Apuntes

de

Redes de Ordenadores

Tema 9

Nivel de Red: IP

Uploaded by

IngTeleco

http://ingteleco.iespana.es
ingtelecoweb@hotmail.com

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

TEMA 9: INTERCONEXIÓN DE REDES: IP

9.1.- INTRODUCCIÓN

El protocolo IP es el componente de red del conjunto de protocolos TCP/IP, cuya función es asegurar el envío de datagramas a través de cualquier combinación de redes intermedias hasta su destino final. Sorprendentemente, y a diferencia de otros protocolos de red, el servicio que ofrece es un servicio de envío de datagramas no orientado a la conexión y no fiable.

Por no fiable entendemos que no garantiza que un datagrama IP llegue correctamente a su destino. No se produce ninguna confirmación por parte del destinatario de la recepción del datagrama ni se desarrolla ningún procedimiento de corrección de error. El grado de fiabilidad del servicio dependerá del proporcionado por la red subyacente. Cualquier fiabilidad añadida debe ser proporcionada por los niveles superiores (por ejemplo TCP).

El término no orientado a la conexión significa que IP no mantiene ninguna información de estado sobre los datagramas sucesivos. Cada datagrama se maneja de forma independiente del resto de los datagramas aunque puedan ser parte de un mismo mensaje; esto significa que los datagramas IP pueden llegar a su destino desordenados, faltar alguno de ellos e incluso llegar repetidos. Si un emisor envía dos datagramas consecutivos al mismo destino, cada uno de ellos es encaminado de forma independiente y puede utilizar diferentes rutas y llegar por lo tanto desordenados; esto a su vez supone una de las grandes ventajas de IP que es su flexibilidad y capacidad de adaptación.

La especificación oficial del protocolo IP (el RFC 791) proporciona tres definiciones importantes:
- La unidad básica de transferencia de datos utilizada en redes TCP/IP (formato de los datagramas).
- Las funciones de encaminamiento que desarrolla IP, eligiendo el camino que seguirá un datagrama hasta alcanzar su destino.
- Las reglas acerca de envío de datagramas (no fiables), que recogen el modo en que los ordenadores y las pasarelas deben procesar los datagramas, cómo y cuándo deben generarse mensajes de error, y las condiciones bajo las cuales deben descartarse los datagramas.

La Internet es un compendio de redes diferentes que comparten un protocolo, o pila de protocolos comunes (IP a nivel de red y sobre todo TCP a nivel de transporte); cada una de estas redes es administrada por una entidad diferente: universidades, redes académicas nacionales, proveedores comerciales (también llamados ISPs, Internet Service Providers), operadores, multinacionales, etc. Como consecuencia de esto las políticas de uso son muy variadas.

Técnicamente a nivel de red la Internet puede definirse como un conjunto de redes o *sistemas autónomos* conectados entre sí que utilizan el protocolo de red IP.
IP es una red de datagramas, no orientada a conexión, con calidad de servicio 'best effort', es decir, no hay calidad de servicio (QoS); no se garantiza la entrega de los paquetes ya que en momentos de congestión éstos pueden ser descartados sin previo aviso por los routers que se encuentren en el trayecto.

9.2.- EL DATAGRAMA IP

Toda información en una red IP ha de viajar en datagramas IP. Esto incluye tanto las TPDUs (Transport Protocol Data Units) de TCP y UDP, como cualquier información de routing que se intercambie en la red (paquetes ECHO, HELLO, PRUNE; de asfixia, etc.).

El tamaño de un datagrama IP se especifica en un campo de dos bytes por lo tanto su valor máximo es de 65535 bytes, sin embargo, muy pocas redes admiten este valor. Normalmente el nivel de enlace de datos no se encarga de fragmentar, por lo que el nivel de red debe adaptar el tamaño de cada paquete para que viaje en una trama; con lo que en la práctica el tamaño máximo de paquete viene determinado por el tamaño máximo de trama característico de la red utilizada. Este tamaño máximo de paquete se conoce como MTU (Maximum Transfer Unit); a continuación damos algunos ejemplos de valores de MTU característicos de las redes más habituales:

Protocolo a nivel de enlace	**MTU**
PPP (valor por defecto)	1500
PPP (bajo retardo)	296
SLIP	1006 (límite original)
X.25	1600 (varía según las redes)
Frame relay	al menos 1600 normalmente
SMDS	9235
Ethernet version 2	1500
IEEE 802.3/802.2	1492
IEEE 802.4/802.2	8166
Token Ring IBM 16 Mb/s	17914 máximo
IEEE 802.5/802.2 4 Mb/s	4464 máximo
FDDI	4352
Hyperchannel	65535
ATM	9180

Tabla 9.1.- Valor de MTU para los protocolos mas comunes a nivel de enlace.

Es bastante normal utilizar 1500 como valor de MTU (teniendo en cuenta la proporción actual de redes basadas en tecnología Ethernet). RFC 791 especifica que cualquier red debe soportar como mínimo un MTU de 68 bytes.

El datagrama tiene dos partes: cabecera y texto. La estructura de la cabecera es la que se muestra en la Figura 9.1; tiene una parte fija de 20 bytes y una opcional de entre 0 y 40 bytes (siempre múltiplo de 4).

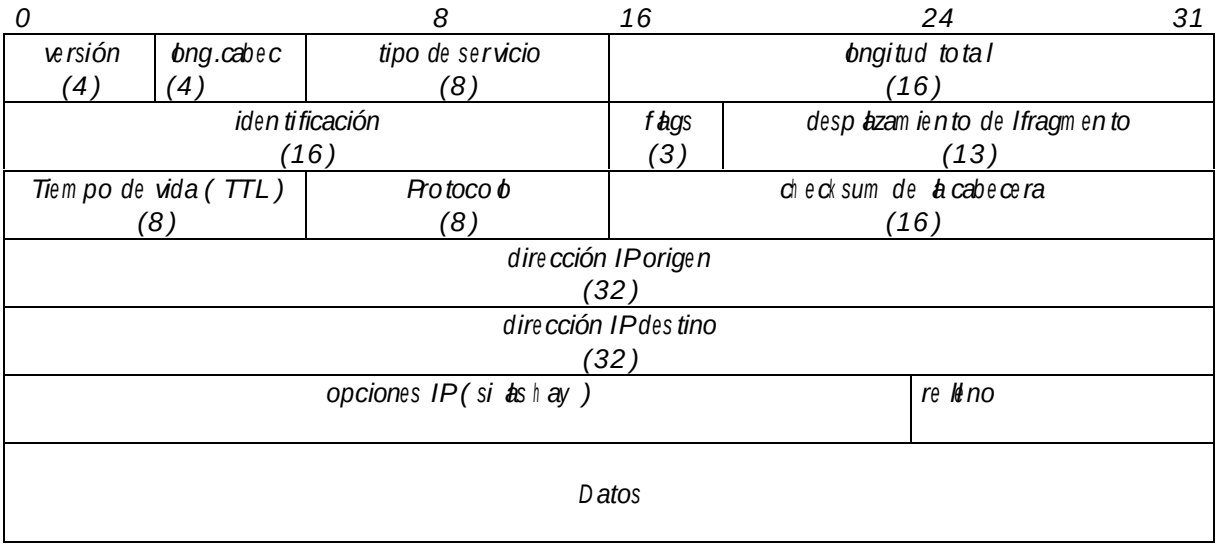

Figura 9.1.- Estructura de la cabecera de un datagrama IP

El campo **versión** permite que coexistan en la misma red y sin ambigüedad datagramas IP de distintas versiones; la versión actualmente utilizada de IP (correspondiente a la estructura de datagrama indicada en la figura anterior) es la 4. Como veremos mas tarde, se empieza a extender el uso de una nueva versión (V6) con una estructura de datagrama diferente.

El campo **longitud de cabecera** especifica ésta medida en palabras de 32 bits, incluyendo el campo de opciones. El valor habitual de este campo, cuando no se utilizan opciones es 5 y su valor máximo es 15, que equivale a 40 bytes de información opcional. Esta limitación supone, en la práctica, la inutilidad de algunas opciones definidas en IP (como la opción de Registro de Ruta). La longitud de la cabecera siempre ha de ser un número entero de palabras de 32 bits, por lo que si la longitud de los campos opcionales no es un múltiplo exacto de 32 bits se utiliza un campo de relleno al final de la cabecera.

El campo **Tipo de Servicio** (TOS) tiene la siguiente estructura:

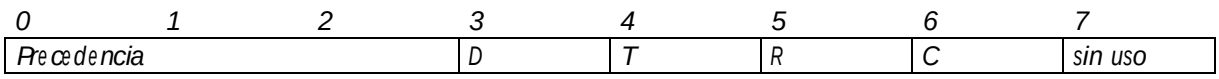

Figura 9.2.- Estructura del campo 'Tipo de servicio'

Los bits de precedencia permiten especificar una prioridad entre 0 y 7 para cada datagrama, pudiendo así indicar la importancia o urgencia de los mismos. Por ejemplo marcando los paquetes normales con prioridad 0 y los "paquetes de asfixia" con prioridad 7. La prioridad puede actuar alterando el orden de los paquetes en cola en los routers, pero no modifica la ruta de éstos. Dada la actual abundancia de nodos gestionados por el usuario final, existe un gran debate sobre la conveniencia de la existencia de un campo prioridad, ya que el usuario podría descubrir que obtiene mejor servicio con alta prioridad y utilizar sistemáticamente el valor 7 para todo tipo de paquetes; en la práctica muchos equipos ignoran este campo y cuando hacen uso de él es únicamente para datagramas transmitidos desde dentro de la subred (es decir, entre routers), que se supone que están libres de esta sospecha.

Los cuatro bits siguientes actúan como flags denominados D, T, R y C respectivamente tienen el siguiente significado:

- D: minimizar el retardo (D=Delay)
- T: maximizar el rendimiento (T=Throughput)
- R: maximizar la fiabilidad (R=Reliability)
- C: minimizar el coste monetario (C=Cost)

Sólo uno de estos cuatro bits puede estar activado, y si todos los bits se encuentran a cero se supone un servicio normal. Los documentos RFC1340 y RFC1349 especifican el modo en que deberían usarse estos bits en todas las aplicaciones estándares, por ejemplo, para telnet se recomienda 1000 (mínimo retardo), para FTP 0100 (máximo rendimiento) y para NNTP (news) 0001 (mínimo costo).

Algunos routers utilizan el subcampo TOS para encaminar los paquetes por la ruta óptima en función del valor especificado (podrían tener una ruta diferente según se desee mínimo retardo o mínimo costo, por ejemplo); también pueden utilizar el valor del campo TOS para tomar decisiones sobre que paquetes descartar en situaciones de congestión (por ejemplo descartar antes un paquete con mínimo costo que uno con máxima fiabilidad). Muchos routers simplemente ignoran este subcampo.

El campo **Longitud Total** especifica la longitud del datagrama completo (cabecera incluida) medida en bytes. Como el campo es de 16 bits, el tamaño máximo de un datagrama IP es de 65.535 bytes.

Si bien es posible enviar datagramas IP de 65.535 octetos la mayoría de los niveles de enlace de datos los fragmentarán. De hecho un sistema conectable a Internet no está obligado a recibir datagramas mayores de 576 octetos.

El campo longitud total es necesario en la cabecera IP porque algunos protocolos del nivel de enlace de datos (por ejemplo Ethernet) rellenan las tramas pequeñas para alcanzar una longitud mínima.

Si el campo longitud total no existiera, el nivel IP no sabría qué parte de una trama Ethernet de 46 octetos sería realmente un datagrama IP.

El campo **Identificación** sirve para identificar de forma única cada uno de los datagramas enviados por un nodo. Normalmente se incrementa en uno cada vez que se envía un datagrama, pero no debe confundirse con un número de sencuencia. Permite al receptor reconocer las partes correspondientes a un mismo datagrama en caso de que éste se haya fragmentado.

La utilización y significado de los siguientes bits correspondientes a los campos **Flags** y **Desplazamiento del Fragmento** se analizan en apartado de la fragmentación de los datagramas IP.

El campo **Tiempo de Vida** (TTL) permite descartar un datagrama cuando ha pasado un tiempo excesivo viajando por la red y es presumiblemente inútil. En el diseño original se pretendía que el valor de este campo (que inicialmente podía valer por ejemplo 64) se decrementara en cada router en un valor igual al tiempo en segundos que el paquete había empleado en esa parte del trayecto, restando como mínimo 1 en cualquier caso. En la práctica medir tiempos en una red es mucho más difícil de lo que parece (los relojes de los routers han de estar muy bien sincronizados, cosa que hoy en día no ocurre), por lo que todas las implementaciones se limitan sencillamente a restar 1 al valor de TTL de cada paquete que pasa por ellos, sin analizar el tiempo que el paquete ha invertido en el salto. Como de cualquier forma hoy en día es muy raro que un paquete tarde más de un segundo en cada salto esto está aproximadamente de acuerdo con el diseño original. El valor inicial de TTL de un paquete fija el número máximo de saltos que podrá dar, y por tanto debería ser suficientemente grande como para que pueda llegar a su destino. Cuando este campo alcanza el valor 0, el router que reciba el datagrama lo descartará, notificando su eliminación al emisor mediante el envío de un mensaje ICMP. Este procedimiento evita que los datagramas puedan entrar en rutas circulares y permanecer indefinidamente en la red en caso de corrupción de las tablas de encaminamiento de los routers.

El campo **Protocolo** especifica a que protocolo del nivel de transporte corresponde el datagrama. La tabla de protocolos válidos y sus correspondientes números son controlados por el IANA (Internet Assigned Number Authority) y se especifican en un RFC denominado 'Assigned Numbers', que se actualiza regularmente; el vigente actualmente es el RFC 1700. Algunos de los posibles valores del campo protocolo son los siguientes:

Valor	Protocolo	Descripción
0		Reservado
1	ICMP	Internet Control Message Protocol
2	IGMP	Internet Group Management Protocol
3	GGP	Gateway-to-Gateway Protocol
4	IP	IP en IP (encapsulado)
5	ST	Stream
6	TCP	Transmission Control Protocol
8	EGP	Exterior Gateway Protocol
17	UDP	User Datagram Protocol
29	ISO-TP4	ISO Transport Protocol Clase 4
38	IDRP-CMTP	IDRP Control Message Transport Protocol
80	ISO-IP	ISO Internet Protocol (CLNP)
88	IGRP	Internet Gateway Routing Protocol (Cisco)
89	OSPF	Open Shortest Path First
255		Reservado

Tabla 9.2.- Ejemplo de valores y significados del campo protocolo en un datagrama

Obsérvese, como curiosidad, que el valor 4 está reservado al uso de IP para transportar IP, es decir al encapsulado de un datagrama IP dentro de otro.

El campo **checksum** sirve para detectar errores producidos en la cabecera del datagrama; no es un CRC sino el complemento a uno en 16 bits de la suma complemento a uno de toda la cabecera (incluidos los campos opcionales si los hubiera), tomada en campos de 16 bits; para el cálculo el campo checksum se pone a sí mismo a ceros. Este campo permite salvaguardar a la red de un router

que alterara los campos de cabecera de un datagrama, por ejemplo por un problema hardware. El campo checksum se ha de recalcular en cada salto, ya que al menos el TTL cambia. Esto supone un serio inconveniente desde el punto de vista de rendimiento en routers con mucho tráfico.

ICMP, IGMP, UDP y TCP utilizan todos el mismo algoritmo de Checksum, si bien TCP y UDP incluyen varios campos de la cabecera IP a la hora de calcularlo. En el RFC 1071 pueden encontrarse las técnicas de implementación del campo Checksum de Internet. Dado que un router generalmente sólo cambia el campo TTL (decrementandolo en 1) cuando envía un datagrama recibido puede actualizar de forma incremental el Checksum en lugar de efectuar el cálculo del mismo con la cabecera completa (RFC 1141).

Los campos **dirección de origen** y **dirección de destino** corresponden a direcciones IP según el formato que veremos más adelante.

9.3.- FRAGMENTACIÓN Y REENSAMBLADO DE DATAGRAMAS IP

El protocolo IP no limita el tamaño mínimo de los datagramas, ni garantiza que datagramas mayores sean enviados sin fragmentación. El emisor puede elegir cualquier tamaño de datagrama que crea apropiado; la fragmentación y el reensamblado se producen de forma automática. Cada uno de los fragmentos tiene el mismo formato que el datagrama original con una cabecera que duplica la mayor parte de la información de la cabecera original.

Una vez que un datagrama ha sido fragmentado, los fragmentos viajan como datagramas separados hasta el destino final, donde son reensamblados. Esto supone que, una vez que el datagrama ha sido fragmentado al atravesar una red con una MTU pequeña, los fragmentos viajarán de forma independiente a través de las redes situadas a continuación de aquella aún cuando dichas redes pudieran tener capacidad para transportar fragmentos mayores, lo cual producirá ineficiencias. Por otro lado si alguno de los fragmentos se pierde el datagrama no podrá ser reensamblado; para controlar esto último, el receptor arranca un temporizador de reensamblado cuando recibe el primero de los fragmentos, si el temporizador expira antes de que lleguen todos los fragmentos el sistema descartará los fragmentos recibidos sin procesar el datagrama. La probabilidad de que se pierda un datagrama aumenta cuando se produce fragmentación dado que la perdida de un fragmento supone la perdida del datagrama completo.

La fragmentación se controla mediante tres campos de la cabecera del datagrama IP:

- **Identificación**
- **Flags**
- **Desplazamiento del fragmento.**

El campo Flags contiene tres bits que que intervienen en el proceso de fragmentación:

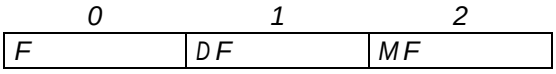

Figura 9.3.- Estructura del campo Flags

El bit **F**, cuando está a 1 indica que el datagrama es en realidad un fragmento.

El bit **DF** (Don't Fragment) cuando está a 1 indica a los routers que no deben fragmentar el datagrama para reenviarlo, ya que el receptor no está capacitado para reensamblarlo. Por ejemplo, si un ordenador arranca su sistema operativo a través de la red solicitará que el ejecutable correspondiente se le envíe desde algún servidor a través de la red como un único datagrama (ya que en ese estado él aun no está capacitado para reensamblar datagramas). Si un datagrama con el bit DF puesto no puede pasar por una red el router lo rechazará con un mensaje de error al emisor. Existe una técnica para averiguar el MTU de una ruta (denominada 'path MTU discovery') que consiste en enviar un datagrama grande con el bit DF puesto al destino deseado; si se recibe un mensaje de error se envía otro mas pequeño, hasta que el emisor averigua a base de tanteos cual es el valor de MTU de la ruta correspondiente, y a partir de ahí puede utilizarla para todos los datagramas sin riesgo de que sean fragmentados en el camino (siempre y cuando la ruta no cambie sobre la marcha).

El bit **MF** (More Fragments) a 1 especifica que este datagrama es realmente un fragmento de un datagrama mayor, y que no es el último. Si está a 0 indica que este es el último fragmento (o bien que el datagrama original no esta fragmentado).

El campo **Desplazamiento del Fragmento** sirve para indicar, en el caso de que el datagrama sea un fragmento de un datagrama mayor, en que posición del datagrama mayor empieza este fragmento. La fragmentación se realiza siempre en unidades de tamaño múltiplo de 8 bytes (la unidad elemental de fragmentación), por lo que este campo en realidad cuenta los bytes de 8 en 8. Al ser su longitud de 13 bits el número máximo de fragmentos es de 8192, que da cabida a la longitud máxima de un datagrama (8192 x 8 = 65536). Los fragmentos pueden llegar desordenados, por lo que el último fragmento puede llegar al receptor sin que haya recibido aun todos los fragmentos; la información *desplazamiento del fragmento* junto con *longitud* del último fragmento (identificado porque tiene el bit MF a 0) le permite al receptor calcular la longitud total del datagrama original (que sería *fragment_offset*8 + longitud*). Supongamos que construimos un datagrama de 4000 bytes de datos (es decir 4000 bytes sin contar la cabecera IP) que ha de pasar por una red cuya MTU es de 1500 bytes (PPP por ejemplo); el resultado de la fragmentación sería el siguiente:

Paquete Original:	Id=X	L=4020	DF=0	MF=0	Offset=0

Fragmento 1:	Id=X	L=1500	DF=0	MF=1	Offset=0
Fragmento 2:	Id=X	L=1500	DF=0	MF=1	Offset=185
Fragmento 3:	Id=X	L=1060	DF=0	MF=0	Offset=370

Una vez que un datagrama es fragmentado ya no será unido nuevamente hasta que llegue a su destino. Puede suceder que un datagrama tenga que ser fragmentado en ruta a un tamaño máximo determinado y mas tarde haya de pasar por otra red de MTU aun menor, con lo que tendrá que ser fragmentado de nuevo; por ejemplo si el fragmento 2 anterior ha de pasar por una red con MTU de 296 bytes (PPP con bajo retardo) el resultado será:

Fragmento 2 Original:	Id=X	L=1500	DF=0	MF=1	Offset=185

Fragmento 2a:	Id=X	L=292	DF=0	MF=1	Offset=185
Fragmento 2b:	Id=X	L=292	DF=0	MF=1	Offset=219
Fragmento 2c:	Id=X	L=292	DF=0	MF=1	Offset=253
Fragmento 2d:	Id=X	L=292	DF=0	MF=1	Offset=287
Fragmento 2e:	Id=X	L=292	DF=0	MF=1	Offset=321
Fragmento 2f:	Id=X	L=140	DF=0	MF=1	Offset=355

Obsérvese que aunque la MTU es en este caso de 296 los datagramas generados nunca son mayores de 292 bytes por la condición de que la parte de datos de los fragmentos siempre debe ser múltiplo de 8 bytes. Al final solo el último fragmento de todos los que lleguen al receptor tendrá puesto a 1 el bit MF.

9.4.- OPCIONES DEL DATAGRAMA IP

Los campos opcionales de la cabecera no siempre están soportados en los routers y se utilizan muy raramente; de estos podemos destacar los siguientes:

Registro de ruta: Esta opción pide a cada router por el que pasa este datagrama que anote en la cabecera su dirección, con lo que al recibir el datagrama se dispondrá de una traza de la ruta seguida para fines de prueba o diagnóstico de problemas. Debido a la limitación en la longitud de la cabecera como máximo pueden registrarse 9 direcciones, lo cual es insuficiente en algunos casos.

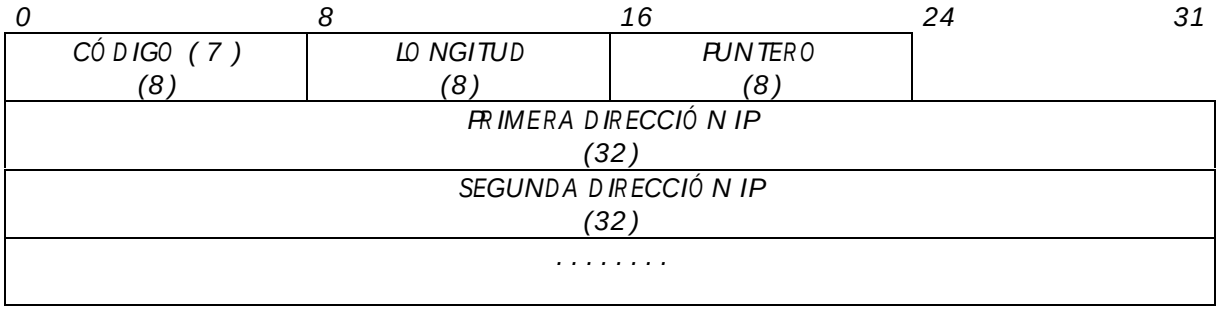

Figura 9.4.- Formato de la opción Registro de Ruta

El campo **código** identifica la opción (7). El campo **longitud** especifica la longitud total de la opción tal y como aparece en el datagrama IP, incluyendo los primeros 3 octetos; El campo **puntero**, es un índice que muestra donde almacenar la siguiente dirección IP; su valor mínimo es 4, que es el puntero de la primera dirección IP. A medida que se registran nuevas direcciones en la lista, el valor del campo puntero aumenta para indicar dónde comenzaría la siguiente dirección a registrar. Para añadir su propia dirección a la lista, el sistema compara en primer lugar los campos puntero y longitud, si el puntero es mayor que la longitud, la lista está llena, y por lo tanto el sistema reencamina el datagrama sin insertar su dirección. Si la lista no está llena, el sistema insertará su dirección IP en la posición especificada por el campo puntero e incrementará en 4 dicho campo.

Cuando llega un datagrama, el sistema de destino debe extraer y procesar la lista de direcciones IP. Habitualmente los sistemas ignoran la ruta registrada, es por tanto evidente que hace falta un acuerdo entre emisor y receptor para que las rutas se registren y para que estás sean analizadas.

Encaminamiento desde el Origen: permite al emisor especificar la ruta que debe seguir el datagrama hasta llegar a su destino. Por ejemplo, para probar el rendimiento de una red particular, los administradores del sistema pueden utilizar el encaminamiento desde origen para forzar que los datagramas IP atraviesen dicha red, aún cuando los routers normalmente elegirían un camino distinto.

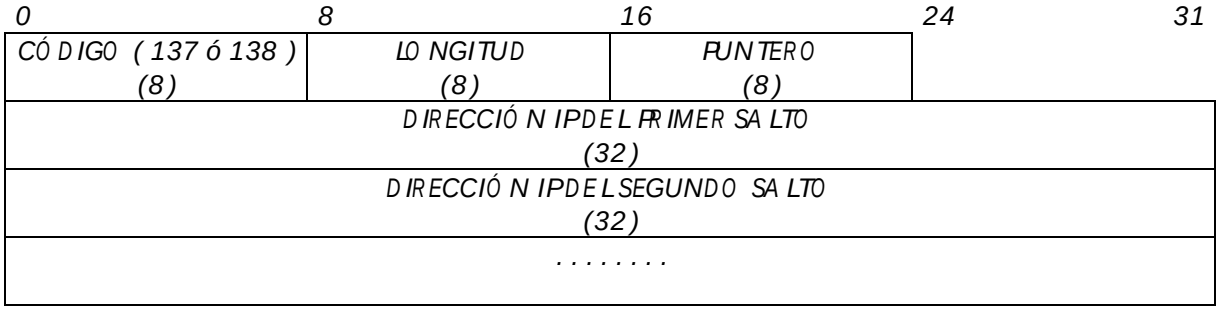

Figura 9.5 Formato de la opción Encaminamiento desde el Origen

Existen dos variantes: *encaminamiento estricto desde el origen*, que permite especificar la ruta exacta salto a salto, de modo que si algún paso de la ruta no es factible por algún motivo se produzca un error. El *Encaminamiento flexible desde el origen* no precisa detallar todos los saltos, pudiendo existir saltos intermedios no especificados entre dos direcciones sucesivas de la lista.

El formato de una opción de encaminamiento desde el origen es una replica de la opción de registro de ruta. Cada router que reencamina un datagrama examina los campos longitud y puntero para ver si la lista se ha agotado. Si ha sido así, el puntero será mayor que la longitud y el router encamina el datagrama hacia su destino como habitualmente. Si la lista no está agotada, el router tomará la siguiente dirección IP indicada por el puntero y encaminará el datagrama utilizando dicha dirección.

Timestamp: esta opción actúa de manera similar a registro de ruta, pero además de anotar la dirección IP de cada router atravesado se anota en otro campo de 32 bits el instante en que el datagrama pasa por dicho router. El uso de dos campos de 32 bits acentúa aún mas el problema antes mencionado del poco espacio disponible para grabar esta información.

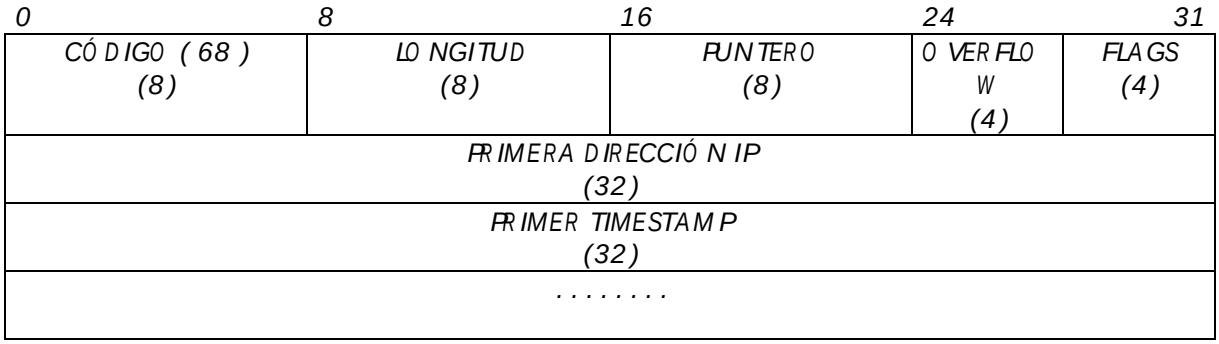

Figura 9.6 Formato de la opción Timestamp

Como en las opciones anteriores, los campos código, longitud y puntero se utilizan para especificar la clase y número de opción, el espacio reservado para la opción y la posición del siguiente elemento no utilizado. El campo **overflow** indica el número de routers que no han podido registrar su hora debido a que la opción era demasiado pequeña. Finalmente, los valores del campo **flags** controlan el formato exacto de la opción e indican el modo en que debe proporcionar su registro un router. Sus valores pueden ser los siguientes:

flags	Significado
0	Registrar solo la hora, omitir las direcciones IP
1	Anteponer a cada hora una dirección IP
3	Las direcciones IP son especificadas por el emisor; un router solo registra su hora si la siguiente dirección IP de la lista coincide con la suya.

Tabla 9.3

Cada registro proporciona la hora y fecha en que el router recibió el datagrama, expresado en milisegundos a partir de media noche, UTC. Si este formato de representación no está disponible en el router, puede emplear cualquier representación de tiempo, pero debe activar entonces el bit de orden más elevado del Timestamp para indicar que es un valor no estándar.

La limitaciones de la opción de registro de ruta son todavía peores en el caso de esta opción. Si registramos direcciones IP y hora del sistema, sólo podremos almacenar 4 de estos pares. La opción de registrar la hora sin la dirección IP es poco menos que inútil porque no nos proporciona ninguna información útil de a que corresponde cada registro.

Cuando se fragmenta un datagrama, el cada router replica a unas opciones IP en todos lo fragmentos mientras que pone otras sólo en el primero. Esto se controla mediante el **bit de copia** del campo código, especificando si debe copiarse en todos o solo uno de los fragmentos.

Consideremos por ejemplo la opción de registro de ruta, dado que cada fragmento es tratado de forma independiente no existe ninguna garantía de que todos los fragmentos vayan a seguir el mismo camino hasta su destino. Si todos los fragmentos contuvieran la opción de registro de ruta, el destinatario recibiría diferentes rutas en cada fragmento. Esto no produciría una ruta única para el datagrama reensamblado, es por ello que IP especifica que la opción registro de ruta solo debe copiarse en uno de los fragmentos.

La opción de encaminamiento desde el origen, sin embargo, que especifica cómo debe viajar un datagrama a través de la Internet, es sin embargo necesaria en todos los fragmentos, para que todos ellos sigan la misma ruta.

9.5.- DIRECCIONAMIENTO IP

Cuando se estandarizó IP en Septiembre de 1981, la especificación requería que a cada sistema conectado a una red basada en IP le fuera asignada una dirección de 32 bits única. Si un nodo dispone de varias interfaces físicas (cosa habitual en los routers) cada una de ellas deberá tener necesariamente una dirección IP distinta si se desea que sea accesible para este protocolo. Es posible además, y en algunas situaciones resulta útil, definir varias direcciones IP asociadas a una misma interfaz física.

Las direcciones IP tienen una estructura jerárquica de direccionamiento de dos niveles, tal y como aparece en la Figura 9-7. La primera parte de la dirección identifica la red a la que está conectado el sistema, mientras que la segunda parte identifica el nodo ("host") particular dentro de dicha red.

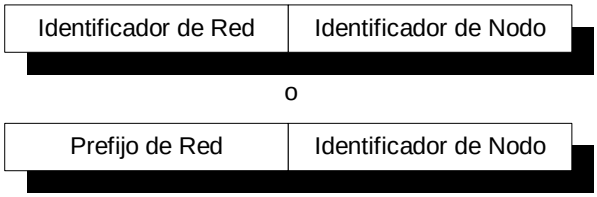

Figura 9.7: Estructura de las Direcciones IP

Recientemente, se ha comenzado a denominar el campo Identificador de Red como "prefijo de red". Todos los nodos de una red dada comparten el mismo prefijo de red, si bien deben tener un Identificador de Nodo único. De igual modo, dos nodos que se encuentren en redes distintas deben tener prefijos de red diferentes, aunque puedan compartir el mismo Identificador de Nodo.

9.5.1.- Notación decimal y máscara

Para hacer más sencilla la interpretación de las direcciones IP a los humanos, generalmente se expresan mediante cuatro números decimales separados por puntos. Esta notación divide las direcciones de 32 bits en cuatro campos de 8 bits cada una y especifica el valor de cada campo de manera independiente como un número decimal, separando los campos mediante puntos. La Figura 9-8 muestra cómo se expresa una dirección en notación decimal.

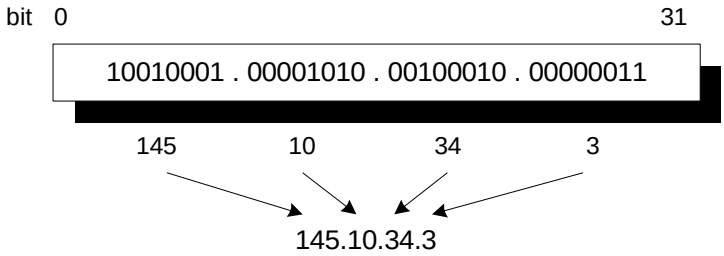

Figura 9-8: Notación Decimal

Tradicionalmente, para identificar la separación entre el prefijo de red y la identificación de nodo se ha utilizado una notación denominada **máscara**, consistente en poner a 1 los bits que corresponden al prefijo de red y a 0 los que corresponden al identificador de nodo. Así, por ejemplo, la máscara correspondiente a una dirección en la cual el prefijo de red sean los primeros 16 bits y el identificador de nodo los últimos 16 la máscara sería 255.255.0.0.

9.5.2.- Clases Primarias

En el diseño inicial de la Internet se reservaron los ocho primeros bits para la red, dejando los 24 restantes para el nodo; se creía que con 254 redes habría suficiente para una red experimental que era fruto de un proyecto de investigación del Departamento de Defensa americano. Sin embargo, ya desde el principio se vio que esto resultaba insuficiente, por lo que se reorganizó el espacio de direcciones reservando una parte para poder definir redes más pequeñas. Para dar mayor flexibilidad

y permitir diferentes tamaños se optó por dividir el rango de direcciones en tres partes adecuadas para redes grandes, medianas y pequeñas, conocidas como redes de clase A, B y C respectivamente. En la literatura aparece referenciado este direccionamiento como “classful addressing” puesto que el espacio de direcciones está dividido en clases, grupos o categorías predefinidas. Cada clase fija el límite entre el prefijo de red y el identificador de nodo en un punto diferente dentro de la dirección de 32 bits. Los formatos de las direcciones pertenecientes a estas clases es el representado en la Figura 9-9.

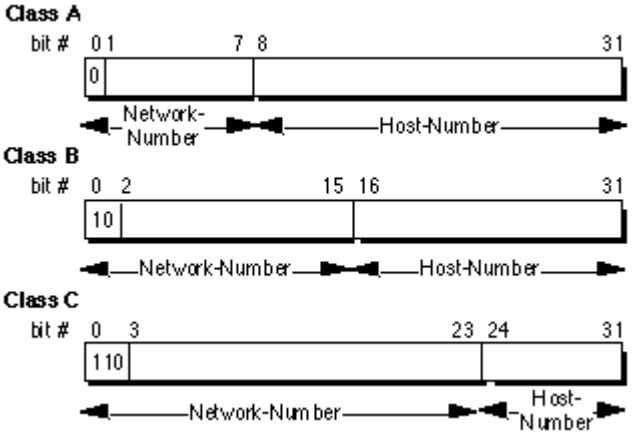

Figura 9-9: Estructura de las Direcciones IP

Uno de los aspectos fundamentales de este direccionamiento IP es que cada dirección contiene la clave que identifica el punto de separación entre el prefijo de red y el identificador de nodo. Esto simplificó el sistema de encaminamiento en los primeros años de Internet de modo que los protocolos de encaminamiento originales no necesitaban información adicional para separar el prefijo de red.

9.5.2.1.- Redes de Clase A (Prefijo /8)

Una red de clase A (que corresponde a las redes originalmente diseñadas) se caracteriza por tener a 0 el primer bit de dirección (el bit de orden más alto). El prefijo de red ocupa los 7 bits siguientes y el identificador de nodo los últimos 24 bits. En la actualidad se hace referencia a las direcciones de red de Clase A como “ /8s ” dado que tienen un prefijo de red de 8 bits.

Pueden definirse un total de 126 (2^7 -2) redes de clase A. Este número se debe a que existen dos identificadores de red que deben reservarse; La red 0.0.0.0 se utiliza como ruta por defecto y la red 127.0.0.0 para la función de “loopback” (más adelante se verá el significado de estos dos conceptos). Cada una de estas redes soporta un máximo de 16,777,214 (2^{24} -2) nodos. Igualmente, este cálculo contempla que los identificadores de nodo todo-0 ("esta red") y todo-1 ("difusión") no pueden ser asignados a nodos individuales.

Dado que los bloques de direcciones /8 contienen 2^{31} (2,147,483,648) direcciones individuales y que el espacio de direcciones IPv4 contiene un máximo de 2^{32} (4,294,967,296) direcciones, el espacio de direcciones /8 es 50% del espacio total de direcciones unicast de IPv4.

9.5.2.2.- Redes de Clase B (Prefijo /16)

Una dirección de clase B tiene un prefijo de red de 16 bits. Los dos bits de mayor orden tienen el valor 1-0, el campo red ocupa los 14 bits siguientes, y el identificador de nodo los 16 últimos bits. Se hace referencia a las direcciones de clase B como “ /16 “, dado que su prefijo de red es de 16 bits.
Puede haber 16382 redes clase B con 65534 nodos cada una.

Pueden definirse un máximo de 16,384 (2^{14}) redes, cada una de ellas con un total de 65,534 (2^{16} -2) nodos. Puesto que el bloque de direcciones /16 contiene 2^{30} (1,073,741,824) direcciones, representa el 25% del espacio total de direcciones unicast de IPv4.

9.5.2.3.- Redes de Clase C (Prefijo /24)

Una dirección de clase C tiene un prefijo de red de 24 bits. Los tres bits de mayor orden tienen el valor 1-1-0, el campo red ocupa los 21 bits siguientes, y el identificador de nodo los 8 últimos bits. Las redes de clase C se conocen también como " /24 ", dado que su prefijo de red es de 24 bits.

Pueden haber un máximo de 2,097,152 (2^{21}) redes /24, hasta con 254 (2^8 -2) nodos por red. Como el bloque de direcciones /24 contiene 2^{29} (536,870,912) direcciones, representa el 12.5% del espacio total de direcciones unicast de IPv4.

9.5.2.4.- Otras Clases

Además de las clases anteriores, existen dos clases adicionales.

- Las direcciones (no redes) de clase D, cuyos primeros cuatro bits de mayor orden tienen el valor 1-1-1-0, que se utilizan para dar soporte al multicasting IP. En estas direcciones, el grupo multicast viene definido por los 28 bits siguientes.
- Por último, las direcciones de clase E, cuyos primeros bits tiene el valor 1-1-1-1-0, y está reservada para usos futuros.

9.5.2.5.- Resumen

De los valores de los primeros bits de cada una de las clases antes mencionadas se puede deducir el rango de direcciones que corresponde a cada una de ellas, aunque también suele utilizarse la máscara para identificar la porción de la dirección que corresponde al prefijo de red y al identificador de nodo. Así pues, en la práctica es inmediato saber a que clase pertenece una dirección determinada sin más que saber el primer byte de su dirección. La Tabla 9-4 resume toda la información esencial sobre los tipos de direcciones de Internet.

Clase	Primeros Bits	Bits red/nodo	Nº Redes	Nº Nodos	Máscara	Rango de direcciones
A /8	0	7/24	126	16777214	255.0.0.0	1.xxx.xxx.xxx 127.xxx.xxx.xxx
B /16	10	14/16	16384	65534	255.255.0.0	128.0.xxx.xxx 191.255.xxx.xxx
C /24	110	21/8	2097152	254	255.255.255.0	192.0.0.0 223.255.255.255
D	1110			268435454 (1)		224.0.0.0 239.255.255.255
E	1111					240.0.0.0 255.255.255.255

[1] En este caso no se trata de número de nodos, sino de número de grupos multicast.

Tabla 9-4: Rangos Decimales de cada Clase de Direcciones

9.5.3.- Asignación de Direcciones

La asignación de direcciones válidas de Internet la realizan los NICs (Network Information Center). Al principio había un NIC para toda Internet, más tarde se crearon NICs regionales por continentes; actualmente muchos países tienen un NIC propio; en España el NIC es administrado por RedIRIS.

Existen unas reglas y convenios en cuanto a determinadas direcciones IP que es importante conocer:

1. La dirección **255.255.255.255** indica broadcast en la propia red, cualquiera que sea ésta.

2. La dirección **0.0.0.0** identifica al nodo actual.

3. La dirección con el **identificador de nodo todo a ceros** se utiliza para indicar la red misma, y por tanto no se utiliza para ningún nodo.

4. La dirección con el **identificador de nodo todo a unos** se utiliza como la dirección broadcast de la red indicada, y por tanto no se utiliza para ningún nodo.

5. La dirección con el **prefijo de red todo a ceros** identifica a un nodo en la propia red, cualquiera que esta sea; por ejemplo si queremos enviar un datagrama al primer nodo (0.1) de una red clase B podemos utilizar la dirección 0.0.0.1. Esto permite enviar datagramas sin saber en qué red nos encontramos, aunque es preciso conocer si es de clase A, B o C para saber que parte de la dirección es red y que parte es nodo.

6. Las redes **127.0.0.0, 128.0.0.0, 191.255.0.0, 192.0.0.0** y el rango de **240.0.0.0** en adelante están reservados y no deben utilizarse.

7. La dirección **127.0.0.1** se utiliza para pruebas loopback; todas las implementaciones de IP devuelven a la dirección de origen los datagramas enviados a esta dirección sin intentar enviarlos a ninguna parte.

8. Las redes **10.0.0.0, 172.16.0.0 a 172.31.0.0, y 192.168.0.0 a 192.168.255.0** están reservadas para redes privadas ('intranets') por el RFC 1918; estos números no se asignan a ninguna dirección válida en Internet y por tanto pueden utilizarse para construir redes, por ejemplo detrás de un cortafuego, sin riesgo de entrar en conflicto de acceso a redes válidas de la Internet.

Como consecuencia de las reglas 3 y 4 siempre hay dos direcciones inútiles en una red. Por ejemplo, si tenemos la red 200.200.200.0 (clase C) tendremos que reservar la dirección 200.200.200.0 para denotar la red misma, y la dirección 200.200.200.255 para envíos broadcast a toda la red; dispondremos pues de 254 direcciones para nodos, no de 256.

9.5.3.1.- Limitaciones imprevistas en el direccionamiento IP

Los diseñadores originales nunca supusieron que Internet podría crecer hasta alcanzar su dimensión actual. Muchos de los problemas que sufre Internet hoy en día tienen su origen en las decisiones que se tomaron inicialmente durante los años de formación de Internet.

- Durante los primeros años de Internet, el espacio de direcciones, aparentemente ilimitado, fue asignado a las organizaciones basándose en sus peticiones en lugar de hacerlo basándose en sus necesidades reales. Como resultado, las direcciones se asignaron libremente a aquellos que las pedían sin preocuparse por el agotamiento eventual de las direcciones IP.

- La decisión de estandarizar las direcciones a una longitud de 32 bits supuso que el total de direcciones disponibles era de sólo 2^{32} (4,294,967,296). Si se hubieran contemplado direcciones ligeramente más largas se habría incrementado exponencialmente el número de direcciones eliminando el actual problema de escasez de las mismas.

- Los límites entre los prefijos de red y el identificador de nodo de las clases A, B y C eran sencillas de entender e implementar, pero no permiten una asignación eficiente de un espacio de direcciones limitado. Los problemas surgen, principalmente, como consecuencia de la falta de una clase de red pensada para organizaciones de tamaño medio. Una dirección de clase C (/24), que soporta 254 nodos es demasiado pequeña, mientras que una de clase B (/16) que soporta 65,534 nodos es demasiado grande. En el pasado, se han asignado direcciones de clase B a organizaciones con varios cientos de nodos, desaprovechando la mayor parte de las direcciones asignadas. Desafortunadamente, esto ha conducido a un agotamiento prematuro del espacio de direcciones de clase B. El único espacio de direcciones disponible para organizaciones medianas en la actualidad es el correspondiente a direcciones de clase C lo que tiene un impacto muy negativo en el aumento de tamaño de las tablas de encaminamiento globales de Internet.

La historia de Internet en los últimos años se ha centrado en una serie de pasos encaminados a solucionar estos aspectos problemáticos del direccionamiento y han permitido el continuado crecimiento global de Internet.

La solución a largo plazo a estos problemas será la utilización de la próxima versión de IP (IPv6). Sin embargo, mientras la comunidad Internet espera a esta nueva versión, IPv4 ha sufrido algunas modificaciones de modo que Internet puede continuar proporcionando la conectividad universal que todos esperamos.

9.5.4.- Subredes

En 1985, RFC 950 definió un procedimiento estándar para soportar la división en subredes (“Subnetting”) de una dirección de Clase A, B o C. El direccionamiento de subred se introdujo para solucionar algunos de los problemas que estaban empezando a surgir en Internet como consecuencia de la jerarquía de direccionamiento en dos niveles de cada una de las clases:

- Las tablas de encaminamiento de Internet estaban creciendo sustancialmente.
- Los administradores locales debían pedir un nuevo identificador de red cada vez que querían instalar una nueva red en su organización, aún cuando no hubieran agotado el espacio de direcciones del que disponían.

Ambos problemas fueron atacados añadiendo otro nivel de jerarquía en la estructura de direcciones IP. En lugar de la jerarquía de dos niveles, el direccionamiento de subred soporta una jerarquía de tres niveles. La Figura 9-10 ilustra la idea básica del direccionamiento de subred, que es dividir el identificador de nodo estándar de las clases básicas en dos partes, un identificador o número de subred y el identificador de nodo en dicha subred.

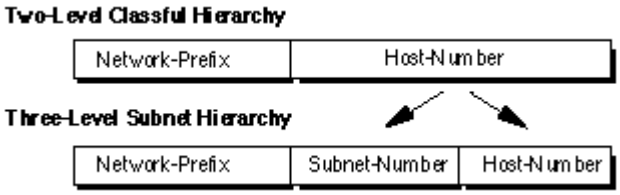

Figura 9-10: Jerarquía de Direccionamiento de Subred

El direccionamiento de subred atajó el problema del crecimiento de las tablas de encaminamiento asegurando que la estructura de subred de una red no es visible fuera de la red privada de una organización. La ruta desde Internet hasta cualquier subred de una dirección IP dada es siempre la misma, independiente de en qué subred se encuentre el nodo de destino. Esto es así porque todas las subredes de una red dada utilizan el mismo prefijo de red, aunque diferentes números de subred. Los routers dentro de una organización privada necesitan diferenciar entre subredes individuales, pero en lo que respecta a los routers de Internet, todas las subredes de una organización están recogidas en una sola fila de la tabla de encaminamiento .Esto permite al administrador local introducir una complejidad arbitraria en su red privada sin afectar al tamaño de las tablas de encaminamiento de Internet.

El direccionamiento de subred también soluciona el problema de la necesidad de un prefijo de red por cada red que una organización desee instalar. La organización que dispone de un espacio de direcciones IPv4 es libre de asignar distintos números de subred a cada una de sus redes internas. Esto permite a las organizaciones instalar subredes adicionales sin necesidad de obtener nuevos prefijos de red.

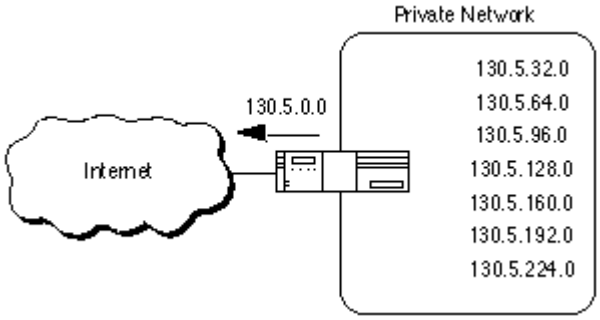

Figura 9-11: El direccionamiento de subred reduce el tamaño de las tablas de encaminamiento

En la Figura 9-11, una organización con varias redes lógicas utiliza direccionamiento de subred con una dirección de Clase B (/16). El router que conecta la organización con Internet acepta todo el tráfico dirigido a la red 130.5.0.0, reenviándolo a las subredes interiores basándose en el tercer octeto de la dirección.

La utilización de subredes dentro de una red privada tiene varias ventajas:

- El tamaño de las tablas de encaminamiento globales de Internet no crece porque el administrador local no necesita obtener espacios de direcciones adicionales y por lo tanto la ruta para todas las subredes están combinadas en una sola entrada de la tabla de encaminamiento.

- El administrador local tiene la flexibilidad de crear subredes adicionales sin tener que solicitar nuevos prefijos de red de Internet.

- El cambio de rutas dentro de una red privada no afecta a las tablas de encaminamiento de Internet puesto que los routers desconocen la accesibilidad de cada subred individual, sólo conocen la accesibilidad de la red global.

9.5.4.1- Prefijo de Red Extendido

Los router de Internet sólo utilizan el prefijo de red de la dirección de destino par encaminar el tráfico a un entorno dividido en subredes. Los router dentro de este entorno subdividido utilizan un prefijo de red extendido para encaminar el tráfico entre las subredes individuales. El prefijo de red extendido está compuesto por el prefijo de red y el identificador de subred.

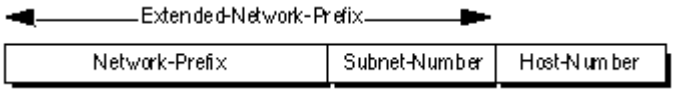

Figura 9-12**:** Prefijo de red extendido

El prefijo de red extendido ha sido identificado tradicionalmente por la máscara de subred. Los bits de la máscara de subred y los de la dirección IP tienen una correspondencia uno a uno. Los bits que ocupan posiciones que corresponden al prefijo de red extendido están puestos a 1 , mientras que los bits que corresponden al identificador de nodo están puestos a 0. Esto se ilustra en la Figura 9-12, donde para la dirección /16 130.5.0.0 se utiliza la máscara 255.255.255.0 con lo que se incluye el tercer octeto de la dirección en el prefijo de red extendido.

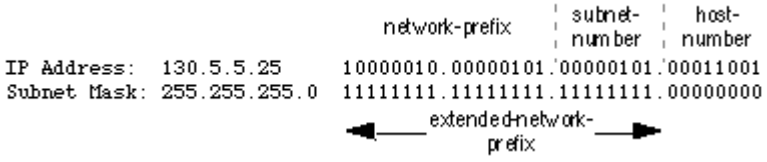

Figura 9-13: Máscara de Subred

Los estándares que describen protocolos de encaminamiento modernos se refieren a menudo a la longitud del prefijo de red extendido en lugar de a la máscara de subred. La longitud del prefijo es igual al número de bits a uno contiguos en la máscara de subred tradicional Esto significa que especificar la dirección de red 130.5.5.25 con una máscara de subred 255.255.255.0 es equivalente a expresarla como 130.5.5.25/24. La notación /<longitud de prefijo> es más compacta y fácil de entender que la escritura tradicional de la máscara de subred en formato decimal.

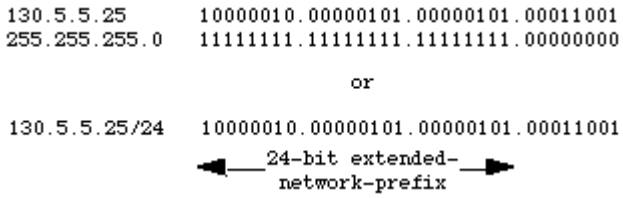

Figura 9-14: Longitud de Prefijo extendido de red

Sin embargo, es importante tener en cuenta que los protocolos de encaminamiento modernos todavía hacen uso de las máscaras de subred. No hay ningún protocolo de encaminamiento estándar en Internet que utilice un campo de un octeto en sus cabeceras para transportar la longitud del prefijo de red extendido, y en su lugar transportan los cuatro octetos de la máscara de subred.

9.5.4.2- Asignación de direcciones de subred

Al crear subredes hay dos direcciones de cada subred que quedan automáticamente reservadas: las que corresponden al identificador de nodo todo a 0s y todo a 1s; éstas se emplean para designar la subred y para la dirección broadcast, respectivamente. Así si la red 156.134.0.0 /24 se subdivide con la máscara 255.255.255.0 se crean 256 subredes tipo *156.134.subred.nodo*, siendo *156.134.subred.0* la dirección que identifica a toda la subred y *156.134.subred.255* la dirección broadcast de la subred. Por tanto, el número de nodos de una subred es siempre dos menos que el número de direcciones que abarca, por ello no tiene sentido crear subredes con la máscara 255.255.255.254 en las que el identificador de nodo tendría un bit, pues no quedarían direcciones útiles.

Ejemplo 1

Consideremos una organización a la que se le ha asignado el identificador de red 193.1.1.0/24 y que necesita definir seis subredes. La subred mayor debe soportar 25 nodos.

Definición de la Máscara de Subred/Prefijo de Longitud extendido.

El primer paso es determinar el número de bits necesario para definir las 6 subredes. Dado que un identificador de red sólo puede ser dividido en un número de subredes que sea potencia de dos resulta imposible dividir la red exactamente en 6 subredes, por lo que elegiremos 8 (2^3) subredes, disponiendo de 2 subredes sin utilizar reservadas para el crecimiento futuro.

Serán necesarios 3 bits para diferenciar las ocho subredes entre sí ($8 = 2^3$). Así, dividimos una dirección de Clase C (/24), de modo que el prefijo de red extendido tendrá una longitud 27 bits (/27). La máscara de Subred será por tanto 255.255.255.224 en notación decimal, como se aprecia en la Figura 9-15.

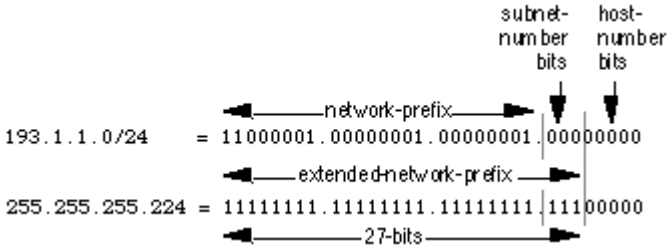

Figura 9-15: Ejemplo 1 – Definición de la Máscara de Subred / Longitud de Prefijo extendido

Un prefijo de red extendido de 27 bits deja 5 bits para definir el identificador de nodo en cada subred. Esto supone que cada subred /27 representa un bloque contiguo de 2^5 (32) direcciones individuales IP. Sin embargo, dado que las direcciones todo 0s y todo 1s no pueden asignarse hay un total de 30 (2^5 -2) identificadores de nodo asignables en cada subred.

Definición de cada identificador de Subred

Las ocho subredes en este ejemplo estarán numeradas de la 0 a la 7 y sus direcciones son los siguientes:

Subred 0	*11000001.00000001.00000001.**000*** 00000	193.1.1.0/27
Subred 1	*11000001.00000001.00000001.**001*** 00000	193.1.1.32/27
Subred 2	*11000001.00000001.00000001.**010*** 00000	193.1.1.64/27
Subred 3	*11000001.00000001.00000001.**011*** 00000	193.1.1.96/27
Subred 4	*11000001.00000001.00000001.**100*** 00000	193.1.1.128/27
Subred 5	*11000001.00000001.00000001.**101*** 00000	193.1.1.160/27
Subred 6	*11000001.00000001.00000001.**110*** 00000	193.1.1.192/27
Subred 7	*11000001.00000001.00000001.**111*** 00000	193.1.1.224/27

Tabla 9-5: Ejemplo 1 – Definición de cada Identificador de Subred

donde los dígitos en negrita y cursiva identifican los 3 bits que representan el identificador de subred. Una forma sencilla de chequear si las subredes son correctas es asegurarse de que todas ellas sean múltiplos de la dirección de la subred 1. En este caso, todas las subredes son múltiplos de 32: 0, 32, 64, 96, ...

Definición del identificador de nodo para cada Subred

En este caso disponemos de 5 bits en el campo identificador de nodo de cada dirección de subred. Esto significa que cada subred representa un bloque de 30 direcciones de nodo (2^5 -2 = 30). Los nodos de cada subred están numerados desde el 1 al 30.

Subred 2	*11000001.00000001.00000001.010*000000	193.1.1.64/27
Nodo 1	*11000001.00000001.00000001.010***00001**	193.1.1.65/27
Nodo 2	*11000001.00000001.00000001.010***00010**	193.1.1.66/27
Nodo 3	*11000001.00000001.00000001.010***00011**	193.1.1.67/27
Nodo 4	*11000001.00000001.00000001.010***00100**	193.1.1.68/27
Nodo 5	*11000001.00000001.00000001.010***00101**	193.1.1.69/27
	. . . . .	
Nodo 15	*11000001.00000001.00000001.010***01111**	193.1.1.79/27
Nodo 16	*11000001.00000001.00000001.010***10000**	193.1.1.80/27
	. . . . .	
Nodo 27	*11000001.00000001.00000001.010***11011**	193.1.1.91/27
Nodo 28	*11000001.00000001.00000001.010***11100**	193.1.1.92/27
Nodo 29	*11000001.00000001.00000001.010***11101**	193.1.1.93/27
Nodo 30	*11000001.00000001.00000001.010***11110**	193.1.1.94/27

Tabla 9-6: Direcciones de nodo de la Subred 2

La Tabla 9-6 muestra las direcciones de nodo válidas para la Subred 2 de este ejemplo. La parte en itálica de cada dirección identifica el prefijo de red extendido, mientras que la parte negrita identifica el campo identificador de nodo.

La Tabla 9-7 muestra las direcciones de nodo válidas para la Subred 6 de este ejemplo. La parte en itálica de cada dirección identifica el prefijo de red extendido, mientras que la parte negrita identifica el campo identificador de nodo.

Subred 6	*11000001.00000001.00000001.110*00000	193.1.1.192/27
Nodo 1	*11000001.00000001.00000001.110***00001**	193.1.1.193/27
Nodo 2	*11000001.00000001.00000001.110***00010**	193.1.1. 194/27
Nodo 3	*11000001.00000001.00000001.110***00011**	193.1.1.195/27
Nodo 4	*11000001.00000001.00000001.110***00100**	193.1.1.196/27
Nodo 5	*11000001.00000001.00000001.110***00101**	193.1.1.197/27
	. . . . .	
Nodo 15	*11000001.00000001.00000001.110***01111**	193.1.1.207/27
Nodo 16	*11000001.00000001.00000001.110***10000**	193.1.1.208/27
	. . . . .	
Nodo 27	*11000001.00000001.00000001.110***11011**	193.1.1.219/27
Nodo 28	*11000001.00000001.00000001.110***11100**	193.1.1.220/27
Nodo 29	*11000001.00000001.00000001.110***11101**	193.1.1.221/27
Nodo 30	*11000001.00000001.00000001.110***11110**	193.1.1.222/27

Tabla 9-7: Direcciones de nodo de la Subred 6

Definición de la dirección de broadcast de cada subred

La dirección de broadcast de la Subred 2 es la dirección 11000001.00000001.00000001.010 *11111 (193.1.1.95). Apreciar que la dirección de broadcast de la Subred 2 es exactamente una unidad menos que la dirección base de la Subred 3 (193.1.1.96). Esto es siempre el caso, la dirección de la Subred n es uno menos que la dirección base de la subred (n+1).*

La dirección de broadcast de la Subred 6 es simplemente la dirección todo 1s: 11000001.00000001.00000001.110 *11111 (193.1.1.223). De nuevo, la dirección de broadcast de la Subred 6 es exactamente uno menos que la dirección base de la Subred 7 (193.1.1.224).*

Del mismo modo que los valores todo 0s o todo 1s del identificador de nodo están reservados con un significado especial, el valor todo 0s y todo 1s del número de subred también son especiales. Cuando se definieron las subredes en el RFC 950 se prohibió el uso de las subredes todo 0s y todo 1s. Si se desconoce la máscara de subred o la longitud del prefijo extendido (algo habitual en las notificaciones de ruta de algunos protocolos, como veremos más adelante) resultaría ambiguo el identificador 193.1.1.0 que no podría asegurarse si representa a la red entera 193.1.1.0/24 o a la subred 193.1.1.0/27 como puede apreciarse en la Figura 12.

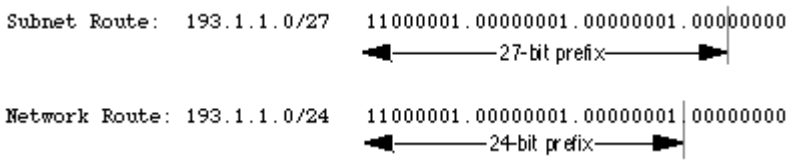

Figura 9-16: Diferenciación entre una Subred todo ceros y la Red entera

De igual modo, en una subred todo 1s, si se desconoce la máscara de subred o la longitud del prefijo, no sería posible determinar si un broadcast se refiere a toda la red (todas las subredes) o a la subred todo 1s. Por ejemplo, existiría ambigüedad con la dirección 193.1.1.255 que se utilizaría tanto para el broadcast de la red completa 193.1.1.0/24 como para la subred todo 1s 193.1.1.224/27, tal y como se muestra en la Figura 13.

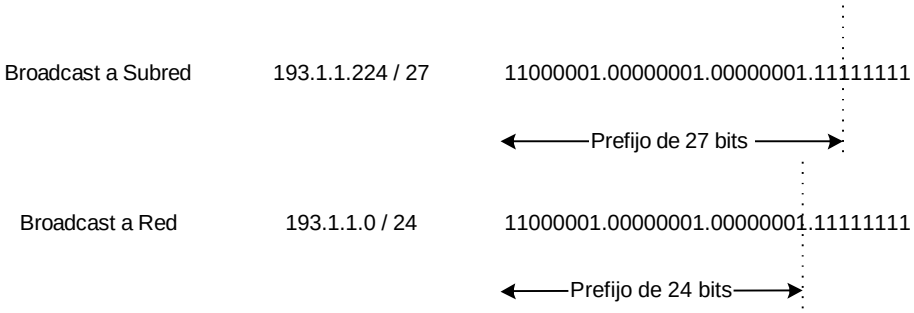

Figura 9-17: identificación de un broadcast a la subred todo 1s y la red entera

Por consiguiente en el número de subred también se pierden siempre dos direcciones, y tampoco tendría sentido crear máscaras con el número de subred de un bit, tales como 255.255.128.0 (longitud de prefijo 17) en el caso de una red clase B.

Con el desarrollo de protocolos de encaminamiento que proporcionan la máscara o la longitud del prefijo con cada ruta, el espacio definido por las subredes todo 0s y todo 1s ha vuelto a ser utilizable pese a las cautelas planteadas en el RFC 950. Como resultado, los fabricantes permiten al usuario configurar los dispositivos especificando “Subred-zero” indicando que se consideren válidas las subredes especificadas por el identificador todo 0s y todo 1s. De este modo es posible aprovechar mejor el espacio de direcciones disponible, dividiendo, por ejemplo, una red de Clase B o Clase C por la mitad en dos subredes mediante las máscaras 255.255.128.0 o 255.255.255.128 respectivamente, cosa que no sería posible ateniéndose a las especificación RFC 950. Hay que tener en cuenta que para utilizar esta opción todos los routers de la organización deben ser capaces de interpretar, aprender y reencaminar el tráfico correctamente a estas subredes.

Otro aspecto a considerar es que en la definición inicial del direccionamiento de subred en el RFC 950 se permitía crear subredes con máscara no contigua, por ejemplo emplear en una dirección de clase B la máscara 255.255.0.255. En este caso, el nodo vendría especificado por el tercer byte y la subred por el cuarto. Dado que esta práctica sólo complicaba la comprensión de las subredes y hacía menos eficiente el proceso en los routers, actualmente la norma exige que la máscara sea contigua.

Bits de subred	Número de subredes	Nº subredes (subred cero)	Bits de nodo	Long prefijo de subred	Número de nodos	Máscara
0	0	0	16	/16	65534	255.255.0.0
1	0	2	15	/17	32766	255.255.128.0
2	2	4	14	/18	16382	255.255.192.0
3	6	8	13	/19	8190	255.255.224.0
4	14	16	12	/20	4094	255.255.240.0
5	30	32	11	/21	2046	255.255.248.0
6	62	64	10	/22	1022	255.255.252.0
7	126	128	9	/23	510	255.255.254.0
8	254	256	8	/24	254	255.255.255.0
9	510	512	7	/25	126	255.255.255.128
10	1022	1024	6	/26	62	255.255.255.192
11	2046	2048	5	/27	30	255.255.255.224
12	4094	4096	4	/28	14	255.255.255.240
13	8190	8192	3	/29	6	255.255.255.248
14	16382	16384	2	/30	2	255.255.255.252
15	32766	32768	1	/31	0	255.255.255.254
16	65532	65536	0	/32	0	255.255.255.255

Tabla 9-8: Direccionamiento de Subred de una dirección de Clase B

Para terminar de clarificar el uso de máscaras para crear subredes resumimos en la Tabla 9-8 las posibles subredes y máscaras que se pueden utilizar con una red clase B:

En el caso de una clase C sería según se recoge en la Tabla 9-9:

Bits de subred	Número de subredes	Nº subredes (subred cero)	Bits de host	Long prefijo de subred	Número de nodos	Máscara
0	0	0	8	65534	254	255.255.255.0
1	0	2	7	32766	126	255.255.255.128
2	2	4	6	16382	62	255.255.255.192
3	6	8	5	8190	30	255.255.255.224
4	14	16	4	4094	14	255.255.255.240
5	30	32	3	2046	6	255.255.255.248
6	62	64	2	1022	2	255.255.255.252
7	126	128	1	510	0	255.255.255.254
8	254	256	0	254	0	255.255.255.255

Tabla 9-9: Direccionamiento de Subred de una dirección de Clase B

Ejemplo 2

Supongamos una organización a la que se le ha asignado la dirección de red 140.25.0.0/16 y necesita crear un conjunto de subredes que alberguen 60 nodos cada una.

Definición de la máscara de subred / Longitud del prefijo de red extendido

El primer paso es determinar el número de bits necesario para definir 60 nodos en cada subred. Será necesario que el administrador defina bloques de 62 (2^6 -2) direcciones de nodo. Sin embargo esto sólo proporcionaría dos direcciones libres para futuros crecimientos; si el administrador considera que esto es insuficiente debería elegir bloques de 126 (2^7 -2) direcciones de nodo, resultando 66 direcciones no asignadas en cada subred. Para esta última definición se necesitan 7 bits para el identificador de nodo.

El siguiente paso es determinar la longitud del prefijo de subred extendido / máscara. Dado que se necesitan 7 bits de los 32 para el identificador de nodo, el prefijo extendido debe ser /25 (25 = 32-7), que puede expresarse mediante la 255.255.255.128, como se ilustra en la Figura 9-18.

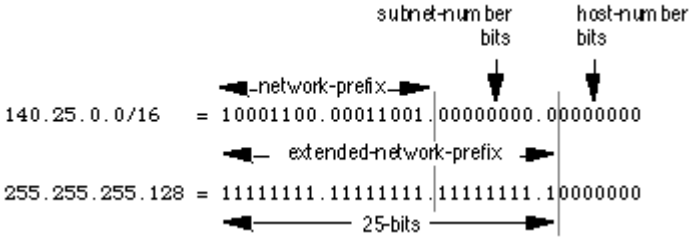

Figura 9-18: Ejemplo 2 – Definición de la Máscara de Subred / Longitud de Prefijo Extendido

Dado que 2^9 = 512, 9 bits permiten la definición de 512 subredes. Dependiendo de las necesidades de la organización, el administrador podría haber elegido asignar bits adicionales al identificador de nodo (permitiendo más nodos en cada subred) y reducir el número de bits de identificador de subred (disminuyendo el número total de subredes que pueden definirse).

Aunque este ejemplo crea un número mayor de subredes, resulta interesante porque ilustra lo que sucede con la representación decimal de una dirección de subred cuando los bits del identificador de subred se extienden más allá de los límites de un octeto. Debería mencionarse que el mismo tipo de confusión puede producirse cuando el identificador de nodo se extiende más allá de los límites de un octeto.

Definición de cada identificador de subred

Las 512 subredes pueden numerarse de la 0 a la 511, tal y como se recoge en l a Tabla 9-10.

Red	10001100.00011001.00000000.00000000	140.25.0.0 /16
Subred 0	*10001100.00011001.**00000000.0**0000000*	140.25.0.0 /25
Subred 1	*10001100.00011001.**00000000.1**0000000*	140.25.0.128 /25
Subred 2	*10001100.00011001.**00000001.0**0000000*	140.25.1.0 /25
Subred 3	*10001100.00011001.**00000001.1**0000000*	140.25.1.128 /25
Subred 4	*10001100.00011001.**00000010. 0**0000000*	140.25.2.0 /25
Subred 5	*10001100.00011001.**00000010.1**0000000*	140.25.2.128 /25
Subred 6	*10001100.00011001.**00000011.0**0000000*	140.25.3.0 /25
Subred 7	*10001100.00011001.**00000100.0**0000000*	140.25.3.128 /25
Subred 8	*10001100.00011001.**00000100.1**0000000*	140.25.4.0 /25
Subred 9	*10001100.00011001.**00000101.0**0000000*	140.25.4.128 /25
	. . . . .	
Subred 510	*10001100.00011001.**11111111.0**0000000*	140.25.255.0 /25
Subred 511	*10001100.00011001.**11111111.1**0000000*	140.25.255.128 /25

Tabla 9-10: Ejemplo 1 – Definición de cada Identificador de Subred

Debe apreciarse el modo en que los números secuenciales de subred no parecen serlo cuando se expresan en notación decimal, lo cual puede originar confusión, ya que en estos casos, la notación decimal obscurece en lugar de esclarecer el esquema de numeración de subred.

Definición de identificadores de nodo de cada Subred

En este ejemplo hay 7 bits en cada identificador de nodo de cada subred, por lo que cada una de ellas puede contener 126 direcciones de nodo. La Tabla 9-11 contiene las direcciones válidas para la Subred 3.

Subred 3	*10001100.00011001.**00000001.1**0000000*	140.25.1.128 /25
Nodo 1	*10001100.00011001.00000001.1***0000001**	140.25.1.129 /25
Nodo 2	*10001100.00011001.00000001.1***0000010**	140.25.1.130 /25
Nodo 3	*10001100.00011001.00000001.1***0000011**	140.25.1.131 /25
Nodo 4	*10001100.00011001.00000001.1***0000100**	140.25.1.132 /25
	. . . . .	
Nodo 62	*10001100.00011001.00000001.1***0111110**	140.25.1.190 /25
Nodo 63	*10001100.00011001.00000001.1***0111111**	140.25.1.191 /25
Nodo 64	*10001100.00011001.00000001.1***1000000**	140.25.1.192 /25
	. . . . .	
Nodo 123	*10001100.00011001.00000001.1***1111011**	140.25.1.251 /25
Nodo 124	*10001100.00011001.00000001.1***1111100**	140.25.1.252 /25
Nodo 125	*10001100.00011001.00000001.1***1111100**	140.25.1.253 /25
Nodo 126	*10001100.00011001.00000001.1***1111110**	140.25.1.254 /25

Tabla 9-11: Direcciones de nodo de la Subred 3

Definición de la dirección de broadcast para cada subred

La dirección de broadcast para la subred 3 es 10001100.00011001.00000001.1 ***1111111*** *(140.25.1.255), de acuerdo con la regla anteriormente citada, una unidad inferior a la dirección base de la subred 4 (140.25.2.0).*

9.5.4.3.- Consideraciones de diseño de Subredes

La creación de un plan de direccionamiento requiere una planificación cuidadosa por parte del administrador. Hay cuatro cuestiones clave que deben contestarse antes de adoptar un diseño:

1) ¿Cuántas subredes precisa la organización en la actualidad?
2) ¿ Cuántas subredes precisará la organización en el futuro?
3) ¿Cuántos nodos hay en la subred más grande de la organización en la actualidad?
4) ¿Cuántos nodos habrá en la subred más grande de la organización en el futuro?

El primer paso en el proceso de planificación es tomar el número máximo de subredes necesarias y redondearlo a la potencia de dos más próxima. Por ejemplo, si una organización precisa 9 subredes, 2^3 (8) no proporciona suficiente espacio de direccionamiento de subred, de modo que el administrador deberá redondear a 2^4 (16). Cuando se realice este cálculo resulta crítico que el administrador reserve espacio para el crecimiento futuro. Por ejemplo, si se necesitan 14 subredes hoy, la elección de 16 podría no ser suficiente cuando se necesite en un futuro instalar 17 subredes. En este caso, hubiera sido necesario prever un crecimiento mayor y seleccionar 2^5 (32) como número máximo de subredes.

El segundo paso es asegurarse de que hay suficientes direcciones de nodo para la subred de la organización más grande. Si la subred más grande necesita 50 nodos sería necesario redondear como mínimo a 2^6 (64) el espacio de identificadores de nodo.

El último paso es asegurarse de que la asignación de direcciones de la organización proporciona suficientes bits para albergar el plan de direcciones de subred elaborado. Por ejemplo, si la organización dispone de una sola dirección de Clase B (/16), pueden destinarse 4 o 5 bits para el número de subred y 12 u 11 más para el identificador de nodo. Sin embargo, si la organización dispone de varias direcciones de Clase C (/24), sería necesario dividir cada una de las direcciones en cuatro subredes (utilizando 2 bits) y construir la red combinando las subredes de 3 prefijos de red diferentes de Clase C (/24). Una solución alternativa podría ser utilizar prefijos de red del espacio de direcciones privado (RFC 1918) para conectividad interna y utilizar un Traductor de Direcciones de Red (NAT, Network Address Translator) para proporcionar acceso externo a Internet.

9.5.5.- Máscaras de Subred de Longitud Variable (VLSM)

En 1987 el RFC 1009 especificó el modo de subdividir una red utilizando más de una máscara de subred. Cuando a una red IP se le asigna más de una máscara de subred se denomina a ésta "máscara de subred de longitud variable" dado que los prefijos extendidos de red tienen longitudes diferentes. Esta técnica, sin embargo sólo puede utilizarse, cuando los protocolos de encaminamiento distribuyen la máscara de subred con cada ruta, de modo que en protocolos como RIP-1 es una técnica no válida.

9.5.5.1.- Uso eficiente del espacio de direcciones IP asignado

VLSM permite un uso más eficiente del espacio de direcciones asignado a una organización. Uno de los mayores problemas de la limitación de soportar sólo una máscara de subred es que una vez que se elige la máscara, esto fija el número de subredes. Por ejemplo, suponiendo que el administrador decide configurar la red 130.5.0.0/16 con un prefijo de red extendido /22.

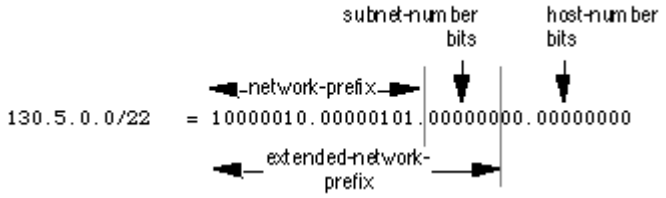

Figura 9-19: 130.5.0.0/16 con un prefijo de red extendido /22

Una red /16 con un prefijo de red extendido /22 permite 64 subredes (2^6), cada una de las cuales soporta un máximo de 1,022 nodos (2^{10} -2). Esto es correcto si la organización desea instalar un número mayor de subredes, pero ¿qué sucede con las redes pequeñas que sólo tienen 20 o 30 nodos? Dado que una red subdividida solo podría tener una máscara, el administrador tenía que asignar los 20 o 30 nodos a una subred con un prefijo de 22. Esta asignación desaprovecharía aproximadamente 1,000 direcciones en cada una de las pequeñas redes instaladas. Limitar la asociación de una red con una sola máscara no permite un uso eficiente y flexible del espacio de direcciones de una organización.

Una solución a este problema era permitir asignar a una red subdividida varias máscaras de subred. Supongamos que en el ejemplo anterior, el administrador puede también configurar la red 130.5.0.0/16 con un prefijo de red extendido /26, como en la Figura 9-20. Una dirección de red /16 con un prefijo de red extendido /26 permite 1024 subredes (2^{10}), cada una de ellas con un máximo de 62 nodos (2^6 -2). El prefijo /26 sería ideal para redes pequeñas con menos de 60 nodos, mientras que el prefijo /22 sería adecuado para redes grandes con más de 1000 nodos.

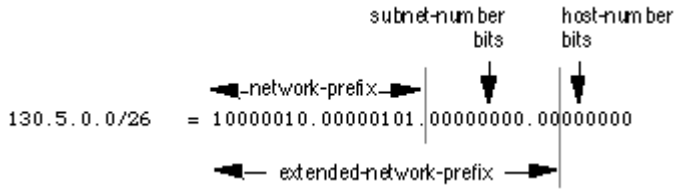

Figura 9-20: 130.5.0.0/16 con un Prefijo de red extendido /26

9.5.5.2.- Agregación de Rutas

VLSM permite también la división recursiva del espacio de direcciones de una organización de modo que pueda ser reensamblado y agregado para reducir la cantidad de información de encaminamiento al máximo nivel. Conceptualmente, una red es primeramente dividida en subredes, algunas de las subredes son posteriormente divididas en subredes y algunas de estas sub-subredes se dividen en sub^2-subredes. Esto permite que la estructura detallada de la información de encaminamiento de una subred esté oculta a los routers de otros subredes.

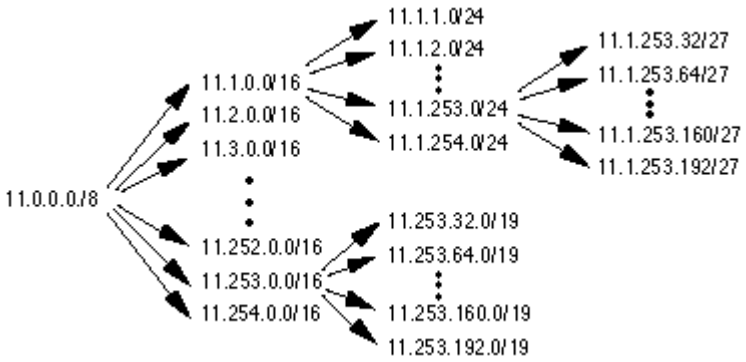

Figura 9-21: VLSM permite la división recursiva de un Prefijo de Red

En la Figura 9-21, puede verse como la red 11.0.0.0/8 está configurada inicialmente con un prefijo de red extendido /16. La subred 11.1.0.0/16 está a continuación configurada con un prefijo de red extendido /24 y la subred 11.253.0.0/16 con un prefijo /110. Como se ve el proceso recursivo no requiere que se asigne el mismo prefijo de red extendido a cada nivel de recursión. Además, la subdivisión recursiva del espacio de direcciones puede extenderse tan lejos como sea necesario.

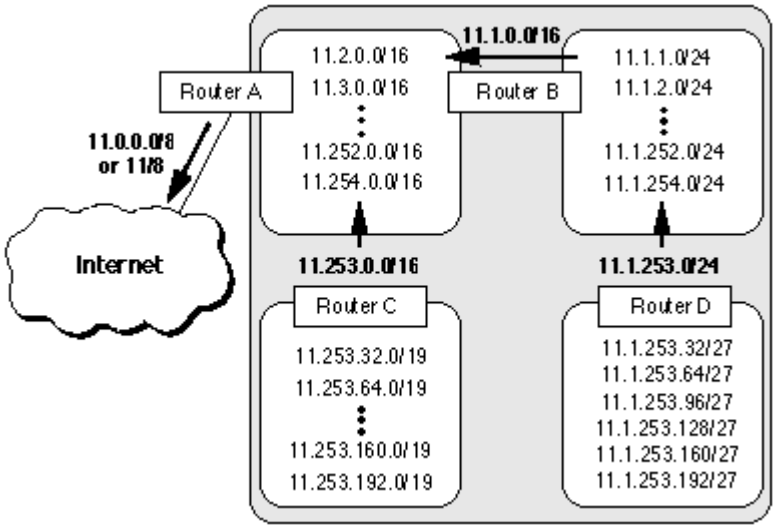

Figura 9-22: VLSM permite agregar rutas y reducir el tamaño de las tablas de encaminamiento

La Figura 9-22 ilustra como una asignación bien de VLSM puede reducir el tamaño de las tablas de encaminamiento de una organización. Como puede verse el Router D puede resumir las 6 subredes a las que proporciona acceso en una sola entrada (11.1.253.0/24) y como los Router B Y C pueden hacer lo mismo (11.1.1.0/16, 11.253.0.0/16). Finalmente, dado que la estructura de subred no es visible desde el exterior, el Router A notifica una sola ruta en las tablas de encaminamiento de Internet 11.0.0.0/ 8 (o 11/8).

9.5.5.3.- Consideraciones en el Diseño VLSM

Cuando se procede a un diseño VLSM, el administrador debe hacerse las mismas preguntas que en el diseño tradicional de subredes, con idénticas decisiones de diseño en cada nivel de la jerarquía:

1) ¿Cuántas subredes precisa la organización en la actualidad?
2) ¿ Cuántas subredes de este nivel precisará la organización en el futuro?
3) ¿Cuántos nodos hay en la subred más grande de este nivel en la actualidad?
4) ¿Cuántos nodos habrá en la subred más grande de este nivel en el futuro?

En cada nivel, el diseñador debe asegurarse de dejar suficientes bits extra para soportar el número de subentidades en los siguientes niveles de recursión.

Supongamos que una red está distribuida en tres localizaciones; esto supone que serán necesarios 3 bits para la división en subredes ($2^3 = 8$) para permitir un cierto crecimiento en el futuro. En cada una de estas localizaciones, habrá probablemente un segundo nivel de subdivisión para identificar cada edificio y finalmente, dentro de cada edificio, puede haber un tercer nivel de subdivisión para identificar cada uno de los grupos individuales de trabajo. Siguiendo este modelo jerárquico, el nivel superior, y el nivel más bajo están determinados por el "número máximo de subredes/número máximo de usuarios por subred" en cada edificio.

El establecimiento de esquema jerárquico precisa un estudio detallado. Es esencial que el administrador trabaje recursivamente hacia abajo en la jerarquía hasta alcanzar el nivel inferior. En este nivel, debe asegurarse de que las subredes son capaces de albergar el número de nodos necesarios. Cuando se finaliza el plan de direcciones, las direcciones de cada localización deben ser agregadas un único bloque de direcciones que prevenga el crecimiento de las tablas de encaminamiento.

9.5.5.4.- Requerimientos para la utilización de VLSM

La utilización correcta de VLSM tiene tres prerrequisitos :

- Los protocolos de encaminamiento deben incluir información sobre el prefijo de red extendido en sus notificaciones.
- Todos los routers deben utilizar un algoritmo de envío basado en la "coincidencia más larga ".
- Para conseguir agregación de rutas, las direcciones deben ser asignadas de modo que sean topológicamente significativas.

Los protocolos de encaminamiento como OSPF, RIP-2, I-IS-IS, etc. permiten el uso de VLSM al incluir notificaciones del prefijo extendido de red o el valor de la máscara con cada ruta notificada.

Por otro lado, hoy en día, todos los router implementan algoritmos de encaminamiento consistentes basados en la "coincidencia más larga". El uso de VLSM supone que el conjunto de redes asociadas con el prefijo de red extendido pueden tener relaciones de subconjunto. Una ruta con un prefijo de red extendido describe un conjunto más pequeño de destinos que la misma ruta con un prefijo de red extendido más corto. Como resultado, un ruta con un prefijo de red extendido más largo se dice "más específica" que otra con un prefijo más corto. Los routers debe utilizar rutas con el prefijo de red extendido más cuando reencaminan tráfico.

Por ejemplo, si un router debe analizar un paquete dirigido a la dirección 11.1.2.5 y existen en su tabla tres prefijos de red (11.1.2.0/24, 11.1.0.0/16, y 11.0.0.0/8), el router debería elegir la ruta 11.1.2.0/24, porque su prefijo tiene la coincidencia de más número de bits con la dirección de destino.

```
     Destination     11.1.2.5     = 00001011.00000001.00000010.00000101

*    Route #1        11.1.2.0/24  = 00001011.00000001.00000010.00000000
     Route #2        11.1.0.0/16  = 00001011.00000001.00000000.00000000
     Route #3        11.0.0.0/8   = 00001011.00000000.00000000.00000000
```

Figura 9-23: La mejor coincidencia es aquella que contiene el prefijo más largo (más específica)

Hay una consideración muy importante a realizar en este punto: aunque que la dirección de destino se corresponde con las tres rutas, debe ser asignada a un nodo conectado a la subred 11.1.2.0/24. Si fuera asignada a un nodo conectado a la subred 11.1.0.0/16 o 11.0.0.0/8, el sistema de encaminamiento nunca enviará tráfico hacia el nodo dado que el algoritmo de "coincidencia más larga" supone que el nodo es parte de la subred 11.1.2.0/24. Esto significa que debe tenerse especial cuidado en la asignación de las direcciones a los nodos para asegurarse de que todos ellos sean alcanzables.

El encaminamiento jerárquico precisa que las direcciones asignadas reflejen la topología real de la red con el fin agrupar grupos de direcciones correspondientes a un zona en una sola ruta. Si las direcciones carecen de sentido topológico no se podrá realizar la agregación y por lo tanto no se conseguirá una reducción de las tablas de encaminamiento.

Ejemplo VLSM

Supongamos una organización a la que se la ha asignado la dirección de red 140.25.0.0/16 y que pretende implementar VLSM de acuerdo al plan indicado en la Figura 9-24.

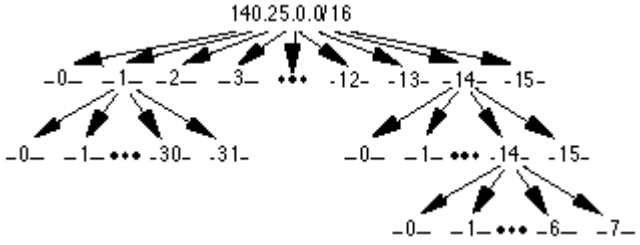

Figura 9-24: Ejemplo de estrategia de direcciones VLSM

El primer paso es dividir la dirección base de red en 16 bloques de direcciones de igual tamaño. A continuación, la Subred 1 se divide en 32 bloques de igual tamaño y la Subred 14 se divide en 16 bloques de direcciones similares. Finalmente, la Subred 14-14 se divide en 8 bloques de tamaño similar.

El primer paso es dividir la dirección base en 16 bloques como se muestra en la Figura 9-25.

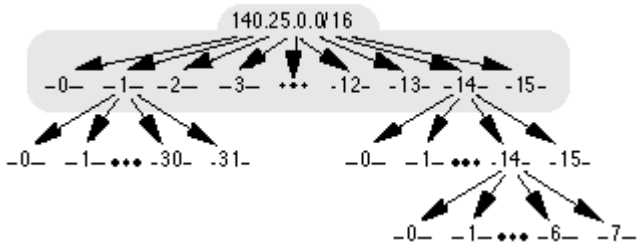

Figura 9-25: Definición de las 16 Subredes de 140.25.0.0/16

Se precisan 4 bits para identificar unívocamente cada una de las subredes. Esto significa que la organización necesitará un prefijo de red extendido /20 para definir las 16 subredes de 140.25.0.0/16. Cada una de éstas representa un bloque contiguo de 2^{12} (o 4,096) direcciones de nodo.

Red	10001100.00011001.00000000.00000000	140.25.0.0 /16
Subred 0	*10001100.00011001.**0000**0000.00000000*	140.25.0.0 /20
Subred 1	*10001100.00011001.**0001**0000.00000000*	140.25.16.0 /20
Subred 2	*10001100.00011001.**0010**0000.00000000*	140.25.32.0 /20
	. . . . .	
Subred 13	*10001100.00011001.**1101**0000.00000000*	140.25.208.0 /20
Subred 14	*10001100.00011001.**1110**0000.00000000*	140.25.224.0 /20
Subred 15	*10001100.00011001.**1111**0000.00000000*	140.25.240.0 /20

Tabla 9-12: Identificadores de las 16 subredes del primer nivel

Examinemos la asignación de identificadores de nodo en la Subred 3, tal y como aparece en la Figura 9-26.

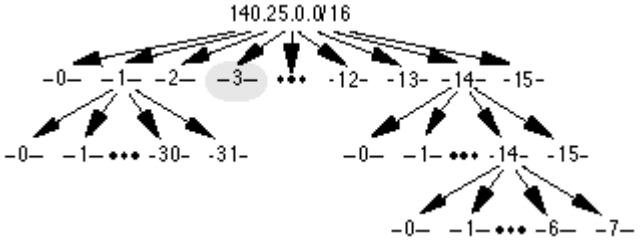

Figura 9-26: Definición de las direcciones de nodo para la subred Subred 3 (140.25.48.0/20)

Dado que el espacio reservado para el identificador de nodo en la Subred 3 es de 12 bits, hay 4,094 direcciones de nodo válidas (2^{12} -2) en el bloque de direcciones. Los nodos están numerados de 1 a 4,094. Las direcciones de nodo válidas para la Subred 3 son las indicadas en la Tabla 9-13.

Subred 3	*10001100.00011001.0011*0000.00000000	140.25.48.0 /20
Nodo 1	*10001100.00011001.0011***0000.00000001**	140.25.48.1 /20
Nodo 2	*10001100.00011001.0011***0000.00000010**	140.25.48.2 /20
	. . . . .	
Nodo 4093	*10001100.00011001.0011***1111.11111101**	140.25.63.253 /20
Nodo 4094	*10001100.00011001.0011***1111.11111110**	140.25.63.254 /20

Tabla 9-13: Identificadores de nodo de la Subred 3

La dirección de broadcast de la Subred 3 es 10001100.00011001.0011 ***1111.11111111*** *= 140.25.63.255, exactamente una unidad inferior a la dirección base de la Subred 4 (140.25.64.0). Una vez dividida la dirección base de red en 16 subredes hay que subdividir la Subred 14 en otros 16 bloques iguales como se indica en la Figura 9-27.*

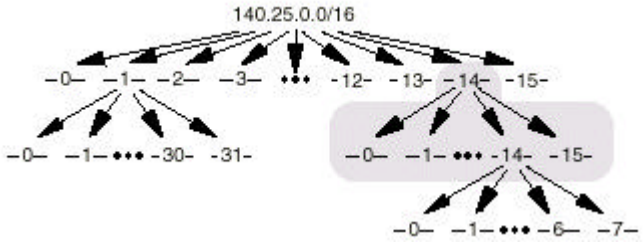

Figura 9-27: Define las Sub-Subredes de la Subred 14 (140.25.224.0/20)

Se necesitan 4 bits para identificar cada una de estas subredes, por lo que se utilizará /24 como longitud del prefijo de red extendido. La Tabla 9-14 muestra las 16 subredes del bloque 140.25.224.0/20, numeradas de 0 a 15.

Subred 14	*10001100.00011001.1110*0000.00000000	140.25.224.0 /20
Subred 14-0	*10001100.00011001.1110***0000**.00000000	140.25.224.0 /24
Subred 14-1	*10001100.00011001.1110***0001**.00000000	140.25.225.0 /24
Subred 14-2	*10001100.00011001.1110***0010**.00000000	140.25.226.0 /24
	. . . . .	
Subred 14-14	*10001100.00011001.1110***1110**.00000000	140.25.238.0 /24
Subred 14-15	*10001100.00011001.1110***1111**.00000000	140.25.2310.0 /24

Tabla 9-14: Identificadores de las 16 subredes del primer nivel

La asignación de identificadores de nodo a la Subred 14-3 (140.25.227.0/24) indicada en la Figura 9-28 contaría con 8 bits disponibles.

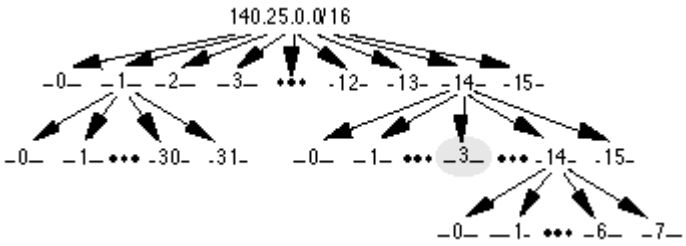

Figura 9-28: Definición de los identificadores de nodo para la Subred #14-3 (140.25.227.0/24)

Cada bloque representará 254 direcciones de nodo válidas (2^8 -2), estando numerados estos desde 1 a 254, tal y como se indica en la Tabla 9-15

Subred 14-3	*10001100.00011001.11100011*.00000000	140.25.227.0 /24
Nodo 1	*10001100.00011001.11100011*.**00000001**	140.25.227.1 /24
Nodo 2	*10001100.00011001.11100011*.**00000010**	140.25.227.2 /24
Nodo 3	*10001100.00011001.11100011*.**00000011**	140.25.227.3 /24
	. . . . .	
Nodo 253	*10001100.00011001.11100011*.**11111101**	140.25.253.0 /24
Nodo 254	*10001100.00011001.11100011*.**11111110**	140.25.254.0 /24

Tabla 9-15: Identificadores de nodo de la Subred 3

La dirección de broadcast de la Subred 14-3 es 10001100.00011001.11100011.***11111111*** *(140.25.227.255). Una vez subdividida la Subred 14 en 16 subredes, como se indica en la Figura 9-29, la Subred 14-14 debe dividirse a su vez en 8 bloques de direcciones de igual tamaño.*

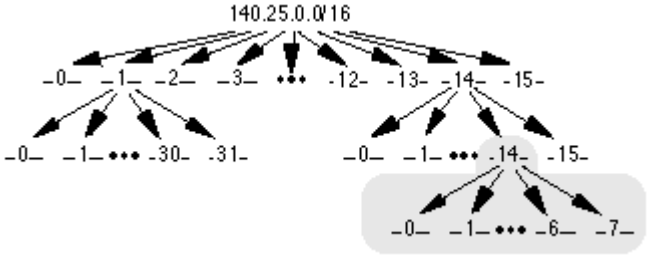

Figura 9-29: Define la Subredes para la Subred 14-14 (140.25.238.0/24)

Serán necesarios 3 bits adicionales para identificar las 8 Subredes, lo que significa que la organización utilizará /27 como longitud del prefijo de red extendido. La Tabla 9-16 contiene los bloques de direcciones de las 8 Subredes de la dirección 140.25.238.0/24:

Subred 14-14	*10001100.00011001.11101110*.00000000	140.25.238.0 /24
Subred 14-14-0	*10001100.00011001.11101110*.**000**00000	140.25.238.0 /27
Subred 14-14-1	*10001100.00011001.11101110*.**001**00000	140.25.238.32 /27
Subred 14-14-2	*10001100.00011001.11101110*.**010**00000	140.25.238.64 /27
Subred 14-14-3	*10001100.00011001.11101110*.**011**00000	140.25.238.96 /27
Subred 14-14-4	*10001100.00011001.11101110*.**100**00000	140.25.238.128 /27
Subred 14-14-5	*10001100.00011001.11101110*.**101**00000	140.25.238.160 /27
Subred 14-14-6	*10001100.00011001.11101110*.**110**00000	140.25.238.192 /27
Subred 14-14-7	*10001100.00011001.11101110*.**111**00000	140.25.238.224 /27

Tabla 9-16: Direcciones de las Subredes comprendidas dentro de Subred 14-14

La asignación de direcciones de nodo a la Subred 14-14-2 (140.25.238.64/27) tal y como se indica en la Figura 9-30 dispondrá de 5 bits para el identificador de nodo.

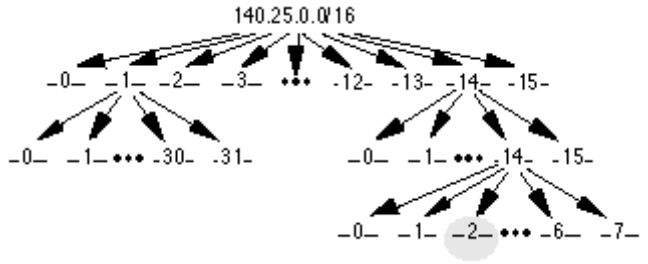

Figura 9-30: Definición de las direcciones de nodo de la Subred 14-14-2 (140.25.238.64/27)

Cada bloque de direcciones contendrá 30 direcciones de nodo válidas tal y como se indica en la Tabla 9-17

Subred 14-14-2	*10001100.00011001.11101110.010*00000	140.25.238.64 /27
Nodo 1	*10001100.00011001.11101110.010***00001**	140.25.238.65 /27
Nodo 2	*10001100.00011001.11101110.010***00010**	140.25.238.66 /27
	. . . . .	
Nodo 29	*10001100.00011001.11101110.010***11101**	140.25.238.93 /27
Nodo 30	*10001100.00011001.11101110.010***11110**	140.25.238.94 /27

Tabla 9-17: Identificadores de nodo de la Subred 3

La dirección de difusión de la Subred 14-14-2 es 10001100.00011001.11011100.010 ***11111*** *(140.25.238.95).*

9.5.6.- CIDR (Classless Inter-Domain Routing)

En 1992 el crecimiento exponencial de Internet estaba haciendo plantearse a los miembros del IETF la posibilidad de que el sistema de encaminamiento de Internet fuera escalable y soportara el crecimiento futuro. Estos problemas estaban relacionados con:

- El práctico agotamiento del espacio de direcciones de Clase B.
- El rápido crecimiento del tamaño de las tablas de encaminamiento globales de Internet.
- El eventual agotamiento del espacio de direcciones de 32 bits de IPv4 .

El crecimiento previsto de Internet hacía ver claro que los dos primeros problemas iban a ser críticos en 1994 o 1995. La respuesta fue el desarrollo del concepto de Encaminamiento Inter-Dominio Sin Clases (CIDR - Classless Inter-Domain Routing). La solución al tercer problema, que se plantea a medio plazo, ha sido el desarrollo de una nueva versión de IP (Ipv6).

CIDR fue documentado oficialmente en Septiembre de 1993 en RFC 1517, 1518, 1519, y 1520. CIDR contempla dos hechos importantes que benefician el sistema de encaminamiento global de Internet:

- Elimina el concepto tradicional de direcciones de Clases A, B y C, permitiendo una asignación más eficiente del espacio de direcciones que permitirá continuar creciendo a Internet hasta que esté disponible IPv6.
- Soporta la agregación de rutas en una sola entrada de las tablas de encaminamiento donde ésta puede representar el espacio de encaminamiento de quizás miles de rutas tradicionales basadas en clases. Esto permite que una sola entrada especifique el modo de encaminar tráfico a muchas direcciones de red individuales, reduciendo la cantidad de información de encaminamiento en los routers del backbone de Internet y facilita la administración de las mismas.

Sin la rápida implantación de CIDR en 1994 y 1995, las tablas de encaminamiento de Internet habrían excedido de las 70,000 rutas (en lugar de las actuales 30,000+) y probablemente se hubiera colapsado Internet.

CIDR ha eliminado el concepto tradicional de Clases A, B, y C y lo ha reemplazado por un concepto generalizado de “prefijo de red”. Los routers utilizan el prefijo de red en lugar de los tres primeros bits de la dirección para determinar el punto de división entre el identificador de red y el identificador de nodo. Como resultado, CIDR soporta el uso de redes de tamaño arbitrario en lugar de los identificadores estándares de 8, 16 o 24 bit asociados con las direcciones de Clases.

En el modelo CIDR cada elemento de información de encaminamiento se notifica con una máscara o longitud del prefijo. Por ejemplo, una red con 20 bits de identificador de red y 12 bits de identificador de nodo sería notificado con /20 como prefijo de red.

Por ejemplo, todos los prefijos con longitud /20 representan la misma cantidad de espacio de direcciones (2^{12} o 4,096 direcciones de nodo). Además un prefijo /20 puede ser asignado a una dirección de red tradicional de clase A, B, C. La Figura 9-31 muestra cómo cada uno de los bloques /20 10.23.64.0/20, 130.5.0.0/20, y 200.7.128.0/20 representan 4,096 direcciones de nodo.

Traditional A	10.23.64.0/20	00001010.00010111.01000000.00000000
Traditional B	130.5.0.0/20	10000010.00000101.00000000.00000000
Traditional C	200.7.128.0/20	11001000.00000111.10000000.00000000

Figura 9-31: Bloques de Direcciones /20

La Tabla 9-18 proporciona información sobre los bloques de direcciones CIDR más utilizados. Podemos ver que la asignación /15 puede especificarse también utilizando la notación tradicional de máscara 255.254.0.0. La asignación /15 contiene un bloque contiguo de 128K (131,072) direcciones IP que pueden interpretase como 2 direcciones de red de Clase B o 512 redes de Clase C.

CIDR prefix-length	Dotted-Decimal	# Individual Addresses	# of Classful Networks
/13	255.248.0.0	512 K	8 Bs or 2048 Cs
/14	255.252.0.0	256 K	4 Bs or 1024 Cs
/15	255.254.0.0	128 K	2 Bs or 512 Cs
/16	255.255.0.0	64 K	1 B or 256 Cs
/17	255.255.128.0	32 K	128 Cs
/18	255.255.192.0	16 K	64 Cs
/19	255.255.224.0	8 K	32 Cs
/20	255.255.240.0	4 K	16 Cs
/21	255.255.248.0	2 K	8 Cs
/22	255.255.252.0	1 K	4 Cs
/23	255.255.254.0	512	2 Cs
/24	255.255.255.0	256	1 C
/25	255.255.255.128	128	1/2 C
/26	255.255.255.192	64	1/4 C
/27	255.255.255.224	32	1/8 C

Tabla 9-18: Bloques de Direcciones CIDR

9.5.6.1.- Implicaciones en los nodos del uso de CIDR

Es importante tener en cuenta que el uso de CIDR puede tener serias implicaciones en los nodos. Dado que la mayoría de los nodos se basa en direcciones de clases tradicionales, el interfaz de usuario no permite configurarlos con un máscara más corta que la máscara "natural" de las direcciones de clase. Por ejemplo, pueden existir problemas potencias si se desea utilizar la dirección 200.25.16.0 con prefijo /20 para definir una red capaz de soportar 4,094 (2^{12} -2) nodos. El software de las estaciones es posible que no permita configurar una tradicional Clase C (200.25.16.0) con una máscara /20, dado que la máscara natural de la Clase C (200.25.16.0) es /24. Si, por el contrario, el software del nodo soporta CIDR, permitirá configurar máscaras más cortas

Sin embargo, no habrá problemas con los nodos cuando se quiera utilizar una dirección tradicional de Clase C (200.25.16.0/20) como un bloque de 14 redes /24 puesto que los nodos no-CIDR interpretarán su dirección local /24 como una Clase C. De igual modo una dirección de Clase B como 130.14.0.0/16 podría considerarse como un bloque de 255 redes /24 donde cada nodo interpretará cada red /24 como subredes de una red /16. Si el software del nodo soporta la configuración de máscaras más cortas que las esperadas, el administrador tiene una flexibilidad tremenda en el diseño de la red y la asignación de direcciones.

9.5.6.2.- Asignación eficiente de direcciones

¿Cómo conduce todo esto a una asignación más eficiente del espacio de direcciones de Ipv4?

En un entorno de Clases, un Proveedor de servicio Internet (ISP) sólo puede asignar direcciones /8, /16 o /24. En un entorno CIDR, el ISP puede extraer un bloque de su espacio de direcciones registradas que se ajuste específicamente a las necesidades de su cliente, proporcionando espacio adicional para su crecimiento y sin desperdiciar un recurso tan escaso como son las direcciones.

Supongamos que un ISP haya recibido el bloque de direcciones 206.0.64.0/18, que representa 16,384 (2^{14}) direcciones IP que pueden ser interpretadas como 64 /24s. Si un cliente necesita 800 direcciones de nodo, en lugar de asignarle un dirección de Clase B (y desperdiciar unas 64,700 direcciones) o 4 direcciones de Clase C (e introducir 4 rutas en las tablas de encaminamiento globales de Internet), el ISP podría asignar al cliente el bloque 206.0.68.0/22 de 1,024 (2^{10}) direcciones IP (4 redes /24 contiguas). La eficiencia de esta asignación se aprecia en la Figura 9-32.

ISP's Block:	11001110.00000000.01000000.00000000	206.0.64.0/18
Client Block:	11001110.00000000.01000100.00000000	206.0.68.0/22
Class C #0:	11001110.00000000.01000100.00000000	206.0.68.0/24
Class C #1:	11001110.00000000.01000101.00000000	206.0.69.0/24
Class C #2:	11001110.00000000.01000110.00000000	206.0.70.0/24
Class C #3:	11001110.00000000.01000111.00000000	206.0.71.0/24

Figura 9-32: CIDR permite una asignación eficaz del espacio de direcciones

Ejemplo de asignación de direcciones CIDR

En este ejemplo vamos a suponer que un ISP posee el bloque de direcciones 200.25.0.0/16. Esto representa 65, 536 (2^{16}) direcciones IP (o 256 redes /24). De este bloque se desea asignar el bloque de direcciones 200.25.16.0/20, que representa 4,096 (2^{12}) direcciones IP (o 16 redes /24).

Bloque de Direcciones	11001000.00011001.00010000.00000000	200.25.16.0/20

En un entorno de Clases tradicionales, el ISP está obligado a utilizar el bloque /20 como 16 redes individual /24s.

Red 0	11001000.00011001.0001**0000**.00000000	200.25.16.0/24
Red 1	11001000.00011001.0001**0001**.00000000	200.25.17.0/24
Red 2	11001000.00011001.0001**0010**.00000000	200.25.18.0/24
Red 3	11001000.00011001.0001**0011**.00000000	200.25.110.0/24
Red 4	11001000.00011001.0001**0100**.00000000	200.25.20.0/24
	. . . . .	
Red 13	11001000.00011001.0001**1101**.00000000	200.25.210.0/24
Red 14	11001000.00011001.0001**1110**.00000000	200.25.30.0/24
Red 15	11001000.00011001.0001**1111**.00000000	200.25.31.0/24

Tabla 9-19

Si se contempla el bloque de direcciones /20 como un "pastel", en un entorno tradicional sólo puede se dividido en 16 sectores de igual tamaño, tal y como se muestra en la Figura 9-32.

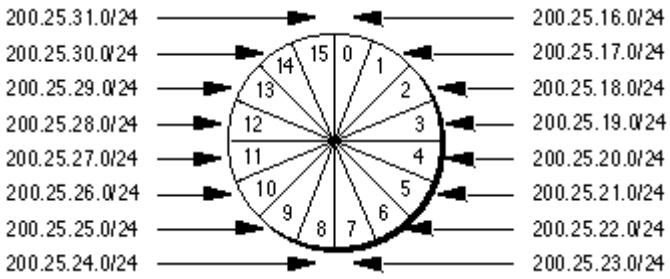

Figura 9-32: Dividiendo el espacio – Entorno de Clases tradicionales

Sin embargo, en un entorno sin clases, el ISP es libre de cortar el pastel del modo en que desee. Podría dividirlo en 2 partes (cada una con la mitad del espacio de direcciones) y asignar una porción a la Organización A, a continuación dividir la otra mitad en partes iguales (cada una con ¼ del espacio total de direcciones) y asignar una porción a la Organización B, y finalmente dividir el espacio sobrante en otras dos partes (cada una de 1/8 del espacio de direcciones) y asignárselas a las Organizaciones C y D, tal y como se muestra en la Figura 9-33. Cada una de las organizaciones es libre de asignar el espacio de direcciones dentro de su "intranet" como le parezca.

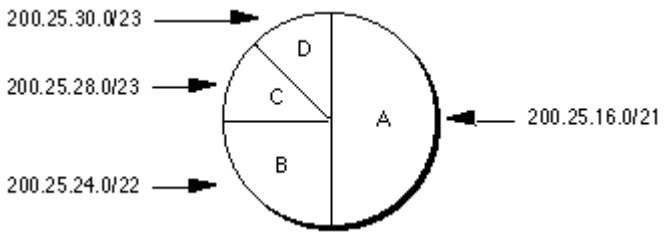

Figura 9-33: Dividiendo el espacio – Entorno CIDR

Paso 1: Dividir el bloque de direcciones 200.25.16.0/20 en dos partes iguales. Cada una representa la mitad del espacio de direcciones, es decir 2,048 (2^{11}) direcciones IP.

Bloque ISP	11001000.00011001.00010000.00000000	200.25.16.0/20
Organización A	11001000.00011001.00010000.00000000	200.25.17.0/21
Reservado	11001000.00011001.00011000.00000000	200.25.24.0/21

Paso 2: Dividir el bloque reservado 200.25.24.0/21 en dos mitades iguales. Cada bloque representará un cuarto del espacio total, es decir 1,024 (2^{10}) direcciones IP.

Reservado	11001000.00011001.00011000.00000000	200.25.24.0/21
Organización B	11001000.00011001.00011000.00000000	200.25.24.0/22
Reservado	11001000.00011001.00011100.00000000	200.25.28.0/22

Paso 3: Dividir el bloque de direcciones reservadas 200.25.28.0/22 en dos bloques de igual tamaño. Cada uno representa 1/8 del espacio de direcciones total, es decir 512 (2^{9}) direcciones IP.

Reservado	11001000.00011001.00011100.00000000	200.25.28.0/22
Organización C	11001000.00011001.00011100.00000000	200.25.28.0/23
Organización D	11001000.00011001.00011110.00000000	200.25.30.0/23

9.5.6.3.- CIDR y VLSM

CIDR y VLSM son básicamente la misma cosa, dado que ambos permiten dividir recursivamente un espacio de direcciones IP en partes más pequeñas. La diferencia es que con VLSM, la recursión se realiza sobre un espacio de direcciones previamente asignado a una organización y resulta invisible al resto de Internet. Sin embargo, CIDR, permite la asignación recursiva de un bloque de direcciones por una entidad de Registro de Internet a un ISP de alto nivel, a un ISP de nivel medio, a un ISP de bajo nivel y finalmente a la red privada de un organización.

Al igual que VLSM, para el uso de CIDR son necesarios tres requisitos:

- Los protocolos de encaminamiento deben incluir información sobre el prefijo de red en cada notificación.
- Todos los routers deben implementar un algoritmo consistente basado en la “coincidencia más larga”.
- Para conseguir una agregación de rutas, las direcciones deben ser asignadas de modo que sean topológicamente significativas.

9.5.6.4.- Control del crecimiento de las tablas de encaminamiento de Internet

Otro beneficio importante de CIDR es el papel que juega en el control del crecimiento de las tablas de encaminamiento de Internet. La reducción de la información de encaminamiento requiere que Internet se divida en dominios de direccionamiento. Dentro de un dominio, se dispone de información detallada acerca de todas las redes que residen en el dominio; fuera de un dominio de direccionamiento, sólo se notifica el prefijo de red común. De este modo una sola entrada de la tabla de encaminamiento puede especificar numerosas direcciones de red individuales.

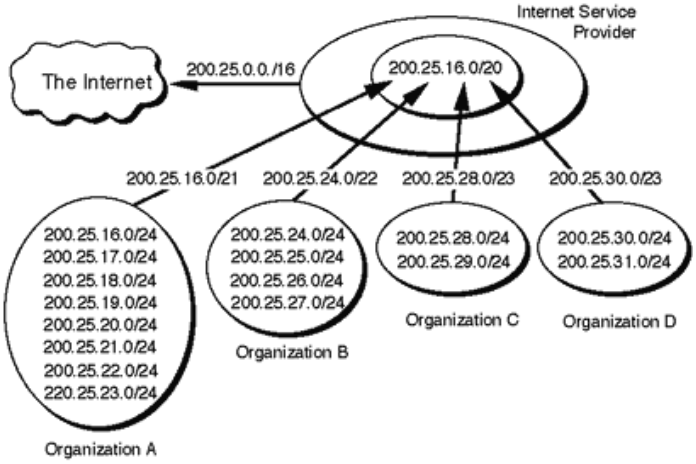

Figura 9-34: CIDR Reduce el Tamaño de las tablas de encaminamiento.

La Figura 9-34 muestra el modo de en que CIDR reduce el tamaño de las tablas de Internet. Supongamos que una porción del bloque de direcciones del ISP (200.25.16.0/20) se haya has asignado como se ha descrito en el ejemplo anterior. La Organización A agrega 8 direcciones /24 en una sola notificación (200.25.16.0/21), la Organización B agrega 4 direcciones /24 (200.25.24.0/22), la Organización C agrega 2 direcciones /24 (200.25.28.0/23), y la Organización D agrega 2 direcciones /24 (200.25.30.0/23). Finalmente, el ISP puede presentar las 256 redes /24s en Internet con una sola notificación (200.25.0.0/16).

La utilización de CIDR permitirá que el número de redes individuales de Internet se expanda minimizando el número de rutas de las tablas de encaminamiento.

9.5.6.5.- Jerarquía geográfica en la asignación de direcciones

Otra medida adoptada junto con CIDR es la introducción de una ordenación geográfica en la asignación de direcciones, asignando rangos iniciales por continentes:

194.0.0.0 a 195.255.0.0 para Europa
198.0.0.0 a 199.255.0.0 para Norteamérica
200.0.0.0 a 201.255.0.0 para Centro y Sudamérica
202.0.0.0 a 203.255.0.0 para Asia y la zona del Pacífico

A su vez dentro de cada uno de estos rangos se ha dado una parte a cada país, y dentro de éste un rango a cada ISP. Con esta distribución regional de los números es posible simplificar las entradas de las tablas de encaminamiento, de modo que, por ejemplo, un router en Japón puede poner una sola entrada en sus tablas indicando que todos los paquetes dirigidos a las redes 194.0.0.0 hasta 195.255.0.0 vayan a la interfaz que da acceso a Europa.

9.5.7.- Nuevas Soluciones para el escalado del espacio de direcciones de Internet

A pesar del éxito de CIDR, el crecimiento en los últimos años ha sido de tal magnitud que los problemas han reaparecido en forma de crecimiento exponencial de rutas de encaminamiento. El IETF ha continuado sus esfuerzos por desarrollar soluciones que permitan solucionar estos problemas, permitiendo un continuado crecimiento y la escalabilidad de Internet.

9.5.7.1.- Devolución de prefijos de red IP no utilizados

El RFC 1917 pide que la comunidad Internet devuelva los bloques de direcciones no utilizados a la IANA (Internet Assigned Numbers Authority) para su redistribución. Esto incluye los identificadores de red no utilizados, las direcciones de red que nunca serán conectadas a Internet por seguridad y las de los sitios que están utilizando sólo un pequeño porcentaje de su espacio de direcciones. El RFC 1917 también pide a los ISP que devuelvan los prefijos de red no utilizados que están fuera de sus bloques de direcciones asignadas. Esta petición ha tenido poco éxito porque muchas organizaciones contemplan estas direcciones como un activo.

9.5.7.2.- Asignación de direcciones para Internet privadas

El RFC 1918 pide que las organizaciones hagan uso de espacio de direcciones Internet privadas para los nodos que necesitan conectividad IP dentro de su red corporativa pero no requieren conexiones externas a Internet. Para este propósito, la IANA ha reservado, tal y como hemos mencionado anteriormente los siguientes bloques de direcciones:

10.0.0.0 - 10.255.255.255 (10/8 prefijo)
172.16.0.0 - 172.31.255.255 (172.16/12 prefijo)
192.168.0.0 - 192.168.255.255 (192.168/16 prefijo)

Una organización que elige utilizar direcciones de estos bloques reservados puede hacerlo sin contactar con la IANA o un registro Internet. Dado que estas direcciones no se inyectan nunca en el sistema de encaminamiento global de Internet, el espacio de direcciones puede ser utilizado simultáneamente por diferentes organizaciones.

La desventaja de utilizar este esquema de direccionamiento es que requiere un Traductor de Direcciones de Red (NAT - Network Address Translator) para el acceso a Internet. Sin embargo, el uso de un espacio de direcciones privado y NAT hace más sencillo a los clientes cambiar de ISP sin necesidad de renumerar sus estaciones. Los beneficios de este esquema de direcciones para Internet es que reduce la demanda de direcciones IP de modo que las organizaciones grandes pueden necesitar solo un bloque pequeño del espacio de direcciones IPv4.

9.5.7.3.- Asignación de Direcciones del espacio de Direcciones Reservado de Clase A

El draft Internet, "Observations on the use of Components of the Class A Address Space within the Internet" <draft-ietf-cidrd-classa-01.txt>, analiza la asignación de la mitad superior del espacio de direcciones de Clase A, actualmente reservado, mediante registros delegados. A medida que la demanda de direcciones IP continúe creciendo, parece que puede ser necesario asignar eventualmente el espacio de direcciones 64.0.0.0/2. Hay que tener en cuenta que el espacio 64.0.0.0/2 es enorme y representa el 25% del espacio de direcciones IPv4.

9.5.7.4.- Implicaciones de las Políticas de Asignación de Direcciones

El draft Internet, "Implications of Various Address Allocation Policies for Internet Routing" <draft-ietf-cidrd-addr-ownership-07.txt>, discute los aspectos fundamentales ha tener en cuenta mientras Internet desarrolla una nueva política de gestión y asignación de direcciones unicast. El draft compara los beneficios y limitaciones de una política de "propiedad de direcciones" con una política de "préstamo de direcciones".

La "propiedad de direcciones" supone que cuando se asigna un bloque de direcciones a una organización, permanece asignada a dicha organización mientras ésta desee mantenerla. Esto significa que el bloque de direcciones es "portable" y que la organización debería ser capaz de utilizarlo para conseguir acceso a Internet independientemente de dónde estuviera conectada a Internet la organización.

Por el contrario, el "préstamo de direcciones" supone que una organización obtiene un bloque de direcciones sobre la base de un préstamo. Cuando éste termina, la organización no puede seguir usando el bloque de direcciones, debe obtener nuevas direcciones y renumerar sus nodos.

Como hemos visto, el encaminamiento jerárquico necesita que las direcciones reflejen la topología de la red para poder permitir la agregación de rutas. El draft sugiere que hay dos problemas fundamentales que pueden impedir el direcionamiento jerárquico y el modelo de encaminamiento soportado por CIDR:

- La existencia continuada de rutas pre-CIDR que no pueden ser agregadas.
- Las organizaciones que cambian de ISP y continúan utilizando direcciones del bloque del ISP previo. El nuevo ISP no puede agregar el bloque de direcciones antiguo como parte de su agregación, de modo que inyecta una ruta de excepción en Internet. Si el número de excepciones crece continuamente, reducirá las ventajas de CIDR e impedirá la escalabilidad del sistema de encaminamiento de Internet.

El draft concluye con la recomendación de que los grandes proveedores, que pueden expresar sus destinos con un solo prefijo, deberían recibir bloques siguiendo el modelo de "propiedad de direcciones". Sin embargo, las asignaciones a sus clientes deberían seguir el modelo de "préstamo de direcciones". Esto supone que cuando una organización cambia de proveedor, se cancela le préstamo y debe renumerar sus nodos.

Este draft generó una gran discusión en la comunidad Internet sobre el concepto de propiedad y su significado en el encaminamiento global.

9.5.7.5.- "Procedures for Internet/Enterprise Renumbering" (PIER)

En el debate entre "propiedad de direcciones" y "préstamo de direcciones", es claro que la renumeración puede resultar un factor crítico. "Procedures for Internet/Enterprise Renumbering " (PIER) es un grupo de trabajo del IETF encargado de la tarea de desarrollar una estrategia de renumeración.

El RFC 1916 es una petición de PIER para que la comunidad Internet proporciona ayuda en el desarrollo de una serie de documentos que describen el modo en que una organización podría proceder a renumerar su red. El objetivo final de estos documentos es proporcionar educación y experiencia práctica a la comunidad Internet.

9.6. ENCAMINAMIENTO DE DATAGRAMAS IP

El encaminamiento es una de las funciones principales de IP y consiste en encontrar la ruta para alcanzar el destinatario final de cada datagrama. Esta función de encaminamiento es desarrollada cooperativamente por el nodo emisor y cada uno de los routers hacia los que se encamina el datagrama hasta alcanzar el nodo destino.

La Figura 9-35 muestra una visión simplificada de los procesos llevados a cabo en el nivel IP. Los datagramas que deben ser encaminados habrán sido generados bien en el propio nodo o en otro. En esté último caso, el sistema debe estar configurado para actuar como un router, o de otro modo los datagramas recibidos a través de los Interfaces de red que no sean nuestros serán borrados.

El objetivo de este apartado es analizar el **mecanismo de encaminamiento** IP, es decir, el modo en que IP toma las decisiones de encaminamiento para hacer llegar un datagrama hasta su destino a través de Internet. Además del mecanismo de encaminamiento, existe otra problemática conocida como **política de encaminamiento,** asociada a aspectos tales como el modo de intercambiar información de encaminamiento con routers vecinos y el modo de seleccionar cual es la ruta más adecuada hacia un destino que analizaremos en capítulos posteriores.

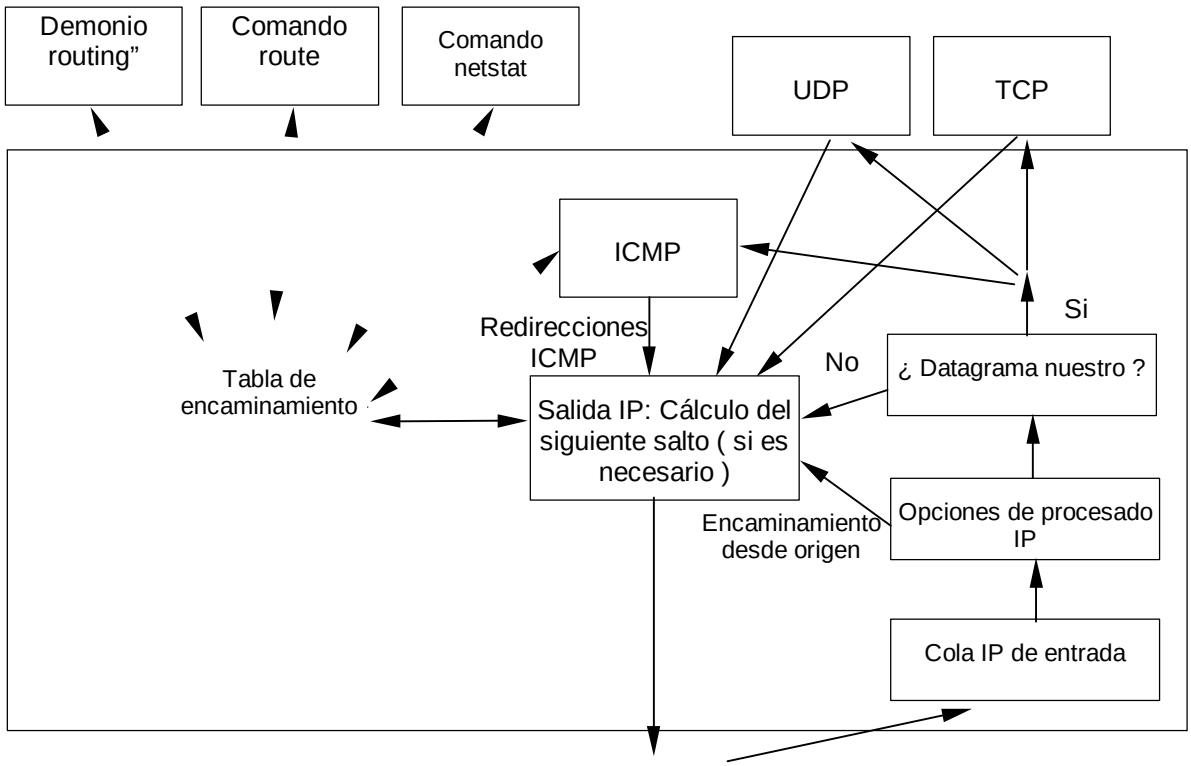

Figura 9.35

La **tabla de encaminamiento** constituye el elemento central del mecanismo de encaminamiento, debiendo utilizarse cada vez que se debe encaminar un datagrama. La actualización de la tabla se lleva a cabo por el demonio “routing”, aunque con mucha menos frecuencia de lo que IP la consulta (posiblemente una vez cada 30 segundos); además esta tabla también puede ser actualizada cuando se reciben mensajes ICMP de redirección, como veremos más adelante.

Para elegir el mejor camino hacia un cierto destino, el protocolo de encaminamiento debería examinar aspectos tales como la carga de la red, la longitud de los datagramas y el tipo de servicio especificado en la cabecera del datagrama. Sin embargo, la mayoría de las aplicaciones internet son más sencillas y seleccionan las rutas basándose exclusivamente en suposiciones físicas acerca del camino más corto.

Tanto los nodos emisores como los routers participan en el encaminamiento IP. El nodo debe tomar una decisión de encaminamiento cuando elige dónde enviar un datagrama, escogiendo el router al que entrega el mismo. Cada uno de los routers que recibe un datagrama debe reenviarlo, cuando no está dirigido a ellos, buscando en sus tablas de encaminamiento la mejor ruta para alcanzar el destino final del datagrama.

Podemos dividir el encaminamiento en dos clases: **encaminamiento directo** y **encaminamiento indirecto**. El **encaminamiento directo** consiste en la transmisión de un datagrama de un sistema directamente a otro, y constituye la base de toda comunicación Internet. Dos sistemas pueden hacer uso de encaminamiento directo sólo si están conectados al mismo sistema de transmisión físico. El **encaminamiento indirecto** se desarrolla cuando el destino no está directamente conectado a la misma red que el nodo de origen, obligando al emisor a enviar el datagrama a un router para su reenvío.

9.6.1. Encaminamiento Directo

La transmisión de un datagrama IP entre dos nodos conectados a una misma red física no involucra a ningún router. El nodo que envía un datagrama a otro nodo que está conectado a su misma red física puede hacerle llegar el mismo directamente mediante el envío a la red de una trama conteniendo el datagrama.

Para saber si el destinatario se encuentra conectado a su misma red, el emisor compara el **identificador de red** de la dirección IP de destino con el **identificador de red** de su propia dirección IP, cuando ambos coinciden, los dos nodos se encuentran conectados a la misma red física.

Las acciones que desarrolla en este caso el nodo emisor son las siguientes:

- Encapsula el datagrama en una trama de enlace de datos agregándole la cabecera correspondiente en función del protocolo de enlace de datos que deba emplear.
- Efectúa una resolución de la dirección IP en una dirección de enlace de datos, bien directamente a partir de su caché ARP, bien enviando una petición ARP para obtener la equivalencia.
- Envía la trama resultante a la red, directamente a su destino.

El encaminamiento directo es un caso especial de encaminamiento, donde los datagramas no atraviesan ningun router intermedio.

9.6.2. Encaminamiento Indirecto

Resulta mucho mas complejo que el anterior, puesto que el emisor debe identificar un router al que enviar el datagrama y éste debe enviar el datagrama hacia la red de destino, bien directamente o atravesando otros routers intermedios.

Los routers en TCP/IP forman una estructura cooperativa. Los datagramas pasan de router en router hasta alcanzar uno que pueda enviar el datagrama de forma directa hasta el nodo de destino final.

9.6.3. Tabla de Encaminamiento

Cada sistema dispone de una **tabla de encaminamiento** que contiene información acerca de **posibles destinos y el modo de alcanzarlos**. Cada vez que un nodo o un router desea transmitir un datagrama, consulta su **tabla de encaminamiento** para decidir adonde enviarlo.

Como ya hemos visto, las direcciones IP se asignan habitualmente de modo que las máquinas conectadas a una misma red física compartan el mismo identificador de red. Por ello, las tablas de encaminamiento sólo necesitan contener identificadores de red y no direcciones IP completas, de modo que resultan más pequeñas y el algoritmo de encaminamiento más eficiente. El tamaño de las tablas de encaminamiento depende del número de redes de Internet, y por lo tanto sólo crece cuando se añaden nuevas redes, siendo el tamaño de las tablas y su contenido independiente del número de ordenadores individuales conectados a cada red.

El hecho de elegir las rutas según la red de destino y no la dirección completa tiene ciertas servidumbres. Todo el tráfico dirigido a una red sigue el mismo camino, de modo que aún cuando existan varios caminos no podrán ser utilizados concurrentemente. Sólo el último router del camino que está directamente conectado con el nodo de destino puede determinar si el ordenador existe o está operativo. Por otro lado, los datagramas pueden seguir caminos distintos desde un origen hasta su destino y en el camino de vuelta, ya que los routers emplean sus tablas de manera independiente.

Una tabla de encaminamiento se compone de una serie de entradas, cada una de las cuales representa una ruta hacia un cierto destino y cuya información es la siguiente:

- **Dirección IP de destino**: puede ser una dirección de red o de nodo.
- **Dirección IP del siguiente salto.** Corresponde al primer router en la ruta hacia del destino, con el que se comparte red física.
- **Flags.** Son 5, y proporcionan información adicional sobre una ruta.
  - **U**: La dirección está conectada y por tanto la ruta está activa.
  - **G**: La dirección se corresponde con un router (**ruta indirecta**). Si este flag no aparece, el destino está conectado directamente (**ruta directa**).
  - **H**: La dirección se corresponde con la dirección completa de un nodo. Si este flag no aparece, el destino es una dirección de red: una identificador de red, o una combinación de red y subred.

D: La dirección fué creada mediante un mensaje ICMP de **redirección**.
M: La dirección fué modificada mediante un mensaje ICMP de **redirección**.

- **Identificación del interfaz** de red al que debe pasarse un datagrama para su transmisión. Los routers e incluso los nodos pueden estar conectados a varias redes a través de distintos interfaces; por ello es preciso no sólo saber a qué dirección debemos enviar un datagrama, sino por cuál de los distintos interfaces del sistema debemos hacerlo.

Un router sólo conoce el **siguiente salto** hacia el destino y no el camino completo hacia el mismo, excepto en los casos de los destinos conectados a su misma red. Las entradas e la tabla de encaminamiento siempre señalan a routers que pueden ser alcanzadas dentro de la misma red.

Una técnica utilizada habitualmente en las tablas de encaminamiento es consolidar múltiples entradas en una sola entrada por defecto. Esto supone que el mecanismo de encaminamiento IP busca en primer lugar en la tabla de encaminamiento la red de destino, si no aparece ninguna ruta hacia dicha red, el router envía el datagrama a un router configurado como **ruta por defecto**. El encaminamiento por defecto es especialmente útil cuando existe un conjunto pequeño de direcciones locales y sólo una conexión al resto de la Internet, como en el caso de sistemas conectados a una sola red física y que sólo pueden alcanzar una pasarela, por lo demás, una de las situaciones más habituales.

Aunque los nodos contienen también tablas de encaminamiento, sus tablas contienen la información mínima necesaria. La idea es obligar a los nodos a ceder la mayor parte de las funciones de encaminamiento a los routers, que son quienes conocen realmente las rutas hacia todos los destinos, de esta forma se simplifica el mantenimiento de las tablas de los usuarios evitando mantener rutas que nunca serán utilizadas.

La complejidad de la tabla de encaminamiento de un nodo depende de la topología de la red a la que tiene acceso:

- Un nodo aislado y no conectado a ninguna red tiene una tabla de encaminamiento con una sola entrada para el **interface de loopback**. En este caso, TCP/IP se emplear para comunicar distintas aplicaciones residentes en el mismo nodo.
- Un nodo conectado a una sola red de área local, y que sólo puede acceder a los nodos de dicha red tiene una tabla de encaminamiento con dos entradas: una para el **interface de loopback** y otra para la dirección IP de la red local.
- Un nodo que puede alcanzar otras redes a través de un unico router, además de las entradas anteriores incluye una dirección por defecto correspondiente al único router al que tiene acceso.
- Finalmente, en casos más complejos, se pueden añadir otras direcciones IP correspondientes enlaces o routers, sean nodos o redes específicas.

En el momento del arranque del sistema se inicializa la tabla de encaminamiento, incluyendo con la inicialización de cada interfaz una ruta directa para dicho interfaz. Para enlaces punto a punto y para el interfaz de loopback la ruta es hacia un nodo, para interfaces de difusión (Ethernet) la ruta es hacia dicha red. Las rutas a nodos o redes no directamente conectadas deberán introducirse en la tabla de encaminamiento manualmente o, generalmente, a partir de un fichero de configuración. Otro modo de inicializar esta tabla es utilizar mensajes ICMP de descubrimiento de rutas para encontrar la pasarela por defecto.

La mayor parte de las implementaciones de IP permiten establecer rutas hacia un nodo concreto. De esta manera el administrador de la red tiene un cierto control sobre el uso de la red y el control de acceso por motivos de seguridad.

9.6.4. Mecanismo de Encaminamiento IP

Conceptualmente el encaminamiento IP es simple, especialmente para un nodo. Si los nodos emisor y destino comparten un mismo enlace o una misma red, el datagrama IP será enviado directamente al destino. En caso contrario el nodo debe enviar el datagrama a un router por defecto, y dejar que el dicho router envíe el datagrama a su destino final. Este esquema tan simple es el que se utiliza en la mayor parte de las configuraciones.

IP puede recibir datagramas bien desde los protocolos de los niveles superiores como TCP, UDP, ICMP y IGMP que soliciten el envío datagramas, o bien desde el interfaz de red.

Cuando se recibe un datagrama desde el interfaz de red, IP comprueba en primer lugar la dirección IP de destino, si es una de sus propias direcciones o una dirección de difusión, entrega el datagrama al protocolo de nivel superior especificado en el campo protocolo de la cabecera IP. Si el datagrama no está destinado a ninguna de estas direcciones, bien se descarta el datagrama, o si el nodo está configurado como router, lo trata como un datagrama saliente y lo reenvía.

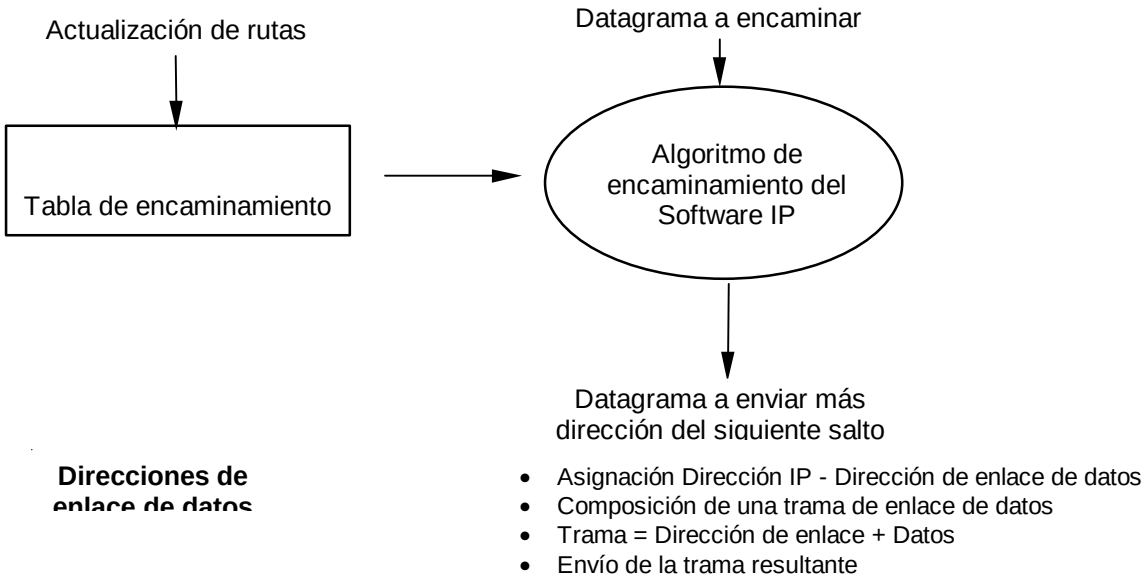

Figura 9.36

El encaminamiento IP se fundamenta en una transmisión basada en saltos. Como ya hemos visto, la tabla de encaminamiento no contiene rutas completas hacia ningún destino, excepto aquellos destinos que están directamente conectados al nodo emisor. Todo lo que proporciona la tabla de encaminamiento IP es la dirección del siguiente router al que debe enviarse el datagrama. El siguiente router está directamente conectado al nodo emisor y se supone que está realmente más próximo al destino que el nodo emisor.

Podemos resumir el mecanismo de encaminamiento IP en las siguientes reglas:

1. Se extrae el identificador de red de la dirección IP de destino del datagrama. Si coincide con el identificador de red de una de las direcciones IP del nodo emisor se enviará directamente el datagrama por el interfaz especificado en la tabla de encaminamiento.
2. Se busca en la tabla de encaminamiento una entrada que coincida con la dirección IP completa de destino (tanto identificador de red como de nodo). Si existe, se envía el paquete al router especificado en la tabla de encaminamiento (**siguiente salto**).
3. Se busca en la tabla de encaminamiento una entrada que se ajuste al identificador de red de la dirección de destino. Si aparece, se envía el paquete al router especificado. Todos los nodos de la red de destino se gestionan con una sola entrada en la tabla de encaminamiento. La comprobación del indentificador de red debe tener en cuenta las posibles mascaras de subred.
4. Se busca en la tabla de encaminamiento una entrada por defecto. Si se encuentra, se envía el datagrama al router especificado como ruta por defecto.

Si no se cumple ninguno de los siguientes pasos anteriores, no es posible enviar el datagrama. Cuando un nodo genera un datagrama que no es posible enviar, generalmente se devuelve un mensaje de error de “nodo no alcanzable” o “red no alcanzable” a la aplicación que generó el datagrama, y si es un router el que recibe el datagrama que no puede reenviarse lo descarta generando un mensaje ICMP de destino no alcanzable al nodo emisor del mismo.

El mecanismo de encaminamiento IP no altera el datagrama original. Las direcciones origen y destino que figuran en la cabecera del datagrama no se modifican. El algoritmo de encaminamiento busca la dirección IP a la cual debería enviarse el datagrama, que generalmente es la dirección de un router; esta dirección se conoce como la **dirección del siguiente salto,** pero debe tenerse en cuenta que esta dirección no se almacena dentro del datagrama.

Si el datagrama puede enviarse directamente, la dirección obtenida de la tabla por el mecanismo de encaminamiento debería ser la misma que la dirección de destino del datagrama.

9.6.5. Manejo de datagramas entrantes

Cuando un datagrama IP llega a un nodo, y si la dirección de enlace de datos que figura en la cabecera de la trama coincide con la dirección de enlace del nodo, el interfaz de enlace de datos lo envía al nivel IP para su procesamiento.

Si la dirección de destino IP coincide con la dirección IP del nodo, el nivel IP aceptará el datagrama y lo pasará al protocolo adecuado (TCP, UDP, ICMP, …) para su procesamiento. Si por el contrario la dirección IP de destino no coincide con la dirección propia, el datagrama se descartará en el caso de un nodo o bien se reencaminará en el caso de un router. Un nodo también acepta datagramas cuya dirección de destino IP se una **dirección de difusión**.

A diferencia de los nodos, los routers efectúan reenvío de datagramas, mediante el mecanismo de encaminamiento presentado anteriormente, hasta que el datagrama llegue a su destino final.
El datagrama habrá alcanzado su destino final cuando la dirección IP de destino coincida con la dirección IP del nodo que lo ha recibido, entonces, IP pasará el datagrama al protocolo adecuado de nivel superior para su procesamiento (en función del protocolo de la cabecera del datagrama IP).

Si el datagrama no ha alcanzado su destino final y es recibido por un router, el protocolo IP reenviará el mensaje al siguiente router en función de su tabla de encaminamiento. Antes de ello, decrementa el campo **Tiempo de Vida** de la cabecera IP, descartando los datagramas cuando este campo alcanza el valor cero, reenviándolo mientras sea positivo es positivo.

El hecho de diferenciar las funciones de encaminamiento desarrolladas por un nodo o un router se debe a las siguientes razones:

- Si un nodo recibe un datagrama dirigido a otro nodo se debe a que ha habido algún mal funcionamiento en el esquema de direccionamiento, o el encaminamiento Internet, el problema no sería detectado si el nodo tomara acciones correctivas reenviando del datagrama hacia su destino correcto.
- Los routers emplean protocolos especiales para informar de los errores y para mantener actualizadas las rutas de sus tablas de encaminamiento, mientras que por simplicidad los nodos no lo hacen, es por ello que las funciones de corrección de los mecanismos de encaminamiento deben ser limitados a los nodos configurados como router.

¿ Que ocurre si en un nodo no hay rutas por defecto y no se ha encontrado una ruta para un destino dado? Si el datagrama fué generado el propio nodo, se enviará un error a la aplicación que envió el datagrama, que puede ser "nodo no alcanzable" o "red no alcanzable", en caso contrario, el router descartará el datagrama generando un mensaje ICMP de error de nodo o red no alcanzable

9.7.- IP V6

Si bien la adopción de medidas paliativas como CIDR puede dar un respiro momentáneo en el problema de las direcciones y tablas de routing, es evidente que la Internet no puede seguir mucho tiempo en esta situación. Aunque sea el más importante, la escasez de direcciones es sólo uno de los problemas que acechan a la Internet en su estado actual.

En un intento por resolver lo mejor posible todos los problemas detectados en el protocolo IP el IETF empezó a trabajar ya en 1990 en una nueva versión de IP con el fin de producir un protocolo de nivel de red mucho más avanzado que soportara nuevas exigencias como acceso remoto a datos, videoconferencias, distribución de audio, control remoto de dispositivos y que mejorara la seguridad para usos comerciales. Los objetivos planteados fueron los siguientes:

- **Direcciones**: Establecer un espacio de direcciones que no se agote en el futuro previsible.
- **Eficiencia**: Reducir el tamaño de las tablas de routing. Simplificar la cabecera de los datagramas IP de modo que los routers puedan procesar los paquetes más rápidamente.
- **Seguridad**: Ofrecer mecanismos que permitan incorporar fácilmente en el protocolo medidas de seguridad (privacidad y validación) mediante técnicas criptográficas.
- **Tipo de servicio**: Manejar mejor los diferentes tipos de servicio, en especial para poder ofrecer garantías de calidad de servicio y para el transporte de datos en tiempo real.
- **Multicasting**: Facilitar el uso de multicasting.
- **Movilidad**: Permitir la movilidad de un host sin necesidad de cambiar su dirección.
- **Evolución**: Contemplar un mecanismo que permita extender el protocolo en el futuro.
- **Compatibilidad**: Permitir la coexistencia del protocolo nuevo con el viejo.

El IETF hizo una convocatoria pública (RFC 1550) en julio de 1992 para que se presentaran propuestas de como podría ser el nuevo protocolo. De las propuestas iniciales se fueron descartando las consideradas menos interesantes, y finalmente la propuesta finalmente adoptada fue un híbrido de dos de las presentadas.

Durante el período en que se celebraban en el seno del IETF las discusiones sobre las diferentes propuestas el nuevo protocolo se denominó IPng (IP next generation, de la famosa serie de televisión Star Trek cuya segunda parte recibió esa denominación). Finalmente el nombre oficial elegido fue IP Versión 6 o abreviadamente IPv6 (el IP actualmente en uso es Versión 4, y la versión 5 se había utilizado ya para un protocolo experimental denominado ST, STream protocol).

En el mercado existen implementaciónes de IPv6 para hosts y para routers, aunque su uso está muy limitado todavía.

El nuevo protocolo coincide con IPv4 en ofrecer un servicio CLNS de entrega de datagramas. Se contemplan sin embargo algunas opciones que, con la ayuda de otros protocolos (RSVP o ATM por ejemplo) permiten convertirlo en un servicio con calidad de servicio.

IPv6 no es realmente compatible con IPv4, ya que utiliza un formato diferente para la cabecera, pero lo es con todos los demás protocolos de Internet. Para la implantación del nuevo protocolo se ha previsto un plan de migración gradual en el que vayan apareciendo 'islas' IPv6 donde se utilice el nuevo protocolo y se aprovechen sus ventajas; para la interconexión de esas islas a través del backbone IPv4 se prevé utilizar túneles, de forma similar a lo que se hace actualmente con el tráfico MBone. La red experimental que se utiliza para las pruebas de IPv6 se conoce como 6Bone. En España ya hay una red 6Bone funcionando entre varias universidades españolas.

Los protocolos de routing se han tenido que modificar para tener en cuenta las características propias y el nuevo formato de direcciones que utiliza IPv6; así se ha creado por ejemplo RIPv6 y OSPFv6.

Muy brevemente algunas de las principales virtudes de IPv6 son las siguientes:

- Las direcciones son de 16 bytes, lo cual da un espacio de direcciones increíblemente grande, suficiente con creces para todo futuro uso previsible por muy ineficiente que sea la asignación de las mismas.
- La cabecera se simplifica, pasando de 13 a 7 campos y acelerando así el proceso en los routers.
- A la vez que se reduce la cabecera se ha mejorado el soporte de los campos opcionales haciendo más fácil añadir opciones adicionales.
- Con el fin de mejorar el rendimiento no se permite la fragmentación de los datagramas en los routers.
- La seguridad (validación y privacidad) es ahora una parte fundamental del protocolo, servicio que puede ser utilizado por cualquier aplicación que no contemple aspectos de seguridad. En particular se soportan servicios de autenticación y privacidad.

- Es posible asociar datagramas con una clase de servicio particular, realizándose el encaminamiento de los datagramas en base a estas clases. En particular esto resulta importante para soportar servicios en tiempo real y especificar niveles de prioridad para determinar la estrategia de descarte en caso de congestión.

9.7.1.- La Cabecera en IPV6

Probablemente el rango y formato de direcciones es lo más llamativo de IPv6, pero no es en modo alguno la única nueva funcionalidad. Para mostrar la funcionalidad del protocolo vamos a describir la estructura de la cabecer de los datagramas IPv6:

La cabecera más simple de IPv6 tiene un tamaño fijo de 40 bytes, tal y como aparece en la figura

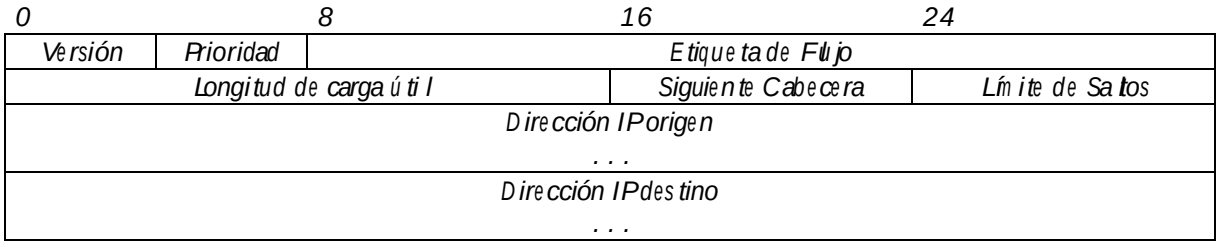

Figura 9.37.- Formato de la cabecera de un datagrama IP Versión 6

El campo **versión** vale 6.

El campo **prioridad** permite al emisor especificar la prioridad que desea para cada paquete que envía en relación con otros paquetes del mismo origen. De hecho, este campo permite al emisor identificar dos características separadas relacionadas con la prioridad. En primer lugar los paquetes son clasificados como parte de tráfico para el que el emisor está proporcionando control de congestión o no; y en segundo lugar, a los paquetes se le asigna uno de los ocho niveles de prioridad relativa dentro de cada clasificación.

Tráfico con control de congestión. Se refiere al tráfico para el que el emisor reacciona como respuesta a una congestión, por ejemplo TCP. En este caso es aceptable un retraso variable en la entrega de los paquetes e incluso la recepción desordenada de los mismos. IPv6 define las siguientes categorías de tráfico con control de congestión con prioridades decrecientes:

- Tráfico de control de Internet: El más importante, especialmente en momentos de gran congestión. Por ejemplo, protocolos de encaminamiento como OSPF y BGP necesitan recibir actualizaciones sobre las condiciones del tráfico parar ajustrar las rutas y evitr la congestión. Los protocolos de gestión como SNMP necesitan notificar la congestión a las aplicaciones de gestión para realizar una reconfiguración y hacer frente a la congestión.
- Tráfico interactivo: Se trata de las conexiones on-line entre usuarios y nodos. La eficiencia del usuario depende críticamente una rápida respuesta durante las sesiones interactivas, de modo que debe minimizarse el retraso.
- Transferencia masiva atendida: En estas aplicaciones que suponen la transferencia masiva de información el usuario debe esperar a que la transferencia finalice. Esta categoría difiere de la anterior en que el usuario está preparado para aceptar una espera más importante que durante un diálogo interactivo. Buenos ejemplos son FTP o HTTP.
- Transferencia masiva no atendida: En estas aplicaciones no se espera una entrega instantánea. Generalmente el usuario no está esperando que la transferencia se complete sino ocupado en otras tareas. El mejor ejemplo de esta categoría es el correo electrónico.
- Tráfico de relleno: Este tráfico se gestiona desatendidamente cuando otros tipos de tráfico ya se han envido. Las noticias USENET constituyen un buen ejemplo.
- Tráfico sin caracterizar: Si el nivel de aplicación no proporciona ninguna información sobre la prioridad de la información, se le asignará el valor más bajo de prioridad.

Tráfico sin control de congestión. Este tráfico precisa de una velocidad de transferencia y retraso en la entrega constantes. Como ejemplos podemos citar el video y audio en tiempo real, en los cuales no tiene sentido retransmitir paquetes descartados. También se han reservado ocho niveles de prioridad

para este tipo de tráfico, desde la prioridad más baja, 8 (con más probabilidad de descarte), hasta la más alta, 15 (menos probabilidad de descarte). En general, el criterio de asignación de prioridad será el modo en que se degradará la calidad del tráfico recibido en caso de descarte de paquetes. Por ejemplo, el audio de baja fidelidad como una conversación telefónica recibiría una alta prioridad, puesto que la pérdida de unos pocos paquetes será rápidamente perceptible en forma de clicks y zumbidos en la línea. Sin embargo, una señal de vídeo de alta fidelidad contiene suficiente redundancia para que la pérdida de unos pocos paquetes pueda pasar desapercibida y por lo tanto recibirá una prioridad relativamente baja.

Debe tenerse en cuenta que no existe ninguna relación de prioridad implícita entre las prioridadades de ambos tipos de tráfico y ésta sólo tiene sentido dentro de cada categoría.

IPv6 define un flujo como una secuencia de paquetes enviados desde un origen particular hacia un destino (unicast o multicast) para el cual el emisor desea un tratamiento especial por parte de los routers que intervienen (por ejemplo tráfico en tiempo real o con reserva de capacidad vía RSVP). El campo **etiqueta de flujo** permite identificar los paquetes que pertenecen dicha secuencia.; esto en cierto modo es una aproximación a las redes de circuitos virtuales. El 'circuito virtual' queda identificado por la combinación del valor de este campo y de la dirección origen. Se espera que las implementaciones aprovechen la tecnología de red subyacente para ofrecer una calidad de servicio adecuada a las necesidades; por ejemplo si los datagramas discurren por una red ATM puede haber varios circuitos virtuales con diferentes características de calidad de servicio; en función del valor que apareciera en el campo 'etiqueta de flujo' el router decidiría por que circuito virtual debe enviar el datagrama.

Desde el punto de vista del emisor, un flujo será generalmente una secuencia de paquetes que son generados desde una aplicación y que tienen los mismos requisitos de servicios. Un flujo puede comprender una sola o múltiples conexiones TCP; un ejemplo de esto último sería una aplicación típica de transferencia múltiple de ficheros que hacen uso de varias conexiones TCP. Una misma aplicación puede generar uno o varios flujos. Un ejemplo de esto sería una conferencia multimedia, que podría tener un flujo para audio y otro para ventanas gráficas, cada uno con requisitos de transferencia diferententes en términos de velocidad de transferencia, retraso, y variación de retraso.

Desde el punto de vista del router, un flujo es una secuencia de paquetes que comparten atributos con efecto el el modo en que son tratados por el router. Esto incluye el camino, asignación de recursos, requisitos de descarte, registro, y atributos de seguridad. El router puede tratar los paquetes de diferentes flujos de modo diferente haciendo uso de varias técnicas, como asignar diferentes espacios de almacenamiento en buffers, dando precedencias diferentes en términos de reencaminamiento, y pididiendo diferentes calidades de servicios de las subredes.

Ninguna etiqueta de flujo tiene significado especial; el manejo especial para un paquete deberá ser indicado de otro modo. Por ejemplo, un emisor puede negociar o pedir un tratamiento especial de los routers mediante un protocolo de control o en las cabeceras de extensión del paquete.

En principio, todos los requisitos de usuario para un flujo particular podrían ser definidos en una cabecera de extensión incluida en cada paquete. Si queremos dejar abierto el concepto de flujo para incluir una amplia variedad de requisitos, este diseño podría originar grandes cabeceras de paquetes. La alternativa, adoptada en IPv6 es la etiqueta de flujo, estando los requisitos de flujo definidos con anterioridad a que comience el flujo de paquetes y asignando una única etiqueta a todos ellos. En este caso, el router debe almacenar la información de los requisitos de cada uno de los flujos.

Existen una serie de reglas aplicables a las etiquetas de flujo:

- Los nodos o routers que no soporten el campo etiqueta de flujo deben poner el campo a 0 cuando originan un paquete, pasar el campo sin modificación cuando lo reenvían e ignorar el campo cuando lo reciben.
- Todos los paquetes originados por un emisor determinado con la misma etiqueta de flujo deben tener la misma dirección de destino, dirección de origen, prioridad, y cabeceras de extensión salto a salto y encaminamiento (si están presentes). El objetivo es que un router pueda decidir cómo encaminar un paquete inspeccionando simplemente la etiqueta sin necesidad de examinar el resto de la cabecera.

- El emisor asigna una etiqueta a un flujo. La etiquetas nuevas pueden ser elegidas aleatoria y uniformemente entre 1 y 224, con la restricción de no reutilizar una etiqueta para un nuevo flujo durante el tiempo de vida de un flujo existente.

El campo **longitud de carga útil** indica el tamaño del paquete en bytes, excluidos los 40 bytes de la cabecera. El valor máximo es 65535, que será el tamaño máximo de paquete IPv6.

El campo **siguiente cabecera** indica si esta cabecera esta seguida por alguna de las cabeceras opcionales. Si no hay cabeceras opcionales este campo indica el protocolo del nivel de transporte al que pertenece este paquete utilizando los mismos códigos numéricos que en IPv4 (ver tabla 5.4), en caso contrario identifica el tipo de cabecera que aparece a continuación.

El campo **límite de saltos** equivale al antiguo campo TTL, pero se ha decidido ponerle un nombre que refleje su uso real. Dado que el campo tiene 8 bits el máximo número de saltos que puede especificarse es 255. Algunos opinaban que este valor era demasiado bajo y que en algunos casos podría plantear problemas. Actualmente ya hay situaciones en la Internet donde se esta próximo a los 64 saltos. El límite de saltos se ajusta a un valor máximo por el emisor del paquete y es decrementado en 1 por cada router que reencamina el paquete. Este se descarta cuando el valor alcanza el valor 0.

Por último, los campos **dirección de origen** y **dirección de destino** corresponden a las direcciones de 16 bytes que hemos visto.

Los campos que han desaparecido de la cabecera IPv4 son los siguientes:

- El campo *Longitud de cabecera* carece de sentido puesto que la cabecera tiene longitud fija.

- El campo *protocolo* no aparece pues su función la desempeña el campo *siguiente cabecera*.

- El campo *checksum* se ha suprimido ya que no parece justificada la pérdida de rendimiento que supone su cálculo en cada salto ante la rara eventualidad de que se produzca un error en este nivel, teniendo en cuenta que normalmente tanto el nivel de enlace como el de transporte realizan su propia comprobación de errores.

- Todos los campos relativos a fragmentación han desaparecido, porque en IPv6 sólo se permite la fragmentación en origen, y ésta se controla con cabeceras adicionales; antes de enviar un datagrama el emisor debe realizar un *Path MTU Discovery*, mediante el cual averigua el tamaño máximo de paquete que puede enviar a un destino determinado. Si durante el envío las condiciones cambian y el MTU se reduce (por ejemplo porque la ruta cambie dinámicamente a otra que no permite un valor tan grande del MTU) el router que recibe el datagrama demasiado grande devuelve un mensaje de error indicando al emisor que reenvíe a partir de ese momento la información creando paquetes más pequeños. Además en IPv6 cualquier nodo y cualquier red deben aceptar paquetes de 576 bytes al menos, lo cual hace menos probable la fragmentación.

9.7.2.- Cabeceras extendidas

En IPv4 la cabecera podía contener un número bastante limitado de campos opcionales, principalmente para diagnóstico de problemas de routing; debido a su corta extensión esos campos eran poco utilizados. En IPv6 se ha habilitado un mecanismo más flexible y eficiente; las cabeceras opcionales aparecen como cabeceras adicionales al continuación de la cabecera estándar, identificadas por el valor de los campos *siguiente cabecera*. Cada cabecera de extensión también contiene su propio cambo siguiente cabecera, lo que permite un número casi ilimitado de extensiones y por lo tanto soportar nuevas facilidades en el futuro. Se han definido las siguientes cabeceras:

La cabecera *salto-a-salto* indica información que debe ser examinada por todos los routers por los que pase este datagrama. De esta forma se permite que los routers intermedios prescindan de analizar cabeceras que no les incumben, aumentando con esto la eficiencia. Hasta ahora solo se le ha definido a esta cabecera una opción, que permite especificar datagramas de longitud superior a 64 KB; estos datagramas (que pueden llegar a tener hasta 4 GB) se conocen como *jumbogramas*.

La cabecera *encaminamiento* realiza las funciones combinadas de los encaminamientos desde el origen estricto y flexible de IPv4; el máximo número de direcciones que pueden especificarse es de 24. El formato de la cabecera es el siguiente:

Siguiente Cabecera	*Tipo de encam. = 0*	*Nº de Direcciones*	*Siguiente Dirección*
Reservado	*Bit de máscara estricto/flexible*		
Dirección 0			
Dirección 1			

Figura 9-38: Formato de la cabecera de encaminamiento

El campo *siguiente dirección* es incrementado por cada router para apuntar a la dirección del siguiente router que hay que visitar.

La cabecera *fragmentación* se utiliza para fragmentar un paquete usando un mecanismo similar al de IPv4, con la diferencia de que en IPv6 el emisor del paquete es el único autorizado a fragmentarlo, cosa que hará cuando un paquete sea rechazado por alguno de los routers por los que debía pasar por ser excesivamente grande; no se permite la fragmentación 'en ruta', lo cual también ha simplificado notablemente la complejidad de proceso en los routers.

La cabecera *autenticación* ofrece mediante técnicas criptográficas un mecanismo de firma digital por el cual el receptor de un paquete puede estar seguro de que quien se lo ha enviado es quien dice ser.

La cabecera *carga útil de seguridad encriptada* permite el envío de información encriptada para que solo pueda ser leída por el destinatario.

La cabecera *opciones de destino* contiene información para su examen por el nodo de destino.

9.7.3.- Direcciones en IPv6

Una dirección IPv6 está compuesta por 16 bytes. Los primeros bits identifican el tipo de dirección, de manera análoga a IPv4. Existen 21 clases diferentes de direcciones, pero no todas tienen asignado el mismo rango, y la mayoría están reservadas para usos futuros.

Se contempla la asignación de varias direcciones a un mismo interfaz, pudiendo utilizarse cualquiera de las direcciones unicasta para identificar unívocamente al nodo. La combinación de direcciones más largas y de direcciones múltiples para un interfaz permiten aumentar la eficiencia del encaminamiento. En IPv4, las direcciones no tienen una estructura que facilite el encaminamiento y por lo tanto los routers necesitan mantenet grandes tablas de encaminamiento. El uso de direcciones más larga permite la agregación de las mismas por jerarquías de rec, proveedor de acceso, proximidad geográfica, pertenencia a una coroporación, etc. Dicha agregación permite reducir el tamaño de las tablas y por lo tanto el tiempo de consulta a las mismas. El uso de múltiples direcciones permitiría a un subscriptor que utiliza varios proveedores de acceso a través de la misma interfaz tener direcciones separadas agregadas en el espacio de direccionamiento de cada proveedor. El formato de la dirección se indica por los primeros bits de acuerdo con la siguiente tabla:

Espacio de Asignación	Prefijo (binario)	Fracción del espacio de direcciones
Reservado para compatibilidad con IPv4	0000 0000	1/256
No asignado	0000 0001	1/256
Reservado for NSAP Allocation	0000 001	1/128
Reservado for IPX Allocation	0000 010	1/128
No asignado	0000 011	1/128
No asignado	0000 1	1/32
No asignado	0001	1/16
No asignado	001	1/8
Dirección unicast basada en el proveedor	010	1/8
No asignado	011	1/8
Reservada para unicast basada en la Geografía	100	1/8
No asignado	101	1/8
No asignado	110	1/8
No asignado	1110	1/16
No asignado	1111 0	1/32
No asignado	1111 10	1/64
No asignado	1111 110	1/128
No asignado	1111 1110 0	1/512
Direcciones para uso de enlace local	1111 1110 10	1/1024
Direcciones para uso de ubicación local	1111 1110 11	1/1024
Direcciones multicast	1111 1111	1/256

Tabla 9-20. Asignación de direcciones.

El primer campo de una dirección IPv6 es el prefijo formato de longitud variable que identifica la categoría de la dirección. La Tabla 9-20 indica la asignación actual de direcciones.

Se ha previsto un rango específico para direcciones IPv4. De esta forma cualquier dirección IPv4 puede incluirse en un datagrama IPv6.

IPv6 considera tres tipos de direcciones:

- Unicast: Un identificador para un solo interfaz. Un paquete enviado a una dirección unicast se envía al interfaz identificado por dicha dirección.
- Anycast: Un identificdor para un conjunto de interfaces (generalmente pertenecientes a nodos diferentes). Un paquete enviado a una dirección anycast se entrega a uno de los interfaces identificados por dicha dirección (el más próximo de acuerdo con la medida de distancia de los protocolos de encaminamiento).
- Multicast: Un identificador para un conjunto de interfaces (generalmente pertenecientes a nodos diferentes). Un paquete enviado a una dirección multicast se entrea a todos los interfaces identificados por la dirección.

9.7.3.1.- Direcciones Unicast

Las direcciones unicast pueden estructurarse de varias formas: globales basadas en el proveedor ("provider-based global"), de enlace local ("link-local"), de ubicación local ("site-local"), "IPv4-compatible IPv6", y "loopback".

Una dirección unicast "provider-based global" proporciona direccionamiento global en el universo completo de nodos conectados. La dirección tiene cinco campos tras el formato prefijo.

3	*m bits*	*n bits*	*o bits*	*p bits*	*125-m-n-o-p bits*
010	*ID registro*	*ID Proveedor*	*ID suscriptor*	*ID subred*	*ID Interfaz*

Figura 9-39: Dirección unicast

la autoridad de registro identificada en la cabecera asignará identificadores únicos a los proveedores que a su vez proporcionarán identificadores únicos a sus suscriptores a los que proporcionan acceso a Internet. Estas organizaciones dividirán los bits restantes entre las direcciones de subred y de interfaz.

- ID Registro: Identifica la autoridad de registro que asigna la porción de proveedor de la dirección.
- ID Proveedor: Un proveedor de servicio internet que asigna la porción del suscriptor de la dirección.
- ID Suscriptor: Distingue entre múltiples suscriptores que utilizan el mismo identificador de proveedor en sus dirección.
- ID Subred: Un grupo de nodos topologicamente conectados dentro de la red del subscriptor.
- ID Interfaz: Identifica un interfaz de nodo entro el grupo de interfaces identificados por el prefijo de subred.

En la actualidad no hay una longitud fija asociada a cada uno de estos campos. Sin embargo, desde el punto de vista de un administrador de red, los identificadores de Subred e Interfaz son los elementos que le conciernen.

El subscriptor puede tratar estos campos de varios modos. Una posibilidad en una instalación basada en una LAN es utilizar un campo de interfaz de 48 bits y utilizar la dirección MAC IEEE 802 para dicho campo. Los bits restantes serían el campo ID Subred, identificando una LAN particular en la instalación del subscriptor. Este medida (utilizada ya en redes OSI y ATM) permite la autoconfiguración de los sistemas: el equipo fija la parte local de su dirección (los últimos seis bytes) y el router le informa de la parte de red (los primeros diez), con lo que la conexión de nuevos equipos a una red es realmente 'Plug-and-Play'. Un red de una empresa que estuviera conectada a través de un proveedor podría cambiar de proveedor y cambiar su dirección de red sin más que cambiar en el router la parte de red de la dirección. Los ordenadores obtendrían el nuevo prefijo del router y le añadirían cada uno la parte de host correspondiente. El uso de una dirección MAC garantiza que la dirección sea siempre única. La autoconfiguración también facilita grandemente la movilidad de equipos.

o bits	*p bits*	*48 bits*
Prefijo de lsuscriptor	*ID subred*	*Dirección IEEE 802*

Figura 9-40

IPv6 contempla direcciones unicast de uso local. Los paquetes con dichas direcciones sólo pueden encaminarse localmente, es decir, dentro de una subred o conjunto de subredes de un subscriptor dado.

Se han definido dos tipos de direcciones de uso local:

- Las de enlace local que se utilizan sobre un enlace local o una subred para autoconfiguración de la dirección, descubrimiento de vecinos o cuando no hay routers presentes. No pueden integrarse en el esquema de direccionamiento global.

10 bits	*n bits*	*118-n bits*
1111111010	*0*	*ID Interfaz*

Figura 9-41

- Las de ubicación local se emplean cuando no se desea registrarse para obtener una dirección Internet oficial. Su formato permite una migración fácil si porsteriormente desea integrarse en un esquema de direcciones global, puesto que la porción significativa de la dirección está confinada en los bits menos significatios no utilizados en las direcciones globales.

10 bits	*n bits*	*m bits*	*118-n-m bits*
1111111010	*0*	*ID Subred*	*ID Interfaz*

Figura 9-42

Cuando una organización está preparada para su conexión global, los bits que preceden al identificador de Subred se reemplazan por un prefijo global (p.e., 010 + registry ID + provider ID + subscriber ID).

Un aspecto clave en el despliegue de IPv6 es la transición desde IPv4 a IPv6, puesto que necesariamente habrá un período de coexistencia para el cual se han pensado las direcciones IPv6 compatibles IPv4. Esta forma de direccionar consiste en utilizar una dirección IPv4 de 32 bits en los 32 bits menos significativos de la dirección con un prefijo de 96 ceros.

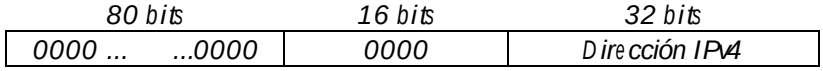

Figura 9-43

Esto permite a los sistemas de encaminamiento IPv4 transportar, haciendo uso de una técnica conocida como “tunel” , tráfico IPv6 mientras dure la transición entre ambas versiones. En esencia, un paquete IPv6 está encapsulado dentro de un datagrama IPv4. Con el uso de estas direcciones, el "tunneling" está automatizado en el sentido de que la dirección de destino IPv4 puede derivarse de la dirección IPv6 address, evitando la necesidad de configurar explícitamente el mapeo entre direcciones IPv6 y direcciones IPv4.

La dirección unicast 0:0:0:0:0:0:0:1 se denomina dirección loopback address. Puede emplearlos un nodo para envirarse un paquete IPv6 a sí mismo; estos paquetes nunca son enviados a la red permaneciendo dentro del nodo.

9.7.3.2.- Direcciones Anycast

Una dirección “anycast” permite a un emisor especificar que desea contactar con un nodo cualquiera de un grupo mediante una sola dirección. Un paquete con una de estas direcciones será encaminado al interfaz más próximo de acuerdo con la medida de distancia del router. Un ejemplo de uso de estas direcciones es la especificación dentro de una cabecera de encaminamiento de una dirección intermedia a lo largo de una ruta. La dirección anycast podría referirse al grupo de routers asociados con un proveedor particular o una subred, indicando de este modo que el paquete debe encaminarse a través del proveedor o internet del modo más eficiente. También permite, por ejemplo, acceder a un servidor multihomed haciendo balance de carga entre varias interfaces, o por aquella que está mas cerca del solicitante, o para comunicarse con cualquiera de los routers de la red propia cuando se desea acceder a ella desde el exterior (“mobile computing”).

Las direcciones anycast se asignan del mismo espacio de direcciones que las unicast. Así, los miembros de cualquier grupo anycast deben configurarse para reconocer dicha dirección, y los routers para mapear una dirección anycast a un grupo de direcciones de interfaz unicast.

Está predefinida una forma particular de dirección anycast, la dirección anycast router-subred. El campo prefijo de subred identifica una subred específica. El prefijo de subred identifica una subred específica. Por ejemplo, en un espacio de direcciones globales basadas en proveedor, el prefijio de subred es de la forma (010 + ID registro + ID proveedor + ID subscriptor + ID subnet). La dirección anycast es idéntica a la unicast para un interfaz de esta subred, con la porción del identificador de interfaz puesta a ceros. Cualquier paquete enviado a esta dirección será entrefagado a un router de la subred; todo lo que se necesita es insertar el ID interfaz correcto en la dirección anycast para componer la dirección unicast de destino.

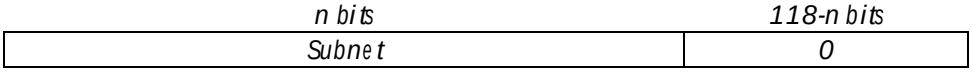

Figura 9-44

9.7.3.3.- Direcciones Multicast

IPv6 incluye la capacidad de direccionar un grupo predefinido de interfaces con una sola dirección multicast. Un paquete con una dirección multicast será entregado a todos los miembros del grupo. El formato de las direcciones multicast IPv6 es más complejo que el de las de IPv4.

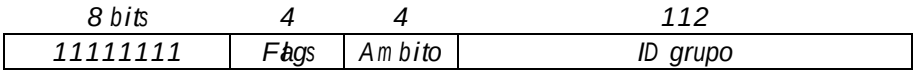

Figura 9-45

En la actualidad el campo flags consta de 3 ceros seguidos por el bit T:

- T = 0: Indicata un dirección permanentemente asignada (well-known) por la autoridad de numeración global de internet.
- T = 1: Indicata una dirección multicast no permanentemente asignada o transitoria.

El campo Ambito (scope) se utiliza para limitar el alcance del grupo multicast. Sus valores posibles son:

0: reserved
1: node -local
2: link-local
3: unassigned
4: unassigned
5: site-local
6: unassigned
7: unassigned
8: organization-local
9: unassigned
10: unassigned
11: unassigned
12: unassigned
13: unassigned
14: global
15: reserved

El ID grupo identifica a un grupo multicast, permanente o transitorio, dentro de un cierto ámbito. En el caso de direcciones permanentes, la dirección en sí es independiente del campo ámbito, pero dicho campo limita el área de entrega de un paquete concreto. Por ejemplo, si al grupo "NTPservers" se le asigna un dirección multicast permanente con un ID grupo 43 (hex):
FF05:0:0:0:0:0:0:43 señala a todos los servidores NTP con la misma ubicación que el emisor.
FF0E:0:0:0:0:0:0:43 indica a todos los servidores NTP de internet.

Las direcciones multicast no permanentemente asignadas sólo tienen significado dentro del ámbito especificado, permitiendo que se reutilice el mismo ID grupo con diferentes interpretaciones en redes distintas.

La capacidad de multicasting resulta muy útil en muchos contextos. Por ejemplo, prmite a nodos y routers enviar mensajes de descubrimiento de vecinos sólo a aquellas máquinas que están registradas para recibirlos, eliminando la necesidad de que el resto tengan que examinar y descartar paquetes irrelevantes. O también, crear direcciones de broadcast de subred en una LAN, asingnando un ámbito de enlace local a un ID grupo configurado en todos los nodos de la misma.

No se ha previsto ninguna dirección específica para broadcast, ya que esto se considera un caso particular de multicast.

9.7.3.4.- Notación

La escritura de direcciones de 16 bytes usando el sistema tradicional resulta muy farragosa; por ejemplo:

128.0.0.0.0.0.0.1.35.69.113.137.171.205.239

Para evitarlo se ha diseñado una notación en la que las direcciones se escriben como 8 grupos de 4 dígitos hexadecimales separados por dos puntos; por ejemplo la dirección anterior se escribiría:

8000:0000:0000:0000:0123:4567:89AB:CDEF

Dado que muchas direcciones contendrán gran cantidad de ceros se ofrece la posibilidad de utilizar una notación abreviada en la que los ceros a la izquierda pueden omitirse, y además si uno o más grupos tienen todos los dígitos a cero se pueden omitir poniendo en su lugar dobles dos puntos. Así por ejemplo la dirección anterior se escribiría:

8000::123:4567:89AB:CDEF

Para evitar ambigüedad la notación abreviada :: sólo puede utilizarse una vez en una dirección.

Por último, para facilitar la escritura de direcciones IPv4 se prevé también el uso de la notación decimal si se desea utilizando puntos, por ejemplo :

::147.156.11.11

Con 16 bytes es posible crear 2^{128} direcciones, equivalente a aproximadamente 3×10^{38}; aunque el reparto produce despilfarro es evidente que IPv6 no andará escaso de direcciones en mucho tiempo.

9.8.- ARP

Las direcciones IP que se han presentado en el capítulo anterior sólo tienen significado en el nivel de red dentro del modelo TCP/IP. En cualquier caso, la comunicación se desarrolla en ultimo termino por redes físicas haciendo uso de interfaces que efectúan los envíos y direcciones basándose en sus direcciones de enlace de datos, cuyo formato y significado variarán de una subredes a otras.

Esto supone que cada vez que IP transmite un datagrama debe efectuarse la asignación de una dirección de enlace de datos a la dirección IP de destino del datagrama, lo que se conoce como **resolución de direcciones IP**.

Todo nodo conectado a una red IP debe ser capaz de efectuar resolución de direcciones IP, lo que supone que deberia almacenar tablas con emparejamientos entre direcciones del nivel de enlace de datos y direcciones IP, implementando alguna función “hash” para la búsqueda de la equivalencia. Sin embargo, el mantenimiento de una tabla estática en redes grandes resultaría imposible de mantener (teniendo en cuenta que la asignación puede cambiar). Con el fi de permitir la la configuración automática de correspondencias entre direcciones MAC y dirección IP se diseñó el protocolo denominado ARP (Address Resolution Protocol) que opera en medios que soporte difusión.

Este protocolo está situado a caballo entre los niveles de Red y de Enlace de Datos, y dependiendo de los autores podemos encontrarlos ubicado en uno u otro. En realidad su función es la de convertir direcciones de red en direcciones de enlace de datos y por lo tanto sus funciones no corresponden ni a uno ni a otro sino que permite la interoperación entre ambos niveles. Sin embargo, para enviar sus mensajes hace uso de paquetes del nivel de enlace de datos por lo que en buena lógica debe ser ubicado por encima de éste.

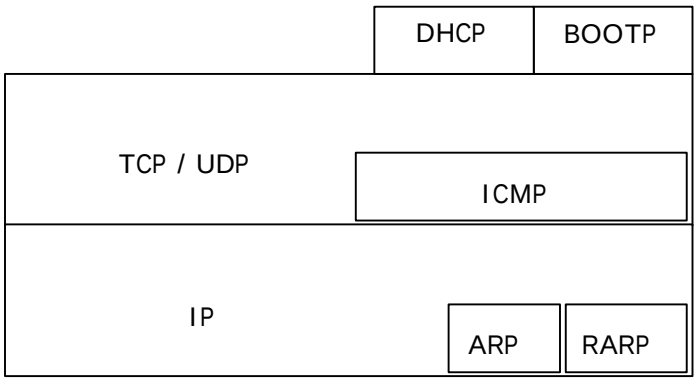

Figura 9.46

Veamos el modo de funcionamiento del protocolo:

Supongamos que desde un PC conectado a una red local ethernet, configurado con la dirección IP 160.230.7.7 y la máscara 255.255.0.0 se desea iniciar una sesión de terminal remoto con un servidor cuya dirección IP sabe que es 160.230.12.1. Por comparación de la dirección de destino con la propia y su máscara se deduce que el ordenador de destino se encuentra en su misma red local (o en alguna otra conectada mediante puentes o conmutadores LAN con la suya). El protocolo ARP genera una trama *broadcast ARP* con la pregunta ¿quién tiene la dirección 160.230.12.1?; la trama broadcast ARP se encapsula para su envío en una trama broadcast ethernet (dirección de destino toda a unos); al ser broadcast la trama es recibida y procesada por todas las máquinas de la red (e incluso retransmitida a través de los conmutadores o puentes MAC locales o remotos si los hubiera). Finalmente una máquina, y solo una, se reconoce como propietaria de la dirección IP solicitada, y ésta será la única que responda; la respuesta puede ser una trama unicast, puesto que el servidor ya conoce la dirección MAC del PC que pregunta puesto que iba escrita en la trama ethernet; la respuesta incluye la dirección MAC del servidor, por lo que a partir de ese momento el primero de los PCs puede comunicar con el segundo mediante tramas unicast, para no molestar a los demás ordenadores de la red.

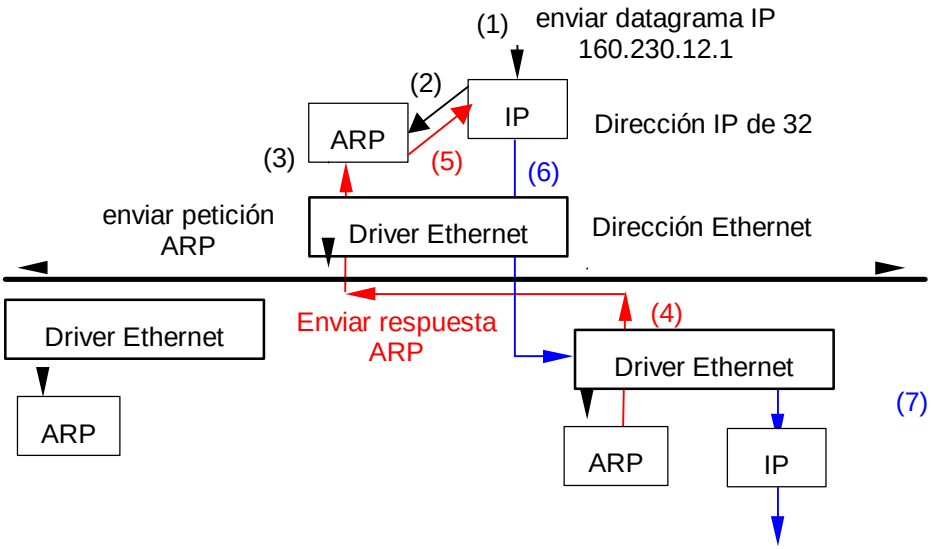

Figura 9.47

Como podemos ver, cada vez que se precise enviar un datagrama puede averiguarse la dirección MAC de destino preguntando por ella directamente al destinatario en lugar de tener que mantener una tabla con todas las equivalencias.

Las direcciones resueltas vía ARP se mantienen un cierto tiempo en cache, con el fin de no emplear repetidamente ARP preguntando por la misma dirección de enlace de datos. En realidad cuando se manda una trama ARP *todas* las máquinas de la red, no sólo la destinataria, aprovechan este

mensaje para 'fichar' al emisor, anotando en su cache la asociación dirección IP- dirección LAN. De esta forma si más tarde necesitan contactar con dicha máquina podrán hacerlo directamente, sin necesidad de enviar antes una trama ARP broadcast (esto último no es de implementación corriente en todos los sistemas operativos).

Las resoluciones obtenidas se mantienen en la caché ARP durante un tiempo de unos 20 minutos, pasados los cuales, si debe enviarse un nuevo datagrama a la misma dirección IP habrá que repetir el envío del mensaje ARP para obtener la dirección de enlace del destinatario.

9.8.1.- FORMATO DE PAQUETES ARP

La figura muestra el formato de las tramas de **petición y respuesta ARP**, cuando se emplan en una red Ethernet para resolver direcciones IP (si bien ARP es general y puede emplearse para resolver cualquier tipo de dirección lógica, no sólo IP, sobre cualquier tipo de red que soporte difusión).

<table>
<tr><td colspan="2">0 15</td><td>16 31</td></tr>
<tr><td colspan="2">Hardware (1)</td><td>Protocolo (0x0800)</td></tr>
<tr><td>Longitud hardware</td><td>Longitud protocolo</td><td>Operación</td></tr>
<tr><td colspan="3">Dirección ethernet emisor</td></tr>
<tr><td colspan="2">Dirección ethernet emisor</td><td>Dirección IP emisor</td></tr>
<tr><td colspan="2">Dirección IP emisor</td><td>Dirección ethernet destino</td></tr>
<tr><td colspan="3">Dirección ethernet destino</td></tr>
<tr><td colspan="3">Dirección ip destino</td></tr>
</table>

Figura 9.48

Los campos **Protocolo** y **Hardware** sirven para especificar el tipo de direccionamiento de nivel de red y de enlace de datos respectivamente. En el caso de direcciones IP, el campo Protocolo tiene el valor 0x800 y en el caso de Ethernet el campo Hardware tiene el valor 1.

Los campos **Longitud Hardware** y **Longitud de Protocolo** especifican el tamaño en octetos de las direcciones de enlace de datos y de red respectivamente. En Ethernet serán 6 y 4 respectivamente.

El campo **Operación** especifica si la trama es una petición (1) o una respuesta (2).

Los siguientes cuatro campos son las direcciones de enlace de datos y de red del emisor y del destinatario respectivamente. En las respuestas los cuatro campos estarán rellenos, mientras que en la pregunta el campo **Dirección Ethernet de Destino** estará vacío.

9.8.2.- PROXY ARP

El Proxy ARP permite a un “router” contestar peticiones ARP efectuadas en alguna de sus redes y dirigidas a un nodo situado en otra de las redes a la cual pertenece. El objetivo de Proxy ARP era solucionar en parte el uso ineficiente del espacio de direcciones IP, antes de la aparición de CIDR. Por ejemplo, una organización que deseara conectar 2048 nodos a una red Ethernet necesitaría una dirección de clase B (desaprovechando 63486 direcciones) para poder asignar la dirección IP a cada nodo, sin embargo, y debido a las características del cableado Ethernet, necesitará dividir la red en dos partes separadas al no soportar más de 1024 nodos; la forma más natural de hacer esto es colocar un “bridge” o un “router” entre ellas. En el caso de utilizar un “router” necesitaría dos direcciones de clase B y no una sola, desaprovechando todavía una cantidad mayor de direcciones, Proxy ARP permite hacer uso de la misma dirección de clase B en ambos lados del “router”, reduciendo la necesidad de direcciones distintas. Se utilizan algunos bits para identificar las distintas subredes, si bien, sólo el “router” que opera con Proxy ARP debe interpretarlos.

En una red que hace uso de Proxy ARP, los nodos emisores utilizan ARP para obtener la dirección MAC del nodo de destino, tanto si están en su mismo segmento de la red como si no lo están (dado que ellos no perciben la división de su red). El “router” proxy contestará a la petición incluyendo una de sus direcciones MAC. El emisor de la petición, cuando reciba la respuesta ARP interpretará que el “router” es el nodo destino, cuando en realidad éste se encuentra “al otro lado” del “router”; esto es posible porque los nodos no efectúan ninguna comprobación sobre la validez de la dirección MAC contenida en una respuesta ARP. El emisor encapsulará los datagramas IP en datagramas Ethernet dirigiéndolos a una de las direcciones MAC del “router”, y éste a su vez retransmitirá el datagrama que recibe hacia la red correcta sustituyendo su dirección MAC por la del destinatario final.

El resultado final es que el “router” actúa como apoderado o representante (proxy) del nodo destino, y de ahí su nombre. También se conoce como "ARP promiscuo" o "ARP hack", nombres que vienen del efecto de ocultar a los nodos la existencia de dos redes físicas separadas, proporcionando la apariencia de una única red.

9.8.3.- ARP gratuito

Se produce cuando un nodo envía una petición ARP buscando su propia dirección de enlace. Generalmente se emplea al configurar el interfaz de red durante el arranque.

El ARP gratuito proporciona dos funciones, por un lado permite a un nodo determinar si otro nodo está ya configurado con la misma dirección IP y por otro cuando un nodo cambia su dirección de enlace de datos, el envío de un paquete ARP gratuito hace que cualquier nodo que tenga una entrada en su caché para la vieja dirección de enlace actualice su entrada de la caché ARP.

9.9.- RARP

A veces se plantea el problema inverso a ARP, es decir hallar la dirección IP de una determinada dirección LAN. Por ejemplo, cuando se arranca una estación de trabajo 'diskless', es decir, que tiene su disco de arranque en otra estación, ésta desconoce todo lo relativo a su configuración de red (incluida la dirección IP) excepto la dirección MAC que está registrada en su tarjeta de red local.

Para esto se diseño RARP (Reverse Address Resolution Protocol), que consiste en que la estación emita una trama broadcast indicando su dirección LAN y solicitando alguien le informe de cual es la dirección IP que le corresponde. En este caso una máquina en la red local (el servidor RARP) atenderá la petición, consultará en sus tablas, y devolverá la dirección IP correspondiente.

El formato de los mensajes RARP es el mismo que el de los mensajes ARP.

Generalmente, cada estación tiene asignado un servidor RARP primario que responde a sus peticiones. Si no se recibe contestación a la petición se volverá a enviar la misma y entonces contestará a la petición alguno de los servidores configurados como secundarios.

Como la estación solicitante emite una trama broadcast el servidor RARP puede estar en la misma LAN o en otras que estén conectadas con esta por puentes o conmutadores LAN, pero no puede haber routers entre ella y el servidor RARP (los routers filtran los paquetes broadcast de las LANs). Por otro lado, el protocolo RARP solo permite al servidor enviar la dirección IP del cliente en el paquete, cuando sería interesante aprovechar para enviar en el mismo datagrama una serie de parámetros de configuración a la estación que arranca.

9.10.- ICMP (Internet Control Message Protocol)

Como hemos visto anteriormente hay situaciones en las que un router debe enviar un mensaje de error al emisor de un datagrama ante una situación de error o algún otro evento extraordinario al emisor. El protocolo ICMP, compuesto por un conjunto de mensajes de control es el mecanismo de que disponen los routers para comunicar dichas situaciones.

ICMP se considera a menudo como parte del nivel de red, porque su función es la de comunicar mensajes de error relacionados con el nivel de red así como otras condiciones de atención. ICMP, sin embargo, no hace a IP fiable, esto será responsabilidad de los niveles superiores. La especificación de este protocolo se encuentra en RFC 792.

Los mensajes ICMP pueden ser empleados bien por el nivel IP, por el nivel de transporte (TCP o UDP) o incluso por el nivel de aplicación.

Los mensajes ICMP se trasmiten encapsulados dentro de datagramas IP, como se muestra en el figura 9.49, y se envían haciendo uso del encaminamiento IP, sin embargo, ICMP no se considera un protocolo de alto nivel sino una parte de IP. La razón de utilizar IP para enviar mensajes ICMP es que pueden necesitar atravesar varias redes físicas, y es el protocolo IP quien asegura el tránsito de los datagramas a través de Internet.

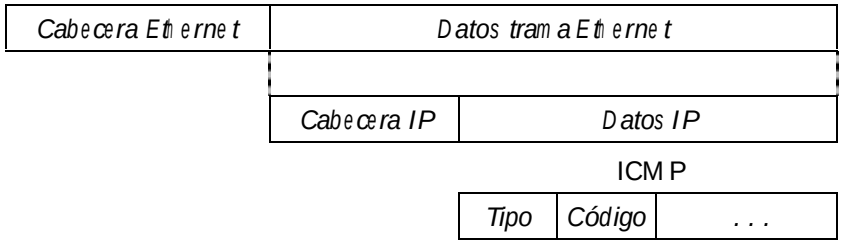

Figura 9.49

El formato básico de los mensajes ICMP aparece en la figura 9.50, si bien, salvo los primeros cuatro octetos, el formato del resto del mensaje difiere de uno a otro según el contexto.

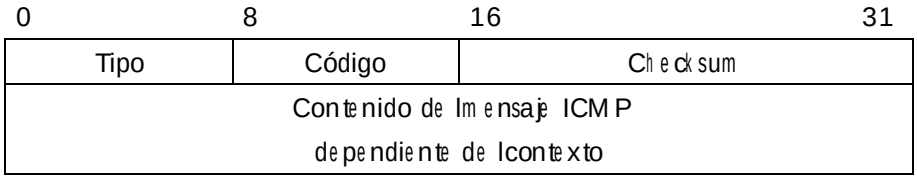

Figura 9.50

Los campos **Tipo** y **Código** sirven para identificar los distintos mensajes ICMP utilizados. Más adelante, cuando presentemos los distintos mensajes ICMP, analizaremos los valores que pueden tener estos campos.

El campo **Cheksum** sirve para comprobar el mensaje completo ICMP, el algoritmo que se utiliza es el mismo que se describió en la cabecera IP, y es un campo requerido.

9.10.1.- TIPOS DE MENSAJES ICMP

Los distintos tipos de mensajes ICMP en uso están recogidos en la tabla 9-21, donde además se indican los valores de los campos tipo y código que emplean.

Las dos últimas columnas de esta tabla especifican si los mensajes ICMP corresponden a **Peticiones** o a **Respuestas/errores**. Es necesario hacer esta distinción porque los mensajes de error ICMP tienen un tratamiento especial, ya que no se generan nunca como respuesta a:

- Otro mensaje de error ICMP; sí como respuesta a una petición ICMP.
- Un datagrama destinado a una dirección de difusión IP o a una dirección multicast IP.
- Un datagrama enviado como difusión del nivel de enlace.
- Un fragmento de un datagrama distinto del primero.
- Un datagrama cuya dirección fuente no define a un simple nodo (como una dirección de origen a ceros, a unos, una dirección de loopback, de difusión o multicast).

Cuando se envíe un mensaje de error ICMP, el mensaje siempre contiene la cabecera IP y los primeros ocho octetos del datagrama IP que originaron el mensaje de error. Ésto permite al módulo ICMP receptor asociar el mensaje con un determinado protocolo (TCP/UDP) a partir del campo protocolo de la cabecera IP, y un proceso de usuario particular (a partir de los números de puerto).

tipo	cod	descripción	petición	error
0	0	Respuesta de echo	x	
3		**Destino no alcanzable**		x
	0	Red no alcanzable		x
	1	Nodo no alcanzable		x
	2	Protocolo no alcanzable		x
	3	Puerto no alcanzable		x
	4	Fragmentación necesaria, bit no fragm. Activado		x
	5	Fallo en ruta desde origen		x
	6	Red de destino desconocida		x
	7	Nodo de destino desconocido		x
	9	Red destino administrativamente prohibida		x
	10	Nodo destino administrativamente prohibido		x
	11	Red no alcanzable por TOS		x
	12	Nodo no nalcanzable por TOs		x
	13	Comunicación administrativamente pohibida por filtrado		x
	14	Violación de la precedencia del nodo		x
	15	Limitación de precedencia en efecto		x
4	0	Source quench		x
5		**Redirección**		x
	0	Redirección para una red		x
	1	Redirección para un nodo		x
	2	Redirección para un TOS y red		x
	3	Redirección para un TOS y nodo		x
8	0	Petición de eco	x	
9	0	Notificación de pasarela	x	
10	0	Petición de pasarela	x	
11		**Tiempo excedido para un datagrama**		
	0	TTL igual a 0 durante el tránsito		x
	1	TTL igual a 0 durante el reensamblado		x
12		**Problema parametrizable**		
	0	Cabecera IP errónea		x
	1	Falta de una opción necesaria		x
13	0	Petición de timestamp	x	
14	0	Respuesta de timestamp	x	
17	0	Petición de máscara de subdirección	x	
18	0	Respuesta de máscara de subdirección	x	

Tabla 9.21

Petición de eco y respuesta (echo request y reply - tipos 0 y 8)

ICMP permite comprobar si un determinado destino es alcanzable o no mediante el envío de mensajes de **petición de eco (tipo 8) y respuesta de echo** (tipo 0).

Cualquier nodo o “router” puede enviar mensajes ICMP de **petición de eco** dirigidos a cualquier destino. El nodo que recibe esta petición devuelve un mensaje de **respuesta de eco** al emisor original de la petición.

La **respuesta de eco** contiene un area de datos opcional y una copia de los datos enviados en la petición con el fin de que el emisor pueda comprobar que los datos han circulado correctamente en su viaje de ida y vuelta.

La **petición y respuesta de eco** permiten comprobar si es posible alcanzar un cierto destino y si éste responde. Ambos mensaje viajan en datagramas IP, y por lo tanto, verifican la mayor parte del sistema de transporte:

- El software IP del nodo origen encamina correctamente el datagrama hacia su destino.
- Los routers intermedios operan con normalidad y son capaces de encaminar el datagrama correctamente hasta su destino final.
- El nodo destinatario está conectado y tanto el software IP como ICMP operan correctamente.
- La rutas en los routers en el camino de retorno son correctas.

El formato de estas tramas ICMP es el indicado en la figura 9.51 El formato de la petición y de la respuesta coinciden, si bien el número de código será respectivamente 8 ó 0.

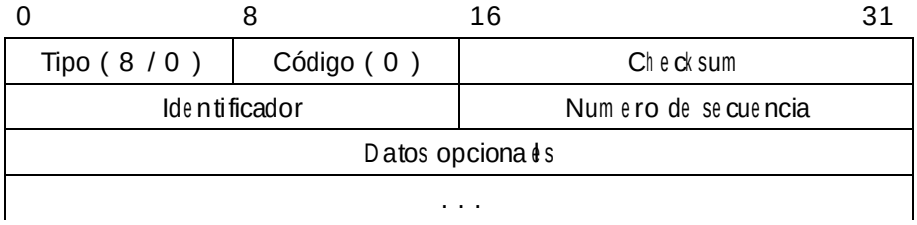

Figura 9-51

El campo **identificador** identifica al proceso emisor que ha generado la petición. Hay que tener en cuenta que dado que los sistemas son multiusuario y multiproceso es probable que varios programas lancen peticiones de eco simultaneamente y haya que identificar a qué proceso debe entregarse cada una de las respuestas recibidas por el nodo.

El **número de secuencia** permite relacionar cada respuesta con su petición correspondiente. Es un número secuencial que comienza en 0 y es incrementado cada vez que se envía una nueva petición.

El cliente puede incluir un campo de **datos opcionales** si lo desa en la petición de eco. El servidor repite el campo **identificador,** el campo **número de secuencia** y los **datos opcionales** enviados por el cliente en la respuesta, pudiendo incluir nuevos datos si así lo desea.

En la mayoría de los sistemas operativos, los usuarios pueden enviar mensajes de petición de eco haciendo uso del comando "**ping**". Este comando envía una serie de peticiones de eco, se encarga de capturar las respuestas correspondientes a dichas peticiones y finalmente proporciona estadísticas de los datos enviados y recibidos, de los datos perdidos y del tiempo empleado por los datagramas en viajar entre ambos extremos. Ping imprime el número de secuencia de cada paquete devuelto, permitiéndonos ver si los paquetes son erróneos, cambiados de orden o duplicados.

$ ping locis.loc.gov
PING locis.loc.gov: 64 byte packets
64 bytes from 140.147.254.3: icmp_seq=0. time=2954. ms
64 bytes from 140.147.254.3: icmp_seq=1. time=2520. ms
64 bytes from 140.147.254.3: icmp_seq=3. time=2258. ms
64 bytes from 140.147.254.3: icmp_seq=2. time=3561. ms
64 bytes from 140.147.254.3: icmp_seq=4. time=1680. ms
64 bytes from 140.147.254.3: icmp_seq=5. time=1210. ms
----locis.loc.gov PING Statistics----
8 packets transmitted, 6 packets received, 25% packet loss
round-trip (ms) min/avg/max = 1210/2363/3561

Mediante este comando, un usuario puede comprobar la accesibilidad de un nodo, y haciendo uso de las opciones de registro de ruta y de registro de "timestamp" analizar la trayectoria seguida por los paquetes hacia su destino.

El programa **ping** que envía la petición se denomina **cliente** y el nodo que contesta **servidor**. Las implementaciones TCP/IP incluyen el servidor ping directamente en el núcleo.

Destino no alcanzable (destination unreachable - tipo 3)

Si bien IP propociona un servicio no fiable, los datagramas no se descartan alegremente. Siempre que un error impide a un “router” encaminar o enviar un datagrama, el “router” envía un mensaje ICMP de **destino no alcanzable** al emisor de dicho datagrama y sólo entonces elimina el datagrama. Si bien hay definidos mensajes correspondientes a diferentes situaciones por las cuales no es posible reencaminar un datagrama, no todos los errores de encaminamiento son notificados por parte de los “routers”.

Los mensajes de error ICMP deben incluir la cabecera IP (incluyendo todas las opciones) del datagrama que generó el error junto con los primeros ocho octetos que siguen a la cabecera IP. Una de las razones para que la cabecera IP del datagrama que originó el error sea devuelta es porque contiene el campo de protocolo que permite a ICMP saber cómo interpretar los ocho octetos que siguen (identificar si corresponden a UDP o a TCP). Cuando analicemos las cabeceras TCP y UDP veremos que los números de puertos de origen y de destino están contenidos dentro de los primeros ocho octetos de la cabecera. El formato general de los mensajes ICMP de **destino no alcanzable** es mostrado en la figura.

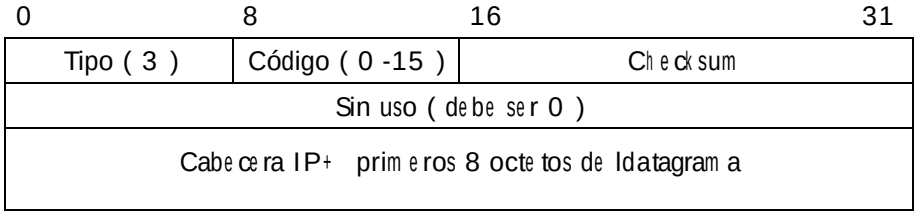

Figura 9.52

Como puede apreciarse hay varios mensajes de **destino no alcanzable**, con valores del campo **código** que van desde 0 a 15, relacionados con el error que propició la eliminación del datagrama.

Red no alcanzable	El router no encuentra una ruta hacia la red especificada en la dirección IP de destino. Estos mensajes sólo son pueden ser generados por un “router”, e indican que es necesaria una comprobación de la dirección IP de destino o de las tablas de encaminamiento de los “routers” que han intervenido. Dado que el mensaje contiene la parte inicial del datagrama que originó el problema, el emisor puede saber exactamente qué dirección es no alcanzable, y puede averiguar el punto de la red en que se produjo el error extrayendo la dirección IP de origen del datagrama IP que contenía el error ICMP.
Nodo no alcanzable	El datagrama descartado ha alcanzado un “router” que está conectado directamente a la red de destino, pero éste no puede comunicarse con el nodo de destino. El nodo puede estar desconectado o simplemente la dirección IP no existe.
Protocolo no alcanzable	En este caso el datagrama que originó el error ha alcanzado el nodo de destino, pero el protocolo identificado en la cabecera IP no está disponible en él.
Puerto no alcanzable	El servicio del nivel de aplicación al cual va dirigido el datagrama no está disponible. Este tipo de mensajes es generado en el nodo de destino por uno de los protocolos del nivel de transporte (UDP o TCP).

Fragmentación necesaria y bit de no fragmentación activado.	Se genera cuando un “router” necesita fragmentar un datagrama para retransmitirlo y el **bit de no fragmentación** está activado. Este error es poco usual, dado que el bit de no fragmentación sólo suele ser empleado por estaciones sin disco que arrancan haciendo uso de TFTP. También pueden emplearse estos mensajes de error para determinar la MTU más pequeña en el camino hacia un destino; este **mecanismo de descubrimiento de MTU** permite al “router” que envía el mensaje de error incluir la MTU del interfaz de salida en los octetos 7 y 8, tal y como se indica en la figura 9.10.5
Fallo en el encaminamiento desde el origen.	Se produce cuando no puede encaminarse un datagrama IP en cuya cabecera figura la ruta que debe seguir. Por alguna razón, el “router” que genera el mensaje de error no es capaz de hacer llegar el datagrama al siguiente “router” del campo opciones.
Red de destino desconocido.	Lo genera un “router” cuando puede determinar a partir de su software de nivel de enlace de datos que la red de destino no existe.
Nodo de destino desconocido.	Lo genera un “router” cuando determina a partir de su software de enlace de datos que el nodo de destino no existe.
Red no alcanzable por el Tipo de Servicio especificado.	Es generado por un “router” cuando no está disponible una ruta hacia la red de destino especificada con el Tipo de Servicio solicitado o su valor por defecto.
Nodo no alcanzable por el Tipo de Servicio especificado	Lo genera un “router” cuando la ruta hacia el destino no se ajusta al Tipo de Servicio solicitado o a su valor por defecto.
Comunicación administrativamente prohibida.	Generado cuando un “router” no puede encaminar un datagrama debido a reglas de filtrado de paquetes, por ejemplo por razones de seguridad como un firewall.
Violación de la precedencia de un nodo.	Es enviado por el primer “router” en el camino a un nodo para indicar que la precedencia solicitada no está permitida por la combinación de nodo o red de origen/destino, el protocolo de nivel superior y el puerto de origen/destino.
Limitación de la precedencia en efecto.	Indica que el gestor de red ha configurado un nivel mínimo de precedencia necesario para utilizar esta ruta; el datagrama se enviará con un nivel de precedencia menor que el necesario.

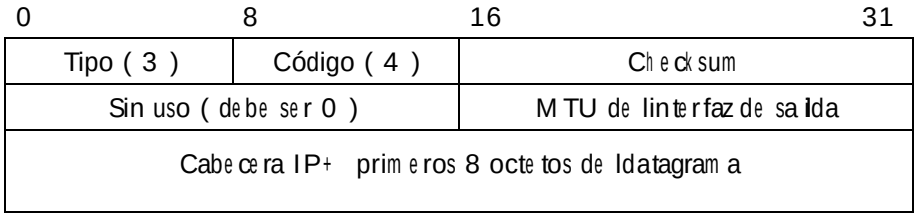

Figura 9.53

Control de congestion y de flujo (source quench - tipo 4)

Dado que IP es un protocolo no orientado a la conexión, los “routers” no pueden hacer una reserva de memoria con el fin de poder almacenar los datagramas recibidos desde un cierto origen. Como consecuencia, pueden saturarse con el tráfico, una condición conocida como **congestión**. Es importante entender que la congestión puede producirse por dos motivos completamente diferentes. El primero, que una estación de alta velocidad genere tráfico más deprisa de lo que una red pueda transferirlo; incluso la estación estuviera conectada a una red local de alta velocidad, los datagramas pueden tener que atravesar una red de baja velocidad dentro de Internet. La congestión se producirá en el “router” conectado a esta red de baja velocidad porque los datagramas llegarán mucho más deprisa de lo que pueden ser reenviados. En segundo lugar, si muchos ordenadores necesitan simultáneamente enviar datagramas a través de un mismo camino, el “router” que da acceso a dicha ruta puede experimentar una congestión, aunque no sea un solo emisor quien esté causando el problema.

Cuando los datagramas llegan demasiado deprisa para que un “router” los procese, éste los encola temporalmente en memoria. Si los datagramas son parte de una ráfaga pequeña, este almacenamiento resolverá el problema, si por el contrario el tráfico es continuo, el “router” agotará su memoria y deberá descartar los datagramas que lleguen. Estos sistemas utilizan los mensajes ICMP de **retención de la emisión** (“**source quench”**) para impedir la congestión, ya que son una petición para que el emisor reduzca la velocidad a la que genera datagramas. Generalmente los “routers” congestionados envían un mensaje de retención por cada datagrama que descartan. Los “routers” pueden utilizar también técnicas de control de congestión más sofisticadas; algunos monitorizan el tráfico entrante y retienen a los emisores que tienen las velocidades de transmisión de datagramas más altas. Otros intentan evitar la congestión enviando peticiones de retención a medida que sus colas de almacenamiento empiezan a crecer demasiado antes de que se sobrepasen.

No hay ningún mensaje ICMP para invertir el efecto de un mensaje de **retención de la emisión**, en su defecto, se reduce la velocidad a la que envían datagramas al “router” congestionado hasta que se dejan de recibir mensajes de retención; a partir de entonces, y gradualmente, se incrementa la velocidad mientras no reciba nuevas peticiones de retención.

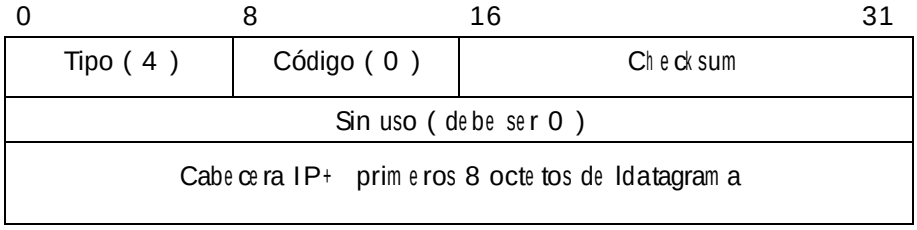

Figura 9.54

Esta técnica ha demostrado ser muy poco efectiva, ya que no mejora el rendimiento de las redes muy ocupadas y simplemente añade nueva carga al procesador y a la red congestionada. Por ello, los últimos RFC indican que los “routers” no deben generar mensajes “source quench” aún cuando descarten datagramas y los nodos deben aceptar los mensajes si bien no es necesario que tomen ninguna acción correctiva.

Peticiones de cambio de ruta (redirect - tipo 5)

Las tablas de encaminamiento de los nodos en Internet permanecen generalmente estáticas durante largos periodos de tiempo, inicializadas en el arranque a partir del fichero de configuración. Sin embargo, la topología de la red puede cambiar, bien temporal o permanentemente, de forma que las tablas de encaminamiento de los nodos y “routers” resulten incorrectas. Los “routers” intercambian información de encaminamiento periódicamente para acomodarse los cambios en la red y mantener sus rutas actualizadas, empleando protocolos de encaminamiento, pero no así los nodos. Como norma general se supone que los “routers” conocen rutas correctas, y que los nodos comienzan con una información de encaminamiento mínima y aprenden nuevas rutas a partir de la información contenida en las tablas de encaminamiento de los “routers”.

Cuando un “router” detecta que un nodo está utilizando una ruta no óptima (ha enviado un datagrama a un “router” que no es el más adecuado en la ruta hacia el destino final), le envía un mensaje de ICMP de **petición de cambio de ruta**, para que modifique en su tabla de encaminamiento la ruta hacia dicho destino. El “router”, por supuesto, reenvía el datagrama a su destino con el fin de que éste no se pierda.

La ventaja del esquema de redirección ICMP es su simplicidad: permite a un nodo arrancar conociendo la dirección de un solo “router” en su red local como ruta por defecto. Este “router” inicial devuelve mensajes ICMP de **peticion de cambio de ruta** cuando el nodo le envía datagramas para los cuales existe una ruta mejor. La tabla de encaminamiento del nodo permanece pequeña si bien contiene rutas óptimas para todos los destinos en uso.

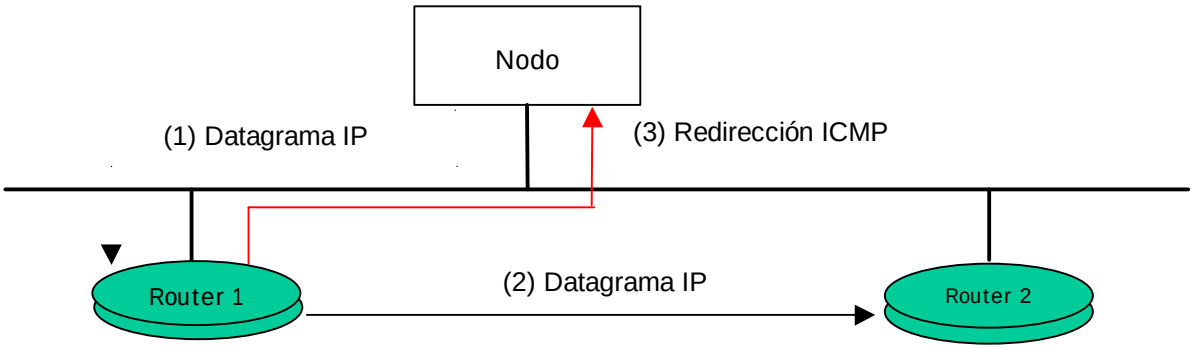

Figura 9-55

Los mensajes de petición de cambio de ruta no solucionan el problema de propagar rutas de modo general dado que están limitados a interacciones entre “routers” y nodos conectados directamente a la misma red.

El formato de los mensajes de **redirección** ICMP es el siguiente:

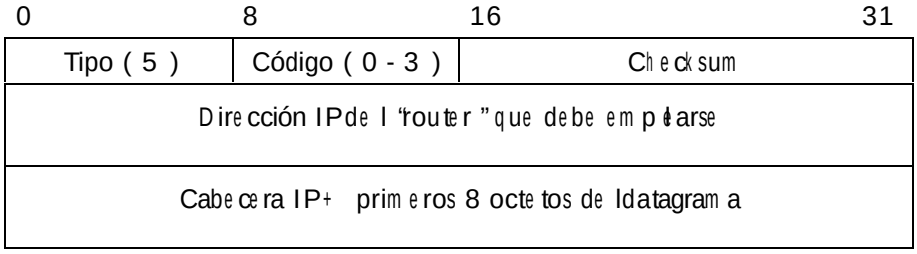

Figura 9-56

El campo D**irección IP** contiene la dirección IP del “router” que el nodo debería utilizar para alcanzar el destino que aparece en la cabecera del datagrama. Así, un nodo que recibe un mensaje de redirección ICMP examina los primeros octetos del datagrama para determinar cual era su dirección de destino y efectúa la actualización de su tabla de encaminamiento. El campo **Código** especifica además cómo interpretar la dirección de destino, basándose en los siguientes valores :

Código	Significado
0	Redireccionar datagramas hacia una red de destino
1	Redireccionar datagramas hacia un nodo de destino
2	Redireccionar datagramas hacia una red con un TOS
3	Redireccionar datagramas hacia un nodo con un TOS

Tabla 9-22

Como norma general, los “routers” sólo envían peticiones de redirección ICMP a nodos y no a otros “routers”, puesto que estos utilizan sus propios protocolos para actualizar la información de encaminamiento.

Hay tres direcciones IP involucradas en el proceso de redirección ICMP:

- La dirección IP que causó el cambio de dirección (contenida en la cabecera IP incluido en el campo de datos).
- La dirección IP del “router” que envió el cambio de dirección (contenida en la cabecera IP del datagrama que contiene el mensaje de redirección ICMP).
- La dirección IP del “router” más correcto para alcanzar el destino (contenida en el campo correspondiente del mensaje ICMP).

Existen una serie de comprobaciones que debe efectuar el “router” emisor de un mensaje ICMP de petición de cambio de ruta antes de enviarlo:

- El interfaz de salida del datagrama reenviado debe ser igual al de entrada
- La ruta en uso en la propia lista de encaminamiento para el datagrama saliente no debe haber sido creada o modificada por un mensaje de redirección ICMP, y no debe ser la ruta por defecto del “router”.
- El datagrama no debe emplear la opción de encaminamiento desde el origen.
- El propio “router” debe estar configurado para enviar redirecciones.

Igualmente, el nodo receptor de un mensaje de redirección ICMP debe llevar a cabo una serie de comprobaciones antes de modificar su tabla de encaminamiento.

- El nuevo “router” debe estar en una red directamente conectada a él.
- El mensaje de redirección debe proceder del nodo actual en la ruta hacia dicho destino.
- La redirección no debe indicar al propio nodo como “router”.
- La ruta que esta siendo modificada debe ser una ruta indirecta.

Detección de rutas circulares (time exceeded - tipo 11)

Puesto que los “routers” Internet calculan el siguiente salto utilizando tablas locales, pueden producirse errores en las tablas de encaminamiento que generen rutas circulares para algún destino. Si un datagrama entrase en una ruta circular permanecería en la red indefinidamente. Para evitar ésto, cada cabecera IP contiene el campo **tiempo de vida** (TTL)**,** que los “routers” decrementan siempre que procesan un datagrama, descartando todos aquellos que reciben con el valor de dicho campo a cero.

Siempre que un “router” descarta un datagrama porque el campo **tiempo de vida** ha alcanzado el valor CERP envía un mensaje ICMP de **tiempo excedido** utilizando el formato indicado en la figura.

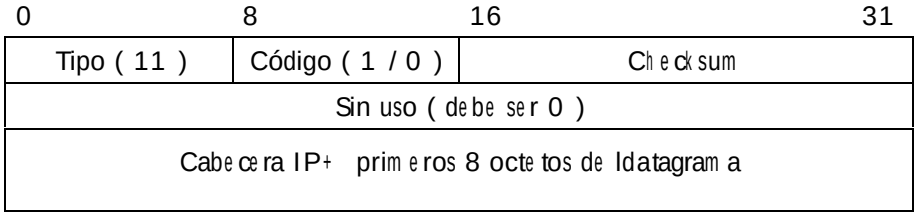

Figura 9.57

El campo **Código** permite diferenciar la situación anterior (Código = 0) de una segunda condición en la que se genera un mensaje ICMP de **tiempo excedido** por parte del nodo de destino. Cuando el nodo de destino recibe el primer fragmento de un datagrama arranca un temporizador, si éste expira antes de recibir todos los fragmentos que lo componen, descartará todos los recibidos y enviará un mensaje ICMP de este tipo con el campo Código con valor 1.

Notificación de rutas (tipo 9)

Los mensajes ICMP de notificación de rutas permiten a los “router” presentarse a los nodos de una red. El formato de los mensajes es el siguiente:

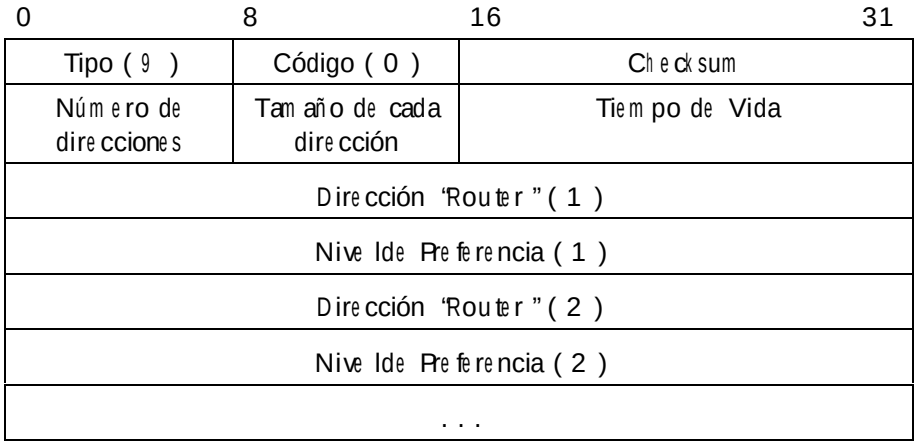

Figura 9.58

El campo **número de direcciones** indica cuántas direcciones de "routers" se notifican en el mensaje. El campo **tamaño de cada dirección** es el número de palabras de 32 bits empleada para cada dirección, y es siempre 2. El campo **Tiempo de Vida** especifica el número de segundos que las direcciones notificadas en un mensaje deben ser consideradas válidas (por defecto 1800 segundos).

Por cada dirección IP notificada se incluyen dos campos, la dirección IP, que corresponde a la dirección IP del "router" en el interfaz desde el que se envía el mensaje, y el **nivel de preferencia, que** indica la preferencia de esta dirección como una dirección por defecto con respecto a otras direcciones de la misma subred; los valores mayores indican direcciones con mayor preferencia.

Estas notificaciones se envían a intervalos regulares entre 7 y 10 minutos, o bien como respuesta a un mensaje ICMP de solicitud de rutas. Además, si en un "router" se deshabilita uno de sus interfaces se transmite una **notificación de ruta** con el campo **tiempo de vida** puesto a 0, indicando que esa ruta ya no será válida en adelante. Generalmente estos mensajes se envían a la dirección multicast 224.0.0.1 (todos los sistemas de esta subred).

Estos mensajes no contienen información acerca de las rutas alcanzables por este "router", ya que una vez escogido el "router" por defecto, los nodos irán aprendiendo nuevas rutas en base a mensajes ICMP de petición de cambio de ruta.

Petición de router (tipo 10)

El mensaje ICMP de **petición de router** es difundido por cualquier nodo que desea averiguar qué "routers" están disponibles en su red local. Pueden enviarlo en cualquier momento, pero lo más habitual es hacerlo en el momento del arranque. El formato del mensaje es el siguiente.

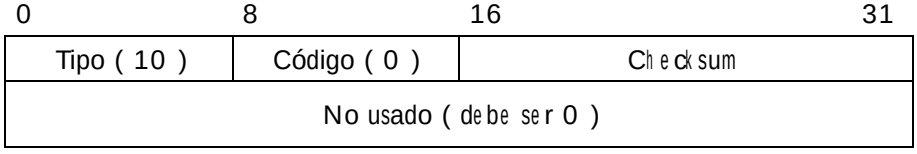

Figura 9.59

Los mensajes se envían generalmente a la dirección multicast 224.0.0.2 (todos los "routers" de esta subred). Todos los "router" que reciben este mensaje deben contestar mediante un mensaje ICMP de **notificación de ruta.** Además, los nodos suelen recibir estos mensajes a intervalos de 7 a 10 minutos. A partir de estas notificaciones, o bien cuando expira el temporizador de tiempo de vida de la dirección, se actualiza la ruta por defecto de la tabla de encaminamiento

Problema parametrizable (parameter problem - tipo 12)

Cuando un "router" o un nodo encuentran problemas en un datagrama no contemplados por los mensajes de error ICMP anteriores (p.e. una opción de la cabecera IP incorrecta) envían un mensaje de problema **parametrizable** al emisor. El mensaje, cuyo formato se indica en la figura 9.60, sólo se envía cuando el problema es tan severo que el datagrama debe ser descartado.

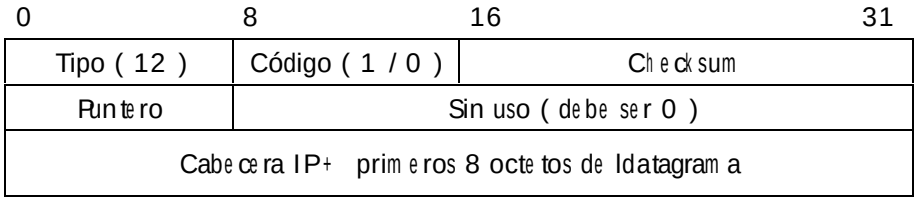

Figura 9.60

Para asegurar que el mensaje no sea ambiguo, el emisor utiliza el campo **Puntero** de la cabecera del mensaje para identificar el octeto del datagrama que originó el error (1 señalaría que el error corresponde al campo TOS y 20 que corresponde al primer octeto del campo opciones).

El **código** 1 se utiliza para informar de la falta de una opción necesaria en el datagrama, y en este caso el campo Puntero no se utiliza.

Sincronización del reloj (timestamp - tipos 13 y 14)

Los nodos que operan en Internet, si bien pueden comunicarse, generalmente operan de forma independiente, manteniendo cada uno de ellos su propia noción del tiempo local. Los relojes pueden diferir de forma importante y confundir a los usuarios de sistemas distribuidos. Existen varios protocolos dentro de TCP y IP que pueden utilizarse para sincronizar relojes. Una de las técnicas más sencillas es la utilización de mensajes ICMP para obtener la hora de otro sistema.

Cuando un sistema envía un mensaje de **petición de "timestamp"** a otro, solicita a este segundo sistema que devuelva el valor actual de la hora de su sistema. El sistema receptor devuelve un mensaje de **contestación de "timestamp"**. El formato de estos mensajes es el indicado en la figura 9.61.

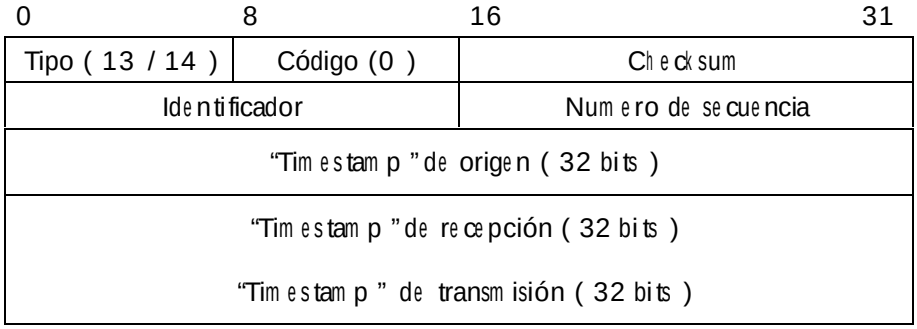

Figura 9.61

El valor recomendado para la respuesta es el número de milisegundos a partir de medianoche, del tiempo universal coordinado (UTC). Un aspecto curioso de los mensajes ICMP es que proporciona una resolución de milisegundos mientras que otros métodos para obtener la hora característicos de los sistemas UNIX simplemente proporcionan una resolución de segundos. Por el contrario, uno de los inconvenientes de estos mensajes es que sólo se proporciona la hora a partir de medianoche, y por lo tanto el nodo que pregunta debe conocer el día por algún otro medio.

El campo **tipo** identifica el mensaje como una petición (13) o una respuesta (14); los campos **identificador** y **número de secuencia se interpretan igual que** en el resto de los mensajes ICMP. Los campos restantes especifican distintos registros de horas del sistema emisor y receptor. El

campo **"timestamp" de origen** lo completa el emisor antes de transmitir el mensaje; el receptor rellena el campo **"timestamp" de recepción** inmediatamente después de recibir la petición, y el campo **"timestamp" de transmisión** cuando envía la respuesta

Los nodos utilizan estos tres campos para hacer estimaciones del retraso entre ellos y para sincronizar sus relojes. Dado que la respuesta incluye el campo **"timestamp" de origen**, el sistema puede calcular el tiempo que necesita la petición para viajar hasta su destino ser transformado en una respuesta y volver; el sistema puede calcular el tiempo de transito en la red y a partir de ello estimar las diferencias entre los relojes local y remoto.

Debemos tener en cuenta que el tiempo que necesita la respuesta para viajar hasta el emisor puede variar sustancialmente del que necesitó la pregunta, y por lo tanto las estimaciones pueden ser muy imprecisas.

Obtencion de la mascara de subred (tipos 17 y 18)

Cuando los nodos utilizan Direccionamiento de subred, algunos de los bits de la parte de nodo de su dirección IP identifican una subred. Para participar en un esquema de direccionamiento de subred los nodos necesitan saber qué bits de los 32 que componen la dirección corresponden a identificación de subredes y cuáles corresponden a la identificación de nodo. La información necesaria para interpretar la dirección está representada mediante una dirección de 32 bits denominada mascara de su red.

La **petición de máscara de subred** ICMP es empleada por los sistemas sin disco para conseguir su máscara de subred en el momento del arranque a partir de otro nodo, que le enviará la respuesta correspondiente. Estos sistemas difunden su petición ICMP de un modo similar al uso de RARP para conseguir la dirección IP. Existe un método alternativo para estos sistemas sin disco que es el uso del protocolo BOOTP, que veremos más adelante. La figura muestra el formato de estos mensajes.

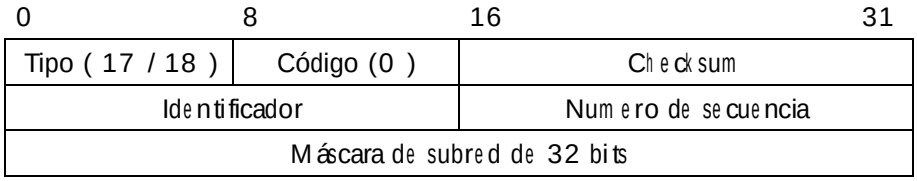

Figura 9.62

El campo **Tipo** especifica si un mensaje es una respuesta (18) o una petición (17). La respuesta contendrá la mascara de direccionamiento de subred en el campo correspondiente. Como en los casos anteriores, los campos **identificador** y **número de secuencia** del mensaje ICMP pueden tomar cualquier valor elegido por el emisor, y estos mismos valores se devuelven en la respuesta, permitiendo que el emisor relacione peticiones con respuestas.

El campo **máscara de subred** no existe en la petición y en la respuesta contiene la máscara solicitada.

9.11.- DNS: SISTEMA DE DOMINIO DE NOMBRES

En una red TCP/IP, los nodos pueden identificarse bien mediante una dirección IP o mediante un nombre único (por conveniencia del usuario).

Las direcciones IP tienen longitud fija, 32 bits y por su estructura su tratamiento resulta sencillo por parte de los ordenadores. Además, están íntimamente ligadas al encaminamiento de los datagramas, incluyendo la información que precisan los routers para hacer llegar los datagramas a su destino final. Por el contrario, los nombres tienen una longitud variable, un carácter mnemónico y no contienen información acerca de la ubicación de los nodos, ni por lo tanto para la búsqueda de la ruta que debe seguir un datagrama.

9.11.1.- Espacio de nombres y resolución

El espacio de nombres define el conjunto de nombres posibles y tiene una estructura jerárquica frente al espacio plano que componen las direcciones IP.

El sistema de dominio de nombres es una base de datos distribuida que mantiene colecciones de relaciones entre nombres y direcciones. Es una base de datos distribuida, porque no hay ningún lugar único en Internet que conozca toda la información, en su lugar cada organización mantiene su propia base de datos de información.

El mecanismo de resolución de nombres permite a las aplicaciones obtener la dirección IP a partir del nombre de un nodo proporcionado por el usuario, permitiendo de este modo la comunicación entre clientes y servidores.

La implementación del mecanismo de resolución de nombres es el Servidor de Nombres. Cada organización debe disponer de un Servidor de Nombres que contiene la información acerca de la relación entre nombres y direcciones IP de los nodos de la misma.

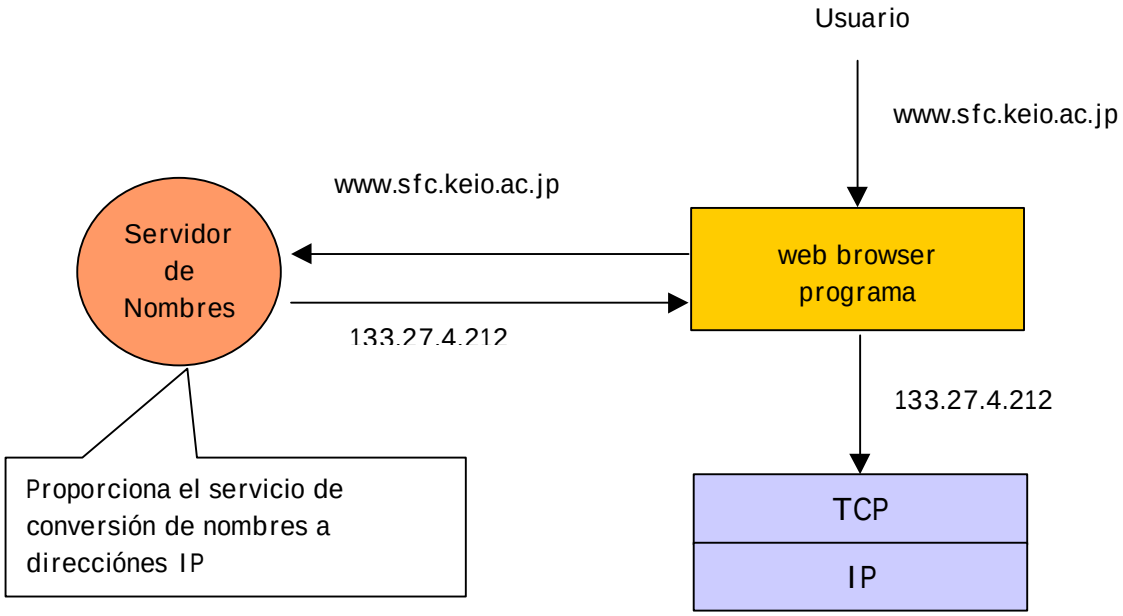

Figura 9.63

Desde el punto de vista de la aplicación, el acceso al Servidor de Nombres se efectúa mediante un “resolver”. Es importante tener en cuenta que el resolver es generalmente parte de la aplicación, y no del sistema operativo como los protocolos TCP y IP.

Todas las cuestiones relativas a la gestión del Espacio de Nombres tales como quién debe asignar nombres a los nodos, qué clase de datos deben mantener los Servidores de Nombres, así como las relativas a la resolución de nombres en internet, como el formato de los mensajes o la

responsabilidad de resolución de un nombre vienen descritas en dos RFCs. El RFC1034 especifica los conceptos y utilidades que proporciona DNS, y el RFC1035 detalla su implementación y especificación.

9.11.2.- ARBOL DE DOMINIO

El espacio de nombres es jerárquico, tal y como se muestra en la figura.

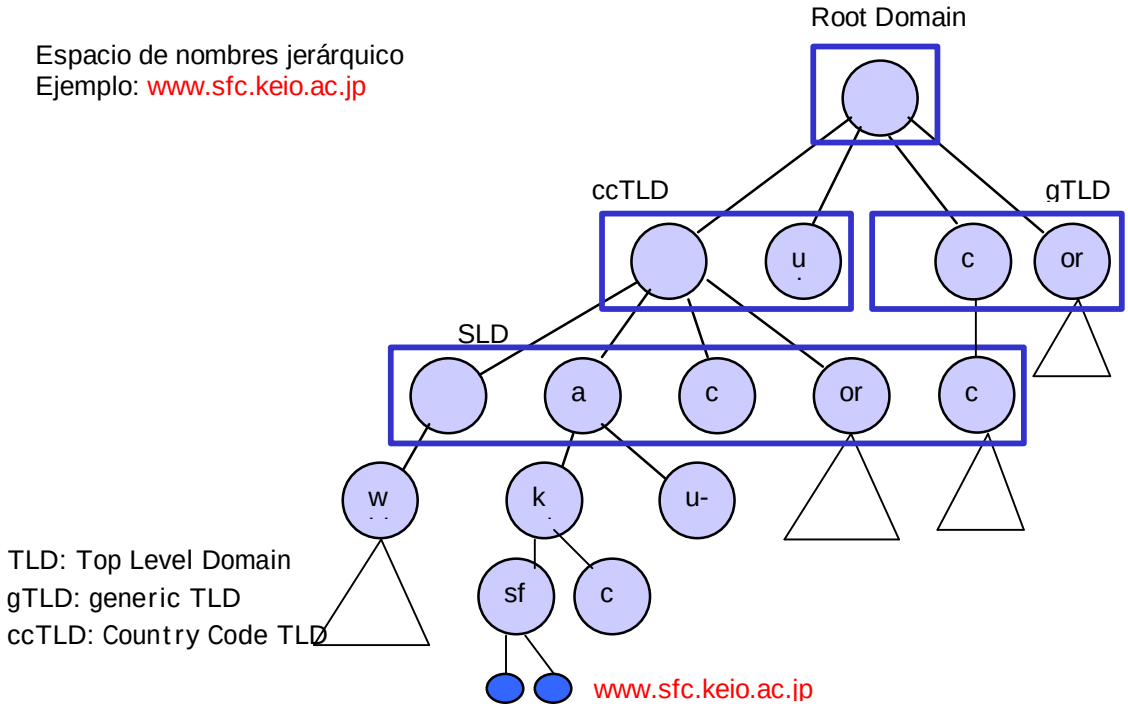

Figura 9.64

El Dominio Raíz es un nodo especial con una etiqueta nula. El nombre de dominio de un nodo cualquiera del árbol es la lista de etiquetas, comenzando en dicho nodo que lleva hasta la raíz, utilizando puntos para separar las distintas etiquetas. Cada nodo del árbol debe tener un nombre de dominio único con un máximo de 63 caracteres.

Un nombre de dominio terminado con "." se denomina un nombre de dominio absoluto o nombre de dominio completamente cualificado (FDQN). Si el nombre no termina con un punto se supone que el nombre necesita ser completado, y la forma de hacerlo dependerá del software empleado.

Los dominios del más alto nivel se dividen en dos áreas (veremos más adelante que hay una tercera:

- Los siete dominios de tres caracteres se dominan dominios genéricos u organizacionales.
- Todos los dominios de dos caracteres se basan en códigos de país y se denominan dominios de país o dominios geográficos.

La tabla siguiente contiene la clasificación normal de los siete dominios genéricos.

dominio	Descripción
com	Organizaciones comerciales
edu	Instituciones educacionales
gov	otras organizaciones gubernamentales us
int	Organizaciones internacionales
mil	Organizaciones militares us
net	Redes
Org	otras organizaciones

Tabla 9-23

Cada Dominio es mantenido por una autoridad de registro de dominio. El dominio raíz es mantenido por ICANN/IANA. Los dominios de más alto nivel (TLD) son mantenidos por los registros TLD; por ejemplo, el dominio .COM gTLD es mantenido por NSI, el dominio .ES ccTLD por ESNIC. A su vez, estos registros delegan sus subdominios a registros inferiores; por ejemplo, ESNIC delega .deusto en la Universidad de Deusto.

El arbol de dominio está dividido en unidades de administración de datos denominados zonas. Cada zona está implementada por un conjunto de Servidores de Nombres. Cada zona es administrada separadamente, y puede a su vez ser subdividida en nuevas zonas. Así como .dominio. es un límite administrativo de nombres, .zona. es un límite administrativo de mapeo de nombres.

Una vez que la autoridad para registrar nombres y direcciones a una zona se delega, corresponde al responsable de dicha zona proporcionar un conjunto de Servidores de Nombres para ella. También es posible que un mismo Servidor de Nombres pueda ser responsable de la resolución de los mismos para varias zonas.

Cada vez que se instala un nuevo sistema en una zona, el administrador del Servidor de Nombres debe registrar en su Base de Datos el nombre y la dirección IP asignados al nuevo sistema.

Cuando el servidor de nombre no contiene la información que se le solicita, debe contactar con otro servidor de nombres, por la naturaleza distribuida de la base de datos. Sin embargo, cada servidor no sabe cómo contactar con cualquier otro servidor de nombres; en vez de esto, cada servidor de nombres debe saber cómo contactar con los servidores de nombres raíz. Hay ocho servidores de nombres raíz y todos los servidores primarios deben saber las direcciones IP de cada uno de dichos servidores. Estos servidores raíz conocen los nombres y la localización (la dirección IP) de cada servidor de nombres autorizado para todos los dominios de segundo nivel. Esto supone un proceso iterativo : el servidor de nombres que pregunta debe contactar con un servidor raíz, el servidor raíz le dirá que contacte con otro servidor y así sucesivamente.

9.11.3.- REGISTROS DE RECURSOS

Cada Servidor de Nombres mantiene información que relaciona cada nombre con un valor almacenada en lo que denominan Registros de Recursos. Cada uno de estos registros contiene la siguiente información:

- Nombre
- Valor
- Tipo
- Clase
- TTL

El contenido de nombre y valor no es necesariamente un nombre de nodo o una dirección IP, sino que como veremos más adelante depende del tipo del registro.

El valor del campo clase es IN, que corresponde a Internet, lo que permite a otras entidades definir nuevas clases.

El contenido de TLL es opcional e indica durante cuanto tiempo es válido un registro.

Los tipos de registro son los siguientes:

Tipo	Valor	Descripción	Nombre	Valor
A	1	“IP Adress”	Nombre de nodo	Dirección IP
NS	2	“Name Server” Utilizado para la creación de alias	Nombre de dominio	Nombre del Servidor de Nombres de dicho dominio.
CNAME	5	“Canonic Name”	Nombre de nodo	Nombre canónico
HINFO	13	“Host info”	Nombre de nodo	Información general acerca del nodo.
MX	15	“Mail eXchange”	Nombre de dominio	Nombre del Servidor de Correo de dicho dominio.
SOA		“Start of Authority”	Nombre de dominio	Información acerca de los Servidores de Nombres y los datos que mantienen.
PTR	12	“Domain Name PoinTeR” Utilizado para las búsquedas inversas.	Dirección IP	Nombre de nodo
AXFR	252	Petición de Transferencia de Zona		
ANY	253	Petición todos los registros		

Tabla 9-24

A modo de ejemplo, vamos a listar registros de recursos de tipo A de diversos servidores de nombres relacionados con el ejemplo de la Figura 9.1 para ilustrar su organización jerárquica.

RR en los Servidores de Nombres de la “Zona Raíz”
<jp, ns1.nic.ad.jp, NS, IN>
< ns1.nic.ad.jp, 202.12.30.33, A, IN>

RR en los Servidores de Nombres de la “Zona JP”
<ad.jp, ns0.nic.ad.jp, NS, IN>
<ns0.nic.ad.jp, 202.12.30.131, A, IN>
<ac.jp, ns0.nic.ad.jp, NS, IN>
<ns0.nic.ad.jp, 202.12.30.131, A, IN>
<keio.ac.jp, ns0.keio.ac.jp, NS, IN>
<ns0.keio.ac.jp, 133.27.4.121, A, IN>
<wide.ad.jp, ns.wide.ad.jp, NS, IN>
<ns.wide.ad.jp, 203.178.136.63, A, IN>

RR en los Servidores de Nombres de la “Zona keio.ac.jp”
<sfc.keio.ac.jp, ns1.sfc.keio.ac.jp, NS, IN>
<ns1.sfc.keio.ac.jp, 133.27.4.2, A, IN>
<cc.keio.ac.jp, kogwy.cc.keio.ac.jp, NS, IN>
<kogwy.cc.keio.ac.jp, 131.113.1.1, A, IN>

RR en el Servidor de Nombres “sfc.keio.ac.jp”
<ccz02.sfc.keio.ac.jp, 133.27.4.212, A, IN>
<www.sfc.keio.ac.jp, ccz02.sfc.keio.ac.jp, CNAME, IN>
<sfc.keio.ac.jp, mail.sfc.keio.ac.jp, MX, IN>

9.11.4.- RESOLUCIÓN DE NOMBRES

Es el Servidor de Nombres de cada red quien atiende y se encarga de contestar las preguntas de todos los nodos de dicha red; por lo tanto, cada nodo debe saber cómo alcanzar al Servidor de Nombres Local.

Servidores de Nombres

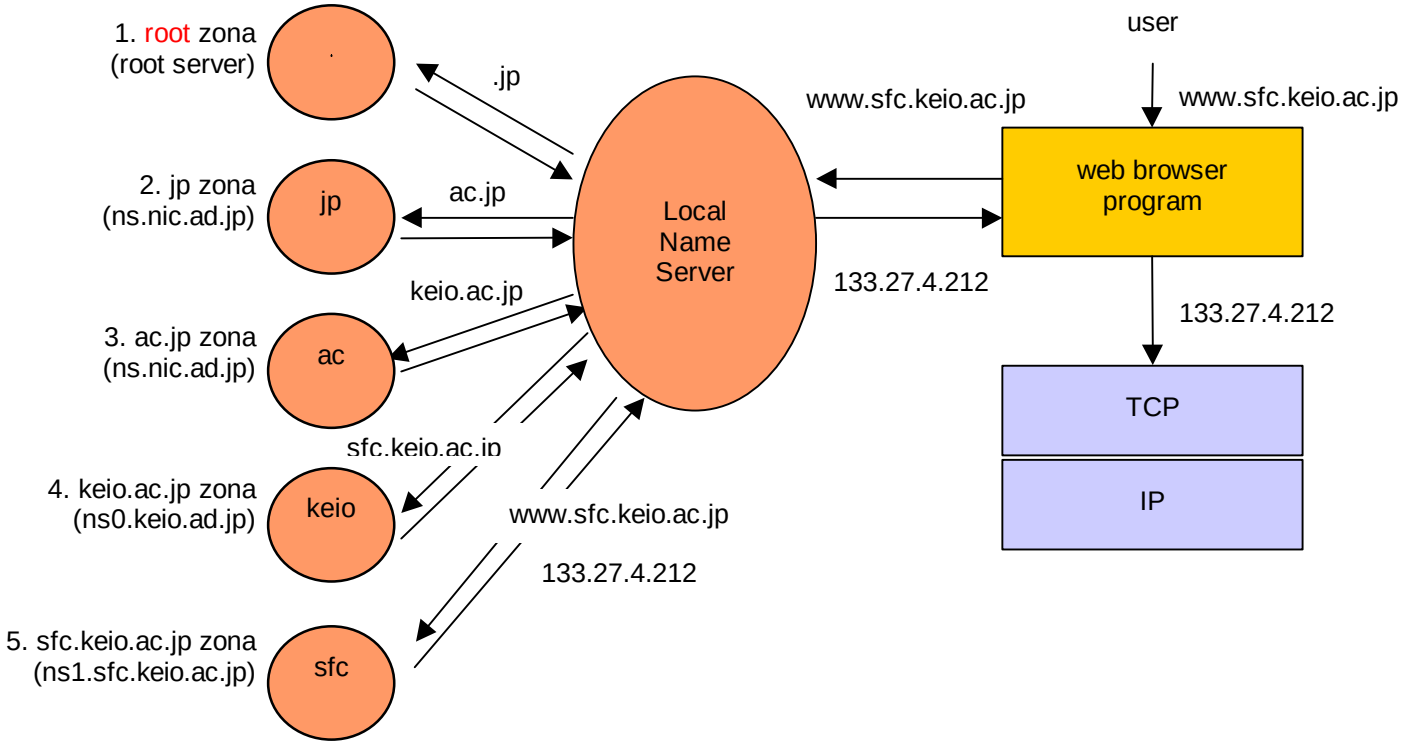

Figura 9.65

Sin embargo, el Servidor de Nombres Local sólo mantiene información sobre las correspondencias entre direcciones y nombres de su zona y por lo tanto, cuando reciba una pregunta referente a un nombre correspondiente a otra zona tendrá que localizar el Servidor de Nombres autorizado para la zona a la que correponsde el nombre preguntado. Para ello, y en primer lugar, el Servidor de Nombres Local dirigirá toda pregunta a uno de los Servidores de Nombres Raíz, cuya dirección deberá por lo tanto conocer. Resulta, por lo tanto, imprescindible que exista conectividad entre un Servidor Local y un Servidor Raíz para la búsqueda de nombres.

La lista completa de los Servidores Raíz es la siguiente:

Organización	Ciudad	Tipo	URL
InterNIC	Herndon,VA, US	com	http://www.internic.org
ISI	Marina del Rey,CA, US	edu	http://www.isi.edu/
PSInet	Herndon,VA, US	com	http://www.psi.net/
UMD	College Park,MD, US	edu	http://www.umd.edu/
NASA	Mt View, CA, US	usg	http://www.nasa.gov/
ISC	Palo Alto, CA, US	com	http://www.isc.org/
DISA	Vienna, VA, US	usg	http://nic.mil/
ARL	Aberdeen, MD, US	usg	http://www.arl.mil/
NORDUnet	Stockholm, SE	int	http://www.nordu.net/
(TBD)	(colo w/ A)	()	http://www.iana.org/
RIPE	London, UK	int	http://www.ripe.net/
(TBD)	(colo w/ B)	()	http://www.iana.org/
WIDE	Tokyo, JP	int	http://www.wide.ad.jp/

Tabla 9-25

9.11.5.- REDUNDANCIA Y RENDIMIENTO

Con el fin de asegurar el servicio de resolución de nombres, además de asignar un servidor de nombres primario a una zona es preciso asignar uno o más servidores de nombres secundarios a la misma. Los servidores primarios y secundarios deben ser independientes y redundantes de modo que la disponibilidad del servicio de nombres en la zona no pueda ser afectado por un fallo en un solo punto. La principal diferencia entre los servidores primarios y secundarios es que el primario contiene una copia maestra de los datos de la zona en sus discos, mientras que los secundarios mantienen una copia de ella que obtienen a partir del primario. Estas copias se mantienen sincronizadas mediante un proceso conocido como transferencia de zona.

Cada vez que se añade un nuevo nodo a una zona, el administrador agrega la correspondiente información (nombre y dirección IP) al disco del sistema primario; el servidor primario de nombre recibe una notificación para releer sus ficheros de configuración. Los servidores secundarios consultan a los primarios regularmente (generalmente cada tres horas) y si el servidor primario contiene nuevos datos, el servidor secundario obtendrás los nuevos datos utilizando una transferencia de zona.

Otro elemento fundamental del DNS es la “caché”. La existencia de memoria caché en los Servidores de Nombres Locales permite mejorar el rendimiento de las búsquedas, reduciéndo el tráfico en Internet. Los Servidores Locales almacenan en su memoria caché los RR utilizados recientemente, así como un registro de dónde obtuvo la equivalencia; de esta forma podrá contestar a preguntas acerca de dicha equivalencia posteriores sin tener que efectuar las consultas.Cada una de las entradas mantenidas en la caché tiene un tiempo de vida marcado en la contestación que proporcionó el Servidor de Nombres con autoridad en la zona.

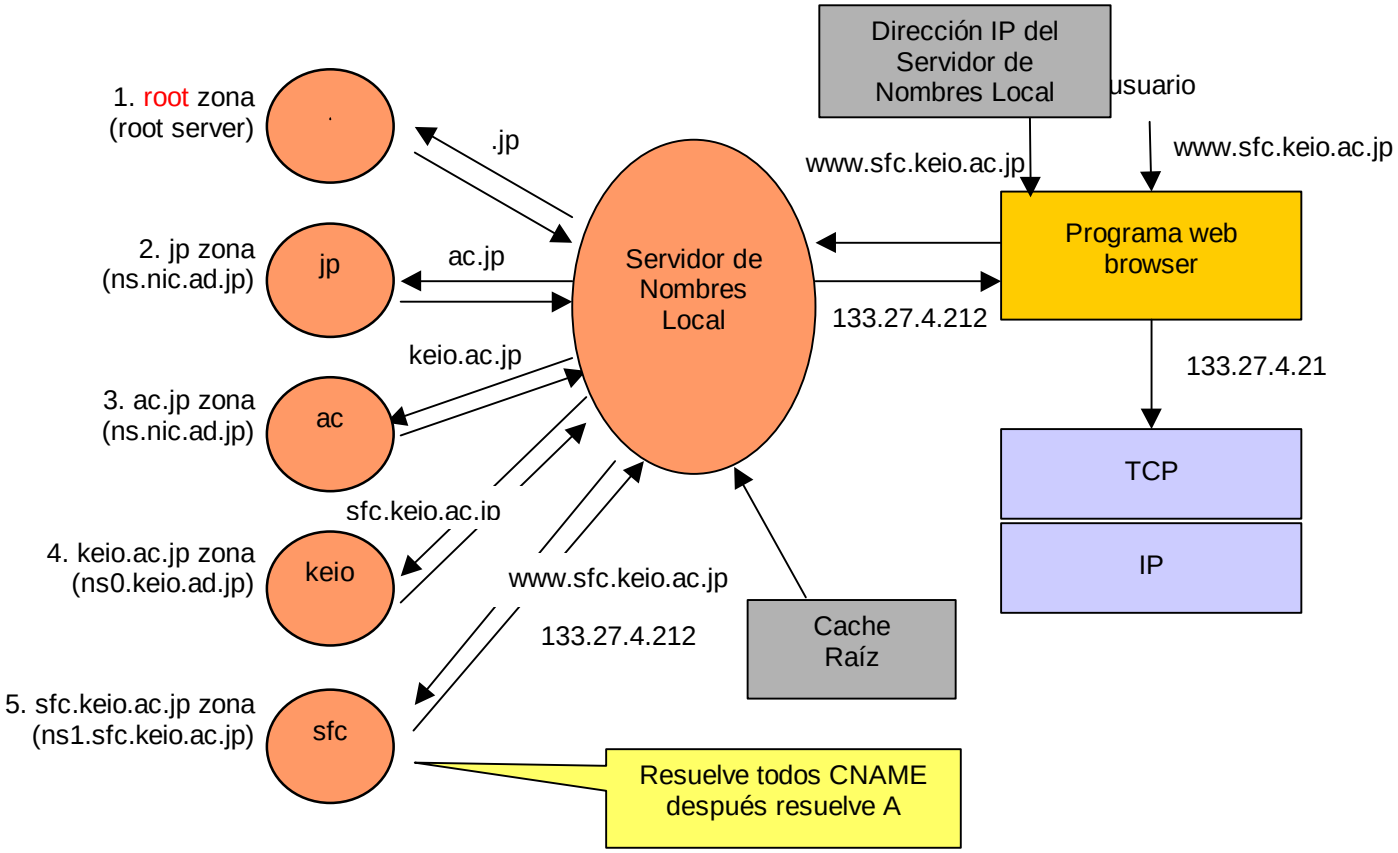

Figura 9.66

Es bueno que los Servidores de Nombres estén distribuidos topológicamente para mejorar el rendimiento. Estos servidores pueden a su vez ser autorizados o no autorizados sobre una zona, en este caso, es mejor que los servidores autorizados estén registrados en la zona superior mientras que los no autorizados no estén en la zona superior, siendo su uso generalmente interno.

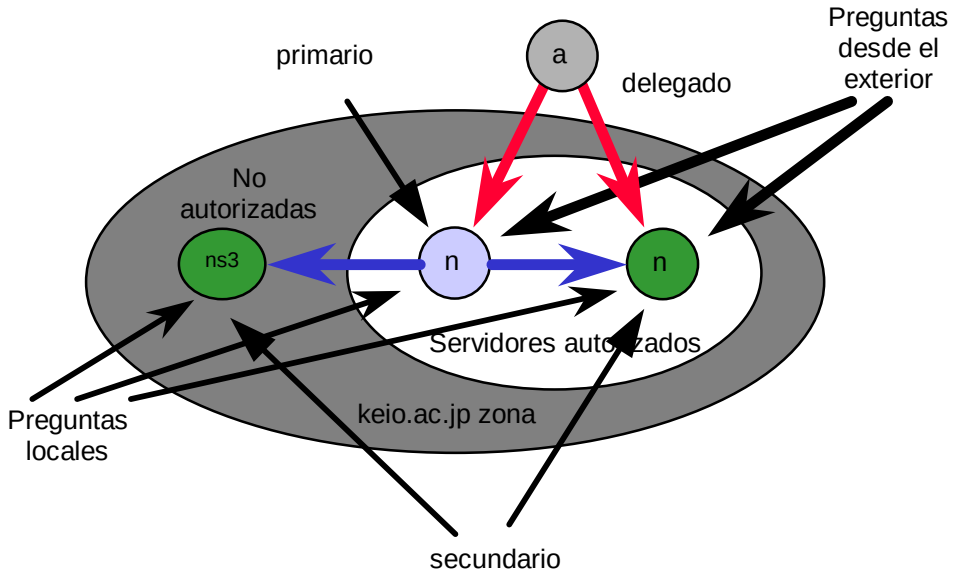

Figura 9.67

9.11.6.- FORMATO DE LOS MENSAJES DNS

Hay un solo mensaje DNS definido tanto para las consultas como para las respuestas tal y como se muestra en la figura.

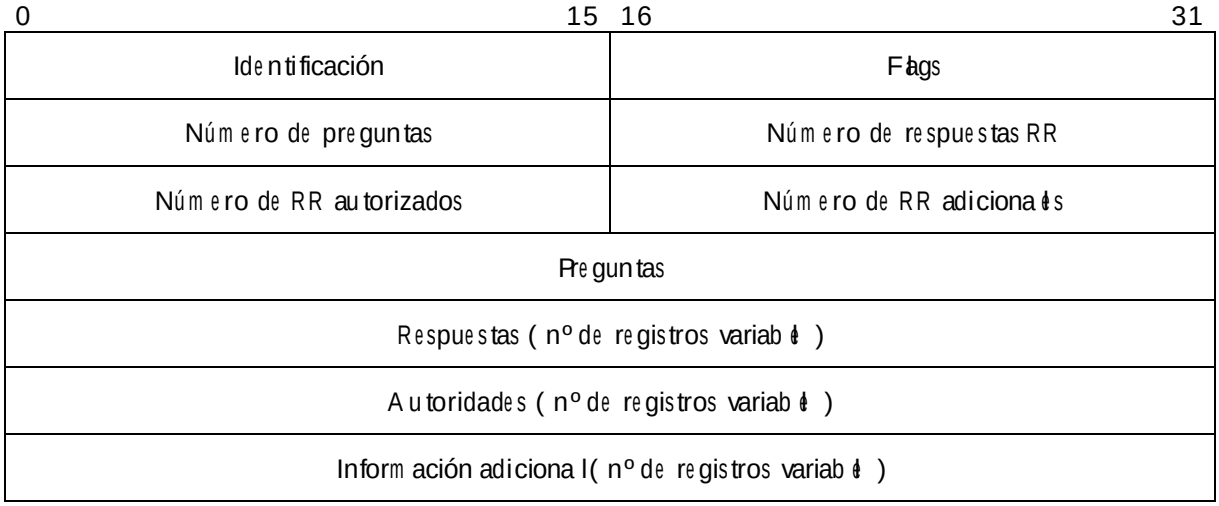

Figura 9.68

El mensaje tiene una cabecera fija de 12 bytes seguida por cuatro campos de longitud variable.

El campo **identificación** es fijado por el cliente y devuelto por el servidor en su respuesta, lo que permite al cliente relacionar las respuestas recibidas con las preguntas enviadas.

El campo **flags** está dividido en varios subcampos tal y como muestra la figura.

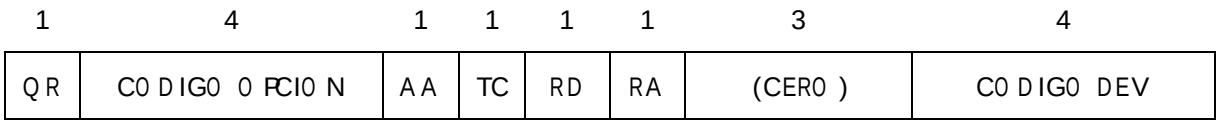

Figura 9.69

QR identifica si el mensaje es una pregunta (0) o una respuesta (1).

El **código opción** es un campo de 4 bytes. El valor normal es 0 que corresponde a una pregunta estándar, el valor 1 representa una pregunta inversa (cuál es el nombre que corresponde a una dirección IP) y el 2 corresponde a una petición del estado del servidor.

AA significa respuesta autorizada. Supone que el servidor de nombre tiene autoridad para el dominio que figura en la sección de preguntas.

TC significa truncado. El uso de UDP supone que el tamaño máximo de los datagramas será de 512 bytes, lo que puede originar que en algunos casos las respuestas estén truncadas. En este caso se activa este bit y se devuelven los primeros 512 byte de la respuesta.

RD indica si se debe tratar la pregunta de forma recursiva o iterativa. Cuando tiene el valor 1, el servidor de nombres debe manejar la pregunta de forma recursiva. Cuando tiene el valor 0, y el servidor de nombres no dispone de una respuesta, devuelve una lista de otros servidores de nombres a los cuales preguntar para obtener una respuesta.

RA indica que está disponible la opción recursiva en el Servidor de Nombres. El byte se pone a 1 en la respuesta si el servidor soporta recursión; la mayoría de los Servidores de Nombres proporcionan tratamiento recursivo, a excepción de los Servidores Raíz.

Código Devolución contiene un código del error devuelto. Los valores más habituales son 0 que significa no hay error, y 3 que significa error de nombre. Sólo un Servidor de Nombres con autoridad en el dominio especificado puede devolver un mensaje de error de nombre referente a dicho dominio.

Los siguientes cuatro campos de 16 bytes cada uno de ellos, especifican el número de entradas existentes en los cuatro campos de longitud variable que completan el mensaje, que contienen respectivamente las preguntas, los Registros de Recursos (RR) devueltos, los RR autorizados y los RR adicionales. En una pregunta, el número de cuestiones es generalmente 1 y los otros tres valores están a 0. En una respuesta, el número de respuestas es al menos 1 y los otros dos contadores pueden ser 0 o no.

9.11.6.1.- FORMATO DEL CAMPO PREGUNTAS

El campo **preguntas** generalmente contiene una única pregunta con el siguiente formato.

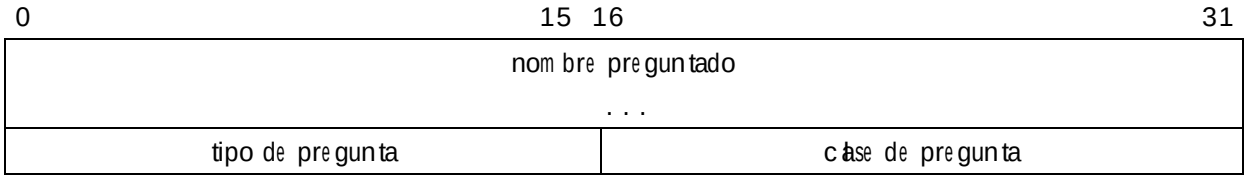

Figura 9.70

El campo que figura en la pregunta es una secuencia de una o más etiquetas, cada una de las cuales comienza con un contador de un octeto que especifica el número de octetos que le siguen componiendo la etiqueta. El nombre está terminado por un octeto de ceros, que es una etiqueta de una longitud de 0 que corresponde a la etiqueta de la raíz. Cada octeto que mide la longitud de la etiqueta puede valer de 0 a 63. A diferencia de otros formatos de mensajes de otros protocolos, este campo puede finalizar con una longitud distinta de 32 byte, y no se emplea ningún relleno.

A cada pregunta le sigue el **Tipo** y **Clase**, que identifica el tipo de registro (RR) y la clase al cual se refiere la misma y cuyos valores coinciden con los indicados en la Tabla.

9.11.6.2.- FORMATO DEL CAMPO RESPUESTAS

Los tres últimos campos del mensaje DNS, los campos RESPUESTA, AUTORIDAD e INFORMACIÓN ADICIONAL tienen el mismo formato que se muestra en la figura.

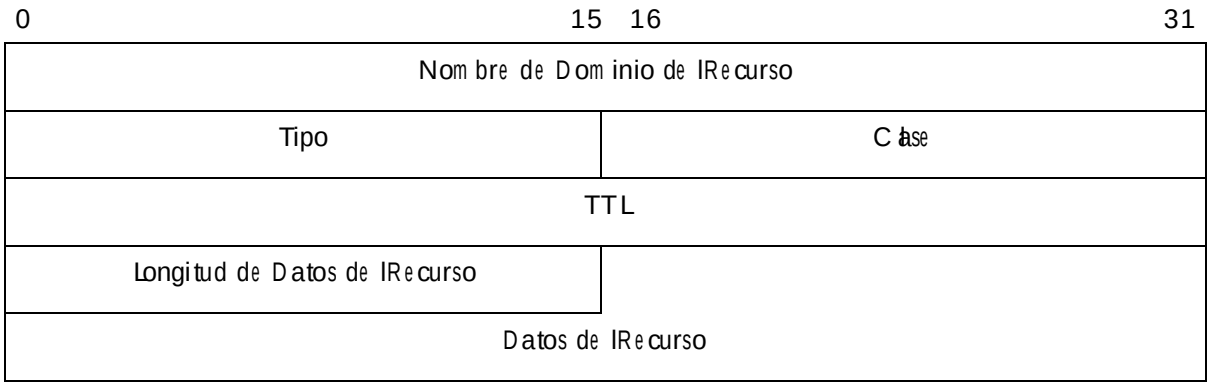

Figura 9.71

El campo Nombre de Dominio es el nombre al cual corresponden los datos de recursos siguientes; tiene el mismo formato que el descrito para el campo Nombre en la sección pregunta.

El campo Tipo especifica uno de los códigos de tipo RR, estos valores son los mismos que para el tipo de pregunta descrito anteriormente.

El campo TTL es el número de segundos que el registro de recurso debe ser almacenado en la memoria “caché” por el cliente, a menudo toma el valor de dos días.

El campo Longitud de Datos del Recurso especifica la longitud del campo Datos del Recurso. El formato de estos datos depende del tipo (por ejemplo, para un recurso de tipo IN es una dirección IP de cuatro octetos). A continuación examinaremos el formato y objeto de algunos otros registros.

9.11.7.- RR MX

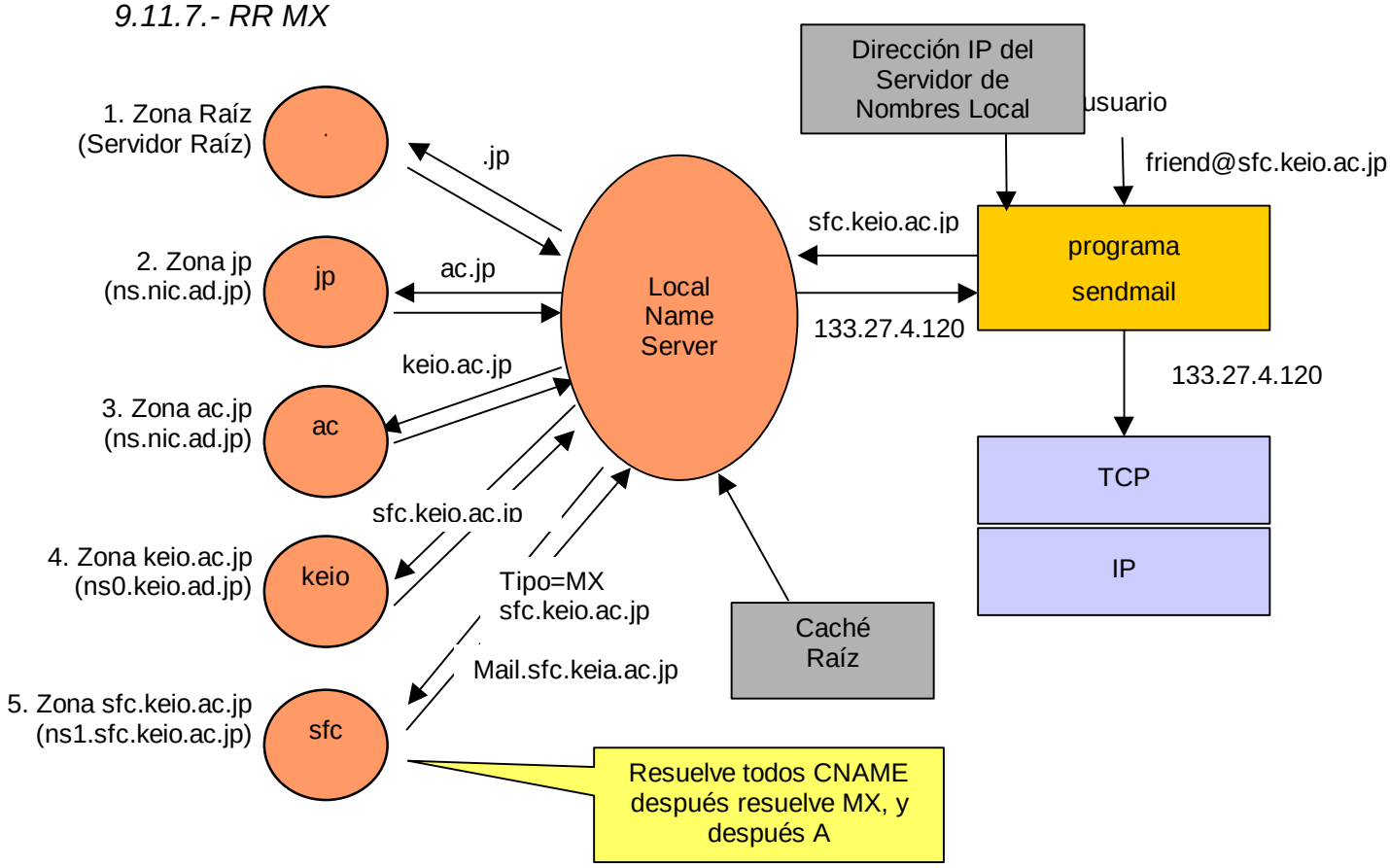

Figura 9.72

Los registros MX tienen están ordenados por “preferencia” que se indica en números, indicando mayor prioridad los números más bajos.

9.11.8.- RR PTR

Se utilizan para la resolución inversa, es decir la búsqueda de la relación entre una dirección IP y su nombre correspondiente. Existe un dominio especial denominado “in-addr.arpa” para llevar a cabo esta resolución.

En primer lugar tenemos que tener en cuenta que cuando un servidor está autorizado en una zona del árbol de dominio, también lo está sobre la porción del mismo de la rama in-addr.arpa correspondiente a la dirección IP de red de su organización. La organización de la rama del árbol de dominio in-addr.arpa se hace comenzando inmediatamente debajo del nivel in-addr.arpa con el primer octeto de la dirección IP, el siguiente nivel es el segundo octeto de la dirección IP y así sucesivamente. También hemos de tener en cuenta que los nombres se escriben desde la parte inferior del árbol de dominio hacia arriba.

Un ejemplo de registro PTR sería el siguiente :

< 7.140.178.203.in-addr.arpa, shonan.sfc.wide.ad.jp, PTR, IN>

Si no hubiera una rama separada del árbol DNS para gestionar la traducción de dirección a nombre, no habría más remedio para efectuar la traducción inversa que comenzar desde la raíz del árbol y probar cada uno de los dominios de máximo nivel, lo que llevaría días o semanas dado el tamaño actual de Internet.

Si falla la traducción de la dirección IP al nombre es posible que se produzcan problemas al hacer uso de ciertos servicios de internet y acceder a algunos lugares públicos, el motivo es que algunos IRCs deniegan el acceso a nodos con dicha dirección.

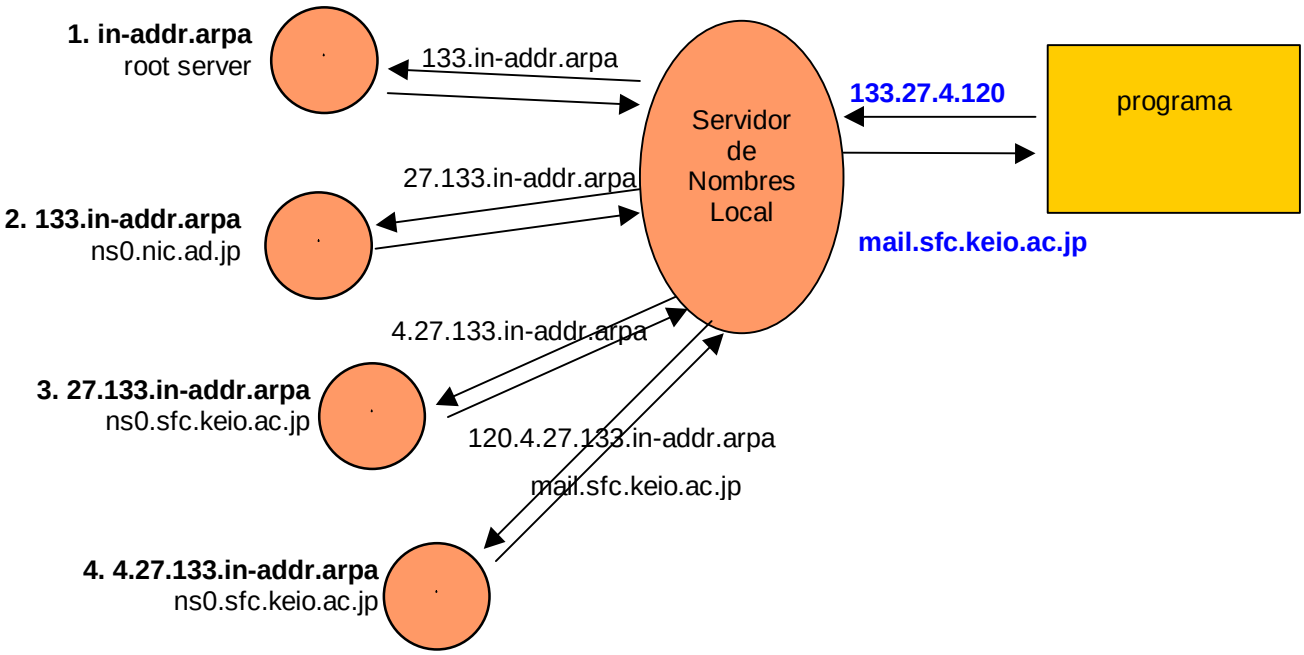

Figura 9.73

En principio PTR está diseñado para utilizarlo con un modelo de espacio de direcciones de Clases A, B, C, D:

192.0.2.* (2.0.192.in-addr.arpa) a la organización X,
192.0.3.* (3.0.192.in-addr.arpa) a la organización Y
192.0.4.* (4.0.192.in-addr.arpa) a la organización Z

¿Cómo se puede compatibilizar el esquema anterior de preguntas inversas (PTR) con el uso de CIDR?

192.0.2.0/25 a la organización X,
192.0.2.128/26 a la organización Y
192.0.2.192/26 a la organización Z

El RFC2317 “Delegación in-addr.arpa sin clases“ trata acerca del modo de resolver esta cuestión.

Utilizando CNAME en la zona superior:

<0/25.2.0.192, ns.A.domain, NS,IN>
< 1.2.0.192.in-addr.arpa, 1.0/25.2.0.192.in-addr.arpa, CNAME, IN>
< 2.2.0.192.in-addr.arpa, 2.0/25.2.0.192.in-addr.arpa, CNAME, IN>

en el Servidor de Nombres en el dominio A

<1.0/25.2.0.192, host1.ns.A.domain, A,IN>
<2.0/25.2.0.192, host2.ns.A.domain, A,IN>

9.11.9.- SOA RR

La información acerca del modo en que se mantienen sincronizados los servidores primarios y secundarios en una zona se encuentra en los registros SOA, junto con otra información acerca del período de validez de la información. Veamos un ejemplo:

```
sfc.wide.ad.jp   IN   SOA   shonan.sfc.wide.ad.jp. root.sfc.wide.ad.jp. (
                           1999112901        ; Serial
                           1800                    ; Refresh
                           900                     ; Retry
                           3600000           ; Expire
                           10800             ; TTL   )
```

El Servidor primario para la zona sfc.wide.ad.jp es shonan.sfc.wide.ad.jp y la dirección de correo electrónico de su administrador es root@sfc.wide.ad.jp.

Los Servidores Secundarios refrescan su copia cada “refresh” segundos (sólo si se actualiza el número “serial”) . Si falla la transferencia de zona, el secundario intenta refrescar su copia cada “retry” segundos.

La zona se considera fuera de servicio si después de “expire” segundos no se obtiene ninguna respuesta.

Los registros mantenidos en la caché acerca de esta zona son válidos, por defecto, durante “TTL” segundos.

9.12.- BOOTP

Para ofrecer las facilidades mencionadas anteriormente se diseñó BOOTP (BOOTstrap Protocol), que tiene fundamentalmente las siguientes ventajas sobre RARP:

- El mensaje inicial se envía utilizando UDP; al ser este un protocolo de transporte reconocido por IP puede ser enviado a través de routers; esto permite mayor flexibilidad en la ubicación del servidor; en particular éste puede estar en una ubicación remota respecto al cliente; en la LAN del cliente debe haber al menos un retransmisor (relay) Bootp que se ocupe de redirigir el paquete UDP al servidor remoto (en este caso dicho reenvío se hace en modo unicast).

- El formato de un mensaje BOOTP permite enviar muchos parámetros IP al cliente, no únicamente la dirección IP. Entre estos se encuentran por ejemplo la máscara de subred, el MTU, rutas estáticas, el valor por defecto del parámetro TTL, etc.

BOOTP fue el primer protocolo estándar para arranque automático de ordenadores en TCP/IP.

RARP permite centralizar en una o unas pocas máquinas (los servidores RARP) la información de direcciones IP de todas las máquinas de una red. BOOTP permite además mantener centralizados multitud de parámetros de configuración de red de cada máquina, y debido a su mayor flexibilidad de uso ha desplazado prácticamente por completo a RARP. Esta facilidad de configuración centralizada resulta especialmente útil en redes grandes, ya que permite hacer de manera cómoda cambios en la configuración de máquinas de una red sin tener que ir de una en una haciendo los cambios localmente.

9.12.1.- Formato de los mensajes BOOTP

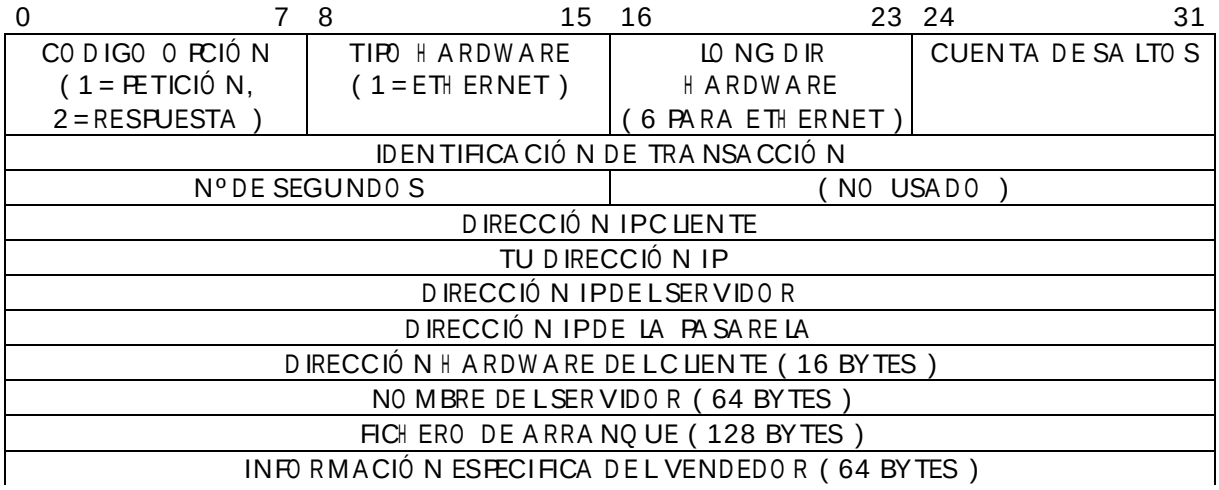

Figura 9.74

El campo **código opción** identifica peticiones y respuestas (1 y 2 respectivamente). El campo **tipo de hardware** identifica el protocolo de enlace de datos que se emplea (toma el valor 1 para Ethernet, al igual que en ARP), de la misma forma el campo **longitud hardware** especifica la longitud de las direcciones de enlace de datos (cuyo valor es 6 octetos en Ethernet).

El campo **contador de saltos** es inicializado por el cliente y lo emplean los servidores proxy como veremos más adelante.

El identificador de transacción es un entero de 32 bits que empleado para relacionar respuestas con peticiones. El número de segundos indica el tiempo pasado en el cliente desde que empezó a intentar arrancar, este valor sirve a los servidores secundarios para no contestar hasta que haya transcurrido un tiempo sin que el servidor primario haya contestado.

Las **dirección IP del cliente** puede ser completada por éste si ya la conoce, en caso contrario el servidor la incluirá en el campo **tu dirección IP**, a la vez que incluye la suya en el campo **dirección IP del servidor**. Si se utiliza un servidor proxy incluirá su dirección en el campo **dirección IP de la pasarela**.

El cliente debe incluir su **dirección hardware** para que el servidor puede efectuar la búsqueda de la dirección IP. El campo **nombre del servidor** es una cadena que puede ser completada por el servidor de manera opcional, al igual que el **nombre del fichero de arranque**, donde se incluiría el nombre completo del directorio en que se encuentra.

9.13.- DHCP.

Tanto RARP como BOOTP requieren una asignación estática biunívoca entre direcciones MAC y direcciones IP; hacen falta tantas direcciones IP como ordenadores vayan a utilizar el protocolo TCP/IP. Existen situaciones en las que esto no es conveniente, por ejemplo:

- Una empresa con 500 ordenadores quiere conectarse a la Internet, de forma que cualquiera de ellos pueda salir al exterior; para esto dispone de una clase C; se sabe que estadísticamente nunca habrá más de 250 ordenadores simultáneamente conectados a la Internet, por lo que en principio una clase C sería suficiente; utilizando BOOTP no puede ofrecerse el servicio a más de 254 ordenadores, al tener que efectuar una asignación estática de las direcciones MAC a direcciones IP.

- En un congreso se habilita una sala para la conexión a la Internet de los ordenadores portátiles de los participantes; la sala dispone de una red local con 50 puntos de enganche, por lo que se han asignado para su utilización 50 direcciones IP; con BOOTP sería preciso averiguar de antemano las direcciones MAC que tendrán los ordenadores de los participantes, y asignar direcciones IP a los primeros 50 solicitantes únicamente.

Claramente en estas situaciones es necesario un mecanismo más flexible de asignación de direcciones IP que el ofrecido por BOOTP.

Para resolver estos problemas el IETF diseñó en 1993 el protocolo DHCP (Dynamic Host Configuration Protocol), descrito en el RFC1541, muy similar al BOOTP pero permite una asignación dinámica de direcciones IP.

DHCP se encuentra dividido en dos partes: un protocolo para enviar información de configuración específica desde un servidor a una máquina, y un mecanismo de asignación de direcciones IP. Esta basado en el modelo cliente-servidor, dónde un servidor DHCP es el que proporciona a las máquinas los parámetros de configuración necesarios, entre ellos una dirección IP en 'alquiler' para poder trabajar; el tiempo que dura el alquiler es negociado entre cliente y el servidor en el momento de establecer la conexión, y puede variar entre unos pocos minutos o duración indefinida.

En cuanto a la asignación de direcciones IP, ésta se puede llevar a cabo de 3 modos distintos:

- Asignación automática: el servidor DHCP asigna una dirección IP fija y permanente a la máquina, sin que el administrador intervenga.
- Asignación dinámica: el servidor DHCP asigna una dirección IP temporal a la máquina, sin que el administrador intervenga.
- Asignación manual: al administrador configura en el servidor DHCP la dirección IP a asignar a cada máquina.

De estos tres mecanismos, sólo la asignación dinámica permite la reutilización de direcciones IP. El cliente DHCP, que forma parte del sistema operativo de la máquina contacta con un servidor DHCP solicitándole la asignación de una dirección IP. El servidor DHCP, que gestiona una lista con las direcciones IP que puede asignar, reserva una de las direcciones IP que están libres para la máquina y le envía un mensaje notificándosela. El cliente la recibe y configura su pila de protocolos TCP/IP con la misma. Este proceso se realiza como parte de la inicialización del sistema de la máquina.

Este proceso es útil cuando una máquina no requiere una dirección IP fija continuamente o cuando hay más máquinas que direcciones IP disponibles, permitiendo que distintas máquinas utilicen la misma dirección IP en distintos instantes de tiempo.

La mayor flexibilidad de DHCP le ha convertido en el protocolo preferido para la configuración remota de ordenadores en una red local. Con DHCP se mejora notablemente la seguridad y fiabilidad de una red; también se simplifican las albores de administración de la red.

Un inconveniente de DHCP frente a RARP o BOOTP es que, al no haber una asignación permanente de direcciones IP, si se desea rastrear un problema pasado cierto tiempo y sólo se dispone de la dirección IP resulta más difícil (a veces imposible) averiguar que ordenador o usuario ha sido el causante del problema.

9.13.1.- FORMATO DEL MENSAJE Y FUNCIONAMIENTO

Los clientes solicitan sus parámetros de configuración, de los cuales el más importante es la dirección IP, al servidor, incluyendo en el mensaje un campo Identificador de Cliente unívoco, mediante el cuál el servidor puede localizarle para enviar la respuesta. El valor de este campo suele ser la dirección hardware de la conexión a la red del cliente, ya que es un valor único.

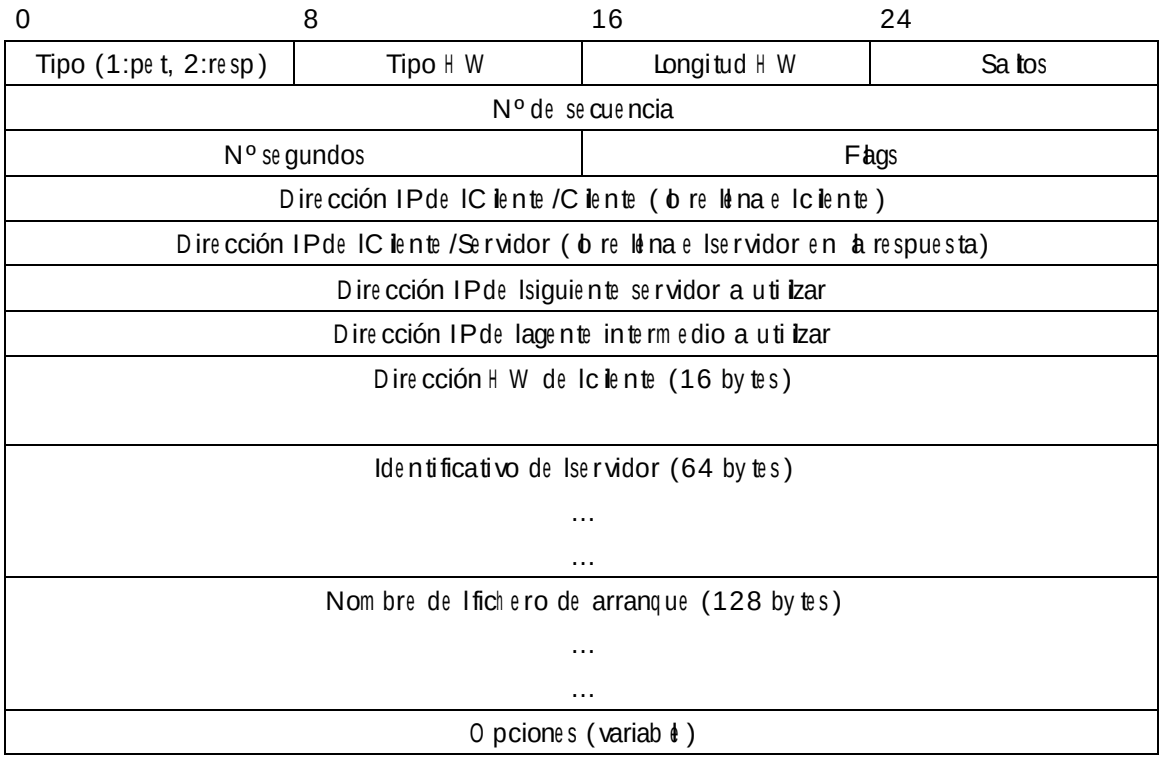

Figura 9.75

Los campos más importantes son el Tipo de mensaje (1: Petición y 2: Respuesta), el tipo y longitud de la dirección hardware (como en ARP) y el campo de dirección IP que rellena el servidor.

Los servidores DHCP siempre intentan asignar la misma dirección IP a una máquina en diferentes peticiones DHCP a lo largo del tiempo, siempre y cuando dicha dirección esté libre. Esta asignación se considera una especie de préstamo, pudiendo la máquina renovarlo en un momento dado mediante un mensaje DHCP de petición, rellenando en este caso el campo Dirección IP de Cliente/Cliente con la dirección asignada actualmente cuyo préstamo esta a punto de finalizar.

El campo Nº segundos indica el número de segundos transcurridos desde que el cliente inició el proceso de petición o renovación de dirección IP.

Toda la información sobre el tiempo de alquiler deseado por el cliente, si se trata de una solicitud de renovación, si se trata de un mensaje de devolución de dirección IP, etc. se incluye en el campo opciones.

El proceso de configuración de una máquina mediante DHCP es el siguiente:

1. El cliente difunde un mensaje para conocer que servidores DHCP hay disponibles, incluyendo en el campo Opciones, valores sugeridos de tiempo de alquiler, dirección IP, etc.
2. Los servidores DHCP responden ofreciéndole al cliente distintas configuraciones.
3. El cliente escoge el servidor más conveniente y difunde un mensaje de solicitud DHCP rellenando el campo Identificativo del Servidor con el valor del servidor elegido.
4. Todos los servidores reciben el mensaje, los no elegidos lo ignoran sabiendo que el cliente no ha aceptado su ofrecimiento.
5. El servidor elegido atiende la petición del cliente enviándole los parámetros de configuración apropiados.
6. El cliente DHCP recibe el mensaje con el valor de los parámetros y se autoconfigura.