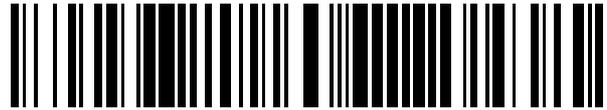


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 671 627**

51 Int. Cl.:

H04L 12/26 (2006.01)

H04L 29/06 (2006.01)

G06F 15/173 (2006.01)

G06F 9/50 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **21.06.2013** **E 13305854 (5)**

97 Fecha y número de publicación de la concesión europea: **28.03.2018** **EP 2816775**

54 Título: **Sistema y procedimiento IMS en nube elástico y escalable**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
07.06.2018

73 Titular/es:

ALCATEL LUCENT (100.0%)
Site Nokia Paris Saclay Route de Villejust
91620 Nozay, FR

72 Inventor/es:

PIANESE, FABIO y
JANSSENS, NICO

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 671 627 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

Sistema y procedimiento IMS en nube elástico y escalable

La presente invención está relacionado con un procedimiento y un sistema IMS en nube elásticos y escalables y a un procedimiento IMS en nube elástico y escalable.

5 Una tendencia actual en la evolución de los sistemas de telecomunicaciones ha sido la sustitución de elementos funcionales implementados en hardware o como máquinas autónomas con funciones lógicas ejecutadas por componentes software. La arquitectura actual de los sistemas telefónicos basados en la especificación del Subsistema Multimedia de Protocolo Internet (IMS), especifica las interconexiones y protocolos lógicos para la interoperabilidad de dichas funciones software separadas que cooperan para suministrar los servicios requeridos. La arquitectura IMS ha, por tanto, "virtualizado" el entero sistema telefónico con independencia de un implementación hardware concreta.

10 Sin embargo, la forma en que el IMS actualmente se implementa y gestiona sigue empleando las prácticas normales de los operadores de telecomunicaciones. Aunque ya no son estrictamente necesarias, las funciones lógicas todavía se implementan y se ejecutan en plataformas hardware separadas, ya sea en máquinas independientes o en hojas / núcleos independientes de máquinas compartidas. Las principales razones de estas elección es el temor de interferencias no deseadas entre piezas diferentes de software y la costumbre de asignar máquinas individuales a una función bien definida.

15 La Figura 1 ilustra una forma de realización ejemplar de un sistema de la técnica anterior. En adelante, el término "trabajadora" se utilizará para indicar una entidad lógica programable que puede ejecutarse en unos recursos informáticos en red sobre los cuales se ha instanciado una función o servicio IMS. Por ejemplo, la Función de Control de Sesión de Llamada Proxy (P-CSCF) puede ser instanciada sobre una entidad lógica programable formando de esta manera una trabajadora P-CSCF.

20 De nuevo con referencia a la Figura 1, una P-CSCF y una trabajadora de Función de Control de Sesión de Llamada de Servicio (S-CSCF) son desplegadas para soportar un establecimiento de llamada básica de soporte entre dos piezas de un equipo de usuario (UE).

25 Cada UE 1, 2 se registra por medio de la trabajadora 3, 4 P-CSCF con la trabajadora 5 de S-CSCF, que almacena la dirección del UE en un almacenamiento de datos (Servidor de Abonado Inicial - HSS) 6. Cuando un abonado que llama inicia una sesión, su UE envía una solicitud a la trabajadora P-CSCF, la cual, a su vez, la envía a la S-CSCF para determinar el emplazamiento del llamado. A continuación, la solicitud es transmitida de nuevo a la trabajadora P-CSCF la cual, a su vez, la transmite al destinatario de la llamada.

30 Al establecer una llamada entre dos UEs, las solicitudes SIP relacionadas y las respuestas siguen la misma vía fijada. Para ser precisos, los servidores IMS con memoria del estado de la llamada, por ejemplo la P-CSCF, la Función de Control de Sesión de Llamada Interrogante (I-CSCF) y las instancias S-CSCF pueden marcar mensajes con las cabeceras ROUTE y VIA para garantizar todas las solicitudes relacionadas (sesión) y la visita de respuestas de las mismas instancias de servidor - las cuales actualmente se adaptan al estado de las sesiones afectadas. Aunque la implementación de dichas "sesiones adheridas" pueden tener ciertas ventajas, presenta diversos inconvenientes. Por ejemplo, al retroceder en la escala una agrupación de trabajadoras IMS con memoria del estado, el agente de control que dirige esta acción debe accionar una trabajadora hasta un estado de ejecución segura, en el que las sesiones en marcha han sido establecidas o interrumpidas, antes de que se pueda detener con seguridad esa trabajadora. Esto puede llevarse a cabo esperando hasta que una trabajadora haya completado todas las sesiones en marcha, o transfiriendo el procesamiento de las sesiones en marcha a otras instancias de trabajadoras. Así mismo, al suprimir la escala de una agrupación de trabajadoras IMS con memoria del estado, las nuevas trabajadoras solo volverán a ser utilizadas cuando se establezcan nuevas sesiones. Esto provoca la pérdida del equilibrio de las cargas de trabajo lo que, a su vez, puede provocar que un servicio de vigilancia concluya falsamente que la nueva trabajadora se está comportando de manera incorrecta.

35 El documento titulado "Flujo abierto virtualizado de estructura en red" autoría de Hideyuki Shimonishi: Shuji Ishii; y publicado el 19.4.2010, describe una arquitectura para unas infraestructuras en red virtualizadas, que incluye un análisis de los cuatro requisitos fundamentales, que son apertura, virtualización, modularización y programabilidad, y en base a dichos requisitos, el documento propone una arquitectura y un modelo para un OS de red. Un inconveniente de la solución conocida, por tanto, es que los recursos informáticos en red no son utilizados de manera eficiente. Por ejemplo, si en una red IMS no hay situadas muchas llamadas de usuario, la mayoría de los recursos pueden estar inactivos, mientras que podrían haber sido utilizados para ejecutar otras tareas o liberados del conjunto para reducir el consumo de energía y otros costes variables.

40 Es, por tanto, un objeto de la presente invención proporcionar un sistema que ofrezca una manera para asignar de manera flexible los componentes funcionales requeridos sobre los recursos informáticos contraídos de manera dinámica.

Con este fin, la presente invención proporciona, en un primer aspecto, un sistema de Subsistema Multimedia de Protocolo Internet (IMS) de acuerdo con la reivindicación 1.

5 De acuerdo con la invención, el sistema IMS comprende unos recursos informáticos en red y un repositorio de códigos de servicio para almacenar un código para implementar al menos un servicio de red, en el que el al menos un servicio de red comprende un servicio de repositorio de estados que ofrece un almacenamiento de un estado global correspondiente a una función de red.

El sistema IMS comprende además un repositorio de códigos de función, estando la función de red configurada para utilizar el servicio de repositorio de estados para asegurar el reparto de estado global entre las instancias de la misma función de red dentro del sistema.

10 Así mismo, el sistema IMS comprende una pluralidad de entidades lógicas programables que son ejecutadas sobre las fuentes informáticas en red, en la que cada una de las entidades está o bien en un estado no instanciado o bien en un estado instanciado después de que un código del repositorio de códigos de función o del repositorio de códigos de servicio haya sido desplegado sobre ellas.

15 El sistema IMS comprende además un controlador para mantener una visión constante del estado de las entidades lógicas no instanciadas y para determinar un estado de viabilidad global de los recursos informáticos en red. El controlador está también configurado para desplegar un código del repositorio de códigos de función y / o del repositorio de códigos de servicio sobre una entidad lógica programable no instanciada o para interrumpir y liberar una entidad lógica programable instanciada.

20 El sistema IMS puede comprender además una unidad de vigilancia que incorpore un servicio de vigilancia para hacer posible la realización de pruebas comparativas de las entidades lógicas programables instanciadas y una unidad elástica que incorpora un servicio elástico para utilizar el controlador para modificar una pluralidad de unidades lógicas instanciadas pertenecientes a la función de red y / o al servicio de red en base a una entrada procedente de la unidad de vigilancia relativa a esa función de red y / o a ese servicio de red.

25 Si las pruebas comparativas indican que solo están disponibles unas pocas instancias, y que estas instancias están completamente ocupadas o cargadas, pueden ser instanciadas nuevas funciones y / o servicios.

Un problema de los sistemas existentes es que se requiere que permanezca activa una función de red IMS determinada mientras las sesiones que se están gestionando están activas. Esto ya no se considera deseable en cuanto impide la distribución igualitaria de la carga a través de la red y la reconfiguración oportuna del sistema IMS.

30 Sin embargo, si se instancia una nueva función o servicio, requiere que la información asuma las sesiones en curso en las cuales están implicadas otras funciones o servicios. Para afrontar este problema, la presente invención propone el uso de un servicio de repositorio de estados que ofrezca un almacenamiento de un estado global correspondiente a una función de red.

35 El estado global es una información de estados que necesita ser compartida entre dos o más de las instancias de una función. Comprende una función de estados a la que todas las instancias necesitan acceder y / o modificar (configuración principal, claves criptográficas, etc.), la información de estados que es compartida entre múltiples instancias, pero no todas ellas (por ejemplo, la configuración regional, los parámetros de anchura del sitio, etc.). La información de sesión de usuario pertenece a esta categoría. Otro tipo de estado compartido puede también ser incluido en esta definición. Diferentes tipos de estado global pueden requerir diferentes grados de garantías de constancia para la correcta operación de la función de red.

40 De acuerdo con la invención, una función de red nuevamente instanciada extraerá el estado global correspondiente a esa función del repositorio de estados, haciendo que trabaje de manera similar a los de los ejemplos actualmente existentes de las funciones de red.

Debe destacarse que una pluralidad de funciones de red idénticas puede utilizar el mismo servicio de repositorio de estados para compartir un estado global perteneciente a esta función de red.

45 La unidad de vigilancia puede determinar la carga de una función de red y / o de un servicio de red. Así mismo, la unidad elástica puede ser configurada para modificar el número de entidades lógicas instanciadas pertenecientes a la unidad de red y / o al servicio de red cuando la carga sobrepase un intervalo operativo predeterminado. El intervalo operativo predeterminado puede estar compuesto por un límite superior, por encima del cual el número de entidades lógicas instanciadas se incrementará, y un límite inferior, por debajo del cual se reducirá el número de
50 entidades lógicas instanciadas.

La unidad elástica puede también ser configurada para tener en cuenta el estado de la viabilidad de los servicios informáticos en red obtenidos del controlador para determinar si hay que modificar el número de entidades lógicas instanciadas. Por ejemplo, si no hay más recursos disponibles, el número de entidades lógicas instanciadas no puede incrementarse. De modo similar, si hay una cantidad suficiente de recursos disponibles, la decisión de

incrementar las entidades lógicas instanciadas puede adoptarse en un periodo previo. De modo similar, el intervalo operativo mencionado puede depender del estado de la viabilidad.

5 El repositorio de códigos de servicio puede comprender un código para implementar el servicio de vigilancia y el servicio elástico. Así mismo, la unidad de vigilancia puede comprender una instanciación del servicio de vigilancia sobre una entidad lógica programable y la unidad elástica puede comprender una instanciación del servicio elástico sobre una entidad lógica programable.

10 El sistema IMS puede además comprender una unidad de predicción que incorpore un servicio de predicción para predecir las subidas de la carga de la función de red y / o del servicio de red. En este caso, la unidad elástica puede ser configurada para modificar el número de unidades lógicas instanciadas pertenecientes a la unidad de red y / o al servicio de red en base a una salida del servicio de predicción. El servicio de predicción hace posible un incremento o una reducción preventiva del número de entidades locales instanciadas. Por ejemplo, este número puede incrementarse antes de las 0: 00 AM o en el día de año nuevo.

15 El repositorio de códigos de servicio puede comprender un código para implementar el servicio de predicción. De modo similar a la unidad de vigilancia, el servicio de predicción puede ser una instanciación del servicio de predicción sobre una entidad lógica programable.

20 El controlador puede comprender una base de datos que incorpore dentro de ella el nombre de una función de red y / o de un servicio de red y una dirección y un estad correspondientes a cada unidad lógica instanciada sobre la cual se ha instanciado la función de red y / o el servicio de red. Esta base de datos mantiene el seguimiento de cuál sea la función o el servicio instanciado sobre qué entidad lógica y qué información de dirección, por ejemplo una dirección IP u otro identificador de red utilizar para acceder a la función o servicio.

El acceso de solo lectura a la base de datos para otras funciones o servicios dentro del sistema IMS puede ser autorizado utilizando una unidad de consulta con un servicio de consulta.

25 El repositorio de códigos de función y / o el repositorio de códigos de servicio puede comprender un código para una pluralidad de funciones de red y / o de servicios de red, respectivamente. En este caso, la unidad elástica puede ser configurada para modificar el número de entidades lógicas instanciadas pertenecientes a cada unidad de red y / o servicios de red basados individualmente en una entrada procedente de la unidad de vigilancia relativa a la función de red y / o al servicio de red. Por consiguiente, para algunas funciones de red o servicios de red, el número de instancias dentro del sistema IMS puede incrementarse mientras que respecto de otras funciones o servicios, el número poder reducirse, dependiendo en cada caso de la carga para esas funciones o servicios.

30 De acuerdo con un segundo aspecto, la presente invención proporciona un procedimiento de cambio de escala elástico de un sistema de Subsistema Multimedia de Protocolo Internet (IMS).

El procedimiento de acuerdo con la invención comprende las etapas de:

proveer unos servicios informáticos en red;

35 almacenar un código para implementar al menos un servicio de red en un repositorio de códigos de servicios, en el que el al menos un servicio de red comprende un servicio de repositorio de estados que ofrece un almacenamiento de un estado global correspondiente a una función de red;

almacenar un código para implementar al menos una función de red en un repositorio de códigos de función, estando la función de red configurada para utilizar el servicio de repositorio de estados para asegurar el reparto de estados global entre unas instancias de la función de red dentro del sistema;

40 proveer una pluralidad de entidades lógicas programables que sean ejecutadas sobre las fuentes informáticas en red, estando cada una de las entidades o bien en un estado no instanciado o bien en un estado instanciado después de que un código del repositorio de códigos de función o del repositorio de códigos de servicio haya sido desplegado sobre ellas;

45 mantener una visión constante del estado de las entidades lógicas instanciadas y determinar un estado de disponibilidad global de los recursos informáticos en red; y

modificar una pluralidad de entidades lógicas instanciadas pertenecientes a la función de red y / o servicio de red desplegando un código del repositorio de códigos de función y / o del repositorio de códigos de servicio sobre una entidad lógica programable no instanciada o para interrumpir y liberar una entidad lógica programable instanciada.

50 El procedimiento puede también comprender la realización de pruebas comparativas de las entidades lógicas programables instanciadas. En este caso, la modificación antes mencionada de un número de entidades lógicas instanciadas pertenecientes a la función de red y / o servicio de red se puede basar en la realización de pruebas comparativas relativas a esa función de red y / o servicio de red.

De acuerdo con un tercer aspecto, la presente invención proporciona un paquete de sistema de Subsistema Multimedia de Protocolo Internet (IMS) que comprende un medio de almacenamiento legible por ordenador, en el que el medio comprende un código que debe ser desplegado sobre una pluralidad de entidades lógicas programables que son ejecutadas sobre las fuentes informáticas en red. Cada una de las entidades está o bien en un estado no instanciado o bien en un estado instanciado después de que un código haya sido desplegado sobre ellas.

De acuerdo con la invención, el medio de almacenamiento comprende un código para implementar al menos un servicio de red, en el que el al menos un servicio de red comprende un servicio de repositorio de estados que ofrece el almacenamiento de un estado global correspondiente a una función de red. Comprende además un código para implementar la función de red en el que la función de red está implementada para configurar el servicio de repositorio de estados para asegurar el reparto de estado global entre instancias de la misma función de red dentro del sistema.

El medio también comprende un código para implementar un servicio controlador para mantener una visión constante del estado de las entidades lógicas instanciadas y para determinar un estado de la disponibilidad global de los recursos informáticos en red, en el que el servicio controlador está también configurado para desplegar código sobre una entidad lógica programable no instanciada o para interrumpir y liberar una entidad lógica programable instanciada.

El medio puede también comprender un código para implementar un servicio de vigilancia para hacer posible la comparación de pruebas comparativas de las entidades lógicas programables instanciadas, y un código para implementar un servicio elástico que esté configurado para utilizar el servicio de controlador para modificar una pluralidad de entidades lógicas no instanciadas pertenecientes a la unidad de red y / o al servicio de red en base a una entrada procedente de la unidad de vigilancia relativa a esa función de red y / o ese servicio de red.

El experto en la materia deberá apreciar que el paquete puede comprender una pluralidad de medios de almacenamiento diferentes sobre los cuales se distribuya el código mencionado anteriormente.

A continuación, se describirá la invención con mayor detalle y con referencia a los dibujos en los que:

La figura 1 ilustra un procedimiento de establecimiento de llamada de acuerdo con una forma de realización de la técnica anterior;

la figura 2 ilustra los componentes lógicos de una forma de realización del sistema de acuerdo con la invención; y

las figuras 3 y 4 ilustran la ejecución del establecimiento de llamada ilustrado en la figura 1 que utiliza una forma de realización del sistema de acuerdo con la invención.

Esquemáticamente, cada función de red, $F(i)$ presenta un estado global, $S(i)$, que determina el comportamiento futuro de cada función con respecto a los episodios de red que recibirá de las demás funciones y / o de los usuarios del sistema. El estado global, $S(i)$ debe ser compartido entre todas las instancias de la misma aplicación, mientras que algún(os) estado(s) $s(i)$ puede(n) también existir que no necesite ser compartido, en cuanto su ausencia podría traducirse en unos fallos recuperables tomados en combinación con el protocolo ordinario. Diferentes funciones de red no comparten necesariamente un estado, y pueden incorporar diferentes existencias acerca del nivel de constancia, disponibilidad y latencia máxima con la cual el estado global deba ser accesible.

Para todas las acciones macroscópicas iniciadas por los usuarios, por ejemplo el establecimiento de una llamada de vídeo o la distribución de un mensaje multimedia, se desplegará la función de red que desencadenará una serie de acciones y de actualización es de estados en una pluralidad de funciones. La actividad de usuario se traduce en un grado variable de actividad de un lado a otro de las funciones de red, según se determina mediante la complejidad de procesamiento y por los detalles del protocolo que cada función debe implementar. Las diferencias entre el papel y el consumo de recursos de las funciones de red determinan un comportamiento de cambio de escala diferente de cada función sometida a una carga de usuario creciente.

La presente invención divulga una arquitectura o sistema en la que un conjunto de recursos M en red pueden ser dinámicamente instanciados para ejecutar una pluralidad deseada de trabajadoras W . Una trabajadora es una instancia software que puede ser programada para ejecutar unas piezas de código arbitrarias que pueden implementar, entre otras, las funciones de red, F , que pueda ser implementada de varias formas conocidas, puede implementar técnicas establecidas de apalancamiento como por ejemplo multitarea, programación en tiempo real, virtualización, acotación de seguridad ("*sandboxing*"), etc. Cuando la demanda del sistema para una función resulta demasiado elevada (por ejemplo cuando está determinada por un controlador consciente de la aplicación, o bien centralizado o distribuido, una trabajadora procedente del conjunto es instanciada para convertirse en una instancia de la función deseada mediante la carga del código procedente del repositorio de códigos.

El repositorio de códigos proporciona un almacenamiento y una recuperación de piezas nominadas / versionadas de código que implementan las funciones de red F. Un código es considerado inmutable, y el repositorio podría implementar unas técnicas de gestión de versión y configuración como las adoptadas ampliamente en la técnica.

5 El conjunto de recursos M está también configurado para acoger una pluralidad de servicios D de amplitud del sistema que proporcionen entre otras un conjunto de instalaciones unificadas para el almacenamiento de estado global. Las funciones de red están diseñadas para utilizar una o más de las instalaciones de almacenamiento de estado D con el fin de asegurar el reparto de estado global entre todas las instancias con las garantías de constancia, exclusión mutua, tolerancia a los fallos, etc., requeridas. Los servicios D pueden ejecutarse sobre un conjunto de recursos de hardware dedicados o, en una forma de realización preferente, como una segunda categoría de programas software que sean ejecutados por el conjunto de trabajadoras W.

10 La figura 2 ilustra los componentes lógicos del sistema. Las trabajadoras 7, a la derecha, se representan como cuadrados. Las trabajadoras son entidades lógicas programables que ejecutan unos servicios informáticos en red, por ejemplo sistemas operativos (virtualizados), intérpretes de máquinas OS, máquinas virtuales, etc. La relación de uno a uno se asume que existe entre trabajadoras y recursos, y se supondrá que hay hasta M trabajadoras en el sistema. Las trabajadoras pueden existir en dos estados, notablemente no instanciados (representados como un cuadrado vacío) e instanciados (cuadrados ocupados) después de que se haya ejecutado sobre ellas el código apropiado entre un repositorio. El dibujo muestra unas trabajadoras instanciadas que ejecutan un código de función (por ejemplo, los cuadrados marcados con F(1) y F(2)) y un código de servicio (marcado D(1), D(2), y SR).

15 La figura 2 ilustra también unos repositorios 8, a la izquierda, representados como unos cilindros ocupados. Los repositorios son instalaciones de almacenamiento que pueden ser implementadas mediante medios técnicos conocidos, por ejemplo sistemas de archivo (distribuidos), almacenamientos de valor clave (distribuidos), bases de datos de memoria (distribuidas), etc.

20 Tipos especiales de un código de servicio que ejecutado sobre una trabajadora incluyen, por ejemplo, el código que implementa una funcionalidad de almacenamiento de red, esto es, el Repositorio de Estados 9 como un servicio ejecutado sobre las trabajadoras 10 inicializadas u otros servicios en red, por ejemplo la función 11 de controlador (distribuida), el servicio 12 de consulta (distribuido), etc. El código para todos los servicios está contenido en el "repositorio de códigos de servicios" y desplegado a petición a través de las trabajadoras bajo la supervisión del controlador.

25 El controlador 11 mantiene una información de estados actualizada acerca del despliegue de las trabajadoras y del estado de disponibilidad de los recursos informáticos. Con este fin, pueden ejecutarse unas mediciones a lo largo de la red (para una latencia de ancho de banda, y otras cantidades relevantes) y consultar periódicamente los nodos de trabajadoras acerca de su estado. La información así obtenida, junto con el conocimiento específico del sistema / función / servicio, es utilizado como base para decidir el despliegue de un código de función y servicio sobre las trabajadoras no instanciadas, y la interrupción y liberación de (a) trabajadora(s) instanciada(s) específica(s). Esto se puede implementar como un sistema centralizado o distribuido.

30 El servicio 12 de consulta puede ser requerido para otras funciones o servicios para obtener la dirección de red de otras funciones y servicios desplegadas a lo largo de los M recursos informáticos. Puede ser implementado como sistema centralizado o distribuido (utilizando técnicas conocidas, por ejemplo tablas hash distribuidas).

35 Al desplegar las trabajadoras P-CSCF y S-CSCF elásticas, el controlador instancia nuevas instancias de trabajadoras F(1) y F(2) que ejecutan el código apropiado según es recuperado del repositorio de códigos de función. El controlador puede también instanciar las trabajadoras D(1) y D(2) para ejecutar uno o más de los siguientes servicios:

- 40 • un servicio de vigilancia para hacer posible la realización de pruebas comparativas y el perfilado de las trabajadoras IMS así como de los servicios global del sistema empleados (por ejemplo los sistemas de almacenamiento de datos compartidos).
- 45 • un servicio de predicción para predecir las subidas repentinas de carga (en base a las predicciones de carga de llamadas a corto plazo combinadas con las predicciones basadas en historial para anticipar los cambios de carga futuros). Este servicio opcional hace posible unas trabajadoras IMS de provisión proactiva, y permite reducir la cantidad de sobreaprovisionamiento de las trabajadoras limitando al tiempo el riesgo de provisionamiento defectuoso de las trabajadoras (lo que, a su vez, puede provocar violaciones de LSA e incrementos del riesgo de perder clientes).
- 50 • un servicio elástico que implementa el bucle de retroalimentación responsable para cada tipo de trabajadora IMS para (a) decidir cuántas nuevas trabajadoras crear, (b) qué trabajadoras existentes liberar y (c) cuánto tiempo estarán inactivas.
- 55 • un servicio de mantenimiento que proporciona acceso a las diversas trabajadoras IMS operativas, por ejemplo para accionar una trabajadora IMS a un estado de ejecución segura antes de ser liberada.

Con el fin de operar un servicio IMS escalable, se puede introducir un conjunto de servicios globales del sistema anteriormente mencionados que puedan ser utilizados para proporcionar elasticidad y para mejorar la tolerancia global a los fallos y las propiedades de equilibrio de la carga de un sistema de ejecución.

5 Para permitir la liberación de trabajadoras que se requiera de manera inmediata y para compartir de manera igualitaria la carga cuando se añadan nuevas trabajadoras, las trabajadoras P-CSCF y S-CSCF recuperan y / o almacenan el estado de ejecución procedente de y / o sobre un sistema de almacenamiento compartido antes y después del procesamiento de una solicitud / respuesta que está relacionada con o puede afectar al (a los) estado(s) global relevante. En la figura 2, se muestra el despliegue de unas trabajadoras de repositorios de estados redundantes SR para dar servicio a las trabajadoras P-CSCF y S-CSCF.

10 Finalmente, la funcionalidad de llamada IMS es operativa, y puede soportar el establecimiento de llamadas básicas y de otras características telefónicas IMS.

Las figuras 3 y 4 ilustran los mecanismos subyacentes que inspeccionan la operación básica del sistema IMS en nube, específicamente los mecanismos que permiten la creación y coordinación de una multiplicidad de trabajadoras para proporcionar la funcionalidad IMS estándar deseada.

15 La figura 3 muestra el encaminamiento de una llamada de servicio para la función apropiada, por ejemplo en el contexto de una ejecución de flujo de llamadas entre el UE 1 y el UE 2 de la figura 1. Cuando un llamante UE 1 inicia una sesión con destino a un receptor UE 2, un mensaje de red, por ejemplo un mensaje SIP, llega hasta un punto 13 de entrada al sistema IMS que invoca el servicio 12 de consulta en la etapa S1 para encontrar una trabajadora F(1) P-CSCF que sea capaz de procesar su solicitud. A continuación, en la etapa S2, el mensaje es transmitido hasta la
 20 función F(3) P-CSCF integrada por F(1). En la etapa S3, F(1), a su vez, accede al servicio 12 de consulta para encontrar un supuesto del repositorio de estados SR. En la etapa S4, F(1) extrae y actualiza la información de estados global relevante acerca de las sesiones de telefonía activa del llamante UE 1. El estado contenido en el SR permitirá que la función F(1) P-CSCF contacte con la dirección de la trabajadora 5 apropiada de la S-CSCF, integrada por la trabajadora F(2), que sea capaz de procesar la solicitud y, así mismo, de notificar, en la etapa S5, el
 25 UE 1 del progreso de la llamada de acuerdo con las especificaciones IMS y las máquinas de estado de protocolo. Después de que la trabajadora F(2) recibe esta solicitud en la etapa S6, accederá, en la etapa S7, a su vez, al estado almacenado en el SR y actualizarlo en base a las instrucciones recibidas de la F(1) y sobre los protocolos y las máquinas de estado IMS relevantes. Contactará con la F(1) en la etapa S8 y, posiblemente con otras funciones IMS, por ejemplo las funciones HSS 6 y / u otras funciones P-CSCF, para alcanzar en último término el UE 2. Al final
 30 del completo procedimiento mostrado de establecimiento de llamada, una llamada puede así establecerse entre el UE 1 y el UE 2

La figura 4 ilustra el supuesto en el que el sistema IMS recibe una nueva demanda para una función previamente no instanciada, por ejemplo una solicitud de establecimiento para una nueva portadora de datos para vídeo destinada a ser recibida por el UE 1.

35 El punto 13 de entrada invoca el servicio 12 de consulta en la etapa S9 para encontrar una trabajadora que proporcione una función que actualmente no esté instanciada en el sistema y que se desee para la nueva portadora de datos. En este caso, el servicio 12 de consulta no puede encontrar una instancia trabajadora para procesar la solicitud. En la etapa S10, el servicio 12 de consulta da instrucciones al controlador 11 para instanciar, en la etapa S11, una nueva función F(3) de procesamiento multimedia apropiada. Una vez que la nueva trabajadora F(3) está
 40 operativa, el servicio de consulta hace retornar en la etapa S12 la dirección de la trabajadora de función F(3) hasta el punto 13 de entrada, y el mensaje procedente del UE 1 es finalmente encaminado hasta la F(3) en la etapa S13. La F(3) a continuación procederá con la apropiada ejecución del flujo accediendo al SR para recuperar, como se ha mostrado anteriormente, la información de estados global apropiada acerca del UE 1 y posiblemente de las direcciones de red de las demás funciones de red IMS que necesita para contactar, siguiendo las especificaciones
 45 IMS relevantes y las máquinas de estados de protocolo.

La presente invención hace posible la instalación de funciones y / o servicios deseados, proporcionando con ello una arquitectura flexible que es capaz de ajustarse a las fuentes informáticas requeridas. Proporciona un sistema capaz de instanciar nuevas funciones y servicios en el momento en que se requieran. Por tanto, las fuentes informáticas no tienen que ser reservadas *ab initio*.

50 Con la adición de servicios opcionales, como por ejemplo el servicio elástico y el servicio de vigilancia, se puede añadir funcionalidad al sistema. Más concretamente, estos servicios permiten que el sistema IMS aumente o se reduzca elásticamente dependiendo de la capacidad requerida y / o de la distribución de la carga. Si la capacidad requerida es baja, los recursos pueden resultar disponibles para otros usos. Esto permite una gestión más eficiente de los recursos informáticos de lo que es posible en otros sistema de la técnica anterior.

55 Debe resultar evidente para el experto en la materia que pueden realizarse diversas formas de realización a las formas de realización descritas en la presente memoria sin apartarse de alcance de la invención el cual se define por las reivindicaciones adjuntas.

REIVINDICACIONES

1.- Un sistema de Subsistema Multimedia de Protocolo Internet (IMS), que comprende:

unas fuentes informáticas en red;

5 un repositorio de códigos de servicio para almacenar un código para implementar al menos un servicio de red, en el que dicho al menos un servicio de red comprende un servicio de repositorio de estados que permite el almacenamiento de un estado global correspondiente a una función de red;

un repositorio de códigos de función para almacenar un código para implementar la función de red, estando dicha función de red configurada para usar el servicio de repositorio de estados para asegurar el reparto de estado global entre unas instancias de la misma función de red dentro del sistema;

10 una pluralidad de entidades lógicas programables que son ejecutadas sobre los recursos informáticos en red, estando cada una de dichas entidades o bien en un estado no instanciado o en un estado instanciado, después de que el código del repositorio de códigos de función o de que el repositorio de códigos de servicio ha sido desplegado sobre aquellas;

15 un controlador para mantener una visión constante del estado de las entidades lógicas instanciadas y para determinar un estado de disponibilidad global de los recursos informáticos en red, estando el controlador también configurado para desplegar un código del repositorio de códigos de función y / o del repositorio de códigos de servicio sobre una entidad lógica programable no instanciada o para interrumpir y liberar una entidad lógica programable instanciada;

en el que el sistema comprende además:

20 una unidad de vigilancia que incorpora un servicio de vigilancia para posibilitar la realización de pruebas comparativas de las entidades lógicas programables instanciadas;

25 una unidad elástica que incorpora un servicio elástico para utilizar el controlador para modificar una pluralidad de entidades lógicas instanciadas pertenecientes a la función de red y / o al servicio de red en base a una entrada de la unidad de vigilancia relativa a la función de red y / o al servicio de red.

2.- El sistema IMS de acuerdo con la reivindicación 1, en el que la unidad de vigilancia determina la carga de una función de red y / o de un servicio de red, y en el que la unidad elástica está configurada para modificar el número de entidades lógicas instanciadas pertenecientes a la unidad de red y / o al servicio de red cuando la carga sobrepasa un intervalo operativo predeterminado.

30 3.- El sistema IMS de acuerdo con la reivindicación 2, en el que la unidad elástica está también configurada para tomar en cuenta el estado de viabilidad de los recursos informáticos en red obtenidos por el controlador para determinar si hay que modificar dicho número de entidades lógicas instanciadas.

35 4.- El sistema IMS de acuerdo con cualquiera de las reivindicaciones 1-3, en el que el repositorio de códigos de servicio comprende un código para implementar el servicio de vigilancia y el servicio elástico, y en el que la unidad de vigilancia comprende una instanciación del servicio de vigilancia sobre una entidad lógica programable y en el que la unidad elástica comprende una instanciación del servicio elástico sobre una entidad lógica programable.

40 5.- El sistema IMS de acuerdo con cualquiera de las reivindicaciones 1-4, en el que el sistema comprende además una unidad de predicción que incorpora un servicio de predicción para pronosticar las subidas repentinas de carga de la función de red y / o del servicio de red, en el que la unidad elástica está configurada para modificar el número de entidades lógicas instanciadas pertenecientes a una función de red y / o a un servicio de red en base a una salida del servicio de predicción.

6.- El sistema IMS de acuerdo con la reivindicación 5, en el que el repositorio de códigos de servicios comprende un código para implementar el servicio de predicción, y en el que la unidad de predicción es una instanciación del servicio de predicción sobre una entidad lógica programable.

45 7.- El sistema IMS de acuerdo con cualquiera de las reivindicaciones 1-6, en el que el repositorio de códigos de función y / o el repositorio de códigos de servicio comprende un código para una pluralidad de funciones de red y / o de servicios de red, respectivamente, y en el que la unidad elástica está configurada para modificar el número de las entidades lógicas instanciadas pertenecientes a cada función de red y / o de servicios de red individualmente sobre la base de una entrada de la unidad de vigilancia relativa a esa función de red y / o servicio de red.

50 8.- El sistema IMS de acuerdo con cualquiera de las reivindicaciones precedentes en el que el controlador comprende una base de datos que incorpora, almacenado en su interior, el nombre de una función de red y / o de un servicio de red, y una dirección y un estado correspondientes para cada entidad lógica instanciada sobre la cual dicha función de red y / o dicha red ha sido instanciada.

- 9.- El sistema IMS de acuerdo con la reivindicación 8, que comprende además una unidad de consulta que incorpora un servicio de consulta para permitir el acceso a la base de datos para otras funciones o servicios dentro del sistema IMS.
- 5 10.- Un procedimiento para el cambio de escala elástico de un sistema de Subsistema Multimedia de Protocolo Internet (IMS), que comprende:
- proveer unos recursos informáticos en red;
 - almacenar un código para implementar al menos un servicio de red en un repositorio de códigos de servicio, en el que dicho al menos un servicio de red comprende un servicio de repositorio de estados que ofrece el almacenamiento de un estado global correspondiente a una función de red;
 - 10 almacenar un código para implementar al menos una función de red en un repositorio de códigos de función, estando dicha función de red configurada para utilizar el servicio de repositorio de estados para asegurar un reparto de estado global entre instancias de la misma función de red dentro del sistema;
 - 15 proveer una pluralidad de entidades lógicas programables que son ejecutadas sobre los recursos informáticos en red, estando cada una de dichas entidades, o bien en un estado no instanciado, o bien en un estado instanciado después de que el código del repositorio de códigos de función o el repositorio de códigos de servicio ha sido desplegado sobre aquellas;
 - mantener una visión constante de estado de las entidades lógicas instanciadas y determinar un estado de disponibilidad global de los recursos informáticos en red;
 - 20 modificar una pluralidad de entidades lógicas instanciadas pertenecientes a la función de red y / o al servicio de red para desplegar el código del repositorio de códigos de función y / o el repositorio de códigos de servicio sobre una entidad lógica programable no instanciada o para determinar la interrupción y la liberación de una entidad lógica programable instanciada;
 - comprendiendo además el procedimiento:
 - la realización de pruebas comparativas de las entidades lógicas programables instanciadas;
 - 25 en el que dicha modificación de una pluralidad de entidades lógicas instanciadas pertenecientes a la función de red y / o al servicio de red se basa en dicha realización de pruebas comparativas relacionadas con esa función de red y / o ese servicio de red.
- 11.- El procedimiento de acuerdo con la reivindicación 10, que comprende además determinar la carga de la función de red y / o del servicio de red, en el que dicha modificación comprende modificar el número de entidades lógicas instanciadas pertenecientes a la función de red y / o al servicio de red cuando la carga sobrepasa un intervalo operativo predeterminado.
- 30 12.- El procedimiento de acuerdo con la reivindicación 10 u 11, que comprende además tomar en cuenta el estado de disponibilidad de los recursos informáticos en red para determinar si hay que modificar dicho número de entidades lógicas instanciadas.
- 35 13.- El procedimiento de acuerdo con cualquiera de las reivindicaciones 10 - 12, que comprende además la predicción de subidas repentinas de carga de la función de red y / o del servicio de red, comprendiendo dicha modificación modificar el número de entidades lógicas instanciadas pertenecientes a la función de red y / o al servicio de red también sobre la base de una salida del servicio de predicción.
- 40 14.- El procedimiento de acuerdo con cualquiera de las reivindicaciones 10 - 13, en el que el repositorio de códigos de función y / o el repositorio de códigos de servicio comprende un código para una pluralidad de funciones de red y / o de servicios de red, respectivamente, comprendiendo dicha modificación modificar una pluralidad de entidades lógicas instanciadas pertenecientes a cada función de red y / o servicio de red individualmente basadas en la realización de pruebas comparativas de esa función de red y / o ese servicio de red.
- 45 15.- Un paquete de un sistema de Subsistema Multimedia de Protocolo Internet (IMS) que comprende un medio de almacenamiento legible por ordenador, y comprendiendo dicho medio un medio para ser desplegado sobre una pluralidad de entidades lógicas programables que son ejecutadas sobre los recursos informáticos en red, estando cada una de dichas entidades o bien en un estado no instanciado o bien en un estado instanciado después de que el código ha sido desplegado sobre aquellas, comprendiendo el medio de almacenamiento:
- 50 un código para implementar al menos un servicio de red, en el que dicho al menos un servicio de red comprende un servicio de repositorio de estados que ofrece el almacenamiento de un estado global correspondiente a una función de red;

un código para implementar la función de red, estando dicha función de red configurada para utilizar el servicio de repositorio de estados para asegurar el reparto de estado global entre las instancias de la misma función de red dentro del sistema;

5 un código para implementar un servicio de controlador para mantener una visión constante del estado de las entidades lógicas instanciadas y para determinar un estado de disponibilidad global de los recursos informáticos en red, estando dicho servicio de controlador configurado además para desplegar un código del repositorio de códigos de función y / o del repositorio de códigos de servicio sobre una entidad lógica programable no instanciada o para interrumpir y liberar una entidad lógica programable instanciada;

10 el paquete Sistema de Subsistema Multimedia de Protocolo Internet (IMS) que comprende un código para implementar un servicio de vigilancia para posibilitar la realización de pruebas comparativas de las entidades lógicas programables instanciadas y un código para implementar un servicio elástico que está configurado para utilizar el servicio de controlador para modificar una pluralidad de entidades lógicas instanciadas pertenecientes a la función de red y / o al servicio de red sobre la base de una entrada por parte de la unidad de vigilancia relativa a esa función de red y / o a ese servicio de red.

15 16.- El paquete de Sistema de Subsistema Multimedia de Protocolo Internet (IMS) de acuerdo con la reivindicación 15, que comprende además un código para implementar un servicio de predicción para pronosticar unas subidas repentinas de carga de la función de red y / o del servicio de red, en el que el servicio elástico está configurado para modificar el número de unidades lógicas instanciadas pertenecientes a la función de red y / o al servicio de red en base a una salida del servicio de predicción.

20

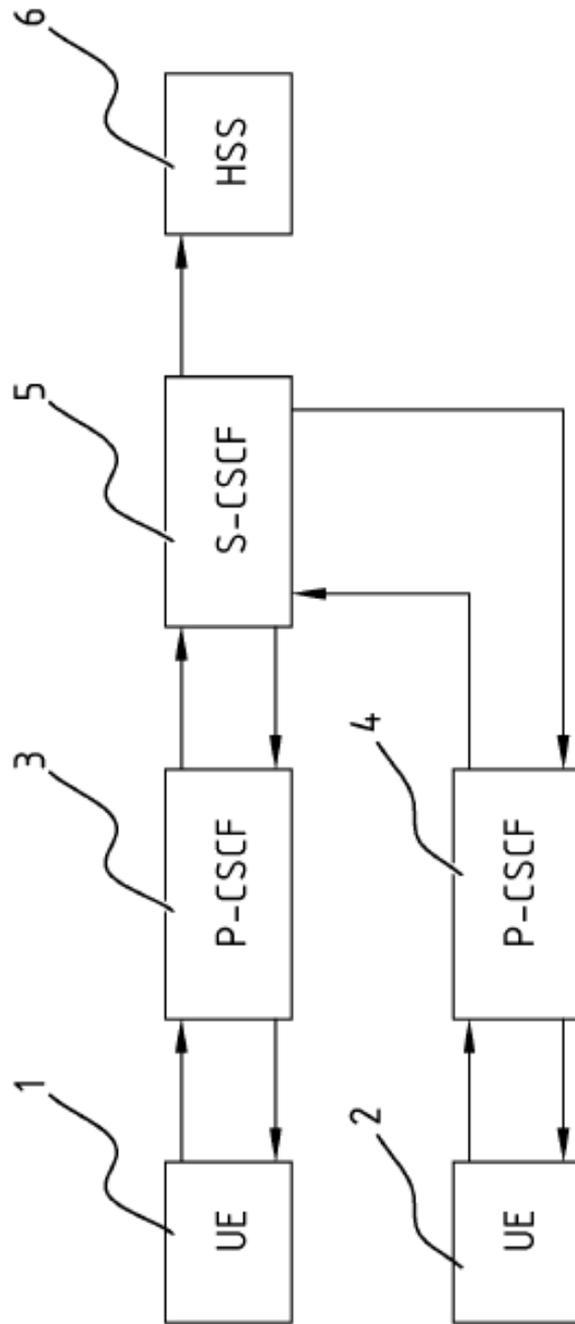


FIG. 1

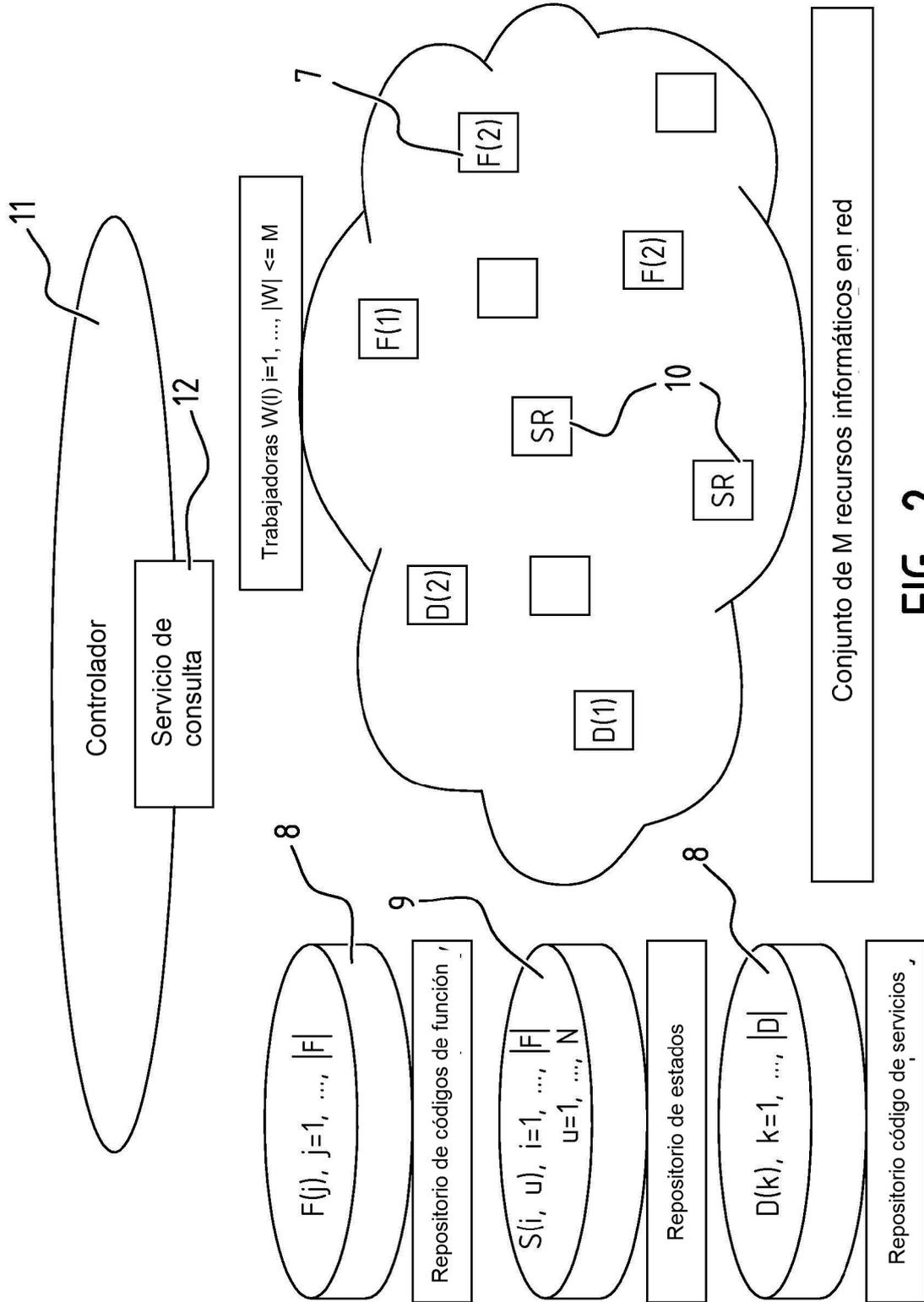


FIG. 2

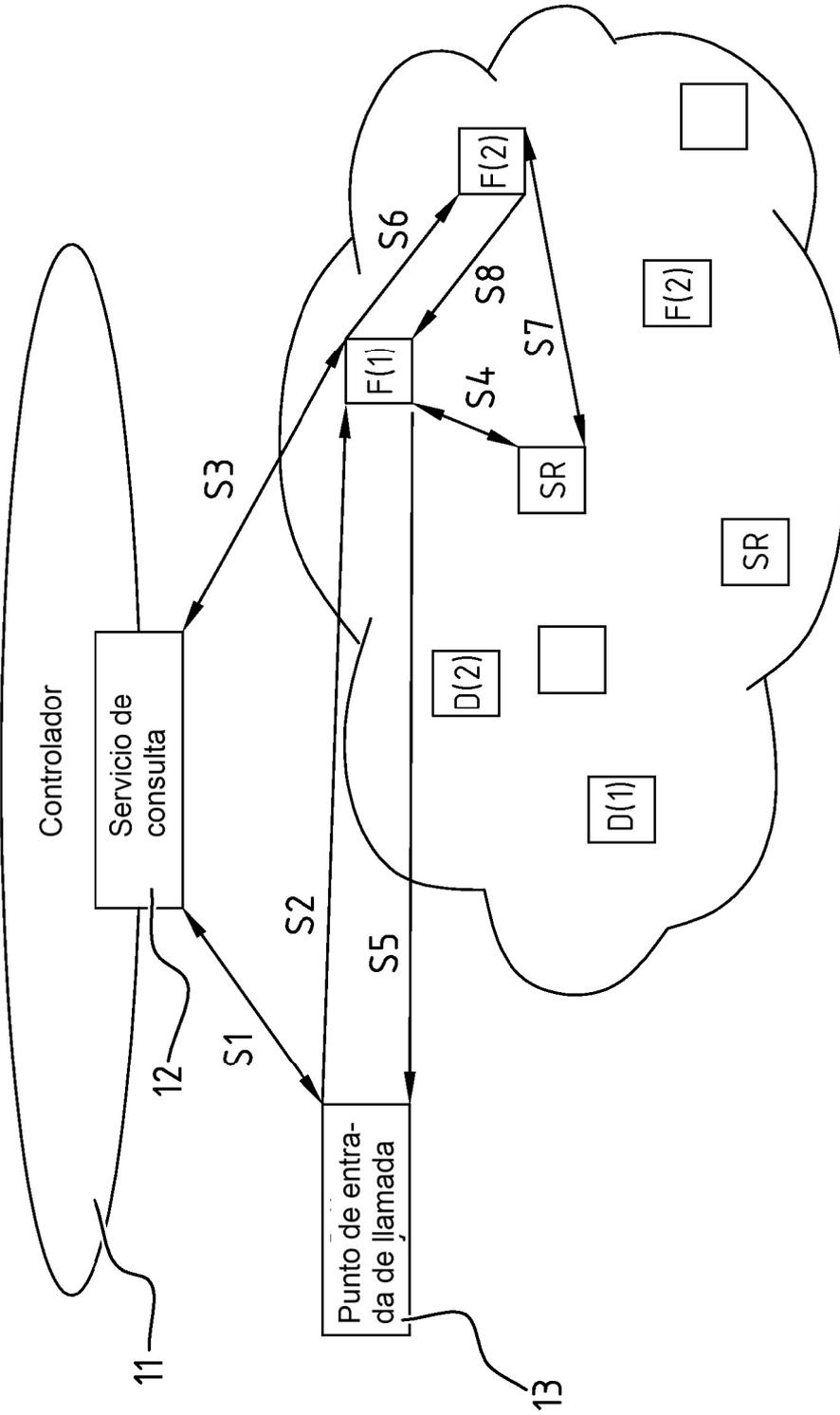


FIG. 3

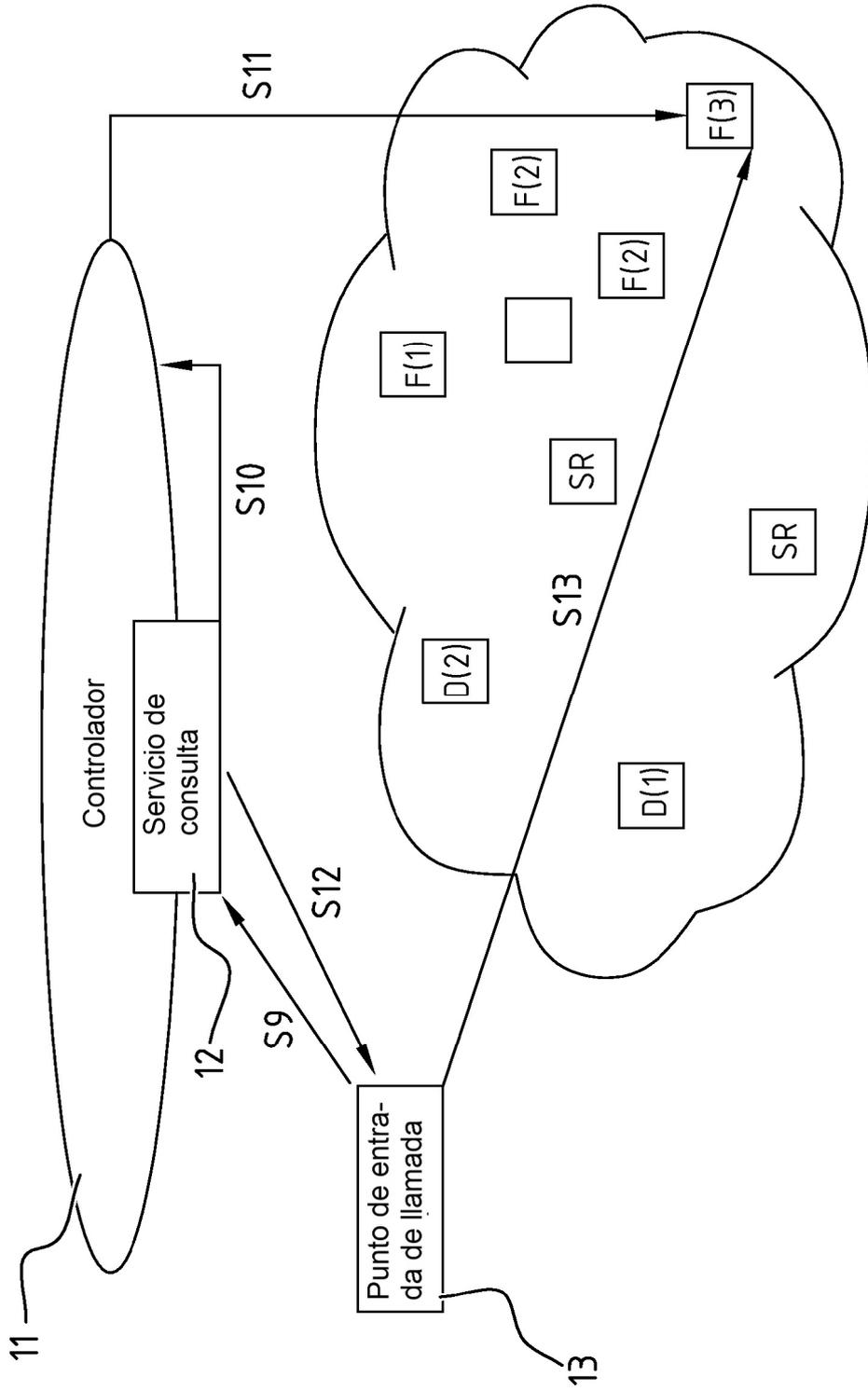


FIG. 4