UNIÓN INTERNACIONAL DE TELECOMUNICACIONES

UIT-T Y.1291

(05/2004)

SECTOR DE NORMALIZACIÓN
DE LAS TELECOMUNICACIONES
DE LA UIT

SERIE Y: INFRAESTRUCTURA MUNDIAL DE LA INFORMACIÓN, ASPECTOS DEL PROTOCOLO INTERNET Y REDES DE LA PRÓXIMA GENERACIÓN

Aspectos del protocolo Internet – Arquitectura, acceso, capacidades de red y gestión de recursos

Marco arquitectural para el soporte de calidad de servicio en redes de paquetes

Recomendación UIT-T Y.1291

RECOMENDACIONES UIT-T DE LA SERIE Y

INFRAESTRUCTURA MUNDIAL DE LA INFORMACIÓN, ASPECTOS DEL PROTOCOLO INTERNET Y REDES DE LA PRÓXIMA GENERACIÓN

INFRAESTRUCTURA MUNDIAL DE LA INFORMACIÓN	
Generalidades	Y.100–Y.199
Servicios, aplicaciones y programas intermedios	Y.200–Y.299
Aspectos de red	Y.300–Y.399
Interfaces y protocolos	Y.400–Y.499
Numeración, direccionamiento y denominación	Y.500–Y.599
Operaciones, administración y mantenimiento	Y.600–Y.699
Seguridad	Y.700–Y.799
Características	Y.800–Y.899
ASPECTOS DEL PROTOCOLO INTERNET	
Generalidades	Y.1000–Y.1099
Servicios y aplicaciones	Y.1100–Y.1199
Arquitectura, acceso, capacidades de red y gestión de recursos	**Y.1200–Y.1299**
Transporte	Y.1300–Y.1399
Interfuncionamiento	Y.1400–Y.1499
Calidad de servicio y características de red	Y.1500–Y.1599
Señalización	Y.1600–Y.1699
Operaciones, administración y mantenimiento	Y.1700–Y.1799
Tasación	Y.1800–Y.1899
REDES DE LA PRÓXIMA GENERACIÓN	
Marcos y modelos arquitecturales funcionales	Y.2000–Y.2099
Calidad de servicio y calidad de funcionamiento	Y.2100–Y.2199
Aspectos relativos a los servicios: capacidades y arquitectura de servicios	Y.2200–Y.2249
Aspectos relativos a los servicios: interoperabilidad de servicios y redes en las redes de próxima generación	Y.2250–Y.2299
Numeración, denominación y direccionamiento	Y.2300–Y.2399
Gestión de red	Y.2400–Y.2499
Arquitecturas y protocolos de control de red	Y.2500–Y.2599
Seguridad	Y.2700–Y.2799
Movilidad generalizada	Y.2800–Y.2899

Para más información, véase la Lista de Recomendaciones del UIT-T.

Recomendación UIT-T Y.1291

Marco arquitectural para el soporte de calidad de servicio en redes de paquetes

Resumen

Esta Recomendación proporciona un marco arquitectural para el soporte de calidad de servicio en redes de paquetes. El marco arquitectural es un conjunto de mecanismos de red genéricos (o bloques de construcción de QoS) utilizados para controlar la respuesta del servicio de red a una petición de servicio, que puede ser específica de un elemento de red, o para la señalización entre elementos de red, o para administrar y controlar tráfico a través de una red. Los bloques de construcción están distribuidos en tres planos lógicos (plano de control, plano de datos y plano de gestión) y pueden utilizarse combinadamente en forma de diversos métodos para proporcionar el efecto colectivo satisfactorio de una calidad de funcionamiento del servicio, variante, requerida por una gama de aplicaciones como la transferencia de ficheros y la conferencia multimedia.

Orígenes

La Recomendación UIT-T Y.1291 fue aprobada el 7 de mayo de 2004 por la Comisión de Estudio 13 (2001-2004) del UIT-T por el procedimiento de la Recomendación UIT-T A.8.

PREFACIO

La UIT (Unión Internacional de Telecomunicaciones) es el organismo especializado de las Naciones Unidas en el campo de las telecomunicaciones. El UIT-T (Sector de Normalización de las Telecomunicaciones de la UIT) es un órgano permanente de la UIT. Este órgano estudia los aspectos técnicos, de explotación y tarifarios y publica Recomendaciones sobre los mismos, con miras a la normalización de las telecomunicaciones en el plano mundial.

La Asamblea Mundial de Normalización de las Telecomunicaciones (AMNT), que se celebra cada cuatro años, establece los temas que han de estudiar las Comisiones de Estudio del UIT-T, que a su vez producen Recomendaciones sobre dichos temas.

La aprobación de Recomendaciones por los Miembros del UIT-T es el objeto del procedimiento establecido en la Resolución 1 de la AMNT.

En ciertos sectores de la tecnología de la información que corresponden a la esfera de competencia del UIT-T, se preparan las normas necesarias en colaboración con la ISO y la CEI.

NOTA

En esta Recomendación, la expresión "Administración" se utiliza para designar, en forma abreviada, tanto una administración de telecomunicaciones como una empresa de explotación reconocida de telecomunicaciones.

La observancia de esta Recomendación es voluntaria. Ahora bien, la Recomendación puede contener ciertas disposiciones obligatorias (para asegurar, por ejemplo, la aplicabilidad o la interoperabilidad), por lo que la observancia se consigue con el cumplimiento exacto y puntual de todas las disposiciones obligatorias. La obligatoriedad de un elemento preceptivo o requisito se expresa mediante las frases "tener que, haber de, hay que + infinitivo" o el verbo principal en tiempo futuro simple de mandato, en modo afirmativo o negativo. El hecho de que se utilice esta formulación no entraña que la observancia se imponga a ninguna de las partes.

PROPIEDAD INTELECTUAL

La UIT señala a la atención la posibilidad de que la utilización o aplicación de la presente Recomendación suponga el empleo de un derecho de propiedad intelectual reivindicado. La UIT no adopta ninguna posición en cuanto a la demostración, validez o aplicabilidad de los derechos de propiedad intelectual reivindicados, ya sea por los miembros de la UIT o por terceros ajenos al proceso de elaboración de Recomendaciones.

En la fecha de aprobación de la presente Recomendación, la UIT ha recibido notificación de propiedad intelectual, protegida por patente, que puede ser necesaria para aplicar esta Recomendación. Sin embargo, debe señalarse a los usuarios que puede que esta información no se encuentre totalmente actualizada al respecto, por lo que se les insta encarecidamente a consultar la base de datos sobre patentes de la TSB.

© UIT 2005

Reservados todos los derechos. Ninguna parte de esta publicación puede reproducirse por ningún procedimiento sin previa autorización escrita por parte de la UIT.

ÍNDICE

Página

Recomendación UIT-T Y.1291

Marco arquitectural para el soporte de calidad de servicio en redes de paquetes

1 Alcance

Esta Recomendación proporciona un marco arquitectural para el soporte de calidad de servicio en redes de paquetes. El marco arquitectural se basa en un conjunto de bloques de construcción de calidad de servicio distribuidos en tres planos lógicos (plano de control, plano de datos y plano de gestión) para controlar la calidad de funcionamiento de la red, incluso en caso de contienda por recursos de la red. En último término, los bloques de construcción ayudan a producir el "efecto colectivo de calidad de funcionamiento del servicio que determina el grado de satisfacción del usuario del servicio".

2 Referencias

Las siguientes Recomendaciones del UIT-T y otras referencias contienen disposiciones que, mediante su referencia en este texto, constituyen disposiciones de la presente Recomendación. Al efectuar esta publicación, estaban en vigor las ediciones indicadas. Todas las Recomendaciones y otras referencias son objeto de revisiones por lo que se preconiza que los usuarios de esta Recomendación investiguen la posibilidad de aplicar las ediciones más recientes de las Recomendaciones y otras referencias citadas a continuación. Se publica periódicamente una lista de las Recomendaciones UIT-T actualmente vigentes. En esta Recomendación, la referencia a un documento, en tanto que autónomo, no le otorga el rango de una Recomendación.

– Recomendación UIT-T E.360.1 (2002), *Encaminamiento orientado a la calidad de servicio y métodos de ingeniería de tráfico conexos para redes multiservicios basadas en el protocolo Internet, modo de transferencia asíncrono y multiplexación por división en el tiempo*.

– Recomendación UIT-T E.360.2 (2002), *Encaminamiento orientado a la calidad de servicio y métodos de ingeniería de tráfico conexos – Métodos de encaminamiento de llamadas y de encaminamiento de la conexión*.

– Recomendación UIT-T E.360.3 (2002), *Encaminamiento orientado a la calidad de servicio y métodos de ingeniería de tráfico conexos – Métodos de gestión de recursos orientada a la calidad de servicio*.

– Recomendación UIT-T E.360.4 (2002), *Encaminamiento orientado a la calidad de servicio y métodos de ingeniería de tráfico conexos – Métodos y requisitos de la gestión de tablas de encaminamiento*.

– Recomendación UIT-T E.360.5 (2002), *Encaminamiento orientado a la calidad de servicio y métodos de ingeniería de tráfico conexos – Métodos de encaminamiento de transporte*.

– Recomendación UIT-T E.360.6 (2002), *Encaminamiento orientado a la calidad de servicio y métodos de ingeniería de tráfico conexos – Métodos de gestión de capacidad*.

– Recomendación UIT-T E.360.7 (2002), *Encaminamiento orientado a la calidad de servicio y métodos de ingeniería de tráfico conexos – Requisitos operacionales de ingeniería de tráfico*.

– Recomendación UIT-T E.361 (2003), *Soporte de encaminamiento de la calidad de servicio para el interfuncionamiento de las clases de calidad de servicio con diversas tecnologías de encaminamiento*.

– Recomendación UIT-T E.860 (2002), *Marco de un acuerdo de nivel de servicio*.

– Recomendación UIT-T G.114 (2003), *Tiempo de transmisión en un sentido*.

– Recomendación UIT-T G.1000 (2001), *Calidad de servicio de las comunicaciones: Marco y definiciones*.

– Recomendación UIT-T G.1010 (2001), *Categorías de calidad de servicio para los usuarios de extremo de servicios multimedios*.

– Recomendación UIT-T I.350 (1993), *Aspectos generales de calidad de servicio y de calidad de funcionamiento en las redes digitales incluidas las redes digitales de servicios integrados*.

– Recomendación UIT-T J.112 (1998), *Sistemas de transmisión para servicios interactivos de televisión por cable*.

– Recomendación UIT-T J.162 (2004), *Protocolo de señalización de llamada de red para la prestación de servicios dependientes del tiempo por redes de televisión por cable que utilizan módems de cable*.

– Recomendación UIT-T J.163 (2004), *Calidad de servicio dinámica para la prestación de servicios en tiempo real por las redes de televisión por cable que utilizan módems de cable*.

– Recomendación UIT-T J.170 (2002), *Especificación de la seguridad de IPCablecom*.

– Recomendación UIT-T J.174 (2002), *Calidad de servicio interdominio IPCablecom*.

– Recomendación UIT-R M.1079-2 (2003), *Requisitos relativos a la calidad de funcionamiento y servicio en las redes de acceso a telecomunicaciones móviles internacionales-2000 (IMT-2000)*.

– Recomendación UIT-T X.805 (2003), *Arquitectura de seguridad para sistemas de comunicaciones de extremo a extremo*.

– Recomendación UIT-T Y.1221 (2002), *Control de tráfico y control de congestión en las redes basadas en el protocolo Internet*.

– Recomendación UIT-T Y.1540 (2002), *Servicio de comunicación de datos con protocolo Internet – Parámetros de calidad de funcionamiento relativos a la disponibilidad y la transferencia de paquetes del protocolo Internet*.

– Recomendación UIT-T Y.1541 (2002), *Objetivos de calidad de funcionamiento de red para servicios basados en el protocolo Internet*.

3 Definiciones

En esta Recomendación no se definen nuevos términos.

4 Abreviaturas, siglas o acrónimos

En esta Recomendación se utilizan las siguientes abreviaturas, siglas o acrónimos.

DiffServ Servicios diferenciados (*differentiated services*)

DQoS Calidad de servicio dinámica (*dynamic QoS*)

IETF Grupo de tareas especiales de Ingeniería en Internet (*Internet Engineering Task Force*)

IntServ Servicios integrados (*integrated services*)

LSP Trayecto conmutado por etiquetas (*label switched path*)

MPLS Conmutación por etiquetas multiprotocolo (*multi-protocol label switching*)

MTA Adaptador de terminal multimedios, (multimedia) (*multimedia terminal adaptor*)

QoS Calidad de servicio (*quality of service*)

RSVP Protocolo de reserva de recursos (*resource reservation protocol*)

SLA Acuerdo de nivel de servicio (*service level agreement*)

UIT-T Unión Internacional de Telecomunicaciones – Sector de Normalización de las Telecomunicaciones

5 Introducción

En último término, el objetivo de la calidad de servicio (QoS, *quality of service*) es el soporte de las características y propiedades de aplicaciones específicas. Sin embargo, las necesidades de las distintas aplicaciones pueden ser muy diferentes. Por ejemplo, en telemedicina, la exactitud de los datos entregados es más importante que el retardo global o la variación del retardo de los paquetes (por ejemplo, la fluctuación de fase), mientras que en telefonía IP, la fluctuación de fase y el retardo son de capital importancia y deben reducirse al mínimo. Varias Recomendaciones UIT-T tratan la QoS. La Rec. UIT-T E.800 define la QoS como "el efecto colectivo de la *calidad de funcionamiento del servicio*, que determina el grado de satisfacción del usuario del servicio". Dado que la Rec. UIT-T E.800 considera que el soporte, operabilidad, posibilidad de ser servido (*serviceability*), y seguridad son, todos ellos, parte de la calidad de funcionamiento del servicio, esta definición de QoS tiene un vasto alcance. Ampliando el concepto de QoS E.800, la Rec. UIT-T G.1000 descompone la *calidad de funcionamiento del servicio* (o calidad de servicio) en componentes funcionales y la vincula a la calidad de funcionamiento de red tal como se define en las Recomendaciones UIT-T I.350, Y.1540, e Y.1541. Como complemento a la Rec. UIT-T G.1000, que presenta un marco, la Rec. UIT-T G.1010 describe requisitos de aplicación centrados en el usuario de extremo en términos de amplias categorías (tales como interactivo, tolerante a los errores). En lo que respecta a aplicaciones o parámetros de calidad de funcionamiento específicos, entre las normas conexas, la Rec. UIT-R M.1079-2 define los requisitos de calidad vocal y calidad de datos, así como de calidad de funcionamiento, de extremo a extremo, para las redes de acceso IMT-2000, en tanto que la Rec. UIT-T G.114 especifica los límites para los tiempos de transmisión y las conexiones a través de una red digital.

Para proporcionar la calidad de funcionamiento de red requerida es necesario situar ciertos mecanismos en la red. Estos mecanismos deberán controlar y proporcionar las diversas respuestas de servicio red, incluso en caso de que se produzcan contiendas por recursos de red. RFC 2990 del IETF resume como sigue las posibles características de la respuesta de servicio controlada a una determinada petición de servicio: *consistente y predecible*, *a un nivel igual o superior a un mínimo garantizado*, o *establecido previamente.* Por ejemplo, en caso de contienda por un recurso de red, o de congestión, para mantener la respuesta de servicio esperada es necesario disponer de una diversidad de medios que permitan trabajar en diferentes escalas de tiempo, que van desde aquellas que proporcionan una cuidadosa planificación de red basada en patrones de tráfico en un largo periodo de tiempo, hasta las que utilizan una atribución de recursos y un control de admisión diferenciales basados en las condiciones de carga existentes en la red en ese momento. Estos y otros mecanismos (por ejemplo, un método de señalización para indicar el nivel deseado de calidad de funcionamiento de la red) están en el centro del marco arquitectural para el soporte de la QoS. En particular, esta Recomendación identifica un conjunto de mecanismos de red QoS genéricos y proporciona una estructura para ellos. En último término, los mecanismos de red habrán de utilizarse combinadamente para producir el efecto colectivo satisfactorio de una calidad de funcionamiento de servicio, variante, requerida por una amplia gama de aplicaciones. La independencia con respecto a la aplicación del marco arquitectural identificado la distingue de arquitecturas QoS específicas de la aplicación como la definida en la Rec. UIT-T H.360, que es específica de aplicaciones multimedios.

6 Bloques de construcción QoS

Esencial para el marco arquitectural QoS es un conjunto de mecanismos de red genéricos para controlar la respuesta de servicio de red a una petición de servicio, la cual puede ser específica de un elemento de red, o para señalización entre elementos de red, o para controlar y administrar tráfico a través de la red. (Obsérvese que los bloques de construcción no deben considerarse elementos de extremo a extremo.) Como se muestra en la figura 1, los bloques de construcción están distribuidos en tres planos:

- Plano de control, que contiene mecanismos que trabajan sobre las vías por las que se transmite tráfico de usuario. Estos mecanismos incluyen control de admisión, encaminamiento QoS, y reserva de recursos.
- Plano de datos, que contiene mecanismos que trabajan con tráfico de usuario directamente. Estos mecanismos incluyen gestión de memorias tampón, prevención de congestión, marcado de paquetes, puesta en cola y calendarización, clasificación de tráfico, aplicación de políticas de tráfico y conformación de tráfico.
- Plano de gestión, que contiene mecanismos que se ocupan de los aspectos de operación, administración, y gestión de la red. Estos mecanismos incluyen acuerdo de nivel de servicio (SLA, *service level agreement*), restablecimiento de tráfico, metraje y registro, y aplicación de políticas.

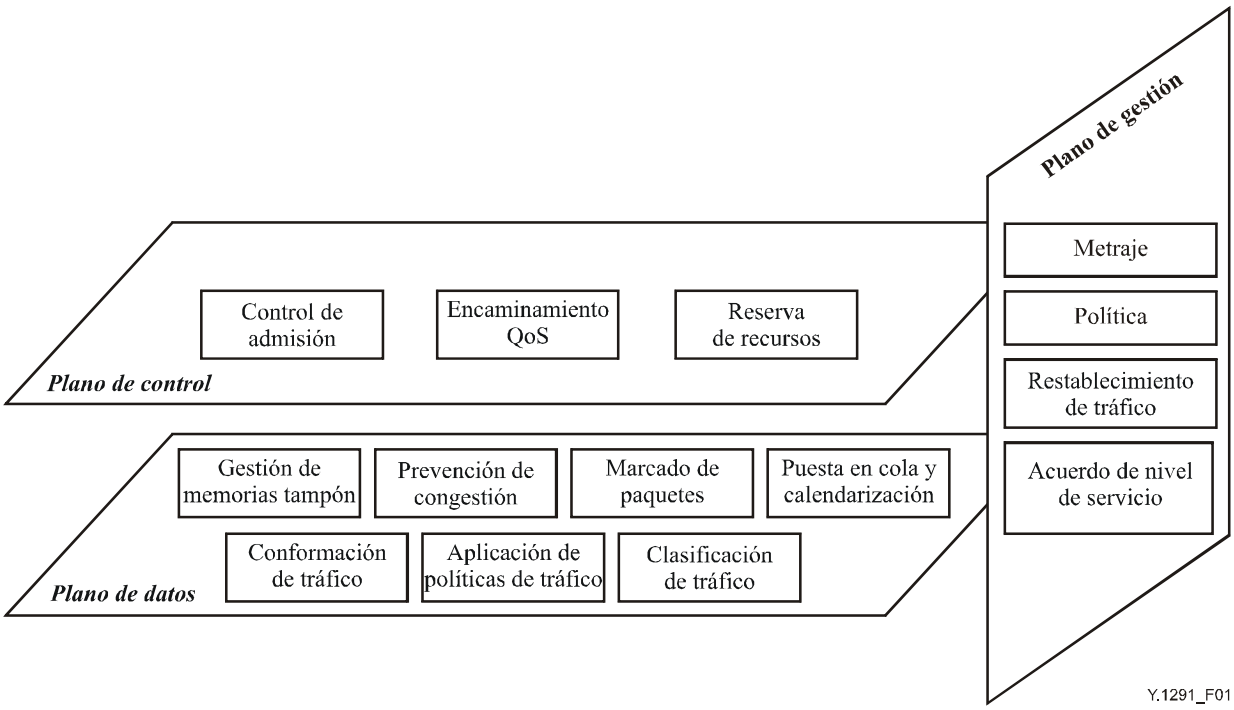

Figura 1/Y.1291 – Marco arquitectural para soporte de QoS

Un bloque de construcción QoS puede ser especifico de un nodo de red (como en el caso de la gestión de una memoria tampón) o aplicable a un segmento de red (como en el caso del encaminamiento QoS). Este último, en particular, requiere señalización entre nodos de red, independientemente de que formen parte de un segmento de red que sea de extremo a extremo, de extremo a borde, de borde a borde, o de red a red. Puede haber señalización en cualquiera de los tres planos lógicos. Cuando la señalización se produce en el plano de control o en el de gestión, entraña la utilización de un protocolo de señalización. Debido a sus propiedades únicas, esta Recomendación trata la señalización como parte de las interacciones entre bloques de construcción QoS y la examina en la sección correspondiente.

Es importante observar que el marco arquitectural QoS es un marco lógico y no condiciona la realización de un bloque de construcción. Por tanto, un bloque de construcción puede implementarse en forma distribuida o centralizada. En las siguientes cláusulas se describen con más detalle los bloques de construcción de cada uno de los planos.

7 Mecanismos del plano de control

7.1 Control de admisión

Este mecanismo controla el tráfico que habrá de admitirse en la red. Normalmente, los criterios de admisión se basan en la aplicación de una determinada política [IETF RFC 2753]. El hecho de que un tráfico sea o no admitido depende de un acuerdo de nivel de servicio preestablecido. Además, la decisión puede depender de que exista en la red una cantidad suficiente de recursos adecuados disponibles, de modo que el nuevo tráfico admitido no sobrecargue la red y degrade el servicio que se presta al tráfico existente. Desde el punto de vista del proveedor de servicio, se debe admitir el mayor volumen de tráfico posible siempre que se mantenga el mismo nivel de QoS (incluidas la calidad de funcionamiento en cuanto a las transacciones y las expectativas) en cuanto a la fiabilidad y la disponibilidad del servicio.

Los métodos de admisión de llamada relacionados con la calidad de funcionamiento con respecto a las transacciones suelen basarse en parámetros o mediciones. Los métodos basados en parámetros obtienen los límites de caso más desfavorable para un conjunto de medidas (por ejemplo, pérdida de paquetes, retardo y fluctuación de fase) a partir de parámetros de tráfico y son apropiados para proporcionar QoS *rígida* para servicios en tiempo real. Este método se utiliza típicamente tras una petición de reserva de recurso con el fin de securizar un recurso necesario para un consiguiente flujo de tráfico. En el apéndice I se presenta un ejemplo de método QoS en el que se utiliza este tipo de control de admisión.

En cambio, con el método basado en mediciones, para tomar una decisión de admisión se realizan mediciones del tráfico existente. Este método no garantiza un caudal, ni límites rígidos para la pérdida de paquetes, el retardo y la fluctuación de fase, y es adecuado para proporcionar una QoS *flexible* o relativa. Por lo general, este método requiere una mayor utilización de recursos que el basado en parámetros. En el apéndice II se presenta en forma resumida un método QoS experimental basado en mediciones. Obsérvese que, en principio, se puede utilizar un método mixto en que se utilicen mediciones para actualizar los recursos disponibles según el método basado en parámetros.

El control de admisión también puede utilizarse para satisfacer las exigencias de fiabilidad/disponibilidad del servicio en un periodo especificado para los tipos de transacción deseados negociados en el SLA. Concretamente, la fiabilidad/disponibilidad del servicio deseada puede solicitarse como un nivel de prioridad para el control de admisión que, a su vez, determina el establecimiento de una "conexión" o enlace como por ejemplo un LSP. Las políticas de control de admisión dan preferencia a flujos continuos de tráfico (por ejemplo, para comunicaciones de urgencia) que un proveedor de servicio considera más críticas en condiciones de congestión. La prioridad para el control de admisión es una forma de dar preferencia a la admisión de un LSP de alto nivel prioridad con respecto a los LSP de niveles de prioridad inferiores.

En el anexo A se especifican con más detalle los niveles de prioridad para el control de admisión.

7.2 Encaminamiento QoS

En una definición estricta, por encaminamiento QoS ha de entenderse la selección de un trayecto que satisface los requisitos de QoS de un flujo. Lo más probable es que el trayecto seleccionado no sea el tradicional trayecto más corto. En función de la especificidad y de la cantidad de medidas de QoS que intervengan, el trabajo de cálculo requerido para la selección del trayecto puede llegar a alcanzar niveles prohibitivos según aumenta el tamaño de la red. De aquí que en los esquemas de encaminamiento QoS utilizados en la práctica generalmente se consideran casos en que se tiene en cuenta una sola medida de QoS (por ejemplo, el ancho de banda o el retardo) o dos medidas de QoS (por ejemplo, costo-retardo, costo-ancho de banda, ancho de banda-retardo)[1]. Para reducir aún más la complejidad del cálculo del trayecto pueden utilizarse diversas estrategias de encaminamiento existentes. Según el procedimiento empleado para mantener la información de estado y para realizar la búsqueda de trayectos factibles, existen estrategias tales como encaminamiento atendiendo al origen, encaminamiento distribuido, y encaminamiento jerárquico [Chen]. Además, según la forma en que se traten múltiples medidas de QoS, existen estrategias como las de ordenación de medidas y filtrado secuencial en las que pueden hacerse negociaciones entre una calidad óptima global y una reducción de la complejidad del cálculo [IETF RFC 2386].

El proceso de selección de trayecto implica el conocimiento de los requisitos y características QoS del flujo, así como de información (que cambia frecuentemente) sobre la disponibilidad de los recursos de la red (expresados en términos de medidas estándar como el ancho de banda disponible y el retardo). El conocimiento suele obtenerse y distribuirse con el auxilio de protocolos de señalización. Por ejemplo, RSVP [IETF RFC 2205] puede utilizarse para transportar los requisitos y características de un flujo, y extensiones de OSPF definidas en IETF RFC 2676 para disponibilidad de recursos. En comparación con el encaminamiento por el trayecto más corto, que selecciona rutas óptimas basándose en una medida relativamente constante (es decir, la cuenta de saltos, o el costo), el encaminamiento QoS suele entrañar cálculos de trayecto más frecuentes y complejos y más tráfico de señalización [Apostolopoulos].

Es importante señalar que el encaminamiento QoS sólo proporciona un medio para determinar un trayecto que probablemente puede responder a la calidad de funcionamiento solicitada. Para garantizar una calidad de funcionamiento en un trayecto seleccionado hay que utilizar el encaminamiento QoS conjuntamente con reserva de recursos para reservar los recursos de red necesarios a lo largo del trayecto.

El encaminamiento QoS puede también generalizarse para aplicarlo a la ingeniería de tráfico. (Cuando los patrones de tráfico cambian lentamente en un largo periodo de tiempo y los flujos de tráfico presentan una granularidad bruta, la ingeniería de tráfico comprende gestión del tráfico, gestión de la capacidad, medición y modelado del tráfico, modelado de la red, y análisis de la calidad de funcionamiento.) Con esta finalidad, la selección de encaminamiento tiene en cuenta una diversidad de constricciones tales como atributos de tráfico, constricciones de red, y constricciones de política [IETF RFC 3272]. Tal encaminamiento QoS generalizado se denomina también encaminamiento basado en constricción, que puede permitir que la selección de trayecto contornee puntos congestionados (o comparta carga) y mejorar la utilización de la red global, así como automatizar la aplicación de las políticas de tráfico.

La serie de Recomendaciones UIT-T E.360.x describe, analiza y recomienda métodos para controlar la respuesta de la red a demandas de tráfico y otros estímulos, como por ejemplo fallos de enlaces o de nodos. Específicamente, los métodos tratados en la serie de Recomendaciones E.360.x incluyen encaminamiento de llamada y de conexión, gestión de recursos QoS, gestión de tablas de encaminamiento, encaminamiento de transporte dinámico, gestión de capacidad, y requisitos de

---

[1] Obsérvese que algunas de estas medidas son aditivas y otras son limitativas. Por ejemplo, retardo y costo son aditivos, ancho de banda es limitativo. Estas consideraciones son importantes a la hora de diseñar algoritmos de encaminamiento implementables.

funcionamiento. La Rec. UIT-T E.361 especifica también funciones de encaminamiento QoS y parámetros asociados, como atribución y protección del ancho de banda, prioridad de encaminamiento, prioridad de puesta en cola, e identificación de la clase de servicio. Además, la Rec. UIT-T E.361 prescribe medios para señalizar parámetros de encaminamiento QoS a través de redes que emplean diferentes tecnologías de encaminamiento.

7.3 Reserva de recursos

Este mecanismo reserva, a petición, recursos de red requeridos para la prestación de la calidad de funcionamiento de red deseada. El que se acepte una petición de reserva de recursos está estrechamente ligado al control de admisión. Son por tanto aplicables todas las consideraciones relativas al control de admisión. Sin embargo, por lo general, una condición necesaria para que se acepte una petición de reserva de recursos es que la red tenga recursos suficientes.

La naturaleza precisa de una reserva de recursos depende de los requisitos de calidad de funcionamiento de la red y del método de red concreto que se utilice para cumplirlos. Por ejemplo, en el método *IntServ*, lo importante son los flujos símplex y se caracterizan en términos de parámetros que describen un cubo de testigos, y las reservas iniciadas por el receptor se efectúan a petición, de acuerdo con requisitos de velocidad de cresta para garantizar que se respetarán los límites de retardo. Abstracción hecha de los aspectos específicos, para los proveedores de servicio es importante poder tarificar por el uso de recursos reservados. Por tanto, la reserva de recursos necesita el soporte de autenticación, autorización, y contabilización y liquidación entre diferentes proveedores de servicio. La reserva de recursos se realiza típicamente mediante un protocolo diseñado para una determinada finalidad, como RSVP [IETF RFC 2205].

La reserva de recursos puede considerarse como una funcionalidad distribuida o centralizada. La diferencia entre la disponibilidad de un recurso en un momento dado y la predicha es una cuestión de suma importancia y, a la hora de poner el nodo, el enlace y otros recursos a disposición de la aplicación solicitante, debe tenerse el cuidado de utilizar la información más reciente.

8 Mecanismos del plano de datos

8.1 Gestión de las colas (o memorias tampón)

La gestión de las colas o memorias tampón trata de resolver la situación siguiente: cuando hay varios paquetes en espera de transmisión, qué paquetes se almacenan y qué paquetes se eliminan (se descartan o desechan). Un importante objetivo de la gestión de colas es minimizar el tamaño de la cola en estado estacionario, sin que por ello resulte subutilizado el enlace, y evitar el fenómeno de bloqueo cuando un solo flujo o conexión ocupa íntegramente el espacio de la cola [IETF RFC 2309]. Los esquemas de gestión de colas se diferencian principalmente en los criterios seguidos para la eliminación de paquetes y en cuanto a la determinación de los paquetes que se eliminan. La utilización de múltiples colas introduce una variación adicional en los esquemas, por ejemplo, en cuanto a la manera en que los paquetes se distribuyen entre las colas.

Un criterio que se sigue usualmente para la eliminación de paquetes es que la cola haya alcanzado su tamaño máximo. Se eliminan paquetes cuando la cola está llena. Qué paquetes habrán de eliminarse depende de las modalidades de la eliminación que se utilicen, por ejemplo:

- La "eliminación al final" rechaza el último paquete que llega en ese momento. Esta es la estrategia más usual.
- La "eliminación al principio" conserva el paquete que llega en ese momento, a expensas del paquete situado al principio de la cola, que es eliminado.
- La "eliminación aleatoria" conserva el paquete que llega en ese momento a expensas de un paquete que se selecciona al azar entre los que están en la cola. Este esquema puede ser costoso pues exige un recorrido a través de la cola.

Un esquema de eliminación de paquetes según el cual sólo se eliminan paquetes cuando la cola está llena tiende a mantener la cola en el estado lleno durante un periodo de tiempo relativamente largo, lo que puede tener consecuencias catastróficas en caso de tráfico en ráfagas. Hay esquemas que utilizan un criterio más dinámico que no se basa en el tamaño máximo fijo de la cola, por lo que se puede efectuar una gestión activa de la cola. Un método conocido es el denominado Pronta detección aleatoria (RED, *random early detection*) [Floyd], que también ayuda a tratar el problema de la cola llena y evitar la congestión. RED elimina los paquetes (entrantes) por un procedimiento probabilístico, basándose en un tamaño de cola promedio estimado. La probabilidad de eliminación aumenta a medida que aumenta el tamaño de cola promedio estimado. Dicho sea de otra forma, si la cola ha estado mayormente vacía en el pasado reciente, se tiende a conservar los paquetes entrantes; en cambio, si la cola ha estado relativamente llena recientemente, los paquetes entrantes probablemente serán eliminados. Específicamente, RED utiliza dos niveles para el tamaño de cola promedio. Uno especifica el tamaño de cola promedio por debajo del cual no se eliminan paquetes; el otro especifica el tamaño de cola promedio por encima del cual se eliminan todos los paquetes. Para una cola de un tamaño promedio entre los dos umbrales, la probabilidad de eliminación del paquete es proporcional al tamaño promedio. Naturalmente, la eficacia de RED depende de la forma en que se dan valores a los parámetros. No hay un conjunto único de parámetros que funcione bien para todos los tipos de tráfico y escenarios de congestión. Por tanto, han surgido variantes de RED, por ejemplo:

- RED de flujo (FRED, *flow RED*) [Lin *y otros*, 1997], que introduce control adicional en RED al proporcionar un tratamiento de eliminación diferencial de flujos en base a su utilización de la memoria tampón. Si la cantidad de paquetes que pasan del flujo a la cola es menor que un umbral específico del flujo, un paquete que llega en ese momento en ese mismo flujo no será eliminado. En otro caso, se somete a un tratamiento de eliminación en el que se favorecen los flujos en que haya menos paquetes en la memoria tampón. En comparación con RED, FRED es más flexible en cuanto a la protección de los flujos de modo que no utilicen una parte menor o mayor que cierta proporción justa del espacio de la memoria tampón y del ancho de banda del enlace.
- RED ponderado, que introduce control adicional en RED al proporcionar un tratamiento de eliminación diferencial de flujos en base a su nivel de prioridad. Cuanto más alto es el nivel de prioridad de un paquete, tanto menor es la probabilidad de que sea eliminado.

8.2 Prevención de la congestión

Se produce congestión en una red cuando el tráfico rebasa o se aproxima al volumen de tráfico que la red puede manejar, y esto se debe a la falta de recursos tales como el ancho de banda del enlace y el espacio de memoria tampón. Un signo de congestión es, por ejemplo, que las colas del encaminador (o conmutador) están siempre llenas y los encaminadores comienzan a eliminar paquetes. La eliminación de paquetes ocasiona retransmisiones, como consecuencia de las cuales aumenta el tráfico y se agrava la congestión. La reacción en cadena puede conducir que la red deje de funcionar y que su caudal sea nulo. En esta situación, intuitivamente se piensa en utilizar memorias tampón de gran tamaño para evitar la congestión debida a la falta de espacio de memoria tampón. Nagle [1987] demostró que no sucede esto, sino lo contrario. Los largos tiempos de espera de los paquetes debido a los grandes tamaños de las memorias tampón provocan la retransmisión de paquetes, que provoca la congestión. La prevención de la congestión trata de proporcionar medios más eficaces para mantener la carga de la red por debajo de su capacidad, de manera que pueda funcionar con un rendimiento aceptable y no sufra los efectos de la congestión.

Un esquema típico de prevención de la congestión se pone en marcha cuando el emisor reduce el volumen de tráfico que entra en la red al recibir una indicación de que se está produciendo (o está próxima a producirse) una congestión de la red [Jacobson, 1988]. A menos que haya una indicación explícita, la pérdida de paquetes o la expiración de un temporizador normalmente se considera una indicación implícita de congestión de la red. La forma en que la fuente del tráfico reacciona para

reducir su volumen depende de la especificidad de los protocolos de transporte. En un protocolo basado en ventana como TCP, esto se realiza disminuyendo por un factor multiplicativo el tamaño de la ventana.

En una situación ideal, en el origen de una reducción del tráfico está un cliente cuyo nivel de prioridad para el control de admisión no es crítico. Esto permite que un tráfico con nivel de prioridad más alto continúe recibiendo el servicio normal.

Cuando la congestión se atenúa, el emisor aumenta cuidadosamente el tráfico.

Para evitar los posibles retardos excesivos debidos a retransmisiones tras las pérdidas de paquetes se han creado recientemente esquemas de notificación explícita de congestión (ECN, *explicit congestion notification*). IETF RFC 3168 especifica un esquema ECN para IP y TCP, entre otros esquemas de gestión activa de las memorias tampón. En este esquema, una congestión de red incipiente se indica marcando paquetes, sin eliminarlos. Al recibir un paquete que ha sufrido los efectos de la congestión, un anfitrión capaz de funcionar con el esquema ECN responde esencialmente como si se hubiera eliminado el paquete.

8.3 Puesta en cola y calendarización

En resumen, este mecanismo controla qué paquetes habrán de seleccionarse para transmisión por un enlace de salida. El tráfico entrante se retiene en un sistema de puesta en cola que, típicamente, comprende varias colas y un calendarizador. Para gobernar dicho sistema se utiliza una disciplina de puesta en cola y calendarización. Hay tres métodos esenciales:

- Puesta en cola según el principio de primero en entrar, primero en salir: Los paquetes se introducen en una cola única y son tratados en el mismo orden en que fueron introducidos en la cola.
- Puesta en cola sobre una base justa: Los paquetes se clasifican en flujos y se asignan a colas dedicadas a flujos respectivos. Las colas son entonces atendidas según el método de "ciclo completo". Las colas vacías se saltan. La puesta en cola sobre una base justa se conoce también por puesta en cola flujo por flujo, o puesta en cola basada en flujos.
- Puesta en cola por nivel de prioridad: Los paquetes se clasifican primero, y después se introducen en colas con diferentes niveles de prioridad. Los paquetes son calendarizados a partir de la cabecera de una cola dada, solamente si todas las colas de nivel de prioridad más alto están vacías. Dentro de cada una de esas colas con diferentes niveles de prioridad, los paquetes son calendarizados según el principio de primero en entrar, primero en salir.
- Puesta en cola ponderada sobre una base justa: Los paquetes se clasifican en flujos y se asignan a colas dedicadas a flujos respectivos. A una cola se asigna un porcentaje del ancho de banda de salida de acuerdo con el ancho de banda que necesita el flujo correspondiente. Al distinguir entre paquetes de longitud variable, este método también impide que a flujos de paquetes más grandes se les atribuya más ancho de banda que a los flujos de paquetes más pequeños.
- Puesta en cola basada en clases. Los paquetes se clasifican en diversas clases de servicio y después se asignan a colas que a su vez han sido asignadas a las clases de servicio, respectivamente. A cada cola se puede asignar un porcentaje diferente del ancho de banda de salida y se atiende según el método de "ciclo completo". Las colas vacías se saltan.

8.4 Marcado de paquetes

Los paquetes pueden marcarse de acuerdo con las clases de servicio específicas que recibirán en la red, paquete por paquete. El marcado de paquetes lo realiza típicamente un nodo de borde e implica la asignación de un valor a un determinado campo del encabezamiento de un paquete, en una forma normalizada. (Por ejemplo, el tipo de servicio en el encabezamiento IP o los bits EXP del encabezamiento cuña MPLS [IETF RFC 3032] se utilizan para codificar comportamientos

observables externamente de encaminadores en el método *DiffServ* [IETF RFC 2474] o *MPLS-DiffServ* [IETF RFC 3270].) Si la marca la efectúa un anfitrión, deberá ser comprobada y, si es necesario, cambiada por un nodo de borde. Algunas veces pueden utilizarse valores especiales para marcar paquetes no conformes, que podrán ser eliminados posteriormente debido a congestión. También es posible promover o retrotraer paquetes en base a resultados de mediciones.

Tanto si el marcado de paquetes lo ha efectuado un anfitrión o un nodo de borde, los criterios para ello tienen que estar aprovisionados o configurados dinámicamente. En el caso de configuración dinámica puede utilizarse el protocolo común abierto de servicio de políticas (IETF RFC 2748) o el RSVP. En el caso del RSVP, la entidad que efectúa el marcado puede utilizarlo para interrogar la red sobre el marcado que habrá de aplicarse a paquetes pertenecientes a un determinado flujo [IETF RFC 2996].

8.5 Clasificación de tráfico

La clasificación de tráfico puede hacerse a nivel de flujo o a nivel de paquete. En el borde de la red, la entidad responsable de la clasificación de tráfico típicamente examina multicampos (como las cinco tuplas asociadas con un flujo IP) de un paquete y determina el agregado a que pertenece el paquete y el respectivo acuerdo de nivel de servicio.

8.6 Aplicación de políticas de tráfico

La aplicación de políticas de tráfico tiene por objeto determinar si el tráfico que se está presentando es salto por salto y observa las políticas o contratos prenegociados. Los paquetes no conformes suelen eliminarse. Se podrá notificar a los emisores los paquetes que se hayan eliminado y determinar las causas determinadas, así como asegurar por SLA la futura observancia.

8.7 Conformación de tráfico

La conformación de tráfico tiene por objeto controlar la velocidad y el volumen del tráfico que entra en la red. La entidad responsable de la conformación de tráfico almacena en memoria tampón los paquetes no conformes hasta que logre que el agregado respectivo sea conforme con el tráfico. El tráfico así obtenido no está tan caracterizado por ráfagas como el original, y es más predecible. A menudo es necesario conformar el tráfico cursado entre nodos de egreso y de ingreso.

Para efectuar la conformación de tráfico existen dos métodos fundamentales: el método del cubo no estanco y el método del cubo de testigos. El método del cubo no estanco emplea una función conocida por este nombre para regular la velocidad del tráfico que sale de un nodo. Cualquiera que sea la velocidad del flujo entrante, el cubo no estanco mantiene el flujo saliente a una velocidad constante. Todo paquete en exceso que desborde el cubo se descarta. Este método se caracteriza por dos parámetros que generalmente son configurables por el usuario: el tamaño del cubo no estanco y la velocidad de transmisión.

El método del cubo de testigos, por otro lado, no es tan rígido en la regulación de la velocidad del tráfico que sale del nodo. Permite que los paquetes salgan tan rápidamente como entran, siempre que haya *testigos* suficientes. Los testigos se generan a cierta velocidad y se depositan en el cubo de testigos hasta que éste se llena. El consumo de un testigo permite que cierto volumen de tráfico (es decir, cierta cantidad de octetos) salga del nodo. No pueden transmitirse paquetes si no hay testigos en el cubo. Sin embargo, pueden consumirse varios testigos de una sola vez para permitir el paso de ráfagas. Este método, a diferencia del método del cubo no estanco, no tiene una política de descarte. Deja que sea la gestión de la memoria tampón quien se encargue de los paquetes si el cubo se llena. Este método se caracteriza por dos parámetros que generalmente son configurables por el usuario: el tamaño del cubo de testigos y la velocidad de generación de testigos.

Los métodos del cubo no estanco y del cubo de testigos pueden utilizarse conjuntamente. En particular, el tráfico puede conformarse primeramente mediante el método del cubo de testigos, y después mediante el método del cubo no estanco para suprimir las ráfagas no deseadas. También es posible utilizar dos cubos de testigos dispuestos en cascada.

9 Mecanismos del plano de gestión

9.1 Acuerdo de nivel de servicio

Un acuerdo de nivel de servicio (SLA) representa típicamente el acuerdo entre un cliente y un proveedor de un servicio que especifica el nivel de disponibilidad, posibilidad de ser servido, calidad de funcionamiento, operación u otros atributos del servicio. Puede incluir aspectos como los de fijación de precios que son de naturaleza comercial. La parte técnica del acuerdo se denomina especificación de nivel de servicio (SLS, *service level specification*) [IETF RFC 3198], e incluye específicamente un conjunto de parámetros y sus valores que, juntos, definen el servicio ofrecido por una red al trafico de un cliente. Los parámetros SLS pueden ser genéricos, como los definidos en la Rec. UIT-T Y.1540 o específicos de la tecnología como la calidad de funcionamiento y los parámetros de tráfico utilizados en *IntServ* o *DiffServ*. En general, la Rec. UIT-T E.860 define una red SLA general para un entorno multivendedor.

9.2 Metraje y registro de tráfico

El metraje tiene por objeto supervisar las propiedades temporales (por ejemplo, la velocidad) de un flujo continuo de tráfico cotejándolo con el perfil de tráfico convenido. Implica la observación de características de tráfico en un punto dado de la red y la toma y almacenamiento de la información de tráfico con miras a su análisis y acciones consiguientes. En función del nivel de conformidad, un dispositivo de metraje puede disponer que se aplique a un tren de paquetes un tratamiento necesario (por ejemplo, eliminación o conformación).

9.3 Restablecimiento de tráfico

El restablecimiento se define aquí en su sentido lato como la respuesta mitigante de una red en condiciones de fallo y debe considerarse en múltiples capas. En la parte inferior de la pila estructurada en capas, unas redes ópticas son ahora capaces de proporcionar una funcionalidad de protección dinámica de anillos y mallas y una funcionalidad de restablecimiento a nivel de longitud de onda. En la capa de SONET/SDH, la disponibilidad se proporciona mediante conmutación de protección automática (APS, *automatic protection switching*) así como arquitecturas de anillos y mallas con restablecimiento automático. ATM proporciona capacidades similares. El reencaminamiento se utiliza tradicionalmente en la capa IP para restablecer el servicio tras fallos de los enlaces y nodos, y puede ser de extremo a extremo o local (reencaminamiento rápido). El reencaminamiento en la capa IP se produce después de un periodo de convergencia del encaminamiento, que puede durar un lapso de segundos a minutos. MPLS proporciona ahora recuperación en la capa IP antes de la convergencia.

Hay dos tipos de fallos de red:

- Fallo de nodo: Fallo de un elemento de red (por ejemplo, tarjeta de encaminador) en un nodo de red u oficina. Este tipo de fallo suele tratarse previendo en el diseño características de redundancia en elementos de red para reducir al mínimo los efectos del fallo. Los fallos catastróficos como las interrupciones del suministro de energía y desastres naturales pueden, sin embargo, paralizar un nodo de red completo, en cuyo caso el tráfico de extremo a extremo puede reencaminarse por enlaces de reserva diseñados alrededor del nodo fallido.

- Fallo de enlace de transporte: Fallo de un enlace (por ejemplo, T1, OC-3) que conecta dos nodos de red. Los fallos de enlace suelen producirse como consecuencia de un fallo de un elemento del enlace (por ejemplo, de una tarjeta de línea) (que puede causar la interrupción de un solo enlace) o, lo que es más grave, la ruptura de una fibra (que puede perturbar un gran número de enlaces). Los proveedores de servicio pueden prever en sus diseños una capacidad de reserva adicional para mitigar el efecto de tales fallos y restablecer los flujos de tráfico hasta que se elimine el fallo.

Obsérvese que algunos de estos términos son, por lo general, específicos de la capa y se deben estudiar cuidadosamente las múltiples capas que entran en juego en el diseño global. Por ejemplo, un fallo de enlace en la capa física podría afectar a muchos enlaces y trayectos en la capa IP.

Al igual que en el caso del control de admisión, ciertos flujos continuos de tráfico relacionados con servicios críticos pueden requerir niveles de restablecimiento más altos que otros. Un proveedor de servicio debe planificar niveles adecuados de recursos de reserva de manera que los SLA de QoS se cumplan en condiciones de restablecimiento. Son parámetros típicos para la medición de las posibilidades de restablecimiento del servicio el tiempo hasta el restablecimiento y el porcentaje de restablecimiento del servicio. Los detalles de los niveles de prioridad figuran en el anexo A.

9.4 Políticas

Las políticas son conjuntos de reglas que suelen establecerse para administrar, gestionar y controlar el acceso a recursos de red. Pueden ser específicas de las necesidades del proveedor de servicio o reflejar el acuerdo entre el cliente y el proveedor de servicio, lo que puede incluir requisitos de fiabilidad y disponibilidad durante un periodo de tiempo, y otros requisitos de QoS. Los proveedores de servicio pueden implementar mecanismos en el plano de control y en al plano de datos basándose en políticas. Algunas posibles aplicaciones son: encaminamiento de políticas (direccionamiento de un flujo de paquetes hacia un puerto de destino sin una tabla de encaminamiento), políticas de filtrado de paquetes (marcado y eliminación de paquetes en base a una política de clasificador), registro cronológico de paquetes (que permite a los usuarios registrar cronológicamente flujos de paquetes especificados) y políticas relacionadas con la seguridad.

Diversos eventos pueden provocar decisiones de política. Algunos están relacionados con el tráfico y otros no. Los detalles dependen generalmente de la especificidad de cada aplicación. IETF RFC 2748, por ejemplo, especifica un protocolo simple de indagación y respuesta que puede utilizarse para intercambiar información de política entre un servidor de políticas (o punto de decisión de política) y su cliente (o punto de aplicación de política).

10 Interacciones entre bloques de construcción

En una solución QoS de carácter general suelen utilizarse múltiples bloques de construcción a través del plano de control, del plano de datos y del plano de gestión. Por tanto, es necesario intercambiar parámetros entre los diversos bloques de construcción. Estos parámetros incluyen la calidad de funcionamiento desde el punto de vista de las transacciones en el nivel de paquetes (por ejemplo, retardo y pérdida de paquetes) y expectativas en cuanto a la fiabilidad/disponibilidad del servicio en forma de niveles de prioridad de tráfico para determinadas funciones de red como el control de admisión y restablecimiento de tráfico. Son ejemplos de mecanismos para transportar los valores de estos parámetros la señalización y la consulta de bases de datos.

10.1 Señalización QoS

La señalización QoS se utiliza principalmente para transportar requisitos de calidad de funcionamiento de la aplicación (o de la red), reservar recursos de red a través de la red, o descubrir rutas QoS. Según que la información de señalización forme o no parte del tráfico de datos asociado, la señalización QoS puede efectuarse dentro o fuera de banda:

– Dentro de banda: La señal QoS forma parte del tráfico de datos asociado, que suele presentarse en un determinado campo del encabezamiento (por ejemplo, el campo TOS en IPv4, como en *DiffServ* y 802.1p) de los paquetes de datos. La señalización dentro de banda se efectúa en el plano de datos, por lo que no introduce tráfico adicional en la red, ni entraña retardos de establecimiento para el tráfico de datos. Naturalmente, tal tipo de señalización no es adecuado para la reserva de recursos, ni para el encaminamiento QoS, que deben realizarse previamente, antes de la transmisión de datos.

– Fuera de banda: La señal QoS, que es transportada por paquetes especializados, está separada del tráfico de datos asociado. Además, la señalización QoS puede realizarse salto por salto o de extremo a extremo. En el caso de la señalización salto por salto (que se representa como caso B en la figura 2), la información de señalización probablemente sea modificada por nodos intermedios. En cambio, en el caso de la señalización de extremo a extremo (representado como caso A en la figura 2), la información de señalización no es modificada por nodos intermedios. Por consiguiente, la señalización fuera de banda introduce tráfico suplementario en la red y se traduce en una tara a los efectos de la calidad de funcionamiento deseada de la red, que se proporciona. Por otro lado, entraña la utilización de un protocolo de señalización y ulterior procesamiento por encima de la capa de red, lo que tiende a que las respuestas no sean tan rápidas como en el caso de la señalización dentro de banda. No obstante, la señalización fuera de banda, por naturaleza, se presta a ser utilizada para reserva de recursos o encaminamiento QoS.

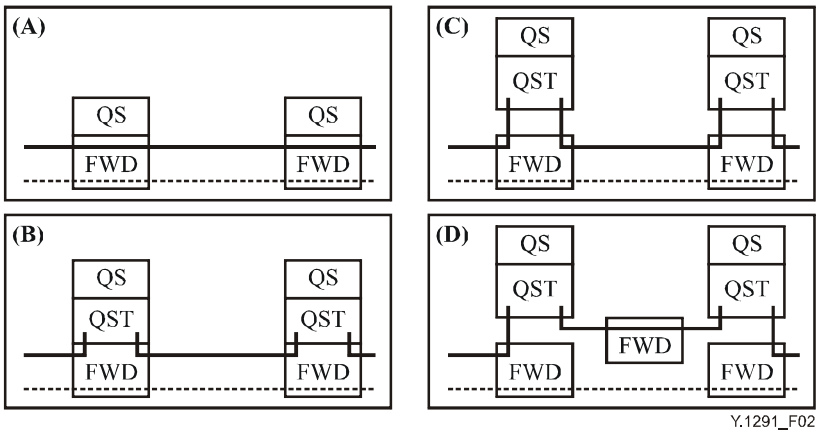

Figura 2/Y.1291 – Ilustración de diferentes formas de señalización QoS

De manera similar, según que el trayecto de señalización esté o no estrechamente ligado al trayecto de datos asociado, la señalización QoS puede considerarse como acoplada al trayecto o desacoplada del trayecto:

– Acoplada al trayecto: Los mensajes de señalización QoS se encaminan solamente a través de los nodos que pueden existir en el trayecto de datos. Por definición, la señalización dentro de banda está acoplada al trayecto; en cambio, la señalización fuera de banda puede o no estarlo. La señalización acoplada al trayecto implica que los nodos de señalización tienen que estar coubicados con encaminadores. Este montaje tiene, por un lado, la ventaja de un costo reducido del procesamiento de la señalización global (pues apoya las tareas de encaminamiento de la capa de red), pero por otro lado tiene el inconveniente de ser inflexible cuando se trata de elevar el nivel de los encaminadores o integrar entidades de

control (por ejemplo, servidores de políticas) que no están en el trayecto de datos (o métodos de encaminamiento no tradicionales). Esto significa que si un mecanismo acoplado al trayecto implica un protocolo de señalización, los encaminadores tienen que soportar el protocolo y poder procesar mensajes de señalización conexos. Un ejemplo de un protocolo de señalización acoplada al trayecto es RSVP.

− Desacoplada del trayecto: Los mensajes de señalización QoS se encaminan a través de nodos que se supone que no forman parte del trayecto. Como tal, sólo la señalización fuera de banda puede estar desacoplada del trayecto. La señalización desacoplada del trayecto implica que la entidad que termina la señalización QoS debe ser especializada y estar separada de la entidad reenviante, que normalmente está situada en encaminadores. A diferencia de la señalización acoplada al trayecto, tiene la ventaja de la flexibilidad para desplegar y elevar el nivel de los nodos de señalización independientemente de los encaminadores, o integrar entidades de control que no forman parte del trayecto de datos, pero tiene el inconveniente de que introduce complejidad y aumenta el costo de las tareas globales de procesamiento y explotación. En los casos C y D de la figura 2 se ilustra con más detalle la señalización desacoplada del trayecto.

10.2 Señalización intraplano

Queda en estudio.

10.3 Señalización interplanos

Plano de control y plano de datos

Correspondencia de la clase de QoS Y.1541 con DSCP

En el apéndice VI/Y.1541 se presenta una asociación de las clases de QoS de dicha Recomendación con Diffserv por cada comportamiento de dominio (PDB, *per domain behaviors*):

- PDB basado en PHB con reenvío acelerado: clases 0 y 1 de Y.1541.
- PDB basado en PHB con reenvío asegurado: clases 2, 3 y 4 de Y.1541.
- PDB basado en PHB de mejor esfuerzo (por defecto): clase 5 de Y.1541.

11 Consideraciones relativas a la seguridad

En general, la Rec. UIT-T X.805 proporciona una arquitectura de seguridad de red que ayuda a examinar las propiedades de seguridad de bloques de construcción QoS y de soluciones QoS, así como a diseñar salvaguardas para los mismos.

11.1 Plano de datos

En el plano de datos el tráfico se procesa típicamente de acuerdo con la información contenida en el encabezamiento del paquete. Los paquetes pueden marcarse asignando un valor a cierto campo del encabezamiento, o clasificarse en base a la información contenida en varios campos del encabezamiento (por ejemplo, cinco tuplas IP). La conformación de tráfico, aplicación de políticas y puesta en cola pueden entonces realizarse en base a la clasificación y al marcado de los paquetes. La integridad de los encabezamientos de paquetes, como tal, es esencial para la validez y seguridad de un método QoS. Hay que evitar la creación, modificación y manipulación maliciosas de la información contenida en los encabezamientos de los paquetes.

También es importante señalar que aunque un anfitrión o cualquiera de los nodos de red pueden hacer o cambiar una marca en un paquete, es conveniente que sea un nodo de borde el que lo haga. Por lo general, un nodo de borde tiene una relación de confianza con nodos medulares. Por tanto, una marca hecha por un anfitrión debe ser comprobada y puede ser modificada por un nodo de borde.

11.2 Plano de control y plano de gestión

El plano de control y el plano de gestión se encargan del tráfico a nivel de flujo o de agregado. Asimismo, un flujo se identifica y describe, por ejemplo, mediante cinco tuplas IP o una etiqueta MPLS en el encabezamiento del paquete, la cual es constante durante el ciclo de vida del flujo.

El control de admisión realizado en los nodos de borde ayuda a prevenir los ataques por impostura y la congestión resultante del tráfico no autorizado. Los nodos de borde pueden ser puestos en una relación de confianza por los nodos medulares y pueden tener una visión de la utilización de los recursos de la red global. Tanto si se realiza en forma centralizada o distribuida, el control de admisión debe comprender autenticación y autorización.

La reserva de recursos está estrechamente ligada al control de admisión. Una petición de reserva de recursos puede ser iniciada por un anfitrión de extremo o por un nodo que soporta el servicio, situado en la red. Las peticiones de reserva maliciosas pueden tener por consecuencia una cantidad excesiva de reservas ilegales, la extinción de los recursos y la denegación del servicio. Es conveniente tener salvaguardas que prevengan esas peticiones maliciosas.

En general, mecanismos de seguridad de la red como cortafuegos y dispositivos de detección de intrusión pueden ayudar a proteger las interfaces de red, tanto cuando interviene como cuando no interviene la QoS. También las entidades encargadas de la autenticación deben tener salvaguardas contra los ataques por denegación del servicio.

11.3 Señalización QoS

Para la protección contra ataques por intercepción, modificación y fabricación, la señalización QoS debe utilizar mecanismos de autenticación y de integridad, como RIPEMD160 o SHA-1 (algoritmo de troceado securizado 1). La utilización de mecanismos de seguridad puede influir en la calidad de funcionamiento. Puesto que, normalmente, el tráfico de señalización es mucho menor que el tráfico de datos, las repercusiones de la calidad de funcionamiento de la red debidas a la señalización fuera de banda securizada (o desacoplada del trayecto) deben ser menores que las de la señalización dentro de banda securizada (o acoplada al trayecto). Además, las entidades encargadas de la señalización deben tener salvaguardas contra los ataques por denegación del servicio.

12 Ejemplos de métodos

Para ilustrar como los bloques de construcción QoS interactúan según diversos métodos QoS, en esta cláusula se describen cuatro métodos normalizados: servicios integrados (*IntServ*), servicios diferenciados (*DiffServ*), conmutación por etiqueta multiprotocolo (MPLS), y QoS dinámica IPCablecom. (Obsérvese que RFC 2998 integra los métodos IntServ y DiffServ.) Puesto que están surgiendo, y se encuentran en evolución, métodos más amplios, se presentan ejemplos en los apéndices I y II.

12.1 IntServ

El método *IntServ* (véase, por ejemplo, [IETF RFC 1633]), cuyo objetivo principal es el soporte de aplicaciones sensibles al funcionamiento en tiempo real, se ha elaborado partiendo de la noción de que un flujo al que se da servicio a una velocidad algo mayor que su velocidad de datos tiene un retardo sujeto a un límite y que la red puede garantizar el límite de retardo de un flujo en el caso de una reserva de recurso flujo por flujo. Según este método, una aplicación, antes de enviar datos, primero señaliza a la red la petición del servicio deseado, incluyendo sus especificidades, tales como su perfil de tráfico y sus requisitos de ancho de banda y retardo. Después, la red determina si puede asignar recursos adecuados (por ejemplo, ancho de banda o espacio de memoria tampón) para proporcionar la calidad de funcionamiento deseada de la petición de servicio. Sólo después de haber sido aceptada la petición puede la aplicación comenzar a enviar datos. Mientras la aplicación respete su perfil de tráfico, las red cumple su compromiso de servicio manteniendo el estado flujo

por flujo y utilizando disciplinas avanzadas de puesta en cola (por ejemplo, puesta en cola ponderada sobre una base justa) para compartición de enlace. Los bloques de construcción de importancia para el método *IntServ* incluyen control de admisión, puesta en cola, reserva de recurso, clasificación de tráfico y aplicación de políticas de tráfico. En particular, el protocolo de señalización RSVP se utiliza para reservar recursos. La red puede aceptar o rechazar una petición de reserva de recursos mediante control de admisión basado en la disponibilidad de los recursos. Cuando una petición de recursos es aceptada se instalan los estados pertinentes en los nodos sensibles a RSVP. Los bloques de construcción interactúan mediante el acceso a la información de estado y a otros objetos de datos aprovisionados (por lo que son relativamente estáticos).

12.2 DiffServ

La noción en que se funda el método *DiffServ* es la de tratar un paquete en base a su clase de servicio tal como está codificada en su encabezamiento IP. El proveedor de servicio establece con cada usuario un acuerdo de nivel de servicio (o especificación de nivel de servicio), el cual, entre otras cosas, especifica qué cantidad de tráfico un usuario puede enviar dentro de una clase de servicio dada. El tráfico consiguiente se clasifica (paquete por paquete) en un flujo agregado, o en una clase, dentro de un pequeño número, de flujos agregados o de clases, y se somete a la aplicación de políticas en el borde de la red del proveedor de servicio. Una vez que el tráfico ha entrado en la red, los encaminadores le dan un tratamiento diferenciado. En contraste con el método *IntServ*, el tratamiento se no da flujo por flujo, sino únicamente a la clase de servicio indicada. La red global se establece de manera que se cumplan todos los acuerdos de nivel de servicio. Los bloques de construcción pertinentes (que incluyen gestión de memorias tampón, marcado de paquetes, acuerdo de nivel de servicio, metraje y registro de tráfico, aplicación de políticas de tráfico, conformación de tráfico, y calendarización) interactúan unos con otros de una manera relativamente estática, esencialmente mediante objetos de datos aprovisionados.

12.3 MPLS

Inicialmente desarrollado con miras al interfuncionamiento entre las redes IP y ATM (o relevo de trama), MPLS [IETF RFC 3031] obtiene importantes ganancias en la velocidad de reenvío de paquetes mediante el uso de etiquetas cortas, similares a las de la capa 2. Una vez que el paquete ha entrado en la red MPLS, se le asigna de una sola vez una clase de equivalencia para el reenvío (FEC, *forward equivalence class*), que se codifica como una cadena de longitud fija, que se conoce por una etiqueta. Cuando el paquete se reenvía al salto siguiente, se envía la etiqueta conjuntamente. En el salto siguiente, la etiqueta se utiliza como un índice en una tabla preconfigurada para identificar el nuevo salto siguiente y una nueva etiqueta. La antigua etiqueta se sustituye por la nueva y el paquete se reenvía al nuevo salto siguiente. El proceso continúa hasta que el paquete llega a su destino. Dicho sea en otras palabras, el reenvío de paquetes en MPLS está totalmente gobernado por la etiqueta, en virtud de lo cual los paquetes a que se ha asignado la misma FEC son reenviados de la misma forma. Por otra parte, las etiquetas sólo son significativas para el par de encaminadores que comparten un enlace, y sólo en un sentido de transmisión: de un emisor a un receptor. El receptor, no obstante, elige la etiqueta y negocia su semántica con el emisor por medio de un protocolo de distribución de etiquetas. MPLS en su forma básica es particularmente útil en ingeniería de tráfico. Para proporcionar un soporte de QoS explícito, MPLS utiliza ciertos elementos de los métodos *IntServ* y *DiffServ*. El protocolo de distribución de etiquetas, por ejemplo, puede basarse en un protocolo de reserva de recursos [IETF RFC 3209]. Con este protocolo, los recursos de red requeridos, así como un trayecto con conmutación por etiqueta, pueden reservarse de esta forma durante su fase de establecimiento para garantizar la QoS de los paquetes transmitidos por el trayecto. Además, al utilizar la etiqueta y ciertos bits EXP del encabezamiento cuña (*shim header*) que transporta la etiqueta para representar las clases de servicio diferenciadas, paquetes con la misma FEC pueden ser objeto del tratamiento *DiffServ* [IETF RFC 3270]. Los bloques de construcción pertinentes para MPLS incluyen gestión de memorias tampón, marcado de paquetes, encaminamiento QoS, puesta en cola, reserva de recursos, clasificación de tráfico y conformación

de tráfico. Estos bloques interactúan mediante la información de estado del trayecto conmutado por etiqueta, instalada en cada nodo MPLS por un protocolo de distribución de etiquetas y mediante objetos de datos aprovisionados.

12.4 QoS dinámica IPCablecom

Para el soporte de aplicaciones multimedios (multimedia) interactivas a través de una red de acceso IPCablecom, la Rec. UIT-T J.163 especifica un método basado en una reserva de recursos dinámica flujo por flujo. La red de acceso conecta el adaptador de terminal multimedios (MTA, *multimedia terminal adaptor*) al nodo de acceso definido en la Rec. UIT-T J.112. En la red J.112 se atribuyen recursos para cada flujo individual asociado con una sesión de aplicación, por cada abonado, sobre la base de una autorización y autenticación.

En el centro del método de QoS dinámica están las puertas de QoS dinámica (DQoS, *dynamic QoS*) y el controlador de puerta. Mediante el protocolo común abierto de servicio de aplicación de políticas (COPS, *common open policy service protocol*) de acuerdo con RFC 2748, el controlador de puerta controla la existencia y el funcionamiento de las puertas.

Las puertas DQoS se implementan en el nodo de acceso entre la red J.112 y un espinazo IP utilizando las funciones de clasificación y filtrado de paquetes J.112. Una puerta DQoS es por naturaleza unidireccional y es una entidad lógica asociada con una sesión. Si una puerta está "cerrada", los datos que transitan por la red de acceso J.112 pueden ser eliminados o, simplemente, recibir el servicio de tipo mejor esfuerzo, lo que dependerá de la política aplicada por el proveedor.

El controlador de puerta se implementa en el servidor de gestión de llamadas, que normalmente gestiona sesiones multimedia iniciadas por MTA mediante la señalización de llamada controlada por la red (definida en la Rec. UIT-T J.162) o señalización de llamada distribuida (definida en IETF RFC 3261). El controlador se encarga de tomar decisiones en cuanto a la aplicación de políticas sobre si se crea o se abre una puerta. La apertura de una puerta implica control de admisión al recibirse una petición de gestión de recursos (por medio de RSVP) y reserva de los recursos que necesite la red. Debe señalarse que la reserva de recursos se efectúa en dos fases. Al final de la primera fase, los recursos están reservados pero todavía no están disponibles por los MTA. Sólo después de terminada la segunda fase se abren las puertas de los nodos de acceso y los recursos se ponen a disposición de los MTA. El modelo de reserva y compromiso asegura que los recursos están disponibles antes de que señalice a la parte terminadora que se está iniciando una sesión, y que los recursos se comprometen únicamente cuando se piden.

Los bloques de construcción pertinentes para el método DQoS de IPCablecom incluyen principalmente control de admisión, puesta en cola, reserva de recursos, clasificación de tráfico, aplicación de políticas de tráfico y políticas. Los protocolos de señalización RSVP y COPS se utilizan para reservar y comprometer recursos. La red puede aceptar o rechazar una petición de reserva mediante control de admisión basado en la disponibilidad de los recursos o en la aplicación de una política. Cuando se acepta una petición de reserva se instalan los estados pertinentes en los nodos sensibles al protocolo RSVP. Los bloques de construcción interactúan ganando acceso a la información de estado y a otros objetos de datos aprovisionados.

Anexo A

Niveles de prioridad de tráfico

Las expectativas de calidad de servicio (QoS) para servicios prestados por redes de paquetes pueden considerarse desde dos puntos de vista. Los objetivos de calidad de funcionamiento para paquetes de transacción (por ejemplo, pérdida de paquetes y retardo) se rigen por las clases de transacción especificadas en Recomendaciones UIT-T como la Y.1541 para servicios IP e I.356 para servicios ATM. Estas clases abarcan una amplia gama de servicios que incluyen aplicaciones de voz, datos y multimedia. Los parámetros asociados definen niveles aceptables de calidad de funcionamiento (por ejemplo, paquetes perdidos) para cada clase de transacción. Las expectativas de fiabilidad expresadas como un nivel de prioridad se relacionan con el establecimiento del enlace o de la "conexión" como un trayecto conmutado por etiqueta (LSP, *label switched path*) con conmutación por etiqueta multiprotocolo (MPLS, *multi-protocol label switching*) a través del cual se puede encaminar una transacción por paquete en la red. Entre los mecanismos utilizados para alcanzar estos objetivos de QoS están los métodos de encaminamiento de llamada y conexión y los métodos para la atribución de recursos QoS tales como ancho de banda, encaminamiento por nivel de prioridad, puesta en cola por nivel de prioridad, y restablecimiento de transporte. El objeto de este anexo es la fiabilidad de esos trayectos LSP, expresada en forma de un nivel de prioridad, y la necesidad de especificar niveles de prioridad para la señalización de la QoS.

Los niveles de prioridad de tráfico desempeñan un papel importante a la hora de proporcionar a los clientes de redes de telecomunicaciones un servicio con una fiabilidad/disponibilidad aceptables. Por ejemplo, las comunicaciones de urgencia requieren el más alto nivel de prioridad del control de admisión disponible en condiciones que implican desastres naturales y ataques terroristas. En las redes telefónicas públicas conmutadas (RTPC) de hoy en día, este nivel de prioridad es único. La fiabilidad/disponibilidad deseada puede solicitarse como un nivel de prioridad para una determinada función de red que, as u vez, determina el establecimiento de un LSP. Dos funciones de red para consideraciones de prioridad en redes de paquetes en evolución son las siguientes:

- Control de admisión de conexión: Las políticas de control de admisión dan preferencia a flujos continuos de tráfico que el proveedor de servicio considera más críticos (por ejemplo, comunicaciones de urgencia) en condiciones de congestión. La prioridad por control de admisión es una forma de dar preferencia con el fin de admitir los LSP de nivel de prioridad más alto antes que los LSP de nivel de prioridad más bajo.
- Restablecimiento: El restablecimiento se define aquí con un sentido lato como la respuesta mitigante de la red en condiciones de fallo. Entre los posibles métodos de recuperación tras fallo está la conmutación de protección automática para la protección de línea/trayecto y los métodos de restablecimiento con malla compartida. Los flujos continuos de tráfico de servicios críticos pueden pedir el restablecimiento con un nivel de prioridad más alto. Esos flujos continuos de trafico pueden encaminarse a través de un LSP que tenga el nivel de prioridad de restablecimiento "marcado" apropiadamente.

El establecimiento de prioridades de tráfico debe permitir una máxima flexibilidad para la implementación desde la perspectiva de los proveedores de servicio. Los niveles de prioridad deben cumplir los siguientes requisitos:

- El número de clases de prioridad debe ser pequeño, con el fin de asegurar la escalabilidad.
- Se debe evitar la división de cada una de las clases de prioridad, con el fin de asegurar la simplicidad.
- Los niveles de prioridad son relativos y no están asociados a parámetros específicos (por ejemplo, tiempo hasta el restablecimiento), ni a sus valores.

- Se debe permitir a los proveedores de servicio escoger, para sus ofertas de servicio, el número de niveles de prioridad entre los que forman el conjunto disponible. En consecuencia, pueden establecer acuerdos de nivel de servicio (SLA) para cualquier tratamiento dado de la clase de prioridad, con sus clientes, incluyendo otros proveedores de servicio (interfaz red-red).

Para el tráfico de servicio al cliente, en el control de admisión de conexión existen cuatro niveles de prioridad:

- Crítico: Nivel de prioridad único reservado para tráfico de comunicaciones de urgencia para todos los proveedores de servicio, nacionales e internacionales.
- Alto: Son ejemplos de estos servicios los servicios gubernamentales, los de clientes comerciales importantes, los de redes privadas virtuales.
- Normal: Son ejemplos de estos servicios los servicios vocales residenciales.
- Mejor esfuerzo: Son ejemplos de estos servicios los prestados por el proveedor de servicio Internet

Para restablecimiento existen tres niveles de prioridad: alto, normal, y mejor esfuerzo. Los ejemplos de servicios antes indicados son aplicables en este caso; las comunicaciones de urgencia solicitarían alta prioridad.

Como se ha dicho anteriormente, un proveedor de servicio puede hacer ofertas de servicios con un determinado nivel de prioridad en base a las capacidades de red disponibles y las necesidades de los clientes. Por ejemplo, un proveedor de servicio puede optar por ofrecer servicios con los cuatro niveles de prioridad definidos para control de admisión de la conexión, y con sólo dos niveles de prioridad para restablecimiento: alto y normal. Por otro lado, un proveedor ISP puro puede optar por ofrecer solamente los niveles de prioridad crítico y mejor esfuerzo para el control de admisión de la conexión, y sólo el nivel de prioridad mejor esfuerzo para restablecimiento.

Apéndice I

Método QoS generalizado basado en control de recursos independiente

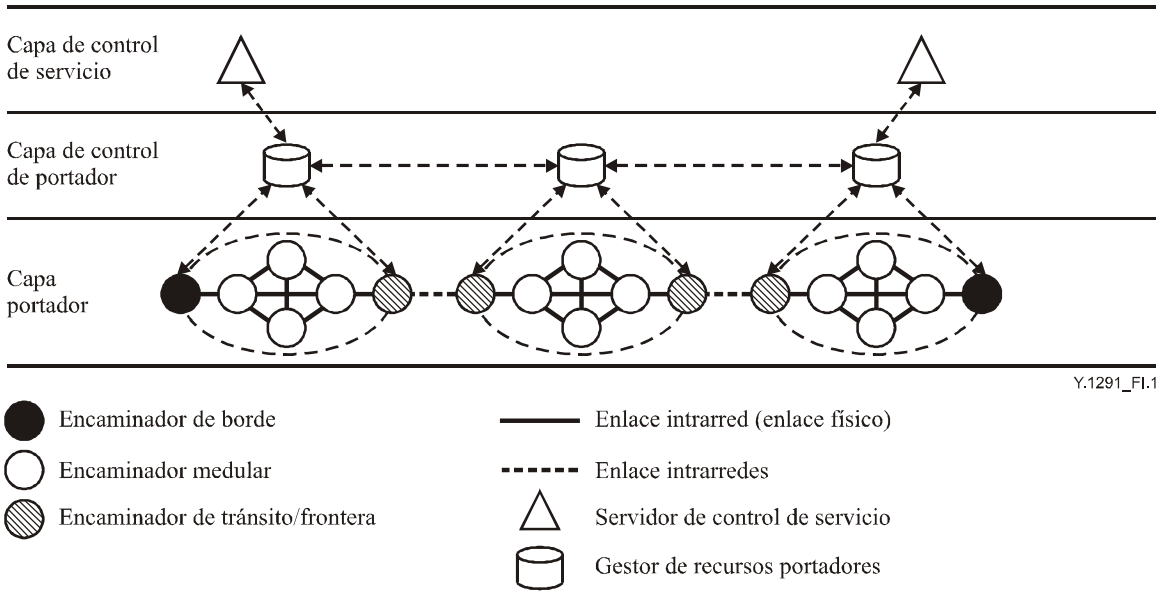

Figura I.1/Y.1291 – Método QoS generalizado basado en control de recursos independiente

Para soportar servicios cuyos requisitos de calidad de funcionamiento varían, a través de una sola red medular IP, y garantizar la QoS de servicios en tiempo real con conexión (como la telefonía IP), se ha desarrollado un método QoS generalizado basado en control de recursos independiente, que se muestra en la figura I.1. El método integra MPLS, DiffServ, ingeniería de tráfico y gestión de políticas.

Los servicios que requieren QoS garantizada se dividen en categorías establecidas de acuerdo con tipos de servicio generales (por ejemplo, voz) los niveles de tratamiento de la QoS (por ejemplo, EF). Para facilitar el manejo y asegurar la estabilidad de la red, la red medular IP de un proveedor de red se divide en múltiples dominios administrativos. Esta división es flexible y puede no coincidir con la división en dominios de encaminamiento. Por ejemplo, un dominio administrativo puede ser tan pequeño que sólo contenga un encaminador de borde, o tan grande que contenga una red de operador completa.

Un gestor de recursos portadores (BRM, *bearer resource manager*) es una función de control de recursos independiente que gestiona todos los recursos portadores en cada dominio administrativo y podría implementarse en una o varias unidades. El BRM registra y mantiene la base de datos de topología y recursos de red (NTRD, *network topology and resource database*). Con el apoyo de las NTRD, el BRM efectúa la selección de trayecto intradominio, atribución de recursos y control de admisión para un flujo de servicio. Los BRM de diferentes dominios interactúan mediante señalización para efectuar el control de recursos para flujos de aplicación interdominios. Además, un BRM puede tener también funciones como la de gestión de política, gestión de SLA, metraje de tráfico LSP, e interfaz cos servidores AAA.

Diversos servidores de control de servicio (SCS, *service control servers*) se encargan de controlar diversas peticiones de servicio (por ejemplo, señalización de llamada vocal), identificación de los puntos de origen y de terminación de cada petición de servicio, traducción de número (o nombre) en una dirección IP, y después enviar las peticiones de recursos al BRM del dominio de origen. En el caso de servicios que deben cumplir requisitos de QoS pero que no tienen servidores de control de servicio, como en los servicios punto a punto, los anfitriones pueden iniciar una petición de servicio QoS mediante el RSVP u otros protocolos de señalización QoS. En el presente caso, RSVP sólo es utilizado por los anfitriones para solicitar la garantía de QoS, y los encaminadores no están obligados a soportar el RSVP para reserva de recursos flujo por flujo. El equipo desplegado para procesar peticiones de servicio QoS de anfitriones puede considerarse un caso particular de SCS.

Un BRM recibe peticiones de recursos del SCS dentro de su dominio administrativo o de otro BRM. Las procesa y las notifica en retorno al SCS. Al mismo tiempo, si se admite una petición de recursos de flujo de servicio, el BRM notifica la identificación del flujo, el trayecto y los atributos QoS a los encaminadores de borde de ingreso. El encaminador de borde de ingreso identifica, clasifica, marca, aplica la política, conforma, y encapsula los paquetes de un flujo con la información de QoS especificada por el BRM.

En el caso de flujos de servicio que pasan por múltiples proveedores de red, generalmente hay pasarelas de aplicación y encaminadores fronterizos entre diferentes proveedores de red que se interconectan a través de los recursos del enlace fijo y los SLA interredes especificados. Diferentes proveedores de red pueden desplegar diferentes mecanismos QoS en sus redes. En este caso, los BRM sólo gestionan los recursos de enlace intrarred, en tanto que las pasarelas de aplicación o los encaminadores fronterizos gestionan los recursos de enlace interredes atendiendo a los SLA especificados y una pasarela de aplicación o encaminador fronterizo actúa como el encaminador de borde de ingreso o de egreso.

Los bloques de construcción pertinentes para este método comprenden todos los bloques indicados en la figura 1. El BRM actúa como un plano de control y de gestión físicamente independiente. Los bloques de construcción interactúan esencialmente mediante señalización flujo por flujo y en base a una gestión de recursos LBN (red de portador lógico) por LBN. Existe evidentemente una interfaz de señalización entre el plano de datos y el plano de control.

I.1 Flexibilidad de implementación para redes de paquetes con soporte de MPLS

En este caso se supone que la MPLS sensible a DiffServ está soportada en las redes medulares IP.

La tecnología MPLS LSP se despliega para preaprovisionar una red de portador lógico (LBN, *logical bearer network*) para cada clase de servicio a través de la red IP subyacente, manual o automáticamente, mediante el protocolo RSVP-TE o CR-LDP. En el caso de flujos de servicio pertenecientes a una clase de servicio, la selección de trayecto, atribución de recursos, control de admisión y reenvío de etiqueta se tratan dentro de la misma LBN. La planificación de la topología y la reserva de ancho de banda de cada LBN depende de los datos de metraje y pronóstico de tráfico, políticas administrativas y SLA, que pueden ajustarse automática o manualmente para la protección de LSP, cambios de capacidad u optimización de la calidad de funcionamiento de la red de acuerdo con las constricciones de ingeniería de tráfico.

Dentro del recurso restante de las redes de paquetes subyacentes, el tráfico BE sin requisitos de QoS todavía se encamina y reenvía por métodos convencionales de encaminamiento y reenvío IP con o sin DiffServ.

El BRM registra y mantiene una base de datos de topología y recursos de red (NTRD) separadamente para cada LBN. Con el apoyo de las NTRD y políticas, el BRM efectúa la selección de trayecto intradominio, atribución de recursos y control de admisión para un flujo de servicio dentro de su correspondiente LBN. En cuanto al recurso restante de las redes de paquetes subyacentes, el BRM podría también efectuar atribución de recursos y control de admisión.

La información de trayecto QoS para un flujo especificado por BRM es una pila de etiquetas multicapa que representa un conjunto LSP concatenado. El encaminador de borde encapsula los paquetes con las etiquetas contenidas en esta pila, lo que a su vez hace que los encaminadores de tránsito intermedios reenvíen los paquetes de un flujo a lo largo de un trayecto especificado atendiendo a la pila de etiquetas y el nivel de prioridad especificado.

I.2 Flexibilidad de implementación para redes de paquetes sin soporte de MPLS

En este caso, el control de admisión y la reserva de recursos se aplican dinámicamente con la reserva de recursos enlace por enlace, y no se requiere que la capa de portador tenga la capacidad MPLS. El encaminamiento y reenvío de todo el tráfico se realiza bajo el control de protocolos de encaminamiento IP tradicionales e IP Diffserv.

El BRM se despliega para gestionar directamente todos los recursos de enlace físico dentro de cada dominio administrativo. El BRM contiene y mantiene una base de datos de topología y recursos de red (NTRD). Con el apoyo de la información contenida en la NTRD, el BRM se encarga de la consulta de rutas, reserva de recursos enlace por enlace y control de admisión para cada flujo que requiere garantía de QoS. Si se admite un flujo con un alto nivel de prioridad, no interferirá a otros flujos de tráfico.

I.3 Flexibilidad de implementación para control de recursos distribuido

En este caso, las LBN son enlaces virtuales (llamados tuberías de QoS) entre pares de encaminadores de borde (ER, *edge routers*) de ingreso-egreso en un dominio de red. Se establece una tubería de QoS para transportar flujos agregados de un servicio o clase de QoS determinados.

Si la función BRM está implementada en encaminadores de borde (ER), el control de recursos flujo por flujo es distribuido a los bordes. La función de control de recursos (RCF, *resource control function*) en ER mantiene la tabla de estados de los recursos de las correspondientes tuberías de QoS y en consecuencia efectúa control de admisión y atribución de recursos. También procesa la señalización QoS.

Las tuberías de QoS se ajustan manual o automáticamente a un plazo medio o un plazo largo, lo que puede realizar el sistema de gestión de red.

Apéndice II

Esquema de elevación del nivel de prioridad

El esquema de elevación del nivel de prioridad (PPS, *priority promotion scheme*) es un nuevo esquema para control de tráfico que está todavía en una etapa experimental. Brevemente, el PPS utiliza una forma de control de admisión para obtener QoS de extremo a extremo en una red basada en paquetes. Las principales aplicaciones para este tipo de esquema son servicios multimedios interactivos como voz por IP, videocharla (*video chat*), y videoconferencia. Específicamente, el esquema se basa en la medición de extremo a extremo de recursos de red por sistemas de extremo. Antes de establecer una sesión, o incluso durante una sesión, el sistema de extremo fuente detecta, mide o sondea la disponibilidad de nuevos recursos de red enviando paquetes con un nivel de prioridad inferior en una unidad al de los paquetes normales. El resultado es la modificación del valor del punto de código DiffServ (DSCP, *DiffServ code point*) de los paquetes IP subsiguientes: el nivel de prioridad se eleva para establecer firmemente la sesión, se baja para dejar recursos que se utilizan en sesiones existentes, o en otro caso se ajusta para que el número de paquetes no rebase la capacidad disponible. Se supone que la red, es decir, los enlaces de salida de los encaminadores o los conmutadores L2 sólo soportan la forma de control de prioridad clase por clase que acompaña a la arquitectura DiffServ. Haciendo que todos los sistemas de extremo adopten el comportamiento antes descrito se obtiene la QoS de extremo a extremo sin necesidad de que los estados se mantengan flujo por flujo por flujo en los nodos de red.

BIBLIOGRAFÍA

[IETF RFC 1633] BRADEN (R.), *et al*.: Integrated Services in the Internet Architecture: an Overview, junio de 1994.

[IETF RFC 2205] BRADEN (R.), *et al*.: Resource ReSerVation Protocol (RSVP) – Version 1 Functional Specification, septiembre de 1997.

[IETF RFC 2309] BRADEN (R.), *et al*.: Recommendations on Queue Management and Congestion Avoidance in the Internet, abril de 1998.

[IETF RFC 2386] CRAWLEY (E.), *et al*.: A Framework for QoS-based Routing in the Internet, agosto de 1998.

[IETF RFC 2474] NICHOLS (K.), *et al*.: Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers, diciembre de 1998.

[IETF RFC 2748] DURHAM (D.), *et al*.: The COPS (Common Open Policy Service) Protocol, enero de 2000.

[IETF RFC 2753] YAVATKAR (R.), *et al*.: A Framework for Policy-based Admission Control, enero de 2000.

[IETF RFC 2990] HUSTON (G.): Next Steps for the IP QoS Architecture, noviembre de 2000.

[IETF RFC 2996] BERNET (Y.): Format of the RSVP DCLASS Object, noviembre de 2000.

[IETF RFC 2998] BERNET (Y.), *et al*.: A Framework for Integrated Services Operation over Diffserv Networks, noviembre de 2000.

[IETF RFC 3031] ROSEN (E.), *et al*.: Multiprotocol Label Switching Architecture, enero de 2001.

[IETF RFC 3032] ROSEN (E.), *et al*.: MPLS Label Stack Encoding, enero de 2001.

[IETF RFC 3198] WESTERINEN (A.), *et al*.: Terminology for Policy-Based Management, noviembre de 2001.

[IETF RFC 3209] AWDUCHE (D.), *et al*.: RSVP-TE: Extensions to RSVP for LSP Tunnels, diciembre de 2001.

[IETF RFC 3261] ROSENBERG (J.), *et al*.: SIP: Session Initiation Protocol, junio de 2002.

[IETF RFC 3270] LE FAUCHEUR (F.), *et al*.: Multi-Protocol Label Switching (MPLS) Support of Differentiated Services, mayo de 2002.

[IETF RFC 3272] AWDUCHE (D.): Overview and Principles of Internet Traffic Engineering, mayo de 2002.

[Jacobson, 1988] JACOBSON (V.): Congestion Avoidance and Control*, Proceedings of ACM SIGCOMM'88*, pp. 314-329, agosto de 1988.

[Lin *et al*., 1997] LIN (D.), MORRIS (R.): Dynamics of Random Early Detection*, Proceedings of ACM SIGCOMM'97*, pp. 127-138, septiembre de 1997.

[Chen] CHEN (Shigang), NAHRSTEDT (Klara): An Overview of Quality-of-Service Routing for the Next Generation High-Speed Networks: Problems and Solutions*, IEEE Network, Special Issue on Transmission and Distribution of Digital Video*, Vol. 12, No. 6, pp. 64-79, noviembre/diciembre de 1998.

[Apostolopoulos] APOSTOLOPOULOS (D.), *et al*.: Intra domain QoS Routing in IP Networks: A Feasibility and Cost Benefit Analysis, *IEEE Network*, Vol. 13, No. 5, pp. 42, septiembre/octubre de 1999.

[Floyd] FLOYD (S.), JACOBSON (V.): Random Early Detection Gateways for Congestion Avoidance, *IEEE/ACM Transactions on Networking*, Vol. 1, No. 4, pp. 397-413, agosto de 1993.

[Nagle] NAGLE (J.): On Packet Switches with Infinite Storage. *IEEE Trans. on communications*, Vol. COM-35, pp. 435-438. abril de 1987.

SERIES DE RECOMENDACIONES DEL UIT-T

Serie A	Organización del trabajo del UIT-T
Serie B	Medios de expresión: definiciones, símbolos, clasificación
Serie C	Estadísticas generales de telecomunicaciones
Serie D	Principios generales de tarificación
Serie E	Explotación general de la red, servicio telefónico, explotación del servicio y factores humanos
Serie F	Servicios de telecomunicación no telefónicos
Serie G	Sistemas y medios de transmisión, sistemas y redes digitales
Serie H	Sistemas audiovisuales y multimedios
Serie I	Red digital de servicios integrados
Serie J	Redes de cable y transmisión de programas radiofónicos y televisivos, y de otras señales multimedios
Serie K	Protección contra las interferencias
Serie L	Construcción, instalación y protección de los cables y otros elementos de planta exterior
Serie M	RGT y mantenimiento de redes: sistemas de transmisión, circuitos telefónicos, telegrafía, facsímil y circuitos arrendados internacionales
Serie N	Mantenimiento: circuitos internacionales para transmisiones radiofónicas y de televisión
Serie O	Especificaciones de los aparatos de medida
Serie P	Calidad de transmisión telefónica, instalaciones telefónicas y redes locales
Serie Q	Conmutación y señalización
Serie R	Transmisión telegráfica
Serie S	Equipos terminales para servicios de telegrafía
Serie T	Terminales para servicios de telemática
Serie U	Conmutación telegráfica
Serie V	Comunicación de datos por la red telefónica
Serie X	Redes de datos y comunicación entre sistemas abiertos
Serie Y	**Infraestructura mundial de la información, aspectos del protocolo Internet y Redes de la próxima generación**
Serie Z	Lenguajes y aspectos generales de soporte lógico para sistemas de telecomunicación

Impreso en Suiza
Ginebra, 2005