



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 356 622**

51 Int. Cl.:
H04Q 11/04 (2006.01)
H04L 12/56 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **02776701 .1**
96 Fecha de presentación : **20.09.2002**
97 Número de publicación de la solicitud: **1428408**
97 Fecha de publicación de la solicitud: **16.06.2004**

54 Título: **Transmisión distribuida de informaciones en una red de comunicaciones sin conexión orientada a paquetes.**

30 Prioridad: **20.09.2001 DE 101 46 349**
04.10.2001 DE 101 48 893
14.12.2001 DE 101 61 508
14.12.2001 DE 101 61 546
14.12.2001 DE 101 61 547

45 Fecha de publicación de la mención BOPI:
11.04.2011

45 Fecha de la publicación del folleto de la patente:
11.04.2011

73 Titular/es:
NOKIA SIEMENS NETWORKS GmbH & Co. KG.
St. Martin Strasse 76
81541 München, DE

72 Inventor/es: **Schrodi, Karl**

74 Agente: **Zuazo Araluze, Alexander**

ES 2 356 622 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

DESCRIPCIÓN

Transmisión distribuida de informaciones en una red de comunicaciones sin conexión orientada a paquetes.

El objeto de la solicitud se encuentra, entre otros, también en el ámbito de redes unificadas para el transporte fiable de información codificada digitalmente para servicios de datos, voz, audio/video y otros servicios y aplicaciones, cumpliendo las correspondientes exigencias de Quality of Service (calidad del servicio) específicas del servicio y de la aplicación hasta la comunicación interactiva en tiempo real. Las solicitudes de patente DE 10146349.9, DE 10161508.6, DE 10161546.9, DE 10161547.7 se encuentran también en este ámbito.

En el pasado se han configurado dos tipos esenciales de redes de comunicaciones para transmitir informaciones alojadas en flujos de transporte: redes de datos orientadas a paquetes y redes de voz orientadas a la línea. Las mismas se diferencian entre otros en sus distintas exigencias de Quality of Service (QoS).

La Quality of Service o calidad del servicio se define de diferentes maneras en función del contexto y en consecuencia se evalúa con distintas métricas en cada caso. Ejemplos conocidos de métricas para medir la calidad del servicio son la cantidad máxima transmisible de informaciones (Bandwidth o anchura de banda), la cantidad de informaciones transmitidas, la cantidad de informaciones no transmitidas (Loss Rate), el retardo - dado el caso promediado - en la transmisión (Transmission Delay), la desviación - dado el caso promediada - de la distancia, por lo demás usual, entre dos transmisiones de información (Delay Jitter, Interarrival Jitter) o la cantidad de informaciones no admitidas en absoluto para la transmisión (Blocking Rate).

Los servicios se denominan en redes multimedia también "aplicaciones multimedia". Bajo red multimedia se entiende al respecto una red en la que se realiza un conjunto de servicios diferentes. En sentido estricto se entiende bajo este concepto en particular una red de banda ancha, de servicios integrados (B-ISDN = Broadband Integrated Services Digital Network), en la que los flujos de transporte que resultan cuando se utilizan los servicios pueden transmitirse mediante un mecanismo de transporte unificado, preferiblemente orientado a paquetes. En el concepto aplicación multimedia se incluyen al respecto tanto servicios como la telefonía usual (en redes de IP orientadas a paquetes denominadas también "Voice over IP (VoIP)", como también servicios como fax, teleconferencia, videoconferencia, video sobre demanda (Video on Demand, VoD) y similares.

Las redes (de voz) orientadas a la línea están diseñadas para transmitir flujos de transporte en los que están alojadas las informaciones (de voz) que fluyen continuamente. En el mundo especializado se denominan las mismas también "conversación", "llamada" o "sesión". La transmisión de las informaciones se realiza al respecto usualmente con una elevada calidad del servicio y seguridad. Por ejemplo, para la voz es importante un retardo (Delay) mínimo sin oscilaciones del tiempo de retardo (Delay-Jitter), ya que la voz necesita cuando se reproduce en el aparato receptor un flujo de información continuo. Por ello una pérdida de información no puede compensarse repitiendo la transmisión de la información no transmitida y da lugar en el aparato receptor usualmente a un ruido seco perceptible acústicamente. En el mundo especializado se denomina la transmisión de voz en general también "servicio (de transmisión) en tiempo real" o bien "servicio de tiempo real" (Realtime-Service).

Se logra una baja tasa de bloqueo (Blocking Rate) por ejemplo mediante el correspondiente dimensionado y planificación de las redes de voz. Un Delay o bien Delay Jitter bajo y muy constante se consigue en la transmisión común de varios flujos de transporte a través de un canal común usualmente utilizando un multiplexado estático (en el tiempo), denominado también TDM (Time Division Multiplex). Al respecto se segmentan los flujos de transporte en el emisor en unidades homogéneas de una longitud fija - también denominadas ranuras de tiempo o bien Timeslots y se transmiten entrelazadas en el tiempo. La asignación de las ranuras de tiempo a los correspondientes flujos de transporte se indica mediante su posición dentro del canal. Tras la transmisión común pueden así asignarse en el receptor las ranuras de tiempo a sus correspondientes flujos de transporte y caso necesario también volver a ensamblarse de nuevo en los flujos de datos iniciales. Como consecuencia la capacidad de transmisión de los flujos de datos no está sometida en la transmisión orientada a la línea a oscilación alguna, sino que está fijada en un valor predeterminado (por ejemplo 64 kbps en las actuales redes telefónicas ISDN).

Las redes (de datos) orientadas a paquetes están diseñadas para transmitir flujos de transporte configurados como flujos de paquetes, que en el mundo especializado se denominan también "flujos de paquetes de datos". Al respecto no tiene que garantizarse usualmente una calidad del servicio elevada. Por ejemplo, para un e-mail no es necesario ningún retardo (Delay) mínimo sin oscilaciones en el tiempo de retardo (Delay-Jitter), ya que un e-mail no tiene que reproducirse en el receptor en tiempo real. Aquí es importante más bien una transmisión sin errores del e-mail. Una pérdida de información se compensa por lo tanto usualmente mediante una nueva transmisión (Retransmission) de la información no transmitida o transmitida incorrectamente. El Delay de un e-mail varía en consecuencia en función de la frecuencia de retransmisión. También el Delay Jitter es en consecuencia más bien elevado. En el mundo especializado se denomina la transmisión de datos por ello también transmisión sin condiciones de tiempo real o bien "Non-Realtime-service".

Básicamente en las redes de datos orientadas a paquetes no hay ninguna tasa de bloqueo (Blocking Rate). En principio se transmiten siempre todos los paquetes de todos los flujos de transporte. La transmisión de los flujos de transporte se realiza no obstante, ya con un grado de carga moderado de una red de datos, con retardos en el

tiempo significativamente oscilantes, ya que los distintos paquetes se transmiten usualmente en la secuencia de su acceso a la red, es decir, los retardos son tanto más grandes cuanto más paquetes han de transmitirse en una red de datos. Una transmisión común de varios flujos de transporte a través de un canal común se logra usualmente utilizando un multiplexado estadístico (en el tiempo). Entonces se transmiten los paquetes de los flujos de transporte en el emisor según reglas estadísticas entrelazados en el tiempo. La norma podría por ejemplo prescribir que los paquetes se transmitan en la secuencia de su llegada (Best Effort o lo mejor que se pueda). Cuando llegan simultáneamente varios paquetes, se transmite uno de ellos, mientras que el resto - con la consecuencia de un elevado Delay Jitter - se memoriza transitoriamente (buffer) durante un breve tiempo. Cuando llegan a la vez más paquetes que los que pueden ser memorizados transitoriamente, se rechazan los paquetes excedentes. La asociación de los paquetes a los correspondientes flujos de transporte se indica mediante una información de asociación en el Overhead (compuesto por Header o cabecera y/o Trailer cola) de los paquetes. Tras la transmisión común, pueden así asociarse en el receptor los paquetes a sus correspondientes flujos de transporte. La capacidad de transmisión de los flujos de transporte no está sometida en la transmisión orientada a paquetes a ninguna limitación, sino que puede presentar básicamente (dentro de la capacidad del canal común) en cada momento un valor distinto cualquiera.

En el curso de la convergencia de redes de datos orientadas a la línea, orientadas a la voz y a paquetes, se realizarán los servicios de transmisión de la voz, y en el futuro también los servicios de banda más ancha como por ejemplo la transmisión de informaciones de imágenes en movimiento (VoD, videoconferencia), en redes (multimedia) de servicios integrados orientadas a paquetes, denominadas también "redes de voz-datos", es decir, la transmisión de los servicios en tiempo real transmitidos hasta ahora usualmente orientados a la línea se realiza en una red de voz-datos convergente en flujos de paquetes. Estos se denominan también "flujos de paquetes en tiempo real". Entonces resulta el problema de que para una realización orientada a paquetes de un servicio en tiempo real es necesaria una elevada calidad de servicio y seguridad, para que la misma sea comparable cualitativamente a una transmisión orientada a línea, mientras que las redes de datos actuales (orientadas a paquetes) y en particular Internet no prevén ningún mecanismo adecuado para garantizar una elevada calidad del servicio.

Las exigencias a la calidad del servicio en las redes de servicios integrados orientadas a paquetes rigen en general para todos los tipos de redes. Las mismas son independientes de la configuración concreta de la orientación a paquetes. Los paquetes pueden en consecuencia estar configurados como paquetes de Internet, X.25 o bien Frame-Relay (transmisión de tramas), pero también como células ATM. Los mismos se denominan también a veces "mensajes", sobre todo cuando un mensaje se transmite en un paquete. Los flujos de paquetes de datos y los flujos de paquetes en tiempo real son al respecto ejemplos de ejecución de flujos de transporte transmitidos en redes de comunicaciones. Los flujos de transporte se denominan también "enlaces", precisamente también en redes orientadas a paquetes en las que se utiliza una técnica de transmisión sin conexión. Por ejemplo se realiza una transmisión de información en TCP/IP con ayuda de los llamados Flows o flujos, mediante los que el emisor y el receptor (por ejemplo Web Server o servidor de la web y Browser o navegador) se unen entre sí pese al carácter de sin conexión de IP a un nivel lógico abstracto, es decir, abstraído lógicamente significan también los Flows enlaces. Sólo es decisivo para un enlace que antes de la transmisión se establezca el enlace, en el que se logra un contexto que permanece al menos durante la transmisión. Tras la transmisión puede realizarse un establecimiento explícito del enlace. No obstante, son posibles también mecanismos implícitos como por ejemplo un Time Out (tiempo de interrupción) del enlace tras un espacio de tiempo fijado sin transmisión.

La red de datos más conocida actualmente es Internet. Internet es una red de datos abierta (de transporte a larga distancia) con interfaces abiertas para el enlace de redes de datos (al menos locales y regionales) de distintos fabricantes. La característica principal reside por lo tanto hasta ahora en que pone a disposición una plataforma de transporte independiente del fabricante. Los mecanismos adecuados para garantizar la calidad del servicio juegan un papel secundario y por lo tanto apenas existen.

La convergencia de telecomunicación (también denominada redes de voz) y del mundo de datos clásico (también denominado redes de datos) hasta las redes y servicios basados en IP (Internet Protocol), es en relación con la técnica de IP una tarea difícil, ya que la misma está diseñada como red de datos orientada a paquetes sobre todo como transmisión "Best Effort" (lo mejor que se pueda) y como mucho prevé el cumplimiento de acuerdos del nivel de servicio "Service Level Agreement, SLA" más bien inconcretos, mientras que en la telecomunicación se encuentran en primer plano exigencias muy rigurosas en cuanto a la QoS, la fiabilidad, la disponibilidad y la seguridad de la red y los servicios. El "mundo de Internet" reacciona a esta tarea con múltiples principios de solución cada vez más complejos y costosos, pero no ha encontrado hasta ahora todavía ninguna solución completa que también sea manejable y soportable en el aspecto económico.

Las exigencias QoS de un servicio o de una aplicación a una red pueden estar definidas mediante diversos criterios, de los que mencionaremos algunos a modo de ejemplo:

- la curva característica de flujo de la información codificada digitalmente, es decir, la anchura de banda necesaria o bien curva característica de anchura de banda (anchura de banda fija, anchura de banda variable, por ejemplo con el valor medio, valor de cresta, factor de capacidad de ráfaga otros parámetros característicos) y la sensibilidad frente a pérdidas de información,

- la curva característica de retardo, es decir, las repercusiones de un retardo absoluto (tiempo de recorrido desde la fuente hasta el sumidero de la información) y la sensibilidad frente a oscilaciones del tiempo de recorrido o de retardo (evidentemente pueden transformarse las oscilaciones del retardo mediante "compensación" en retardo absoluto, pero esto la mayoría de las veces es muy costoso),
- 5
- la consistencia en el tiempo o "invarianza en el tiempo" necesaria o no necesaria para la información transmitida, es decir, si las unidades de información deben ser suministradas en la misma secuencia exacta en la que han sido aportadas o no (dado el caso debe considerarse también la capacidad o incapacidad de niveles de servicio y aplicación más elevados).

Las consecuencias que resultan de distintas exigencias QoS las indicaremos en base a dos ejemplos:

- 10
- I. Las aplicaciones de audio/video orientadas unidireccionalmente (por ejemplo "Streaming Video", video reproducido en tiempo real) exigen ciertamente una presentación en tiempo real en el receptor, pero en la mayoría de los casos es intrascendente si el retardo absoluto es entonces de 1/100, 1 ó 5 segundos, siempre que tras comenzar el paso del video exista continuidad. Esta tolerancia al retardo podría utilizarse por ejemplo para compensar por medio de repeticiones pérdidas de información y mejorar así la calidad. Alternativamente podría transmitirse también con redundancia (mayor anchura de banda), para así compensar posibles pérdidas de datos.
- 15
- II. La comunicación en tiempo real (voz, video,...) interactiva, es decir, bidireccional entre personas debe tener en cuenta la capacidad de reacción y el comportamiento típico en comunicación y diálogo de las personas. Aquí debe limitarse el retardo absoluto (y con ello naturalmente también las oscilaciones del retardo) a unos pocos cientos de milisegundos (por ejemplo 200 ms). Por otro lado, pueden tolerarse entonces en determinadas condiciones tasas de pérdidas algo más elevadas, ya que la capacidad del cerebro humano para "aplanar irregularidades" en la voz y en la percepción visual es muy acusada y en el diálogo la atención para percibir pequeños defectos más bien se reduce ligeramente. No obstante, más complejos son los diálogos en tiempo real entre máquinas. Entonces debe pretenderse en determinadas condiciones tanto que la información sea completa como también que los retardos sean reducidos hasta las proximidades del límite físico debido a la distancia espacial (tiempo de recorrido aprox. 5 ms por cada 1000 km de distancia).
- 20
- 25

Si están definidas las exigencias de QoS, entonces puede una red, siempre que tenga aún reservas en estas zonas, utilizar sin problemas éstas también para compensar déficits en otra zona. Describiremos esta compensación en base a dos ejemplos:

- 30
- I. Si una aplicación tolera pérdidas de información relativamente altas, entonces pueden reducirse las oscilaciones del retardo rechazándose unidades de información afectadas por un elevado retardo. A la inversa, pueden utilizarse también naturalmente mayores oscilaciones del retardo para lograr pérdidas inferiores, lo cual no obstante da lugar a grandes memorias buffer.
- 35
- II. Si el máximo de las oscilaciones del retardo se encuentra por debajo de la distancia mínima en el tiempo entre las unidades de información aportadas (la llamada "red rápida"), entonces pueden excluirse problemas con la consistencia en el tiempo de la información transmitida. Si se prevén medidas para restablecer esta consistencia en el tiempo, entonces pueden tolerarse oscilaciones del retardo relativamente grandes, siempre que no se sobrepase el marco del retardo absoluto admisible.

40

Además de la QoS, también es un parámetro esencial la disponibilidad general de los servicios, que depende en gran medida de la red y de sus características. Cuando hay una falta, por ejemplo cuando fallan componentes individuales o líneas de enlace, ¿se dispone de una vía sustitutoria y en cuánto tiempo puede disponerse de ella? ¿Se presentan interrupciones perceptibles por el usuario y cuánto duran las mismas? ¿debe intervenir dado el caso el operador de la red o incluso el propio usuario de alguna forma para restablecer el servicio?. La fiabilidad de la red en sí y la forma en que puede contribuir a cubrir casos de falta, y dado el caso al restablecimiento de las aplicaciones, es al respecto de gran importancia.

45

Una red unificada debe por lo tanto enjuiciarse cualificadamente con condiciones marco del tipo aquí indicado y naturalmente debe ser también lo más eficiente posible, es decir, debe poder realizarse con un coste y gasto lo menores posible.

Las técnicas de red conocidas cumplen las citadas condiciones en el mejor de los casos sólo parcialmente.

- 50
- 1) El principio más sencillo es la acreditada técnica de la conmutación de línea, en la que para cada relación de comunicaciones se conecta un (en el caso bidireccional o bien en relaciones múltiples también dado el caso dos o más) enlace dedicado (a veces denominado también "ruta" algo a la ligera) con una anchura de banda fijamente asignada y absolutamente reservada. Tales enlaces están configurados bien explícitamente como líneas físicas individuales (p.e. hilos de cobre) o bien como canales (virtuales) en los llamados sistemas de transmisión o conmutación, que permiten un aprovechamiento múltiple de líneas físicas. También es posible una mezcla de tramos parciales configurados distintos. La velocidad de datos posible en un tal enlace viene determinada por su propia anchura de banda, o bien la que tiene asignada y el tiempo de retardo para el
- 55

- transporte está compuesto por el "Propagation Delay" (retardo de propagación), es decir, el tiempo de recorrido sobre la línea, que depende de la distancia, y los "Switching Delays" (retardos de conmutación), es decir, los tiempos de procesamiento inherentes que resultan al "conmutar" la información ("datos") codificada digitalmente en los nodos de red (switches). "Conmutación" significa aquí el traslado de la información ("datos") desde una determinada línea/canal que llega hasta una línea/canal de salida determinada al establecerse el enlace. Ambos componentes del retardo (Delay) pueden suponerse constantes por lo general (es decir, durante el funcionamiento de los sistemas sin perturbaciones) mientras dura una relación de comunicaciones (con la ruta interconectada o bien "enlace" existente). Cuando no hay perturbaciones, está prescrita así y puede lograrse para todas las aplicaciones la misma QoS casi "óptima" (ninguna pérdida de información, retardo constante, por lo general relativamente pequeño, ninguna confusión). Pero a cambio debe estar conectado (y reservado) el enlace permanentemente mientras dura la relación de comunicación, incluso cuando la aplicación sólo la utilice muy poco (p.e. sólo esporádicamente). La fiabilidad/disponibilidad puede mejorarse conmutándose en caso de falta lo más inmediatamente posible a un enlace alternativo, puesto a disposición previamente (se necesita el doble de capacidad) o conectándose inmediatamente un enlace sustitutorio (retardo y coste, sobre todo cuando muchos enlaces se ven afectados simultáneamente por un fallo).
- 2) La técnica de conmutación de paquetes pretende un mejor aprovechamiento de los recursos (anchura de banda) mediante una utilización conjunta (sharing) flexible de líneas y canales (dado el caso virtuales) o bien medios de conmutación y transmisión mediante varias relaciones de comunicación. Representantes de ello conocidos y actuales son p.e. la técnica ATM orientada al enlace con paquetes de longitud fija (también denominados "células") y la técnica IP sin conexión con paquetes de longitud variable.
- A) La técnica ATM se impulsa también en la ITU-T bajo este concepto y con la finalidad de lograr una "ISDN de banda ancha" (B-ISDN). ATM tiene mecanismos para incluso con recursos muy escasos (anchuras de banda disponibles) proporcionar un amplio espectro de clases de servicio con QoS definida y garantizada (en promedio estadístico). No obstante, debido a ello los sistemas y redes resultantes son complejos y costosos. El dimensionado y el funcionamiento exigen personal especializado de alta cualificación. ATM funciona orientado al enlace, con una red de rutas y canales "virtuales", asociados entre sí jerárquicamente. Para múltiples clases de servicio diferentes pueden reservarse anchuras de banda individualmente para cada enlace y también "garantizarse" en función de la estadística de transporte tomada como base. Para ello se utilizan distintos mecanismos de colas y de planificación (Scheduling), que se ajustan en cada nodo por cada ruta y canal (enlace) mediante los correspondientes parámetros. Mediante unas normas de aceptación de dimensionado y de enlaces de fina granularidad, pueden limitarse en función de reglas estadísticas pérdidas de información y las componentes variables de los retardos de conmutación (Switching-Delays), que vienen determinados esencialmente por las colas. Una confusión entre unidades de información no es de esperar debido a la orientación a enlaces, cuando no hay perturbaciones. Como consecuencia de la orientación a enlaces, han de recorrerse de nuevo durante el tratamiento de casos de falta todos los mecanismos inherentes. Las ideas básicas son al respecto muy similares a menudo a las de la técnica de conmutación de líneas.
- B) La técnica de IP es más bien un principio pragmático, que se ha impuesto en el mundo de los datos debido a sus sencillos mecanismos básicos. La misma ha hecho en los últimos años enormes progresos, con lo que los sistemas y redes que se basan en ella son comparables en cuanto a su potencia (velocidad de datos, eficiencia del control) a sistemas basados en la técnica ATM. El éxito de la técnica de IP se basa en una gran proporción en que una gran parte de los servicios y aplicaciones se implantan ya en el aparato terminal en protocolos de Internet (IP) orientados a paquetes. Actualmente se pronostica que el crecimiento en los servicios basados en IP será también en el futuro un múltiplo de los correspondientes a otras tecnologías, con lo que parece probable una amplia migración de todos los servicios hacia el transporte mediante redes basadas en IP. Contrariamente a las redes ATM, funcionan las redes IP sin conexión y realizan solamente un servicio "Best Effort" (lo mejor que se pueda), con el que incluso con un dimensionado generoso de las redes apenas son posibles predicciones y desde luego no es posible ninguna garantía de la QoS alcanzable.
- C) Además se conocen hasta ahora los siguientes principios de solución:
- Se utiliza una red ATM como red núcleo (core). Los Edge Devices (equipos de primera línea) transforman los flujos de datos de IP en enlaces ATM de clases de servicio adecuadas y el transporte se realiza en los correspondientes enlaces en la red ATM. Son problemáticos la escalabilidad, la complejidad y el coste de establecimiento y funcionamiento (ver antes, técnica ATM). Esta solución ayuda más bien en el core. Cuando se utiliza (adicionalmente) en el Access (acceso), los inconvenientes son los mismos. Una alternativa en el Access es la solución siguiente.
 - Se utiliza un protocolo de señalización y se establecen sobre la red de IP enlaces con anchuras de banda reservadas (servicios integrados; Integrated Services-IntServ, RSVP). Este principio de solución es posible básicamente tanto punto a punto (End-to-End, E2E), es decir, de aparato terminal a aparato terminal, como también para secciones parciales. El mismo puede aplicarse para cada flujo de comunicación o (en el core) también para flujos de comunicación agregados. No obstante, es

complejo, costoso, no escalable (coste del control) e ineficiente, es decir, muy similar a la técnica ATM.

5 c) MPLS: Este principio se apoya en la técnica ATM. Se establecen rutas (enlaces) en la red, a través de las que se enruta selectivamente el transporte de flujos individuales (por lo general agregados). Se propone para QoS a menudo junto con RSVP y DiffServ (ver a continuación bajo el punto d) y puede realizarse también sobre la base de un transporte ATM. Reincide en cuanto a complejidad en la de los mecanismos orientados al enlace con todas las consecuencias ya indicadas (desde el control de la banda ancha hasta la vigilancia de que el enlace existe), es decir, la complejidad es similar a la de la técnica ATM. Debe, juntamente con el principio de solución DiffServ, reducir sobre todo la problemática allí expuesta (control selectivo del tráfico mediante rutas).

10 d) "Differentiated Services (DiffServ)": Los paquetes de datos se clasifican y marcan en el Edge Device (equipo de primera línea), en base a su pertenencia a determinados servicios, aplicaciones o relaciones de comunicación, etc. Adicionalmente puede o debería realizarse un control y vigilancia de acceso (referido al flujo), por ejemplo en cuanto a disponibilidad de recursos y cumplimiento de la curva característica de anchura de banda y QoS indicada. Los paquetes siguen entonces la ruta predeterminada por su información de cabecera de paquetes (por ejemplo dirección de destino) y los protocolos de enrutamiento a través de la red, tratándose (por ejemplo priorizándose) en cada nodo según su marca con el correspondiente "Per Hop Behaviour" (comportamiento por salto). El principio DiffServ deja la libertad del Per Hop Behaviour dentro de un "Single Routing Domain" (dominio de un solo enrutamiento) por ejemplo a la (sub)red de un operador, pero exige un tratamiento "Edge" completo entre tales dominios (subredes). El principio DiffServ no puede evitar cuellos de botella temporales y/o locales, ya que por lo general no se tienen en cuenta ni se realiza coordinación con las rutas predeterminadas por los protocolos de enrutamiento. Por lo general los paquetes con el mismo destino seguirán la misma ruta ajustada desde el momento en que se encuentran en un nodo. Esto puede traer como consecuencia considerables desequilibrios en las cargas así como cuellos de botella en las redes con los retardos (de colas) correspondientemente grandes, hasta llegar incluso a pérdidas de paquetes. La ingeniería de las redes y rutas sigue siendo una tarea compleja, agravada por el aspecto de la fiabilidad y disponibilidad (por ejemplo cambio de ruta en caso de falta).

15 D) Básicamente puede pensarse también en casi todas las combinaciones entre estos principios y también en gran parte han sido ya objeto del debate. Todos estos principios tienen en común que los mismos (a excepción de DiffServ) se apoyan básicamente en rutas y anchuras de bandas reservadas a lo largo de las rutas y dado el caso otros recursos. Incluso un principio DiffServ puro sigue basándose al menos en rutas predeterminadas por protocolos de enrutamiento. Ello implica por lo general un gran coste administrativo para la preparación y la organización (estática) de rutas en la red o bien el correspondiente coste de control elevado para la elección dinámica y la conmutación de las vías. Además, deben mantenerse disponibles en cada nodo de la red equipos de memoria para almacenar las informaciones específicas de rutas y enlaces, que en caso de falta podrían perderse o bien deberían reconfigurarse por otras vías. Incluso en un principio DiffServ puro sigue el transporte las rutas predeterminadas por el protocolo de enrutamiento, que por lo tanto deben ser dimensionadas y vigiladas muy cuidadosamente. No obstante, por lo general no son exactamente previsibles ni todas las oscilaciones en la afluencia de tráfico ni todas las reacciones de los protocolos de enrutamiento cuando sobreviene cualquier suceso en la red.

20 Una tarea de la invención reside en mostrar una vía de cómo servicios que mantienen de manera fiable y eficiente sus exigencias de QoS específicas pueden realizarse de manera sencilla, pragmática y económica en una red de servicios integrados, orientada a paquetes y en particular basada en IP.

25 Esta tarea se resuelve mediante los objetos reivindicados. Se propone una amplia distribución del tráfico en la red. La misma provoca entre otros una QoS bastante equilibrada para todos los servicios y aplicaciones con carácter de Best Effort. La distribución del transporte correspondiente a la invención posibilita también el paso más allá de un Single Routing Domain hasta una solución total global.

30 Un procedimiento para distribuir el transporte en una red de comunicaciones se ha descrito ya en el documento DE 199 23 245 A. Allí se da a conocer un procedimiento de enrutamiento en el que para transmitir la información para cada flujo de datos se elige en cada caso, al menos parcialmente de manera arbitraria, una de varias rutas adecuadas.

35 El documento de patente US 5,953,337 describe un procedimiento para distribuir células ATM de un enlace entre varias rutas existentes entre dos nodos. Mediante la distribución del transporte debe mejorarse la calidad de la transmisión y dificultarse la escucha por parte de terceros.

40 El documento EP 1 133 112 A indica un procedimiento para la distribución de la carga de transporte en una red de comunicaciones en el que la carga de los enlaces y nodos de la red se determina centralmente y se transmiten a los nodos de la red las informaciones de distribución del transporte deducidas a partir de ello.

Todos estos principios tienen en común que los mismos ciertamente distribuyen el transporte existente, pero lo hacen sobre vías más o menos predeterminadas y no dejan prácticamente a los distintos nodos de la red situados en el camino hacia el destino prácticamente ninguna libertad de decisión. Además, no se da a conocer ninguna medida para tener en cuenta posibles exigencias de QoS, con lo que los principios de solución no llegan más allá del citado carácter de Best Effort.

Un aspecto esencial de la invención reside en prescindir de esquemas mentales tradicionales ya asimilados, por ejemplo en el sentido de acometer características como QoS y fiabilidad en su subjetividad y no ligarlos sólo a un ruta, sino definirlos como una característica global de la solución de la red, que gana entonces bastante en cuanto a autonomía y debido a ello también es más económica de operar. Considerada la QoS en el nivel de la red, sólo permite su representación en un funcionamiento sin conexión. Esta consideración correspondiente a la invención se basa entre otros en las siguientes consideraciones:

a) La calidad del servicio (QoS) es un concepto relativo. Incluso cuando se transmite información con la técnica de conmutación por línea, no pueden excluirse pérdidas de datos (por ejemplo debidas a perturbaciones (-> errores de bits) o desplazamiento de tramas). No obstante, tales debilidades pueden tolerarse (por ejemplo en la telefonía digital) o bien son compensadas (técnica de datos) por las correspondientes medidas de seguridad en los mismos niveles (por ejemplo mediante redundancia) o niveles más altos (por ejemplo mediante repeticiones). En definitiva es decisiva la percepción (subjetiva) de la calidad por parte del receptor de la información. La comunicación en tiempo real, interactiva, incluyendo a personas por ejemplo se realiza siempre mediante sus órganos sensoriales (de funcionamiento analógico), que pueden perfectamente funcionar con información incompleta (caso contrario no serían posibles con seguridad la telefonía, sobre todo la móvil, las películas y la televisión en su forma actual). Para el control interactivo de máquinas (por ejemplo el telecontrol de robots) las exigencias son claramente más altas en determinadas circunstancias, con lo que aquí puede ser necesaria una consideración más detallada de cada caso individual. Pero en ningún caso pueden sobrepasarse los límites físicos, por ejemplo en cuanto a los tiempos de recorrido que dependen de la distancia.

Por lo tanto, la QoS no exige necesariamente una garantía absoluta (que en ningún caso existe, incluso cuando se utilicen rutas y reservas), sino el cumplimiento de las exigencias específicas del correspondiente servicio desde el punto de vista del receptor de la información. En la transmisión orientada a paquetes esto se refiere sobre todo al tipo y volumen de las posibles pérdidas de información, retardos fijos y/o variables y a la consistencia en el tiempo (secuencia) de la información. La técnica ATM se basa al respecto en nodos de conmutación y tramos de transmisión dimensionados según las reglas de la estadística y el principio de la transmisión orientada a enlaces con los recursos correspondientemente reservados a lo largo de las rutas, asegurándose la distribución correcta de los recursos a lo largo de las rutas mediante mecanismos de formación de colas y de planificación en los nodos de la red ciertamente potentes, pero en consecuencia también complejos.

b) Las modernas redes (de datos) de alta velocidad funcionan en "Wire Speed" (la máxima velocidad que permite el hardware). Las redes basadas en IP como por ejemplo Internet observan ante todo sólo paquetes de datos y tratan a priori a todos por igual: El paquete que ha llegado el primero se retransmite también el primero y si no hay suficientes recursos de transmisión disponibles, entonces se memorizan (puesta en cola, buffer) primeramente los paquetes y cuando ya no se dispone de más espacio de memoria, se rechazan los paquetes excedentes que llegan (principio Best Effort). Los nodos de red en estas redes, los llamados enrutadores, eran al principio ordenadores, es decir, la funcionalidad completa del análisis y retransmisión de los paquetes de datos estaba implementada en programas de software. Correspondientemente eran estas redes hasta hace poco también comparativamente lentas. Con ayuda de memorias tampón (buffer) correspondientemente dimensionadas y mecanismos adecuados de aseguramiento de datos en las capas de protocolo más elevadas, como por ejemplo TCP, pudo lograrse (ciertamente a menudo con un gran retardo) una transmisión que seguía siendo suficientemente fiable y utilizable de información no crítica en el tiempo.

Los progresos tecnológicos posibilitaron la implementación de funciones de enrutador elementales en hardware (ASICs, FPGAs) y abrieron así el camino hacia una rápida retransmisión de los paquetes de datos y con ello también casi en tiempo real sobre líneas de enlace de alta velocidad. Como elemento retardador queda entonces prácticamente sólo la inevitable resolución de conflictos cuando llegan simultáneamente varios paquetes de datos que se enrutan a través de la misma salida. No obstante, estos retardos son cada vez menos significativos al aumentar la anchura de banda (o mejor: la velocidad) de las líneas de enlace entre los enrutadores, porque entonces los tiempos de espera que implica el conflicto son cada vez más pequeños debido a la salida más rápida de los paquetes de datos. Esto es así en particular cuando diversos flujos de transporte pueden diferenciarse por la correspondiente marca y pueden tratarse de manera diferente en la formación de colas y en la planificación (DiffServ, "priorización").

c) Pese a estos progresos tecnológicos, permanecen intactos no obstante aspectos esenciales relevantes para el servicio, como por ejemplo:

- la agregación de flujos de transporte sobre las rutas en la red, con lo que incluso con un control cuidadoso de los flujos de transporte en las entradas de la red, no pueden predecirse y con ello tampoco evitarse a más profundidad en la red cargas desequilibradas, que perjudican entre otros la QoS, o bien

- 5
- el elevado coste y la larga duración que de ello resulta en una reconfiguración de las rutas en caso de falta, con lo que puede verse perjudicada considerablemente la disponibilidad de la red y servicios para los usuarios.

La red de comunicaciones según la invención incluye según esta nueva consideración, correspondiente a la invención, las siguientes características y funcionalidades (idea básica):

- 10
- funciona sin conexión y orientada a paquetes,
 - ofrece múltiples puertos (ports) de entrada y salida,
 - está compuesta por múltiples nodos de red, enmallados entre sí tal que (por lo general) existen múltiples vías entre distintos puertos de entrada y de salida,
- 15
- contiene mecanismos que, teniendo en cuenta el correspondiente destino (puerto de salida) de los paquetes de datos, pretenden en cada momento (es decir, en lo posible en cada punto de decisión en la red) una distribución lo más uniforme posible de la carga de transporte en la red.

Con ello se eliminan en gran medida, tal como se pretendía, los inconvenientes antes citados de las técnicas de red existentes, que resultan en su utilización como red de servicios integrados, orientada a paquetes, y a la vez se logran las ventajas económicas deseadas:

- 20
- La red ha de ser sin conexión y orientada a paquetes. Así puede utilizarse en particular una red basada en IP, ya que la misma cumple las citadas premisas.
 - Una distribución del transporte más uniforme posibilita una utilización óptima de los recursos con la máxima calidad y con ello con el dimensionado más económico posible. Con ello se logra una solución total económica.
- 25
- Una red que opera sin conexión no necesita ninguna prestación de control para establecer y finalizar los enlaces, ninguna búsqueda de vías, ninguna reconfiguración de vías, ninguna recuperación de rutas en caso de falta, etc. Es en consecuencia fácil de controlar y económica de operar, ya que apenas son necesarias intervenciones administrativas y la red casi es autoorganizada.
 - Se evita por definición la agregación de flujos de transporte con el mismo destino, como consecuencia de la distribución, porque incluso flujos de transporte ya agregados se distribuyen en el curso de su transmisión a continuación de nuevo entre diferentes líneas de la red.
- 30
- En caso de falta, es decir, del fallo de una línea de salida, no es necesaria ninguna reconfiguración costosa de los flujos de transporte afectados por el fallo entre rutas sustitutorias. En lugar de ello es suficiente no dirigir flujos de transporte a la línea que ha fallado. Para eliminar la falta es necesario por lo tanto solamente reducir el grado de distribución. No existe la necesidad de una reconfiguración.
- 35
- Finalmente, la solución es en su efecto conjunto evidentemente pragmática, ya que al eliminarse costosos mecanismos de reconfiguración y priorización se reduce significativamente el coste de configuración para la gestión de la red.

40

Pueden añadirse otras características en distinta configuración y composición. Algunas características y combinaciones de características que implican interesantes ventajas se relacionan a continuación con algunas soluciones alternativas posibles:

45

La distribución del transporte tiene el objetivo de lograr una distribución lo más uniforme posible de la carga de tráfico en la red. La misma puede realizarse con diferente granularidad, por ejemplo en base a flujos de transporte agregados, para cada flujo de transporte individual o en base a los distintos paquetes de datos. La distribución será no obstante tanto más eficiente cuanto más fina sea su granularidad. La decisión sobre la distribución debe realizarse en cada nudo de red "ad hoc" y automáticamente. Como criterio de decisión sirve la información que se suministra con los paquetes de datos, por ejemplo una combinación de dirección de origen y destino, dado el caso junto con otra información, que sirve por ejemplo para la asociación a un determinado flujo de transporte. En una distribución sobre la base de flujos de transporte, toman usualmente todos los paquetes de datos que pertenecen al mismo flujo de transporte la misma vía a través de la red. El efecto de incremento de la calidad que provoca la

50

distribución del transporte de evitar cargas desequilibradas, inclusive la sobrecarga de distintas secciones de la red, se da entonces sobre todo cuando existe suficiente masa estadística o flujos de transporte similares (en particular con similares anchuras de banda).

En el caso de una topología predeterminada de una red con una combinación/enmallamiento (teóricamente) regular (ver figura 1), pueden ajustarse previamente de forma más o menos fija informaciones de vías para la distribución del transporte y el "modelo de ramificación" que de ello resulta en los nodos de la red.

5 En una red real, estructurada (ver figura 3) el enmallamiento es normalmente irregular y más bien incompleto. También se producen repetidamente durante el funcionamiento modificaciones de la configuración de la red o de la topología de la red. En el marco de la invención se realiza para ello una actualización (Update) flexible de las vías y modelos de ramificación posibles (según necesidades o regularmente) y/o el nodo deduce a partir de una información modificada sobre las vías nuevos modelos de ramificación. Como mecanismos para distribuir la información sobre las vías puede pensarse en los correspondientes protocolos del entorno de Internet (protocolos de enrutamiento como por ejemplo OSPF, BGP) o variantes/desarrollos posteriores derivados de los mismos. Evidentemente pueden prescribirse también estas informaciones mediante un control de red (siempre de este tipo) o bien un sistema de gestión de la red.

15 En los modelos de ramificación pueden considerarse a la vez otros criterios, como por ejemplo distintas anchuras de banda, distintas distancias al punto de destino, costes de las vías, etc. en los algoritmos para la elección concreta de la vía. Así por ejemplo puede transmitirse, en la distribución por paquetes entre un enlace STM-4 y un enlace STM-1, mediante la correspondiente ponderación, sólo cada quinto paquete sobre el enlace STM-1. La carga puede distribuirse más finamente aún cuando adicionalmente se tienen en cuenta longitudes de paquetes individuales. Una ponderación de los enlaces según los correspondientes criterios es también ventajosa para evitar en una red enmallada compleja la formación de bucles de vías o para limitar por ejemplo oscilaciones en los retardos de los paquetes. Los retardos distintos de los paquetes por diferentes vías pueden dar lugar a una modificación de la secuencia de los paquetes. Ésta se restablece a la salida de la red (resecuenciado), por ejemplo cuando una aplicación lo exige.

25 El carácter de Best Effort de la distribución del transporte según la invención, en el que se ven perjudicados servicios y aplicaciones de carga creciente (en función de sus características y exigencias, de forma más o menos perceptible), puede mejorarse significativamente cuando la carga total de transporte en la red se limite en función de la capacidad efectiva de la red.

30 Por un lado podría entonces considerarse adicionalmente la anchura de banda de los distintos accesos a la red tanto por el lado de entrada como también por el de salida y tenerla en cuenta tanto separadamente como también en el conjunto. Partiendo de las características de transporte estadísticas de los diversos servicios y aplicaciones y basándose en una topología de la red y en la capacidad y potencia de los nodos de red y líneas de enlace, se dimensiona entonces la red (Networking Dimensioning), bajo la hipótesis de determinados modelos de comportamiento de los abonados o bien de determinadas características del transporte que de ello resultan en los accesos a la red, tal que bajo estas condiciones marginales sólo se sobrepasen determinados valores límite de entre los factores que determinan la QoS, como por ejemplo pérdida de paquetes (Loss Rate), retardo de paquetes (delay) u oscilación del retardo de paquetes (Delay Jitter) con una probabilidad estadística suficientemente pequeña y claramente definible.

40 Por otro lado, podría limitarse el tráfico en una red determinada tal que se cumplan las correspondientes condiciones marginales. Para ello se parametrizan correspondientemente todas las relaciones de comunicación y flujos de datos en la red, se detectan individualmente con cada nueva presencia y, en función de la actual situación de carga en la red, se admiten o se rechazan (Admission Control).

45 Así da a conocer el documento EP 1 119 216 A una función de acceso central que pone a disposición de los nodos en el acceso a la red, teniendo en cuenta la capacidad total de transmisión de la red de comunicaciones, una capacidad disponible para transmitir flujos de transporte a la red de comunicaciones. No obstante, entonces no se detectan posibles cuellos de botella que resultan debido a la coincidencia de flujos de datos de distintas entradas en la misma salida.

50 No obstante, ambos mecanismos, sobredimensionamiento y también limitación de acceso, no son en la práctica de por sí practicables ni económicos. Así por ejemplo no puede defenderse económicamente un sobredimensionamiento de las redes para mantener las exigencias de servicios más sensibles teniendo en cuenta los servicios menos sensibles, ni tampoco es practicable un control de admisión (Admission Control) para una gran parte de las aplicaciones tradicionales y también futuras de Internet diseñadas para un entorno Best Effort.

55 Una QoS diferenciada y adaptada a las necesidades del correspondiente servicio se realiza según la invención mediante una diferenciación y división en distintas clases de transporte que se tratan, en particular se priorizan, de manera diferente. La cantidad de clases de tráfico es de al menos dos. Para el tratamiento en los nodos de la red (es decir, en los Queuing Points o puntos de formación de colas), se da preferencia a una estricta priorización, ya que los procedimientos alternativos (p.e. Weighted Fair Queuing, WFQ, espera equitativa ponderada), que también garantizan recursos a clases de servicio de prioridad inferior en cualquier circunstancia, perjudican cuando la carga es elevada la mayoría de las veces al tráfico de mayor prioridad y son por naturaleza mucho más complejos que la estricta priorización, en la que no se transmite tráfico de inferior prioridad mientras esté pendiente de transmitir tráfico de mayor prioridad.

En una red de comunicaciones del tipo antes descrito puede tener la característica de la priorización las siguientes aplicaciones y configuraciones ventajosas: Todos los flujos de datos se reparten en función de sus necesidades en las correspondientes clases de prioridad. La clase más baja se tiene en cuenta en el dimensionamiento de la red (en el marco de la cantidad de tráfico total esperada) y se trata básicamente según el principio del Best Effort. Para todas las relaciones de comunicación o bien flujos de datos en clases de prioridad superiores, se realiza una prueba de admisión (Admission Control) en la entrada de la red (en la dirección de entrada) y a la salida de la red (en la dirección de salida). Para ello se anuncian y evalúan estos flujos de datos en estos dos puntos con los correspondientes parámetros (por ejemplo velocidad media de datos y/o paquetes, velocidad de punta, etc.). Las decisiones a la entrada y la salida son independientes entre sí y sólo cuando ambas decisiones son positivas, se admite el flujo de datos. Como criterio de decisión puede servir por ejemplo un valor de umbral, que se determina en función de la capacidad del puerto, de la capacidad total de la red, de la calidad deseada en cuanto a posibles retardos y pérdidas de paquetes, etc., de la correspondiente clase de prioridad y dado el caso también de otros criterios. También puede pensarse en que existan para cada clase varios valores de umbral en base a distintos parámetros de evaluación, que han de cumplirse todos individualmente o en dependencia uno del otro.

Mediante la prueba de admisión se limita por un lado la cantidad total de tráfico de una determinada clase de prioridad en la red y por otro lado también la correspondiente cantidad de tráfico en cada puerto individual de entrada y salida. Mediante la distribución uniforme del tráfico en la red (idealmente sobre la base de paquetes) y el correspondiente tratamiento preferente, encontrará este tráfico, si los valores de umbral están correctamente ajustados, siempre recursos suficientes (capacidad libre de enlaces, memoria tampón) para poder cumplir con los valores límites de retardo y también de pérdidas correspondientes a sus exigencias de calidad. La red puede entonces funcionar perfectamente a plena carga y económicamente, porque toda la anchura de banda no utilizada por tráfico de mayor prioridad puede ser utilizada en todo momento por tráfico de inferior prioridad.

Idealmente el tratamiento preferente viene provocado por la estricta prioridad, es decir, caso necesario desplazando por completo el tráfico de inferior prioridad. La estricta prioridad aporta un retardo mínimo y pérdidas mínimas. Además, éste es un mecanismo de prioridad claramente más sencillo que por ejemplo el procedimiento Weighted Fair Queuing (espera equitativa ponderada) conocido por el estado de la técnica.

Se vigila el cumplimiento de los parámetros de transporte anunciados de los distintos flujos de datos, porque en el marco de la distribución del tráfico también puede perjudicar considerablemente todo el tráfico en toda la red un único flujo de datos que "ordinariamente se sobrepasa". La función de vigilancia (Traffic Enforcement o aplicación de las normas del tráfico, Policing o vigilancia) puede diseñarse ventajosamente relativamente insensible y con ello económica, porque un desbordamiento leve, aleatorio y de breve duración puede ya compensarse mediante la distribución del tráfico correspondiente a la invención.

La función de vigilancia se utiliza ventajosamente sobre los flujos de datos individuales, tal como los mismos se han anunciado. Como alternativa está prevista una agregación de cualquier índole por cada puerto, en la que solamente se comprueba un límite total y como reacción a un desbordamiento del límite total dentro de un grupo agregado se interviene sin elección y dado el caso transversalmente a través de todos los flujos de datos incluidos. Pueden utilizarse básicamente cualesquiera y naturalmente también todos los mecanismos conocidos y pertinentes (por ejemplo Leaky Bucket o cola limitada) y lo mismo rige también en cuanto a las posibilidades de reacción (rechazo de paquetes, marcado de paquetes, desconexión/bloqueo de flujos de datos,...). Un marcado puede consistir en determinadas circunstancias también en una traslación de los paquetes que vulneran el acuerdo (o mejor de todo el correspondiente flujo de datos) a una clase inferior o la clase Best Effort.

El principio de la distribución del tráfico (en particular cuando la misma se realiza a nivel de paquetes) puede utilizarse también muy ventajosamente para mejorar la fiabilidad y disponibilidad de la red y de los servicios. Para ello es suficiente que los nodos de red, al detectarse una falta (por ejemplo fallo de un enlace, que el nodo contiguo ha fallado) extraigan el o los correspondientes enlaces del haz de ramificación y se continúe la distribución sólo con los restantes enlaces. La decisión puede tomarse cuando se detecta el estado de falta sin demora y autónomamente y cuando la detección es externa, cuando se dispone de la información. Si el dimensionado de la red es suficiente, da lugar una tal reacción en el peor de los casos a un desplazamiento algo mayor del tráfico de Best Effort, pero no produce reducción alguna de la calidad del tráfico de alta prioridad. Esta ventaja especialmente favorable se logra en la combinación correspondiente a la invención de manera perceptible muy sencilla, pragmática y económicamente, ya que - respecto al estado de la técnica descrito al principio - se utilizan mecanismos comparativamente sencillos.

Una variante muy interesante de la idea básica resulta cuando el procedimiento de la distribución del tráfico sólo se utiliza para la(s) clase(s) de transporte de más alta prioridad. El tráfico Best Effort elige su enrutamiento según los principios usuales hoy día en Internet, mientras que el tráfico de más alta prioridad se distribuye uniformemente en la red y casi llena la misma desde abajo. El tráfico Best Effort prácticamente nada en un mar con un grado de llenado moderado del tráfico de mayor prioridad y se ve desplazado cada vez más cuando hay "marea alta". Una ventaja interesante de esta variante consiste en que la solución de QoS puede implantarse complementariamente en redes existentes, mientras que los mecanismos existentes continúan existiendo sin modificación.

El principio presentado puede utilizarse también en una red basada en células, por ejemplo en una red ATM.

La fiabilidad de la red mejora aún más mediante mecanismos de autovigilancia en el enrutador, en particular en el marco del procedimiento de distribución.

5 Para aumentar la fiabilidad en toda la red, se utiliza opcionalmente además un tipo de mecanismos de feedback rápido entre los enrutadores, que permite por ejemplo, en problemas que se han presentado en algún lugar "flujo abajo", distribuir también de otra manera y ya tempranamente el tráfico "flujo arriba".

El control de admisión (Admission Control) está configurado dado el caso tal que cuando hay "overbooking" de una clase de tráfico de más prioridad, ofrece al usuario automáticamente la clase más inferior.

10 La función de resecuenciado (Resequencing) se prevé en general, por ejemplo como función estándar, a la salida de la red. Ventajosamente pueden seguirse utilizando sin variaciones todas las aplicaciones TCP actuales, en las que usualmente no está implementada la función de resecuenciado.

La invención se describirá más en detalle a continuación en base a otros ejemplos de ejecución representados en las figuras.

15 figura 1 una configuración para realizar el procedimiento correspondiente a la invención, configurada a modo de ejemplo para mostrar el principio básico de la distribución del tráfico en una red regular, enmallada a través de diversos escalones (niveles de red),

figura 2 una configuración para realizar el procedimiento correspondiente a la invención, configurada como una de las muchas formas de ejecución posibles más concretas de la invención,

20 figura 3 una configuración para realizar el procedimiento correspondiente a la invención, configurada como red (de datos) real, estructurada, con enmallado irregular y más bien incompleto.

En las figuras se han representado aquellas vías de transmisión a lo largo de las que se transmiten distribuidos los correspondientes flujos de transporte, señalados mediante respectivas puntas de flecha orientadas, con lo que a la vez se muestra la dirección de la transmisión.

Una ejecución de la invención aporta una red de comunicaciones sin conexión, orientada a paquetes,

- 25
- con al menos dos clases de transporte diferentes, de las cuales una se trata como simple tráfico Best Effort, mientras que las otras, de las que al menos hay una, por el contrario (y cuando hay otras varias preferiblemente también entre sí) se priorizan estrictamente,
- 30
- con nodos de la red, desde los cuales se distribuye el tráfico individual y autónomamente con el objetivo de una distribución uniforme de la carga de transporte, preferiblemente en base a paquetes según determinadas reglas entre todas o al menos un conjunto de vías en dirección a su destino (salida de la red),
 - en el que los nodos de la red intercambian/difunden la información a través de rutas disponibles mediante los correspondientes protocolos,
 - en el que los nodos de la red en caso de falta adaptan sin demora y de manera autónoma sus modelos de distribución (del transporte),
- 35
- tal que para los flujos de datos de una clase de tráfico más elevada, de las que menos hay una, realiza preferiblemente en cada entrada y salida un control de admisión sobre la base de determinados parámetros de transporte (que por ejemplo ya no admiten más tráfico de esta(s) clase(s) del tráfico a partir del momento en que se llega una carga total $x\%$ ($x\%$, $(x+d)\%$, $(x+nd)\%$) de la capacidad del puerto,
- 40
- tal que sólo acepta un flujo de datos de la clase de datos superior, de las que al menos hay una, cuando ambos controles de admisión (en el puerto de entrada y en el puerto de salida, independientemente entre sí) han decidido positivamente,
 - tal que en cada entrada se vigilan los parámetros de transporte anunciados de los flujos de datos de la(s) clase(s) de tráfico más alta(s) y dado el caso se interviene con medidas adecuadas, y
- 45
- tal que en cada salida ofrece una función de resecuenciado para su utilización opcional por (todos) los flujos de datos.

50 Otra ejecución de la invención se representa en la red de comunicaciones 100 mostrada en la figura 1. En la red 100 se transmite distribuyéndolo, partiendo de un nodo de transmisión A configurado como nodo de entrada, al menos un flujo de transporte hasta un nodo de transmisión B configurado como nodo de salida. Entonces se realiza la transmisión distribuida en la red 100 tal que se transmite distribuidamente la parte recibida en cada caso por la mayoría de los nodos de transmisión de la red del o de los flujos de transporte a exactamente dos nodos de

transmisión siguientes. Solamente ambos nodos de transmisión dispuestos inmediatamente delante del nodo de salida B transmiten, a falta de vías restantes alternativas sin distribución por toda la red, directamente hasta el nodo de salida B, para que en este las partes transmitidas distribuidas puedan ensamblarse de nuevo para formar los flujos de transporte iniciales. Al nodo de salida B llegan entonces en este caso los flujos de transporte distribuidos desde dos direcciones diferentes. Preferiblemente se consigue en el nodo de salida mediante la correspondiente función de resecuenciado (Resequencing Function) RF un reordenamiento de las partes recibidas de los flujos de transporte hasta su secuencia inicial. Así pueden transmitirse en la red 100 también flujos de transporte entre aplicaciones que deben mantener en la transmisión la secuencia inicial existente antes de la transmisión, sin que para ello sea necesaria una modificación y/o adaptación de las aplicaciones.

En la figura 2 se representa una ejecución alternativa de la invención, configurada como red de comunicaciones 200. En la red 200 se transmite distribuido, partiendo de un nodo de transmisión C configurado como nodo de entrada, al menos un flujo de transporte a un nodo de transmisión D configurado como nodo de salida. Entonces y contrariamente a en la red 100, sólo se utiliza una parte de las vías de transmisión de la red 200 para la transmisión distribuida entre ambos nodos C y D. Esto se basa en que en la red 200 no se utiliza cualquier vía desde el nodo C hasta el nodo D para la transmisión distribuida, sino sólo aquellas vías que son especialmente adecuadas para ello. Éstas son en el presente ejemplo aquellas vías que teniendo en cuenta la topología de la red 200 no dan un rodeo demasiado grande a través de la red 200, con lo que se verían afectadas todas de un retardo en la transmisión (Transmission Delay) que preferiblemente se encuentra dentro de una gama de tolerancia prescrita y relativamente reducida. Las vías que ciertamente conducen igualmente desde el nodo C hasta el nodo D, pero cuyo retardo de la transmisión se desvía demasiado de la gama de tolerancias, no son adecuadas para esta transmisión distribuida descrita a modo de ejemplo.

Los flujos de transporte se transmiten distribuidos preferiblemente en cada nodo de transmisión entre los nodos C y D por paquetes hasta los correspondientes nodos siguientes. Además se realiza la distribución teniendo en cuenta el grado de carga de las vías restantes que quedan en cada caso y/o la longitud de los paquetes transmitidos en cada caso a las vías restantes. Esto da como resultado una transmisión distribuida muy uniforme entre los nodos C y D. Cuando se utiliza este principio entre todos los nodos de entrada y de salida de la red 200, entonces se llena la red 200 desde abajo con una carga de base de transporte distribuido, siendo el grado de carga de los nodos de transmisión y de las vías de transmisión para toda la red 200 aproximadamente igual. Ninguna parte de la red 200 permanece, comparada con el resto de la red 200, durante largo tiempo con sobrecarga.

Si en la red 200 se reparte el tráfico que llega en dos clases de tráfico, dándose preferencia al tráfico de más alta prioridad, que se transmite distribuido, y si se limita el volumen del tráfico de más alta prioridad con ayuda de pruebas de admisión AC y vigilancia del tráfico TE, entonces puede transmitirse el tráfico de más prioridad prácticamente con carácter de tiempo real en la red 200. Para el tráfico de inferior prioridad se alcanza el carácter de Best Effort, descendiendo su calidad al aumentar el tráfico de más alta prioridad y a la inversa. Para la priorización del transporte es totalmente suficiente una estricta priorización. Ésta se caracteriza en comparación con otros mecanismos de priorización conocidos por su especial simplicidad, con lo que su realización puede llevarse a cabo de manera especialmente ventajosa y económicamente en los nodos de transmisión.

Otros aspectos de la invención se representan en la red de comunicaciones 300 de la figura 3. La red 300 incluye varios nodos de transmisión 301-315, estando configurados los nodos de transmisión 301-307 como nodos de entrada y/o de salida. Se transmite distribuidamente, partiendo del nodo de transmisión E configurado al menos como nodo de entrada 301, al menos un flujo de transporte al nodo de transmisión F, configurado al menos como nodo de salida 304, como sigue:

desde el nodo	a través de la vía	hasta el nodo
301	320	308
	321	309
308	322	311
	323	313
309	324	310
	325	314
311	327	303
	329	312
313	330	312
	331	315
310	326	311
	327	313
314	333	313
	334	305
303	329	312
305	335	315
312	336	304
315	337	304

Puede observarse con claridad que desde cada nodo de transmisión entre los nodos E y F del que parte más de una vía restante hasta el nodo de salida F, el tráfico enviado a éste se transmite distribuido entre al menos dos vías restantes.

- 5 Además se representa en la figura 3 cómo cambia el modelo de ramificación cuando falla la vía de transmisión 325. Como consecuencia se borra en el nodo precedente 309 la vía de transmisión 325 de los haces de ramificación memorizados para la transmisión en dirección hacia el nodo de salida F. Por lo tanto ya no se envía más tráfico a los nodos de transmisión 314 y 305 posconectados a la vía de transmisión 325. Los nodos de transmisión 313, 312 y 315, igualmente posconectados, siguen recibiendo no obstante tráfico, que se transmite a través de otras vías que no incluyen la vía de transmisión 325 que ha fallado del modelo de ramificación. Tras el fallo de la vía de transmisión 325 se modifica así en el citado modelo de ramificación como sigue:
- 10

desde el nodo	a través de la vía	hasta el nodo
301	320	308
	321	309
308	322	311
	323	313
309	324	310
	325	314
311	327	303
	329	312
313	330	312
	331	315
310	326	311
	327	313
314	333	313
	334	305
303	329	312
305	335	315
312	336	304
315	337	304

5 Puede observarse claramente que el fallo de la vía de transmisión 325 da lugar solamente a una reducción del modelo de ramificación, pero no resulta necesaria ninguna reconfiguración de la red 300. En particular sigue siendo alcanzable el nodo de salida F a través de dos vías. Es evidente que la invención tiene un alto grado de robustez frente a fallos de nodos de transmisión o vías de transmisión de manera extremadamente pragmática. Cuanto mayor es el grado de enmallamiento de la red de comunicaciones, tanto más vías existen entre los nodos de entrada y los nodos de salida, con lo que incluso cuando falla una gran parte de la red, en la mayoría de los casos sigue existiendo al menos una vía a través de la cual pueden seguir transmitiéndose flujos de transporte. Tiene lugar una interrupción total cuando la red de comunicaciones falla de manera más o menos completa. Pero en este caso 10 ayudaría poco incluso la costosa reconfiguración de las vías conocida por el estado de la técnica. En todo caso la misma sería posible cuando durante el funcionamiento normal hubiera vías clasificadas como inadecuadas y todavía susceptibles de funcionar. En este caso sería posible, con un fallo al menos parcial del modelo de ramificación existente, una reconfiguración de los nodos de transmisión todavía susceptibles de funcionar para formar un nuevo modelo de ramificación menos óptimo. Las transmisiones interrumpidas como consecuencia del fallo pueden 15 entonces reanudarse tras la reconfiguración, siempre que se hayan encontrado vías alternativas.

Subrayemos que la descripción de los componentes de la red de comunicaciones unificada relevantes para la invención no ha de entenderse básicamente de modo limitativo. Para un especialista competente queda claro en particular que los conceptos utilizados han de entenderse funcional y no físicamente. Así pueden realizarse los componentes también en parte o por completo en software y/o distribuidos entre varios equipos físicos.

REIVINDICACIONES

1. Procedimiento para transmitir flujos de transporte en una red de comunicaciones (100, 200, 300) sin conexión, orientada a paquetes, que incluye un conjunto de nodos de transmisión (301-315), que están conectados entre sí tal que existe un conjunto de vías entre los nodos de transmisión,
- 5
- estando configurada al menos una parte de los nodos de transmisión como nodos de entrada (301-307, A, C, E) y/o como nodos de salida (301-307, B, D, F) de la red de comunicaciones y
 - enviándose los flujos de transporte desde los nodos de entrada hasta los nodos de salida,
- con las siguientes etapas:
- los flujos de transporte se dividen en al menos dos clases de transporte, al menos uno sin prioridad y al menos uno con prioridad,

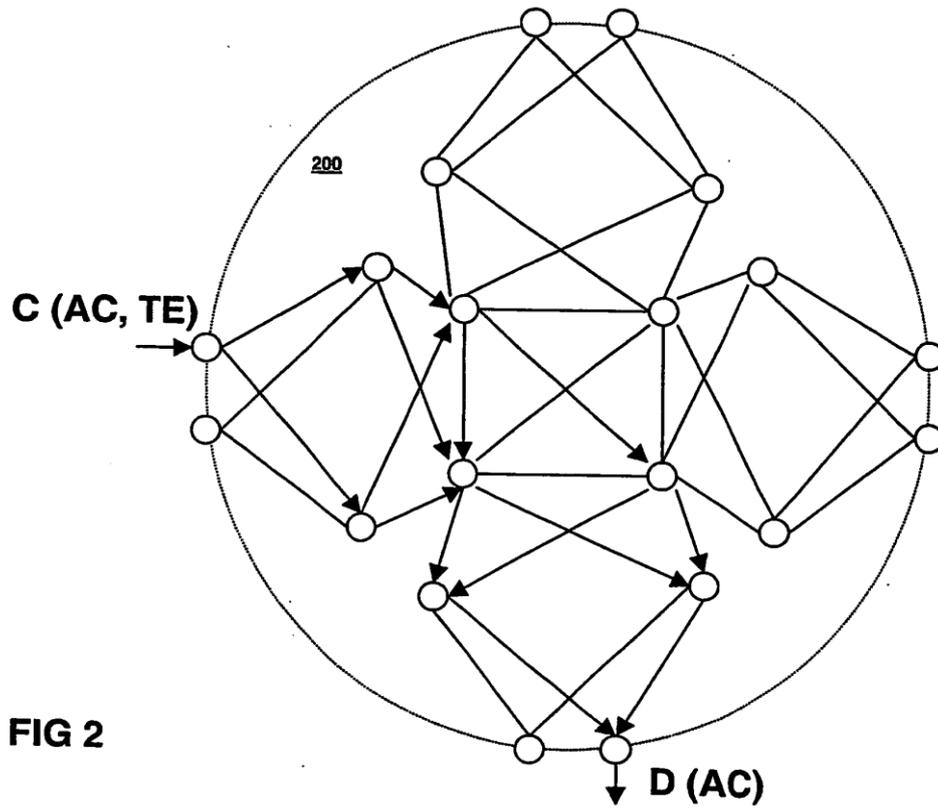
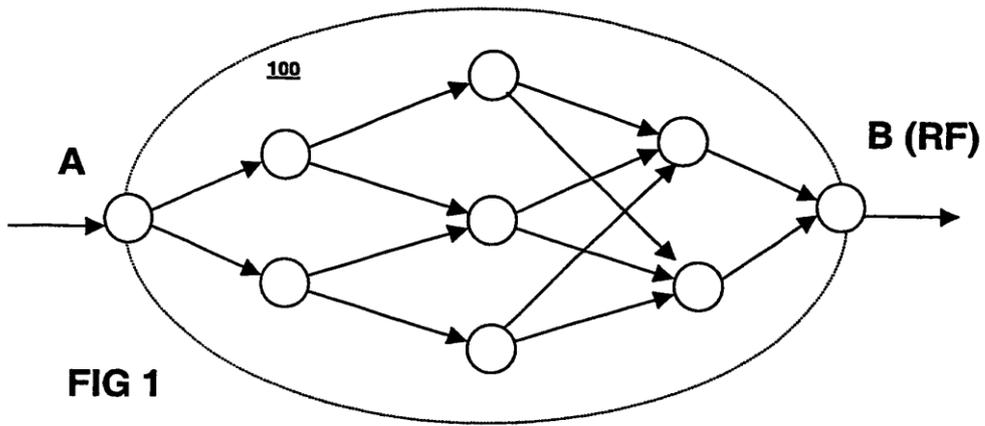
10

 - el flujo entrante de transporte prioritario en la red de comunicaciones se limita con ayuda de una prueba de admisión (AC), que se realiza para cada flujo de transporte prioritario en el correspondiente nodo de entrada y en el correspondiente nodo de salida,
 - para al menos uno de los nodos de transmisión, que recibe al menos una parte de los flujos de transporte enviados (301, 303-305, 308-315), se determina más de una vía restante que parte del mismo hasta al

15

 - desde este nodo de transmisión se transmiten distribuidos los flujos de transporte enviados a este nodo de salida a través de al menos dos de las vías residuales determinadas.
2. Procedimiento según una de las reivindicaciones precedentes,
- 20
- en el que desde al menos uno de los nodos de transmisión para al menos un flujo de transporte hasta un determinado nodo de salida, la distribución se realiza tal que sus paquetes se transmiten distribuidos.
3. Procedimiento según una de las reivindicaciones precedentes,
- en el que desde al menos uno de los nodos de transmisión para al menos un nodo de salida la distribución se consigue tal que los flujos de transporte se transmiten distribuidos hasta este nodo de salida, mientras que los correspondientes paquetes de cada flujo de transporte se distribuyen sin distribuir.
- 25
4. Procedimiento según una de las reivindicaciones precedentes,
- en el que desde al menos uno de los nodos de transmisión para al menos un nodo de salida se logra la distribución tal que se transmiten distribuidas distintas agregaciones de flujos de transporte, mientras que los correspondientes flujos de transporte de cada agregación se transmiten sin distribuir.
5. Procedimiento según una de las reivindicaciones precedentes,
- 30
- en el que las vías restantes se determinan teniendo en cuenta al menos un haz de ramificación, en el que están memorizadas las vías residuales adecuadas para la distribución de los flujos de transporte.
6. Procedimiento según la reivindicación 5,
- en el que una vía restante ciertamente adecuada, pero que ha fallado y/o defectuosa, se señala correspondientemente en los haces de ramificación o bien se borra de los haces de ramificación.
- 35
7. Procedimiento según una de las reivindicaciones precedentes,
- en el que se determinan aquellas vías restantes para una transmisión distribuida que en cuanto a su anchura de banda, su distancia al nodo de partida, sus costes y/o su grado de carga actual, se encuentran dentro una gama de tolerancia predeterminada.
8. Procedimiento según una de las reivindicaciones precedentes,
- 40
- en el que la distribución de los flujos de transporte se realiza teniendo en cuenta el correspondiente grado de carga de las vías restantes adecuadas y/o el correspondiente volumen de las partes del flujo de transporte ya transmitidas por las distintas vías restantes.
9. Procedimiento según la reivindicación 8,
- 45
- en el que para determinar las partes del flujo de transporte se incluye la longitud de los correspondientes paquetes de datos transmitidos.

10. Procedimiento según una de las reivindicaciones precedentes,
 en el que las clases de transporte con prioridad se transmiten estrictamente priorizadas en los nodos de transmisión.
- 5 11. Procedimiento según una de las reivindicaciones precedentes,
 en el que los flujos de transporte con prioridad se transmiten distribuidos.
12. Procedimiento según una de las reivindicaciones precedentes,
 en el que ambas pruebas de admisión se realizan independientemente entre sí y el flujo de transporte analizado sólo se admite cuando ambas pruebas de admisión dan resultado positivo.
- 10 13. Procedimiento según una de las reivindicaciones precedentes,
 en el que se vigila si los flujos de transporte de mayor prioridad cumplen con al menos un parámetro de transporte predeterminado.
14. Procedimiento según la reivindicación 13,
 en el que se toleran desbordamientos del parámetro de transporte que son compensados correspondientemente mediante la distribución del tráfico.
- 15 15. Procedimiento según una de las reivindicaciones 13 ó 14,
 en el que cuando tienen lugar desbordamientos del parámetro de transporte que no son compensados correspondientemente por la distribución del tráfico, se trasmite la parte del tráfico que desborda el parámetro de transporte sin prioridad.
- 20 16. Procedimiento según una de las reivindicaciones 13 a 15,
 en el que al menos para una parte de los puertos (ports) de los nodos de acceso a través de los que se transmiten los flujos de transporte priorizados en la red de comunicaciones, se comprueba si se cumple con al menos un límite total que debe transmitirse como máximo en la red de comunicaciones.
- 25 17. Procedimiento según una de las reivindicaciones 13 a 16,
 en el que al menos para una parte de los flujos de transporte se comprueba individualmente para cada flujo de transporte si se cumple el parámetro del transporte prescrito.
- 30 18. Procedimiento según una de las reivindicaciones precedentes,
 en el que al menos para una parte de los nodos de salida, con ayuda de la correspondiente función de resecuenciado (Resequencing Function) RF asignada, al menos para los flujos de transporte transmitidos distribuidos, se restablece la secuencia inicial tal como la existente antes de su transmisión en la red de comunicaciones.
19. Red de comunicaciones (100, 200, 300) que incluye medios equipados para realizar todas las etapas de un procedimiento según una de las reivindicaciones de procedimiento precedentes.
20. Dispositivo que incluye
- 35
 - medios equipados para realizar aquellas etapas de un procedimiento según una de las reivindicaciones de procedimiento precedentes que logra el dispositivo,
 - medios que están equipados para realizar interacciones prescritas según el procedimiento del dispositivo con otros dispositivos que realizan las restantes etapas del procedimiento.
21. Dispositivo según la reivindicación precedente,
- 40 configurado como producto de programa de ordenador y cuyos medios están configurados como códigos de programa, que están realizados para ejecutar el procedimiento mediante al menos un procesador.



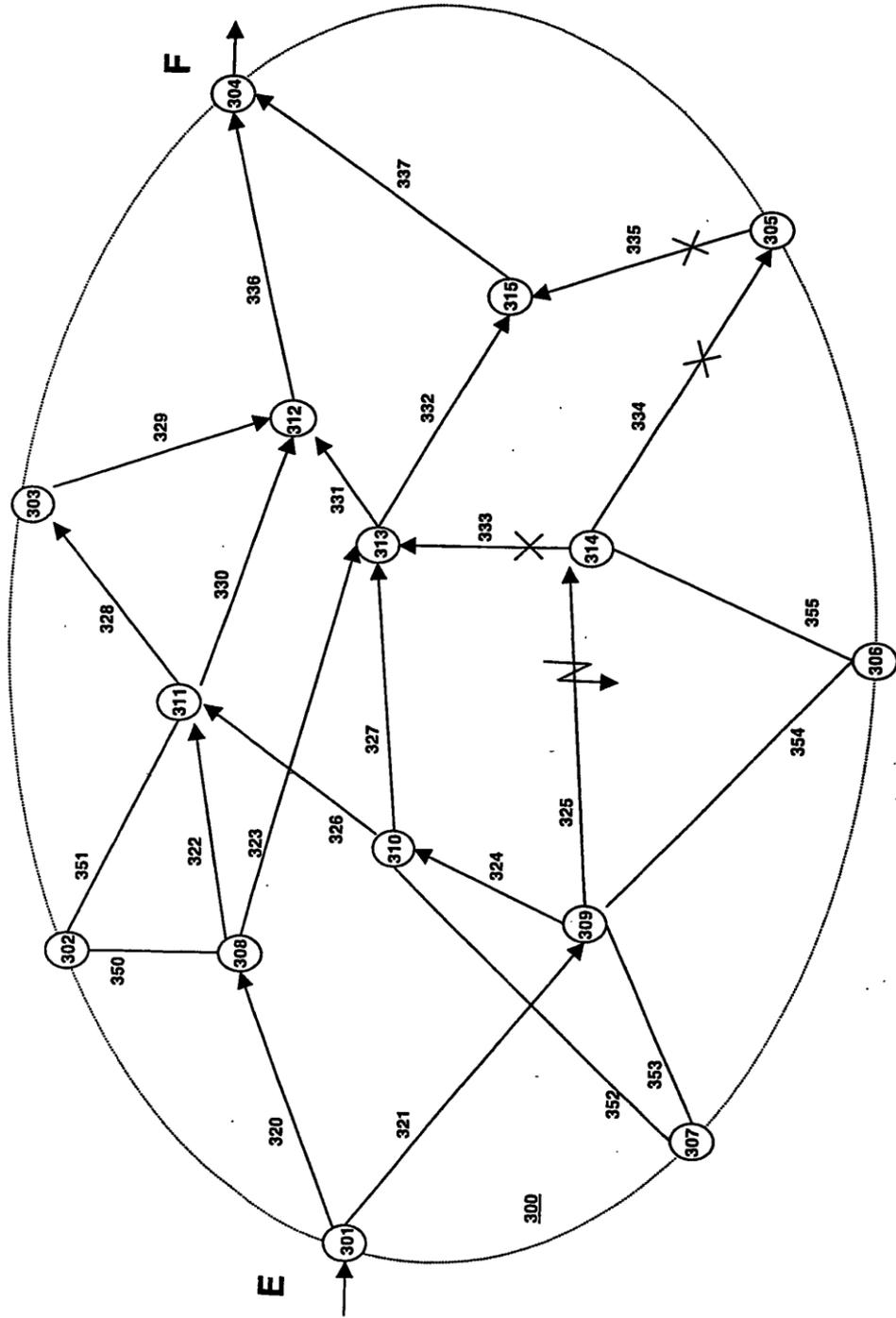


FIG 3