



República Federativa do Brasil
Ministério da Economia
Instituto Nacional da Propriedade Industrial

(11) PI 1009332-0 B1



(22) Data do Depósito: 18/03/2010

(45) Data de Concessão: 12/05/2020

(54) Título: MÉTODO IMPLEMENTADO POR COMPUTADOR E SISTEMA ESTRUTURADO PARA FORNECER FUNCIONALIDADE DE REDE LÓGICA PARA REDES DE COMPUTADOR

(51) Int.Cl.: G06F 15/173.

(30) Prioridade Unionista: 30/03/2009 US 12/414,260.

(73) Titular(es): AMAZON TECHNOLOGIES, INC..

(72) Inventor(es): DANIEL T. COHN; ERIC JASON BRANDWINE; ANDREW J. DOANE.

(86) Pedido PCT: PCT US2010027822 de 18/03/2010

(87) Publicação PCT: WO 2010/114715 de 07/10/2010

(85) Data do Início da Fase Nacional: 08/09/2011

(57) Resumo: MÉTODO IMPLEMENTADO POR COMPUTADOR E SISTEMA ESTRUTURADO PARA FORNECER FUNCIONALIDADE DE REDE LÓGICA PARA REDES DE COMPUTADOR São descritas técnicas para fornecimento da funcionalidade de rede lógica para redes de computador gerenciadas, como, por exemplo, para redes de computador virtuais fornecidas em nome de usuários ou outras entidades. Em algumas situações, um usuário pode configurar a sua ou especificar de outra maneira a topologia de rede para uma rede de computador virtual, como, por exemplo, uma topologia de rede lógica que separa múltiplos nós de computação da rede de computador virtual dentro de múltiplas sub-redes lógicas e/ou que especifica um ou mais dispositivos de rede lógica para a rede de computador virtual. Depois que a topologia de rede é especificada para a rede de computador virtual, a funcionalidade de rede lógica correspondente uma topologia de rede pode ser fornecida de várias maneiras, como, por exemplo, sem implantar fisicamente a topologia de rede para a rede de computador virtual. Em algumas situações, os nós de computação podem incluir nodos de máquina virtual hospedados em uma ou mais máquinas de computação físicas ou sistemas, como, por exemplo, através de um ou mais ou em nome de um ou mais usuários.(...).

“MÉTODO IMPLEMENTADO POR COMPUTADOR E SISTEMA ESTRUTURADO PARA FORNECER FUNCIONALIDADE DE REDE LÓGICA PARA REDES DE COMPUTADOR”

FUNDAMENTOS DA INVENÇÃO

5 Muitas companhias e outras organizações operam redes de computador que interconectam diversos sistemas de computação para suportar suas operações, com os sistemas de computação alternativamente co-localizados (ex., como parte de uma rede local privada) ou então localizados em múltiplas localizações geográficas distintas (ex., conectados através de uma ou mais redes intermediárias privadas ou compartilhadas). Por exemplo,

10 centros de dados que abrigam números significativos de sistemas de computação interconectados se tornaram comuns, como centros de dados privados que são operados por e em nome de uma única organização, bem como centros de dados públicos que são operados por entidades como empresas. Alguns operadores de centros de dados públicos fornecem acesso de rede, energia e dependências para instalação segura de hardware de propriedade de vários clientes, enquanto outros operadores de centros de dados públicos fornecem instalações de “serviço completo” que também incluem recursos de hardware disponibilizados para uso por seus clientes. No entanto, visto que a escala e escopo de centros de dados típicos e redes de computador aumentou, a tarefa de fornecer, administrar e administrar recursos de computação físicos relacionados se tornou cada vez mais complicada.

20 O advento de tecnologias de virtualização para hardware de mercadoria forneceu alguns benefícios com relação à gestão de recursos de computação em grande escala para muitos clientes com necessidades diversas, permitindo que vários recursos de computação fossem compartilhados de modo eficaz e seguro entre diversos clientes. Por exemplo, tecnologias de virtualização como aquelas fornecidas por VMWare, XEN, ou User-Mode Linux

25 podem permitir que uma máquina de computação física única seja compartilhada entre vários usuários fornecendo a cada usuário uma ou mais máquinas virtuais hospedadas pela única máquina de computação física, com cada máquina virtual sendo uma estimulação de software que age como um sistema de computação lógico distinto que fornece aos usuários a ilusão de que são os únicos operadores e administradores de um determinado recurso de

30 computação de hardware, enquanto também fornece isolamento e segurança de aplicativos entre as diversas máquinas virtuais. Além disso, algumas tecnologias de virtualização são capazes de fornecer recursos virtuais que alcançam um ou mais recursos físicos, como uma máquina virtual única com múltiplos processadores virtuais que na verdade alcança múltiplos sistemas de computação físicos distintos.

35 Breve Descrição dos Desenhos

A Figura 1 é um diagrama de rede que ilustra um exemplo de modalidade de configuração e gestão de comunicações entre nós de computação que pertencem a uma rede de

computador virtual.

As Figuras 2A a 2C ilustram exemplos de gestão de comunicações entre nós de computação de uma rede de computador de sobreposição (*overlay*) virtual.

5 A Figure 2D ilustra um exemplo de configuração de endereços de rede de substratos subjacentes para permitir a incorporação de endereços de redes virtuais para uma rede de sobreposição.

A Figura 3 é um diagrama de bloco que ilustra exemplos de sistemas de computação para executar uma modalidade de um sistema para gestão de comunicações entre nós de computação.

10 As Figuras 4A e 4B ilustram um fluxograma de um exemplo de modalidade de uma rotina do Gerenciamento de Sistemas ONM.

As Figuras 5A e 5B ilustram um fluxograma de um exemplo de modalidade de uma rotina de Gerenciador de Comunicação ONM.

Descrição Detalhada

15 São descritas técnicas para fornecer funcionalidade de rede lógica para redes de computador de gestão, como para redes de computador virtuais que são fornecidas em nome de usuários ou outras entidades. Em pelo menos algumas modalidades, as técnicas permitem que um usuário configure ou de forma diversa especifique uma topologia de rede para uma rede de computador virtual sendo fornecida para o usuário, como uma topologia
20 de rede lógica que separa múltiplos nós de computação da rede de computador virtual em múltiplas subredes lógicas e/ou que especifica um ou mais dispositivos de rede lógica que são, cada um, associados a um grupo especificado dos múltiplos nós de computação.

Após uma topologia de rede ser especificada para uma rede de computador virtual, a funcionalidade de rede lógica correspondente à topologia de rede poderá ser fornecida em
25 diversas formas, como sem implementação física da topologia de rede para a rede de computador virtual. Especificamente, em pelo menos algumas modalidades, as comunicações entre múltiplos nós de computação da rede de computador virtual são administradas para emular funcionalidade que seria fornecida por dispositivos de rede lógica especificados se os mesmos estivessem presentes fisicamente e/ou para de outra forma emular funcionalida-
30 de correspondente a uma topologia de rede especificada se esta fosse implementada fisicamente, conforme descrito mais detalhadamente abaixo. Em pelo menos algumas modalidades, algumas ou todas as técnicas descritas aqui são automaticamente realizadas por modalidades de um sistema de gerenciador de redes de sobreposição.

Uma rede local virtual ou outra rede de computador virtual entre múltiplos nós de
35 computação pode ser fornecida de várias formas em várias modalidades, como, por exemplo, criando uma rede de sobreposição usando uma ou mais redes físicas intermediárias que separam os nós de computação múltiplos. Em tais modalidades, a(s) rede(s) física(s)

intermediária(s) podem ser usadas como uma rede de substrato em que a rede de computador virtual é fornecida, com mensagens entre nós de computação da rede de sobreposição de computador virtual sendo entregues à(s) rede(s) física(s) intermediária(s), mas com os nós de computação desconhecendo a existência e uso da(s) rede(s) física(s) intermediária(s) em pelo menos algumas tais modalidades. Por exemplo, os múltiplos nós de computação podem, cada um, ter um endereço de rede de substrato física distinto que corresponde a uma localização do nó de computação dentro da(s) rede(s) física(s) intermediária(s), como um endereço de rede substrato IP ("Protocolo de Internet") (ex., um endereço de rede IP que é especificado em conformidade com IPv4, ou "Protocolo de Internet versão 4," ou em conformidade com IPv6, ou "Protocolo de Internet versão 6," de modo a refletir o protocolo de rede usado pelas redes físicas intermediárias). Em outras modalidades, uma rede de substrato em que uma rede de computador virtual é sobreposta pode incluir ou ser composta de uma ou mais redes de computador virtuais, como outras redes de computador virtuais complementadas por alguém ou mais terceiros (ex., por um operador ou provedor de Internet ou infraestrutura de telecomunicação).

Quando nós de computação são selecionados para participar de uma rede de computador virtual sendo fornecida pelo sistema gerenciador de redes de sobreposição e sendo sobrepostos em uma rede de substrato, cada nó de computação pode receber um ou mais endereços de rede virtual para a rede de computador virtual fornecida que não são relacionados àqueles endereços de rede de substrato de nós de computação, como a partir de uma variedade de endereços de rede virtual usados para a rede de computador virtual fornecida – em pelo menos algumas modalidades e situações, a rede de computador virtual sendo fornecida pode ainda usar um protocolo de rede que é diferente do protocolo de rede usado pela rede de substrato (ex., com a rede de computador virtual usando o protocolo de rede IPv4, e a rede de computador de substrato usando o protocolo de rede IPv6). Os nós de computação da rede de computador virtual intercomunicam usando os endereços de rede virtual (ex., enviando uma comunicação a outro nó de computação de destino especificando aquele endereço de rede virtual do nó de computação de destino como o endereço de rede de destino para a comunicação), mas a rede de substrato pode ser configurada para traçar a rota ou de outra forma encaminhar comunicações baseadas nos endereços de rede de substrato (ex., através de dispositivos roteadores de rede física e outros dispositivos de rede física da rede de substrato). Se assim for, a rede de computador virtual pode ser implementada a partir da borda das redes físicas intermediárias, através da modificação de comunicações que entram nas redes físicas intermediárias para usar endereços de rede de substrato que são baseados no protocolo de rede da rede de substrato, e modificando as comunicações que saem das redes físicas intermediárias para usar endereços de rede virtual que são baseados no protocolo de rede da rede de computador virtual. Detalhes adicio-

nais relacionados ao fornecimento de tal rede de sobreposição de computador virtual são incluídos abaixo.

Em pelo menos algumas modalidades, uma modalidade de um sistema Gerenciador de Rede de sobreposição ("ONM") fornece redes de sobreposição de computador virtual para clientes e outros usuários, como, por exemplo, fornecendo e usando diversos nós de computação que estão em uma ou mais localizações geográficas (ex., em um ou mais centros de dados) e que são interconectados através de uma ou mais redes físicas intermediárias. O sistema ONM pode usar vários módulos gerenciador de comunicação na borda de uma ou mais redes físicas intermediárias para gerenciar comunicações para as várias redes de sobreposição de computador virtual conforme entram e saem das redes físicas intermediárias, e podem usar um ou mais módulos gerenciadores de sistema para coordenar outras operações do sistema ONM. Por exemplo, para permitir os módulos gerenciadores de comunicação para gerenciar comunicações para as redes de sobreposição de computador virtual sendo fornecidas, o sistema ONM pode rastrear e usar várias informações sobre os nós de computação de cada rede de computador virtual, para mapear o endereço de rede física de substrato de cada nó de computação para um ou mais endereços de rede virtual de sobreposição associados ao nó de computação. Tal mapeamento e outras informações podem ser armazenadas e propagadas de várias formas em diversas modalidades, incluindo centralmente ou de forma distribuída, conforme discutido mais detalhadamente abaixo.

Além disso, para fornecer redes de computador virtuais para usuários e outras entidades de forma desejada, o sistema ONM permite aos usuários e outras entidades interagir com o sistema ONM em pelo menos algumas modalidades para configurar uma variedade de tipos de informações para redes de computador virtuais que são fornecidas pelo sistema ONM em nome dos usuários e outras entidades, e pode rastrear e usar tais informações de configuração como parte do fornecimento daquelas redes de computador virtuais. As informações de configuração para uma rede de computador virtual específica tendo múltiplos nós de computação podem incluir, por exemplo, um ou mais da seguinte lista não-exclusiva: uma quantidade dos nós de computação múltiplos para incluir como parte da rede de computador virtual; uma variedade ou outro grupo de múltiplos endereços de rede virtual para associar aos nós de computação múltiplos da rede de computador virtual; endereços de rede virtual específicos para associada a nós de computação ou grupos específicos de nós de computação relacionados; um tipo de pelo menos alguns dos múltiplos nós de computação da rede de computador virtual, para refletir quantidades e/ou tipos de recursos de computação para serem incluídos ou de forma diversa disponibilizados aos nós de computação; uma localização geográfica em que alguns ou todos os nós de computação da rede de computador virtual devam ser localizados; etc. Além disso, as informações de configuração para uma rede de computador virtual podem ser especificadas por um usuário ou outra entidade de

várias formas em várias modalidades, como, por exemplo, executando programa do usuário ou outra entidade que interage com uma API ("interface de programação de aplicativo") fornecida pelo sistema ONM para tal finalidade e/ou através de um usuário que usa interativamente uma GUI ("interface gráfica de usuário") fornecida pelo sistema ONM para aquela finalidade.

Além disso, conforme observado anteriormente, a funcionalidade de rede lógica para redes de computador de gestão pode ser fornecida pelo sistema ONM de diversas formas em várias modalidades, com sistema ONM realizando várias ações para suportar tal funcionalidade de rede lógica. Por exemplo, um usuário ou outra entidade pode interagir com o sistema ONM em pelo menos algumas modalidades para configurar várias informações sobre uma topologia de rede de uma rede de computador virtual específica, e o sistema ONM pode rastrear e usar tais informações de configuração de topologia de rede como parte do fornecimento de tal rede de computador virtual. As informações de configuração de topologia de rede para uma rede de computador virtual podem incluir vários tipos de informações, inclusive a seguinte lista não-exclusiva: um arranjo especificamento dos múltiplos nós de computação da rede de computador virtual, de modo a ter um primeiro sub-conjunto dos múltiplos nós de computação sendo parte de uma primeira sub-rede que é associada a um primeiro dispositivo roteador de rede especificado, e para ter um segundo sub-conjunto dos múltiplos nós de computação sendo parte de uma segunda sub-rede que é associada a um segundo dispositivo roteador de rede especificado, etc; um ou mais dispositivos roteadores de rede especificados ou outros dispositivos de rede que devem operar como parte da rede de computador virtual, para suportar nós de computação específicos e/ou para realizar funções indicadas específicas; etc.

O sistema ONM pode tomar várias ações para suportar uma topologia de rede que é especificada para uma rede de computador virtual específica. Em particular, em pelo menos algumas modalidades, o sistema ONM pode emular funcionalidade de rede lógica que corresponde à topologia de rede específica para uma rede de computador virtual, mas sem implementar fisicamente alguma ou toda a topologia de rede especificada. Como um exemplo, o sistema ONM pode usar múltiplos módulos gerenciadores de comunicação para gerenciar transparentemente comunicações enviadas por e para os nós de computação da rede de computador virtual de modo que emule funcionalidade que seria fornecida por um ou mais dispositivos de rede especificados da topologia de rede se eles fossem implementados fisicamente para a rede de computador virtual e fossem usados para rotear ou de forma diversa encaminhar as comunicações. Além disso, o sistema ONM pode usar múltiplos módulos gerenciadores de comunicação para emular respostas a solicitações de rede feitas por nós de computação na forma de um dispositivo de rede físico local, como para responder a solicitações para pingar, questionários de SNMP ("Protocolo de Gestão de Rede Sim-

ples), etc. Desta forma, o sistema ONM pode fornecer funcionalidade de rede lógica que corresponde a uma topologia de rede especificada para uma rede de computador virtual, mas sem os nós de computação da rede de computador virtual (ou o usuário relacionado ou outra entidade) tendo conhecimento de que a topologia de rede especificada não é fisicamente implementada para a rede de computador virtual. Além disso, conforme descrito mais
5 detalhadamente abaixo, em pelo menos algumas modalidades, múltiplos módulos do sistema ONM podem operar juntos de forma distribuída para fornecer funcionalidade correspondente a um dispositivo de rede lógica específico, de modo que nenhum módulo único ou dispositivo físico seja unicamente responsável por emular um dispositivo de rede lógica específico. Detalhes adicionais relacionados ao fornecimento de uma funcionalidade de rede
10 lógica para uma rede de computador virtual em conformidade com as informações de configuração específica são incluídos abaixo.

Em pelo menos algumas modalidades, os nós de computação entre os quais as comunicações são gerenciadas podem ser sistemas de computação físicos e/ou podem ser
15 máquinas virtuais que são cada uma hospedada em um ou mais sistemas de computação físicos, e as comunicações podem incluir transmissões de dados (ex., mensagens, pacotes, quadros, correntes etc.) em vários formatos. Conforme observado anteriormente, alguns ou todos os nós de computação usados para uma rede de sobreposição de computador virtual fornecida podem em algumas modalidades ser fornecidos pelo sistema ONM para uso pelos
20 usuários, enquanto em outras modalidades alguns ou todos os nós de computação podem, em vez disso, ser fornecidos por um usuário que usa aqueles nós de computação. Além disso, em pelo menos algumas situações, uma modalidade do sistema ONM pode ser parte de, ou de outra forma, associada a um serviço de execução de programa (ou "PES") que executa múltiplos programas em nome de vários clientes ou outros usuários do serviço, como um serviço de execução de programa que usa múltiplos sistemas de computação em
25 várias redes físicas (ex., múltiplos sistemas de computação e redes dentro de um centro de dados). Em pelo menos algumas tais modalidades, as redes de computador virtuais às quais os nós de computação pertencem podem ser selecionadas com base em usuários associados, como, com base nos nós de computação que executam programas em nome de um
30 usuário ou outra entidade. Além disso, em algumas situações, uma modalidade do sistema ONM pode ser parte de, ou de outra forma, associada a um serviço de rede configurável (ou "CNS") que fornece redes de computador privadas configuráveis para vários clientes ou outros usuários do serviço, como, por exemplo, usando técnicas de cloud computing com múltiplos sistemas de computação que são fornecidos em várias redes físicas (ex., múltiplos
35 sistemas de computação físicos e redes dentro de um centro de dados).

Como observado anteriormente, uma rede de computador virtual pode, em algumas modalidades, ser fornecida como uma rede de sobreposição que usa uma ou mais redes

físicas intermediárias como uma rede de substrato, e uma ou mais tais redes de superposição de computador virtual podem ser implementadas através da rede de substrato de várias formas em várias modalidades. Por exemplo, em pelo menos algumas modalidades, as comunicações entre os nodos de uma rede de sobreposição de computador virtual são gerenciadas enviando aquelas comunicações através da rede de substrato sem encapsular as comunicações, como, incorporando informações de endereço de rede virtual para um nó de computação da rede de computador virtual (ex., o endereço de rede virtual do nó de computação de destino) em um espaço de endereço de rede física maior usado para um protocolo de rede de uma ou mais redes físicas intermediárias. Como um exemplo ilustrativo, uma rede de computador virtual pode ser implementada usando endereços de rede IPv4 de 32 bits, e aqueles endereços de rede virtual de 32 bits podem ser incorporados como parte dos endereços de rede IPv6 de 128 bits usados por uma ou mais redes físicas intermediárias, por pacotes de comunicação de novo cabeçalho ou outras transmissões de dados (ex., usando Stateless IP/ICMP Translation, ou SIIT), ou de outra forma modificando tais transmissões de dados para traduzi-las a partir de um primeiro protocolo de rede para o qual são configuradas para um segundo protocolo de rede distinto. Como outro exemplo ilustrativo, tanto a rede de computador virtual como a rede de computador de substrato podem ser implementadas usando o mesmo protocolo de endereçamento de rede (ex., IPv4 ou IPv6), e transmissões de dados enviadas através da rede de sobreposição de computador virtual usando endereços de rede virtual podem ser modificadas para usar diferentes endereços de rede física correspondentes à rede de substrato enquanto as transmissões são enviadas através da rede de substrato, mas com os endereços de rede virtuais originais sendo armazenados nas transmissões de dados modificadas ou de outra forma rastreados para que as transmissões de dados possam ser rearmazenadas à sua forma original quando saírem da rede de substrato. Em outras modalidades, pelo menos algumas das redes de computador de sobreposição podem ser implementadas usando encapsulação de comunicações. Detalhes adicionais relacionados a SIIT são disponíveis em "Solicitação para Comentários 2765 - Stateless IP/ICMP Translation Algorithm", Fevereiro de 2000, em ferramentas <dot>ietf<dot>org<slash>html <slash>rfc2765 (onde <dot> e <slash> são substituídos pelos caracteres correspondentes àqueles nomes), que é incorporado no presente documento por referência integralmente. Mais geralmente, em algumas modalidades, ao implementar uma primeira rede de sobreposição usando uma segunda rede de substrato, um endereço de rede de N-bits que é especificado para a primeira rede de sobreposição em conformidade com um primeiro protocolo de endereçamento de rede pode ser incorporado como parte de outro endereço de rede de M-bits que é especificado para a segunda rede de substrato em conformidade com um segundo protocolo de endereçamento de rede, com "N" e "M" sendo quaisquer números inteiros que correspondem a protocolos de endereçamento de rede.

Além disso, em pelo menos algumas modalidades, um endereço de rede de N-bits pode ser incorporado em outro endereço de rede usando mais ou menos do que N-bits do outro endereço de rede, como se um grupo de endereços de rede de N-bits de interesse pudesse ser representado usando um número menor de bits (ex., com rótulos ou identificadores de L-bits sendo mapeados para endereços de rede de N-bits específicos e incorporados nos outros endereços de rede, onde “L” é menor do que “N”).

Vários benefícios podem ser obtidos pela incorporação de informações de endereço de rede virtual em endereços de rede de substrato para uma rede de substrato física subjacente, incluindo permitindo um de sobreposição da rede de computador virtual na rede de substrato física sem encapsular comunicações ou configurar dispositivos de rede física da rede de substrato física, conforme discutido mais detalhadamente abaixo. Além disso, outras informações podem semelhantemente ser incorporadas no espaço de endereço de rede física maior para uma comunicação entre nós de computação em pelo menos algumas modalidades e situações, como um identificador específico para uma rede de computador virtual específica que inclui aqueles nós de computação (ex., uma rede de computador virtual para um usuário ou outra entidade em cujo nome aqueles nós de computação operam). Detalhes adicionais relacionados ao fornecimento de tais redes de computador virtuais através do uso de redes de sobreposição são incluídos abaixo.

Além disso, além de gerenciar topologias de rede configuradas para redes de computador virtuais fornecidas, o sistema ONM pode usar as técnicas descritas para fornecer vários outros benefícios em várias situações, como limitar comunicações para e/ou a partir de nós de computação de uma rede de computador virtual específica para outros nós de computação que pertencem àquela rede de computador virtual. Desta forma, os nós de computação que pertencem a múltiplas redes de computador virtuais podem compartilhar partes de uma ou mais redes físicas intermediárias, enquanto ainda mantém isolamento de rede para nós de computação de uma rede de computador virtual específica. Além disso, o uso das técnicas descritas também permite que nós de computação sejam facilmente adicionados e/ou removidos de uma rede de computador virtual, para permitir que um usuário modifique dinamicamente o tamanho de uma rede de computador virtual (ex., para modificar dinamicamente a quantidade de nós de computação para refletir uma quantidade de necessidade atual para mais ou menos recursos de computação). Além disso, o uso das técnicas descritas também suporta mudanças a uma rede de substrato subjacente - por exemplo, se a rede de substrato subjacente for expandida para incluir nós de computação adicionais em localizações geográficas adicionais, as redes de computador virtuais novas ou existentes sendo fornecidas poderão sem alterações usar aqueles nós de computação adicionais, uma vez que a rede de substrato subjacente roteará as comunicações para e a partir dos endereços de rede de substrato para aqueles nodos de comunicação adicionais da mesma forma

que para outros nós de computação de rede de substrato existentes anteriormente. Em pelo menos algumas modalidades, a rede de substrato subjacente pode ser de qualquer tamanho (ex., alcançando diversos países ou continentes), independente de latência de rede entre nós de computação em diferentes localizações.

Para fins de ilustração, algumas modalidades são descritas abaixo em que tipos específicos de nós de computação, redes, comunicações, topologias de rede e operações de configuração são realizadas. Estes exemplos são fornecidos para fins de ilustração e são simplificados para serem breves, e as técnicas inventivas podem ser usados em uma ampla variedade de outras situações, algumas das quais são discutidas abaixo.

A Figura 1 é um diagrama de rede que ilustra um exemplo da modalidade de configuração e gestão de comunicações entre nós de computação pertencendo a uma rede de computador virtual, de modo que as comunicações sejam sobrepostas em uma ou mais redes físicas intermediárias de forma transparente para os nós de computação. Neste exemplo, a configuração e gestão das comunicações é facilitada por um módulo gerenciador de sistema e múltiplos módulos gerenciadores de comunicação de um exemplo de modalidade do sistema ONM. O exemplo do sistema ONM pode ser usado, por exemplo, junto com um serviço de execução de programa publicamente acessível (não mostrado) e/ou serviço de rede configurável publicamente acessível (não mostrado), ou em vez disso pode ser usado em outras situações, como, por exemplo, com qualquer uso de redes de computador virtuais em nome de uma ou mais entidades (ex., para suportar múltiplas redes de computador virtuais para diferentes partes de um negócio ou outra organização em uma rede privada para a organização).

O exemplo ilustrativo inclui um exemplo de centro de dados 100 com múltiplos sistemas de computação físicos operados em nome do sistema ONM. O centro de dados de exemplo 100 é conectado a uma internet global 135 externa ao centro de dados 100, que fornece acesso a um ou mais sistemas de computação 145a através da rede privada 140, para um ou mais centros de dados globalmente acessíveis 160 cada um tendo múltiplos sistemas de computação (não mostrados), e a um ou mais outros sistemas de computação 145b. A internet global 135 pode ser, por exemplo, uma rede ou redes publicamente acessíveis (possivelmente operadas por várias partes distintas), como a Internet, e a rede privada 140 podem ser, por exemplo, uma rede corporativa que é completamente ou parcialmente inacessível a partir de sistemas de computação externos à rede privada 140. Os sistemas de computação 145b podem ser, por exemplo, sistemas de computação domésticos ou dispositivos de computação móveis, cada um se conectando diretamente à Internet (ex., através de uma linha telefônica, modem a cabo, Linha Digital de Assinante ("DSL"), rede celular ou outra conexão sem fio etc.).

O centro de dados de exemplo 100 inclui uma série de sistemas de computação fi-

sicos 105^a a 105d e 155a a 155n, bem como um módulo gerenciador de comunicação 150 que opera em um ou mais outros sistemas de computação (não mostrado) para gerenciar comunicações para os sistemas de computação relacionados 155a a 155n, e um módulo gerenciador de sistema 110 que opera em um ou mais sistemas de computação (não mos-
5 trado). Neste exemplo, cada sistema de computação físico hospeda os múltiplos nós de computação da máquina virtual 105a-105d e inclui um módulo gerenciador de comunicação da máquina virtual associada ("VM") {ex., como parte de um monitor hipervisor da máquina virtual para o sistema de computação físico), como um módulo gerenciador de comunicação VM 109a e máquinas virtuais 107a no sistema de comunicação do host 105a, e como um
10 módulo gerenciador de comunicação VM 109d e máquinas virtuais 107d no sistema de computação do host 105d. Os sistemas de computação físicos 155a a 155n não executam nenhuma máquina virtual neste exemplo, e assim podem, cada um, agir como um nó de computação que executa diretamente um ou mais programas de software em nome de um usuário. O módulo gerenciador de comunicação 150 que gerencia comunicações para os
15 sistemas de computação relacionados 155a a 155n podem ter várias formas, como, por exemplo, um dispositivo de computação de proxy, dispositivo de firewall, ou dispositivo de rede (ex., um comutador, roteador, hub, etc.) através dos quais percorrem as comunicações para e provenientes dos sistemas de computação físicos. Em outras modalidades, todos ou nenhum dos sistemas de computação no centro de dados podem hospedar máquinas virtu-
20 ais.

Este exemplo de centro de dados 100 ainda inclui múltiplos dispositivos de rede física, como comutadores 115a-115b, dispositivos roteadores de borda 125a- 125c, e dispositivos roteadores centrais 130a-130c. O comutador 115a é parte de uma sub-rede física que inclui sistemas de computação físicos 105a-105c, e é conectado ao roteador de borda 125a.
25 O comutador 115b é parte de uma sub-rede física distinta que inclui os sistemas de computação físicos 105d e 155a-155n, bem como os sistemas de computação que fornecem o módulo gerenciador de comunicação 150 e o módulo gerenciador de sistema 110, e é conectado ao roteador de borda 125b. As sub-redes físicas estabelecidas pelos comutadores 115a-115b, por sua vez, são conectadas uma a outra e a outras redes (ex., a internet global
30 135) através de uma rede de interconexão intermediária 120, que inclui os roteadores de borda 125a a 125c e os roteadores centrais 130a a 130c. Os roteadores de borda 125a-125c fornecem portas de ligação entre duas ou mais sub-redes ou redes. Por exemplo, o roteador de borda 125a fornece uma porta de ligação entre a sub-rede física estabelecida pelo comutador 115a e a rede de interconexão 120, enquanto o roteador de borda 125c for-
35 nece um gateway entre a rede de interconexão 120 e a internet global 135. Os roteadores centrais 130a-130c gerenciam comunicações na rede de interconexão 120, como por roteamento ou de outra forma encaminhando pacotes ou outras transmissões de dados confor-

me apropriado com base nas características de tais transmissões de dados (ex., informações de cabeçalho incluindo fonte e/ou endereços de destino, identificadores de protocolo etc.) e/ou outras características da própria rede de interconexão 120 (ex., rotas baseadas na topologia de rede física etc.).

5 O módulo gerenciador de sistema e os módulos gerenciadores de comunicação ilustrados podem realizar pelo menos algumas das técnicas descritas para configurar, autir-
zar e de outra forma gerenciar comunicações enviadas para e a partir dos nós de computa-
ção relacionados, inclusive para suportar o fornecimento de várias funcionalidades de rede
lógica para uma ou mais redes de computador virtual que são fornecidas usando vários dos
10 nós de computação. Por exemplo, o módulo gerenciador de comunicação 109a gerencia os
nós de computação da máquina virtual relacionados 107a, o módulo gerenciador de comu-
nicação 109d gerencia os nós de computação da máquina virtual relacionados 107d, e cada
um dos outros módulos gerenciadores de comunicação podem semelhantemente gerenciar
comunicações para um grupo de um ou mais outros de nós de computação relacionados.
15 Os módulos gerenciadores de comunicação ilustrados podem configurar comunicações en-
tre nós de computação para sobrepor uma rede virtual específica sobre uma ou mais redes
intermediárias que são usadas como uma rede de substrato, como através da rede de inter-
conexão 120. Além disso, uma rede virtual específica pode opcionalmente ser estendida
além do centro de dados 100 em algumas modalidades, como se um ou mais centros de
20 dados 160 também fornecessem nós de computação que são disponíveis para uso pelo sis-
tema ONM, e a rede virtual específica inclui nós de computação em dois ou mais centros de
dados em duas ou mais localizações geográficas distintas.

Tais múltiplos centros de dados ou outras localizações geográficas de um ou mais
nós de computação podem ser interconectados de várias formas, inclusive a seguinte: dire-
25 tamente através de uma ou mais redes públicas; através de uma conexão privada, não mos-
trada (ex., uma conexão física dedicada que não é compartilhada com terceiros, u, VPN ou
outro mecanismo que fornece a conexão privada em uma rede pública etc.); etc. Além disso,
embora não ilustrados no presente documento, tais outros centros de dados ou outras loca-
lizações geográficas podem, cada um, incluir um ou mais outros módulos gerenciadores de
30 comunicação que gerenciam comunicações para sistemas de computação naquele centro
de dados ou outra localização geográfica, bem como da internet global 135 ao centro de
dados 100 e quaisquer tais outros centros de dados 160.

Além disso, uma rede de computador virtual específica pode opcionalmente ser es-
tendida além do centro de dados 100 de outras formas em outras modalidades como se um
35 ou mais outros módulos de gerenciador de comunicação no centro de dados 100 fossem
colocados entre o roteador de borda 125c e a internet global 135, ou então baseados em um
ou mais outros módulos gerenciadores de comunicação externos ao centro de dados 100

(ex., se outro módulo gerenciador de comunicação for parte da rede privada 140, para gerenciar comunicações para os sistemas de computação 145a na internet global 135 e rede privada 140; etc.). Assim, por exemplo, se uma organização que opera a rede privada 140
5 desejar estender virtualmente sua rede de computador privada 140 para um ou mais dos nós de computação do centro de dados 100, poderá fazê-lo implementando um ou mais módulos gerenciadores de comunicação como parte da rede privada 140 (ex., como parte da interface entre a rede privada 140 e a internet global 135) - desta forma, os sistemas de computação 145a dentro da rede privada 140 podem se comunicar com aqueles nós de computação de centro de dados como se aqueles nós de computação de centros de dados
10 fossem parte da rede privada.

Desta forma, como um exemplo ilustrativo, um dos nós de computação da máquina virtual 107a no sistema de computação 105a (neste exemplo, nó de computação da máquina virtual 107a1) pode ser parte da mesma rede de computador local virtual como um dos
15 nós de computação da máquina virtual 107d no sistema de computação 105d (neste exemplo, o nó de computação da máquina virtual 107d1), como com o protocolo de rede IPv4 sendo usado para representar os endereços de rede virtual para a rede local virtual. A máquina virtual 107a1 pode em seguida direcionar uma comunicação de saída (não mostrada) ao nó de computação da máquina virtual de destino 107d1, especificando um endereço de rede virtual para tal nó de computação da máquina virtual de destino. O módulo gerenciador
20 de comunicação 109a recebe a comunicação de saída, e, em pelo menos algumas modalidades determina se autoriza o envio da comunicação de saída, como, se baseando em informações obtidas anteriormente sobre o nó de computação da máquina virtual de envio 107a1 e/ou sobre o nó de computação da máquina virtual de destino 107d1 (ex., informações sobre redes virtuais e/ou entidades às quais os nós de computação são relacionados),
25 e/ou por interação dinâmica com o módulo gerenciador de sistema 110 (ex., para obter uma determinação de autorização, para obter algumas ou todas tais informações, etc.). Não distribuir comunicações não-autorizadas a nós de computação, faz com que o isolamento e segurança de rede de redes de computador virtual de entidades aumente.

Se o módulo gerenciador de comunicação 109a determinar que a comunicação de
30 saída é autorizada (ou não realizar tal determinação de autorização), o módulo 109a determinará a localização da rede física real correspondente ao endereço de rede virtual de destino para a comunicação. Por exemplo, o módulo gerenciador de comunicação 109a pode determinar o endereço de rede de destino real para uso para o endereço de rede virtual da máquina virtual de destino 107d1 interagindo dinamicamente com o módulo gerenciador de
35 sistema 110, ou pode ter determinado anteriormente e armazenado tais informações (ex., em resposta a uma solicitação da máquina virtual de envio 107a1 para informações sobre o endereço de rede virtual de destino, como uma solicitação de que a máquina virtual 107a1

especifica usando o Protocolo de Resolução de Endereço, ou ARP). O módulo gerenciador de comunicação 109a em seguida faz o cabeçalho novamente ou de outra forma modifica a comunicação de saída de modo que seja direcionada para o módulo gerenciador de comunicação 109d usando um endereço de rede de substrato real, como se o módulo gerenciador de comunicação 109d fosse associado a uma variedade de múltiplos endereços de rede de substrato reais. As figuras 2A-2D fornecem exemplos de como fazer tal gestão de comunicação em algumas modalidades, inclusive para emular funcionalidade de rede lógica especificada para a rede virtual.

Quando o módulo gerenciador de comunicação 109d recebe a comunicação através da rede de interconexão 120 neste exemplo, obtém o endereço de rede de destino virtual para a comunicação (ex., extraíndo o endereço de rede de destino virtual da comunicação), e determina a qual dos nós de computação da máquina virtual 107d gerenciados pelo módulo gerenciador de comunicação 109d a comunicação é direcionada. O módulo gerenciador de comunicação 109d em seguida determina se a comunicação é autorizada para o nó de computação da máquina virtual de destino 107d1, com exemplos de tais atividades de autorização discutidos mais detalhadamente nos exemplos das Figuras 2A-2D. Se a comunicação for determinada a ser autorizada (ou o módulo gerenciador de comunicação 109d não realizar tal determinação de autorização), o módulo gerenciador de comunicação 109d em seguida faz o cabeçalho novamente ou de forma diversa modifica a comunicação de chegada de modo que seja direcionada ao nó de computação da máquina virtual de destino 107d1 usando um endereço de rede virtual apropriado para a rede de computador virtual, como, através do uso endereço de rede virtual 107a1's do nó de computação da máquina virtual de envio como o endereço de rede da fonte e usando o endereço de rede virtual 107d1's do nó de computação da máquina virtual de destino como o endereço de rede de destino. O módulo gerenciador de comunicação 109d em seguida encaminha a comunicação modificada ao nó de computação da máquina virtual de destino 107d1. Pelo menos em algumas modalidades, antes de encaminhar a comunicação de chegada à máquina virtual de destino, o módulo gerenciador de comunicação 109d pode ainda realizar etapas adicionais relacionadas à segurança, conforme discutido mais detalhadamente em outras partes do presente documento.

Além disso, embora não ilustrado na Figura 1, em algumas modalidades, os vários módulos gerenciadores de comunicação podem tomar outras ações para fornecer funcionalidade de rede lógica correspondente a uma topologia de rede especificada para a rede de computador virtual, por exemplo, gerenciando comunicações entre nós de computação da rede de computador virtual de formas especificadas e respondendo a outros tipos de solicitações enviadas pelos nós de computação da rede do computador virtual. Por exemplo, embora seja separado do nó de computação 107a1 no sistema de computação físico 105a pela

rede de interconexão 120 no exemplo de modalidade da Figura 1, o nó de computação da máquina virtual 107d1 no sistema de computação físico 105d pode ser configurado para ser parte da mesma sub-rede lógica da rede de como o nó de computação 107a1 (ex., para não ser separado por quaisquer dispositivos roteadores lógicos especificados). Em contrapartida, apesar da proximidade física do nó de computação da máquina virtual 107c1 no sistema de computação físico 105c ao nó de computação da máquina virtual 107a1 no sistema de computação físico 105a (i.e., sendo parte da mesma sub-rede física sem quaisquer dispositivos roteadores físicos intervenientes) no exemplo de modalidade da Figura 1, o nó de computação 107c1 pode ser configurado para ser parte de uma sub-rede lógica distinta da rede de computador virtual daquela do nó de computação 107a1 (ex., pode ser configurado para ser separado por um ou mais dispositivos roteadores lógicos especificados, não mostrado). Sendo assim, o exemplo anterior de envio de uma comunicação a partir do nó de computação 107a1 ao nó de computação 107d1 pode ser realizado da forma descrita anteriormente, sem emular as ações de quaisquer dispositivos roteadores lógicos intervenientes (apesar do uso de vários dispositivos roteadores físicos na rede de interconexão de substrato 120 para encaminhar a comunicação), uma vez que os nós de computação 107a1 e 107d1 são configurados para ser parte da sub-rede única na topologia de rede especificada.

No entanto, se o nó de computação 107a1 enviar uma comunicação adicional ao nó de computação 107c1, os módulos gerenciadores de comunicação 109a e/ou 109c nos sistemas de computação do host 105a e 105c poderão realizar ações adicionais que correspondem a um ou mais dispositivos roteadores lógicos especificados configurados na topologia de rede especificada para separar os nós de computação 107a1 e 107c1. Por exemplo, o nó de computação da fonte 107a1 pode enviar a comunicação adicional de modo a direcioná-la inicialmente a um primeiro dos dispositivos roteadores lógicos especificados que é configurado para ser local ao nó de computação 107a1 (ex., incluindo um endereço de hardware virtual no cabeçalho da comunicação adicional que corresponde àquele primeiro dispositivo roteador lógico especificado), sendo esperado que o primeiro dispositivo roteador lógico especificado envie a comunicação adicional para o nó de computação de destino 107c1 através da topologia de rede lógica especificada. Sendo assim, o módulo gerenciador de comunicação 109a pode detectar que o envio da comunicação adicional ao primeiro dispositivo roteador lógico (ex., baseado no endereço de hardware virtual usado no cabeçalho da comunicação adicional), ou de outra forma ficar ciente da topologia de rede configurada para a rede de computador virtual, e pode tomar ações para emular funcionalidade de alguns ou todos os dispositivos roteadores lógicos especificados que são configurados na topologia de rede especificada para separar os nós de computação 107a1 e 107c1. Por exemplo, pode-se esperar que cada dispositivo roteador lógico que encaminha a comunicação adicional tome ações como modificar um valor de salto de TTL ("tempo de vida") para a

comunicação, modificar um endereço de hardware de destino virtual que é especificado para a comunicação para indicar o próximo destino pretendido da comunicação adicional em uma rota ao nó de computação de destino, e/ou de outra forma modificar o cabeçalho de comunicação. Sendo assim, o módulo gerenciador de comunicação da fonte 109a pode realizar

5 algumas ou todas aquelas ações antes de encaminhar a comunicação adicional ao módulo gerenciador de comunicação de destino 109c através da rede de substrato (neste caso, através do dispositivo de comutador físico 115a) para fornecimento ao nó de computação de destino 107c1. Como alternativa, algumas ou todas das ações adicionais para fornecer a funcionalidade de rede lógica para a comunicação adicional enviada podem, em vez disso,

10 ser realizadas pelo módulo gerenciador de comunicação de destino 109c após a comunicação adicional ser encaminhada ao módulo gerenciador de comunicação 109c pelo módulo gerenciador de comunicação 109a. O exemplo da Figura 2C fornece detalhes adicionais a respeito de exemplos de fornecimento de funcionalidade de rede lógica.

Fornecendo funcionalidade de rede lógica usando as técnicas descritas, o sistema

15 ONM fornece vários benefícios. Por exemplo, visto que os vários módulos gerenciadores de comunicação gerenciam a rede virtual de sobreposição e podem emular funcionalidade de dispositivos de rede lógica, os dispositivos de rede especificados e outra topologia de rede não precisam ser implementados fisicamente para redes de computador virtual sendo fornecidas, e assim as modificações correspondentes não são necessárias para a rede de inter-

20 conexão 120 ou comutadores 115a-115b para suportar topologias de rede configuradas específicas. Não obstante, se os nós de computação e programas de software de uma rede de computador virtual tiverem sido configurados para esperar uma topologia de rede específica para a rede de computador virtual, o aspecto daquela topologia de rede poderá, mesmo assim, ser fornecido transparentemente para aqueles nós de computação pelas técnicas des-

25 critas.

As Figuras 2A-2C ilustram outros exemplos com detalhes ilustrativos adicionais relacionados ao gerenciamento de comunicações entre nós de computação que ocorrem através de uma rede de sobreposição em uma ou mais redes físicas, podendo ser usadas pelos nós de computação e redes da Figura 1 ou em outras situações. Especificamente, a Figura

30 2A ilustra vários exemplos de nós de computação 205 e 255 que podem se comunicar usando uma ou mais redes de interconexão intermediárias 250 como uma rede de substrato. Neste exemplo, a rede de interconexão 250 é uma rede de substrato IPv6 em que as redes de computador virtual IPv4 são sobrepostas, embora em outras modalidades a rede de interconexão 250 e as redes de sobreposição de computador virtual possam usar o

35 mesmo protocolo de rede (ex., IPv4). Além disso, neste exemplo de modalidade, os nós de computação são operados em nome de diversas entidades distintas, e um módulo gerenciador de sistema 290 gerencia a associação de nós de computação específicos a entidades

específicos e redes de computador virtual, e rastreia diversas informações de configuração especificadas para as redes de computador virtual. Os nós de computação de exemplo da Figura 2A incluem quatro nós de computação executados em nome de uma entidade de exemplo Z e parte de uma rede de computador virtual correspondente fornecida para a entidade Z, aqueles sendo os nós de computação 205a, 205c, 255a e 255b. Além disso, outros nós de computação são operados em nome de outras entidades e pertencem a outras redes de computador virtual fornecidas, tal como o nó de computação 205b e outros nós de computação 255.

Neste exemplo, os nós de computação 205 são gerenciados por e fisicamente conectados a um módulo gerenciador de comunicação relacionado R 210, os nós de computação 255 são gerenciados por e fisicamente conectados a um módulo gerenciador de comunicação S 260, e os módulos gerenciadores de comunicação ONM 210 e 260 são fisicamente conectados a uma rede de interconexão 250, como é o módulo Gerenciador de Sistemas 290, embora as interconexões físicas entre os nós de computação, módulos e a rede de interconexão não sejam ilustradas neste exemplo. Como um exemplo, os nós de computação 205 podem cada um ser de múltiplas máquinas virtuais hospedadas por um único sistema de comunicação físico, e o módulo gerenciador de comunicação R pode ser parte de um monitor de máquina virtual hipervisor para tal sistema de computação físico. Por exemplo, com referência à Figura 1, os nós de computação 205 podem representar as máquinas virtuais 107a, e os nós de computação 255 podem representar as máquinas virtuais 107d. Sendo assim, o módulo gerenciador de comunicação R pode corresponder ao módulo gerenciador de comunicação 109a da Figura 1, o módulo gerenciador de comunicação S corresponderia ao módulo gerenciador de comunicação 109d da Figura 1, a rede de interconexão 250 corresponderia à rede de interconexão 120 da Figura 1, e o módulo Gerenciador de Sistemas 290 corresponderia ao módulo Gerenciador de Sistemas 110 da Figura 1. Como alternativa, os nós de computação 205 ou 255 podem, em vez disto, ser um sistema de computação físico distinto, tal como para corresponder aos sistemas de computação 155a-155n da Figura 1, ou aos nós de computação em outros centros de dados ou localizações geográficas (ex., sistemas de computação em outro centro de dados 160, sistemas de computação 145a, etc.).

Cada módulo gerenciador de comunicação da Figura 2A é associado a um grupo de múltiplos endereços de rede de substrato física, o qual os módulos gerenciadores de comunicação gerenciam em nome de seus nós de computação relacionados. Por exemplo, o módulo gerenciador de comunicação R é mostrado como sendo associado à faixa de endereço de rede de IPv6 de "::0A:01/72", que corresponde aos endereços de 128 bits (no sistema hexadecimal) de XXXX:XXXX:XXXX:XXxA:0100:0000:0000:0000 até XXXX:XXXX:XXXX:XXxA: 01 FF:FFFF:FFFF:FFFF (representando 2 à energia de 56 ende-

reços de IPv6 únicos), onde cada "X" pode representar qualquer caracter hexadecimal que é apropriado para uma situação específica (ex., com os 64 bits iniciais correspondendo a uma organização e topologia de rede específicas, conforme discutido mais detalhadamente com relação à Figura 2D). A rede de interconexão 250 encaminhará qualquer comunicação com um endereço de rede de destino naquela faixa ao módulo gerenciador de comunicação R - assim, com os 72 bits iniciais da faixa especificada, o módulo gerenciador de comunicação R pode usar os 56 bits restantes disponíveis para representar os nós de computação que gerencia e para determinar como processar comunicações de chegada cujos endereços de rede de destino estão naquela faixa. Communication Manager module R may use the remaining available 56 bits to represent the computing nodes that it manages and to determine how to process incoming communications whose destination network addresses are in that range.

Para fins do exemplo mostrado na Figura 2A, os nós de computação 205a, 205c, 255a e 255b são parte de uma rede de computador virtual única para a entidade Z, e têm endereços de rede virtual IPv4 atribuídos de "10.0.0.2", "10.0.5.1", "10.0.0.3", e "10.1.5.3", respectivamente. Visto que o nó de computação 205b é parte de uma rede de computador virtual distinta para a entidade Y, o mesmo pode compartilhar os mesmos endereços de rede virtual que o nó de computação 205a sem confusão. Neste exemplo, o nó de computação A 205a pretende se comunicar com o nó de computação G 255a, que é configurado neste exemplo para ser parte de uma única sub-rede física de local comum (não mostrada) em uma topologia de rede para a rede de computador virtual, e a rede de interconexão 250 e os módulos gerenciadores de comunicação são transparentes para os nós de computação A e G neste exemplo. Especificamente, apesar da separação física dos nós de computação A e G, os módulos gerenciadores de comunicação 210 e 260 operam para sobrepor a rede de computador virtual para a entidade Z através da rede de interconexão física 250 para comunicações entre aqueles nós de computação, de modo que a falta de uma rede local real seja evidente para os nós de computação A e G.

Para enviar a comunicação do nó de computação G, o nó de computação A troca várias mensagens 220 com o módulo gerenciador de comunicação R 210, apesar de a modalidade ilustrada desconhecer a existência do módulo gerenciador de comunicação R (i.e., o nó de computação A pode acreditar que está transmitindo uma mensagem de transmissão a todos os outros nodos em uma sub-rede local, como, por exemplo, através de um dispositivo de troca especificado que o nó de computação A acredita que conecte os nodos na sub-rede local). Especificamente, neste exemplo, o nó de computação A primeiro envia uma solicitação de mensagem de ARP 220-a que inclui o endereço de rede virtual para o nó de computação G (i.e., "10.0.0.3") e que solicita o endereço de hardware correspondente para o nó de computação G (ex., um endereço MAC de 48 bits). O módulo gerenciador de comu-

nicação R intercepta a solicitação de ARP 220-a, e responde ao nó de computação A com uma mensagem de resposta de ARP forjada 220-b que inclui um endereço de hardware virtual para o nó de computação G.

Para obter o endereço de hardware virtual para o nó de computação G para uso com a mensagem de resposta, o módulo gerenciador de comunicação R primeiro verifica uma armazenagem local 212 de informações que mapeia endereços de hardware virtual para os endereços de redes de substrato físicas reais IPv6 correspondentes, com cada um dos endereços de hardware virtual também correspondendo a um endereço de rede virtual IPv4 para uma rede virtual de entidade específica. Se a armazenagem local 212 não con-

10 ver uma entrada para o nó de computação G (ex., se nenhum dos nós de computação 205 tiver se comunicado anteriormente com o nó de computação G, se uma entrada anterior na armazenagem local 212 para o nó de computação G tiver expirado com base em um tempo de expiração relacionado etc.), o módulo gerenciador de comunicação R interage 225 com o módulo Gerenciador de Sistemas 290 para obter o endereço de rede de substrato física real

15 IPv6 para o nó de computação G em nome do nó de computação A. Especificamente, neste exemplo, o módulo Gerenciador de Sistemas 290 mantém as informações de fornecimento 292 que identificam onde cada nó de computação está de fato localizado e a qual entidade e/ou rede de computador virtual o nó de computação pertence, tal como iniciando a execu-

20 de outra forma obtendo tais informações de fornecimento. Conforme discutido mais detalhadamente com relação à Figura 2B, o módulo gerenciador de sistema determina se a solicitação do módulo gerenciador de comunicação R em nome do nó de computação A para o endereço de rede de substrato física IPv6 real do nó de computação G é válido, inclusive se o nó de computação A é autorizado para se comunicar com o nó de computação G, e sendo

25 assim fornece aquele endereço de rede de substrato física IPv6 real.

O módulo gerenciador de comunicação R recebe o endereço de rede de substrato física IPv6 real para o nó de computação G a partir do módulo Gerenciador de Sistemas 290, e armazena estas informações recebidas como parte de uma nova entrada para o nó de computação G como parte das informações de mapeamento 212 para uso posterior (opcio-

30 nalmente com um tempo de expiração e/ou outras informações). Além disso, neste exemplo, o módulo gerenciador de comunicação R deternubar um endereço de hardware virtual falso para ser usado para o nó de computação G (ex., gerando um identificador que é localmente único para os nós de computação gerenciados pelo módulo gerenciador de comunicação R), armazena aquele endereço de hardware virtual falso junto com o endereço de rede de sub-

35 trato física IPv6 real como parte da entrada das novas informações de mapeamento, e fornece o endereço de hardware virtual falso ao nó de computação A como parte da mensagem de resposta 220-b. Mantendotais informações de mapeamento 212, as comunicações

posteriores do nó de computação A ao nó de computação G podem ser autorizadas pelo módulo gerenciador de comunicação R sem outras interações com o módulo gerenciador de sistema 290, com base no uso do endereço de hardware virtual falso fornecido anteriormente pelo módulo gerenciador de comunicação R. Em outras modalidades, o endereço de hardware usado pelo módulo gerenciador de comunicação R para o nó de computação G pode, então, ser um endereço falso, como se o módulo Gerenciador de Sistemas 290 ainda mantivesse informações sobre endereços de hardware usados pelos vários nós de computação (ex., endereços de hardware virtual designados para nós de computação de máquinas virtuais, endereços de hardware reais designados para sistemas de computação que agem como nós de computação, etc.) e fornece o endereço de hardware usado pelo nó de computação G ao módulo gerenciador de comunicação R como parte das interações 225. Em tais modalidades, o módulo gerenciador de comunicação R pode tomar outras ações se os nós de computação em redes virtuais diferentes usarem o mesmo endereço de hardware virtual, como para mapear cada combinação de endereço de hardware de nó de computação e rede de computador virtual a um endereço de rede de substrato correspondente.

Em outras modalidades, o módulo gerenciador de comunicação R pode interagir com o módulo Gerenciador de Sistemas 290 para obter um endereço de rede de substrato física para o nó de computação G ou de outra forma determinar tal endereço de rede de substrato física em momentos exceto mediante o recebimento de uma solicitação de ARP, como em resposta a qualquer comunicação recebida que é direcionada ao nó de computação G usando o endereço de rede virtual "10.0.0.3" como parte da rede de computador virtual da entidade Z. Além disso, em outras modalidades, os endereços de hardware virtual que são usados podem diferir deste exemplo, como se os endereços de hardware virtual fossem especificados pelo módulo Gerenciador de Sistemas 290, se os endereços de hardware virtual não forem aleatórios e em vez disso armazenarem um ou mais tipos de informações específicas para os nós de computação correspondentes etc. Além disso, neste exemplo, se o nó de computação A não tivesse sido determinado para ser autorizado a enviar comunicações ao nó de computação G₁ pelo módulo gerenciador de sistema 290 e/ou módulo gerenciador de comunicação R, o módulo gerenciador de comunicação R não enviaria a mensagem de resposta 220-b com o endereço de hardware virtual (ex., em vez disso não envia nenhuma resposta ou uma resposta de mensagem de erro).

Neste exemplo, o endereço de rede de substrato física IPv6 real correspondente ao nó de computação G nas interações 225 é "::0B:02:<Z-identifier>:10.0.0.3", onde "10.0.0.3" é armazenado nos últimos 32 bits do endereço IPv6 de 128 bits, e onde "<Z-identifier>" é um identificador de rede de entidade de 24 bits para o nó de computação G correspondente à rede de computador virtual para a entidade Z (ex., conforme anteriormente designado pelo módulo Gerenciador de Sistemas àquela rede para refletir um número aleatório ou algum

outro número correspondente à entidade). Os 72 bits iniciais do endereço de rede IPv6 armazenam a designação "::0B:02", correspondente à sub-rede ou outra porção da rede de interconexão com uma faixa de endereço de rede de "::0B:02/72" à qual o módulo gerenciador de comunicação S corresponde - assim, uma comunicação enviada através da rede de interconexão 250 ao endereço de rede de destino IPv6 "::0B:02:<Z-identifier>:10.0.0.3" será roteada para o módulo gerenciador de comunicação S. Em outras modalidades, o identificador de rede de entidade pode ter outros comprimentos (ex., 32 bits, se o módulo gerenciador de comunicação S tiver uma faixa de endereço de rede relacionada de 64 bits em vez de 56 bits) e/ou pode ter outras formas (ex., pode ser aleatório, pode armazenar vários tipos de informações etc), e os 56 bits restantes usados para a faixa de endereço de rede após a designação "::0B:02" podem armazenar outros tipos de informações (ex., um identificador para uma entidade específica, uma etiqueta ou rótulo para a rede virtual, etc.). Detalhes adicionais relacionados a um exemplo de endereço de rede física real IPv6 para uso com uma rede de computador virtual de sobreposição são descritos com relação à Figura 2D.

Após receber a mensagem de resposta 220-b do módulo gerenciador de comunicação R, o nó de computação A cria e inicia o envio de uma comunicação ao nó de computação G, mostrado na Figura 2A como a comunicação 220-c. Especificamente, o cabeçalho de comunicação 220-c inclui um endereço de rede de destino para nó de computação de destino G que é "10.0.0.3", um endereço de hardware de destino para o nó de computação de destino G que é o endereço de hardware virtual fornecido para o nó de computação A na mensagem 220-b, um endereço de rede da fonte para enviar o nó de computação A que é "10.0.0.2", e um endereço de hardware da fonte para enviar o nó de computação A que é um endereço de hardware real ou falso que foi anteriormente identificado para o nó de computação A. Uma vez que o nó de computação A acredita que o nó de computação G é parte da mesma sub-rede local como o próprio, o nó de computação A não precisa direcionar a comunicação 220-c para quaisquer dispositivos roteadores lógicos intermediários que são configurados na topologia de rede para separar os nós de computação.

O módulo gerenciador de comunicação R intercepta a comunicação 220-c, modifica a comunicação conforme apropriado, e encaminha a comunicação modificada através da rede de interconexão 250 para o nó de computação G. Especificamente, o módulo gerenciador de comunicação R extrai o endereço de rede de destino virtual e o endereço de hardware de destino virtual para o nó de computação G a partir do cabeçalho, e em seguida recupera o endereço de rede de substrato física real IPv6 correspondente àquele endereço de hardware de destino virtual a partir das informações de mapeamento 212. Conforme observado anteriormente, o endereço de rede de substrato físico real IPv6 neste exemplo é "::0B:02:<Z-identifier>:10.0.0.3", e o módulo gerenciador de comunicação R cria um novo cabeçalho IPv6 que inclui aquele endereço de rede de substrato física real como o endereço

de destino. Semelhantemente, o módulo gerenciador de comunicação R extrai o endereço de rede de fonte virtual e o endereço de hardware de fonte virtual para o nó de computação A a partir do cabeçalho da comunicação recebida, obtém um endereço de rede de substrato física correspondente àquele endereço de hardware de fonte virtual (ex., a partir de uma entrada armazenada nas informações de mapeamento 212, interagindo com o módulo gerenciador de sistema 290 para obter tais informações se não obtidas anteriormente etc.) e inclui aquele endereço de rede de substrato física real como o endereço de rede da fonte para o novo cabeçalho IPv6. Neste exemplo, o endereço de rede de substrato física real IPv6 para o nó de computação A é "::0A:01 :<Z-identifier>:10.0.0.2", que, se usado em uma resposta pelo módulo gerenciador de comunicação S em nome do nó de computação G será roteado para o módulo gerenciador de comunicação R para encaminhamento ao nó de computação A. O módulo gerenciador de comunicação R em seguida cria a comunicação 230-3 modificando a comunicação 220-c para substituir o cabeçalho IPv4 anterior pelo novo cabeçalho IPv6 (ex., em conformidade com SIIT), incluindo popular o novo cabeçalho IPv6 com outras informações conforme apropriado para a comunicação (ex., comprimento de carga útil, prioridade de pacote de classe de tráfego etc.)- Assim, a comunicação 230-3 inclui o mesmo conteúdo ou carga útil do comunicador 220-c, sem encapsular a comunicação 220-c dentro da comunicação 230-3. Além disso, o acesso às informações específicas dentro da carga útil não é necessários para tal novo cabeçalho, como para permitir ao módulo gerenciador de comunicação R lide com as comunicações em que a carga útil é criptografada sem necessidade de descriptografar tal carga útil.

Pelo menos em algumas modalidades, antes de encaminhar a comunicação 230-3 do módulo gerenciador de comunicação S, o módulo gerenciador de comunicação R pode realizar uma ou mais ações para determinar se a comunicação 220-c é autorizada a ser encaminhada para o nó de computação G como a comunicação 230-3, tal como com base nas informações de mapeamento 212 incluindo uma entrada válida para o endereço de hardware virtual de destino usado na comunicação 220-c (ex., uma entrada específica para enviar o nó de computação 205a em algumas modalidades, ou então uma entrada correspondente a quaisquer nós de computação 205 em outras modalidades). Em outras modalidades, tal determinação pode não ser realizada pelo módulo gerenciador de comunicação R para cada comunicação de saída, ou então pode ser realizada de outras formas (ex., com base em uma determinação de que o nó de envio e o nó de destino são parte da mesma rede de computador virtual ou são relacionados à mesma entidade ou são de outra forma autorizados a intercomunicar, com base em uma interação com o módulo Gerenciador de Sistemas 290 para obter uma determinação de autorização para a comunicação, etc.).

Após o módulo gerenciador de comunicação R encaminhar a comunicação modificada 230-3 à rede de interconexão 250, a rede de interconexão usa o endereço de rede de

destino IPv6 física da comunicação para rotear a comunicação ao módulo gerenciador de comunicação S. Fazendo isso, os dispositivos da rede de interconexão 250 não usam a porção do endereço de rede de destino que inclui o identificador de rede de entidade incorporado ou o endereço de rede virtual incorporado, e assim não necessitam de qualquer configuração especial para encaminhar tal comunicação, nem mesmo ciência de que uma rede de computador virtual está sendo sobreposta na rede de interconexão física.

Quando o módulo gerenciador de comunicação S recebe a comunicação 230-3 através da rede de interconexão 250, realiza ações semelhantes às aquelas do módulo gerenciador de comunicação R, mas ao contrário. Especificamente, pelo menos em algumas modalidades, o módulo gerenciador de comunicação S verifica se a comunicação 230-3 é legítima e autorizada para ser encaminhada ao nó de computação G, tal como via uma ou mais interações 240 com o módulo Gerenciador de Sistemas. Se a comunicação for determinada a ser autorizada (ou se a determinação de autorização não for realizada), o módulo gerenciador de comunicação S em seguida modifica a comunicação 230-3 conforme apropriado e encaminha a comunicação modificada para o nó de computação G. Detalhes adicionais relacionados à verificação da comunicação 230-3 são discutidos com relação à Figura 2B.

Especificamente, para modificar a comunicação 230-3, o módulo gerenciador de comunicação S recupera as informações das informações de mapeamento 262 que correspondem ao nó de computação G, incluindo o endereço de hardware virtual usado pelo nó de computação G (ou gera tal endereço de hardware virtual se não disponível anteriormente, tal como para um novo modo de computação). O módulo gerenciador de comunicação S em seguida cria a comunicação 245-e modificando a comunicação 230-3 para substituir o cabeçalho IPv6 anterior por um cabeçalho IPv4 novo (ex., em conformidade com SIIT). O novo cabeçalho IPv4 inclui o endereço de rede virtual e o endereço de hardware virtual para o nó de computação G como o endereço de rede de destino e o endereço de hardware de destino para o novo cabeçalho IPv4, o endereço de rede virtual e um endereço de hardware virtual para o nó de computação A como o endereço de rede da fonte e o endereço de hardware da fonte para o novo cabeçalho IPv4, e inclui outras informações conforme apropriado para a comunicação (ex., comprimento total, soma de verificação de cabeçalho etc.). O endereço de hardware virtual usado pelo módulo gerenciador de comunicação S para o nó de computação A pode ser o mesmo que o endereço de hardware usado pelo módulo de Gerenciador de Comunicação R para o nó de computação A, mas em outras modalidades cada módulo gerenciador de comunicação pode manter informações de endereço de hardware separado que não são relacionadas às informações usadas pelos outros módulos gerenciadores de comunicação (ex., se o módulo gerenciador de comunicação S gerou seu próprio endereço de hardware virtual falso para o nó de computação A em resposta a uma solicitação de ARP anterior de um dos nós de computação 255 para o endereço de hardware do nó de compu-

tação A). Assim, a comunicação 245-e inclui o mesmo conteúdo ou carga útil das comunicações 220-c e 230-3. O módulo gerenciador de comunicação S em seguida encaminha a comunicação 245-e para o nó de computação G.

Após receber a comunicação 245-e, o nó de computação G determina o envio de uma comunicação de resposta 245-f ao nó de computação A, usando o endereço de rede virtual da fonte e o endereço de hardware virtual da fonte para o nó de computação A a partir da comunicação 245-e. O módulo gerenciador de comunicação S recebe a comunicação de resposta 245-f, e a processa de forma semelhante àquela anteriormente descrita com relação à comunicação 220-c e módulo gerenciador de comunicação R. Especificamente, o módulo gerenciador de comunicação S opcionalmente verifica se o nó de computação G está autorizado a enviar comunicações para o nó de computação A, e em seguida modifica a comunicação 245-f para criar a comunicação 230-6 gerando um novo cabeçalho IPv6 usando as informações de mapeamento 262. Após encaminhar a comunicação 230-6 à rede de interconexão 250, a comunicação é enviada ao módulo gerenciador de comunicação R, que processa a comunicação de chegada de forma semelhante àquela descrita anteriormente com relação à comunicação 230-3 e o módulo gerenciador de comunicação S. Especificamente, o módulo gerenciador de comunicação R opcionalmente verifica se o nó de computação G está autorizado a enviar comunicações para o nó de computação A e se a comunicação 230-6 de fato foi enviada a partir da localização da rede de substrato do nó de computação G, e em seguida modifica a comunicação 230-6 para criar a comunicação de resposta 220-d gerando um novo cabeçalho IPv4 usando as informações de mapeamento 212. O módulo gerenciador de comunicação R em seguida encaminha a comunicação de resposta 220-d ao nó de computação A. Em outras modalidades e situações, os módulos gerenciadores de comunicação R e/ou S podem manusear comunicações de respostas de forma diferente das comunicações iniciais, tal como para assumir que as comunicações de resposta são autorizadas em pelo menos algumas situações, e para não realizar algumas ou todas as atividades de autorização para comunicações de resposta naquelas situações.

Desta forma, os nós de computação A e G podem intercomunicar usando uma rede de computador virtual com base IPv4, sem qualquer configuração especial daqueles nós de computação para manusear a rede de interconexão de substrato com base IPv6 interveniente real, e a rede de interconexão 250 pode encaminhar as comunicações IPv6 sem qualquer configuração especial de quaisquer dispositivos de rede física da rede de interconexão, com base nos módulos gerenciadores de comunicação sobrepostos à rede de computador virtual através da rede de interconexão física real sem encapsulação de comunicações e usando endereços de rede virtual incorporada nos endereços de rede física de substrato.

Além disso, embora não ilustrado com relação à Figura 2A, pelo menos em algumas modalidades, os módulos gerenciadores de comunicação podem receber e manusear

outros tipos de solicitações e comunicações em nome de nós de computação relacionados. Por exemplo, os módulos gerenciadores de comunicação podem tomar várias ações para suportar capacidades de transmissão e multicast para os nós de computação que eles gerenciam. Como um exemplo, em algumas modalidades, um sufixo de endereço de rede virtual de grupo de multicast pode ser reservado a partir de cada prefixo identificador de rede de entidade para uso na sinalização de comunicações encapsuladas brutas da Camada 2 de redes. Semelhantemente, para comunicações de transmissão e multicast de ligação local, um prefixo /64 do grupo multicast especial pode ser reservado (ex., "FF36:0000::"), enquanto um prefixo de endereço de destino diferente (ex., "FF15:0000::") pode ser usado para outras comunicações de multicast. Desta forma, por exemplo, estruturas de IP de multicast e transmissão podem ser encapsuladas usando um prefixo de 64 bits reservado correspondente para os primeiros 64 bits do endereço IPv6 de 128 bits, com os 64 bits restantes incluindo o endereço de rede IPv4 virtual para o nó de computação de destino de forma semelhante àquela descrita anteriormente. Como alternativa, em outras modalidades, um ou mais tipos de comunicações de transmissão e/ou multicast podem, cada um, ter um rótulo reservado correspondente ou outro identificador que tem um valor ou forma diferente, incluindo usar um número de bits diferente e/ou ser armazenado de forma exceto como um prefixo de endereço de rede. Quando um nó de computação envia uma comunicação de transmissão/multicast, qualquer módulo gerenciador de comunicação com um nó de computação relacionado que foi subscrito àquele grupo de multicast/transmissão seria identificado (ex., com base naqueles módulos gerenciadores de comunicação tendo sido subscritos ao grupo, tal como em resposta a comunicações conjuntas anteriores enviadas por aqueles nodos relacionados), e o módulo gerenciador de comunicação para o nó de computação de envio encaminharia a comunicação a cada um dos módulos gerenciadores de comunicação identificados do grupo, para envio aos seus nós de computação gerenciados apropriados. Além disso, em algumas modalidades e situações, pelo menos algumas comunicações de transmissão ou multicast podem não ser enviadas pelos módulos Gerenciadores de Comunicação, tal como comunicações com um prefixo IPv4 de 224.0/16 ou outro prefixo designado ou outro rótulo ou identificador.

Além de suportar capacidades de transmissão e multicast para nós de computação gerenciados, os módulos gerenciadores de comunicação podem receber e manusear outros tipos de solicitações e comunicações em nome de nós de computação relacionados que correspondem a topologias de rede configuradas para as redes de computador virtual às quais os nós de computação pertencem. Por exemplo, os nós de computação podem enviar várias solicitações que um dispositivo roteador local especificado ou outro dispositivo de rede especificado esperaria manusear (ex., solicitações de ping requests, questionários de SNMP, etc.), e os módulos gerenciadores de comunicação relacionados podem interceptar tais soli-

citações e tomar várias ações correspondentes para emular a funcionalidade que teria sido fornecida pelo dispositivo de rede especificado se este fosse fisicamente implementado.

Além disso, será apreciado que um módulo gerenciador de comunicação pode facilitar comunicações entre vários dos nós de computação que são associados àquele módulo Gerenciador de Comunicação. Por exemplo, com relação à Figura 2A, o nó de computação 205a pode pretender enviar uma comunicação adicional (não mostrada) ao nó de computação 205c. Caso assim, o módulo gerenciador de comunicação R realizaria ações semelhantes àqueles descritas anteriormente com relação ao tratamento da comunicação de saída 220-c pelo módulo gerenciador de comunicação R e o tratamento da comunicação de chegada 230-3 pelo módulo gerenciador de comunicação S, mas sem novo cabeçalho da comunicação adicional para usar um cabeçalho de IPv6 uma vez que a comunicação não percorrerá a rede de interconexão. No entanto, se os nós de computação 205a e 205c forem configurados em uma topologia de rede para a rede de computador virtual ser separada por um ou mais dispositivos de rede lógica, o módulo gerenciador de comunicação R pode tomar ações adicionais para emular a funcionalidade daqueles dispositivos de rede lógica, conforme discutido mais detalhadamente com relação à Figura 2C.

Embora não ilustrados com relação à Figura 2A, em pelo menos algumas modalidades, outros tipos de solicitações e comunicações também podem ser manuseados de várias formas. Por exemplo, em pelo menos algumas modalidades, uma entidade pode ter um ou mais nós de computação que são gerenciados pelos módulos gerenciadores de comunicação e que são parte de uma rede de computador virtual para aquela entidade, e podem ainda ter um ou mais outros sistemas de computação não-gerenciados (ex., sistemas de computação que são diretamente conectados à rede de interconexão 250 e/ou aquele endereço de rede IPv6 de uso nativo) que não tem um módulo gerenciador de comunicação relacionado que gerencia suas comunicações. Se a entidade desejar que aqueles sistemas de computação não-gerenciados sejam parte daquela rede de computador virtual ou de outra forma comunicar com os nós de computação gerenciados da rede de computador virtual, tais comunicações entre os nós de computação e sistemas de computação não-gerenciados poderão ser manuseadas pelo(s) módulo(s) Gerenciador(es) de Comunicação que gerenciam um ou mais nós de computação em pelo menos algumas das tais modalidades. Por exemplo, em tais situações, se tal sistema de computação não-gerenciado for fornecido com um endereço de destino IPv6 real para tal modo de computação (ex., ":::0A:01:<Z-identifier>:10.0.0.2" para o nó de computação A gerenciado neste exemplo), o sistema de computação não-gerenciado pode enviar comunicações ao nó de computação A através da rede de interconexão 250 usando aquele endereço de rede de destino, e o módulo gerenciador de comunicação R encaminharia aquelas comunicações ao nó de computação A (ex., após o novo cabeçalho das comunicações de forma semelhante àquela anteriormente

descrita) se o módulo gerenciador de comunicação R for configurado para aceitar comunicações daquele sistema de computação não-gerenciado (ou de qualquer sistema de computação não-gerenciado). Além disso, o módulo gerenciador de comunicação R pode gerar um endereço de rede virtual falso para corresponder a tal sistema de computação não-gerenciado, mapeá-lo para o endereço de rede de IPv6 real para o sistema de computação não-gerenciado, e fornecer o endereço de rede virtual falso para o nó de computação A (ex., como o endereço fonte para as comunicações encaminhadas para o nó de computação A a partir do sistema de computação não-gerenciado), permitindo assim que o nó de computação A envie comunicações ao sistema de computação não-gerenciado.

Semelhantemente, em pelo menos algumas modalidades e situações, pelo menos alguns nós de computação gerenciados e/ou suas redes de computador virtual podem ser configurados para permitir comunicações com outros dispositivos que não são parte da rede de computador virtual, tal como outros sistemas de computação não-gerenciados ou outros tipos de dispositivos de aplicativos de rede que não tem um módulo gerenciador de comunicação relacionado que gerencie suas comunicações. Em tais situações, se os nós de computação e/ou a rede de computador virtual for configurada para permitir comunicações com os tais outros dispositivos não-gerenciados, tal dispositivo não-gerenciado poderá semelhantemente ser fornecido com o endereço de rede de destino IPv6 real para tal nó de computação (ex., "::0A:01 :<Z-identifíer>:10.0.0.2" para o nó de computação A neste exemplo), permitindo que o dispositivo não-gerenciado envie comunicações ao nó de computação A através da rede de interconexão 250 usando aquele endereço de rede de destino, com o módulo gerenciador de comunicação R em seguida encaminhando aquelas comunicações ao nó de computação A (ex., após o novo cabeçalho das comunicações de forma semelhante àquela descrita anteriormente). Além disso, o módulo gerenciador de comunicação R pode semelhantemente gerenciar comunicações de saída a partir do nó de computação A para tal dispositivo não-gerenciado permitindo que o nó de computação A envie tais comunicações.

Além disso, conforme observado anteriormente, um módulo gerenciador de comunicação gerencia as comunicações para nós de computação relacionados de várias formas, inclusive em algumas modalidades designando endereços de rede virtual aos nós de computação de uma rede de computador virtual, e/ou designando endereços de rede física de substrato para nós de computação gerenciados a partir de uma variedade de endereços de rede física de substrato que correspondem ao módulo gerenciador de comunicação, em outras modalidades, algumas das tais atividades podem, em vez disto, ser realizadas por um ou mais nós de computação da rede de computador virtual, tal como para permitir um que um servidor de DHCP (Protocolo de Configuração de Host Dinâmico) ou outro dispositivo de uma rede de computador virtual especifique os endereços de rede virtual e/ou endereços de

rede física de substrato para nós de computação específicos da rede virtual. Em tais modalidades, o módulo gerenciador de comunicação obtém tais informações de configuração a partir dos dispositivos de rede virtual, e atualiza suas informações de mapeamento consequentemente (e em algumas modalidades pode ainda atualizar um ou mais módulos gerenciadores de sistema que mantêm informações sobre os nós de computação relacionados às redes virtuais). Ainda em outras modalidades, um usuário ou outra entidade relacionada a uma rede de computador virtual pode diretamente configurar nós de computação específicos para usar endereços de rede virtual específicos. Sendo assim, os módulos gerenciadores de comunicação e/ou módulo gerenciador de sistema podem rastrear quais endereços de rede virtual são usados por nós de computação específicos, e semelhantemente atualizar informações de mapeamento armazenadas de acordo.

Além disso, em algumas modalidades e situações, um nó de computação gerenciado por ser tratado como roteador fantasma, com múltiplos endereços de rede virtual relacionados àquele nó de computação gerenciada, e com aquele nó de computação gerenciada que encaminha comunicações a outros nós de computação que correspondem àqueles múltiplos endereços de rede virtual. Em tais modalidades, o módulo gerenciador de comunicação que gerencia as comunicações para aquele nó de computação de roteador gerenciado manuseia as comunicações para e a partir daquele nó de computação de forma semelhante àquela descrita anteriormente. No entanto, o módulo gerenciador de comunicação é configurado com múltiplos endereços de rede virtual que correspondem ao nó de computação de roteador gerenciado, para que as comunicações de chegada para quaisquer dos múltiplos endereços de rede virtual sejam encaminhadas ao nó de computação de roteador gerenciado, e para que as comunicações de saída do nó de computação de roteador gerenciado sejam fornecidas como um endereço de rede física de substrato que corresponde ao nó de computação específico que enviou a comunicação através do nó de computação de roteador gerenciado. Desta forma, os roteadores ou outros dispositivos de rede de um cliente específico ou outra entidade podem ser representados virtualmente para uma rede de computador virtual implementada para tal entidade.

A Figura 2B ilustra alguns dos nós de computação discutidos com relação à Figura 2A, mas fornece detalhes adicionais com relação a algumas ações tomadas pelos módulos gerenciadores de comunicação 210 e 260 e/ou ao módulo gerenciador de sistema 290 para autorizar comunicações entre nós de computação. Por exemplo, após o nó de computação A enviar a mensagem 220-a para solicitar um endereço de hardware para o nó de computação G, o módulo gerenciador de comunicação R pode realizar uma ou mais interações com o módulo gerenciador de sistema 290 para determinar o fornecimento daquelas informações, tal como se baseando em se o nó de computação A é autorizado para comunicar com o nó de computação G, bem como para determinar um endereço de rede física de

substrato para o nó de computação G se baseando na rede de interconexão 250. Se o módulo gerenciador de comunicação R tiver sido obtido anteriormente e armazenado tais informações e permanecer válido (ex., não tiver expirado), as interações 225 não poderão se realizadas. Neste exemplo, para obter o endereço de rede física desejado correspondente

5 ao nó de computação G, o módulo gerenciador de comunicação R envia uma mensagem 225-1 ao módulo Gerenciador do Sistema 290 que inclui os endereços de rede virtual para os nós de computação A e G, e que inclui um identificador de rede de entidade para cada um dos nós de computação, que, neste exemplo, é um identificador de rede de entidade para a rede de computador virtual da entidade Z (ex., um identificador único de 32 bits ou 24

10 bits). Em pelo menos algumas modalidades, o módulo gerenciador de comunicação R pode enviar a mensagem 225-1 aos módulos Gerenciadores de Sistemas 290 usando um endereço de anycast e esquema de rota, de modo que os múltiplos módulos Gerenciadores de Sistemas possam ser implementados (ex., uma para cada centro de dados que inclui módulos gerenciadores de comunicação e nós de computação relacionados) e um destes módulos

15 apropriado (ex., o mais próximo, o mais sub-utilizado etc) é selecionado para receber e tratar a mensagem.

Após o módulo gerenciador de sistema 290 determina se o nó de computação A é autorizado para comunicação com o nó de computação G (ex., se baseando em ter o mesmo identificador de rede de entidade, com base no nó de computação A tendo um identifica-

20 dor de rede de entidade que é autorizado a se comunicar com os nós de computação do identificador de rede de entidade para o nó de computação G, com base em outras informações fornecidas por ou associadas ao nó de computação A indicando que o nó de computação A é autorizado a realizar tais comunicações etc.), o módulo Gerenciador de Sistemas 290 retorna uma mensagem de resposta 225-2 que inclui o endereço de rede física de substrato correspondente ao nó de computação G. Além disso, pelo menos em algumas modalidades,

25 antes de enviar o endereço de rede física real desejado, o módulo Gerenciador de Sistemas 290 pode ainda verificar se o módulo gerenciador de comunicação R é autorizado a enviar a mensagem 225-1 em nome do nó de computação A, tal como se baseando no nó de computação A ser determinado como um dos nós de computação ao qual o módulo gerenciador de comunicação R é associado.

30

Em outras modalidades, o módulo gerenciador de comunicação R pode realizar algumas ou todas as ações descritas como sendo realizadas pelo módulo Gerenciador de Sistemas 290, como para manter informações de fornecimento para os vários nós de computação e/ou para determinar se o nó de computação A é autorizado a enviar comunicações

35 ao nó de computação G, ou então nenhuma determinação de autorização pode ser realizada em algumas ou todas as situações. Além disso, em outras modalidades, outros tipos de determinações de autorização podem ser realizados para uma comunicação entre dois ou

mais nós de computação, tal como se baseando no tipo da comunicação, no tamanho da comunicação, no tempo da comunicação etc.

Conforme observado anteriormente com relação à Figura 2A, após o módulo gerenciador de comunicação S receber a comunicação 230-3 destinada ao nó de computação G através da rede de interconexão 250, o módulo gerenciador de comunicação S poderá realizar uma ou mais interações 240 com o módulo gerenciador de sistema 290 para determinar a autorização daquela comunicação. Especificamente, neste exemplo, para verificar se a comunicação 230-3 é válida e autorizada para ser encaminhada ao nó de computação G, o módulo gerenciador de comunicação S primeiro extrai o endereço de rede de destino IPv6 real e o endereço de rede de fonte IPv6 real do cabeçalho da comunicação 230-3, e em seguida recupera os identificadores de rede de entidade e os endereços de rede virtual incorporados a partir de cada um dos endereços de rede IPv6 extraído. O módulo gerenciador de comunicação S em seguida troca as mensagens 240 com o módulo Gerenciador de Sistemas 290 para obter o endereço de rede física IPv6 real correspondente para o nó de computação de envio A em nome do nó de computação G, incluindo uma mensagem 240-4 que inclui os endereços de rede virtual extraídos para os nós de computação A e G e o identificador de rede de entidade para cada um dos nós de computação. Pelo menos em algumas modalidades, o módulo gerenciador de comunicação S pode enviar a mensagem 240-4 ao módulo gerenciador de sistema 290 usando um endereço de anycast e esquema de rota conforme descrito anteriormente.

O módulo Gerenciador de Sistemas 290 recebe a mensagem 240-4, e retorna uma mensagem de resposta 240-5 que inclui o endereço de rede física de substrato correspondente ao nó de computação A, que, neste exemplo, é "::0A:01 :<Z-identifier>: 10.0.0.2". Conforme discutido anteriormente com relação às mensagens 225-1 e 225-2, em algumas modalidades, o módulo Gerenciador de Sistemas 290 e/ou o módulo gerenciador de comunicação S podem ainda realizar um ou mais outros tipos de atividades de determinação de autorização, tal como para determinar se o nó de computação G é autorizado a se comunicar com o nó de computação A, se o módulo gerenciador de comunicação S é autorizado a enviar a mensagem 240-4 em nome do nó de computação G etc. O módulo gerenciador de comunicação S em seguida verifica se o endereço de rede física retornado na mensagem de resposta 240-5 corresponde ao endereço de rede IPv6 de fonte extraído do cabeçalho da comunicação 230-3, para impedir tentativas de forjar mensagens como sendo do nó de computação A que, na verdade, são enviadas a partir de outros nós de computação em outras localizações. O módulo gerenciador de comunicação S opcionalmente armazena estas informações recebidas a partir da mensagem de resposta 240-5 como parte de uma entrada para o nó de computação A nas informações de mapeamento 262 para uso posterior, junto com o endereço de rede virtual do nó de computação A e um endereço de hardware virtual

para o nó de computação A.

A Figura 2C ilustra outro exemplo do gerenciamento de comunicações de contínuas para a rede de computador virtual descrita com relação às Figuras 2A e 2B, mas com as comunicações sendo gerenciadas para suportar funcionalidade de rede lógica para a rede de computador virtual em conformidade com uma topologia de rede configurada para a rede de computador virtual. Especificamente, a Figura 2C ilustra o nó de computação A, os módulos Gerencadores de Comunicação R e S, o módulo Gerenciador de Sistemas 290 e a rede de interconexão 250 de forma semelhante àquela mostrada nas Figuras 2A e 2B. No entanto, a Figura 2C ainda ilustra informações adicionais em relação ao nó de computação A 205a e o nó de computação H 255b em comparação à Figura 2A, bem como as representações lógicas 270a e 270b de dois dispositivos roteadores especificados que são parte da topologia de rede configurada para a rede de computador virtual mas que não são, de fato, fisicamente implementados como parte do fornecimento da rede de computador virtual. Especificamente, neste exemplo, o nó de computação A está enviando uma comunicação ao nó de computação H, e as ações dos módulos fisicamente implementados 210 e 260 e os dispositivos da rede 250 em de fato enviar a comunicação são mostradas, bem como as ações emuladas dos dispositivos roteadores lógicos 270a e 270b no envio lógico da comunicação.

Neste exemplo, os nós de computação A e H são configurados para ser parte de suas sub-redes distintas da rede de computador virtual, e os dispositivos roteadores lógicos 270a e 270b separam os nós de computação A e H na topologia de rede configurada para a rede de computador virtual. Por exemplo, o dispositivo roteador lógico J 270a pode ser um dispositivo roteador local para o nó de computação A (ex., pode gerenciar uma primeira sub-rede que inclui o nó de computação A), o dispositivo roteador lógico L 270b pode ser um dispositivo roteador local para o nó de computação H (ex., pode gerenciar uma segunda sub-rede distinta que inclui o nó de computação H). Embora os nós de computação A e H são ilustrados como sendo separados por dois dispositivos roteadores na topologia de rede configurada neste exemplo, será apreciado que os dois tais nós de computação podem ser separados por 0, 1 ou mais de 2 dispositivos roteadores em outras situações, e que outros tipos de dispositivos de rede podem separar os nós de computação em algumas situações.

No exemplo da Figura 2C, as informações adicionais que são mostradas para os nós de computação A e H incluem endereços de hardware relacionados àqueles nós de computação para a rede de computador virtual, como endereços de hardware que são designados para os nós de computação pelo módulo Gerenciador de Sistemas 290 e/ou os módulos gerenciadores de comunicação R e S. Especificamente, neste exemplo, o nó de computação A foi atribuído ao endereço de hardware "00-05-02-0B-27-44," e o nó de computação H foi atribuído ao endereço de hardware "00-00-7D-A2-34-11." Além disso, os dis-

positivos roteadores lógicos J e L também foram atribuídos a endereços de hardware, que, neste exemplo, são "00-01-42-09-88-73" e "00-01-42-CD-11-01", respectivamente, assim como endereços de rede virtual, que, neste exemplo, são "10.0.0.1" e "10.1.5.1", respectivamente. Os vários endereços de hardware serão usados como parte do envio da comunicação do nó de computação A ao nó de computação H, e do fornecimento de funcionalidade

Assim, de forma semelhante àquela descrita com relação à Figura 2A, o nó de computação A determina o envio de uma comunicação ao nó de computação H, e consequentemente troca várias mensagens 222 com o módulo gerenciador de comunicação R

210. Especificamente, neste exemplo, o nó de computação A primeiro envia uma solicitação de mensagem de ARP 222-a para informações de endereço de hardware virtual. No entanto, ao contrário do exemplo da Figura 2A em que os nós de computação A e G foram parte da mesma sub-rede lógica, é esperado que as comunicações do nó de computação A para o nó de computação H primeiro passem através de um destino intermediário inicial do dispositivo roteador local J antes de serem encaminhadas ao nó de computação H. Consequentemente, uma vez que o roteador lógico J é o destino intermediário inicial para o nó de computação logicamente remoto H, a solicitação de mensagem de ARP 222-a inclui o endereço de rede virtual para o roteador lógico J (i.e., "10.0.0.1") e solicita o endereço de hardware correspondente para o roteador lógico J. Em outras modalidades, o nó de computação A pode, por sua vez, solicitar informações de endereço de hardware virtual para o nó de computação H diretamente (ex., usando o endereço de rede virtual "10.1.5.3" para o nó de computação H), mas receber o endereço de hardware correspondente para o roteador lógico J.

O módulo gerenciador de comunicação R intercepta a solicitação de ARP 222-a, e obtém um endereço de hardware para fornecer ao nó de computação A como parte da mensagem de resposta de ARP jorjada 222-b. O módulo gerenciador de comunicação R pode determinar o endereço de hardware para o roteador lógico J₁ bem como se o nó de computação H é parte de uma sub-rede lógica distinta a partir do nó de computação A, de várias formas em várias modalidades. Por exemplo, como discutido anteriormente, o módulo gerenciador de comunicação R pode armazenar várias informações de endereço de hardware comoparte das informações de mapeamento 212, e, sendo assim, pode já ter armazenado as informações de endereço de hardware para o roteador lógico J. Se não, no entanto, o módulo gerenciador de comunicação R realiza uma ou mais interações 227 com o módulo gerenciador de sistema 290 para obter informações do módulo 290 correspondentes ao endereço de rede virtual indicada para o roteador lógico J. No entanto, em vez de obter um endereço de rede de substrato correspondente ao endereço de rede virtual indicado, como para o nó de computação G na Figura 2A, o módulo Gerenciador de Sistemas 290 indica

que o endereço de rede virtual corresponde a um dispositivo roteador lógico da topologia de rede configurada, e pode também fornecer informações ao módulo gerenciador de comunicação R que indica as informações de endereço de hardware para o roteador lógico J. Especificamente, o módulo Gerenciador de Sistemas 290 mantém várias informações 294 relacionadas à topologia de rede configurada para as redes de computador virtual que fornece ou de outra forma gerencia, tal como informações sobre dispositivos de rede especificados, e usa estas informações para fornecer informações solicitadas aos módulos Gerenciadores de Comunicação. O módulo gerenciador de comunicação R em seguida armazena as informações recebidas como parte das informações de mapeamento 212 para uso futuro, e desta forma determina se o nó de computação H é parte de uma sub-rede distinta do nó de computação A na topologia de rede configurada. Além disso, o módulo gerenciador de comunicação R fornece ao nó de computação A o endereço de hardware "00-01-42-09-88- 73" correspondente ao roteador lógico J como parte da mensagem de resposta 222-b. Embora a solicitação 222-a e a mensagem de resposta 222-b na verdade passem entre o nó de computação A e o módulo gerenciador de comunicação R da forma discutida, a partir do ponto de vista do nó de computação A, as comunicações 222-a e 222-b são parte das interações lógicas 263 que ocorrem com o dispositivo roteador local J.

Após receber a mensagem de resposta 222-b do módulo gerenciador de comunicação R, o nó de computação A cria e inicia o envio de uma comunicação ao nó de computação H, mostrada na Figura 2C como a comunicação 222-c. Especificamente, o cabeçalho de comunicação 222-c inclui um endereço de rede de destino para o nó de computação H de destino que é "10.1.5.3", um endereço de hardware de destino que é o endereço de hardware virtual para o roteador lógico J fornecido ao nó de computação A na mensagem 222-b, um endereço de rede da fonte para o nó de computação A de envio que é "10.0.0.2", e um endereço de hardware da fonte para o nó de computação A de envio que é um endereço de hardware real ou falso que foi anteriormente identificado para o nó de computação A. A partir do ponto de vista do nó de computação A, a comunicação enviada será manuseada da forma ilustrada para a comunicação lógica 265, e será enviada ao roteador lógico local J como comunicação 265a para encaminhamento baseado no endereço de hardware de destino na comunicação. Se o roteador lógico J fosse fisicamente implementado e recebesse tal comunicação 265a, modificaria o cabeçalho da comunicação 265a e encaminharia a comunicação modificada 265b para o roteador lógico L, que semelhantemente modificaria o cabeçalho da comunicação 265b e encaminharia a comunicação modificada 265c ao nó de computação H. As modificações que o roteador lógico J realizaria para tal comunicação 265a podem incluir diminuir um valor de salto de rede de TTL no cabeçalho e mudar o endereço de hardware de destino para corresponder ao próximo destino, que neste exemplo seria o roteador lógico L. Semelhantemente, as modificações que o roteador lógico L realizaria

para tal comunicação 265b podem ainda diminuir o valor de salto de rede de TTL no cabeçalho e mudar o endereço de hardware de destino para corresponder ao próximo destino, que neste exemplo seria o nó de computação H.

Embora a comunicação 222-c do nó de computação A para o nó de computação H seja logicamente manuseada da forma ilustrada para a comunicação 265, a comunicação 222-c é, de fato, interceptada e manuseada pelo módulo gerenciador de comunicação R. Especificamente, de forma semelhante àquela descrita na Figura 2A para a comunicação 220-c, o módulo gerenciador de comunicação R intercepta a comunicação 222-c, modifica a comunicação conforme apropriado e encaminha a comunicação modificada através da rede de interconexão 250 ao nó de computação H. Para determinar o endereço de rede de substrato a ser usado para encaminhar a comunicação através da rede de interconexão 250, o módulo gerenciador de comunicação R extrai o endereço de rede virtual de destino e o endereço de hardware virtual de destino do cabeçalho da comunicação 222-c. No entanto, com base no endereço de hardware virtual de destino correspondente ao roteador lógico J, o módulo gerenciador de comunicação R determina o uso do endereço de rede virtual de destino para identificar o endereço de rede de substrato de destino, de forma diferente daquela da Figura 2A. Assim, o módulo gerenciador de comunicação R verifica as informações de mapeamento 212 para determinar se um endereço de rede de substrato correspondente ao endereço de rede virtual do nó de computação H foi determinado anteriormente e armazenado. Se não, o módulo gerenciador de comunicação R realiza uma ou mais interações 227 com o módulo gerenciador de sistema 290 para determinar aquelas informações, de forma semelhante às interações 225 da Figura 2A. Conforme discutido mais detalhadamente com relação à Figura 2B, em resposta à mensagem de solicitação de ARP 222-a e/ou comunicação 222-c, o módulo gerenciador de comunicação R e/ou o módulo Gerenciador de Sistemas 290 poderão ainda realizar várias atividades de autenticação opcionais.

Depois que o módulo gerenciador de comunicação R determina o endereço da rede de substrato físico real IPv6 correspondente ao nó H de computação, ele cria um novo cabeçalho IPv6 que inclui esse endereço de rede de substrato físico real como endereço de destino e, similarmente, adiciona um endereço de origem IPv6 para o nó A de computação para o novo cabeçalho. Neste exemplo, o endereço da rede física de substrato correspondente a nó H de computação é semelhante ao nó de computação G, e em particular, é o endereço de rede de substrato IPv6 ":: 0B: 02: <Z-identifier> : 10.1.5.3 ", onde " 10.1.5.3 " é armazenado nos últimos 32 bits do endereço IPv6 de 128 bits, e onde " <Z-identifier> " é um identificador de entidade da rede 24-bit para a rede de computadores virtual. Assim, como com as comunicações enviadas ao nó de computação G, uma comunicação enviada pela rede de interconexão 250 para o endereço de rede do substrato para o nó de computação H será encaminhada para o módulo gerenciador de comunicação S. O módulo gerenciador de

comunicação R em seguida cria uma nova comunicação 232-3 modificando a comunicação 222-c, de modo a substituir o cabeçalho IPv4 anterior com o novo cabeçalho IPv6 (por exemplo, de acordo com SIIT), incluindo o preenchimento do novo cabeçalho IPv6 com outras informações conforme apropriado para a nova comunicação (por exemplo, o comprimento de carga útil, a prioridade de pacotes de classe de tráfego, etc), e encaminha a comunicação 232-3 através da rede de interconexão 250. A rede de interconexão, em seguida, utiliza o endereço da rede de destino do IPv6 físico da comunicação 232-3 para encaminhar a comunicação para o módulo gerenciador de comunicação de S. Quando o módulo gerenciador de comunicação S recebe a comunicação 232-3 através da rede de interconexão 250, que realiza ações semelhantes às descritas na Figura 2A que diz respeito à comunicação 230-3, inclusive para modificar a comunicação para incluir um cabeçalho IPv4 apropriado, e proporcionar a comunicação modificada como comunicação 247-e para o nó de computação H.

Além disso, como observado em outros lugares, o módulo gerenciador de comunicação R e/ou módulo gerenciador de comunicação S toma ainda as medidas necessárias neste exemplo para modificar a comunicação do nó de computação A para o nó de computação H, de tal forma a fornecer funcionalidade de rede lógica correspondente para a topologia de rede configurada para a rede de computadores virtual, incluindo para emular uma funcionalidade que seria fornecida p roteadores lógicos J e L se estivessem fisicamente implantados para a rede de computadores virtual. Por exemplo, como foi previamente discutido, roteadores lógicos J e L realizariam várias modificações para comunicação 265 conforme é encaminhado para o nó de computação H se aqueles roteadores estivessem fisicamente implantados e usados, incluindo para modificar valores de salto de rede TTL e para realizar outras modificações de cabeçalho. De acordo com isso, módulo gerenciador de comunicação R e/ou módulo gerenciador de comunicação S pode realizar diferentes modificações para a comunicação 222-c e/ou 247-e para emular tal funcionalidade dos roteadores lógicos J e L. Portanto, nó de computação H recebe uma comunicação 247-e que aparenta ser comunicação 265c encaminhada por meio da topologia de rede especificada para a rede de computador virtual.

Dessa forma, o sistema ONM pode fornecer funcionalidade de rede lógica correspondente a topologia de rede configurada, sem nenhuma configuração especial dos nós de computação da rede de computador virtual ou dos dispositivos de rede física da rede de interconexão substrato de intervenção, baseada nos módulos gerenciadores de comunicação sobrepondo a rede de computador virtual na rede de interconexão física real de maneira a emular a topologia de rede configurada. Além disso, múltiplos módulos do sistema ONM podem operar juntos de maneira distribuída para fornecer funcionalidade correspondente a um particular dispositivo de rede lógica, assim como com os módulos 210, 260 e 290 ope-

rando juntos no exemplo anterior para emular funcionalidade correspondente a cada um dos dispositivos roteadores lógicos 270a e 270b.

Como foi previamente observado, as informações de configuração que são especificadas para rede de computador virtual podem incluir várias informações de topologia de rede, e vários nós de computação podem ser selecionados para a rede de computador virtual e configurados de acordo com a topologia de rede de várias maneiras. Por exemplo, em algumas modalidades, a seleção de um nó de computação para ser usado em uma rede de computador virtual e/ou para ser usado em uma determinada função em uma topologia de rede configurada pode ser baseada ao menos em parte, em uma localização geográfica e/ou de rede do modo de computação, como um local absoluto, em vez de uma localização relativa a um ou mais outros recursos de computação de interesse (por exemplo, outros nós de computação da mesma rede virtual, recursos de armazenamento a serem usados pelo modo de computação, etc.), como se dentro de uma mínima e/ou máxima distância geográfica especificada ou outro grau de proximidade para outro recurso de computação ou outro local indicado. Além disso, em algumas modalidades, os fatores usados quando se seleciona um nó de computação pode não ser baseado no local, de forma a incluir um ou mais dos seguintes: restrições relacionadas com as capacidades de um modo de computação, tal como um critério relacionado com recurso (por exemplo, uma quantidade de memória, uma quantidade de uso de processador, uma quantidade de largura de banda da rede, e /ou uma quantidade de espaço de disco), e/ou capacidades especializadas disponíveis apenas em um subconjunto de = nós de computação disponíveis; restrições relacionadas com custos, tais como as baseadas em taxas ou custo operacionais com uso de particulares nós de computação; etc.

Vários outros tipos de ações além daquelas discutidas com respeito às figuras 2A-2C podem ser realizadas em outras modalidades, incluindo para tipos de protocolos de direcionamento de rede diferentes de IPv4 e IPv6.

Figura 2D ilustra um exemplo de IPv6 configuração de endereço de rede de substrato físico 272 para uso com as técnicas descritas em algumas modalidades, com o exemplo do endereço da rede sendo configurado de modo a incorporar o endereço de rede virtual e outras informações no endereço de rede de substrato de forma a permitir uma sobreposição de rede de computadores virtual através da rede de computadores de substrato. Como foi previamente discutido, este exemplo de configuração de endereço de rede IPv6 usa o espaço de endereço de rede de 128-bit para armazenar várias informações, com os iniciais 64 bits armazenando uma porção de rede IPv6 do endereço, e com os subseqüentes 64 bits armazenando uma porção de identificador de interface (ou "host") do endereço.

Neste exemplo, a porção inicial de rede de 64-bit do endereço IPv6 inclui um identificador de 32-bit 272a para bits 0 até 31 que corresponde a um identificador corporativo ou

de outra organização designado para tal organização por meio de um operador de registro de Internet registry em nome de Internet Assigned Numbers Authority (neste exemplo, com base em uma atribuição de Regional Internet Registry RIPE NNC₁ ou Reseaux IP Europeens Network Coordination Centre). Por exemplo, em algumas modalidades, uma organização opera uma modalidade do sistema ONM ou outra organização que usa as técnicas descritas pode ter um identificador associado 272a. A porção inicial de rede de 64-bit do endereço também inclui um grupo de 32-bit de informações 272b neste exemplo que corresponde a topologia de um grupo de múltiplos nós de computação (por exemplo, uma sub-rede ou outra porção de rede) fornecidos em nome do grupo cujo identificador é indicado nas informações 272a. Como foi previamente discutida, em ao menos algumas modalidades, a porção inicial de rede de 64-bit do endereço representa um endereço de rede parcial para a rede de substrato que corresponde a um local de relacionados múltiplos de computação, como uma sub-rede ou outra porção da rede de substrato, em particular, a inicial porção de endereço de rede de 64-bit em ao menos algumas modalidades corresponde a um particular módulo gerenciador de comunicação que representa múltiplos associados nós de computação sendo administrados pelo módulo gerenciador de comunicação, tal como baseado no módulo gerenciador de comunicação que administra o intervalo de endereços de rede correspondente a algumas ou todas as porções de endereço do identificador de interface de 64-bit a fim de representar os diversos nós de computação administrados. Em outras modalidades, o endereço de rede parcial pode ser representado com um diferente número de bits (por exemplo, 72) e/ou usando uma parte do endereço diferente de um prefixo.

O grupo de 32-bit de informações de topologia 272b pode representar várias informações de várias formas em diferentes modalidades, com os grupos de informações de topologia 274 e 276 apresentando dois exemplos alternativos de configurações de informações de topologia. Em particular, nos exemplos de 274 e 276, os primeiros dois bits (bits 32 e 33 do endereço geral de IPv6) indica uma particular versão das informações de topologia, de tal forma que o significado dos restantes 30 bits pode mudar com o tempo ou em diferentes situações. Com respeito ao exemplo 274, vários bits são mostrados e cada um indica diferentes localidades geográficas, áreas geográficas dentro dos locais, racks de computador dentro das áreas geográficas, e nodos de sistema de computação física dentro dos racks de computador. Neste exemplo, os 6 bits para as informações de local podem representar 64 valores únicos, os 8 bits para as informações de área podem representar 256 valores únicos para cada valor de local, os 8 bits para as informações de podem representar 256 valores únicos para cada valor de área, e os 8 bits para as informações nó do sistema de computação física podem representar 256 valores únicos para cada valor de rack. De maneira inversa, com respeito ao exemplo 276, apenas informações de rack e de local são mostradas, mas cada um possui adicionais bits a fim de representar aqueles tipos de infor-

mação, de forma a ter 16,384 valores únicos de locais usando os seus 14 bits, e tendo 65,536 valores únicos de rack usando os seus 16 bits para cada valor de local, será apreciado que as informações de topologia podem ser representadas de outras maneiras em outras modalidades.

5 Neste exemplo, a porção do identificador de interface de 64-bit do endereço IPv6 é configurada para armazenar vários tipos de informações, incluindo um identificador de 6-bit 272c que corresponde a um particular slot de nó de computação (por exemplo, um particular nó de computação de máquina virtual em um particular sistema de computação físico correspondente a uma porção inicial de rede de 64-bit do endereço IPv6), dois identificadores
10 de 1-bit 272f e 272g, um identificador de 24-bit 272d para incorporar um identificador de rede de entidade (por exemplo, com referência a uma particular rede de computador virtual), e um identificador de 32-bit 272e para incorporar um Endereço de rede IPv4 (por exemplo, o endereço de rede virtual). Os 6 bits para o identificador do slot podem representar aproximadamente 64 valores únicos, os 24 bits para o incorporado identificador de rede de entidade
15 de podem representar aproximadamente 16.8 milhões de valores únicos, e os 32 bits para o incorporado endereço de rede IPv4 pode representar aproximadamente 4.3 bilhões de valores únicos. Neste exemplo, o identificador de 1-bit 272g (bit 70 do endereço IPv6) representa um bit U/L global/local que em algumas modalidades pode indicar tanto se o endereço é administrado globalmente ou localmente, e o identificador de 1-bit 272f (bit 71 do endereço
20 IPv6) representa um bit I/G individual/grupo que em algumas modalidades pode indicar se o endereço corresponde a um único nó de computação ou a um grupo de múltiplos nós de computação (por exemplo, como parte de uma transmissão ou uma transmissão multiplexada *multicast*). Em ao menos algumas modalidades, o bit I/G é ajustado em zero, e o bit U/L é ajustado em um quando o encaminhamento virtual da comunicação correspondente está
25 sendo usado, tal como para uso em uma sub-rede virtual via roteadores fantasmas de nó de computação e/ou para indicar que uma comunicação de chegada com tal endereço de destino entregue para um nó de computação correspondente ao valor do identificador de slot de 6-bit ao invés de um nó de computação correspondente aos valores do endereço de rede IPv4 de 32-bit incorporado e identificador de rede de entidade de 24-bit. Será apreciado que
30 as informações do identificador de interface podem ser representadas de outras maneiras em outras modalidades.

Como foi previamente observado, o sistema ONM pode ao menos em algumas modalidades estabelecer e/ou manter as redes de computador virtuais por meio da operação de um ou mais módulos gerenciadores de comunicação na borda de uma ou mais redes
35 físicas intermediárias, como pela configuração e administrando de outra forma as comunicações para as redes de computador virtuais. Em algumas situações, um módulo de Gerenciador de Comunicação rastreia ou determina de outra forma as redes de computador virtuais

para as quais os nós de computação associados com os módulos pertencem (por exemplo, com base em entidades em cujo nome as redes de computador virtuais operam) como parte do gerenciamento de comunicações para as redes de computador virtuais. A determinação por meio de um módulo gerenciador de comunicação de uma correspondente rede de computador virtual para um nó de computação pode ser realizada de diversas maneiras em várias modalidades como, por exemplo, pela interação com um módulo gestor de sistema que fornece aquela informação, pelo rastreamento de programas software executando tais nós de computação, pelo rastreamento de entidades associadas com tais nós de computação, etc. Por exemplo, quando um particular nó de computação começa a executar um ou mais programas de software em nome de um usuário, e que o usuário também possui outros programas de software executando outros nós de computação, o novo nó de computação executando os programas do usuário pode ser selecionado para ser associado com a rede de computador virtual dos outros nós de computação. Alternativamente, um usuário ou outra entidade pode especificar uma particular rede de computador virtual na qual um nó de computação pertence, como por exemplo, se a entidade mantém múltiplas redes distintas de computador virtuais entre diferentes grupos de nós de computação. Além disso, em ao menos algumas modalidades, um ou mais módulos de gerenciamento de sistema do sistema ONM pode facilitar a configuração de comunicações entre nós de computação, como, por exemplo, pelo rastreamento de e/ou gerenciamento de quais nós de computação pertencem a quais redes de computador virtuais (por exemplo, com base na execução de programas em nome de um cliente ou outra entidade), e por meio do fornecimento de informações sobre o real endereço de rede de substratos físicos que corresponde aos endereços das redes virtuais usadas para uma particular rede de computador virtual (por exemplo, por um particular cliente ou outra entidade).

Como foi previamente observado, em algumas modalidades, um serviço de execução de programa executa programa de clientes terceirizados usando múltiplos sistemas de computação físicos (por exemplo, em um ou mais centros de dados) que as múltiplas máquinas virtuais de host, com cada uma das máquinas virtuais sendo capaz de executar um ou mais programas para um cliente. Em algumas de tais modalidades, os clientes podem fornecer programas para serem executadas pelo serviço de execução de programas, e podem reservar tempo de execução e outros recursos em instalações de hardware físico ou virtual fornecidos pelo serviço de execução de programas. Além disso, clientes e/ou o serviço de execução de programas pode definir redes de computador virtuais que serão usadas pelo serviço de execução de programas para nós de computação do cliente, para fornecer de forma transparente os nós de computação da rede de computador virtual que aparenta operar em uma exclusiva rede física. Além disso, em algumas modalidades, a rede de computador virtual que é administrado por uma modalidade do sistema ONM pode ser uma rede

de computador configurada fornecida por um serviço de rede configurável. Em algumas de tais modalidades, clientes ou outros usuários podem especificar vários tipos de informações de configuração para as suas redes de computador configuradas fornecidas, como, por exemplo, informações de topologia de rede e/ ou restrições de acesso à rede para a rede de computador fornecida, e podem interagir a partir de um ou mais locais remotos com as suas redes de computador configuradas que foram fornecidas.

Figura 3 é um diagrama em bloco ilustrando um exemplo de sistemas de computação adequados para executar uma modalidade de um sistema para gerenciamento das comunicações entre nós de computação. Em particular, Figura 3 ilustra um grupo 399 de sistemas de computação e inter-rede(s), como, por exemplo, um centro de dados ou outro grupo de nós de computação co-localizados. Em algumas modalidades, alguns ou todos os sistemas de computação do grupo 399 podem ser usados por uma modalidade do sistema ONM para fornecer redes de computador virtuais para usuários ou outras entidades. O grupo 399 inclui um sistema de computador de servidor 300, um sistema de computador de host 350 capazes de executar uma ou mais máquinas virtuais, outros sistemas de computação de host 390 que são semelhantes ao sistema de computação de host 350, e um módulo gerenciador de comunicação ótico 360 que administra sistemas de computação de host 390 e que executa um dos sistemas de computação 390 ou em outro sistema de computação (não mostrado). O sistema de computação do gerenciador de sistema 300 e sistemas de computação de host 350 e 390 são conectados uns com os outros por meio de uma rede interna 380, que inclui um dispositivo de rede 362 e outros dispositivos de rede (não mostrados). A rede 380 pode ser uma rede de interconexão que une múltiplas redes físicas separadas (não mostrado) para o grupo 399 e possivelmente fornece acesso para redes externas (não mostrado) e/ou sistemas, como, por exemplo, outros sistemas de computação 395. No exemplo ilustrado, o dispositivo de rede 362 fornece uma porta de entrada entre a rede 380 e sistemas de computação de host 350 e 390. Em algumas modalidades, dispositivo de rede 362 pode, por exemplo, ser um roteador ou uma ponte.

O sistema de computação 300 opera para configurar e gerenciar redes de computador virtuais dentro do grupo 399, bem como para fornecer outras funções (por exemplo, o fornecimento, inicialização, e execução dos programas nos nós de computação). O sistema de computação 300 inclui um CPU 305, vários componentes de I/O 310, armazenamento 330, e memória 320. Os componentes I/O incluem uma tela de exibição 311, conexão de rede 312, unidade de mídia legível por computador 313, e outros dispositivos de I/O 315 (por exemplo, um mouse, teclado, auto-falantes, etc.).

O sistema de computação de host 350 operar para hospedar uma ou mais máquinas virtuais, como, por exemplo, para uso como nós de computação em redes de computador virtuais (por exemplo, nós de computação que executa programas em nome de vários

usuários). O sistema de computação de host 350 inclui um CPU 352, vários componentes I/O 353, armazenamento 351, e memória 355. Embora não ilustrado aqui, os componentes I/O 353 podem incluir componentes similares aos dos componentes I/O 310. módulo gerenciador de comunicação 356 de máquina virtual e uma ou mais máquinas virtuais 358 estão executando na memória 355, com o modulo 356 gerenciando as comunicações para os nós de computação da máquina virtual associada 358. O módulo gerenciador de comunicação 356 mantém diversas informações de mapeamento 354 sobre o armazenamento relacionado aos nós de computação 358 e outros nós de computação, como, por exemplo, de maneira similar as informações de mapeamento 212 e 262 das Figuras 2A-2B. a estrutura de outros sistemas de computação de host 390 pode ser similar a do sistema de computação de host 350, ou ao invés disso alguns ou todos os sistemas de computação de host 350 e 390 podem atuar diretamente como nós de computação executando programas sem usar máquinas virtuais hospedadas, em um arranjo típico, o grupo 399 pode incluir centenas ou milhares de sistemas de computação de host como, por exemplo, aqueles ilustrados aqui, organizados em um grande número de sub-redes físicas distintas e/ou redes.

Uma modalidade de um módulo gestor de sistema 340 está executando na memória 320 do sistema de computação 300. Em algumas modalidades, o Módulo de Gerenciamento de Sistema 340 pode receber uma indicação de múltiplos nós de computação a serem usados como parte de uma rede de computador virtual (por exemplo, um ou mais nós de computação de máquina virtual no sistema de computação de host 350 ou um ou mais nós de computação usando um dos sistemas de computação de host 390), e em algumas situações pode selecionar o particular modo de computação(s) para a rede de computador virtual. Em alguns casos, informações sobre a estrutura e/ou participação de várias redes de computador virtuais podem ser armazenadas na base de dados de provisionamento 332 no armazenamento 330 pelo módulo 340, e fornecido para os módulos gerenciadores de comunicação em vários momentos. Similarmente, em alguns casos, informações sobre a topologia de rede configurada de várias redes virtuais podem ser armazenadas no banco de dados do dispositivo de rede lógica 334 no armazenamento 330 pelo módulo 340, como, por exemplo, de maneira similar para as informações do dispositivo de rede lógica 294 da Figura 2C, e fornecidas para os módulos gerenciadores de comunicação em vários momentos.

Como discutido em maiores detalhes em outros lugares, os Módulos de Gerenciamento de Comunicações 356 e 360 (e outro módulos gerenciadores de comunicação, não mostrado, que gerencia outros associados nós de computação, não mostrado) e o Módulo de Gerenciamento de Sistema 340 pode interagir de várias maneiras para gerenciar as comunicações entre nós de computação, incluindo para fornecer funcionalidade de rede lógica correspondente a topologias de rede configurada para as redes de computador virtuais fornecidas. Tais interações podem, por exemplo, permitir que os nós de computação 358 e/ou

outros nós de computação se intercomuniem sobre as redes de computador virtuais sem nenhuma configuração especial dos nós de computação, pela sobreposição das redes de computador virtuais através da rede 380 e opcionalmente uma ou mais redes externas (não mostradas) sem nenhuma configuração especial do dispositivo de rede 362 ou outros dispositivos de redes (não mostrados), e sem encapsulamento das comunicações.

Será apreciado que sistemas de computação 300, 350, 390, e 395, e dispositivo de rede 362, são meramente ilustrativos e não se destinam a limitar o escopo da presente invenção. Por exemplo, sistemas de computação 300 e/ou 350 podem ser conectados a outros dispositivos que não são ilustrados, inclusive através de uma ou mais redes externas ao grupo 399, como, por exemplo, a Internet ou através da World Wide Web ("Web"). Mais geralmente, um nó de computação ou outro sistema de computação pode compreender qualquer combinação de hardware ou software que pode interagir e realizar os descritos tipos de funcionalidade, incluindo sem limitação desktop ou outros computadores, servidores de base de dados, dispositivos de armazenamento de rede e outros dispositivos de rede, PDAs, aparelhos celulares, fones sem fio, pagers, organizadores eletrônicos, Aplicativos de internet, sistemas baseados em televisão (por exemplo, usando conversores e/ou gravadores de vídeo pessoais/ digitais), e vários outros produtos de consumo que incluem apropriadas capacidades de comunicação. Além disso, a funcionalidade fornecida pelos módulos ilustrados pode em algumas modalidades ser combinada em menos módulos ou ser distribuída por mais módulos. Similarmente, em algumas modalidades as funcionalidades de alguns módulos ilustrados podem não ser fornecidas e/ou outras funcionalidades podem estar disponíveis.

Também será apreciado que, enquanto vários itens são ilustrados como armazenados na memória ou no armazenamento enquanto são usados, estes itens ou porções deles podem ser transferidos entre a memória e outros dispositivos de armazenamento para fins de gerenciamento da memória e integridade dos dados. Alternativamente, em outras modalidades alguns ou todos os módulos de software e/ou sistemas podem executar na memória em outro dispositivo e se comunicar com os sistemas de computação ilustrados por meio de comunicação inter-computador. Além disso, em algumas modalidades, alguns ou todos os sistemas e/ou módulos podem ser implantados ou fornecidos de outras maneiras, como, por exemplo, ao menos parcialmente em firmware e/ou hardware, incluindo, mas não limitado a, um ou mais circuitos integrados de aplicações específicas (ASICs), circuitos integrados padrão, controladores (por exemplo, executando instruções apropriadas, e incluindo micro-controladores e/ou controladores incorporados), Arranjo de Portas Programável em Campo (FPGAs), dispositivos lógicos programáveis complexos (CPLDs), etc. Alguns ou todos os módulos, sistemas e estruturas de dados também podem ser armazenados (por exemplo, como instruções de software ou dados estruturados) em um meio legível por computador,

como, por exemplo, um disco rígido, uma memória, uma rede, ou um artigo de mídia portátil para ser lido por uma unidade apropriada ou através de uma conexão apropriada. Os sistemas, módulos e estruturas de dados também podem ser transmitidos como sinais de dados gerados (por exemplo, como parte de uma onda transportadora ou outro sinal analógico ou digital propagado) em uma variedade de meios de transmissão legíveis por computador, incluindo meios sem fio ou baseados em cabos/ fios, e podem ter uma variedade de formatos (por exemplo, como parte de um único sinal analógico ou multiplexado, ou como múltiplos distintos pacotes digitais ou estruturas). Produtos desses programas de computador também podem assumir outras formas em outras modalidades. De acordo com isso, a presente invenção pode ser praticada com outras configurações do sistema de computador.

Figuras 4A a 4B são um fluxograma de um exemplo de modalidade de uma rotina do Gerenciador do Sistema ONM 400. A rotina pode ser fornecida por, por exemplo, execução do Módulo de Gerenciamento de Sistema 110 da Figura 1, o Módulo de Gerenciamento de Sistema 290 das Figuras 2A-2C, e/ou o Módulo de Gerenciamento de Sistema 340 da Figura 3, como, por exemplo, to auxiliar no gerenciamento de comunicações entre múltiplos nós de computação através de uma ou mais redes intermediárias, incluindo para gerenciar as comunicações a fim de fornecer funcionalidade de rede lógica correspondente a topologias de rede configurada de redes de computador virtuais, bem como para realizar outros tipos de operações de gerenciamento em algumas situações. Em ao menos algumas modalidades, a rotina pode ser fornecida como parte de um sistema que faz o gerenciamento de comunicações para várias entidades diferentes através de uma rede comum intermediária, com as comunicações configuradas a fim de permitir que cada nó de computação se comunique de forma transparente com outros nós de computação associados usando uma rede de computador virtual privada que é específica para a entidade. Além disso, a rotina pode facilitar a prevenção de comunicações não autorizadas sendo fornecidas para nós de computação de destino, como, por exemplo, pela assistência do módulos gerenciadores de comunicação com determinações de as comunicações são autorizadas.

Na modalidade ilustrada, a rotina começa no bloco 405, onde uma solicitação é recebida. A rotina continua até o bloco 410 para determinar o tipo de solicitação. Caso seja determinado que o tipo de solicitação seja para associar um ou mais nós de computação com uma particular entidade indicada e/ou rede de computador virtual de uma entidade, como, por exemplo, se aqueles nós de computação devem ser parte da rede de computador virtual para a entidade (por exemplo, estão executando ou vão executar um ou mais programas em nome daquela identidade), a rotina continua até o bloco 415 para associar aqueles nós de computação com aquela entidade indicada e rede de computador virtual. Em algumas modalidades, a rotina pode ainda determinar um ou mais nós de computação para serem associados com a entidade indicada e rede de computador virtual, como, por exem-

plo, com base nas informações fornecidas pela entidade indicada, enquanto em outras modalidades a seleção de tais nós de computação e/ou execução de adequados programas naqueles nós de computação podem ser realizadas de outras formas. Além disso, como discutido em maiores detalhes em outros lugares, em algumas modalidades um ou mais dos

5 nós de computação podem ser cada um a máquina virtual que é hospedada por um ou mais sistemas de computação físicos. A rotina continua então até o bloco 420 para armazenar uma indicação do modo de computação(s) e suas associações com a entidade indicada e rede de computador virtual. Em particular, na modalidade ilustrada a rotina armazena uma

10 indicação de um endereço físico de rede de substrato correspondente a um modo de computação, o endereço de rede virtual usado pela entidade para o nó de computação como parte da rede de computador virtual, opcionalmente um endereço de hardware virtual designado para o modo de computação, e uma indicação da entidade associada. Como discutido em maiores detalhes em outros lugares, o endereço físico de rede de substrato correspondente ao nó de computação pode em algumas modalidades ser um endereço de rede de

15 substrato específico para aquele único modo de computação, enquanto em outras modalidades pode ao invés disso se referir a uma sub-rede ou outro grupo de múltiplos nós de computação, como, por exemplo, pode ser gerenciado por um associado módulo gerenciador de comunicação.

Se ao invés disso for determinado no bloco 410 que o tipo de solicitação recebida é

20 uma solicitação para resolução de endereço para o endereço de rede virtual de um nó de computação ou outro dispositivo de rede, como, por exemplo, a partir de um módulo gerenciador de comunicação em nome de um nó de computação gerenciado, a rotina continua ao invés para o bloco 425, onde ele determina se a solicitação está autorizada de uma ou mais formas, como, por exemplo, com base se o nó de computação gerenciado em nome de

25 quem a solicitação é feita é autorizado a enviar comunicações para um nó de computação cuja resolução do endereço de rede virtual é solicitada (por exemplo, com base na rede de computador virtual(s) da qual os dois nós de computação pertencem), com base se o nó de computação gerenciado em nome de quem a solicitação é feita é um nó de computação válido que é atualmente parte de uma rede de computador virtual configurada, e/ou com

30 base se a solicitação é recebida a partir do módulo gerenciador de comunicação que de fato gerencia o indicado nó de computação em nome de quem a solicitação é feita. Se uma solicitação é determinada como sendo autorizada, a rotina continua até o bloco 430, onde ela obtém o endereço de rede virtual de interesse para uma particular rede de computador virtual, como, por exemplo, pode ser identificado com base em um obtido identificador de rede

35 de entidade para a rede de computador virtual ou outro indicador da entidade associada com a rede de computador virtual (por exemplo, uma única marcação alfanumérica ou numérica), como, por exemplo, incluída com a solicitação recebida no bloco 405. A rotina con-

5 continua então até o bloco 435 para recuperar informações armazenadas para o nó de computação que está associada com o endereço de rede virtual para a rede de computador virtual, e em particular para as informações que associam aquele endereço de rede virtual com um endereço físico de rede de substrato para um local de rede que corresponde ao modo de

10 computação, como, por exemplo, pode ser previamente armazenado com respeito ao bloco 420, e opcionalmente para outras informações para o endereço de rede virtual (por exemplo, um endereço de hardware virtual associado, uma indicação sobre se o endereço de rede virtual corresponde a um nó de computação fisicamente implementado com um real endereço de rede de substrato ou ao invés para um dispositivo de rede lógica que não possui um

15 real endereço de rede de substrato, informações sobre uma função ou um estado do dispositivo correspondente ao endereço de rede virtual com respeito às informações de topologia de rede configuradas, etc.). Depois do bloco 435, a rotina continua até 440 para fornecer uma indicação das informações recuperadas para um solicitador. Se o endereço de rede virtual é para um dispositivo de rede lógica que não possui um endereço de rede de substrato físico associado, por exemplo, a rotina pode fornecer uma indicação no bloco 440 referente

20 ao estado ou tipo de dispositivo correspondente ao endereço de rede virtual. Embora não ilustrado aqui, se a determinação no bloco 425 determina que a solicitação não é autorizada, a rotina pode ao invés não executar os blocos 430-440 para aquela solicitação, como, por exemplo, respondendo com uma mensagem de erro para uma solicitação recebida no bloco

405 ou não respondendo àquela solicitação recebida. Além disso, em outras modalidades a rotina pode realizar um ou mais outros testes para validar a solicitação recebida antes de responder com as informações solicitadas, como, por exemplo, para verificar que o nó de computação que iniciou uma solicitação é autorizado para receber aquelas informações.

25 Se ao invés disso for determinado no bloco 410 que a solicitação recebida é para configurar as informações de topologia de rede para uma indicada rede de computador virtual, como, por exemplo, de um usuário associado com aquela rede de computador virtual, a rotina continua até o bloco 470 para receber indicações de informações de topologia de rede e opcionalmente outras informações de configuração para uma indicada rede de computador virtual, como, por exemplo, para especificar um ou mais dispositivos de redes da topologia

30 de rede que será representada pelos y dispositivos de rede lógica cujas funcionalidades são emuladas. No bloco 475, a rotina determina então endereços de hardware virtuais para alguns ou todos os dispositivos de rede especificados, como, por exemplo, para uso posterior por nós de computação na tentativa de enviar comunicações por meio daqueles dispositivos de rede lógica. Depois do bloco 475, a rotina continua até o bloco 480 para armazenar

35 informações sobre os dispositivos de rede lógica e outras informações de topologia de rede configuradas para a indicada rede de computador virtual.

Se ao invés disso for determinado no bloco 410 que a solicitação recebida é de ou-

tro tipo, a rotina continua ao invés para o bloco 485 para executar outra operação indicada conforme apropriado. Por exemplo, em algumas modalidades, a rotina pode receber solicitações para atualizar informações armazenadas sobre particulares nós de computação, como, por exemplo, se um particular nó de computação foi previamente associado com uma particular entidade e/ou rede de computador virtual, mas aquela associação termina (por exemplo, um ou mais programas sendo executados para aquela entidade naquele nó de computação são terminados, o nó de computação falha ou se torna indisponível de alguma outra forma etc.). A rotina pode realizar também uma variedade de outras ações relacionadas com o gerenciamento de um sistema de múltiplos nós de computação, como discutido em maiores detalhes em outros lugares, incluindo determinar automaticamente para iniciar movimentos de particulares nós de computação. Além disso, embora não ilustrado aqui, em outras modalidades a rotina às vezes pode executar ações de outros tipos, como, por exemplo, para executar ocasionais operações de manutenção para analisar e atualizar informações armazenadas conforme apropriado, como, por exemplo, depois que períodos pré-definidos de tempo foram expirados. Além disso, se os problemas de validação possíveis são detectados, como, por exemplo, com respeito a solicitações de resolução de endereço recebido para endereços das redes virtuais, a rotina pode realizar várias ações para sinalizar um erro e/ou realizar outras ações correspondentes conforme apropriado.

Depois dos blocos 420, 440, 480 e 485, a rotina continua até o bloco 495 para determinar se continua, como, por exemplo, até que uma indicação explícita de terminação seja recebida. Se estiver determinada a continuação, a rotina retorna ao bloco 405, e se não continuar ao bloco 499 e termina.

Figuras 5A a 5B são diagramas de fluxo de um exemplo de modalidade de uma rotina de Gerenciador de Comunicação ONM 500. A rotina pode ser fornecida por, por exemplo, execução dos módulos gerenciadores de comunicação 109a, 109b, 109c, 109d e/ou 150 da Figura 1, o módulos gerenciadores de comunicação 210 e/ou 260 das Figuras 2A-2C, e/ou o módulos gerenciadores de comunicação 356 e/ou 360 da Figura 3, como, por exemplo, para gerenciar as comunicações para e de um grupo associado de um ou mais nós de computação a fim de fornecer uma privada rede de computador virtual sobre uma ou mais redes intermediárias compartilhadas, incluindo para determinar se para autorizar comunicações para e/ou dos nós de computação gerenciados, e para apoiar o fornecimento de funcionalidade de rede lógica correspondente as topologias de rede configuradas para redes de computador virtuais.

A rotina começa no bloco 505, onde uma indicação é recebida de um nó de comunicação ou outra mensagem. A rotina continua até o bloco 510 para determinar o tipo de comunicação ou outra mensagem e proceder de acordo com isso. Caso seja determinado no bloco 510 que a mensagem é uma solicitação a partir de um associado nó de computa-

ção gerenciado para resolução de endereço de rede, como, por exemplo, uma solicitação ARP, a rotina continua até o bloco 515 para identificar o endereço de rede virtual de interesse indicado na solicitação. A rotina continua então até o bloco 520 para enviar uma solicitação para um módulo gestor de sistema para resolução do endereço de rede virtual para o indicado endereço de rede virtual para a rede de computador virtual associada com o nó de computação que forneceu uma solicitação como, por exemplo, discutido com relação aos blocos 425-440 das Figuras 4A a 4B. Como discutido em maiores detalhes em outros lugares, a rotina pode em algumas modalidades rastrear informações sobre redes de computador virtuais e/ou entidades associadas com cada nó de computação gerenciado, bem como informações de topologia de rede configuradas para redes de computador virtuais, enquanto em outras modalidades ao menos algumas de tais informações podem ao invés disso ser fornecidas para uma rotina pelos nós de computação e/ou pelo Módulo de Gerenciamento de Sistema, ou ao invés disso o Módulo de Gerenciamento de Sistema pode rastrear e armazenar e aquelas informações sem elas serem fornecidas para ou rastreadas pela atual rotina. Embora não ilustrado aqui, em outras modalidades e situações tais solicitações de resolução de endereço podem ser administradas de outras maneiras. Por exemplo, se um nó de computação sendo gerenciado por um particular módulo gerenciador de comunicação fornece uma solicitação de resolução de endereço para outro nó de computação isso também é gerenciado por aquele módulo gerenciador de comunicação, a rotina pode ao invés disso responder a uma solicitação sem interação com o Módulo de Gerenciamento de Sistema, como, por exemplo, com base em informações localmente armazenadas. Além disso, enquanto na modalidade ilustrada a solicitação recebida é uma solicitação para fornecer um endereço de hardware de camada de ligação do nó de computação que corresponde a um indicado endereço de camada de rede, em outras modalidades a solicitação de resolução de endereço pode ter outras formas, ou nós de computação podem solicitar outros tipos de informações sobre nós de computação que indicaram endereços das redes virtuais.

Na modalidade ilustrada, a rotina a seguir continua para o bloco 525 para receber uma resposta do Módulo de Gerenciamento de Sistema que inclui um endereço físico de rede de substrato e/ou outras informações correspondentes ao indicado endereço de rede virtual (por exemplo, uma indicação de que o endereço de rede virtual corresponde a um dispositivo de rede lógica que não é fisicamente implementado), e armazena informações localmente que mapeiam aquele endereço físico de rede de substrato e/ou outras informações para um único endereço de hardware para uso posterior pela rotina (por exemplo, baseado em um endereço de hardware falso virtual gerado por uma rotina ou fornecido na resposta). A rotina então fornece o endereço de hardware para um nó de comunicação solicitante, que ele usará como parte de comunicações que ele envia para o nó de computação com o indicado endereço de rede virtual. Como discutido em maiores detalhes em outros

lugares, a resposta do endereço físico de rede de substrato que é fornecida pode em algumas modalidades incluir um endereço físico de rede de substrato que é específico para o indicado nó de computação de interesse, enquanto em outras modalidades o endereço físico de rede de substrato pode corresponder a uma sub-rede ou outro grupo de múltiplos nós de computação do qual o indicado nó de computação pertence, como, por exemplo, para 5 corresponder para outro módulo gerenciador de comunicação que gerencia aqueles outros nós de computação. A rotina continua então até o bloco 530 para determinar se os blocos 515-525 foram realizados como parte do gerenciamento de um nó de comunicação de saída, como discutido com relação aos blocos 540-560, e neste caso, continua até o bloco 547.

10 Embora não ilustrado aqui, em algumas modalidades a rotina pode ao invés disso recebe uma resposta de erro a partir de um Módulo de Gerenciamento de Sistema (por exemplo, com base em um nó de computação solicitante não sendo autorizado para comunicar com o indicado nó de computação de destino) ou nenhuma resposta, e neste caso pode não enviar nenhuma resposta para um nó de comunicação solicitante ou pode enviar uma correspondente 15 mensagem de erro para aquele modo de computação.

Se ao invés disso for determinado no bloco 510 que o tipo de comunicação ou outra mensagem é um nó de comunicação de saída de um nó de computação gerenciado por uma rotina para outro indicado nó de computação de destino remoto que não é gerenciado pela rotina, a rotina continua até o bloco 540 para identificar o indicado endereço de hardware para o nó de computação de destino do cabeçalho de comunicação. No bloco 545, a rotina 20 determina então se aquele endereço de hardware de destino é um endereço de hardware previamente mapeado com um endereço físico de rede de substrato correspondente ao nó de computação de destino (ou para uma indicação de que o endereço de hardware de destino corresponde a um dispositivo de rede lógica), tal como foi previamente discutido com relação ao bloco 525. Se não, em algumas modalidades a rotina continua até o bloco 515 25 para realizar os blocos 515-525 para determinar tal endereço de rede físico correspondente para o nó de comunicação de saída, enquanto em outras modalidades tais ações não são realizadas (por exemplo, se o indicado endereço de hardware não for um endereço mapeado, a rotina pode fazer com que o nó de comunicação de saída falhe, como, por exemplo, com uma mensagem de erro de volta para o nó que enviou).

30

Se o indicado endereço de hardware é um endereço mapeado, ou a verificação não for realizada, a rotina continua até o bloco 547 para determinar se o endereço de hardware de destino corresponde a um dispositivo de rede lógica que é parte de uma topologia de rede configurada para a rede de computador virtual. Neste caso, a rotina continua até o bloco 549 para identificar o endereço de rede virtual do cabeçalho de comunicação de destino para uso no direcionamento da comunicação para a destinação pretendida, e então continua 35 até o bloco 555. Em particular, a rotina no bloco 549 determina um endereço físico de rede

de substrato que corresponde ao identificado endereço de rede virtual de destino, como, por exemplo, a partir de informações de mapeamento previamente armazenadas ou por meio da interação com um módulo gestor de sistema para realizar a solicitação de resolução de endereço de rede de maneira similar a dos blocos 520 e 525.

5 Se ao invés disso for determinado no bloco 547 que o endereço de hardware de destino não corresponde ao dispositivo de rede lógica, a rotina continua até o bloco 550 para recuperar o endereço físico de rede de substrato que é mapeado para o endereço de hardware. No bloco 555, a rotina então reescreve o cabeçalho de comunicação de acordo com um protocolo de endereço de rede para uma ou mais redes intermediárias entre os nós

10 de computação de envio e destinação usando o endereço físico de rede de substrato recuperado no bloco 550 ou determinado no bloco 549. Reescrever o cabeçalho pode ainda incluir a mudança de outras informações no novo cabeçalho, incluindo a mudança do endereço de rede virtual para que o nó de computação de envio seja um correspondente endereço físico de rede de substrato, e em ao menos algumas modalidades inclui a modificação da

15 comunicação recebida sem encapsulamento como parte de uma sobreposição da rede de computador virtual sobre uma ou mais redes físicas intermediárias do substrato. Além disso, para uma comunicação cujo endereço de hardware de destino corresponde ao dispositivo de rede lógica, a rotina no bloco 555 pode realizar ainda outras modificações que correspondem ao fornecimento da funcionalidade de rede lógica para emular as ações e funcionalidade que seriam realizadas por um ou mais dispositivos de rede lógica que seriam usados

20 para encaminhar a comunicação para o nó de computação de destino de acordo com a topologia de rede configurada para a rede de computador virtual. No bloco 560, a rotina então facilita o fornecimento de comunicação de saída modificada para o nó de computação de destino, como, por exemplo, iniciando o encaminhamento da comunicação de saída modificada sobre as redes intermediárias do substrato para o nó de computação de destino. Em-

25 bora não ilustrado aqui, em outras modalidades vários tipos adicionais de processamento podem ser realizados para comunicações de nó de saída, como, por exemplo, para verificar que as comunicações são válidas ou autorizadas de outra maneira, de diversas formas (por exemplo, para verificar que o nó de computação de envio é autorizado a enviar comunicações para o nó de computação de destino, como, por exemplo, com base na sua associação

30 com a mesma entidade ou parte da mesma rede de computador virtual, com base nos nós de computação de envio e destinação estando associados com diferentes entidades que estão autorizadas a se intercomunicarem, com base no tipo de comunicação ou outras informações específicas para a comunicação, etc.).

35 Se ao invés disso for determinado no bloco 510 que a mensagem recebida é uma comunicação de nó de chegada para um dos nós de computação gerenciados a partir de um nó de computação externo, a rotina continua até o bloco 565 para identificar o endereço de

rede de substratos físicos para os nós de computação de envio e destinação do cabeçalho de comunicação. Depois do bloco 565, a rotina continua até o bloco 570 para verificar opcionalmente que a comunicação de chegada é válida de uma ou mais formas. Por exemplo, a rotina pode determinar se o endereço físico de rede de substrato para o nó de comunicação de envio é de fato mapeado para um nó de computação que corresponde ao local do endereço físico de rede de substrato fonte, como, por exemplo, com base nas interações com um módulo gestor de sistema e/ou com base em outras informações previamente obtidas e armazenadas pela rotina. Além disso, a rotina pode determinar se o endereço físico de rede de substrato para o nó de comunicação de destino corresponde a um atual nó de computação gerenciado. Embora não ilustrado aqui, se uma comunicação de chegada é determinada como inválida, a rotina pode realizar várias ações não mostrado, como, por exemplo, para gerar um ou mais erros e realizar processamento associado e/ ou deixar cair a comunicação de chegada sem encaminhá-la para o nó de destino indicado. Por exemplo, se a comunicação de chegada indicar um endereço de rede de destino que não corresponde ao atual nó de computação gerenciado, a rotina pode derrubar a comunicação de chegada e/ou iniciar uma mensagem de erro, apesar de que em algumas modalidades tais mensagens de erro não são enviadas para o nó de computação de envio, ou outras ações podem ser tomadas para apoiar o encaminhamento de comunicações para um nó de computação movido que foi previamente apoiado pela rotina.

Na modalidade ilustrada, depois do bloco 570, a rotina continua até o bloco 575 para recuperar o endereço de hardware e o endereço de rede virtual que são mapeados para o endereço de rede de substrato físico de destino, e para reescrever o cabeçalho de comunicação para a rede de computador virtual para que pareça que ele foi enviado para um nó de computação com aquele endereço de rede virtual e endereço de hardware. Por exemplo, em algumas modalidades o endereço de rede virtual de destino pode ser obtido a partir do próprio endereço físico de rede de substrato de destino, como, por exemplo, a partir de um subconjunto dos bits do endereço físico de rede de substrato de destino. Além disso, o endereço de hardware de destino pode ter sido previamente mapeado para o endereço de rede de substrato físico de destino, tal como foi previamente discutido com relação ao bloco 525. Nas situações em que tal mapeamento prévio não ocorreu, a rotina pode ao invés disso realizar os blocos 515- 525 para obter tal informação. A rotina pode similarmente reescrever o cabeçalho de comunicação para a rede de computador virtual para que pareça que ele foi enviado a partir de um nó de computação com uma fonte de endereço de rede virtual e fonte de endereço de hardware correspondente a o nó de computação de envio. Além disso, em ao menos algumas modalidades, a rotina no bloco 575 pode realizar ainda outras modificações para a comunicação de chegada que corresponde ao fornecimento da funcionalidade de rede lógica para emular as ações e funcionalidades que seriam realizadas por um ou

mais dispositivos de rede lógica que teriam sido usados para encaminhar a comunicação para o nó de computação de destino de acordo com a topologia de rede configurada para a rede de computador virtual. Depois do bloco 575, a rotina continua até o bloco 580 para facilitar o fornecimento da comunicação de chegada modificada para o nó de computação de destino, como, por exemplo, iniciando o encaminhamento da comunicação de chegada modificada para o nó de destino.

Se ao invés disso for determinado no bloco 510 que uma mensagem de outro tipo é recebida, a rotina continua até o bloco 585 para executar outra operação indicada conforme apropriado, como, por exemplo, para armazenar informações sobre entidades associadas com particulares nós de computação, armazena as informações sobre topologias de rede configurada para particulares redes de computador virtuais, responde as solicitações e outras mensagens de nós de computação de modo a fornecer funcionalidades de rede lógica correspondentes às topologias de rede configurada para redes de computador virtuais (por exemplo, emulando ações e outras funcionalidades que seriam realizadas por específicos dispositivos de rede lógica se eles estivessem fisicamente implantados), atualização previamente mapeada ou informações armazenadas para refletir as mudanças com respeito aos nós de computação que estão sendo gerenciados ou para remotos nós de computação, etc.

Depois dos blocos 560, 580, ou 585, ou se ao invés disso for determinado no bloco 530 que um processamento não está sendo realizado com relação a uma comunicação de saída, a rotina continua até o bloco 595 para determinar se deve continuar, como, por exemplo, até que uma indicação explícita para terminar seja recebida. Caso seja determinado que continue, a rotina retorna ao bloco 505, e caso não continue, vai para o bloco 599 e termina.

Além disso, várias modalidades podem fornecer mecanismos para os usuários do cliente e outras entidades para interagir com uma modalidade do Módulo de Gerenciamento de Sistema a fim de configurar os nós de computação e suas comunicações. Por exemplo, algumas modalidades podem fornecer um console interativo (por exemplo, um programa de aplicação do cliente fornecendo uma interface de usuário interativa, um navegador baseado em interface de Web, etc.) a partir dos quais os usuários podem realizar o gerenciamento da criação ou deleção de redes de computador virtuais, a configuração de informações de topologia de rede para redes de computador virtuais, e a especificação da participação da rede virtual, bem como funções administrativas mais generalizadas relacionadas com a operação e gerenciamento de aplicações hospedadas (por exemplo, a criação ou modificação de contas do usuário; o fornecimento de novas aplicações; o início, término, ou monitoramento de aplicações hospedadas; a designação de aplicações para grupos; a reserva de tempo ou outros recursos do sistema; etc.). Em algumas modalidades, algumas ou todas as funcionalidades de uma modalidade do sistema ONM pode ser fornecidas em troca de taxas de usu-

ários ou outras entidades, e neste caso os mecanismos para os usuários do cliente e outras entidades para interagirem com uma modalidade do Módulo de Gerenciamento de Sistema podem incluir mecanismos para usuários e outras entidades para fornecer o pagamento e informações relacionadas com o pagamento, bem como monitorar as correspondentes informações sobre o pagamento. Além disso, algumas modalidades podem fornecer um API que permite que outros sistemas de computação e programas programaticamente invocar, pelo menos, algumas das funcionalidades descritas como, por exemplo, APIs fornecidos por bibliotecas ou interfaces de aulas (por exemplo, a serem invocadas pelos programas escritos em C, C++, ou Java) ou de outra forma, e/ou usando protocolos de serviços de rede como, por exemplo, por meio de serviços da Web. Adicionais detalhes relacionadas com a operação dos exemplos de modalidades de um serviço de execução de programa com os quais as técnicas descritas podem ser usadas estão disponíveis no Pedido nº US 11/394,595, depositado em 31 de março de 2006 e chamado "Managing Communication Between Computing Nodes"; Pedido nº US 11/395,463, depositado em 31 de março de 2006 e chamado "Managing Execution of Programs by Multiple Sistema de computações;" Pedido nº US 11/692,038, depositado em 27 de março de 2007 e chamado "Configuring Intercommunications Between Computing Nodes" ; e Pedido nº US 12/332,214, depositado em 10 de dezembro de 2008 e chamado "Providing Access To Configurable Private Computer Networks"; em que cada um destes é incorporados no presente documento por referência em sua forma integral. Além disso, detalhes adicionais relacionados com o gerenciamento de redes virtuais fornecidas que podem ser usadas por ao menos algumas modalidades de um sistema ONM estão disponíveis no Pedido nº US 12/060,074, depositado em 31 de março de 2008 e chamado "Configuring Communication Between Computing Nodes" ; que também está incorporado no presente documento por referência em sua forma integral.

Também será apreciado que, apesar de em algumas modalidades as técnicas descritas serem aplicadas em um contexto de um centro de dados abrigando múltiplas máquinas físicas hospedando máquinas virtuais, outros cenários de implementação também são possíveis. Por exemplo, as técnicas descritas podem ser aplicadas no contexto de uma rede em toda a organização ou redes operadas por um negócio ou outra instituição (por exemplo, universidade) para benefício dos seus empregados e/ou membros. Alternativamente, as técnicas descritas poderiam ser aplicadas por um fornecedor de serviço de rede para melhorar a segurança, disponibilidade e isolamento da rede. Além disso, as modalidades do exemplo podem ser utilizadas dentro um centro de dados ou outro contexto para uma variedade de finalidades. Por exemplo, operadores de centros de dados ou usuários que vendem acesso a aplicações hospedadas para os clientes podem em algumas modalidades usar as técnicas descritas para fornecer isolamento de rede entre os dados de seus clientes e aplicações; as equipes de desenvolvimento de software podem em algumas modalidades use

as técnicas descritas para fornecer isolamento de rede entre várias modalidades que eles usam (por exemplo, desenvolvimento, construção, teste, implantação, produção, etc.); as organizações em algumas modalidades podem usar as técnicas descritas para isolar os recursos de computação utilizados por um grupo de funcionários ou departamento (por exemplo, recursos humanos) a partir dos recursos de computação utilizados por outro grupo de funcionários ou departamento (por exemplo, contabilidade); ou operadores de centros de dados ou usuários que estão implantando uma aplicação de múltiplos componentes (por exemplo, um aplicativo de negócios multifuncional) podem em algumas modalidades usar as técnicas descritas para fornecer decomposição funcional e/ ou isolamento para os vários tipos de componentes (por exemplo, interfaces de usuário da Web, servidores de base de dados, motores de regras de negócio, etc.). Mais geralmente, as técnicas descritas podem ser usadas para virtualizar redes físicas para refletir quase qualquer situação que convencionalmente exigem particionamento físico de diferentes sistemas de computação e/ou networks.

Também será apreciado que em algumas modalidades as funcionalidades fornecidas pelas rotinas discutidas acima podem ser fornecidas de maneiras alternativas, como, por exemplo, sendo divididas entre mais rotinas ou consolidadas em menos rotinas. Similarmente, em algumas modalidades as rotinas ilustradas podem fornecer mais ou menos funcionalidades do que foi descrito, por exemplo, como quando outras rotinas ilustradas, ao contrário, não possuem ou incluem tal funcionalidade respectivamente, ou quando a quantidade de funcionalidade que é fornecida é alterada. Além disso, enquanto várias operações podem ser ilustradas como sendo realizadas de maneira particular (por exemplo, em série ou em paralelo) e/ ou em uma ordem particular, as pessoas versadas na técnica apreciarão que em outras modalidades as operações podem ser realizadas em outras ordens e de outras maneiras. As pessoas versadas na técnica apreciarão também que as estruturas de dados discutidas acima podem ser estruturadas de diferentes maneiras, como, por exemplo, tendo uma única estrutura de dados dividida em múltiplas estruturas de dados ou tendo múltiplas estruturas de dados consolidadas em uma única estrutura de dados. Similarmente, em algumas modalidades ilustradas as estruturas de dados podem armazenar mais ou menos informações do que é descrito como, por exemplo, quando outras ilustradas estruturas de dados ao invés disso não apresentam ou incluem tais informações respectivamente, ou quando a quantidade de tipos de informações que são armazenadas é alterada.

Cláusula 1. Um método para um sistema de computação para fornecer funcionalidade de rede lógica para redes de computador, o método compreende: recebimento de informações de configuração que especificam uma quantidade de múltiplos nós de computação arranjados através de uma topologia de rede especificada para uma primeira rede de computador virtual, a topologia de rede incluindo múltiplas sub-redes em que cada uma in-

clui um subconjunto de múltiplos nós de computação distinto e são cada um associados com um dos múltiplos dispositivos roteadores físicos especificados, múltiplos nós de computação cada um compreendendo um distinto endereço de rede virtual para a primeira rede de computador virtual; implementando automaticamente a primeira rede de computador virtual de acordo com as informações de configuração recebidas pela sobreposição da primeira rede de computador virtual em uma rede de substrato distinta físico que interconecta os múltiplos nós de computação, os múltiplos nós de computação estando cada um localizado em um local distinto na rede de substrato, a implementação da primeira rede de computador virtual incluindo emulação da funcionalidade dos múltiplos dispositivos roteadores físicos sem fornecer fisicamente os múltiplos dispositivos roteadores físicos; e sob controle de um módulo gerenciador de comunicação executando o sistema de computação, automaticamente gerenciando múltiplas comunicações entre os múltiplos nós de computação de tal maneira como para emular a funcionalidade de um ou mais dos múltiplos dispositivos roteadores físicos, cada uma das comunicações sendo a partir de um dos múltiplos nós de computação de envio e sendo designado para um destino remoto um dos múltiplos nós de computação e incluindo o endereço de rede virtual para o nó de destino e incluindo um endereço de hardware virtual para o dispositivo roteador físico associado com uma das sub-redes da qual o nó de computação de envio pertence, o gerenciamento de cada uma das comunicações incluindo: identificação do nó de computação de destino para a comunicação com base no endereço de rede virtual incluído na comunicação, e encaminhamento da comunicação através da rede de substrato para o local do nó de computação de destino na rede de substrato; identificando o dispositivo roteador físico para a comunicação com base no endereço de hardware virtual incluído na comunicação, e modificando a comunicação para incluir informações que o dispositivo roteador físico identificado teria adicionado para a comunicação se o dispositivo roteador físico identificado tivesse encaminhado a comunicação a partir do nó de computação de envio para o nó de computação de destino com base em a topologia de rede; e fornecendo a comunicação modificada para o nó de computação de destino.

Cláusula 2. O método da cláusula 1 em que o módulo gerenciador de comunicação está associado com uma porção dos múltiplos nós de computação de forma que as múltiplas comunicações gerenciadas pelo módulo gerenciador de comunicação são comunicações enviadas pelos nós de computação da porção associada, em que o módulo gerenciador de comunicação gerencia ainda uma primeira sub-rede física da rede de substrato físico da qual os nodos de comunicação da porção associada pertencem, em que os nós de computação da porção associada pertencem a uma ou mais das sub-redes da primeira rede de computador virtual de forma que o módulo gerenciador de comunicação emula as funcionalidades para dois ou mais dos dispositivos roteadores físicos associados com duas ou mais sub-redes, e em que a modificação de cada uma das comunicações gerenciadas inclui mo-

dificação do endereço de hardware virtual incluído na comunicação para ser um endereço de hardware virtual que corresponde para o nó de computação de destino para a comunicação.

Cláusula 3. O método da cláusula 2 em que o módulo gerenciador de comunicação
5 é um dos múltiplos módulos gerenciadores de comunicação de um sistema de gerenciador de rede de sobreposição em que cada um gerencia uma porção associada dos múltiplos nós de computação, e em que a primeira rede de computador virtual é uma das múltiplas redes de computador virtuais que são fornecidas pelo sistema de gerenciamento de rede de sobreposição em nome de múltiplos usuários remotos, as múltiplas redes de computador virtuais cada uma sendo sobreposta na rede de substrato físico e cada uma incluindo múltiplos
10 de uma pluralidade de nós de computação fornecidos pelo sistema de gerenciamento de rede de sobreposição.

Cláusula 4. Um método implementado por computador para fornecimento da funcionalidade de rede lógica para redes de computador, o método compreende: recebimento de
15 informações de configuração para uma primeira rede de computador virtual que inclui múltiplos nós de computação arranjados através de uma topologia de rede especificada, múltiplos nós de computação cada um compreendendo um distinto endereço de rede virtual para a primeira rede de computador virtual, as informações de configuração indicando que um ou mais dispositivos de rede especificados que interconectam os múltiplos nós de computação
20 e que são associados com uma ou mais dos múltiplos nós de computação; e sob controle de um ou mais sistemas de computação configurados, automaticamente fornecendo a primeira rede de computador virtual de acordo com as informações de configuração recebidas pela sobreposição da primeira rede de computador virtual em uma rede de substrato distinta e pela emulação da topologia de rede especificada para múltiplas comunicações entre os múltiplos
25 nós de computação sem implantar fisicamente a topologia de rede especificada, as múltiplas comunicações são enviadas a partir de um dos múltiplos nós de computação de envio para um ou mais dos múltiplos nós de computação que são tanto um nó de computação de destino para a comunicação, cada comunicação incluindo o endereço de rede virtual correspondente a um ou mais nós de computação de destino e incluindo um endereço de
30 hardware virtual associado com um dos dispositivos de rede especificados que é um destino intermediário inicial da comunicação, a emulação da topologia de rede especificada incluindo, para cada uma das múltiplas comunicações; identificando um ou mais nós de computação de destino com base no endereço de rede virtual incluído na comunicação; e para cada um ou mais nós de computação de destino, encaminhamento da comunicação através da
35 rede de substrato para um local do nó de computação de destino; modificando uma comunicação para incluir informações correspondendo a comunicação que foi encaminhada por meio da topologia de rede especificada do nó de computação de envio para o nó de compu-

tação de destino pelo dispositivo de rede especificado cujo endereço associado de hardware virtual é incluído na comunicação; e fornecendo a comunicação modificada para o nó de computação de destino.

5 Cláusula 5. O método da cláusula 4 em que a topologia de rede especificada inclui múltiplas sub-redes lógicas da primeira rede de computador virtual que cada inclui um sub-conjunto de múltiplos nós de computação distinto, e em que um ou mais dispositivos de rede especificados incluem múltiplos dispositivos roteadores lógicos que são associados com uma das múltiplas sub-redes lógicas.

10 Cláusula 6. O método da cláusula 5 em que a emulação da topologia de rede especificada é realizada para comunicações que são enviadas entre nós de computação de envio e nós de computação de destino que são parte de sub-redes lógicas distintas.

15 Cláusula 7. O método da cláusula 4 em que a emulação da topologia de rede especificada inclui representando logicamente um ou mais dispositivos de rede física com um ou mais dispositivos de rede lógica, e emulação da funcionalidade de um ou mais dispositivos de rede lógica no encaminhamento das múltiplas comunicações por meio da topologia de rede especificada.

20 Cláusula 8. O método da cláusula 4 em que a modificação de cada de uma ou mais da múltiplas comunicações para incluir informações correspondendo a comunicação sendo encaminhada por um dispositivo de rede especificado para um nó de computação de destino inclui a substituição na comunicação do endereço de hardware virtual associado com o dispositivo de rede especificado com um endereço de hardware virtual distinto correspondente ao nó de computação de destino para a comunicação e inclui modificação de um cabeçalho da comunicação para refletir a ocorrência de um salto de encaminhamento de rede.

25 Cláusula 9. O método da cláusula 8 em que o fornecimento da primeira rede de computador virtual de acordo com as informações de configuração recebidas inclui: rastreamento de endereços de hardware virtuais para a primeira rede de computador virtual para cada um dos múltiplos nós de computação e cada um dos dispositivos de rede especificados; rastreamento para cada um dos múltiplos nós de computação de um local dos dispositivos de rede especificados na topologia de rede especificada; e para cada uma das múltiplas solicitações dos múltiplos nós de computação para informações que seriam fornecidas pelo dispositivo de rede especificado local de um nó de computação de solicitação que fez a solicitação, respondendo a solicitação de maneira a emular a funcionalidade do dispositivo de rede especificado local de um nó de computação de solicitação que fez a solicitação.

35 Cláusula 10. O método da cláusula 4 em que um ou mais dispositivos de rede especificados incluem múltiplos dispositivos de rede lógica, em que o encaminhamento de cada uma ou mais das múltiplas comunicações por meio da topologia de rede especificada inclui dois ou mais dos múltiplos dispositivos de rede lógica que encaminhariam a comuni-

cação, e em que a modificação de cada uma ou mais comunicações para incluir informações correspondendo a comunicação sendo encaminhada por um dispositivo de rede especificado para um nó de computação de destino inclui a substituição na comunicação do endereço de hardware virtual associado com o dispositivo de rede especificado com um endereço de hardware virtual distinto correspondente ao nó de computação de destino para a comunicação e inclui modificação de um cabeçalho da comunicação para refletir a ocorrência de uma quantidade de múltiplos saltos de encaminhamento de rede correspondente a uma quantidade de dois ou mais dispositivos de rede lógica que encaminhariam a comunicação por meio da topologia de rede especificada.

Cláusula 11. O método da cláusula 4 em que o fornecimento da primeira rede de computador virtual inclui ainda gerenciamento de outras múltiplas comunicações entre os múltiplos nós de computação, cada uma das outras múltiplas comunicações sendo enviada a partir de um dos múltiplos nós de computação de envio para uma destinação de um dos múltiplos nós de computação e incluindo o endereço de rede virtual para o nó de computação de destino e incluindo um endereço de hardware virtual associado com o nó de computação de destino, o gerenciamento de cada uma das outras múltiplas comunicações incluindo encaminhamento da outra comunicação através da rede de substrato para o local do nó de computação de destino e fornecendo a outra comunicação para o nó de computação de destino sem modificação para incluir informações correspondendo a comunicação que foi encaminhada por meio da topologia de rede especificada.

Cláusula 12. O método da cláusula 11 em que a topologia de rede especificada inclui múltiplas sub-redes lógicas da primeira rede de computador virtual que cada inclui um subconjunto de múltiplos nós de computação distinto, em que as múltiplas comunicações estão cada uma entre um nó de computação de envio e ao menos um nó de computação de destino que são parte de duas sub-redes lógicas distintas, e em que as outras múltiplas comunicações estão entre um nó de computação de envio e nó de computação de destino que são parte de uma única sub-rede lógica.

Cláusula 13. O método da cláusula 4 em que a rede de substrato é uma rede física que inclui múltiplos dispositivos interconectados de rede física, em que cada um dos múltiplos nós de computação possui um endereço de rede de substrato associado que corresponde a um local do nó de computação no endereço de rede de substrato e que é diferente do endereço de rede virtual para o modo de computação, e em que o encaminhamento de cada uma das múltiplas comunicações através da rede de substrato para o local do nó de computação de destino inclui envio da comunicação para um dos dispositivos de rede física da rede de substrato para que um dispositivo de rede física use o endereço de rede de substrato para o nó de computação de destino para rotear a comunicação para o local do nó de computação de destino na rede de substrato.

Cláusula 14. O método da cláusula 4 em que as informações de configuração para a primeira rede de computador virtual são recebidas a partir de um usuário em nome de quem a primeira rede de computador virtual é fornecida.

Cláusula 15. O método da cláusula 14 em que os múltiplos nós de computação são
5 cada um uma máquina virtual hospedadas em um dos múltiplos sistemas de computação de host físicos, em que o método é realizado por um módulo gerenciador de comunicação de máquina virtual que executa um ou mais dos sistemas de computação físicos para gerenciar as comunicações pelos nós de computação de máquina virtual hospedados no sistema de computação físico, e em que o módulo gerenciador de comunicação de máquina virtual é
10 um dos múltiplos módulos gerenciadores de comunicação de um sistema de gerenciador de rede de sobreposição que fornecem múltiplas redes de computador virtuais para múltiplos usuários que são cada sobreposição na rede de substrato.

Cláusula 16. Um meio de leitura legível para um computador cujo conteúdo permite um sistema de computação para fornecer funcionalidade de rede lógica para redes de com-
15 putador, realizando um método compreendendo: obtendo informações com relação à topologia de rede especificada lógica para uma primeira rede de computador virtual que inclui múltiplos nós de computação, a primeira rede de computador virtual sendo sobreposta em uma rede de substrato distinta que interconecta os múltiplos nós de computação, cada um dos múltiplos nós de computação com um associado endereço de rede virtual para a primei-
20 ra rede de computador virtual e tendo um distinto e endereço de rede de substrato associado que corresponde a um local do nó de computação na rede de substrato; e automaticamente gerenciando múltiplas comunicações enviadas entre os múltiplos nós de computação de forma a implementar virtualmente a topologia de rede especificada lógica para a primeira rede de computador virtual, as múltiplas comunicações são enviadas a partir de um primeiro
25 nó de computação e incluindo o endereço de rede virtual para um segundo nó de computação de destino, o gerenciamento de cada uma das múltiplas comunicações incluindo: usando o endereço de rede de substrato associado com o segundo nó de computação para encaminhar a comunicação através da rede de substrato para o local do segundo nó de com-
putação na rede de substrato; e antes de fornecer a comunicação para o segundo modo de
30 computação, modificando a comunicação para incluir informações emulando o encaminhamento da comunicação a partir do primeiro nó de computação para o segundo nó de computação através da topologia de rede lógica.

Cláusula 17. O meio legível por computador da cláusula 16 em que as informações obtidas são informações de configuração especificada por um usuário em nome de quem a
35 primeira rede de computador virtual é fornecida, em que a topologia de rede especificada lógica indica um ou mais dispositivos de rede especificados que interconectam os múltiplos nós de computação e que são associados com uma ou mais dos múltiplos nós de computa-

ção, e em que a implementação virtual da topologia de rede especificada lógica para a primeira rede de computador virtual inclui emulação da funcionalidade que seria fornecida por um ou mais dispositivos de rede especificados if um ou mais dispositivos de rede especificados estivessem fisicamente implantados e usados para encaminhar as comunicações dos nós de computação de envio para nós de computação de destino .

Cláusula 18. O meio legível por computador da cláusula 17 em que cada um dos múltiplos nós de computação e cada um ou mais dispositivos de rede especificados possui ainda um endereço de hardware virtual distinto para a primeira rede de computador virtual, em que cada uma das múltiplas comunicações inclui ainda um endereço de hardware virtual para um pretendido destino inicial da comunicação, e em que a modificação de cada uma das múltiplas comunicações inclui ainda determinação de um dos dispositivos de rede especificados para o qual endereço de hardware virtual incluído na comunicação corresponde e adiciona informações para a comunicação que teriam sido adicionadas por um determinado dispositivo de rede durante o encaminhamento da comunicação.

Cláusula 19. O meio legível por computador da cláusula 16 em que o meio legível por computador é ao menos um de memória do sistema de computação no qual os conteúdos são armazenados e um meio de transmissão de dados que armazena um sinal gerado contendo os conteúdos, e em que os conteúdos são instruções que quando executadas fazem com que o sistema de computação realize o método.

Cláusula 20. Um sistema estruturado para fornecer funcionalidade de rede lógica para redes de computador, compreendendo: um ou mais processadores de um ou mais sistemas de computação; e um ou mais Módulos de Gerenciador de Comunicações que são cada um configurados para, quando executados por ao menos um ou mais processadores fornecem a funcionalidade de rede lógica para uma primeira rede de computador virtual que é implementada usando uma ou mais redes intermediárias, uma ou mais redes intermediárias incluindo múltiplos dispositivos de rede física que interconectam fisicamente múltiplos nós de computação da primeira rede de computador virtual, cada um dos múltiplos nós de computação possuindo um associado primeiro endereço de rede virtual para a primeira rede de computador virtual e tendo um distinto e associado segundo endereço de rede que corresponde a uma ou mais redes intermediárias, o fornecimento da funcionalidade de rede lógica incluindo: obtendo informações de configuração com relação a um ou mais dispositivos de rede lógica da primeira rede de computador virtual que interconecta logicamente os múltiplos nós de computação e que são distintos dos dispositivos de rede física de uma ou mais redes intermediárias; e automaticamente gerenciando múltiplas comunicações enviadas entre os múltiplos nós de computação de forma a implementar a funcionalidade dos dispositivos de rede lógica especificados, o gerenciamento de cada uma das múltiplas comunicações incluindo usando o segundo endereço de rede de um nó de computação de destino

para a comunicação para encaminhar a comunicação por meio de ao menos um dos dispositivos de rede física de uma ou mais redes intermediárias para o nó de computação de destino, o gerenciamento de cada uma das múltiplas comunicações inclui ainda a modificação da comunicação para incluir informações emulando o encaminhamento da comunicação por um ou mais dos dispositivos de rede lógica especificados que interconectam logicamente os múltiplos nós de computação .

Cláusula 21 , O sistema da cláusula 20 em que uma ou mais redes intermediárias são uma rede de substrato na qual a primeira rede de computador virtual é sobreposta, em que o associado segundo endereço de rede dos múltiplos nós de computação são endereços de redes de substrato que correspondem aos locais dos múltiplos nós de computação na rede de substrato, em que um ou mais dos dispositivos de rede lógica especificados incluem múltiplos dispositivos de rede lógica especificados que são cada um associados com um subconjunto de múltiplos nós de computação, em que as informações de configuração especificam ainda uma topologia lógica de rede da primeira rede de computador virtual que inclui os múltiplos dispositivos de rede lógica, e em que o fornecimento da funcionalidade de rede lógica inclui emulação da funcionalidade que seria fornecida pelos especificados múltiplos dispositivos de rede lógica se os múltiplos dispositivos de rede lógica especificados estivessem fisicamente implantados e usados para encaminhar as comunicações dos nós de computação de envio para nós de computação de destino .

Cláusula 22. O sistema da cláusula 21 em que cada um dos múltiplos nós de computação e cada um ou mais dispositivos de rede lógica especificados possui ainda um endereço de hardware virtual distinto para a primeira rede de computador virtual, em que cada uma das múltiplas comunicações inclui ainda um endereço de hardware virtual para um pretendido destino inicial da comunicação, em que a modificação de cada uma das múltiplas comunicações inclui ainda a determinação de dispositivos de rede lógica especificados para a o qual o endereço de hardware virtual incluído na comunicação corresponde e adiciona informações para a comunicação que teriam sido adicionadas por um determinado dispositivo de rede lógica mediante encaminhamento da comunicação, e em que um ou mais módulos gerenciadores de comunicação são parte de um gerenciador de rede de sobreposição que fornece a primeira rede de computador virtual e que gerencia automaticamente o fornecimento da funcionalidade de rede lógica para a primeira rede de computador virtual.

Cláusula 23. O sistema da cláusula 20 em que um ou mais módulos gerenciadores de comunicação incluem múltiplos módulos gerenciadores de comunicação em que cada um gerencia as comunicações enviadas de um grupo associado de um subconjunto de múltiplos nós de computação, em que um ou mais sistemas de computação incluem múltiplos sistemas de computação que executam os múltiplos módulos gerenciadores de comunicação, e em que o sistema compreende ainda os múltiplos sistemas de computação, em que cada

um ou mais dos múltiplos sistemas de computação hospeda múltiplas máquinas virtuais em que cada uma fornece um dos múltiplos nós de computação e executando um dos módulos gerenciadores de comunicação como parte de a máquina virtual manager module for o sistema de computação, de forma que os nós de computação fornecidos pelas máquinas virtuais hospedadas do sistema de computação são os grupos associados dos nós de computação par executar o módulo gerenciador de comunicação.

Cláusula 24. O sistema da cláusula 20 em que um ou mais módulos gerenciadores de comunicação inclui instruções de software para execução por ao menos um dos processadores de um ou mais sistemas de computação.

Cláusula 25. O sistema da cláusula 20 em que um ou mais módulos gerenciadores de comunicação cada um consiste de meios para fornecimento da funcionalidade de rede lógica para a primeira rede de computador virtual por: obtenção de informações de configuração com relação a um ou mais dispositivos de rede lógica pós-soviética da primeira rede de computador virtual que interconectam logicamente os múltiplos nós de computação e que são distintos dos dispositivos de rede física de uma ou mais redes intermediárias; e automaticamente gerenciando múltiplas comunicações enviadas entre os múltiplos nós de computação de forma a implementar a funcionalidade dos dispositivos de rede lógica especificados, o gerenciamento de cada uma das múltiplas comunicações incluindo usando o segundo endereço de rede de um nó de computação de destino para a comunicação para encaminhar a comunicação por meio de ao menos um dos dispositivos de rede física de uma ou mais redes intermediárias para o nó de computação de destino, o gerenciamento de cada uma das múltiplas comunicações inclui ainda a modificação de a comunicação para incluir informações emulando o encaminhamento da comunicação por um ou mais dos dispositivos de rede lógica especificados que interconectam logicamente os múltiplos nós de computação,

A partir do seguinte será apreciado que, apesar de modalidades específicas foram aqui descritas para fins de ilustração, várias modificações podem ser feitas sem se desviar do espírito e escopo da invenção. De acordo com isso, a invenção não é limitada, exceto conforme pelas reivindicações anexadas e os elementos citados no presente documento.

Além disso, enquanto certos aspectos da invenção são apresentados abaixo em certas formas de reivindicações, os inventores contemplam os vários aspectos da invenção em qualquer forma de reivindicação disponível. Por exemplo, enquanto apenas alguns aspectos da invenção atualmente são citados como sendo incorporados em um meio de leitura legível para um computador, outros aspectos podem ser incorporados da mesma maneira.

REIVINDICAÇÕES

1. Método implementado por computador para fornecer funcionalidade de rede lógica para redes de computador, o método **CARACTERIZADO** pelo fato de que compreende:

5 receber informações de configuração para uma primeira rede de computador virtual que inclui múltiplos nós de computação (205, 255) arranjos através de uma topologia de rede especificada, os múltiplos nós de computação tendo cada um endereço de rede virtual distinto para a primeira rede de computador virtual, as informações de configuração indicando que um ou mais dispositivos de rede especificados (270) que interconectam os múltiplos nós de computação e que cada um é associado com uma ou mais dos múltiplos nós de
10 computação; e

sob controle de um ou mais sistemas de computação configurados, fornecer automaticamente a primeira rede de computador virtual de acordo com as informações de configuração recebidas pela sobreposição da primeira rede de computador virtual em uma rede de substrato distinta e pela emulação da topologia de rede especificada para múltiplas comunicações entre os múltiplos nós de computação sem implantar fisicamente a topologia de
15 rede especificada, as múltiplas comunicações são, cada uma, enviada a partir de um nó de envio dentre os múltiplos nós de computação (205a) para um respectivo nó de computação de destino (255b) dos múltiplos nós de computação, cada comunicação incluindo um endereço de rede virtual correspondente ao nó de computação de destino e incluindo um endereço de hardware virtual de um dos dispositivos de rede especificados (270a, 270b) que é um
20 destino intermediário inicial da comunicação, a emulação da topologia de rede especificada incluindo, para cada uma das múltiplas comunicações:

identificar o nó de computação de destino com base no endereço de rede virtual incluído na comunicação;

25 encaminhar a comunicação através da rede de substrato para um local do nó de computação de destino;

detectar que a comunicação inclui o endereço de hardware virtual do dispositivo de rede especificado;

modificar a comunicação para emular o dispositivo de rede especificado em que a
30 modificação indica que a comunicação que foi encaminhada por meio da topologia de rede especificada do nó de computação de envio para o nó de computação de destino pelo dispositivo de rede especificado; e

fornecer a comunicação modificada para o nó de computação de destino.

2. Método, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que a
35 topologia de rede especificada inclui múltiplas sub-redes lógicas da primeira rede de computador virtual que cada uma inclui um subconjunto de múltiplos nós de computação distinto, e em que o um ou mais dispositivos de rede especificados incluem múltiplos dispositivos rote-

adores lógicos que são, cada um, associado com uma das múltiplas sub-redes lógicas.

3. Método, de acordo com a reivindicação 2, **CARACTERIZADO** pelo fato de que a emulação da topologia de rede especificada é realizada para comunicações que são enviadas entre nós de computação de envio e nós de computação de destino que são parte de sub-redes lógicas distintas.

4. Método, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que a emulação da topologia de rede especificada inclui representar logicamente o um ou mais dispositivos de rede física com um ou mais dispositivos de rede lógica, e a emulação do dispositivo de rede lógica especificado é realizada durante o encaminhamento das múltiplas comunicações por meio da topologia de rede especificada.

5. Método, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que a modificação de cada uma de uma ou mais das múltiplas comunicações para emular o dispositivo de rede especificado inclui substituir na comunicação do endereço de hardware virtual associada com o dispositivo de rede especificado com um endereço de hardware virtual distinto correspondendo ao nó de computação de destino e modificar um cabeçalho da comunicação para refletir a ocorrência de um salto de encaminhamento de rede.

6. Método, de acordo com a reivindicação 5, **CARACTERIZADO** pelo fato de que o fornecimento da primeira rede de computador virtual conforme as informações de configuração recebidas incluem:

rastrear endereços de hardware virtuais para a primeira rede de computador virtual para cada um dos múltiplos nós de computação e cada um dos dispositivos de rede especificados;

rastrear para cada um dos múltiplos nós de computação um dispositivo local dentre os dispositivos de rede especificados na topologia de rede especificada; e

para cada uma das múltiplas solicitações provenientes dos múltiplos nós de computação para informações que seriam fornecidas pelo dispositivo de rede local especificado do nó de computação de solicitação que fez a solicitação, responder a solicitação de maneira a emular funcionalidade do dispositivo de rede local especificado do nó de computação de solicitação que fez a solicitação.

7. Método, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que um ou mais dispositivos de rede especificados incluem múltiplos dispositivos de rede lógica, em que o encaminhamento de cada uma de uma ou mais das múltiplas comunicações por meio da topologia de rede especificada inclui emular dois ou mais dentre os múltiplos dispositivos de rede lógica que encaminhariam a comunicação, e em que a modificação de cada uma das múltiplas comunicações para emular o dispositivo de rede especificado inclui substituir na comunicação do endereço de hardware virtual associado com o dispositivo de rede especificado com um endereço de hardware virtual distinto correspondendo ao nó de

computação de destino e modificar um cabeçalho da comunicação para refletir a ocorrência de uma quantidade de múltiplos saltos de encaminhamento de rede correspondendo a uma quantidade de dois ou mais dispositivos de rede lógica que encaminhariam a comunicação por meio da topologia de rede especificada.

5 8. Método, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que o fornecimento da primeira rede de computador virtual inclui ainda gerenciar outras múltiplas comunicações entre os múltiplos nós de computação, as outras múltiplas comunicações na primeira rede de computação virtual, em que o gerenciamento de cada uma das outras múltiplas comunicações inclui encaminhar as outras comunicações através da rede de substrato
10 para o local de respectivos nós de computação de destino e fornecer a outra comunicação para os respectivos nós de computação de destino sem modificação para emular o um ou mais dispositivos de rede especificados.

 9. Método, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que a rede de substrato é uma rede física que inclui múltiplos dispositivos interconectados de rede
15 física, em que cada um dos múltiplos nós de computação possui um endereço de rede de substrato associado que corresponde a um local do nó de computação no endereço de rede de substrato e que é distinto do endereço de rede virtual para o nó de computação, e em que o encaminhamento de cada uma das múltiplas comunicações através da rede de substrato para o local do nó de computação de destino inclui enviar a comunicação para um dis-
20 positivo de rede física da rede de substrato para que o dispositivo de rede física use o endereço de rede de substrato para o nó de computação de destino para rotear a comunicação para o local do nó de computação de destino na rede de substrato.

 10. Método, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que as informações de configuração para a primeira rede de computador virtual são recebidas a
25 partir de um usuário em nome de quem a primeira rede de computador virtual é fornecida.

 11. Sistema estruturado para fornecer funcionalidade de rede lógica para redes de computador, **CARACTERIZADO** pelo fato de que compreende:

um ou mais processadores de um ou mais sistemas de computação; e

um ou mais módulos gerenciadores de comunicação (210, 260) que são, cada um,
30 configurados para, quando executados por ao menos um dentre o um ou mais processadores, fornecer funcionalidade de rede lógica para uma primeira rede de computador virtual que é implementado usando uma ou mais redes intermediárias (250), a uma ou mais redes intermediárias incluindo múltiplos dispositivos de rede física (125, 130) que interconectam fisicamente múltiplos nós de computação da primeira rede de computador virtual, cada um
35 dos múltiplos nós de computação (205, 255) possuindo um primeiro endereço de rede virtual associado para a primeira rede de computador virtual e tendo um segundo endereço de rede distinto associado na uma ou mais redes intermediárias, o fornecimento da funcionalidade

de rede lógica incluindo:

obter informações de configuração com relação a um ou mais dispositivos de rede lógica especificados (270a, 270b) da primeira rede de computador virtual que interconectam logicamente os múltiplos nós de computação e que são distintos dos dispositivos de rede física da uma ou mais redes intermediárias; e

automaticamente gerenciar múltiplas comunicações enviadas entre os múltiplos nós de computação de forma a emular funcionalidade dos dispositivos de rede lógica especificados, o gerenciamento de cada uma das múltiplas comunicações incluindo:

usar o segundo endereço de rede de um nó de computação de destino para a comunicação para encaminhar a comunicação por meio de ao menos um dentre os dispositivos de rede física da uma ou mais redes intermediárias para o nó de computação de destino; e

modificar a comunicação para incluir informações emulando o encaminhamento da comunicação por um ou mais dentre os dispositivos de rede lógica especificados que interconectam logicamente os múltiplos nós de computação.

12. Sistema, de acordo com a reivindicação 11, **CARACTERIZADO** pelo fato de que a uma ou mais redes intermediárias são uma rede de substrato na qual a primeira rede de computador virtual é sobreposta, em que os segundos endereços de rede associados dos múltiplos nós de computação são endereços de redes de substrato que correspondem a locais dos múltiplos nós de computação na rede de substrato, em que o um ou mais dispositivos de rede lógica especificados que são, cada um, associados com um subconjunto dos múltiplos nós de computação, em que as informações de configuração especificam ainda uma topologia lógica de rede da primeira rede de computador virtual que inclui os múltiplos dispositivos de rede lógica.

13. Sistema, de acordo com a reivindicação 12, **CARACTERIZADO** pelo fato de que cada um dos múltiplos nós de computação e cada um dos um ou mais dispositivos de rede lógica especificados ainda possuem um endereço de hardware virtual distinto para a primeira rede de computador virtual, em que cada uma das múltiplas comunicações inclui ainda um endereço de hardware virtual para um destino intermediário da comunicação, em que a modificação de cada uma das múltiplas comunicações inclui ainda:

determinar um dentre os dispositivos de rede lógica especificados que corresponde ao endereço de hardware virtual incluído na comunicação; e

adicionar informações para a comunicação que teriam sido adicionadas pelo dispositivo de rede lógica determinado mediante o encaminhamento da comunicação.

14. Sistema, de acordo com a reivindicação 11, **CARACTERIZADO** pelo fato de que o um ou mais módulos gerenciadores de comunicação incluem múltiplos módulos gerenciadores de comunicação em que cada um gerencia comunicações enviadas a partir de

um grupo associado de um subconjunto de múltiplos nós de computação, em que o um ou mais sistemas de computação incluem múltiplos sistemas de computação que executam os múltiplos módulos gerenciadores de comunicação, e em os múltiplos sistemas de computação são, cada um, configurado para hospedar múltiplas máquinas virtuais que fornecem, 5 cada uma, um dos múltiplos nós de computação e executar um dos módulos gerenciadores de comunicação como parte de um modulo de gerenciador de máquina virtual o sistema de computação, de forma que nós de computação fornecidos pelas máquinas virtuais hospedadas do sistema de computação formam um respectivo grupo associado de nós de computação gerenciados por um respectivo módulo gerenciador de comunicação executando no 10 sistema de computação.

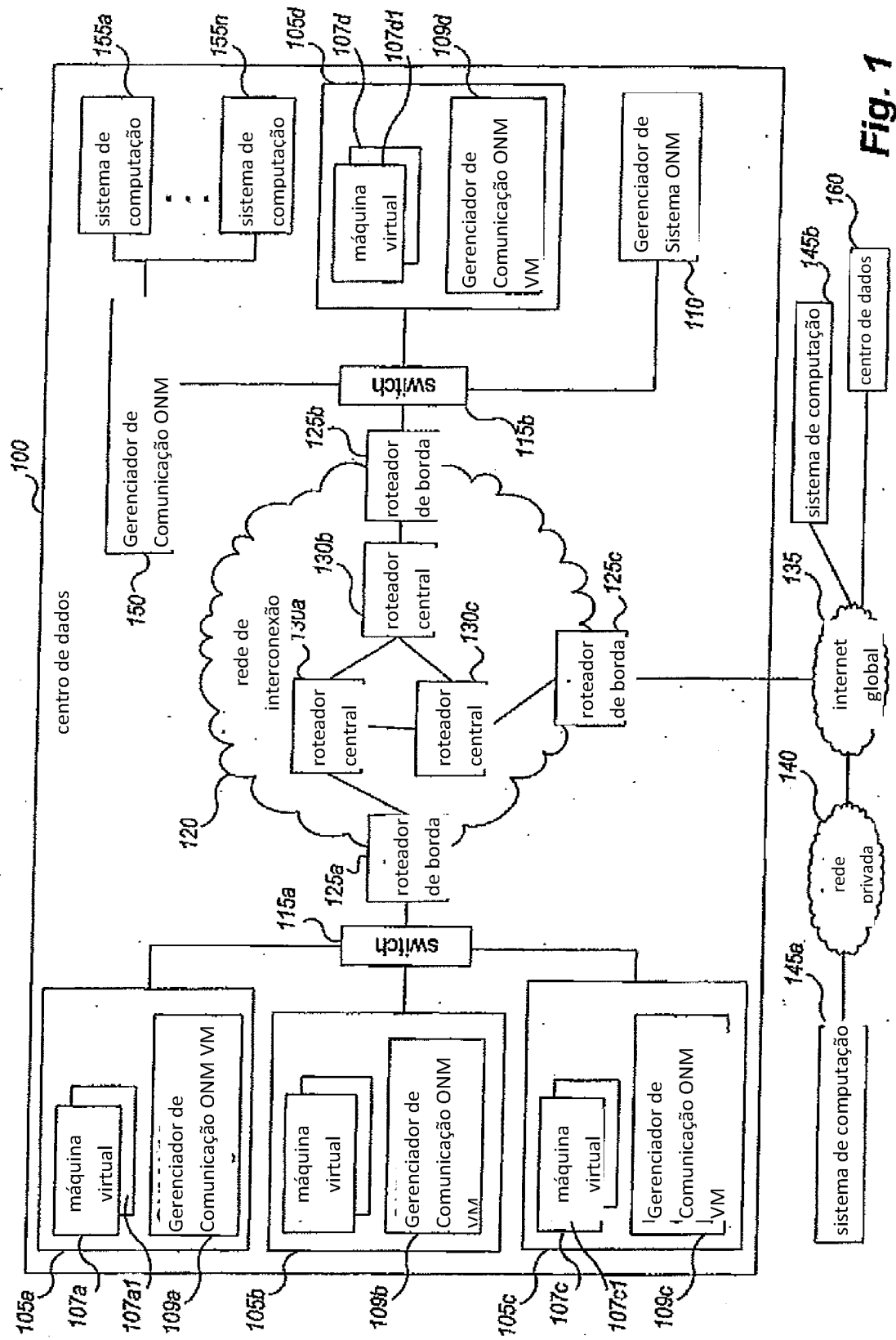


Fig. 1

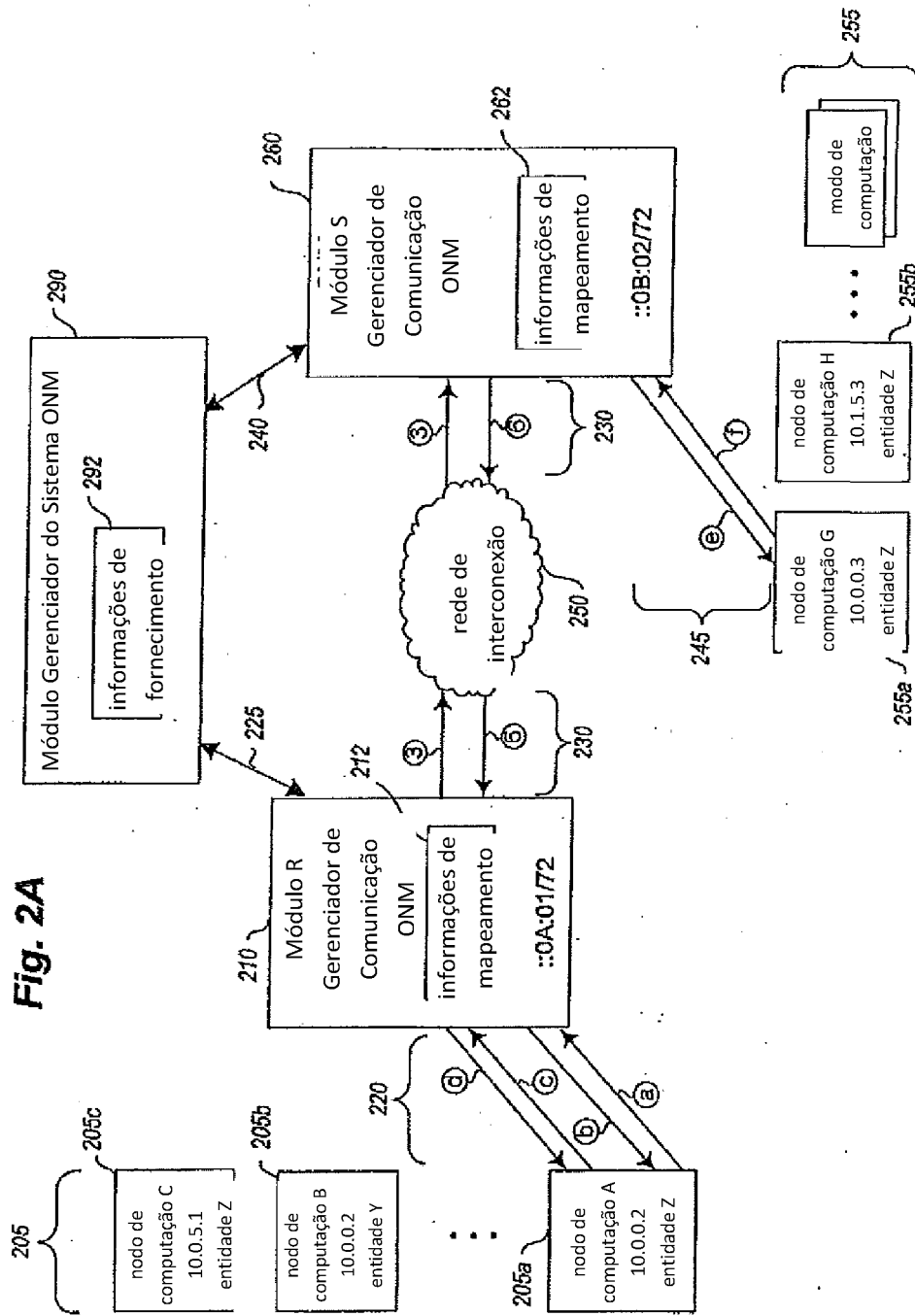


Fig. 2B

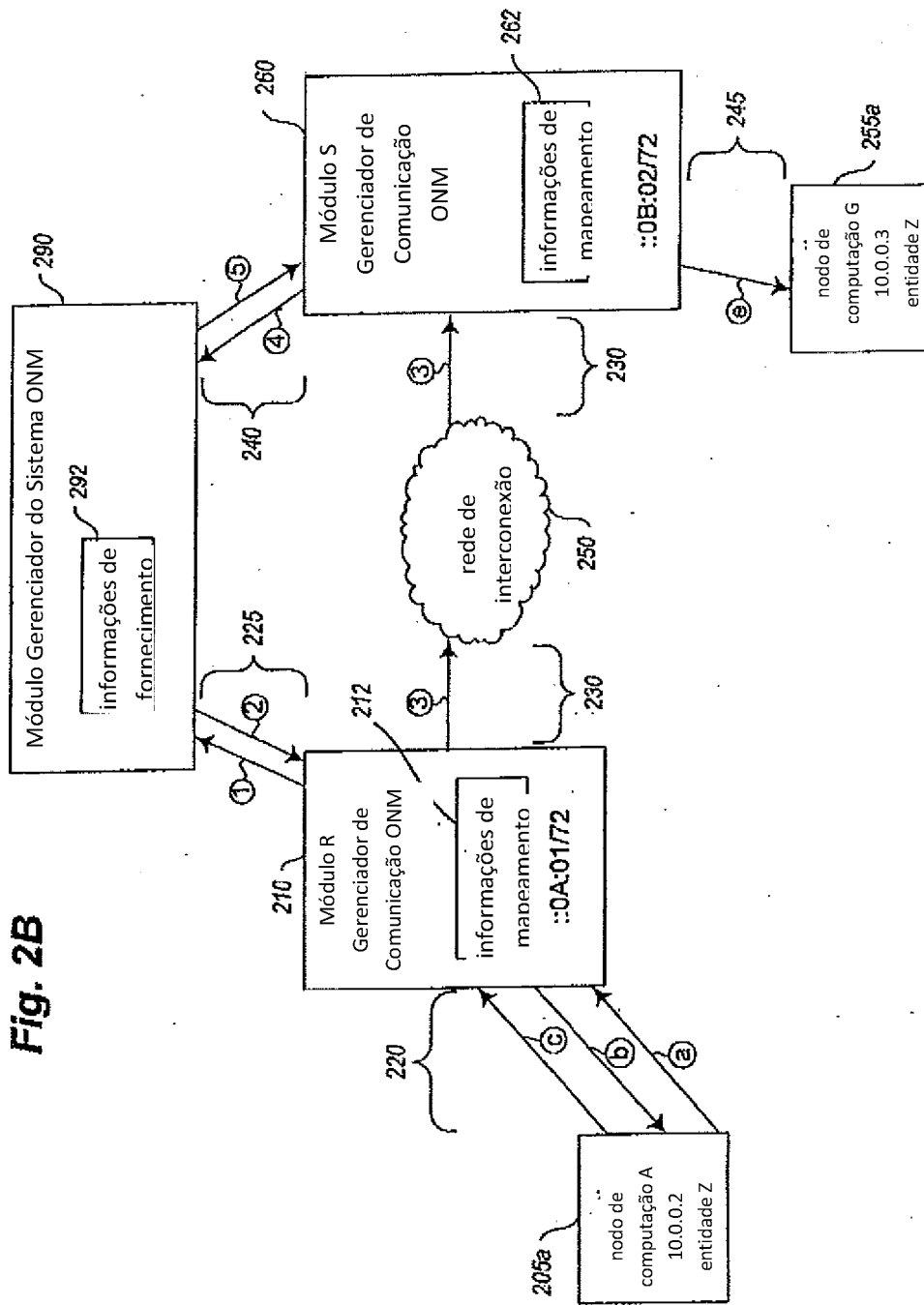
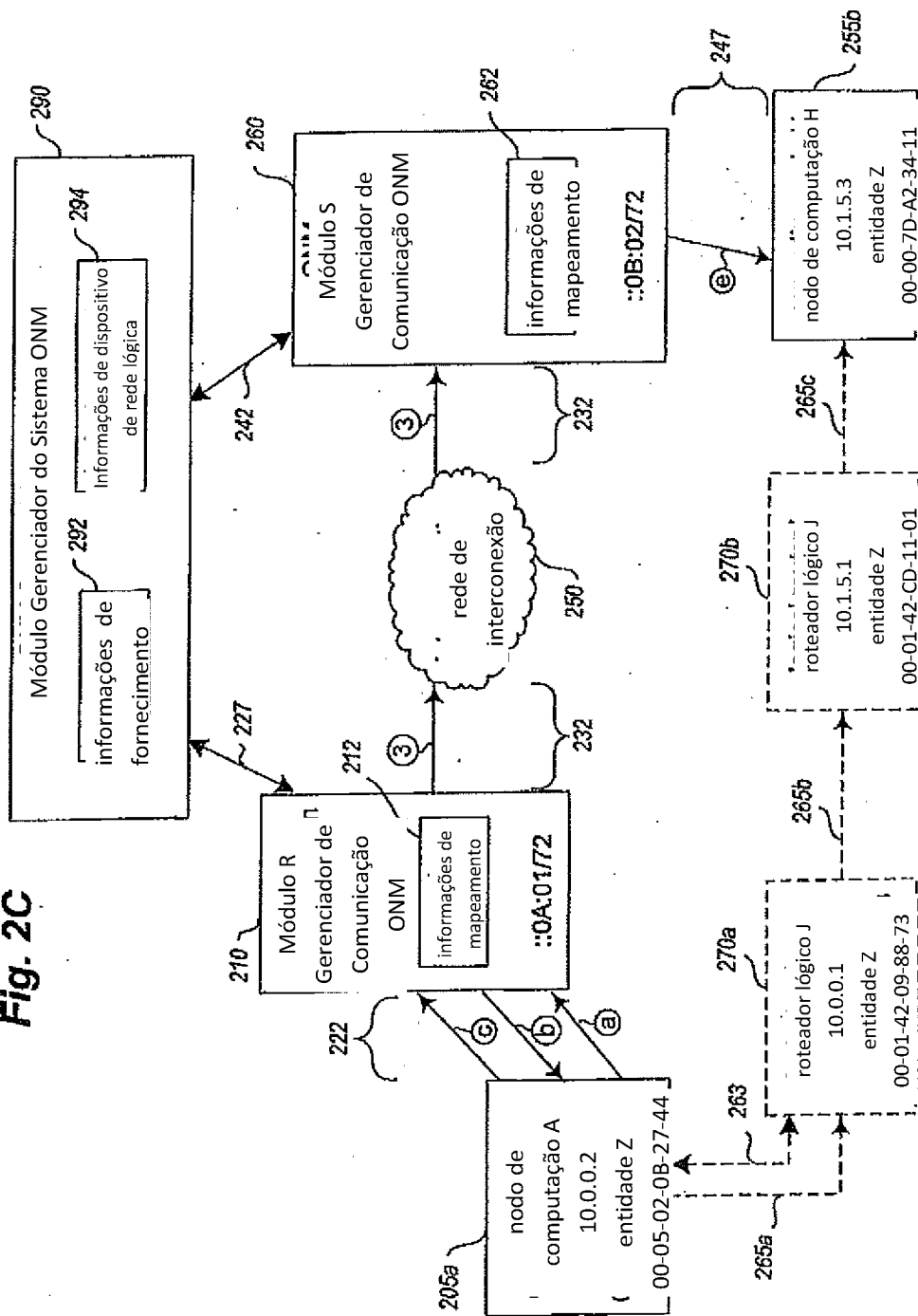


Fig. 2C

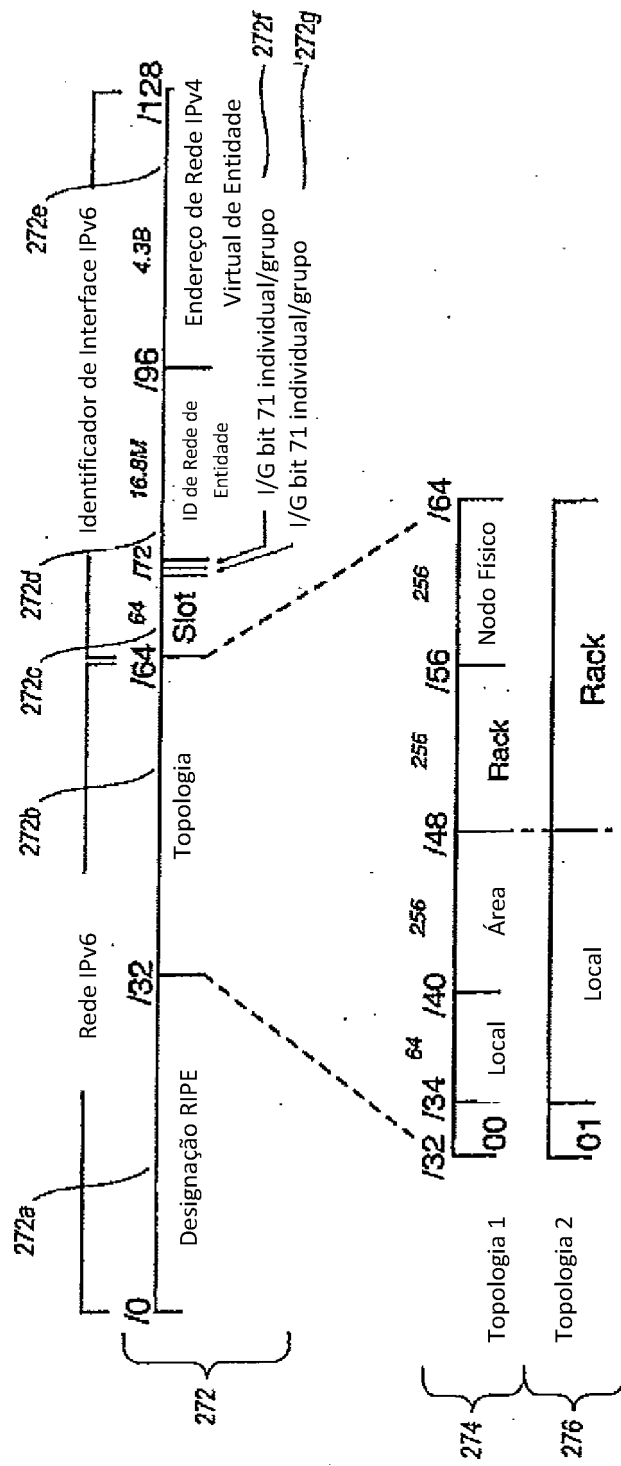


Fig. 2D

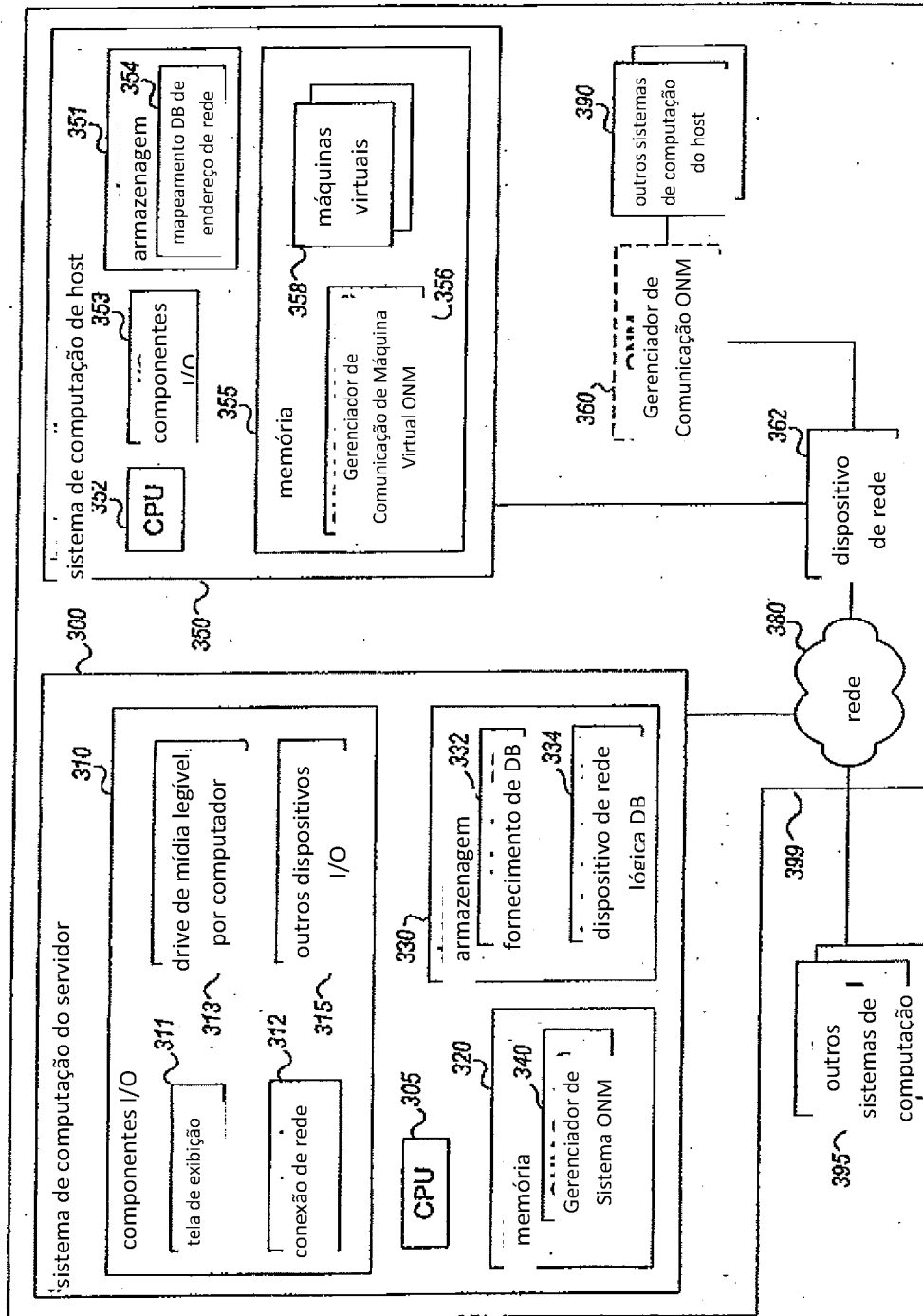


Fig. 3

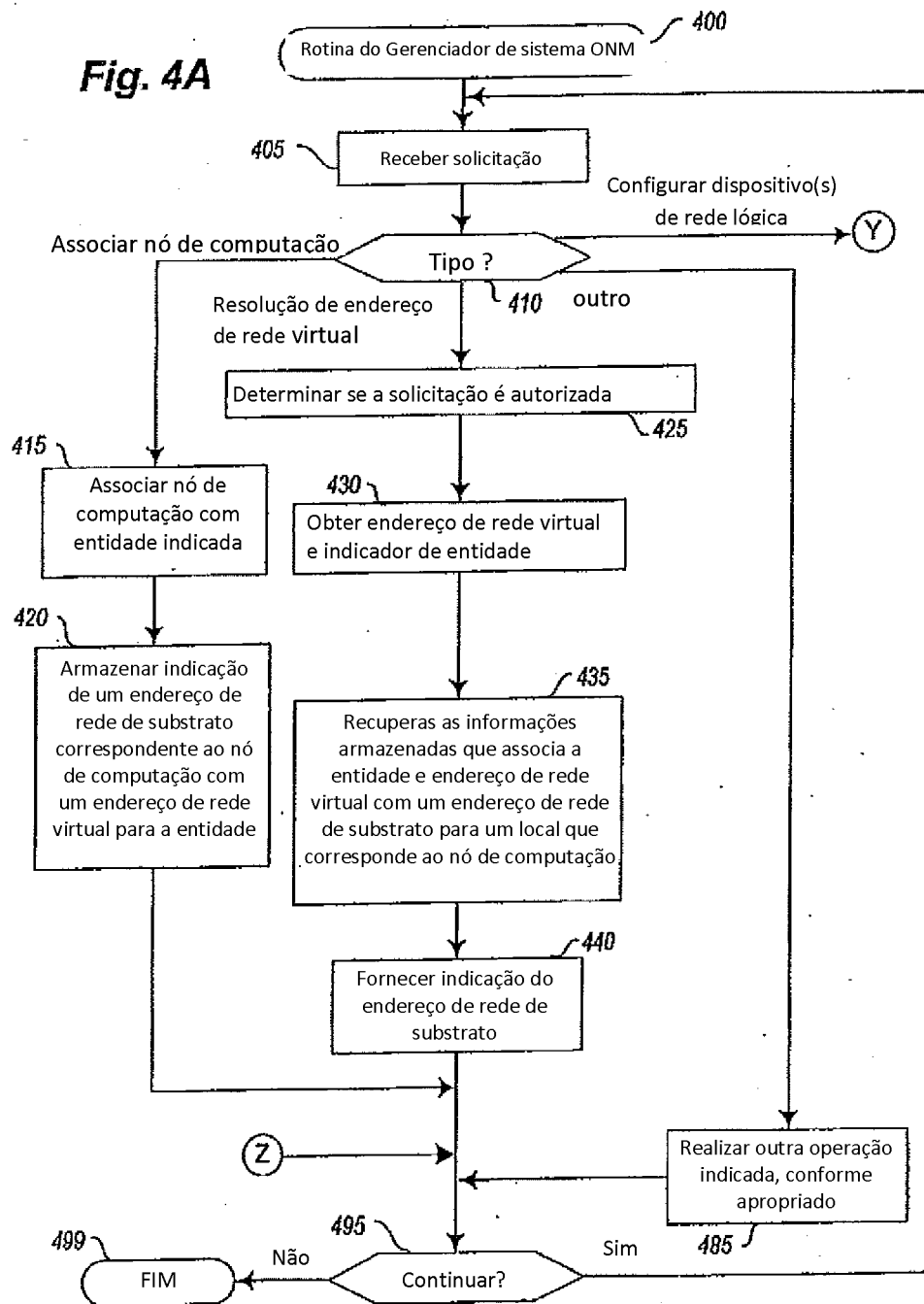
Fig. 4A

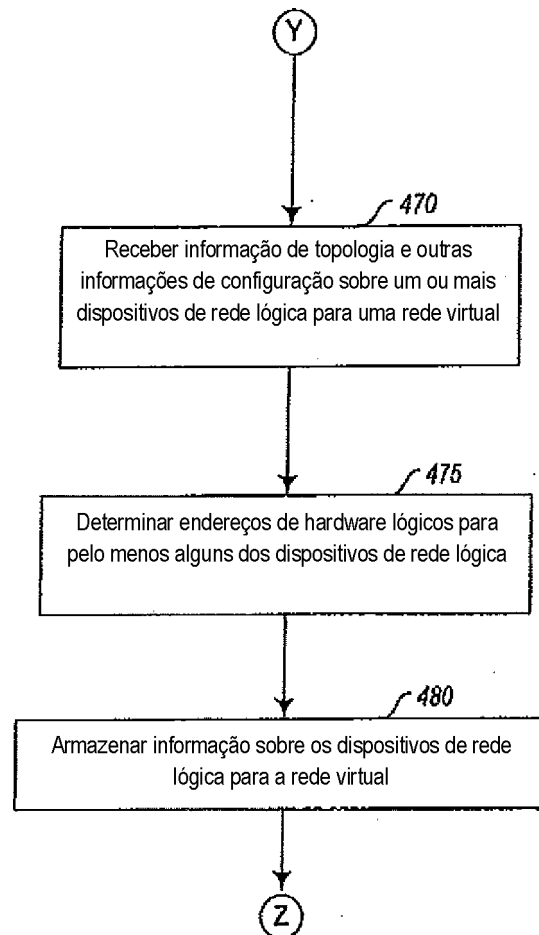
Fig. 4B

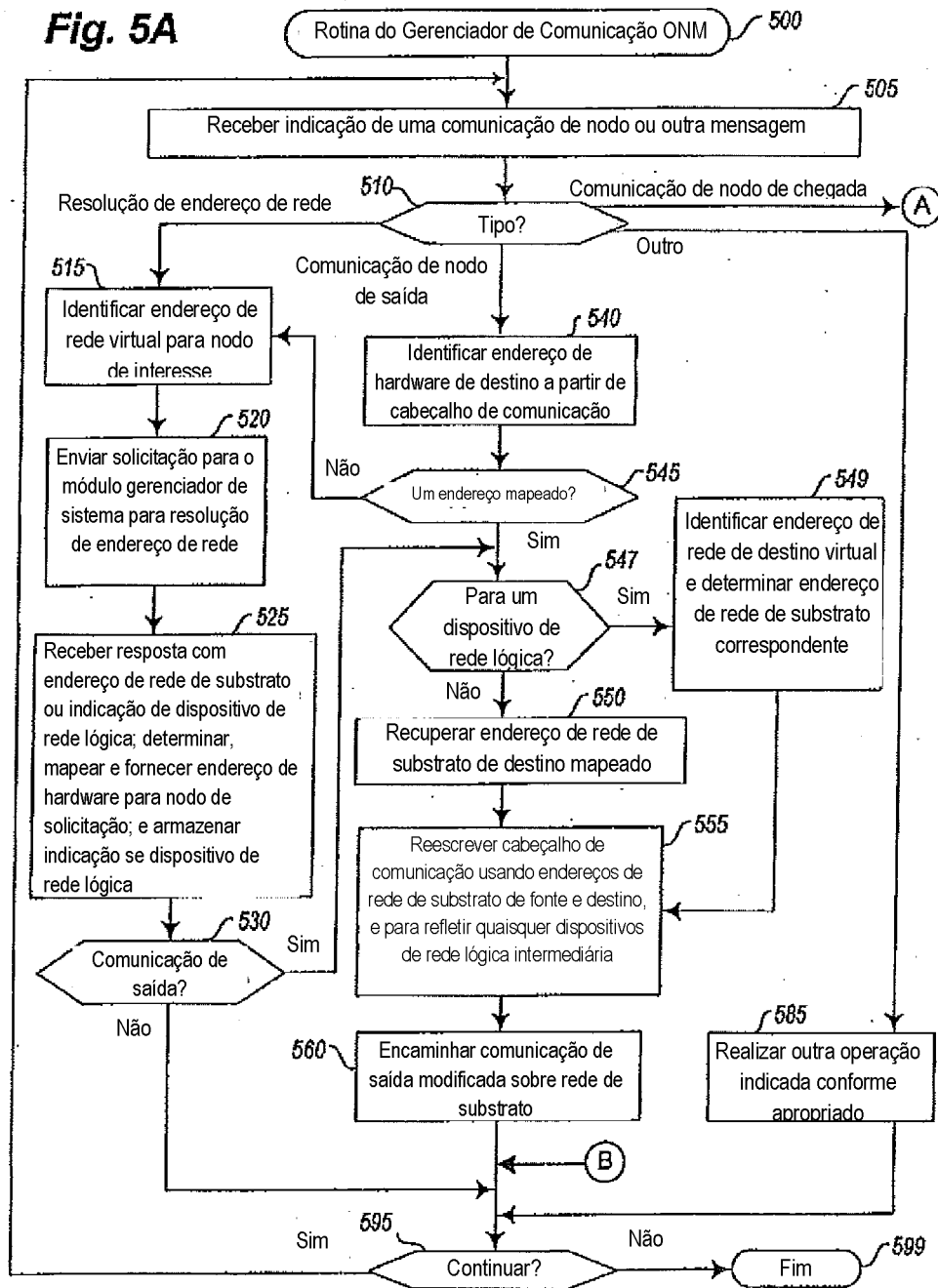
Fig. 5A

Fig. 5B