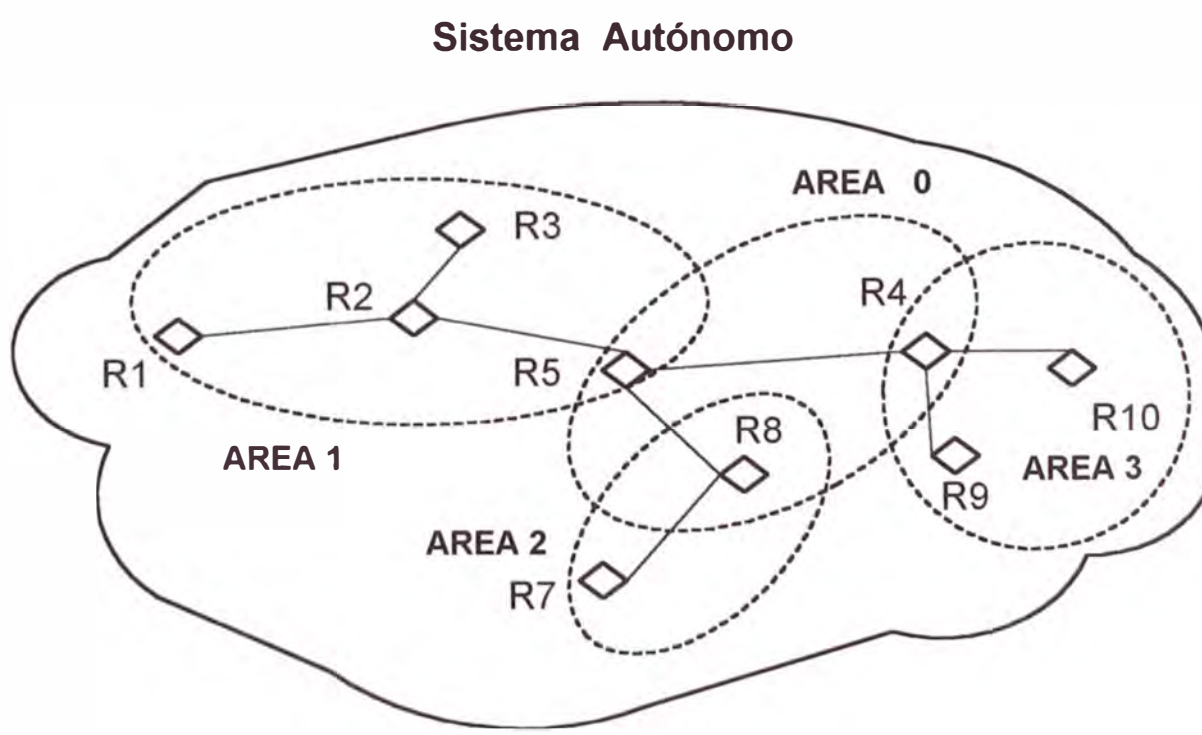


FIG 2.9 División de SA en Areas

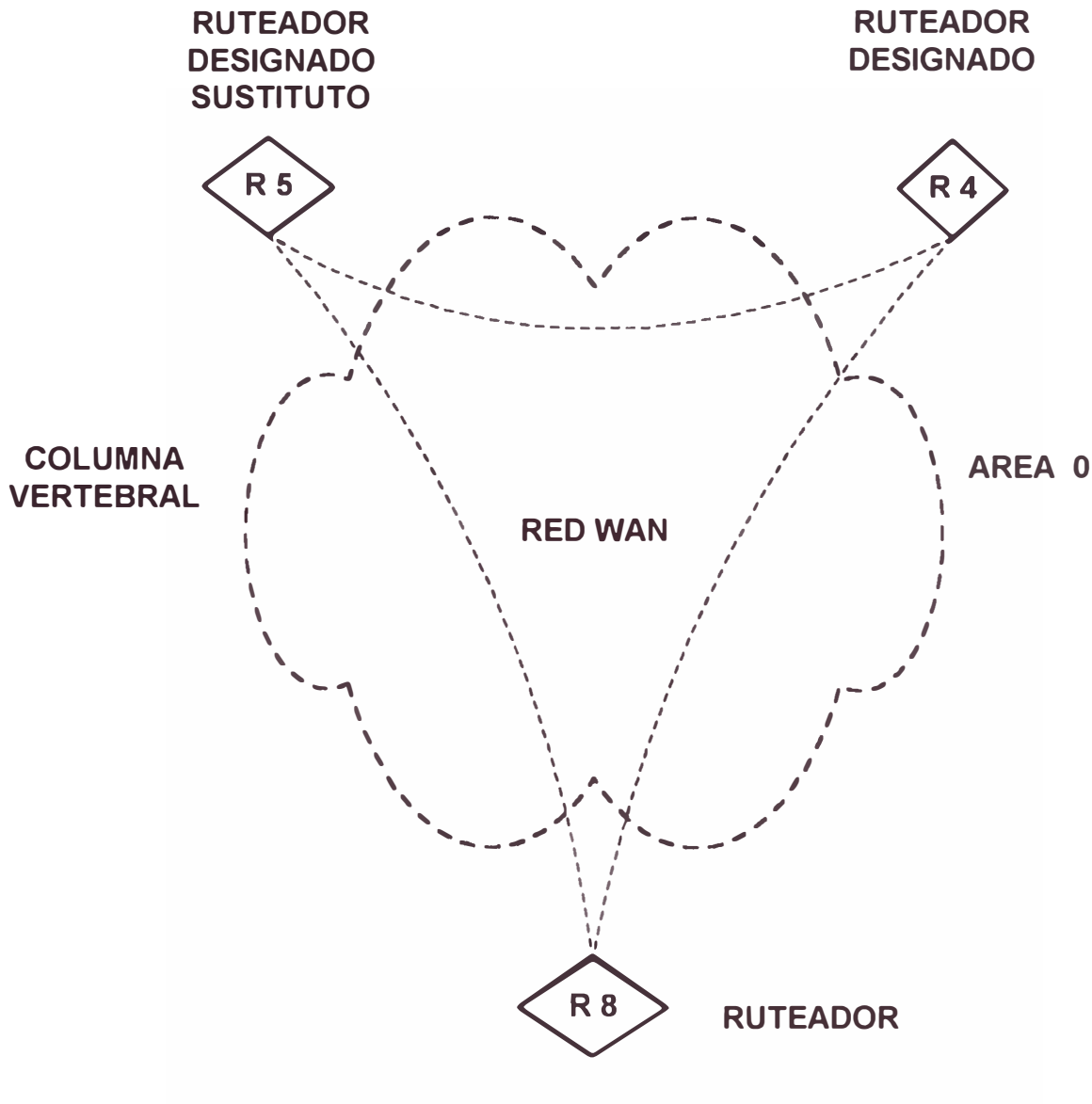


FIG. 2.10 Columna Vertebral Area 0

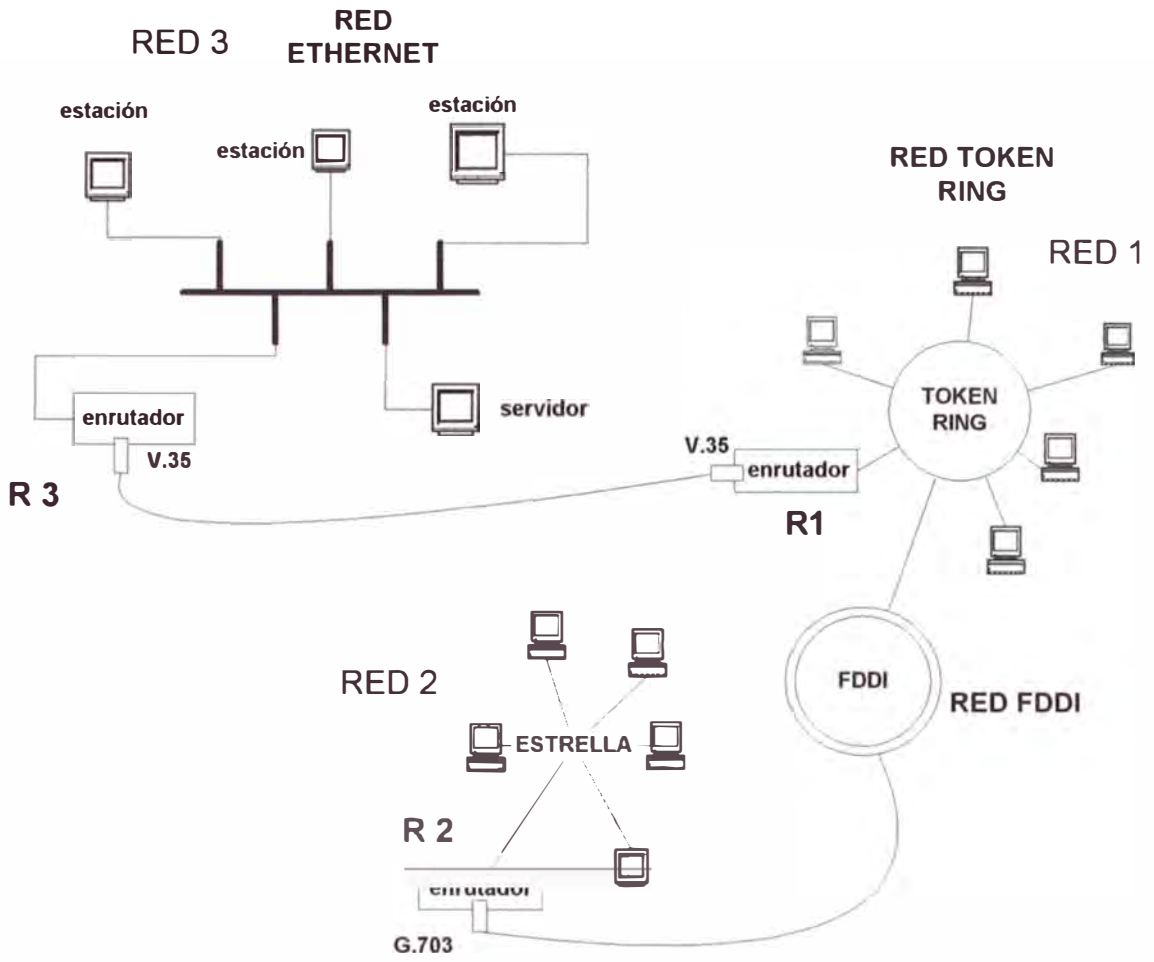


Fig 2.11 Ruteador Designado en OSPF

VERSION	TIPO	TAMAÑO DE PAQUETE
IDENTIFICADOR DE ROUTER		
AREA		
SUMA DE COMPROBACION		AUTORIZACION
AUTENTICACION		
AUTENTICACION		

Fig 2.12 Trama OSPF

CAPITULO III

PROTOCOLOS DE INTERCONEXIÓN DE SISTEMAS AUTÓNOMOS

En los capítulos anteriores hemos tratado los aspectos relacionados, a la interconexión física de las redes de datos antes de abordar los tópicos referentes a la interconexión IP; posteriormente, tratando los aspectos principales del protocolo IP, para luego pasar al estudio de los protocolos de enrutamiento, sus clases, y por último, estudiar dos de los protocolos más representativos de los llamados Protocolos Internos a un Sistema Autónomo IGP, RIP y OSPF. Es momento ahora de enfocar el grupo de los protocolos de enrutamiento exteriores a un Sistema Autónomo EGP.

3.1 Protocolos de Encaminamiento Externo

Al principio de la edad de Internet el número de dispositivos a conectar era ínfimo, si lo comparamos en la actualidad, con los miles de nodos conectados; estamos hablando de decenas o centenas de anfitriones o *"host"*, las cuales estaban agrupados en redes cuya administración era independiente una de otra. Para unir físicamente estas redes, se hizo uso de circuitos de comunicación arrendados a las empresas de telefonía. Estos circuitos dedicados pasaron a constituir la Columna

Vertebral o Backbone de esta inter-red o red de redes, la cual pasó a denominarse comúnmente como Internet; y esto fue así, para diferenciarse de cualquier otra *"internetwork"* que pudiera existir. Recordemos que el termino internet deviene de la naturaleza de ser un sistema de internetworking, es decir de interconexión de redes disímiles; por fuerza de la costumbre el término internetwork se abrevió en la palabra Internet.

Debido a que el funcionamiento de los protocolos IGP, era sumamente rígido, y en muchos casos ineficiente, para inteconectar redes independientes; debido a que no estaban diseñados específicamente para interconectar redes de Sistemas Autónomos entre sí; se optó por el desarrollo de un sistema de protocolo de enrutamiento enfocado a interconectar Sistemas Autónomos; a estos protocolos; se les llamo genéricamente protocolos exteriores EGP (figuras 3.1,3.2). Dos de ellos son el EGP y el BGP; este último es el estándar de facto actual y EGP fue uno de los primeros en emplearse en Internet, actualmente está en desuso.

3.2 Protocolo EGP

El precursor mas conocido de estos protocolos, es el EGP (Exterior Gateway Protocol). Este protocolo fue desarrollado para afrontar la interconexión de un Sistema Autónomo a la columna vertebral de Internet. Su formato (figura 3.3 y 3.4) presenta varios tipos de mensaje entre estos tenemos:

- Solicitud de reconocimiento
- Confirmación de reconocimiento
- Confirmación de estado activo
- Actualización de rutas.

La solicitud de reconocimiento se utiliza cuando se establece al inicio una conexión

entre vecinos EGP, esperándose luego la confirmación de la misma. El mensaje de estado activo confirma la operación normal de un vecino, después de que ambos se han reconocido como vecinos, y el mensaje de actualización de rutas permite el anuncio de redes accesibles.

En la trama el campo versión indica la versión de EGP que se está usando, y el tipo se usa para indicar la clase de mensaje. La información que alcanza EGP a sus vecinos es bastante simple; en resumen, solo anuncia las redes que están accesibles.

Estas prestaciones eran apropiadas para el entorno de su aparición. Sin embargo con el auge de Internet, rápidamente sus limitaciones dieron inicio a la búsqueda de otras soluciones de tipo EGP para la expansiva Internet. Los principales inconvenientes que presenta el EGP son los siguientes:

- EGP solo anuncia una ruta para una red, aun si existieran varias rutas disponibles.
- No reconoce la métrica asociada a la red.
- No es posible el balance d la carga a través de varias conexiones.
- En caso de fallas de enlace la reconfiguración debe hacerse manualmente. Esto se torna mas grave si el ruteador intermedio a otras redes sufre problemas de conexión con una tercera red.

EGP trabaja sobre paquetes IP y se sirve de ICMP para mensajes de estado y no usa ningún algoritmo para calcular rutas, solo anuncia que redes son alcanzables.

La necesidad de mejorar las prestaciones que provee EGP dio origen nuevos desarrollos sobre nuevos protocolos exteriores para Gateways, siendo BGP producto de estas investigaciones.

3.3 Protocolo BGP

BGP hizo su aparición a finales de la década de los 80 en entornos de prueba. Es completamente diferente a EGP; sin embargo, incorpora nuevas prestaciones derivadas de la experiencia acumulada en el empleo de su predecesor (figura 3.5).

Las principales prestaciones del BGP son las siguientes:

- BGP puede operar con redes utilizando topologías con lazos cerrados.
- BGP es capaz de elegir rutas entre varias disponibles; para un mismo destino
- BGP proporciona soporte para CIDR y rutas agregadas.
- BGP soporta rutas alternativas en caso de fallas de enlace, de forma automática; es decir sin intervención humana.
- BGP soporta políticas de enrutamiento para las decisiones en la determinación de rutas, y no solo se basa en las métricas.

La conectividad inicial a nivel de BGP se establece por configuración manual y no de manera automática. Esto permite control total sobre los acuerdos de interconexión, antes durante y después de establecidos estos. La función principal de BGP es intercambiar información respecto a la accesibilidad de una red, esta información esta referida a que Sistema Autónomo debe atravesar el tráfico para alcanzar la red destino, en la forma de un vector (camino,red); con este dato BGP construye un grafo sobre la topología de toda la intered. Como es posible deducir de lo anterior, BGP hace uso de técnicas de algoritmos distancia-vector y estado de enlace, por lo que podría denominársele como un protocolo de tipo híbrido. Otra característica muy importante es que BGP solo propaga los cambios de rutas de forma incremental, reduciendo de esta manera al máximo, el consumo de ancho de banda para alcanzar la convergencia entre los nodos ruteadores BGP. Para garantizar

todo el proceso de comunicación; se recurre a un protocolo de clase confiable como lo es el TCP, destinándose el puerto 179; de esta manera, es posible garantizar el envío y recepción de los paquetes BGP, liberándose al mismo de esta tarea, esto último es esencial para el buen funcionamiento de BGP. Otra característica muy importante es que la política de anuncio de rutas puede controlarse no solo a través del vector (ruta-SA, red destino), si no también a través de políticas determinadas por al administrador y que se configuran de manera manual. También es importante, la facilidad que brinda BGP para intercambiar la información sin requerirse necesariamente de un ruteador para el intercambio de esta entre Sistemas Autónomos. Es decir; se puede tener un nodo que solo cumpla las funciones de BGP, mas no estar encargado de transportar el tráfico de la red.

El BGP opera creando un grafo de SA. El grafo proviene de la información de ruta intercambiada por los ruteadores del BGP entre los SSAA comunicados mediante BGP. El BGP considera a toda la Internet como un gráfico de SSAA, cada uno identificados con un número SA; mientras que; por lo general, los Sistemas Autónomos están conectados entre sí tomando como base una relación de vecindad física. Un ruteador BGP puede configurarse para omitir ruteadores intermedios en el camino de rutas-SA, de manera que dos nodos vecinos BGP no son necesariamente vecinos; desde el punto de vista de físico. El BGP tiene diversas ventajas sobre un protocolo convencional de vector de distancia, como por ejemplo RIP. A continuación un resumen de ellas:

- El BGP envía mensajes sólo si algo cambia y no de manera continua. Obviamente este procedimiento disminuye el uso de ancho de banda.
- El BGP tiene la capacidad de seleccionar libremente la mejor ruta-SA de

enlaces, aun si el sistema presentara varios enlaces físicos.

- El BGP almacena copias de rutas alternativas en caso de que falle la primera ruta, se activa la ruta alternativa, convergiendo rápidamente las tablas de rutas de la red.
- La decisión sobre la ruta puede basarse en las consideraciones políticas y no sólo en el menor número de saltos. Un ruteador BGP establece una relación con otro ruteador BGP, mediante configuraciones manuales y no automáticas. Esto permite garantizar los acuerdos que se firman entre diversos ISP a través de configuraciones especificas de BGP.

3.3.1 Conceptos Generales en el Entorno BGP

Vecinos

Los protocolos de ruta necesitan saber acerca de quién tienen al frente y sobre la forma de cómo intercambiar información de rutas y mensajes de reconocimiento con los demás ruteadores. El BGP no es la excepción y debe, igualmente, considerar estos factores, aparte de número de saltos u otra métrica de enlaces. El BGP implementa soporte de rutas sustentadas en políticas autónomas, Por lo tanto, el saber quiénes tiene al frente en BGP es muy importante, debido a que bien puede tratarse de vecinos externos, es decir que pertenecen a otro Sistema Autónomo y administrado por otra entidad ajena.

Desde un punto de vista técnico, el BGP da soporte a dos tipos de vecinos: El vecino interno, que está en el mismo Sistema Autónomo y el externo, que está en un Sistema Autónomo diferente. Como regla general, los vecinos externos que están juntos uno al lado del otro comparten una misma red. Cabe resaltar que la condición de vecinos no la da la proximidad física; vecinos externos pueden o no estar

fisicamente enfrente, uno del otro o compartir una misma red; pueden localizarse en cualquier lugar dentro de un Sistema Autónomo.

Los Dialogantes del BGP (Speakers)

El BGP emplea el concepto de *"speaker"* para nombrar los dispositivos que se encargan de divulgar información de ruta a nivel o empleando BGP. El speaker puede ser o no un ruteador. Al emplear un conjunto de políticas comunes (acuerdos de rutas descritos anteriormente), los speakers del BGP concuerdan que ruteadores de la frontera del SA servirán como puntos de entrada/salida de tránsito para redes específicas fuera del SA. Esta información se comunica a los ruteadores internos del SA vía el protocolo IGP de rutas interno o mediante configuraciones manuales.

Se les denominan enlaces externos; a las conexiones entre los *"speakers"* BGP de Sistemas Autónomos distintos. Las conexiones BGP entre los *"speakers"* dentro del mismo SA se les denomina enlaces internos.

Comunidades

El BGP puede configurarse para distribuir información acerca de rutas de un grupo de destinos (redes) llamados comunidades. La idea es poder agrupar los destinos en estas comunidades y aplicar políticas de rutas para éstas. Este enfoque simplifica el trabajo del speaker al juntar información de rutas. Asimismo proporciona una herramienta para el administrador de la red que le permite controlar la propagación de la información de rutas.

La operación de la comunidad BGP proporciona una flexibilidad considerable en el manejo de rutas. Por ejemplo, el administrador de la red puede definir la comunidad a la que un destino pertenece. Dado que un destino pertenece o no a una comunidad, la política aplicada en BGP respaldará o no la distribución de la

información de ruta. La operación de las comunidades es un aspecto importante del ruteador del BGP. Nuevamente observamos cómo el BGP difiere de otros protocolos de rutas analizados hasta el momento: La política es la "métrica" fundamental para el BGP, mientras que la métrica que se basa en enlaces es fundamental para los protocolos de IGP, tales como RIP y OSPF.

3.3.2 Propagación de Información en el Entorno BGP

El BGP no elige la mejor ruta por salto mínimo. Veamos el siguiente ejemplo mostrado en la figura 3.5 y 3.6. El BGP descubre que la red 192.168.128.0/20 puede alcanzarse a través del SA1. El SA2 sabe que está en la ruta a la dirección destino anterior, de manera que en la difusión de esta ruta, rechazaría cualquier mensaje que no tuviera al SA2 como nodo intermediario.

Además, vamos a suponer que el enlace entre los ruteadores R4 y R2 falle. En esta situación, el SA4 conoce otra ruta al SA1 a través del SA5 y puede recuperarse de inmediato. No obstante, es posible que el SA5, debido a decisiones políticas, puede no permitir que se use esa ruta, a pesar de que existe conectividad física hasta el destino deseado.

Red Tope

Al igual que otros protocolos de rutas, el BGP emplea el concepto de una red "tope" que es una red por la cual no pasa tráfico externo. Este actúa como una fuente y como un vertedero para el tráfico (envía o recibe tráfico). En la ilustración 3.7, la red del cliente es un tope.

En la figura, el BGP se despliega entre el cliente y el proveedor de servicio (un ISP al que se le designa un número de SA). En este caso, el BGP recibe el nombre de BGP externo (EBGP). Se le denomina Así, porque el BGP puede operar entre los

SA o dentro de un SA. Para esta operación, el proveedor proporciona al cliente un número SA privado (65,412-65,535), pero emplea el número de SA del proveedor para sus operaciones más allá de la interface de interconexión.

Se puede desplegar soporte IGP entre el cliente y el proveedor de servicio. En este caso, el cliente usa RIP, OSPF, u otro protocolo para difundir sus direcciones al proveedor del servicio. Esta operación se establece mediante acuerdos negociados previamente entre ambas partes.

También es posible que el cliente utilice rutas estáticas para conectarse con el proveedor. Después de todo, la red del cliente es de tipo tope; pues sólo hay una vía para acceder desde dentro o fuera de éste.

Arquitectura Basada En Políticas del BGP

Como ya se ha mencionado, uno de los aspectos distintivos del BGP es su arquitectura basada en políticas. En esta parte se proporciona un resumen de la descripción de éste aspecto del BGP (figura 3.8).

El BGP proporciona la capacidad para hacer que se cumplan las políticas sustentadas en varias preferencias y restricciones, las cuales pueden ser consideraciones políticas, económicas y de seguridad. Las políticas no están incorporadas internamente en el protocolo. Más bien, las políticas para el BGP están dadas bajo la forma de directivas de configuración externa.

Las prestaciones del BGP tienen la capacidad de permitirle al administrador de la red (la administración del SA) que ejecute las tareas de "configuración política" cuando se va a crear una determinada política o cambiar una ya existente. Estos cambios afectan la selección de rutas en el ruteador, así como la redistribución de la ruta en el dominio del BGP.

El BGP puede controlar los siguientes aspectos del tráfico en un ruteador:

- BGP puede minimizar el número de SA de tránsito (Puede preferirse rutas de SA más pequeñas)
- Si el BGP determina que dos o más rutas de SA pueden emplearse para llegar a un destino dado, puede emplear una variedad de medios para decidir cuál de las rutas SA candidatas empleará éste. La calidad de un SA puede medirse por parámetros tales como la velocidad del enlace, la capacidad y la congestión.
- Puede dar preferencia de rutas internas sobre las externas.
- El BGP se adhiere a esta importante regla: Un SA comunica a sus SA vecinos sólo las rutas que emplea.

3.3.3 Selección de la Ruta con el BGP

El BGP evalúa las diversas rutas que hay hacia una red destino desde los gateways fronterizos en esa red, luego, selecciona el mejor, aplica las limitaciones más importantes de la política y, finalmente, comunica todo esto a su BGP vecino. Cada SA puede tener sus propios criterios para la evaluación de la ruta.

Sea cual sea la política a implementarse, el BGP crea una base de datos de ruta que consiste en un conjunto de todas las rutas factibles y de la lista de redes a las que se llega por medio de cada ruta. En implementaciones de BGP reales, los criterios para asignar el grado de preferencias para una ruta se especifica en las tareas de configuración, las que incluyen la configuración de los BGP vecinos, la configuración del peso administrativo de una ruta hacia un vecino, la restricción de la información de ruta desde y hacia los vecinos, la configuración de direcciones resumidas, entre los criterios más relevantes.

Seleccionando Una Ruta

E i ten do fuente de información básica para seleccionar una ruta BGP:

1. Información explícita, presente en toda la ruta del SA.
2. Una combinación de información que puede provenir del mismo SA y de la información fuera de este

Entre los criterios para asignar un grado de preferencia para una ruta están:

- Conteo de SA. Las rutas con una cuenta de SA más corta son, por lo general, mejores.
- Consideraciones de las líneas directivas. El BGP respalda las rutas que se basan en políticas sustentadas en la distribución controlada de la información sobre rutas. Un *"speaker"* BGP puede percatarse de algunas limitaciones de las líneas directivas (tanto dentro y fuera de su propio SA) y llevar a cabo una adecuada selección de la ruta. Las rutas que no cumplen con los requerimientos que imponen las políticas ya no se toman más en cuenta.
- La presencia o la ausencia de cierto SA o SA en la ruta. Gracias a la información obtenida por la administración; fuera del alcance del BGP, un SA puede conocer ciertas características del desempeño (por ejm. ancho de banda, MTU) de ciertos SA y puede intentar evitarlas o preferirlas.
- Origen de la ruta. Una ruta que provenga del BGP (por ejm. cuyo punto final sea interno al último SA de la ruta) es, por lo general, mejor que una para la cual parte de la ruta se conozca vía el EGP o por algún otro medio.
- Subconjuntos de ruta SA. Se debería preferir una ruta SA mas especifica, que sea un subconjunto de una ruta SA más larga para el mismo destino. Cualquier

problema en la ruta más corta (tales como una interrupción del servicio) constituirá también un problema en la ruta más larga.

- Dinámica de los enlaces. En lugar de las rutas inestables se prefiere las estables. Obsérvese que este criterio debe emplearse con mucho cuidado para evitar que origine fluctuaciones innecesarias en la ruta. Por lo general, cualquier criterio que dependa de una información dinámica podría causar inestabilidad en la ruta.

Los criterios de elección relativos a ponderación de rutas:

- Si el siguiente salto es inaccesible, no lo tome en cuenta.
- Primero considere los pesos administrativos BGP más grandes.
- Si los ruteadores tienen el mismo peso, considere la ruta que tenga la preferencia más alta.
- Si los ruteadores tienen la misma preferencia local, opte por la ruta originada a partir del ruteador local.
- Si no se ha originado ninguna ruta, opte por la ruta del Sistema Autónomo más corta.
- Si todas las rutas son de la misma extensión de la ruta del Sistema Autónomo, opte por el código de origen más corto (IGP < EGP < INCOMPLETE). El origen indica el protocolo que generó la actualización de la ruta. El BGP considera tres tipos de origen y los prefiere en este orden: IGP primero, luego el EGP y finalmente INCOMPLETE. En el último tipo el BGP recibió un paquete de ruta por redistribución y estas rutas pueden haberse originado en cualquier fuente.
- Si los códigos de origen son los mismos y todas las rutas provienen del mismo Sistema Autónomo, prefiera la ruta con la métrica MED (Multi Exit Discriminator) más baja. Una métrica faltante se considera cero. Se emplea el

MED para dar preferencia a un i tema utónomo que tiene múltiples puntos de entrada.

- Si los MED son los mismos, opte por las rutas externas en lugar de las internas.
- Si la sincronización IGP está deshabilitada y sólo quedan las rutas internas, opte por la ruta a través del vecino más próximo. La sincronización IGP es una operación que tiene un BGP que espera un IGP para enviar información de ruta; antes de comunicar las rutas de tránsito a otros Sistemas Autónomos.
- Opte por la ruta que tenga el mínimo valor de dirección IP para el ID del ruteador BGP.

Modelo de Tabla de Rutas BGP

La ilustración, muestra el modelo de decisión de BGP. Una tabla de ruta BGP se mantiene separado de la tabla de ruta IP. La tabla BGP se usa para determinar cómo es que el ruteador pasa información de ruta. La tabla de ruta IP posee:

(a) Rutas dentro del Sistema Autónomo; y

(b) Rutas Dada por otros Sistemas Autónomos.

La tabla de rutas se emplea para encaminar a los paquetes a su ruta final. El buscador de las Políticas de ingreso es el primer proceso en el ruteador para llevar a cabo un análisis de la ruta. Este contiene filtros que evalúa los prefijos de la dirección, la información de la ruta SA y la información de los atributos BGP para tomar sus decisiones. Asimismo, puede cambiar los atributos de la ruta, lo que afecta a la próxima operación y al proceso de decisión, que determina la ruta que se toma hacia el siguiente nodo.

Los atributos son parámetros que se asocian con las rutas. El BGP los emplea para seleccionar las rutas. El proceso de toma de decisiones se emplea para

determinar la ruta que se seguirá para llegar al destino. Estas rutas se han filtrado a través del buscador de política de ingreso. Si contamos con más de una ruta para llegar al destino, entonces se examina los atributos de las rutas y se escoge la mejor. Las mejores rutas se colocan en la tabla de rutas IP y se comunican a los puntos BGP de este ruteador. El buscador de la política de salida es similar al de entrada; asimismo, éste ejecuta la importante función de no pasar las rutas informadas desde los nodos internos hacia otros nodos internos. Esta operación se le reserva a un IGP. Finalmente; la salida es la información que se envía a los BGP vecinos.

Interacciones con los IGP

Si el SA pasa la información de ruta entre distintos Sistemas Autónomos, el BGP debe estar acorde con la información que estos sistemas proveen. Por ejemplo, si el BGP informa acerca de una ruta antes de que todos los ruteadores en el SA sepan de esta ruta desde el IGP, estos ruteadores podrían recibir el tráfico para el cual no tienen un registro de la tabla de rutas. Para evitar esto, el BGP proporciona un proceso llamado sincronización BGP (o IGP) y espera hasta que el IGP se le comunique la información de ruta a través del SA.

Si el SA no va a ser un SA de tránsito, no es necesario configurar las operaciones de sincronización. Además, el administrador del SA debe tener cuidado de cómo se lleve a cabo la redistribución de la ruta. Por ejemplo, si el BGP se emplea para la redistribución de rutas al IGP, podría suceder que estas rutas se vuelvan luego a redistribuir al BGP.

3.3.4 Sistema Autónomo de Tránsito y Sin Tránsito

Un SA no requiere de crear ni terminar todo el tráfico que fluye hacia éste. Un SA puede ser es una red de tránsito "*multihomed*"; este término significa que el SA

tiene más de una interface conectada a otros SA. El término tránsito quiere decir que el tráfico tiene una dirección fuente o dirección destino externa al SA. El término sin tránsito significa que el SA no permite que el tráfico que no se destine para el SA (o no originado en el SA) pase a través de él.

Algunos dominios de ruta de intranet permiten las operaciones de tránsito y otros no. La decisión de permitir el tráfico de paquetes para pasar por un SA de la organización se sustenta en la política de ruta que tenga éste.

La ilustración 3.9 y 3.10 muestra un ejemplo de un SA sin tránsito "*multihomed*". El SA1 no transmite las informaciones respectivas del SA2 y del SA3 solo transmite las de su propio SA. En consecuencia, el SA2 y el SA3 no saben que pueden contactarse uno con otro por medio del SA1 y no pasará tráfico al SA1, de las redes que radican en estos dos Sistemas Autónomos.

La ilustración 3.11 y 3.12 muestra un SA "*multihomed*" de tránsito. Este SA comunica el tráfico a través de él; es decir, el tráfico que no se crea ni culmina en el SA1. Igualmente, la figura muestra otros aspectos de los SA de tránsito "*multihomed*". La operación BGP que intercambia información entre los Sistemas Autónomos se llama BGP externo o EBGP. Los ruteadores que operan en esta interface se denominan ruteadores de frontera BGP. La comunicación de información también se realiza a través del SA1 para permitir que el SA2 y el SA3 se comuniquen entre ellos con referencia a sus redes. En el caso en que; la operación BGP se realice en el interior del SA, se le denomina BGP interno o IBGP. Los ruteadores que intercambian esta información se denominan ruteadores de tránsito.

Continuando con el análisis de los Sistemas Autónomos de tránsito *"multihomed"*, la ilustración 3.11 muestra cómo la comunicación de las rutas se

realiza por el SA1 al SA2 y al SA3. El hecho de que el SA1 esté comunicando rutas a nombre del SA2 y el SA3, significa que el SA1 acepta ser una red de tránsito para el SA2 y el SA3.

Los ruteadores de tránsito del SA1 también pueden conectarse con los ruteadores sin tránsito en el Sistema Autónomo. Estos no necesitan configurarse con el BGP, pudiendo emplear un IGP como el OSPF.

A fin de evitar los lazos cerrados de información, los ruteadores 3 y 4 no se pasan entre sí esta información. La regla para evitar este potencial lazo cerrado es que los ruteadores no pasen rutas que provengan de los BGP vecinos internos hacia otros BGP vecinos internos.

Por lo tanto, en este ejemplo, el ruteador 2 debe integrarse con todos los ruteadores internos de su propio SA; con la finalidad de pasarse entre ellos información de rutas de sus propias redes. Este planteamiento funciona lo bastante bien en el caso de dominios de rutas pequeños, pero en el caso de sistemas grandes, éste ejerce una gran carga sobre el ruteador principal que tiene que establecer sesiones con todos los demás nodos BGP.

Este nuevo problema planteado se puede encarar de dos formas distintas; las cuales se mencionan a continuación:

(a) Confederaciones de Dominios de Rutas; y

(b) Reflectores de Rutas.

Confederaciones de Dominio de Ruta

Cuando las redes grandes (tales como las redes ISP) se vuelven más grandes, un Sistema Autónomo puede llegar a ser dificil de administrar y controlar. Las configuraciones IBGP pueden tornarse complejas y un ruteador puede tener que

brindar soporte a muchas sesiones internas de BGP. Uno de los métodos para tratar este problema es dividir el SA en partes más pequeñas (figura 3.13). Un Sistema Autónomo puede dividirse en múltiples Sistemas Autónomos y agruparse en una sola confederación. Para el exterior, esta división es transparente y los SA divididos aparecen como si fueran un solo Sistema Autónomo. La ilustración de la figura 3.13 muestra la configuración del dominio de la ruta.

Cada uno de los Sistemas Autónomos está totalmente integrado y tiene conexiones con los demás SA dentro de la confederación. Los puntos de intercambio entre los distintos Sistemas Autónomos, tienen sesiones EBGP; pero intercambian información de ruta con puntos IBGP. Diversos parámetros claves de configuración del ruteador se emplean para permitir que un IBGP se utilice dentro de la confederación; como por ejemplo; el siguiente salto, el MED y la preferencia local.

Observe que EBGP está operando entre los SA más pequeños; dentro de la confederación. Esta configuración se requiere porque cada SA pequeño, emplea un número de SA propio y por lo tanto, se usa enrutamiento EBGP.

Reflectores de Ruta

Otro planteamiento para combatir el problema es el de configurar los ruteadores para dar soporte a un reflector de ruta. Esta operación significa que todos los speakers IBGP no tienen por qué estar completamente integrados. Este planteamiento se representa en la ilustración correspondiente. El ruteador 5 se configura para ser un reflector de ruta y asume la responsabilidad de pasar la información sobre la ruta a un conjunto de IBGP vecinos. Cuando el ruteador recibe del ruteador 4 la información sobre la ruta, le comunica la información al ruteador 6; así de esta manera; las sesiones IBGP entre el ruteador 4 y el 6 se hacen innecesarias

(figura 3.14). Mediante el empleo del reflector de ruta; se ayuda a reducir muchas sesiones. Los puntos BGP internos del reflector de la ruta se dividen en: Puntos clientes y puntos no clientes. El reflector de ruta forma una asociación con sus puntos cliente llamada *"cluster"* y estos ruteadores cliente no tienen por qué estar conectados entre sí. Estos no se comunican con los *"speakers"* IBGP que estén fuera del cluster.

Para saber cómo trabaja un reflector, repasemos a continuación la siguiente lista de operaciones:

- Comunica la información del BGP externo a todos los puntos clientes y no clientes.
- Cualquier información desde un punto no cliente se comunica a todos los clientes.
- Una ruta desde un cliente se comunica a todos los puntos cliente y no cliente, lo que significa que los clientes no tienen que estar conectados.

Luego de revisar el funcionamiento general del BGP; estudiaremos el formato de las tramas del protocolo BGP.

3.3.5 Formato de Mensajes BGP

Se tienen 4 diferentes tipos de mensaje en BGP:

- Apertura *"Open"*
- Notificación *"Notification"*
- Verificación *"Keepalive"*
- Actualización *"Update"*

Todos los mensajes van precedidos por un encabezado común que se muestra en la figura 3.15. El campo marcador se utiliza para la sincronización de mensajes, para la

detección de pérdida de paquetes, o también para validar los mensajes entrantes. El campo longitud; para indicar la longitud del paquete; y el campo tipo de mensaje, para identificar el tipo.

OPEN

El mensaje OPEN se emplea para establecer vecinos BGP. la operación se produce como primer paso antes de que tenga lugar cualquier intercambio de información de BGP es pues, un primer mensaje de contacto entre los ruteadores BGP. Si el mensaje OPEN es recibido por el otro ruteador, éste responde con un mensaje KEEPALIVE. Tan pronto como se produce este primer contacto, los nodos BGP vecinos pueden intercambiar mensajes de UPDATE, KEEPALIVE y NOTIFICATION. Los mensajes OPEN se muestran en la ilustración 3.16.

La primera parte del mensaje contiene el número de versión del BGP y el número del Sistema Autónomo del remitente. El próximo campo es el tiempo de contención, que indica el número de segundos para el valor del cronómetro de contención. Este señala el espacio de tiempo en que los BGP vecinos deberían considerar como válida la información del remitente. El próximo campo es el identificador del remitente BGP. Este valor se determina durante las operaciones de contacto entre los puntos BGP y es el mismo valor para cada interface local y cada punto BGP.

Los dos últimos campos en el paquete son el código de validación y los datos de validación. El campo de código indica el procedimiento de validación empleado para la sesión BGP y el campo de datos de validación depende del campo de código.

UPDATE

El mensaje UPDATE, mostrado en la ilustración 3.17, se emplea para el intercambio de información de enrutamiento. Este comunica una ruta a un punto

BGP y puede, igualmente, retirar del servicio un gran número de rutas que en algún momento sean inalcanzables. Este mensaje contiene dos campos principales y dos de extensión.

El campo de las rutas retiradas contiene una lista de prefijos de direcciones IP de las rutas que están fuera de servicio. El campo del atributo de la ruta es una lista de atributos de ruta, codificadas como: tipo de atributo, longitud del atributo y el valor del atributo. El campo de alcance de la red contiene una lista de prefijos de direcciones IP que la ruta BGP conoce y que puede alcanzar.

Un mensaje UPDATE puede dar información de una ruta como máximo, lo que puede describirse a través de diversos atributos de ruta. También es posible que un mensaje UPDATE pueda proporcionar una lista de múltiples rutas para que se retiren de la tabla de rutas.

KEEPALIVE

Los mensajes KEEPALIVE se emplean entre dos puntos para cerciorarse de que éstos estén en funcionamiento y el enlace hacia ellos esté operativo. El mensaje KEEPALIVE sólo se compone de una cabecera de mensaje BGP. Los mensajes KEEPALIVE se intercambian entre los puntos BGP para renovar el cronómetro de contención.

NOTIFICATION

Este mensaje se emplea cuando se produce un error o la conexión de un punto se cierra y contiene un número de código de error. Luego de la emisión de este, la conexión BGP se cierra. Los mensajes de NOTIFICATION pueden hacer referencia a cualesquiera de los 3 tipos de mensajes anteriores (ver figura 3.18).

3.3.6 Estados de un Nodo BGP

Se define a nivel de BGP un conjunto de estados en los cuales cualquier nodo BGP en funcionamiento atraviesa desde el inicio, para lo cual se hace uso de los diversos mensajes que ofrece BGP para operar y que se han estudiado anteriormente; hasta llegar al punto de poder intercambiar información de enrutamiento sobre el SA. Las figuras 3.19 y 3.20 muestran estos estados. A continuación se pasa a describir estos estados BGP.

IDLE

Es el estado inicial de cualquier nodo bajo BGP, el cual significa que este nodo esta a la espera del evento de inicio, que es el establecimiento de una sesión BGP o el reinicio de esta; que normalmente es ejecutado por el operador o por el nodo remoto. Luego del evento de inicio, BGP renueva sus recursos y espera la conexión TCP del nodo remoto, si la sesión se establece se pasa al siguiente estado. Caso contrario, ante cualquier falla se retorna al estado idle, enviando un mensaje de notificación.

CONNECT

En este estado BGP espera a que el proceso de inicio de la sesión se concluya con éxito de ser así, pasa al estado Open Sent, sino, pasa al estado ACTIVE, si ocurre algún otro evento el estado regresa a IDLE.

ACTIVE

En este estado se esta a la espera del reinicio nuevamente de la sesión a nivel del protocolo de transporte, si esto se completa exitosamente se pasa al estado siguiente. En caso contrario se retorna al estado CONNECT, ante cualquier otro evento generado por el operador o el sistema se pasa al estado IDLE.

OPENSENT

En este estado se espera la llegada de un mensaje OPEN del otro vecino. Al arribo de este se verifica los datos de seguridad e identidad del SA. Si es conforme se envía un mensaje de KEEPALIVE, y se pasa al estado OPENCONFIRM; caso contrario se pasa al estado IDLE.

OPENCONFIRM

En este estado, se esta a la espera del arribo de un mensaje KEEPALIVE en respuesta al enviado previamente. Si llega la confirmación sin problemas se pasa al estado siguiente. Si no llega la confirmación del mensaje se pasa al estado IDLE.

ESTABLISHED

Este estado; que es el estado final de todos los nodos, indica que el proceso de negociación y reconocimiento de pares BGP ha concluido satisfactoriamente. A partir de este estado BGP se encuentran en capacidad de intercambiar información de rutas haciendo uso de los mensajes UPDATE

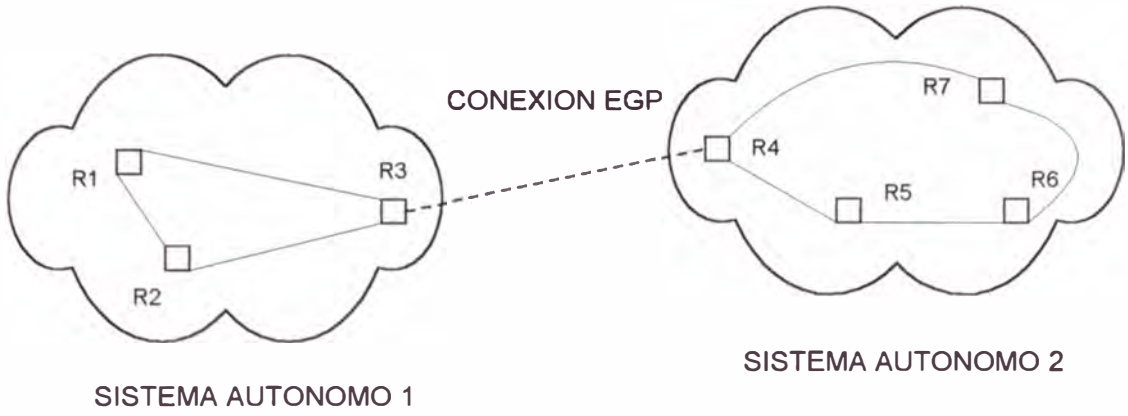


Fig 3.1 SA Conexión EGP

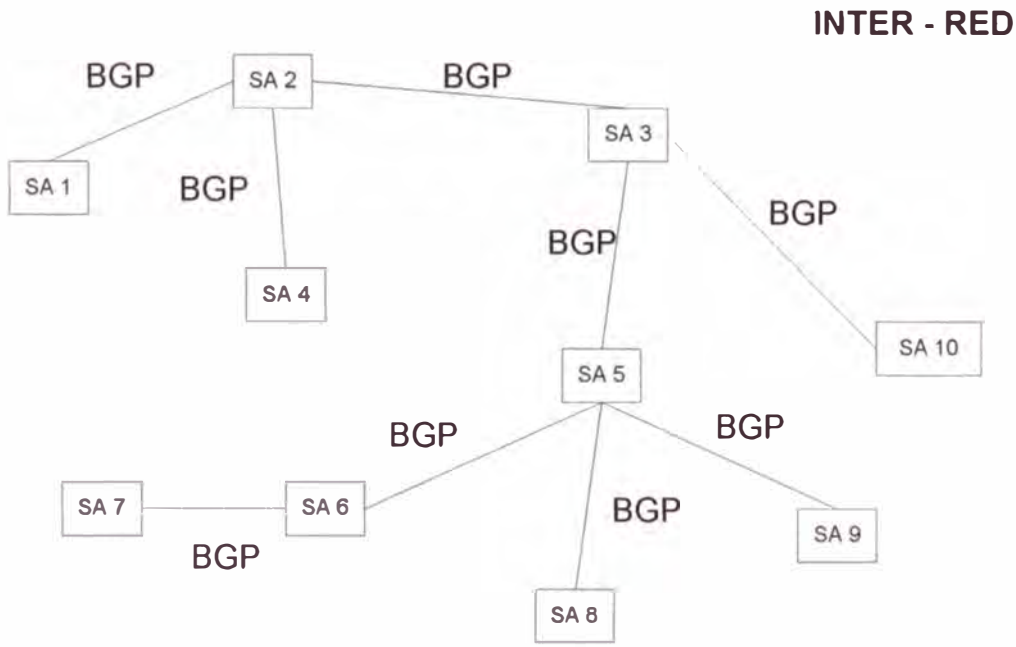


Fig 3.2 SA Conexión BGP

VERSION	TIPO	CODIGO	ESTADO
VERIFICACION DE SUMA		SISTEMA AUTONOMO	
NUMERO DE SECUENCIA		RESERVADO	
NUMERO IP DE RED FUENTE			

Fig 3.3 Trama de sondeo EGP

VERSION	TIPO	CODIGO	ESTADO
VERIFICACION DE SUMA		SISTEMA AUTONOMO	
NUMERO DE SECUENCIA		N GW INT	N GW EXT
NUMERO IP DE RED FUENTE			
NUMERO IP DE RUTEADOR			
N DISTANCIAS			
DISTANCIA 1	N DE REDES		
RED 1 A DISTANCIA 1			
RED 2 A DISTANCIA 2			

Fig 3.4 Trama de rutas EGP

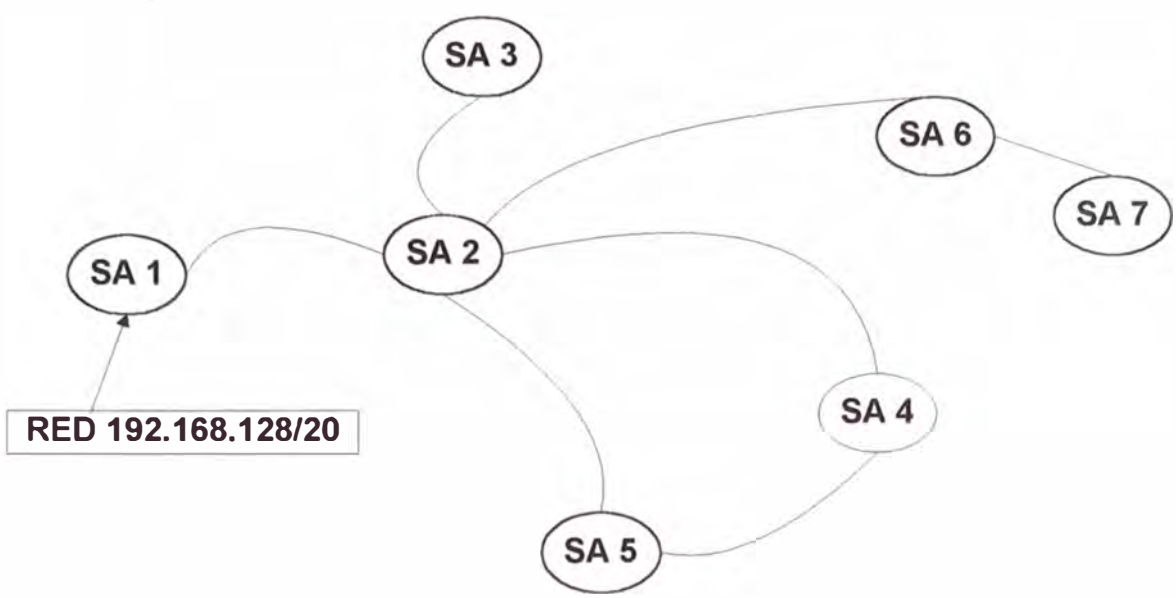


Fig 3.5 SA con BGP

SA 1	ANUNCIA	192.168.128/20	SA1
SA 2	ANUNCIA	192.168.128/20	SA1--SA2
SA 4	ANUNCIA	192.168.128/20	SA1--SA2--SA4
SA 5	RECIBE	192.168.128/20	SA1--SA2
SA 5	RECIBE	192.168.128/20	SA1--SA2--SA4
SA 3	RECIBE	192.168.128/20	SA1--SA2
SA 6	RECIBE	192.168.128/20	SA1--SA2
SA 7	RECIBE	192.168.128/20	SA1--SA2--SA6

Fig 3.6 Tabla de rutas BGP

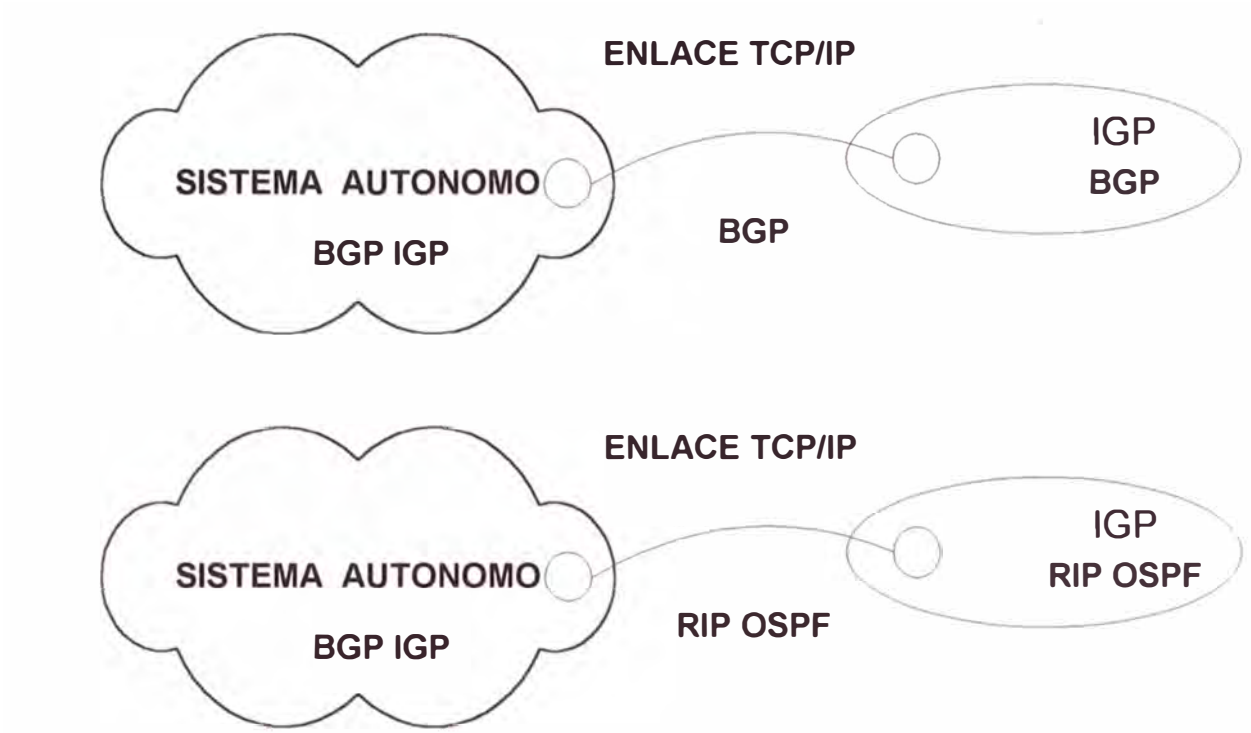


Fig 3.7 Conexión Cliente a ISP

Fig. 3.8 Modelo de decisión BGP

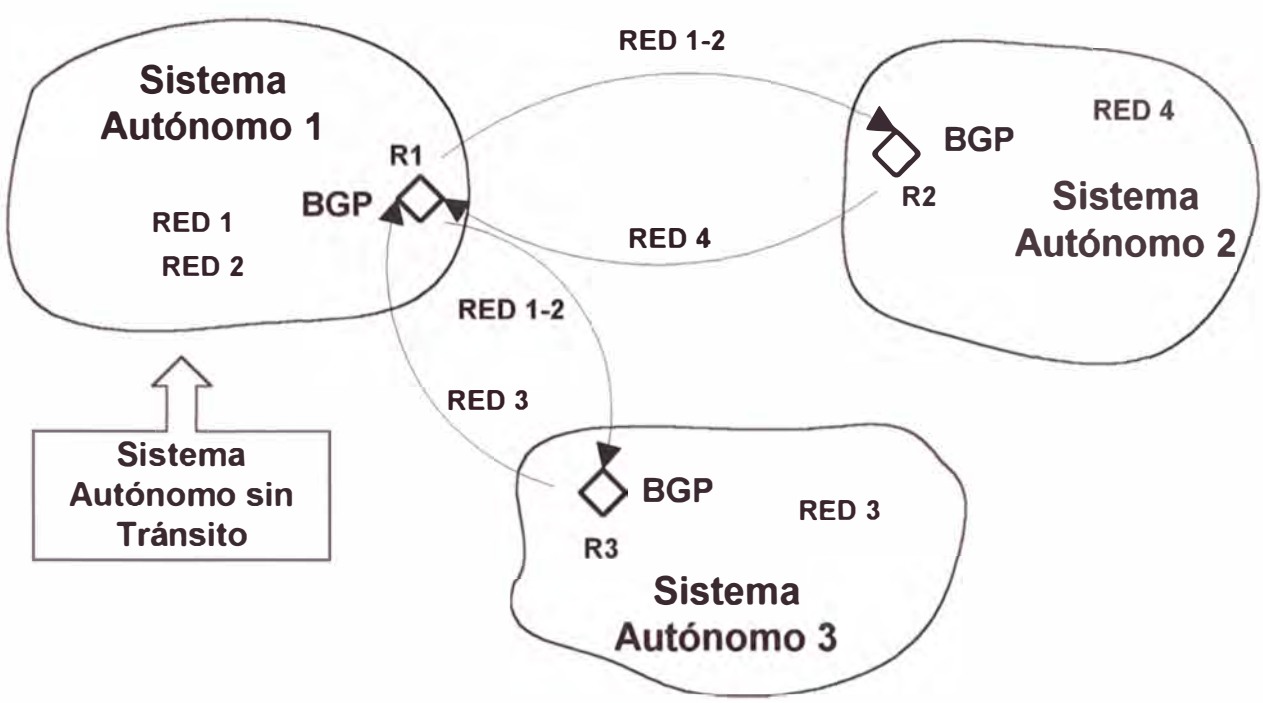


Fig 3.9 Sistema Autónomo sin Tránsito

SISTEMA AUTONOMO	PUBLICA	RECIBE
S A 1	RED 1 SA1	RED 3 SA3
	RED 2 SA1	RED 4 SA2
S A 2	RED 4 SA2	RED 1 SA1
	–	RED 2 SA1
S A 2	RED 3 SA3	RED 1 SA1
	–	RED 2 SA1

Fig 3.10 Tabla de Rutas BGP

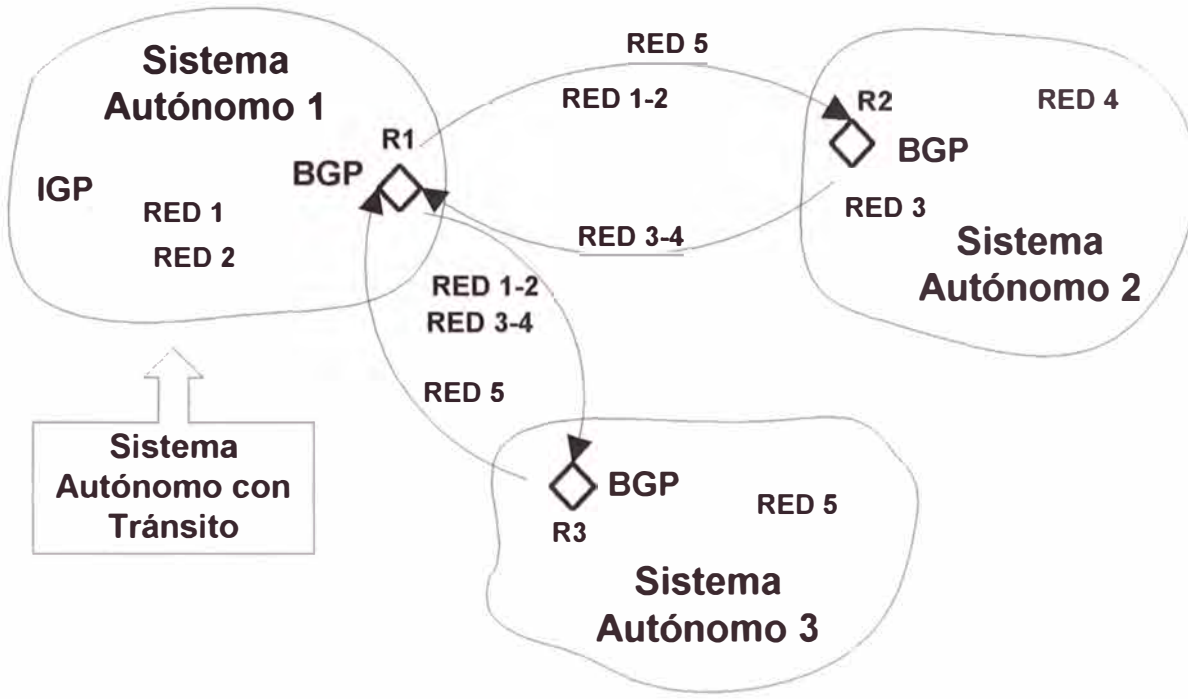


Fig 3.11 Sistema Autónomo con Tránsito

SISTEMA AUTONOMO	PUBLICA	RECIBE
S A 1	RED 1 SA1	RED 3 SA2
	RED 2 SA1	RED 4 SA2
	RED 3 SA1_SA2	RED 5 SA3
	RED 4 SA1_SA2	_
S A 2	RED 3 SA2	RED 1 SA1
	RED 4 SA2	RED 2 SA1
	_	RED 5 SA1_SA3
S A 3	RED 5 SA3	RED 1 SA1
	_	RED 2 SA1
	_	RED 3 SA1_SA2
	_	RED 4 SA1_SA2

Fig 3.12 Tabla de Rutas BGP

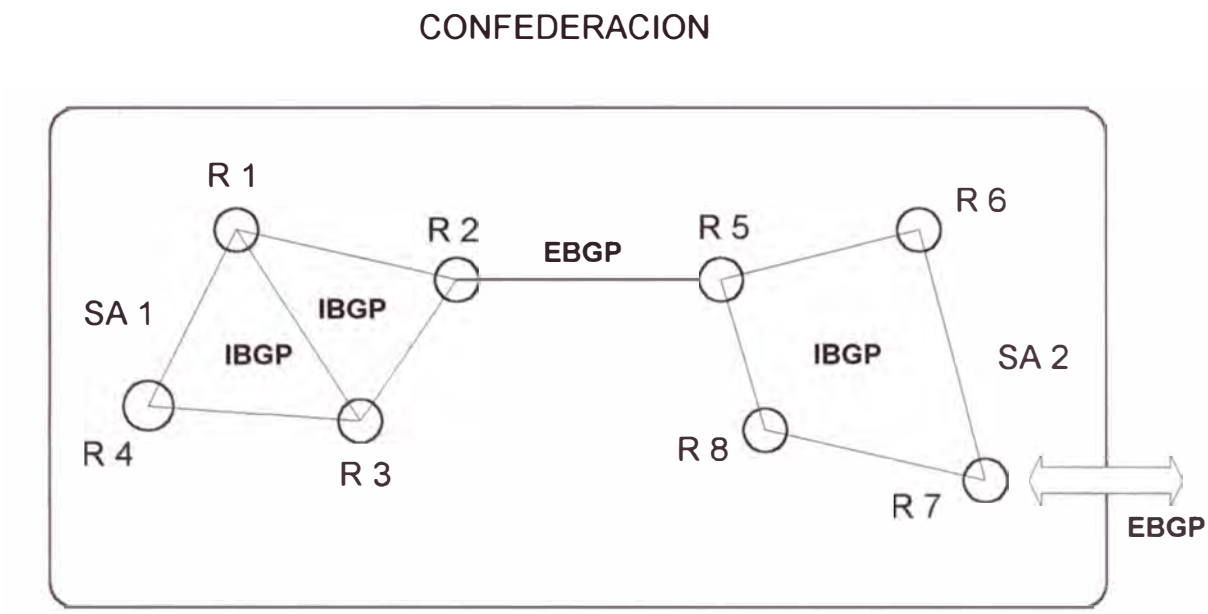


Fig 3.13 Sistema Autónomo Confederado

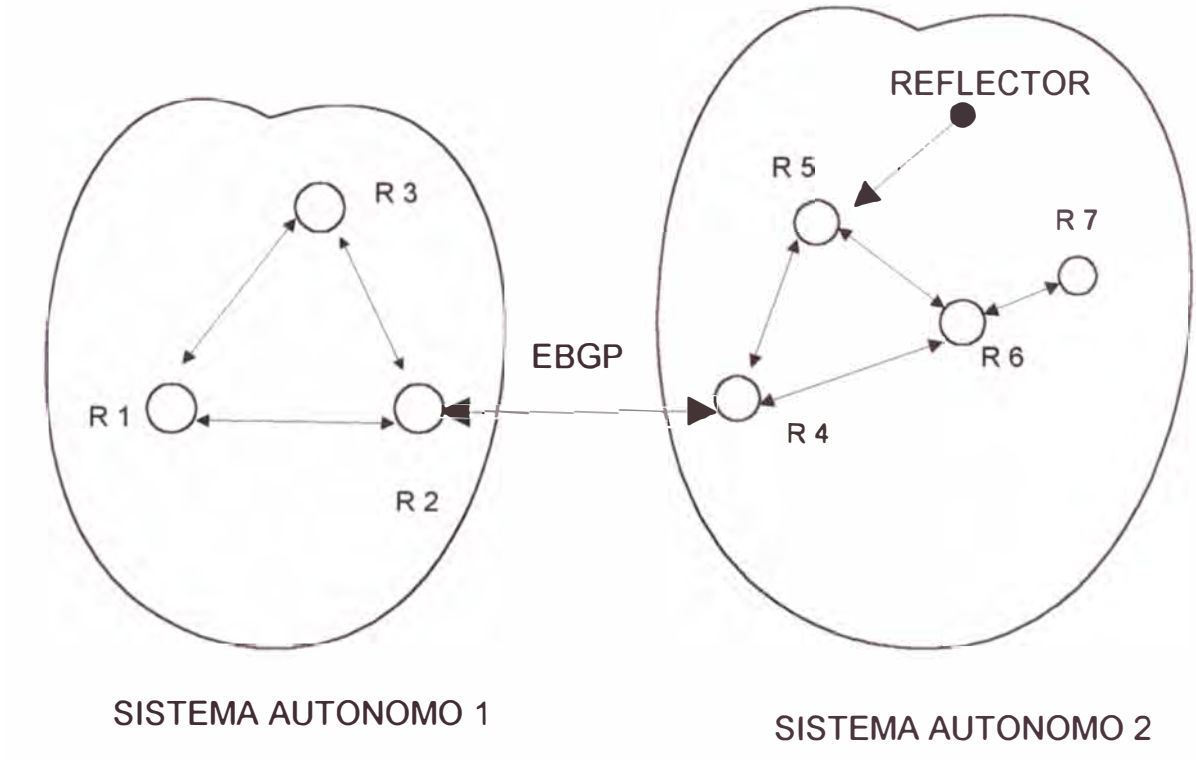


Fig 3.14 Sistema Autónomo con Reflector

Fig 3.15 Cabecera de BGP

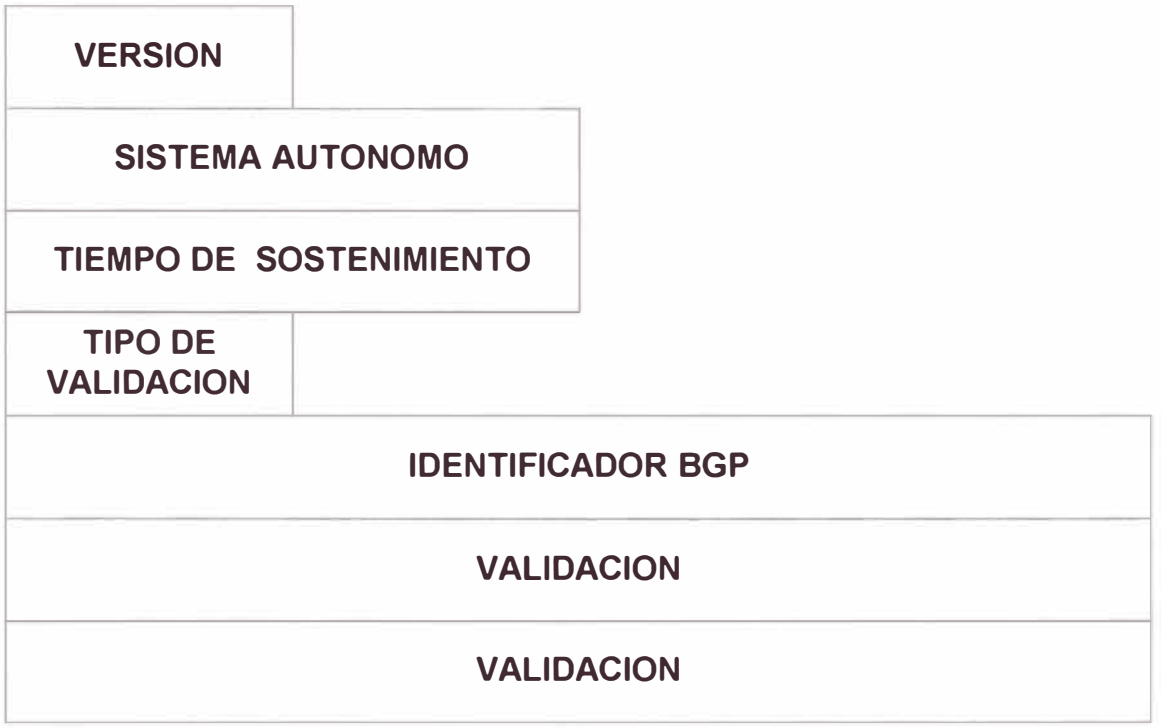


Fig 3.16 Cabecera de BGP Open

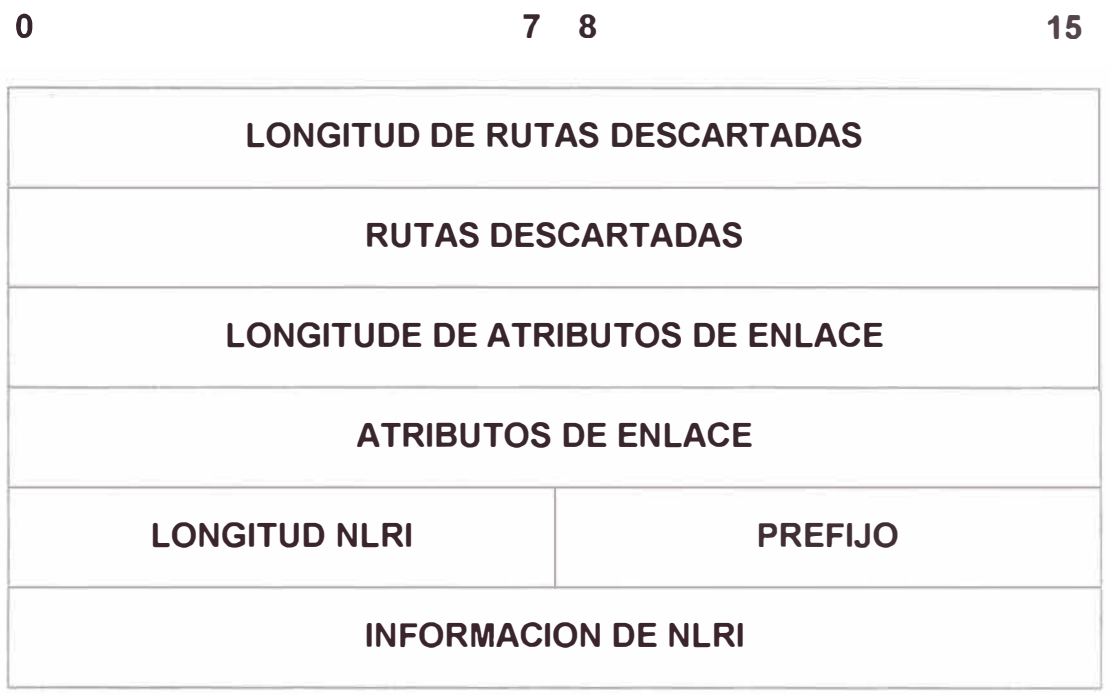


Fig 3.17 Trama de Update

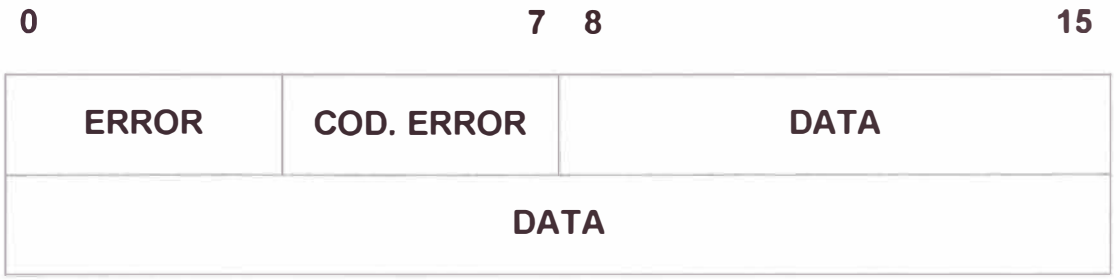


Fig 3.18 Trama de Notificación

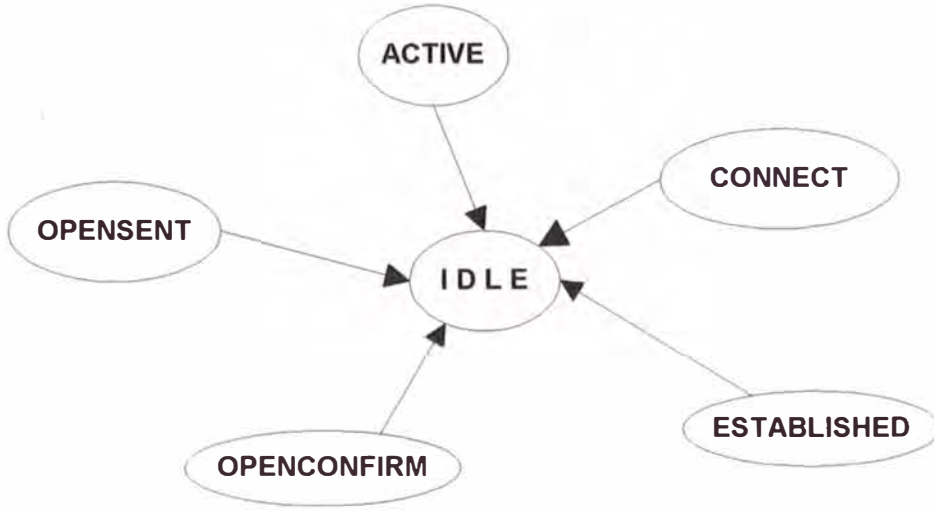


Fig 3.19 Estados de Nodos BGP

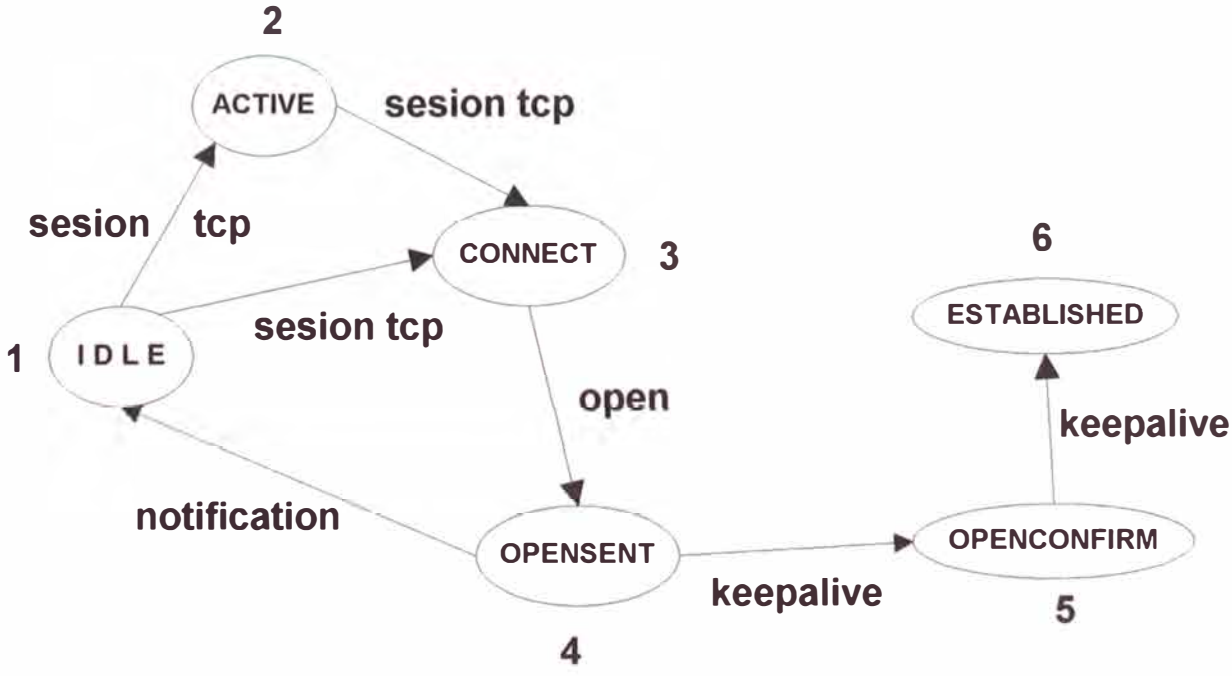


Fig 3.20 Tránsito de Estados BGP

CAPITULO IV

INTERCONEXIÓN EN EL NAP-PERÚ

En los capítulos anteriores se han desarrollado los tópicos necesarios para realizar la interconexión de Sistemas Autónomos; a continuación trataremos los pormenores de los esquemas de interconexión implementados en el NAP-PERÚ; partiendo del estrato físico y ascendiendo hasta el nivel de enlace IP de red.

4.1 Orígenes y Estructura del NAP-PERÚ

La idea de Interconexión de las redes conectadas a Internet en el Perú no es nueva; fue planteada dentro del marco de las relaciones comerciales derivadas luego de la aparición en el medio nacional de la segunda red TCP/IP interconectada a Internet en el Perú; Unired, la cual es parte de Telefónica, que a su vez es filial del grupo Telefónica de España, el año 1997. Hasta entonces la única red IP conectada a Internet era la red de Red Científica Peruana. La necesidad de Interconexión IP fue planteada por parte de RCP a Unired–Infovía tanto en el ámbito privado como público; sin embargo, se argumentaron diversas razones de tipo comercial y técnico, una de ellas referida al acuerdo del tipo de protocolo de enrutamiento IP a utilizarse y otras de tipo económico-comercial, que impidieron la llegada al consenso final.

Luego de varios años; y con la apertura del mercado de telecomunicaciones en nuestro país; al finalizar el plazo que le concedía exclusividad a Telefónica en ciertos sectores de telecomunicaciones en el marco del contrato firmado con el Estado, es que se retoma esta iniciativa. A nivel técnico la discusión estaba zanjada por común acuerdo entorno al uso del protocolo BGP. Lamentablemente razones de índole político-comercial frustraron nuevamente el encuerdo entre Telefónica y RCP en 1998-1999. Tampoco hubo acuerdo alguno entre la tercera empresa que comenzó a brindar acceso a Internet, que en ese entonces era Firstcom, con alguna de las anteriores. Cabe resaltar que la interconexión a nivel físico ya se producía de facto pues al carecer RCP de una red física portadora local propia para llegar a sus usuarios, tanto estos como RCP debían recurrir a Telefónica y Firtscom, alquilando sus conexiones de tipo dedicado o conmutado para comunicarse entre sí.

Finalmente; en el año 2000 con la entrada de Comsat al campo de las telecomunicaciones en el país; y también, la aparición del servicio de Internet de Bellsouth; las conversaciones entre los diferentes proveedores de servicios de Internet, así como por las iniciativas surgidas, tanto de carácter privado y público; entorno a este asunto; era posible vislumbrar que estaban próximos; inexorablemente; acuerdos bilaterales o multilaterales de interconexión entre las diferentes empresas proveedoras de servicios de Internet.

Es así como; dentro del marco propicio al diálogo y al consenso común entre los competidores; así como, en el interés mostrado por el Estado en impulsar la interconexión, entre los distintos proveedores existentes; se plasma en las reuniones del Comité de Telecomunicaciones de la cámara de comercio peruano-americana, AMCHAM presidido entonces por el Ing. Luis Alejos; representante de Comsat

Perú, en el año 2000, comité al cual pertenecían las cuatro empresas extranjeras incluyendo Telefónica, se invita a RCP para que junto a las otras cuatro empresas, se iniciaran conversaciones que condujeron a la puesta en marcha del proyecto Nap-Perú.

4.2 Organización del NAP-PERÚ:

El NAP-PERÚ, es una asociación de carácter privada fundada por las siguientes empresas portadoras proveedoras de servicios de Internet:

AT&T

Bellsouth

Comsat-Perú

Red Científica Peruana

Telefónica

Esta organización consta de dos comités:

El Comité Directivo y El Comité Técnico. Ambos comités están formado por representantes de las empresas integrantes y tienen carácter permanente. La función del Comité Directivo es aprobar las normas y reglamentos así como resolver las controversias y ejecutar los acuerdos tomados; teniendo al Comité Técnico como órgano consultivo.

El Comité Técnico fue el encargado de la elaboración del proyecto así como de su ejecución y puesta en operación; también tiene dentro de sus funciones delinear las políticas de interconexión, seguridad y control que se ejecutan al interior del NAP; que asegure la operatividad del mismo y que sean respetadas por todos sus miembros. El Comité Técnico esta formado por Luis Aliaga (AT&T), Efraín Sarango (Bellsouth), Fernando Torres (Comsat Perú), Luis Chuquicondor (Infoductos-RCP),

Johnny Ramos (Telefónica). El Nap-Perú entró en operación el 10 de mayo del 2001

4.3 Topología de la Red del NAP-PERÚ

En el caso de la interconexión en el NAP peruano se puede encontrar las diferentes muestras de tecnologías disponibles en la diversidad de ofertas de las empresas operadoras que concurren en el NAP. Trataremos los tres estratos de interconexión.

Cada portador es libre de utilizar la tecnología de su elección; siempre y cuando garantice la conectividad a nivel de IP, así como la capacidad inicial de acceso y la ampliación a futuro que sea necesaria. Los accesos remotos WAN no son uniformes. La infraestructura fisica del local del NAP permite la colocación de nodos de acceso para cada integrante, con capacidad suficiente para 20 participantes (un armario por asociado); in situ, dentro de un recinto aislado con sistema de enfriamiento y energía redundante. La red diseñada para permitir el intercambio de tráfico IP entre los concurrentes, esta basada en tecnología Fastethernet conmutada (ver figura 4.1) con una capacidad de 100 Megabits por segundo, con redundancia en línea y distribución de carga; esto permite una capacidad máxima de transmisión de 200 megabits por segundo; por cada nodo concurrente. Los equipos de capa 3 IP son ruteadores con al menos 2 interfaces fastethernet 802.3u, full duplex, y los puertos de acceso WAN que cada operador considere necesarios, los cuales no pueden ser, en sumatoria total, de una capacidad de tráfico menor a lo acordado por el Comité Técnico. En las siguientes secciones se explican los distintos niveles de la red del Nap-Perú

4.3.1 Nivel Físico

Se usan medios fisicos de fibra óptica para la mayoría de los circuitos y también

enlaces digitales de microondas. Inicialmente esta prevista una capacidad de transmisión de 4 megabits por segundo para el enlace de cada portador. Se tiene previsto un segundo escalón de ampliación de capacidad de 8 megabits por segundo en caso de ser necesario. Los asociados están obligados a contar con la escalabilidad necesaria, para soportar la ampliación de capacidad a futuro. Las interfaces fisicas WAN usadas son de tipo E1 o múltiplos así como también terminación Ethernet (AT&T), en realidad no existe restricción alguna al respecto, dejándose a criterio de cada empresa asociada utilizar la que considere más conveniente.

4.3.2 Nivel de Enlace

En el estrato 2 se usan protocolos de tipo punto a punto a punto, como PPP o HDLC, también se usa Ethernet sobre ATM en el acceso remoto WAN. La capacidad mínima de cada enlace es de 4 megabits por segundo, en caso de ser necesario se incrementará esta capacidad la cual se determina por el procedimiento acordado por el Comité Técnico.

4.3.3 Nivel de Red

En el estrato 3 (ver figura 4.2) se realizará la interconexión IP empleando el protocolo de enrutamiento para Sistemas Autónomos BGP-4, por lo cual los ruteadores pertenecientes a los operadores que se instalen en el NAP debe soportar este protocolo.

Esta establecido que no se aplicará ninguna política de restricción, análisis de tipo de trafico, o priorización de tráfico, ni de contenido, al interior de las instalaciones ni de ninguno de los equipos que se instalen en el NAP. La finalidad del NAP es únicamente la de intercambiar tráfico IP en igualdad de condiciones para todos los participantes.

4.4 Interconexión IP en el NAP-PERÚ

El modelo que se usa en el NAP es el modelo ruteador contra ruteador, de la figura 4.3 y 4.4, en un esquema todos contra todos, formando una topología lógica tipo malla. Esto significa que cada ruteador debe mantener permanentemente una sesión, con cada ruteador presente en el NAP, a nivel del protocolo BGP-4. La otra posibilidad es el modelo ruteador contra servidor, este último modelo es apropiado cuando el número de participantes es mas elevado.

Se optó por este modelo, en vista de que la cantidad de participantes era reducida, cinco en total, por la facilidad de configuración y mantenimiento; así como por permitir una fácil descentralización de la administración e implementación de nuevos accesos para los futuros participantes del NAP. Estas son las características más importantes obtenidas gracias a la flexibilidad este modelo.

El esquema básico que se aplica en la configuración del entorno BGP-4 en el NAP son los siguientes:

- Los enrutadores empleados en el NAP soportan el protocolo BGP-4.
- Cada ruteador tiene dos conexiones a través de puertos fastethernet; una con cada conmutador.
- Cada ruteador tiene sesiones BGP-4 establecidas con todos los demás enrutadores participantes.
- Los participantes anuncian a través de las sesiones BGP, únicamente las redes que pertenecen a su propio Sistema Autónomo.
- Los participantes no aceptarán rutas de redes a través de las sesiones BGP, que no hayan sido autorizadas previamente por el Comité Técnico; así sean redes asignadas a los participantes.

- Los participantes no intercambian tráfico únicamente en pares o en grupos que excluyan a alguno de los participantes; intercambian tráfico con todos los participantes.
- Se tiene activado el sistema de automático de aislamiento y reconexión, provisto por las sesiones BGP para el caso de fallas de acceso en la red de cualquier participante.

Para afrontar el problema de escalabilidad, se tiene previsto el esquema mostrado en las figuras 4.5 y 4.6; utilizando un servidor de rutas BGP. Este sistema centraliza las tareas de configuración y administración en el servidor BGP.

Esos son los lineamientos generales implementados en el Nap Perú, para la interconexión a través del protocolo BGP-4. Los detalles específicos en cuanto a la configuración de los enrutadores es manejado por cada participante; así como la marca de equipo que decida emplear. El objetivo principal es la conexión sin restricciones de ninguna naturaleza en igualdad de condiciones para todos los participantes, así como la práctica de reglas y procedimientos que fomenten un ambiente de reciprocidad en la seguridad, confiabilidad, calidad del servicio, confianza y respeto entre los participantes.

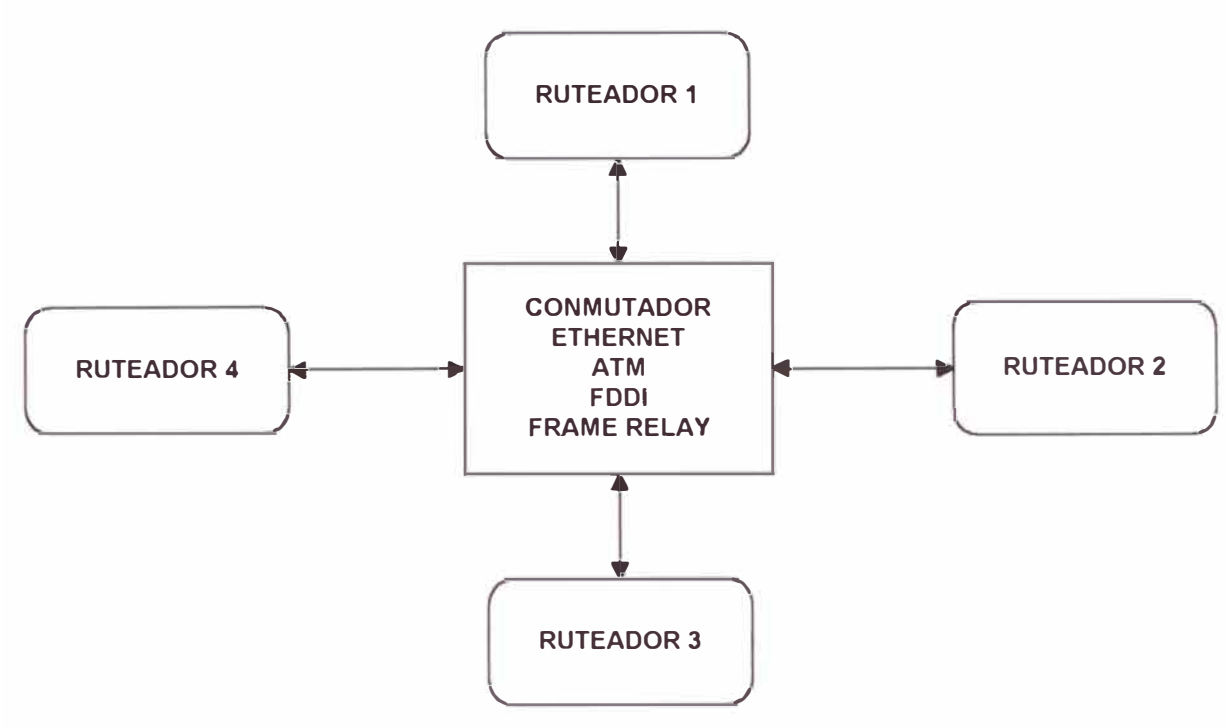


FIG 4.1 Topología Física

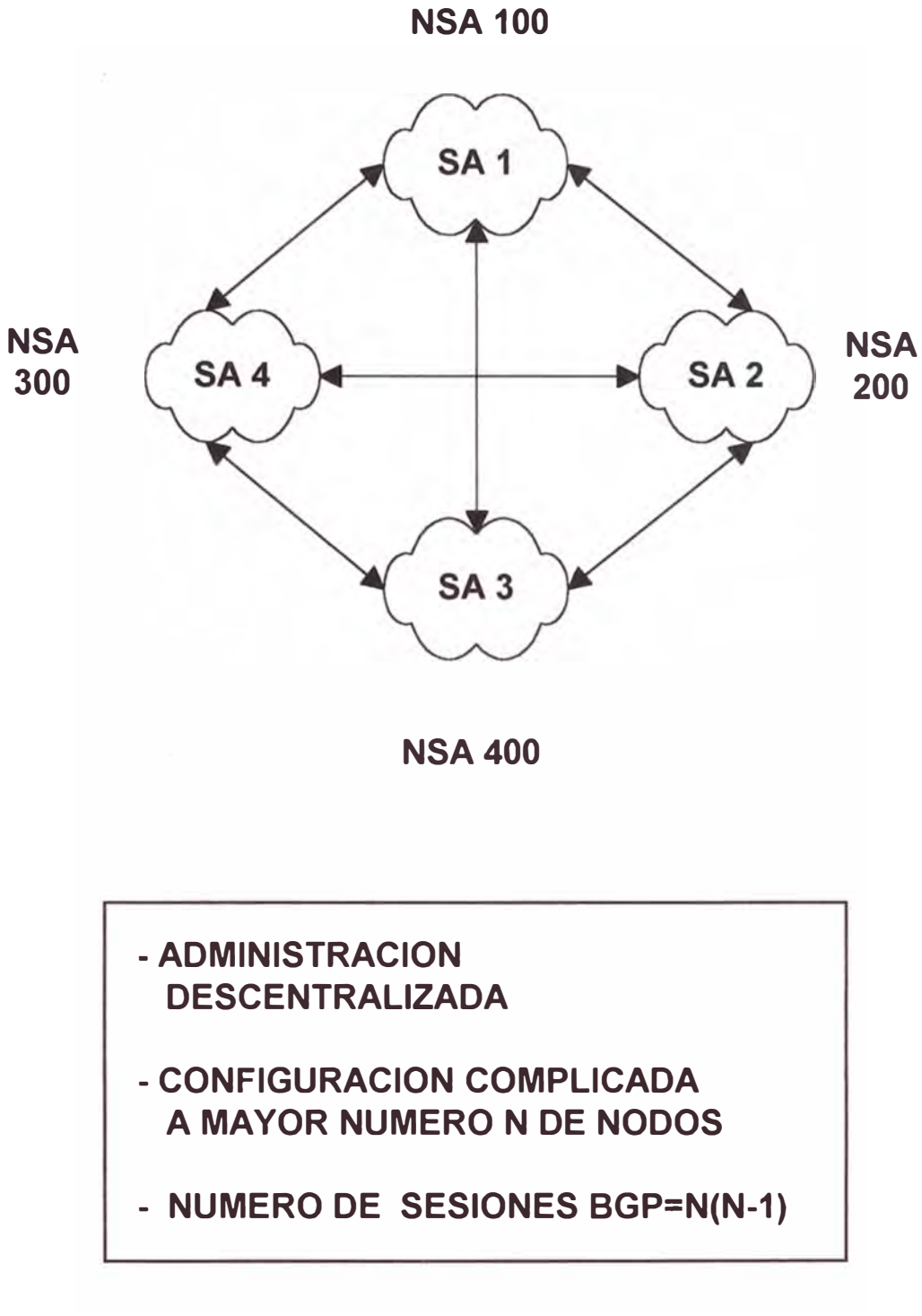


FIG 4.2 Topología BGP

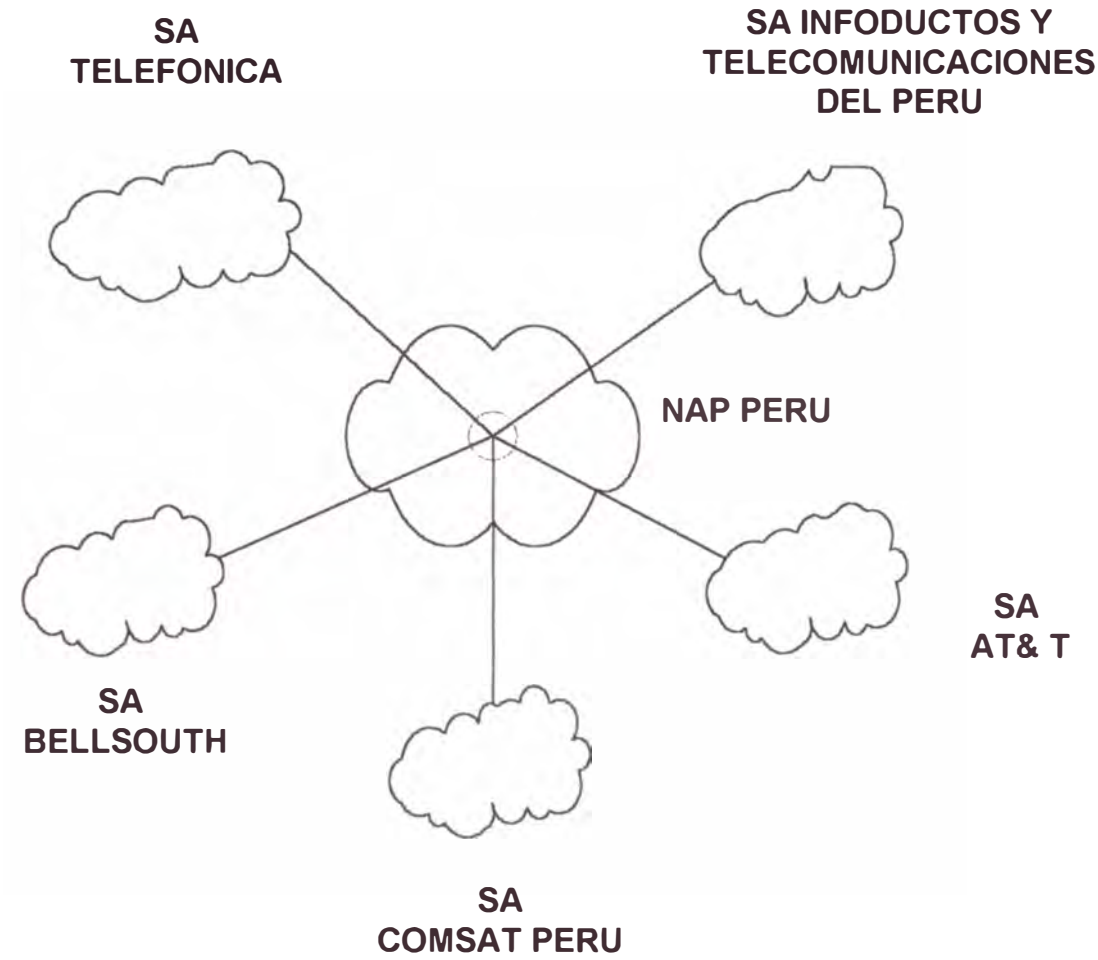


FIG 4.3 Topología Física

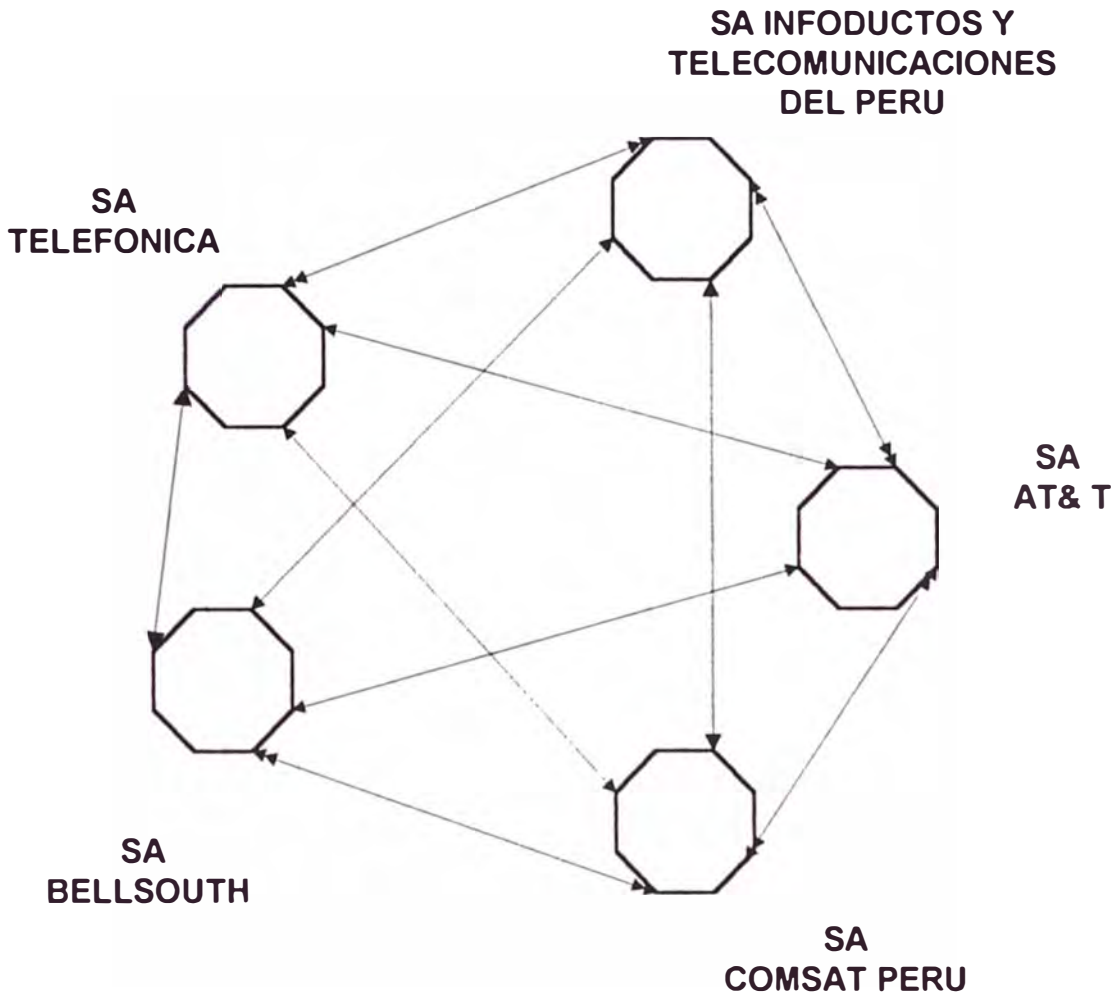


Fig 4.4 Topología BGP

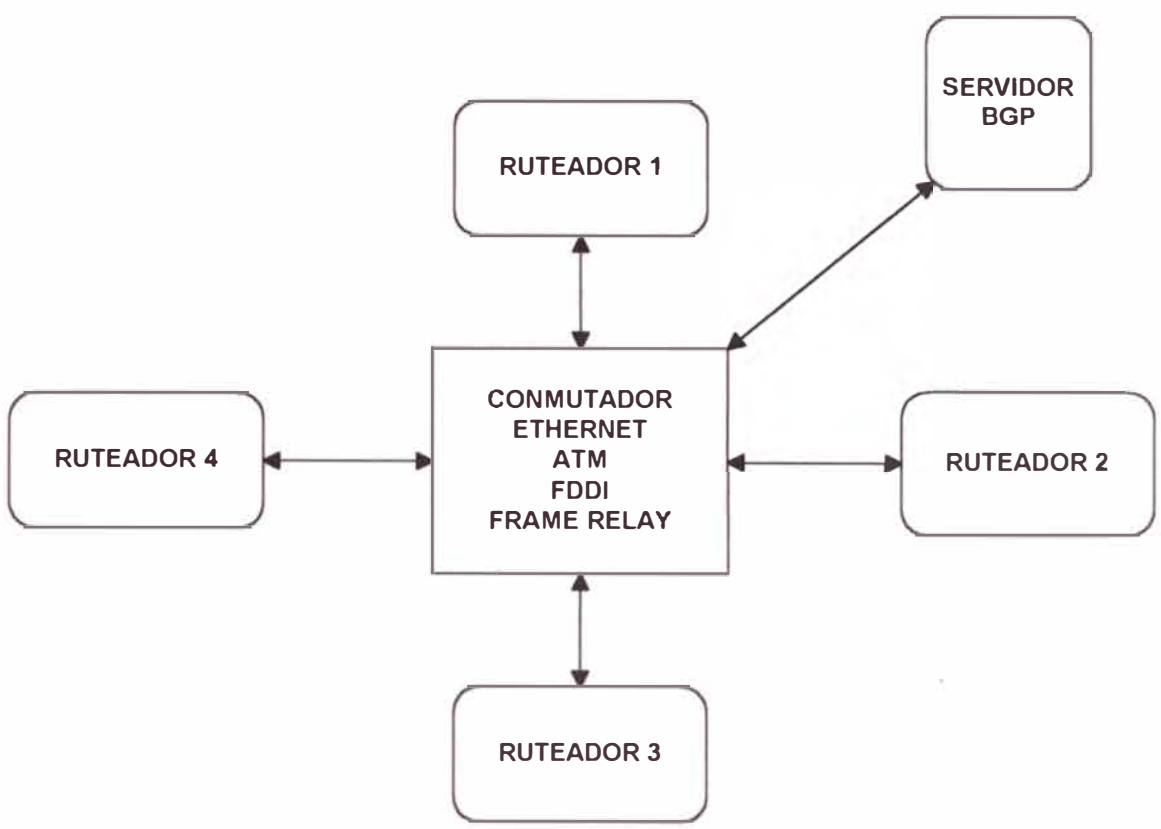
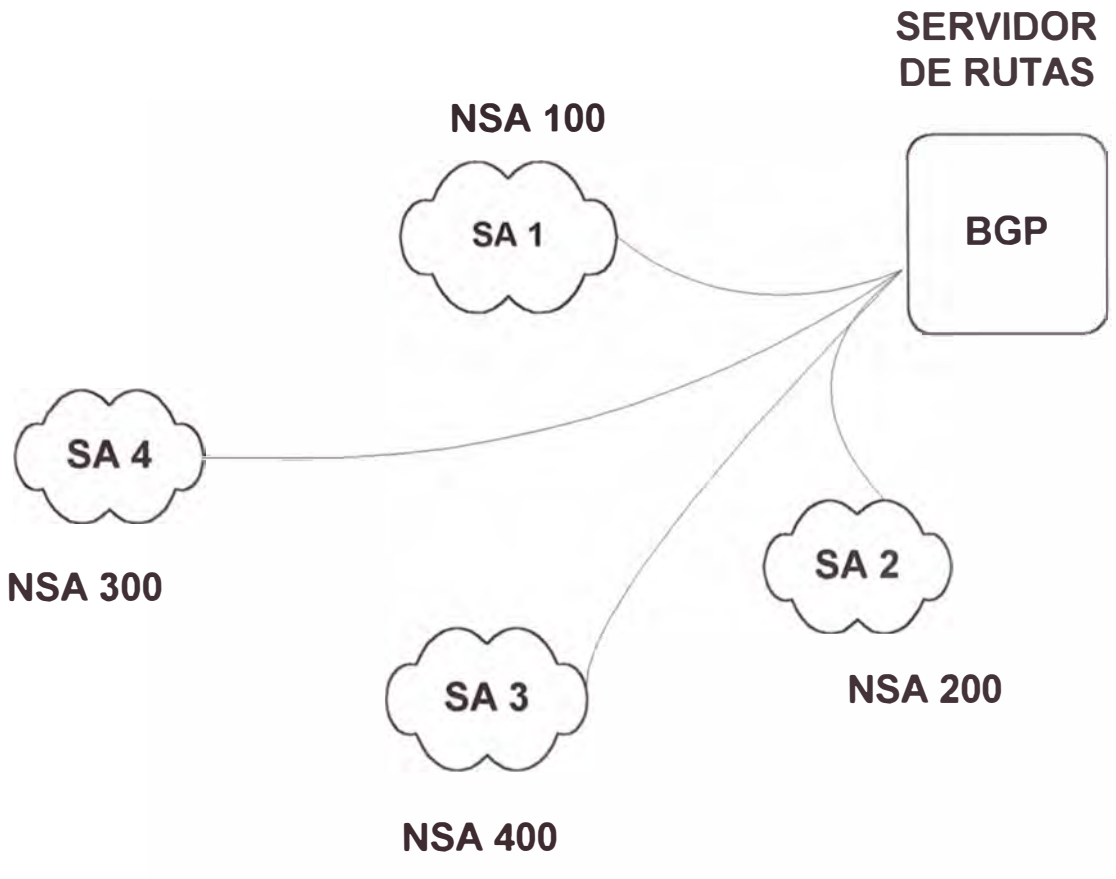


Fig 4.5 Modelo con Servidor BGP



- ADMINISTRACION CENTRALIZADA
- CONFIGURACION SIMPLIFICADA
- MINIMIZA NUMERO DE SESIONES BGP
- NUMERO DE SESIONES=NUMERO DE PARTICIPANTES

Fig 4.6 Topología BGP con Servidor

CONCLUSIONES

A lo largo del desarrollo del texto se han estudiado los aspectos técnicos involucrados en la Interconexión de redes IP, en forma general, para posteriormente, llegar como última etapa, al desarrollo específico de la interconexión de Sistemas Autónomos y su aplicación dentro del Nap Perú.

Es así como se han revisado los aspectos físicos de la interconexión de redes TCP/IP; describiendo la gama de tecnología disponible para este propósito en nuestro país. También se han desarrollado los aspectos principales e inherentes a las funciones del protocolo IP que hacen posible la intercomunicación de redes de diferentes tecnologías; hasta llegar al campo de los protocolos de enrutamiento IP especializados; indispensables para la interconexión.

En este punto; se mostró las posibilidades de 4 de los protocolos más utilizados; RIP, OSPF, EGP y finalmente abordamos en extenso el protocolo de enrutamiento BGP; así mismo en este punto; se desarrollo los distintos modelos que se pueden utilizar para interconectar Sistemas Autónomos en el entorno BGP-4 explicando los pro y contra de cada modelo, que no difieren en gran medida entre sí.

Con todo este material necesario; abordamos el tema específico de interconexión aplicada en el Nap Perú; explicando el entorno desarrollado en el proyecto y la operación del mismo. En este punto es oportuno destacar, que el grupo de especialistas, ingenieros y técnicos involucrados en el planeamiento, ejecución y operación del Nap Perú; son todos nacionales; los cuales han adquirido la experiencia y conocimiento necesarios en el país; quienes han hecho posible la ejecución de este proyecto muy importante para al desarrollo de Internet y las Telecomunicaciones en el Perú; este importante objetivo alcanzado, pone de manifiesto la alta capacidad de la Ingeniería de Redes peruana actual.

ANEXO 1

EL CONJUNTO DE NÚMEROS IP

El conjunto de números IP esta formado por números binarios de 32 cifras y se utiliza para identificar de manera única a cada nodo perteneciente a una red TCP/IP. La regla de oro que se debe cumplir inexorablemente dice que no existen 2 nodos que tengan el mismo número IP. Para facilitar la lectura de estos números; se acostumbra identificarlos por cuatro números decimales, separados por puntos; cada número representa un octeto binario.

Ejemplo de número IP: 192.168.100.55

Cada dirección IP se subdivide en 2 campos claramente definidos: el número de red y el número de nodo. Esta división dio pie a que se identificaran diferentes clases de redes de acuerdo al tamaño del número de red. Estas clases de redes son las siguientes:

Red de Clase A:

El número de red de esta clase esta formado por los 8 primeros dígitos donde el primer dígito es cero. Los siguientes 24 dígitos son asignados a los nodos

pertenecientes a esta red. Dadas estas condiciones podemos deducir que se tienen 127 redes clase A desde la 1 hasta la 127. Cada clase A puede contener 16777216 nodos. Ejemplo de número de esta clase:

126.45.78.115

Red de clase B

El número de red de esta clase esta formado por los 16 primeros dígitos y sus 2 primeros dígitos son fijados en 10; lo cual arroja 16384 redes con 65536 nodos para cada una de las redes de esta clase.

Ejemplo: 160.130.15.15

Red de clase C

Esta clase de red esta formada por 3 octetos de red y un octeto de nodos lo cual arroja 65536 diferentes redes con 256 nodos cada una. Comienza con los dígitos 110.

Ejemplo:

192.168.0.1

Máscara de Red

Cada nodo IP esta definido por su número de nodo y además por su número de máscara de red. Este número de máscara esta formado también por 4 octetos que se pueden escribir mediante 4 números decimales separados por puntos. Sin embargo el conjunto de números correspondiente a la máscara de red es mas limitado y esta asociado a la clase de red a la cual pertenece el nodo IP.

Clases de Máscara:

Existe una Máscara de red por cada clase. Estas son:

Máscara de clase A: 255.0.0.0

Máscara de clase B: 255.255.0.0

Máscara de clase C: 255.255.255.0

La finalidad de la máscara es identificar el número de red al cual pertenece un nodo. Esto se consigue mediante la multiplicación binaria entre la máscara de red y el número IP. Por ejemplo:

Se tiene el número 172.10.1.13 con máscara 255.255.0.0. :

Multiplicando ambos en binario se obtiene 172.10.0.0 que representa él número de red de este nodo.

Existen otras clases de números IP que son las clases D y E; las cuales tienen propósitos distintos a las clases anteriores, y no se utilizan para asignación de nodos de red.

ANEXO 2

REDES PRIVADAS Y SUBREDES

Redes Privadas

Con el propósito de uniformizar criterios de asignación de números para redes TCP/IP llamadas privadas; se decidió reservar bloques de direcciones IP para aquellas redes que no estuvieran conectadas a Internet. Estas son las siguientes:

10.0.0.0 con máscara 255.0.0.0

172.16.0.0 con máscara 255.255.0.0 hasta 172.31.0.0 con máscara 255.255.0.0

192.168.0.0 con máscara 255.255.255.0 hasta 192.168.255.0 con máscara 255.255.255.0

Entonces las redes que están conectadas a Internet directamente; no pueden utilizar ninguna de estas redes; solo deben emplearse en redes privadas. Por lo demás; estas redes privadas no son ruteables en Internet

Subredes

El concepto de subredes apareció cuando se vislumbró que las asignaciones de números IP agotarían la existencia de los mismos; pues se asignaban en razón de

clase de red completa, con lo cual se desaprovechaba una gran cantidad de números IP, en los casos en que la entidad solicitante no tuviera un númerode nodos a conectar que sea próximo al número total de nodos que la clase solicitada brinda.

Como ejemplo, en el caso que una entidad solicita una red clase A y actualmente solo tiene 1000 nodos en toda su red, eso significa que no se estarían usando 16776216 números IP.

Para permitir la utilización de los números IP que se quedan sin emplear; se utiliza el sistema de subredes. Este mecanismo permite que una red de cualquier clase; se pueda dividir en subconjuntos de redes o clases más pequeñas, haciendo uso de la máscara de red adecuada.

Por ejemplo, sea la red 172.18.0.0 con máscara 255.255.0.0, se puede subdividir esta red de Clase B, en 256 redes de tipo Clase C empleando la máscara 255.255.255.0. Veámoslo a continuación:

172.18.0.0 255.255.255.0, 172.18.1.0, 255.255.255.0 172.18.2.0, 255.255.255.0 así sucesivamente hasta obtener la última red 172.18.255.0, 255.255.255.0

Este mismo proceder se puede ampliar para obtener redes aun más pequeñas; cambiando la máscara a aplicar por cualquiera de las siguientes: 255.255.255.128, 255.255.255.192, etc.

A esta forma de enrutar redes más pequeñas se le conoce técnicamente como CIDR Classless Inter Domain Routing. Esta técnica se habilita en los ruteadores y se incorpora en los protocolos de enrutamiento para reutilizar el espacio de números IP que estaba subutilizado hasta comienzos de la década del 90.

Otra técnica que también se incorpora en los enrutadores para facilitar el uso de CIDR se conoce como VLSM Variable Length Subnet Mask. Esta técnica consiste

en permitir el uso de diferentes tipos de máscara de red para identificar y subdividir una sola red. Veámosla con un ejemplo:

Se tiene la red 192.168.1.0, 255.255.255.0 asignada a una red en particular mediante el mecanismo VLSM se puede aplicar otras longitudes de máscara para subdividir la red de la siguiente forma:

192.168.1.0 - 255.255.255.128

192.168.1.128 - 255.255.255.128

De esta manera obtenemos 2 subredes de 128 nodos cada una; usando la misma máscara de red que es de longitud menor a la anterior. Aplicando el mismo criterio podemos obtener subredes menores.

ANEXO 3

LISTA DE ESTADOS Y EVENTOS DE UN NODO BGP

Estados BGP:

1 - Idle

2 - Connect

3 - Active

4 - OpenSent

5 - OpenConfirm

6 - Established

Eventos BGP:

1 - BGP Start

2 - BGP Stop

3 - BGP Transport connection open

4 - BGP Transport connection closed

5 - BGP Transport connection open failed

6 - BGP Transport fatal error

7 - ConnectRetry timer expired

8 - Hold Timer expired

9 - KeepAlive timer expired

10 - Receive OPEN message

11 - Receive KEEPALIVE message

12 - Receive UPDATE message

13 - Receive NOTIFICATION message

Relación de Eventos y Estados en BGP

Evento	Idle 1	Connect 2	Active 3	OpentSent 4	OpenConf 5	Establish 6
1	2	2	3	4	5	6
2	1	1	1	1	1	1
3	1	4	4	1	1	1
4	1	1	1	3	1	1
5	1	3	3	1	1	1
6	1	1	1	1	1	1
7	1	2	2	1	1	1
8	1	1	1	1	1	1
9	1	1	1	1	5	6
10	1	1	1	1 , 5	1	1
11	1	1	1	1	6	6
12	1	1	1	1	1	1 , 6
13	1	1	1	1	1	1

BIBLIOGRAFÍA

1. **INTERNETWORKING CON TCP/IP VOLUMEN I PRINCIPIOS, PROTOCOLOS Y ARQUITECTURA** de Douglas Commer, editorial Prentice Hall, 1996.
2. **LIBRO DE INTERNET** de Douglas Commer, editorial Prentice Hall, 1995.
3. **SISTEMAS DE COMUNICACIONES Y REDES DE ORDENADORES** de John Freer, editorial Anaya Multimedia, 1990.
4. **INTERNET ROUTING ARCHITECTURES** de Bassam Halabi, editorial New Riders Publishing, 1997.
5. **GUÍA DE INTEROPERABILIDAD** de Tom Sheldom, editorial Mc Graw Hill, 1995.
6. **COMUNICACIONES Y REDES DE ORDENADORES** de William Stallings, editorial Prentice Hall, 1997.
7. **OSPF NETWORK DESIGN SOLUTIONS** de Thomas II Thomas, editorial Macmillan Technical Publishing, 1998.

SITIOS DE INTERNET

1. http://www.merit.edu
2. http://www.ietf.net
3. http://www.arin.net
4. http://www.cisco.com
5. http://www.gated.org
6. http://digex.nitrous.net