

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 765 032**

51 Int. Cl.:

G06F 9/50 (2006.01)

G06F 9/455 (2008.01)

G06F 11/34 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **30.05.2012 PCT/US2012/039961**

87 Fecha y número de publicación internacional: **03.01.2013 WO13002936**

96 Fecha de presentación y número de la solicitud europea: **30.05.2012 E 12804593 (7)**

97 Fecha y número de publicación de la concesión europea: **30.10.2019 EP 2726984**

54 Título: **Control de utilización de red**

30 Prioridad:

29.06.2011 US 201113171445

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

05.06.2020

73 Titular/es:

**MICROSOFT TECHNOLOGY LICENSING, LLC
(100.0%)
One Microsoft Way
Redmond, WA 98052 , US**

72 Inventor/es:

FRIES, ROBERT

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 765 032 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

Control de utilización de red

Antecedentes

5 En entornos informáticos modernos, aplicaciones distribuidas pueden consistir en numerosas unidades autónomas que se ejecutan en diferentes anfitriones, cooperando a través de una red, interactuando con clientes tales como navegadores, accediendo a servidores de base de datos, etc. Aplicaciones entre pares, aplicaciones web de múltiples niveles, tablas de función de troceo distribuidas, grupos o agrupaciones de servidores, son ejemplos de aplicaciones distribuidas. Tales aplicaciones distribuidas pueden ejecutarse en la nube, centros de datos, etc.

10 La comunicación de red es necesaria para operación efectiva de tales aplicaciones. En algunos casos, podría ser deseable que una aplicación tenga diversas características de rendimiento de red, por ejemplo, ancho de banda entre componentes, caudal promedio a la Internet o clientes en la misma, ancho de banda promedio por instancia de aplicación, etc. En otras palabras, puede ser deseable tener garantías de rendimiento de red a nivel de aplicación.

15 Sin embargo, hasta la fecha, tales garantías no han sido posibles. Garantías de rendimiento de red se han proporcionado en el nivel de red. Se han usado encaminadores para proporcionar garantías de Calidad de Servicio (QoS), pero estas garantías no se han orientado para aplicaciones distribuidas como un todo. En otras palabras, los medios típicos para asegurar niveles de ancho de banda o de capacidad se han dirigido a redes de datos por sí mismos, en lugar de a aplicaciones. Además, en un entorno en el que una aplicación distribuida puede estar ejecutándose, cuestiones tales como compartición de recursos entre diferentes aplicaciones hacen difícil garantizar recursos de red a lo largo de trayectorias de comunicación a través de una aplicación distribuida.

20 Incluso en casos en los que podría haberse proporcionado aprovisionamiento de recursos de red en el nivel de aplicación, las técnicas usadas para obtener o mantener requisitos de recursos de red han implicado técnicas tal como control de tasa de transmisión, regulación y otros mecanismos de control de tráfico orientados a red. Aunque elementos de aplicación distribuida a menudo se alojan en máquinas virtuales (VM), tecnología de VM no se ha explotado para ayudar a garantizar requisitos de red a nivel de aplicación.

25 A continuación se analizan técnicas relacionadas con el control de recursos de red para aplicaciones distribuidas usando tecnología de máquina virtual (VM).

30 XIAOQIAO MENG Y COL.: "Improving the Scalability of Data Center Networks with Traffic-aware Virtual Machine Placement" INFOCOM, 2010 ACTAS IEEE, IEEE, PISCATAWAY, NJ, Estados Unidos, 14 de marzo de 2010, páginas 1-9, ISBN: 978-1-4244-5836-3, desvela un procedimiento de colocación de máquina virtual consciente de tráfico para mejorar una escalabilidad de red. Máquinas virtuales con gran uso de ancho de banda mutuo se asignan a máquinas de anfitrión en proximidad cercana. El documento presenta un algoritmo de aproximación de dos niveles para resolver un problema de colocación de máquina virtual para tamaños de problemas muy grandes.

35 VIVEK SHRIVASTAVA Y COL.: "Application-aware virtual machine migration in data centers" INFOCOM, 2011 ACTAS IEEE, IEEE, 10 de abril de 2011, páginas 66-70, 001: 10.11 09/INFOCOM.2011.5935247 ISBN: 978-1-4244-9919-9 desvela un esquema de cálculo para incorporar dependencias entre máquinas virtuales y la topología de red subyacente para determinar decisiones de migración de máquinas virtuales. El esquema propuesto ayuda a disminuir adicionalmente tráfico de red en comparación con migración alternativa bien conocida de máquinas virtuales que no es consciente de la aplicación.

Sumario

40 Es el objeto de la presente invención proporcionar procedimientos mejorados para utilización de red.

Este objeto se soluciona mediante la materia objeto de las reivindicaciones independientes.

Se definen realizaciones preferidas en las reivindicaciones dependientes.

45 El sumario siguiente se incluye únicamente para introducir algunos conceptos analizados en la descripción detallada a continuación. Este sumario no es exhaustivo y no se concibe para delinear el ámbito del objeto reivindicado, que se expone mediante las reivindicaciones presentadas en la parte final.

50 La utilización de red de una aplicación arbitraria se controla rastreando estadísticas de uso de red de máquinas virtuales (VM), incluyendo al menos partes de alojamiento de VM de la aplicación. Para control de utilización de red, VM sirven como intermediarios de red para elementos de la aplicación. Una especificación para un requisito de red de la aplicación se evalúa frente a las estadísticas de uso de red. Cuando no se satisface un requisito de capacidad de red a/desde/a través de un elemento de la aplicación, una o más VM se adaptan para satisfacer el requisito. Por ejemplo, una VM puede migrarse desde una ubicación de anfitrión o de red que tiene ancho de banda de red en exceso. O, por ejemplo, puede aumentarse la disponibilidad de ancho de banda de red para una VM bajo requisito y puede disminuirse la disponibilidad de ancho de banda para una VM en una ubicación de anfitrión o de red apropiada. Por lo tanto, pueden satisfacerse requisitos de trayectoria de comunicación a nivel aplicación usando adaptaciones de

VM.

Muchas de las presentes características se explicarán a continuación con referencia a la siguiente descripción detallada considerada en conexión con los dibujos adjuntos.

Breve descripción de los dibujos

5 La presente descripción se entenderá mejor a partir de la siguiente descripción detallada leída a la vista de los dibujos adjuntos, en los que números de referencia similares se usan para designar partes similares en la descripción adjunta.

La Figura 1 muestra una capa de virtualización de ejemplo.

La Figura 2 muestra procedimientos e interacciones de una capa de virtualización en relación con máquinas virtuales e imágenes de máquina virtual.

10 La Figura 3 muestra un ejemplo de una aplicación distribuida.

La Figura 4 muestra otra aplicación distribuida de ejemplo.

La Figura 5 muestra un sistema de especificación de capacidad de red para una aplicación distribuida y de provisión de capacidad de red usando adaptaciones de máquina virtual (VM).

15 La Figura 6 muestra un ejemplo de un anfitrión con mejoras de virtualización para facilitar notificación de red de VM y manipulación de uso de red.

La Figura 7 muestra el sistema de la Figura 5 en relación con el ejemplo aplicación distribuida en la Figura 3.

La Figura 8 muestra anfitriones que alojan una aplicación distribuida que se ejecuta en VM.

La Figura 9 muestra una vista detallada de una parte de la Figura 8.

Las Figuras 10 y 11 muestran especificaciones de red de aplicación de ejemplo.

20 **Descripción detallada**

VISIÓN DE CONJUNTO Y VIRTUALIZACIÓN

Realizaciones analizadas a continuación se refieren al uso de tecnología de VM para proporcionar capacidad de red en la presencia de condiciones de red dinámicas. El análisis comenzará con una vista general de tecnología de virtualización que incluye componentes de virtualización tales como hipervisores. A continuación se analizarán aplicaciones distribuidas y ejemplos de las mismas. Se describirá una vista general de uso de tecnología de VM para aprovisionamiento de red, seguida por un análisis de mejoras de VM relacionadas y explicación detallada de cómo se usan esas mejoras.

La Figura 1 muestra una capa 100 de virtualización de ejemplo. Un ordenador 102 tiene hardware 104, incluyendo una unidad 106 de procesamiento central (CPU), memoria 108, una interfaz 110 de red, almacenamiento 112 no volátil y otros componentes no mostrados, tal como un bus, un adaptador de visualización, etc. La capa 100 de virtualización gestiona y facilita la ejecución de máquinas 114 virtuales. Aunque no se muestra en la Figura 1, cada máquina 114 virtual habitualmente tiene una imagen de disco virtual asociada y un sistema operativo invitado. Por brevedad, el sistema operativo y quizás software de aplicación de una máquina 114 virtual en ocasiones se denominará como un invitado, que se almacena y ejecuta a partir de la imagen de disco virtual asociada con la máquina 114 virtual. Por conveniencia, el término "hipervisor" se usará en el presente documento para referirse a las diversas formas de capas de virtualización. Además, como se analizará a continuación, las máquinas 114 virtuales se usan para alojar elementos de aplicaciones distribuidas.

La capa 100 de virtualización puede ser de cualquier variedad de implementaciones conocidas o futuras, tal como Hyper-V Server (TM), VMWare ESX Server (TM), Xen, Oracle VM (TM), etc. La arquitectura de la capa de virtualización puede ser un tipo alojado, con un supervisor de máquina virtual (VMM) ejecutándose en un sistema operativo anfitrión, o un tipo de metal desnudo con un hipervisor o similar ejecutándose directamente en el hardware 104 del ordenador 102. Como se usa en el presente documento, la expresión "máquina virtual" se refiere a una máquina virtual de tipo sistema que simula cualquier arquitectura de hardware específica (por ejemplo, x86) capaz de ejecutar código nativo para esa arquitectura de hardware; para el invitado, la máquina virtual puede ser casi indistinguible de una máquina de hardware. Máquinas virtuales analizadas en el presente documento no son máquinas virtuales abstractas o de tipo procedimiento tales como máquinas virtuales Java.

La capa 100 de virtualización realiza la función básica de gestión de las máquinas 114 virtuales y compartición del hardware 104 tanto por sí misma como por las máquinas 114 virtuales. Puede usarse cualquiera de una diversidad de técnicas para aislar las máquinas 114 virtuales del hardware 104. En una realización, la capa de virtualización puede proporcionar diferentes entornos aislados (es decir, particiones o dominios) que corresponden a máquinas 114 virtuales. Parte de la capa 100 de virtualización tal como controladores de dispositivo virtuales compartidos, instalaciones de comunicación entre máquinas virtuales y API de gestión de máquina virtual (interfaces de programación de aplicación), pueden ejecutarse en una partición o dominio privilegiado, permitiendo un hipervisor compacto y eficiente. En otras realizaciones, funcionalidad para gestión de máquina virtual y compartición coherente del hardware 104 puede residir en un hipervisor monolítico sobre el metal.

La Figura 2 muestra procedimientos e interacciones de capa 100 de virtualización en relación con máquinas 114 virtuales e imágenes 140 de máquina virtual. La capa 100 de virtualización realiza un procedimiento 142 de inicio y

5 ejecución de una máquina 114 virtual, posiblemente de acuerdo con correspondientes parámetros de configuración de máquina virtual. Cuando se inicia una máquina 114 virtual (VM), la capa de virtualización identifica una imagen 140 de máquina virtual asociada. En la práctica, puede usarse cualquier imagen 140 de máquina virtual por cualquier máquina 114 virtual. La imagen 140 de máquina virtual puede ser un archivo especialmente formateado (por ejemplo, un VHD) en un sistema 141 de archivos de la capa 100 de virtualización. La capa 100 de virtualización carga la imagen 140 de máquina virtual identificada. La máquina 114 virtual iniciada monta y lee la imagen 140 de máquina virtual, quizás buscando un registro de arranque maestro u otra información de arranque, y arranca un sistema operativo invitado que comienza a ejecutarse.

10 La capa 100 de virtualización gestiona la ejecución de la máquina 114 virtual, tratando ciertas llamadas al núcleo del invitado, hiperllamadas, etc., y coordinando el acceso de la máquina 114 virtual al hardware 104 subyacente. A medida que el invitado y su software se ejecutan, la capa 100 de virtualización puede mantener estado del invitado en la imagen 140 de disco virtual; cuando el invitado, o una aplicación ejecutada por el invitado, escribe datos en el "disco", la capa 100 de virtualización traduce los datos al formato de la imagen 140 de disco virtual y escribe en la imagen.

15 La capa 100 de virtualización puede realizar un procedimiento 144 de apagado de la máquina 114 virtual. Cuando se recibe una instrucción para detener la máquina 114 virtual, el estado de la máquina 114 virtual y su invitado se guarda en la imagen 140 de disco virtual, y se borra el procedimiento de máquina 114 virtual (o partición) de ejecución. Una especificación de la máquina 114 virtual puede permanecer para un reinicio posterior de la máquina 114 virtual.

APLICACIONES DISTRIBUIDAS

20 La Figura 3 muestra un ejemplo de una aplicación 160 distribuida. El ejemplo en la Figura 3 es una aplicación de múltiples niveles que tiene un nivel 162 de presentación, un nivel 164 de aplicación y un nivel 166 de datos. Aunque no se muestra, la aplicación distribuida se ejecuta en una red de datos en la que elementos se comunican a través de nodos de comunicación tales como encaminadores, puentes, conmutadores, pasarelas, equilibradores de carga, etc. El nivel 162 de presentación se comprende de instancias 168 de servidor web u otros servidores o aplicaciones que interactúan con clientes que usan la aplicación 160 distribuida. El nivel 164 medio se forma mediante instancias 170 de aplicación que pueden implementar la lógica principal de la aplicación 160 distribuida, traducir entre objetos y datos en una base de datos, traducir peticiones de clientes a consultas de base de datos, etc. El nivel 164 de datos puede tener un número de servidores 172 de base de datos, interfaces de base de datos, almacenes de datos u otras formas de almacenamiento de datos persistente y eficiente. Los equilibradores 174 de carga pueden equilibrar carga a través de los elementos en cada capa. Los elementos en cada nivel, por ejemplo las instancias 168 de servidores web, son duplicados ejecutables o pilas de software que tienen diferentes estados de ejecución.

30 En la operación, la aplicación 160 distribuida recibe peticiones, por ejemplo de clientes, otras aplicaciones, servicios, etc. Las peticiones se tratan por las instancias 168 de servidor web, que pueden iniciar una transacción o de otra manera comunicarse con las instancias 170 de aplicación. A su vez, las instancias 170 de aplicación pueden tener interacciones con los servidores 172 de base de datos.

35 Cada línea continua en la Figura 3 representa una posible trayectoria de comunicación de red. Con respecto a requisitos de red de la aplicación 160 distribuida, cualquiera de las líneas continuas (por ejemplo, la trayectoria 176), o cualesquiera diversos conjuntos de líneas (por ejemplo, el conjunto 178 de trayectorias) entre elementos similares (por ejemplo, todas las líneas entre servidores 168 web y equilibradores 174 de carga), o combinaciones de los mismos, son puntos potenciales en los que podría especificarse, evaluarse o provisionarse un requisito de red.

40 La Figura 4 muestra otra aplicación distribuida de ejemplo; aplicación 200 entre pares. Los pares 202 se ejecutan en diferentes anfitriones (o en diferentes invitados). Cada par 202 generalmente se comporta igual y realiza las mismas funciones, aunque con estados de operación diferentes. La aplicación 200 entre pares podría ser una tabla de función de troceo distribuida, un sistema de almacenamiento distribuido, un servicio de red, etc. La aplicación entre pares puede tener conexiones entre pares 202. Los pares 202 pueden comunicarse con entidades externas a través de la Internet 204. Los pares 202 también podrían usar diversos servicios tal como un servidor 206 de base de datos que proporciona una base 208 de datos relacional. Los pares 202 pueden comunicarse para servicio sustancial de peticiones externas, así como para mantenimiento interno de interconectividad a nivel de par. Una petición externa podría tratarse directamente por un par 202, o ese par podría tener información que le permite pasar la petición a otros pares 202 que dan servicio a la petición y devolver una respuesta.

50 Como en la Figura 3, líneas continuas representan trayectorias de red entre elementos. Líneas entre pares 202 representan trayectorias de red entre pares, y líneas a la aplicación 200 entre pares representan trayectorias a cualquiera o múltiples pares 202. Como en la Figura 3, cualquiera de las líneas (por ejemplo, las trayectorias 210, 212) o conjuntos de líneas (por ejemplo, las trayectorias 214) podría ser objetivos candidatos para especificar un requisito de capacidad de red o ancho de banda, evaluando el requisito frente a condiciones medidas, y ajustando VM para satisfacer el requisito.

APROVISIONAMIENTO DE RED A NIVEL APLICACIÓN

La Figura 5 muestra un sistema de especificación de capacidad de red para una aplicación 230 distribuida y aprovisionamiento de capacidad de red usando adaptaciones de máquina virtual (VM). Un controlador 232 de red

realiza procedimientos, descritos más adelante, para evaluar condiciones de red y necesidades de una aplicación distribuida 232. Se proporciona una especificación 234 de red que especifica requisitos de red (o preferencias) para la aplicación 230 distribuida. La especificación 234 de red incluye requisitos 236 de red para la aplicación 230 distribuida.

5 Un requisito 236 de red puede identificar trayectorias de red a nivel aplicación o conexiones entre elementos de la aplicación 230 distribuida (por ejemplo, conectividad entre elementos sin tener en cuenta encaminadores subyacentes, etc. que proporcionan la conectividad). Por ejemplo, en el ejemplo de la Figura 3, un requisito 236 de red podría especificar la trayectoria 176 o conjunto 178 de trayectorias. En el ejemplo de la Figura 4, un requisito de red podría identificar las trayectorias 210, 212 o trayectorias 214. Tales trayectorias o conexiones se identifican, en el resumen,
10 en términos de los elementos 238 de aplicación de la aplicación distribuida y/o elementos 240 externos que conectan con la aplicación 230 distribuida (elementos de aplicación pueden ser partes de la aplicación 230 distribuida que se ejecuta en diferentes anfitriones o invitados de VM, tal como pares 202, servidores 168 web, servicios de red, etc.). En el ejemplo de la Figura 5, las trayectorias 242 y 244 tienen correspondientes requisitos 236 de red. Un requisito 236 de red también especifica un requisito de recursos de red que corresponde a su trayectoria (o trayectorias, o clases de trayectorias). Por ejemplo, un requisito 236 de red para la trayectoria 242 podría especificar un requisito de ancho de banda específico (o dinámicamente calculado) tal como 50 kilobytes por segundo (KBS). Detalles de especificaciones de red se describen más adelante con referencia a las Figuras 10 y 11.

La Figura 5 también muestra datos 246 de estadísticas de red. Los datos 246 de estadísticas de red almacenan estadísticas 248 de red notificadas desde elementos 238 de aplicación y/o nodos de comunicación de red (por ejemplo, encaminadores, agregadores, pasarelas, etc.). Las estadísticas de red también podrían incluir información de topología de red que usa el controlador 232 de red para construir un modelo de la topología actual de la red de datos subyacente. Tal información de topología podría derivarse a partir de datos de encaminamiento de BGP (protocolo de pasarela de frontera) recopilados o datos de encaminamiento de OSPF (primer trayecto más corto abierto). Obsérvese que no se requieren información de topología de red y un modelo de topología para aprovisionamiento de red.

25 Más específicamente, los datos 246 de estadísticas de red podrían incluir información acerca del ancho de banda de red actual o histórico consumido por los elementos 238 de aplicación, anfitriones que alojan los mismos, VM que alojan elementos de aplicación, y así sucesivamente. En una realización, hipervisores que gestionan VM, en las que se ejecutan los elementos 238 de aplicación, pueden tener un componente que supervisa y notifica uso de red de VM individuales (véase el medidor 270 en la Figura 6), incluyendo posiblemente información que asocia el uso (por ejemplo, un valor de ancho de banda) con una identidad de la VM. En una realización, la actividad de red o consumo de cada VM que aloja un elemento de la aplicación 230 distribuida se notifica y rastrea. En este caso, el sistema realiza seguimiento de qué VM aloja qué elementos 238 de aplicación. En efecto las VM actúan como intermediarios de rastreo de red de los elementos 238 de aplicación. Detalles y ejemplos adicionales de estadísticas 248 de red y datos 246 de estadísticas de red se analizan más adelante.

35 Volviendo a la Figura 5, el controlador 232 de red evalúa los datos 246 de estadísticas de red frente a los requisitos 236 de red de la aplicación 230 distribuida. El controlador 232 de red puede determinar si algunas trayectorias o conjuntos o clases de trayectorias (o combinaciones de los mismos) especificados en los requisitos 236 de red no se satisfacen (o probablemente no se satisfarán). El controlador 232 de red a su vez determina adaptaciones 250 de VM que emite al anfitrión o anfitriones que alojan las VM en las que se están ejecutando los elementos 238 de aplicación.

40 Una adaptación 250 de VM podría ser instrucciones para uno o más componentes de gestión de VM (véase el componente de gestión 274 en la Figura 6) para migrar una VM (o múltiples VM) desde un primer anfitrión a un segundo anfitrión deteniendo la VM en el primer anfitrión, transfiriendo la imagen de disco de la VM al segundo anfitrión (si es necesario), e iniciando la misma en el segundo anfitrión. Una adaptación 250 de VM podría ser una instrucción a un hipervisor para aumentar o disminuir la capacidad de red para una o más de las VM gestionadas por el hipervisor. En una realización, hipervisores tienen lógica (véase el regulador 272 en la Figura 6) para controlar el caudal de red para VM individuales. Mientras regulación y migración son técnicas efectivas y sencillas para el control del uso de red, pueden usarse otras técnicas. Por ejemplo, en una realización, puede reducirse el ancho de banda global reduciendo el número de instancias en una aplicación nivel. En otra realización, se señalizan sistemas operativos invitados para ajustar el uso de red, por ejemplo, simulando una condición de saturación de red al algoritmo de control de tasa de TCP (Protocolo de Control de Transmisión) de un sistema operativo invitado.

50 Con respecto a implementación del controlador 232 de red, el controlador 232 de red podría implementarse en una diversidad de ordenadores. En una realización, el controlador 232 de red se incluye en un sistema de gestión de máquina virtual que gestiona VM a través de una red. Por ejemplo, el System Center Virtual Machine Manager (TM), VMWare VirtualCenter (TM), etc. En otra realización, el controlador 232 de red es una aplicación autónoma que se ejecuta en cualquier anfitrión con conectividad de red. En otra realización más, el controlador 232 de red es en sí mismo una aplicación distribuida. El controlador 232 de red puede implementarse como parte de la aplicación 230 distribuida (con posibles limitaciones en cómo puede gestionarse el ancho de banda). En una realización, el controlador 232 de red gestiona muchas diferentes aplicaciones distribuidas que comparten una misma red y anfitriones. En cuyo caso, puede tener la capacidad de aumentar o disminuir el ancho de banda disponible para una aplicación para ayudar a cumplir con un requisito de otra aplicación.

La Figura 6 muestra un ejemplo de un anfitrión 268 con mejoras de virtualización para facilitar notificación de red de VM y manipulación de uso de red. Un hipervisor 274 realiza funciones de virtualización conocidas, como se ha mencionado anteriormente, para ejecutar las VM 276, 278. Además, el hipervisor 274 tiene un regulador 272 y un medidor 270. El medidor 270 recopila estadísticas acerca del volumen (y/o tasa) de datos de red transmitidos a/desde las VM en el anfitrión 268. En una realización, el hipervisor puede tener un dispositivo de NIC (Tarjeta de Interfaz de Red) virtual para cada VM y se generan estadísticas a través del dispositivo de NIC virtual.

El regulador 272 permite que el hipervisor 274 controle caudal de red de las VM que gestiona. Pueden usarse una diversidad de algoritmos de control. Pueden usarse algoritmos de control de tasa de red conocidos (por ejemplo, retardando paquetes en cola antes de transmisión). En algunos casos, puede reducirse el ancho de banda para algunas VM (por ejemplo, una VM 278 que no es parte de la aplicación distribuida) para permitir el aumento de ancho de banda para una VM 276 que aloja un elemento 238 de aplicación. Obsérvese que el hipervisor o NIC virtual puede inspeccionar paquetes, a medida que pasan a través, para entender la correspondiente trayectoria de red real y decidir qué sesiones regular.

El anfitrión 268 también podría tener un componente 280 de gestión para facilitar la gestión de red. En una realización, el componente 280 de gestión recopila estadísticas desde el medidor 270 y transmite las estadísticas para almacenamiento en los datos 246 de estadísticas de red. El componente 280 de gestión también puede controlar el regulador 272 de acuerdo con instrucciones desde el controlador 232 de red. El componente 280 de gestión también puede efectuar instrucciones para migrar VM en respuesta a adaptaciones de VM desde el controlador 232 de red.

El componente 280 de gestión también puede ser un medio conveniente para permitir que el controlador 232 de red interactúe con diferentes implementaciones de virtualización; pueden proporcionarse diferentes implementaciones de componentes 280 de gestión para respectivas implementaciones de virtualización. Cada una puede implementar una misma API (Interfaz de Programación de Aplicación) dirigida hacia el controlador y tener diferente lógica dirigida hacia la virtualización adecuada para una diferente implementación de virtualización.

La funcionalidad del componente 280 de gestión, regulador 272 y medidor 270 puede incluirse en cualquier lugar en la capa de virtualización que es adecuada para la implementación de virtualización particular que se está usando. Además, estos componentes no necesitan existir como unidades discretas reales; la correspondiente funcionalidad puede realizarse con una diversidad de diseños.

La Figura 7 muestra el sistema de la Figura 5 en relación con la aplicación 160 distribuida de ejemplo en la Figura 3. Los anfitriones 268 tienen VM 276 como se describieron anteriormente, que pueden alojar elementos de aplicación tales como servidores 168 web, instancias 170 de aplicación y servidores 172 de base de datos. Se suponen los medidores 270 y reguladores 272 pero no se muestran. Versiones más pequeñas de anfitriones 268 son las mismas que las versiones más grandes de anfitriones 268. Las versiones más grandes ilustran el detalle de los anfitriones 268 mostrados en la Figura 8. La Figura 8 muestra anfitriones 268 que alojan la aplicación 160 distribuida que se ejecutan en las VM 276. Algunos de los anfitriones residen en una subred 300, que puede un componente de una trayectoria especificada. El controlador 232 de red mantiene información que correlaciona las VM 276 con elementos de aplicación tal como servidores 172 de base de datos, instancias 170 de aplicación, etc. Estadísticas de red de las VM 276 se acumulan en almacenamiento 246 de datos de estadísticas de red. En una realización, nodos 302 de comunicación de red (por ejemplo, encaminadores, pasarelas, agregadores) proporcionan comunicaciones de red para las VM 276.

La Figura 9 muestra una vista detallada de una parte arbitraria de la Figura 8. El medidor 270 de anfitrión 268B acumula estadísticas 320 de medidor de VM 276 que se transmiten y almacenan en los datos 246 de estadísticas de red. Los nodos 302 de comunicación de red también pueden o pueden no enviar estadísticas de red, por ejemplo tasas de ancho de banda de enlaces, estáticas de interfaz, etc. El controlador 332 de red puede realizar un procedimiento 334 para almacenar cualquier estadística de red, que puede implicar recibir estadísticas y almacenar las mismas en los datos 246 de estadísticas de red.

El controlador 232 de red realiza un procedimiento para desplazar carga de tráfico de red dentro de la red. El controlador 232 de red selecciona, a partir de una recopilación 336 de especificaciones de red, una especificación 234A de red que corresponde a la aplicación distribuida. La especificación se evalúa repetidamente a continuación frente a las estadísticas de red que se relacionan con la aplicación distribuida. Siempre que el controlador 232 de red determina que se requiere un ajuste de modo que condiciones reales cumplirán con los requisitos o preferencias de la aplicación distribuida (de acuerdo con la especificación 234A de red), el controlador de red analiza estadísticas de tráfico actual o reciente de la red para identificar uno o más targets para desplazar carga de tráfico de red, posiblemente teniendo en cuenta qué VM alojan qué elementos de la aplicación. Por consiguiente, el controlador 232 de red realiza adaptación de VM en los objetivos.

Respecto a la adaptación de VM, en una realización, el controlador 232 de red, habiendo identificado la VM 276A en el anfitrión 268A como un objetivo, transmite una instrucción 338 de "realizar migración", que puede identificar la VM 172A, el anfitrión 268A de origen, y un anfitrión 268B de destino. En respuesta, los anfitriones 268A, 268B migran la VM 172A al anfitrión 268B. Se conocen técnicas para migración de VM y describen en otra parte. En general, la VM 276A se detiene en el anfitrión 268A, una imagen 338 de disco de la VM detenida se transfiere o hace disponible para el anfitrión 268B, y componentes de virtualización del anfitrión 268B a continuación reinician la VM 276A desde la

imagen 338 de disco recibida/accedida.

En una misma u otra realización, el controlador 232 de red puede determinar que el tráfico de red tiene que desplazarse ajustando la capacidad de red (por ejemplo, bajando y subiendo diversas tasas de transmisión de VM) de VM objetivos. Por ejemplo, el controlador 232 puede determinar que un requisito de trayectoria especificada (por ejemplo, una trayectoria que incluye la VM 276B) no se está satisfaciendo. El controlador 232 de red puede a continuación identificar una o más VM en anfitrión o anfitriones adecuados (por ejemplo, una VM 340 para otra aplicación, que a menudo puede estar en un anfitrión distinto del anfitrión 268B) que en la actualidad puede tener ancho de banda reducido, y a continuación transmite la instrucción 342 para aumentar/disminuir tasas de red o asignaciones de ancho de banda para las VM objetivo. El regulador 272 del anfitrión o anfitriones receptores - anfitrión 268B en este ejemplo - a continuación responde cambiando las asignaciones de las VM objetivo.

En otra realización más, el controlador de red infiere la capacidad de rendimiento de elementos de red (por ejemplo, encaminadores) fuera de los anfitriones supervisando la latencia de las transmisiones bajo diferentes cargas. Por ejemplo, la latencia es una función de ancho de banda, y si con el paso del tiempo se construye una curva de latencia/ancho de banda, es posible usar esa curva (u otra información que correlaciona latencia y ancho de banda) para informar ajustes a colocaciones de VM. También es posible inyectar paquetes adicionales a una NIC virtual para probar el rendimiento de red. Tales paquetes de sondas se inyectarían en el origen y eliminarían en el destino.

Se apreciará que puede usarse cualquiera o ambas de regulación de tasa de VM y migración de VM (o creación, borrado o suspensión). Topología de red o estadísticas notificadas por los nodos 302 de comunicación de red pueden o pueden no informar el procedimiento de toma de decisión del controlador 232 de red. Se apreciará también que la especificación de red y aprovisionamiento en efecto se produce en el nivel de aplicación, ya que la especificación específica trayectorias en términos de elementos de aplicación, estadísticas de red se notifican para y correlacionan con esos elementos, y se realiza análisis y ajuste con respecto a los elementos de aplicación. Por lo tanto, sin alcanzar necesariamente el nivel de red (y sin el privilegio requerido para hacerlo), es posible intentar garantizar tasas/anchos de banda para trayectorias a través de (y posiblemente a/desde) una aplicación distribuida. Dicho esto, en una realización el controlador 232 de red también puede emitir instrucciones a los nodos 302 de comunicación de red como medio adicional para desplazar capacidad de red a donde se necesita. En algunas nubes o centros de datos, ancho de banda de red podría tener diferentes costes en diferentes momentos (por ejemplo, hora del día, día de la semana, etc.), y el sistema puede configurarse para ajustar el uso de red para cumplir límites de costes globales o una función de minimización.

Las Figuras 10 y 11 muestran especificaciones 234A, 234B de red de aplicación de ejemplo. Las especificaciones 234A, 234B de red pueden escribirse en XML (Lenguaje de Marcas Extensible). Los elementos de XML en los ejemplos son autoexplicativos. Considérese que los componentes o elementos de la correspondiente aplicación distribuida pueden definirse, por ejemplo "Internet Gateway" "Tier1App" etc. Especificaciones de capacidad para trayectorias se definen a continuación en términos de los componentes. Por ejemplo, en la especificación 234A de red, una especificación de capacidad indica que el ancho de banda de trayectorias promedio desde Tier1Apps (y por implicación, sus VM) a la InternetGateway debería ser 1,5 MBS (megabytes por segundo). Otra especificación indica que la suma de todas las trayectorias de Tier1-Tier2 deben exceder .5 MBS. En una realización. Una especificación puede ser sensible a la dirección; como se indica por el orden de parámetros en una sentencia de especificación. Las especificaciones de red pueden definir un lenguaje de expresión rica, permitiendo trayectorias complejas y condiciones a especificarse. Pueden especificarse condiciones complejas. También pueden especificarse acciones correctivas. Por ejemplo, un requisito puede especificar una primera condición de trayectoria y otra trayectoria a dirigirse para satisfacer la condición. Un requisito puede especificar múltiples elementos en una trayectoria. Por ejemplo, un ancho de banda desde los componentes A a B a través del componente C.

CONCLUSIÓN

Realizaciones y características analizadas anteriormente pueden realizarse en forma de información almacenada en medios legibles por ordenador o dispositivo volátiles o no volátiles. Esto se considera que incluye al menos medios tales como almacenamiento óptico (por ejemplo, memoria de sólo lectura de disco compacto (CD-ROM)), medios magnéticos, memoria flash de sólo lectura (ROM) o cualquier medio actual o futuro de almacenamiento de información digital. La información almacenada puede ser en forma de instrucciones ejecutables por máquina (por ejemplo, código binario ejecutable compilado), código fuente, código de bytes o cualquier otra información que puede usarse para habilitar o configurar dispositivos informáticos para que realicen las diversas realizaciones analizadas anteriormente. Esto también se considera que incluye al menos memoria volátil tal como memoria de acceso aleatorio (RAM) y/o memoria virtual que almacena información tal como instrucciones de unidad de procesamiento central (CPU) durante ejecución de un programa que efectúa una realización, así como medios no volátiles que almacenan información que permite que un programa o ejecutable se cargue y ejecute. Las realizaciones y características pueden realizarse en cualquier tipo de dispositivo informático, incluyendo dispositivos portátiles, estaciones de trabajo, servidores, dispositivos inalámbricos móviles y así sucesivamente.

REIVINDICACIONES

1. Un procedimiento de especificación y asignación de ancho de banda de red para aplicaciones (160; 230) distribuidas que se ejecutan en una red en máquinas (114; 276) virtuales, VM, que se ejecutan en anfitriones (268) en la red, comprendiendo el procedimiento:
- 5 almacenar especificaciones (234) de red de aplicación, especificando cada especificación de red de aplicación requisitos (236) de red de una aplicación distribuida que tiene elementos de aplicación que se ejecutan en una pluralidad de las VM en una pluralidad de los anfitriones; supervisar uso de red en la red;
- 10 en base a las especificaciones de red de aplicación y la supervisión, determinar por un controlador (232) de red si no se satisface un requisito de ancho de banda de red de algunas trayectorias de red a nivel de aplicación de la aplicación distribuida, en el que el controlador de red evalúa datos (246) de estadísticas de red frente a los requisitos de ancho de banda de red de la aplicación distribuida; y
- 15 satisfacer el requisito de ancho de banda de red por el controlador (232) de red que emite instrucciones a uno o más anfitriones (268) para migrar una o más de las VM en la red y para aumentar o disminuir diversas tasas de transmisión de red de una o más VM gestionadas por un hipervisor (274) en uno de los anfitriones, en el que aumentar o disminuir tasas de transmisión de red de una aplicación distribuida ayuda a cumplir un requisito de otra aplicación distribuida.
2. El procedimiento de acuerdo con la reivindicación 1, en el que una especificación (234) de red de aplicación especifica una pluralidad de trayectorias de comunicación a través de elementos de una correspondiente aplicación (160; 230) distribuida, y el procedimiento comprende adicionalmente migrar una VM (114; 276) desde un primer anfitrión (268A) a un segundo anfitrión (268B) para satisfacer un requisito (236) de red de una de las trayectorias.
- 20 3. El procedimiento de acuerdo con la reivindicación 1, en el que la supervisión comprende capturar estadísticas de red de VM (114; 276) en hipervisores (274) en los anfitriones (268) y transmitir las estadísticas de red a almacenar.
4. El procedimiento de acuerdo con la reivindicación 3, comprendiendo además la supervisión recibir estadísticas de red notificadas por uno o más encaminadores, conmutadores o equilibradores de carga de la red.
- 25 5. El procedimiento de acuerdo con la reivindicación 3, comprendiendo además la supervisión hipervisores (274) que generan e insertan paquetes de sondas para transmisión, reciben los paquetes de sondas, miden una estadística de red de acuerdo con los paquetes de sondas, y eliminan los paquetes de sondas de modo que no se pasan a una máquina virtual.
- 30 6. El procedimiento de acuerdo con la reivindicación 1, comprendiendo además calcular el requisito de ancho de banda de red dinámicamente a base de una función de minimización de costes, en el que, de acuerdo con la función de minimización de costes, costes de ancho de banda varían como una función de tiempo.
7. Un controlador de red de distribución de uso de red entre una pluralidad de anfitriones (268) que alojan máquinas (114; 276) virtuales, VM, que se comunican a través de una red, en el que cada anfitrión tiene un hipervisor (274) que gestiona las VM que se ejecutan en el mismo, disponiéndose el controlador de red para:
- 35 acceder a información de red almacenada por uno o más componentes, comprendiendo la información de red información que especifica requisitos (236) de red de respectivas aplicaciones (160; 230) distribuidas que se ejecutan en las VM, en el que cada aplicación distribuida se ejecuta en una pluralidad de las VM;
- 40 determinar si no se satisface un requisito de ancho de banda de red de algunas trayectorias de red a nivel de aplicación de las aplicaciones distribuidas, en el que datos (246) de estadísticas de red se evalúan frente a los requisitos de ancho de banda de red de la aplicación distribuida; y
- 45 satisfacer el requisito de ancho de banda de red de las aplicaciones emitiendo instrucciones a uno o más anfitriones (268) para migrar VM entre los anfitriones y para aumentar o disminuir diversas tasas de transmisión de red de las VM por los hipervisores, en el que aumentar o disminuir tasas de transmisión de red de una aplicación distribuida ayuda a cumplir un requisito de otra aplicación distribuida.
8. El controlador de red de acuerdo con la reivindicación 7, en el que la información de red comprende una especificación de aplicación para una aplicación (160; 230) distribuida dada, la especificación de aplicación que especifica una trayectoria de comunicación de red que comprende dos componentes que representan respectivos elementos de aplicación de la aplicación distribuida, especificando adicionalmente la especificación de aplicación un requisito (236) de red para la trayectoria de comunicación.
- 50 9. El controlador de red de acuerdo con la reivindicación 7, en el que la aplicación (160; 230) distribuida comprende un primer conjunto de instancias de aplicación que corresponde a un primero de los componentes, y un segundo conjunto de instancias de aplicación que corresponde a un segundo de los componentes.
- 55 10. El controlador de red de acuerdo con la reivindicación 7, dispuesto adicionalmente para almacenar estadísticas de uso de red de VM de VM (114; 276) individuales y usar las estadísticas de uso de red de VM para seleccionar una o más VM para o bien migrarse o bien tener capacidad de red disponible aumentada o disminuida.

11. El controlador de red de acuerdo con la reivindicación 10, dispuesto adicionalmente para disminuir capacidad de red de una VM (114; 276) suspendiendo la VM.

5 12. El controlador de red de acuerdo con la reivindicación 7, dispuesto adicionalmente para identificar una VM (114; 276) particular de acuerdo con uno de los requisitos de red (236), seleccionar una estadística de red de la VM particular que se notificó por un anfitrión (268) que aloja la VM particular, y comparar la estadística de red frente al uno de los requisitos de red.

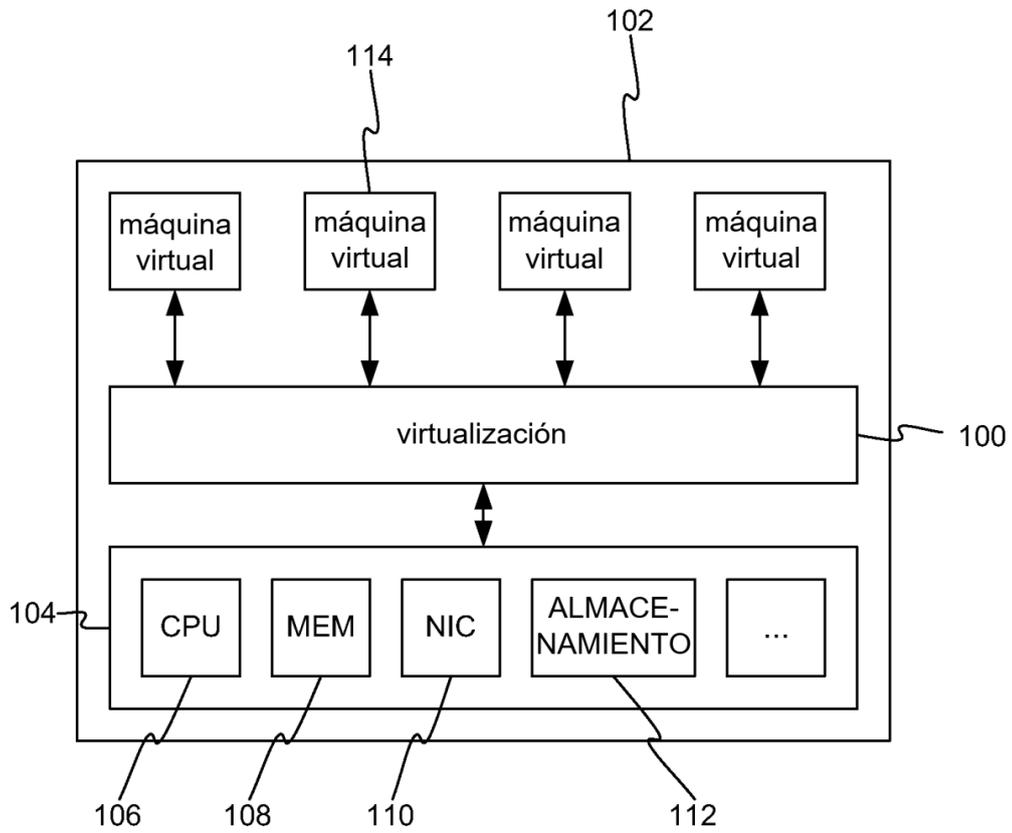


FIG. 1

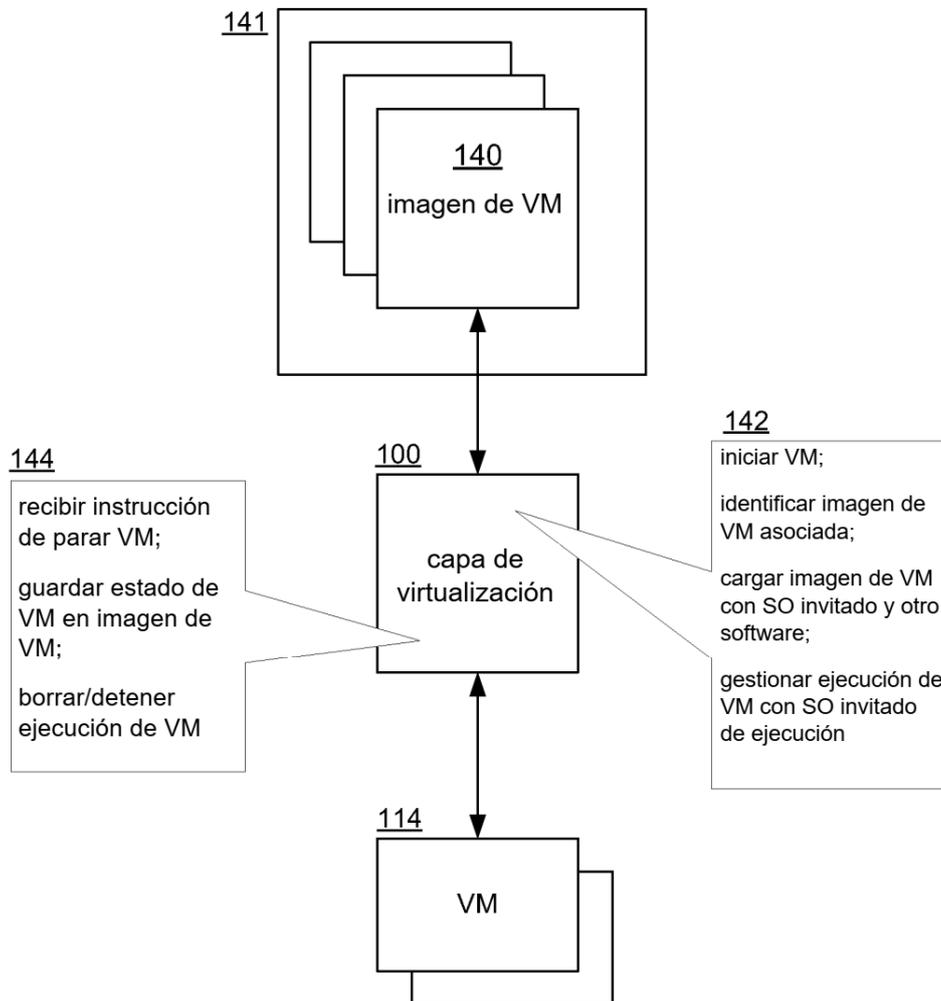
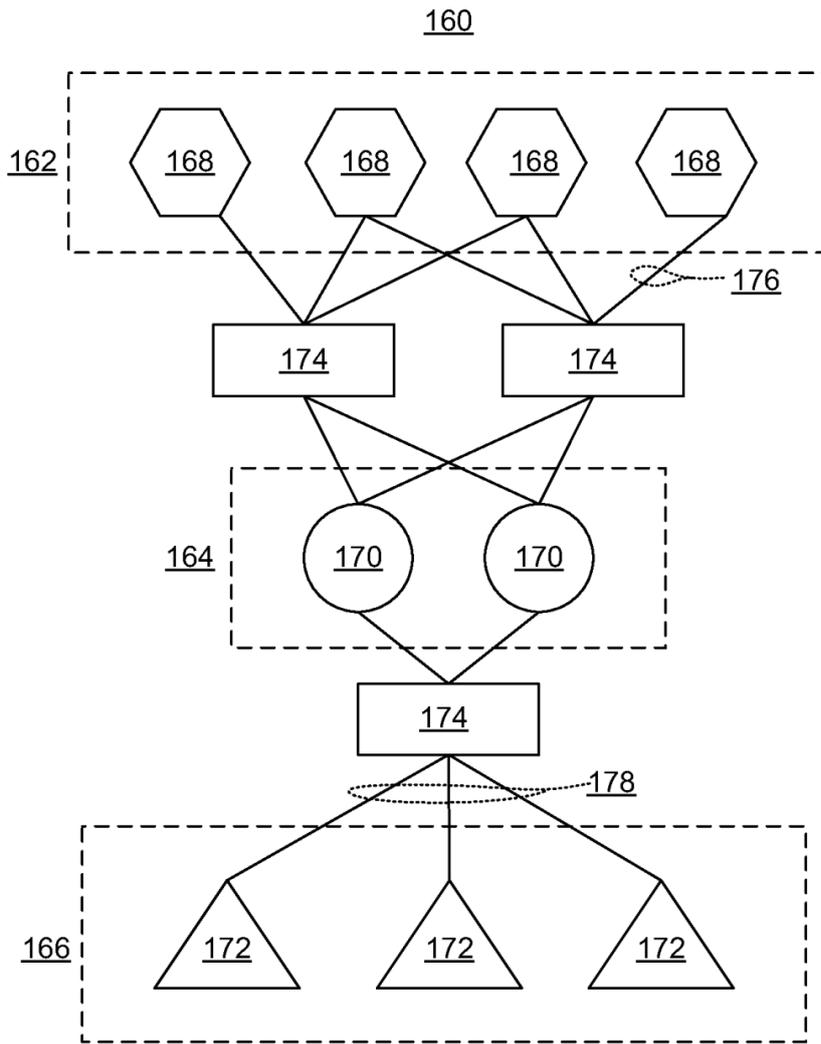


FIG. 2



aplicación distribuida
de ejemplo

FIG. 3

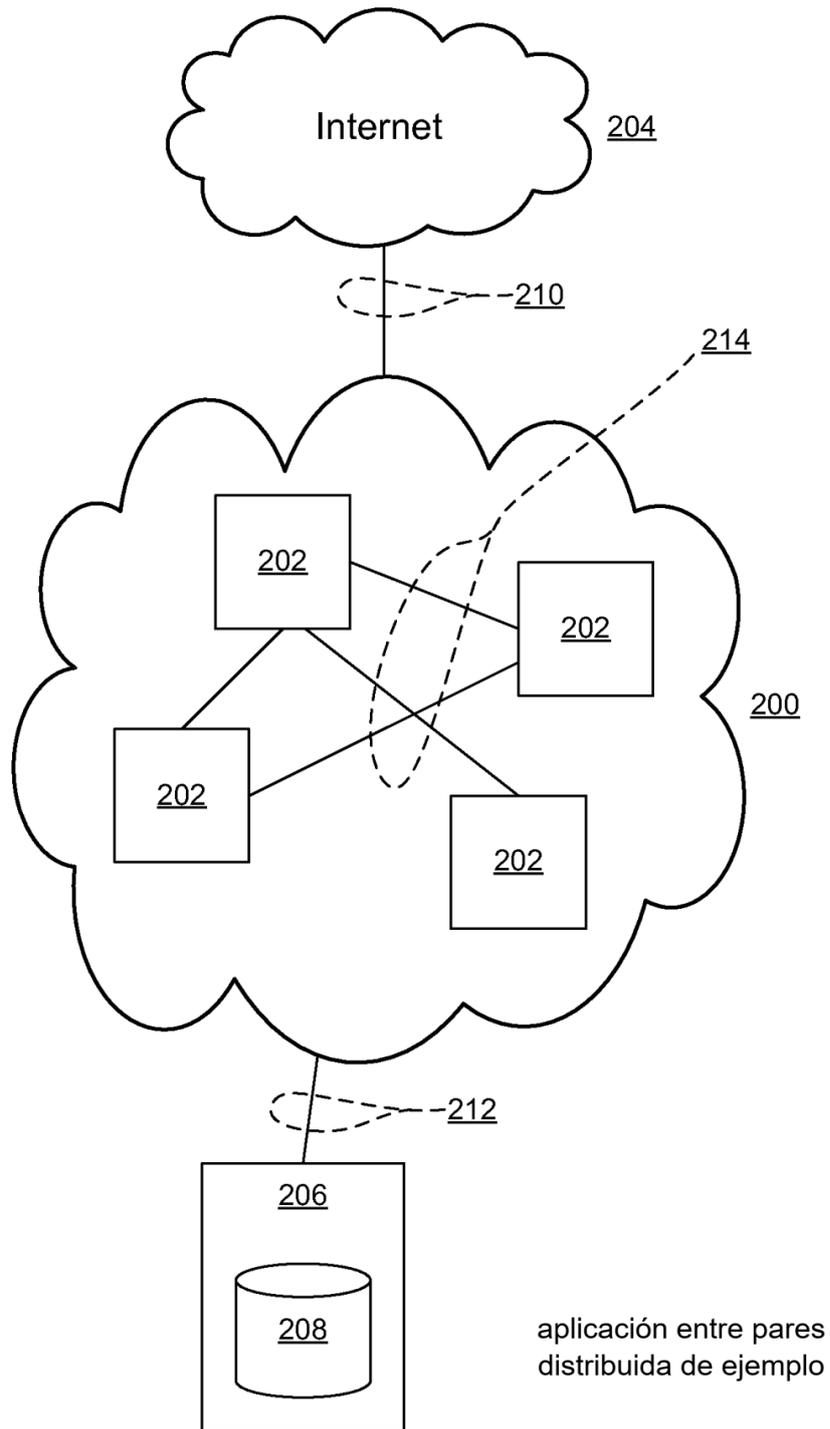


FIG. 4

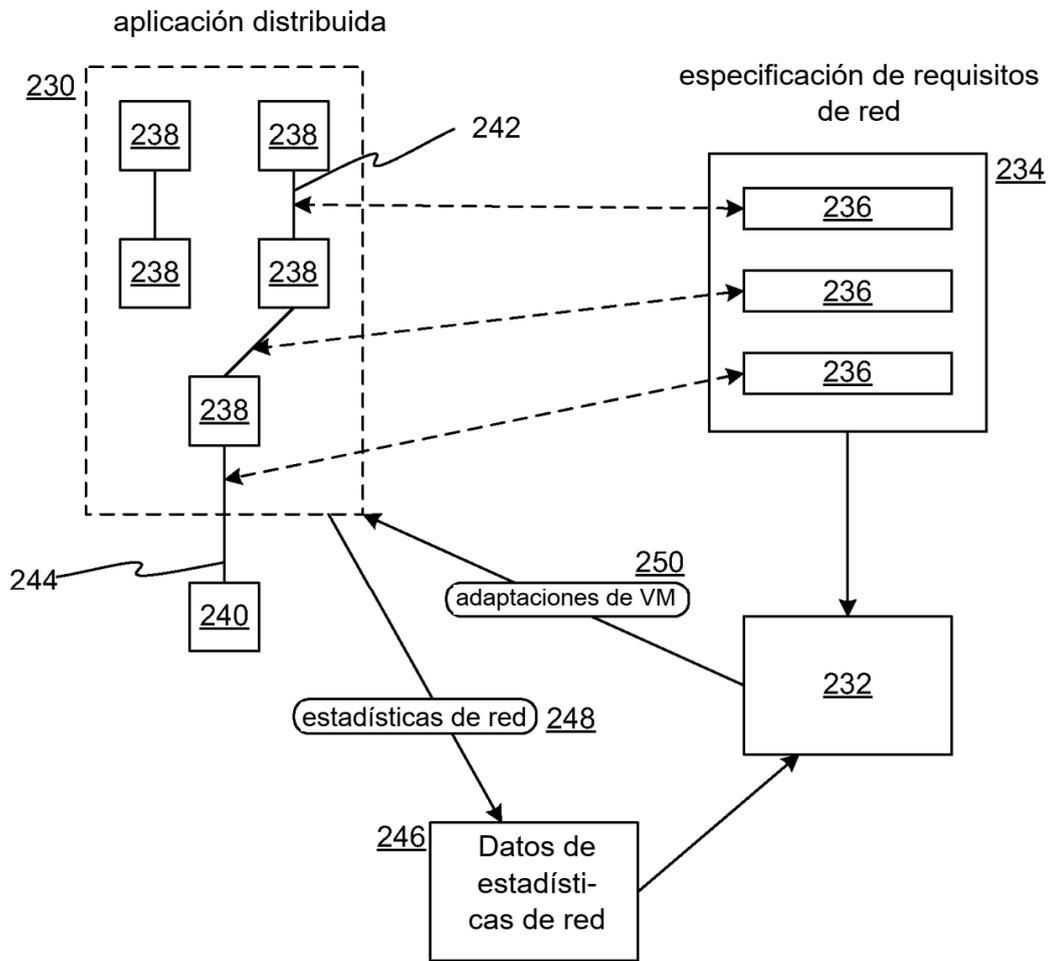


FIG. 5

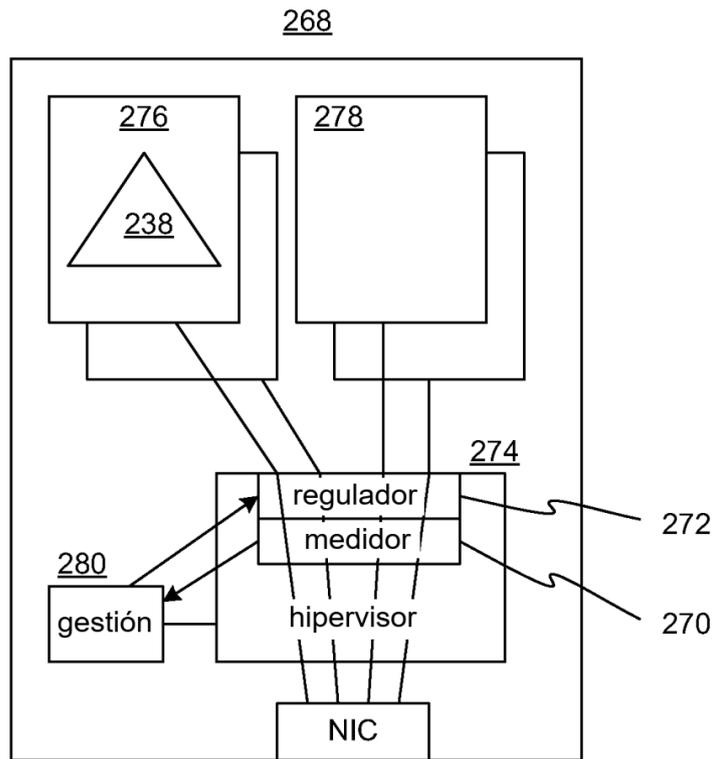


FIG. 6

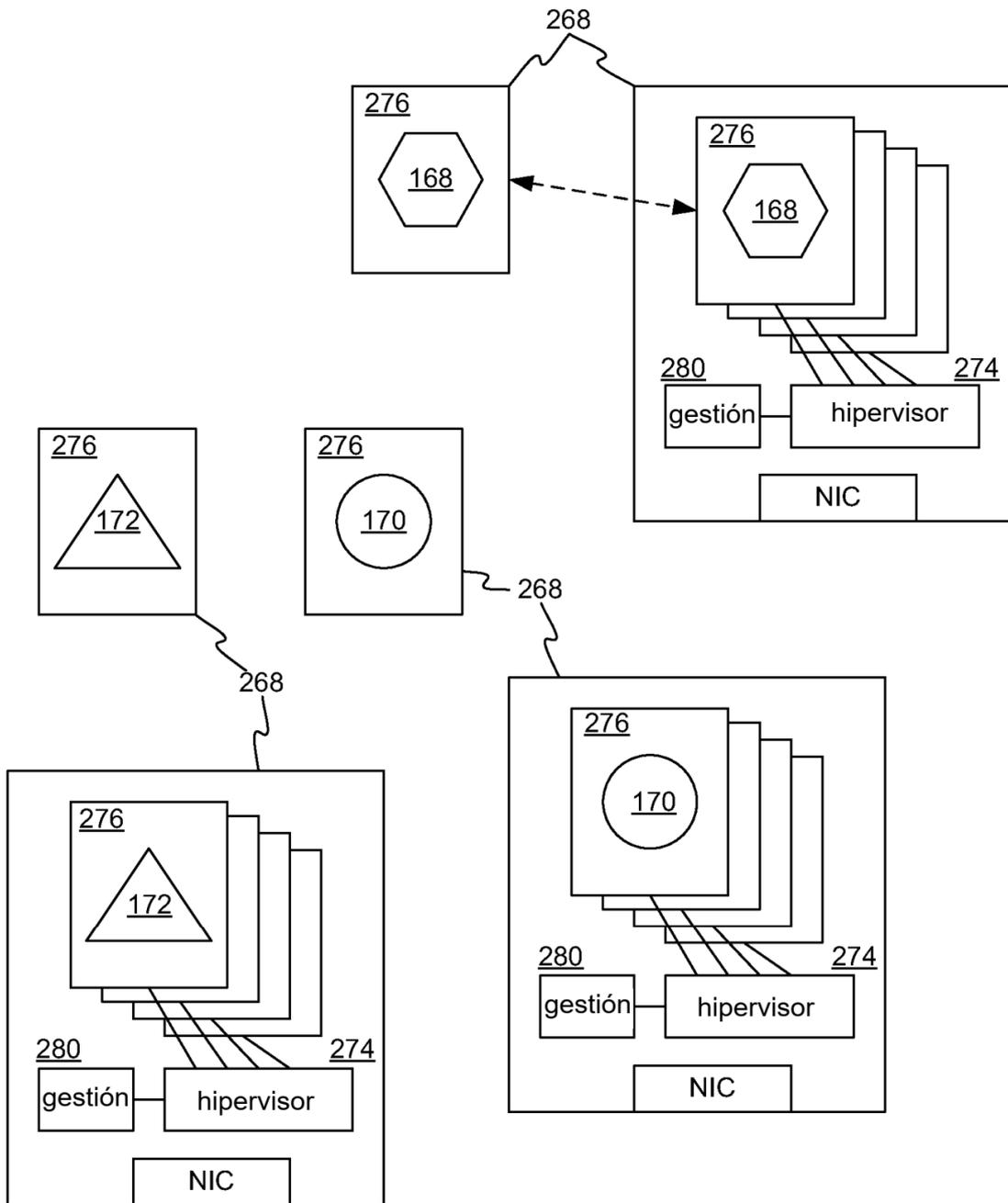


FIG. 7

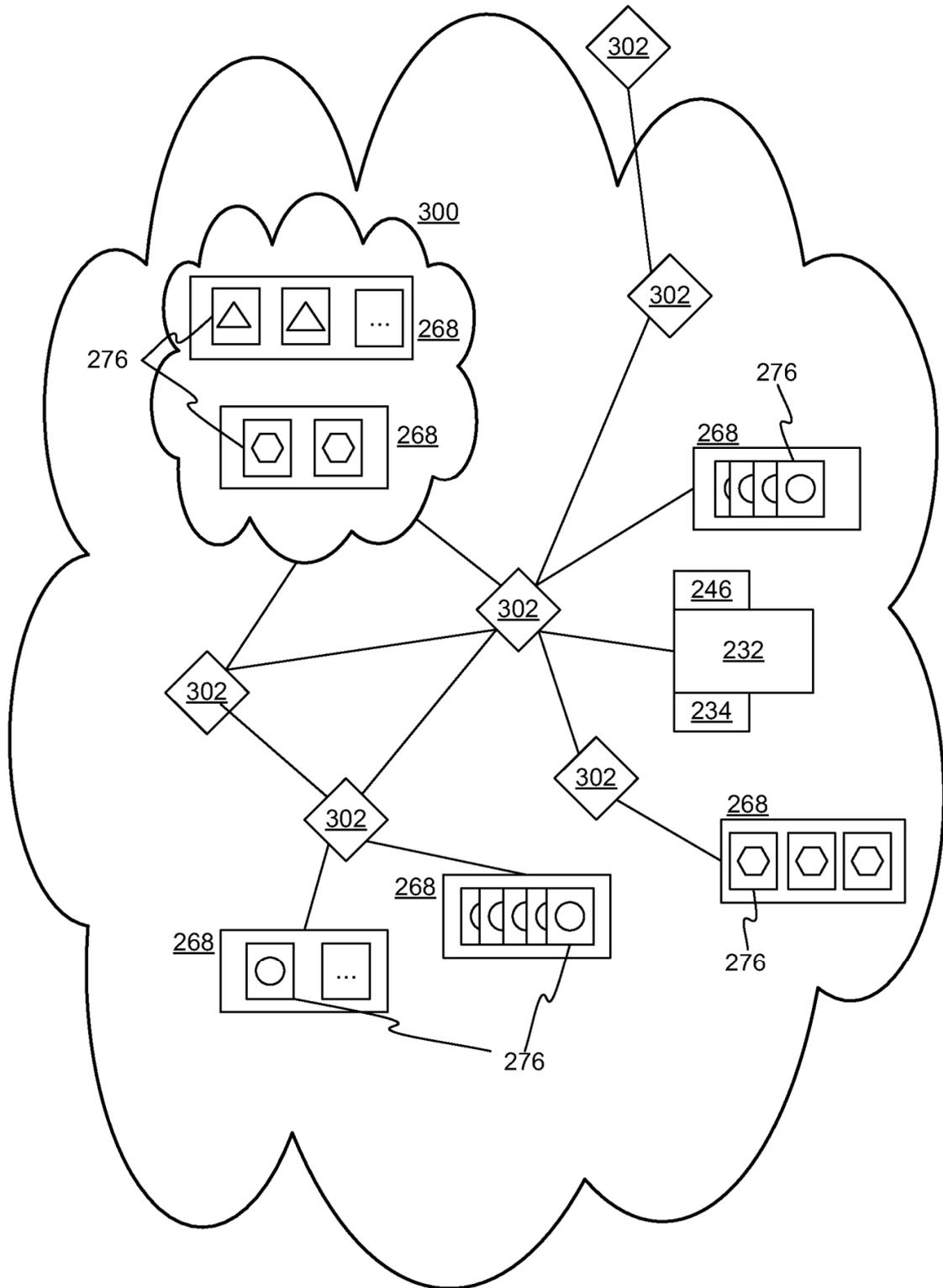


FIG. 8

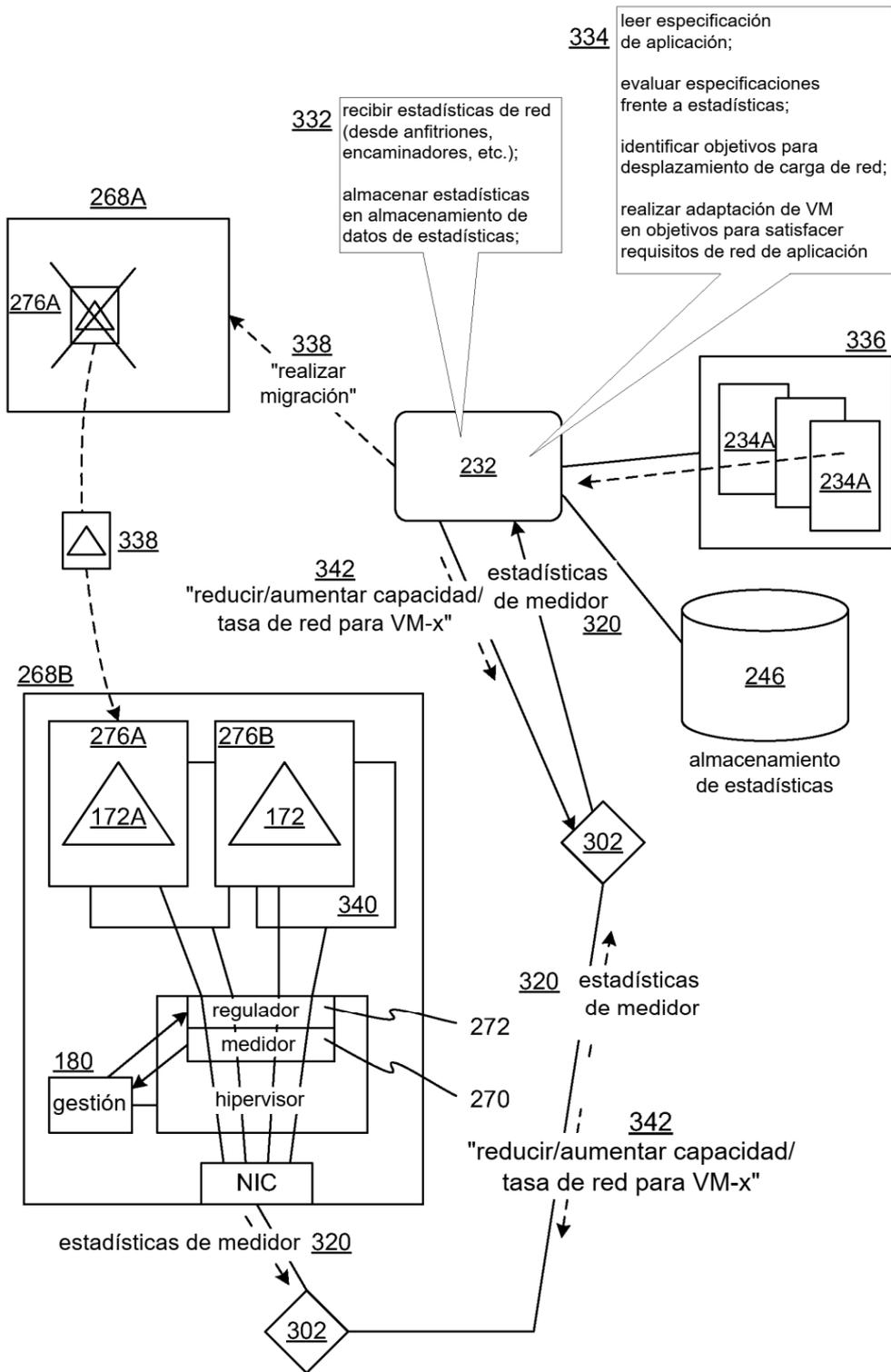


FIG. 9

234A

```

<applicationA>
  <components>
    <component property1...>
      InternetGateway
    </component>
    <component property1...>
      Tier1App
    </component>
    <component property1...>
      Tier2App
    </component>
    <component>
      LoadBalancer
    </component>
    ...
  </components>
  <capacity specifications>
    <capacity specification>
      Min(Average(Tier1App), InternetGateway) = 1.5 MBS
    </capacity specification>
    <capacity specification>
      Sum(Connection(Tier1, Tier2)) > .5 MBS
    </capacity specification>
    <capacity specification>
      Each(Tier2, SQLDB) > .1 MBS
    </capacity specification>
    <capacity specification>
      Min(LoadBalancer, *) > .1 MBS
      Min(*, LoadBalancer) > .4 MBS
    </capacity specification>
  </capacity specifications>
  ...
</ApplicationA>

```

FIG. 10

234B

```
<P2P Application>
  <components>
    <component property1...>
      Gateway
    </component>
    <component property1...>
      Peer
    </component>
    <component property1...>
      Client
    </component>
    ...
  </components>
  <capacity specifications>
    <capacity specification>
      Avg(Peer, Peer) = .2 MBS
    </capacity specification>
    <capacity specification>
      Each(Peer, Gateway) > .1 MBS
    </capacity specification>
    <capacity specification>
      Avg(Client, Gateway, Peer) < 0.05 MBS
    </capacity specification>
  </capacity specifications>
  ...
</P2P Application>
```

FIG. 11