



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

(21) BR 112020005514-0 A2



(22) Data do Depósito: 12/10/2018

(43) Data da Publicação Nacional: 06/10/2020

(54) Título: GATEWAY DE INTERCONEXÃO DIRETA

(51) Int. Cl.: H04L 12/66.

(30) Prioridade Unionista: 12/10/2017 CA 2982147.

(71) Depositante(es): ROCKPORT NETWORKS INC..

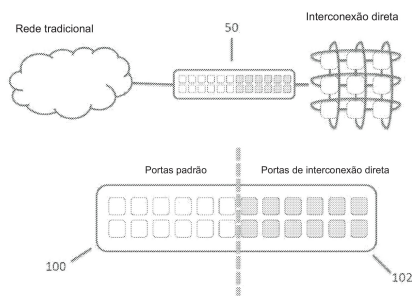
(72) Inventor(es): MATTHEW ROBERT WILLIAMS.

(86) Pedido PCT: PCT IB2018057945 de 12/10/2018

(87) Publicação PCT: WO 2019/073452 de 18/04/2019

(85) Data da Fase Nacional: 19/03/2020

(57) Resumo: Um dispositivo de gateway de rede dedicado que é capaz de ligar em ponte, alternar ou rotear o tráfego de rede entre a rede tradicional e as redes de interconexão direta, compreendendo: um primeiro conjunto de uma ou mais portas de rede tradicionais com um único link por porta, geralmente compreendendo um ou mais conectores SFP + QSFP e QSFP +, tais portas sendo conectadas a comutadores ou dispositivos que formam uma rede tradicional; e um segundo conjunto de uma ou mais portas de interconexão direta com um alto número de links por porta (dois ou mais) geralmente compreendendo um ou mais conectores MXC, MTP e MTO, tais portas sendo conectadas a uma rede de interconexão direta.



Relatório Descritivo da Patente de Invenção para **"GATEWAY DE INTERCONEXÃO DIRETA"**.

CAMPO DA INVENÇÃO

[0001] A presente invenção refere-se a um tráfego de E/O (Input/ Output – Entrada/ Saída) em uma rede de Computadores. Mais especificamente, a presente invenção refere-se a um dispositivo dedicado que liga, alterna ou roteia dados entre redes tradicionais e redes de interconexão direta.

ANTECEDENTES DA INVENÇÃO

[0002] As redes de computadores permitem a uma infinidade de nós rotear ou trocar dados entre si. Como resultado, as redes de computadores são capazes de suportar um imenso número de aplicativos e serviços, como o uso compartilhado de servidores de armazenamento, acesso à World Wide Web, uso de e-mail etc.

[0003] Os próprios nós geralmente podem ser caracterizados em três tipos, com base nas tarefas especializadas que realizam: nós de computação, como servidores tendo CPUs que realizam cálculos (mas que geralmente têm pouco ou nenhum espaço em disco local); nós de E / S que contêm o armazenamento secundário do sistema e fornecem serviços paralelos do sistema de arquivos; e nós de gateway (porta de entrada) que fornecem conectividade para servidores de dados externos e sistemas de armazenamento em massa. Alguns nós podem até servir a mais de uma função, como, por exemplo, lidar com funções de E / S e gateway.

[0004] E/ S para sistemas paralelos e distribuídos, no entanto, tornou-se uma grande preocupação para usuários e projetistas de sistemas de computador. Nesse sentido, enquanto as velocidades de CPUs aumentam a uma taxa exponencial praticamente todos os anos, infelizmente a velocidade dos dispositivos de E / S aumentou em um ritmo mais lento, geralmente devido ao fato de que eles podem ser

mais limitados pela velocidade de componentes mecânicos. O desempenho de E / S, uma medida do tráfego de dados de E / S entre os nós, é, portanto, frequentemente um fator limitante no desempenho da rede. De fato, a incompatibilidade de velocidade entre CPUs e E / S é acentuada em sistemas de computador paralelos e distribuídos, deixando a E / S como um gargalo que pode limitar severamente a escalabilidade. Esse é especialmente o caso quando a rede está envolvida com aplicativos comerciais que envolvem modelagem científica e multimídia, por exemplo, cada um dos quais com enormes requisitos de E / S.

[0005] Redes de interconexão direta, como as divulgadas na publicação do pedido de patente PCT no. WO 2015/027320 A1 (que descreve uma nova topologia de interconexão de toro ou radix superior para conectar nós de rede de maneira semelhante a uma malha em sistemas de computador paralelos), geralmente restringe o tráfego aos nós que fazem parte da interconexão direta. Embora o novo sistema e arquitetura divulgados na Publicação de Pedido de Patente PCT N° WO 2015/027320 A1 sejam particularmente benéficos e práticos para implantação comercial em data centers e data centers em nuvem, a maioria dos data centers em operação atualmente ainda se baseia, infelizmente, em uma arquitetura herdada de três camadas de um sistema tradicional, arquitetura de árvore ou arquitetura centrada no servidor DCell, entre outras. Com os data centers baseados nessas arquiteturas, infelizmente, é indesejável ou impossível que eles ingressem em uma interconexão direta e, portanto, são incapazes de explorar os benefícios dessa topologia de rede. Algumas arquiteturas de interconexão direta da técnica anterior forneceram um sistema em que cada nó, ou um subconjunto de nós (ou seja, nós de gateway), possui conectividade dupla, tanto para a interconexão direta quanto para a rede tradicional, mas é difícil gerenciar e carregar os recursos do dispositivo à medida que ligam ou roteiam entre as duas redes.

[0006] Seria, portanto, desejável ter um gateway de interconexão direta projetado e capaz de permitir a comunicação entre dispositivos de interconexão direta e dispositivos de interconexão não direta. Além disso, seria benéfico ter um gateway que pudesse ajudar a superar algumas das deficiências descritas acima para o tráfego de E / S.

SUMÁRIO DA INVENÇÃO

[0007] Em um aspecto, a presente invenção fornece um dispositivo dedicado, ou seja, um dispositivo de gateway, capaz de fazer uma ponte, alternar ou rotear entre redes de interconexão tradicionais e diretas.

[0008] Em outro aspecto, a presente invenção fornece um dispositivo de gateway altamente gerenciável que pode ser gerenciado por sistemas de gerenciamento de rede.

[0009] Ainda em outro aspecto, a presente invenção fornece um dispositivo de gateway que permite a coordenação de tabelas MAC e respostas ARP, broadcast, multicast e anycast entre várias portas de interconexão direta.

[0010] Em uma modalidade, a presente invenção fornece um dispositivo de gateway de rede dedicado que é capaz de ligar, alternar ou rotear o tráfego de rede entre redes de interconexão tradicionais e diretas, compreendendo: um primeiro conjunto de uma ou mais portas de rede tradicionais com um único link por porta, sendo essas portas conectadas a comutadores ou dispositivos que formam uma rede tradicional; e um segundo conjunto de uma ou mais portas de interconexão direta com dois ou mais links por porta, sendo essas portas conectadas a uma rede de interconexão direta. As portas de rede tradicionais podem compreender um ou mais dos conectores SFP+, QSFP e QSFP+ e podem ser conectadas às portas de comutadores ou roteadores na rede tradicional, enquanto as portas de interconexão direta podem compreender um ou mais conectores MXC, MTP e MTO e po-

dem ser conectadas a um patch panel / hub passivo usado na implementação da rede de interconexão direta. Como alternativa, as portas de interconexão direta podem ser conectadas a seu próprio circuito integrado específico de aplicativo (ASIC) de interconexão direta dedicado ou podem ser conectadas, cada uma, a um ou mais circuitos integrados específicos de aplicativo (ASICs) compartilhados. A função de ligação em ponte, comutação ou roteamento pode ser realizada por um comutador de rede ASIC ou por um controlador de rede ASIC. Os ASICs podem ser capazes de atuar como um nó de interconexão direta com tráfego destinado / originado localmente enviado por uma linha de interface de rede tradicional. Além disso, em outras modalidades, as portas de interconexão direta podem substituir um dispositivo dentro da rede de interconexão direta e podem até ser conectadas a várias redes de interconexão direta.

[0011] Em outra modalidade, a presente invenção fornece um dispositivo de gateway de rede dedicado que é capaz de conectar, alternar ou rotear o tráfego de rede entre redes de interconexão tradicional e direta, em que o referido dispositivo compreende duas portas, a saber, uma primeira porta que é uma porta de interconexão direta que pode ser conectada a uma rede de interconexão direta e uma segunda porta que é uma porta de rede padrão capaz de ser conectada a comutadores ou dispositivos que formam uma rede tradicional. A primeira porta pode compreender um dos conectores MXC, MTP ou MTO e pode ser conectada a um patch panel / hub passivo usado na implementação da rede de interconexão direta. A segunda porta pode compreender um dos conectores SFP+, QSFP ou QSFP+ e pode ser conectada às portas do comutador ou roteador na rede tradicional.

[0012] Ainda em outra modalidade, a presente invenção fornece um dispositivo de gateway de rede dedicado que é capaz de fazer uma ponte ou rotear o tráfego de rede entre uma rede tradicional e uma re-

de de interconexão direta, compreendendo: um primeiro conjunto de portas de rede tradicionais com um único link por porta, sendo essas portas conectadas a dispositivos finais que formam uma primeira rede tradicional; e um segundo conjunto de portas de interconexão direta com dois ou mais links por porta, sendo essas portas conectadas à rede de interconexão direta, em que a referida rede de interconexão direta atua como um backbone (base) que permite que o tráfego de rede seja roteado do dispositivo de gateway de rede dedicado para outro dispositivo de gateway de rede dedicado, o referido outro dispositivo de gateway de rede dedicado, compreendendo: um primeiro conjunto de portas de rede tradicionais com um único link por porta, sendo essas portas conectadas a dispositivos finais que formam uma segunda rede tradicional; e um segundo conjunto de portas de interconexão direta com dois ou mais links por porta, sendo essas portas conectadas à rede de interconexão direta.

[0013] Ainda em outra modalidade, a presente invenção proporciona um dispositivo de gateway de rede dedicado que é capaz de ligar, alternar ou rotear o tráfego de rede entre a rede tradicional e as redes de interconexão direta, compreendendo: um primeiro conjunto de uma ou mais portas de rede tradicionais com um único link por porta, sendo essas portas conectadas a comutadores ou dispositivos que formam uma rede tradicional; um segundo conjunto de uma ou mais portas de interconexão direta com um ou mais links por porta, sendo essas portas conectadas a uma rede de interconexão direta; e uma pluralidade de portas de interconexão direta que são logicamente associadas para atuar como um único nó de interconexão direta.

[0014] Em outra modalidade, a presente invenção fornece um método implementado por computador de ligação em ponte, comutação ou roteamento de tráfego de rede entre uma rede tradicional e uma rede de interconexão direta, compreendendo as etapas de: conexão

do dispositivo de gateway dedicado à rede tradicional e à rede de interconexão direta, o referido dispositivo de gateway dedicado atuando como um ou mais nós dentro da rede de interconexão direta; e encaminhamento do tráfego de rede por meio do dispositivo de gateway entre a rede tradicional e a rede de interconexão direta com base nos cabeçalhos ou no conteúdo do tráfego de rede.

[0015] Em uma modalidade adicional, a presente invenção fornece um método implementado por computador para coordenar qual dispositivo de gateway deve fornecer acesso a um recurso localizado em uma rede tradicional quando o referido recurso é acessível por mais de um dispositivo de gateway, compreendendo as etapas de: (i) recebimento de um tráfego ARP, broadcast, multicast ou anycast em uma porta de interconexão direta, em que o referido tráfego está solicitando acesso ao recurso localizado na rede tradicional e em que a referida porta de interconexão direta está vinculada por um ou mais saltos ao mais de um dispositivos de gateway, cada um dos quais é capaz de fornecer acesso ao recurso; (ii) cálculo de uma porta ideal de dispositivo de gateway dentre os mais de um dispositivos de gateway que devem fornecer acesso ao recurso; (iii) criação de uma associação entre o tráfego, o nó de interconexão direta e a porta ideal calculada do dispositivo de gateway; e (iv) comunicação da associação com cada um dos mais de um dispositivos de gateway para garantir que a porta ideal calculada do dispositivo de gateway forneça acesso ao recurso. A etapa de cálculo da porta ideal do dispositivo de gateway que deve fornecer acesso ao recurso pode compreender a determinar qual das mais de uma porta do dispositivo de gateway está mais próxima da porta de interconexão direta ou pode compreender o emprego de um algoritmo de consenso para garantir a consistência do tráfego. A etapa de comunicação da associação pode ser realizada por um barramento de coordenação dedicado ou compartilhado.

BREVE DESCRIÇÃO DOS DESENHOS

[0016] A modalidade da invenção será agora descrita, a título de exemplo, com referência aos desenhos anexos nos quais:

[0017] A FIGURA 1 é uma visão geral de alto nível de um dispositivo de gateway de acordo com uma modalidade da presente invenção.

[0018] A FIGURA 2 é uma visão geral de um dispositivo de gateway de acordo com uma modalidade da presente invenção, compreendendo um ASIC de comutador de rede tradicional com certas portas conectadas a conectores de rede tradicionais e outras portas conectadas ao seu próprio ASIC de interconexão direta dedicado.

[0019] A FIGURA 2b é uma visão geral de um dispositivo de gateway de acordo com uma modalidade da presente invenção (relacionada à mostrada na Figura 2), em que cada ASIC de interconexão direta pode ser conectado a mais de uma porta de interconexão direta e pode ser conectado a mais de uma porta ASIC do comutador de rede tradicional.

[0020] A FIGURA 2c é uma visão geral de um dispositivo de gateway de acordo com uma modalidade da presente invenção, compreendendo um único ASIC de comutador e interconexão direta que combina as funções de um ASIC de comutador de rede tradicional e um ou mais ASICs de interconexão direta.

[0021] A FIGURA 2d é uma visão geral de um dispositivo de gateway de acordo com uma modalidade da presente invenção, compreendendo uma placa de interface host contendo uma porta de rede tradicional e uma porta de interconexão direta.

[0022] A FIGURA 2e é uma visão geral de um dispositivo de gateway de acordo com uma modalidade da presente invenção, compreendendo uma placa de interface do host em que as funções de um ASIC de interconexão direta e de um ASIC de controlador de rede tra-

dicional são combinadas em um ASIC único de interconexão direta e tradicional.

[0023] A FIGURA 2f é uma visão geral de um dispositivo de gateway de acordo com uma modalidade da presente invenção, em que as portas de interconexão direta contêm um ou mais links por porta e são combinadas em grupos pelo dispositivo de gateway, de modo que cada grupo seja logicamente associado pelo dispositivo de gateway para atuar como um nó único na rede de interconexão direta.

[0024] A FIGURA 3 é uma visão geral de um dispositivo de gateway de acordo com uma modalidade da presente invenção, em que portas de interconexão direta são conectadas a um patch panel / hub passivo.

[0025] A FIGURA 4 é uma visão geral de um dispositivo de gateway de acordo com uma modalidade da presente invenção, em que uma porta de interconexão direta substitui um servidor dentro de uma topologia de interconexão direta.

[0026] A FIGURA 4b exibe uma modalidade como mostrado na Figura 4, em que os links das portas de interconexão direta são conectados a diferentes nós.

[0027] A FIGURA 5 é uma visão geral de uma modalidade da presente invenção, em que as portas de interconexão direta no gateway estão ligadas a diferentes interconexões diretas para permitir ligação em ponte, comutação ou roteamento entre várias interconexões diretas.

[0028] A FIGURA 6 é uma visão geral de uma modalidade da presente invenção, em que todos ou a maioria dos nós na interconexão direta são compostos de portas de gateway e em que todos ou a maioria dos outros dispositivos seriam conectados diretamente a um ou mais gateways.

[0029] A FIGURA 7 fornece um exemplo de como minimizar a dis-

tância média (em saltos) de cada dispositivo na interconexão direta à porta do gateway mais próxima.

[0030] A FIGURA 8 fornece um exemplo de como o gateway pode coordenar qual porta de interconexão direta deve responder a uma solicitação ARP recebida em mais de uma porta na mesma interconexão direta.

[0031] A FIGURA 8b fornece um exemplo de uma modalidade de um mecanismo de coordenação (como na Figura 8), em que a função de coordenação é distribuída pelos ASICs de interconexão direta e cada ASIC de interconexão direta é conectado a um barramento de coordenação usado para comunicar informações e decisões de coordenação.

[0032] A FIGURA 9 fornece uma modalidade de uma árvore lógica que explica como, quando mais de um gateway é conectado ao mesmo toro, os gateways podem coordenar seus conhecimentos sobre a topologia do toro e a resposta ao tráfego ARP, broadcast, multicast e anycast.

[0033] A FIGURA 10 fornece um exemplo de como um nó de gateway intermediário pode processar um pacote em vez de rotear o pacote para outro nó de gateway.

[0034] As FIGURAS 11 e 12 descrevem a operação de uma modalidade preferida de um ASIC de interconexão direta.

DESCRIÇÃO DETALHADA DA INVENÇÃO

[0035] A presente invenção fornece um dispositivo dedicado, nomeadamente um dispositivo de gateway, capaz de fazer uma ligação em ponte, comutar ou rotear entre redes de interconexão tradicionais e diretas. Ao empregar um dispositivo dedicado, os recursos nos nós de interconexão direta não precisam ser sobrecarregados através da ligação em ponte, comutação ou roteamento entre uma rede de interconexão direta e a rede tradicional, minimizando, assim, os impactos no

desempenho de E / S. Além disso, em oposição ao uso da técnica anterior de nós de gateway, o presente dispositivo de gateway é um dispositivo altamente gerenciável que pode ser gerenciado por sistemas de gerenciamento de rede. Além disso, o dispositivo de gateway da presente invenção permite a coordenação de tabelas MAC e respostas ARP, broadcast, multicast e anycast entre várias portas de interconexão direta.

[0036] A figura 1 mostra uma visão geral de alto nível de um dispositivo de gateway 50, de acordo com uma modalidade da presente invenção, compreendendo dois conjuntos de portas. O primeiro conjunto de portas (neste exemplo, as doze portas mais à esquerda) são portas de rede tradicional padrão 100 com um único link por porta (por exemplo, conectores SFP+, QSFP, QSFP+) conectados aos computadores e / ou dispositivos existentes que formam a rede tradicional. Em uma topologia Clos (uma rede de comutação de circuitos de vários estágios), por exemplo, essas portas 100 mais provavelmente seriam conectadas a spine ou super-spine ports. O segundo conjunto de portas (as doze portas mais à direita) são portas de interconexão direta 102 com um alto número de links (dois ou mais) por porta (por exemplo, com conectores MXC / MTP / MTO). Se um patch panel / hub passivo 60 semelhante ao divulgado na Publicação de Pedido de Patente PCT Nº WO 2015/027320 A1 estiver sendo usado para suportar a topologia de interconexão direta, por exemplo, as portas de interconexão direta 102 serão conectadas ao patch panel passivo / hub 60 (veja a Figura 3). O dispositivo de gateway 50 pode estar na forma de um dispositivo montável em rack de 1 unidade de rack (RU) ou mesmo 1/2 RU (ou de outra forma conforme desejado) para economia eficiente de espaço no rack. A Figura 2 mostra uma modalidade composta por um circuito integrado específico de aplicação do comutador de rede tradicional (ASIC) 106, em que no primeiro conjunto de portas de rede tra-

dicional 100 são, cada uma delas, conectadas a conectores de rede tradicional e em que no segundo conjunto de portas de interconexão direta 102 são, cada uma delas, conectadas ao seu próprio ASIC 104 de interconexão direta dedicada. Esses ASICs de interconexão direta dedicada 104 são, de preferência, cada um deles, capazes de atuar como um nó de interconexão direta com tráfego destinado / originado localmente enviado por uma linha de interface de rede tradicional (por exemplo, Ethernet de 100 Gbps). As conexões entre as portas do ASIC 106 de comutador e ASICs de interconexão direta 104 podem ser implementadas usando qualquer interconexão de componente padrão, como 100GBase-KR4, 100GBase-KP4, 40GBase-KR4, 25GBase-KR, 10GBase-KR ou qualquer outro padrão semelhante.

[0037] O ASIC 106 de comutador de rede tradicional contém a funcionalidade padrão de encaminhamento de tráfego de rede, incluindo o aprendizado dos dispositivos acessíveis através de suas portas, o envio de tráfego recebido pela porta de saída apropriada, filtragem de rede, inspeção de tráfego e outras funcionalidades normalmente encontradas na camada 2, na camada 3 e na camada 4 e acima de comutadores, roteadores e pontes de rede. As decisões de encaminhamento podem se basear em um ou mais fatores, incluindo, mas não limitando, entre outros, endereços da camada 2 de origem e destino (MAC), porta de origem, endereços da camada 3 de origem e destino (IPv4, IPv6, etc.), portas da camada 4 de origem e destino e os cabeçalhos e cargas úteis de dados da camada 5 e acima.

[0038] O tráfego de rede recebido da interconexão direta em um ASIC 104 de interconexão direta que tem um destino final acessível através das portas padrão do ASIC 106 de comutador será enviado do ASIC 104 de interconexão direta para o ASIC 106 de comutador, onde a funcionalidade de encaminhamento de tráfego padrão do ASIC 106 de comutador transmitirá o tráfego através da porta padrão apropriada.

[0039] Da mesma forma, o tráfego de rede recebido pelo ASIC 106 de comutador de uma porta padrão que tem um destino final acessível por meio de um ASIC 104 de interconexão direta será encaminhado pelo ASIC 106 de comutador para direcionar o ASIC 104 de interconexão direta.

[0040] Em outra modalidade (não mostrada), e conforme aplicável a todas as modalidades possíveis, deve ser entendido que o ASIC 106 de comutador (e todos os ASICs semelhantes discutidos aqui) pode ser substituído por um arranjo de portas programável em campo (FPGA), em um processador para fins gerais, processador de rede ou qualquer outro dispositivo capaz de executar o encaminhamento de tráfego de rede.

[0041] Em outra modalidade (não mostrada), e conforme aplicável a todas as modalidades possíveis, deve ser entendido que o ASIC 104 de interconexão direta (e todos os ASICs semelhantes discutidos aqui) pode ser substituído por um arranjo de portas programável em campo (FPGA), processador de uso geral, processador de rede ou qualquer outro dispositivo capaz de atuar como um nó em uma rede de interconexão direta.

[0042] A figura 2b mostra outra modalidade onde cada ASIC 104 de interconexão direta pode ser conectado a uma ou mais portas de interconexão direta 102 e pode ser conectado a uma ou mais portas de 106 de comutador de rede tradicional.

[0043] A figura 2c mostra outra modalidade onde um único ASIC 105 de comutador e de interconexão direta combina as funções de ASIC 106 de comutador de rede tradicional e um ou mais ASICs de interconexão direta 104. Mais uma vez, deve ser entendido que o ASIC 105 único de comutador e de interconexão direta (e todos os ASICs similares discutidos aqui) pode ser substituído por um FPGA (field-programmable gate array), processador de uso geral, processa-

dor de rede ou qualquer outro dispositivo capaz de executar o encaminhamento de tráfego de rede e / ou atuando como um nó em uma rede de interconexão direta.

[0044] A Figura 2d mostra outra modalidade de um dispositivo de gateway na forma de uma placa de interface de host 110 contendo uma porta de rede tradicional e uma porta de interconexão direta. A placa de interface do host 110 é projetada para ser conectada a um dispositivo host através de um barramento de comunicação como PCIe, Gen-Z ou CCIX. O dispositivo host pode ser um servidor, computador, disco rígido, unidade de estado sólido, gabinete da unidade ou qualquer outro dispositivo capaz de enviar e receber dados. A porta de rede tradicional é conectada ao ASIC 108 de controlador de rede padrão (ou similarmente a um FPGA (field-programmable gate array), processador de uso geral, processador de rede ou qualquer outro dispositivo capaz de executar o encaminhamento de tráfego de rede, etc.) e a porta de interconexão direta está conectada ao ASIC 104 de interconexão direta. O ASIC 104 de interconexão direta também está conectado ao controlador de rede ASIC 108. Dessa forma, o ASIC 108 de controlador de rede seria capaz de alternar o tráfego entre a rede tradicional e a rede de interconexão direta sem a intervenção do host.

[0045] A Figura 2e mostra uma modalidade de outra placa de interface de host 111, em que as funções do ASIC 104 de interconexão direta e do ASIC 108 de controle de rede são combinadas em um único ASIC 107 de interconexão direta e tradicional. Mais uma vez, deve ser compreendido que o ASIC 107 de interconexão direta e tradicional único (e todos os ASICs semelhantes discutidos aqui) podem ser substituídos por uma field-programmable gate array (FPGA), processador de uso geral, processador de rede ou qualquer outro dispositivo capaz de executar o encaminhamento de tráfego de rede e / ou agir como um nó em uma rede de interconexão direta.

[0046] A Figura 2f mostra outra modalidade em que as portas de interconexão direta 102 contêm um ou mais links por porta e são combinadas em grupos 55 pelo dispositivo de gateway 50. Cada grupo 55 é logicamente associado pelo dispositivo de gateway 50 para atuar como um único nó dentro da rede de interconexão direta.

[0047] Como mostrado na Figura 4, as portas de interconexão direta 102 e os grupos 55 tomariam cada um o lugar de um dispositivo dentro da topologia de interconexão direta.

[0048] Ainda em outra modalidade, se um patch panel / hub passivo 60 não for utilizado na interconexão direta, os links individuais de cada porta de gateway (isto é, as portas de interconexão direta 102) podem ser conectados a dispositivos que fazem parte da interconexão direta. A este respeito, a Figura 4b mostra uma modalidade em que os links de uma das portas de interconexão direta 102 são individualmente conectados aos nós vizinhos dentro da rede de interconexão direta. Na Figura 4b, a interconexão direta assume a forma de um toro bidimensional e os 4 links que compreendem a porta de interconexão direta 102 são numerados 1, 2, 3 e 4. Para formar o toro bidimensional, o link 1 é conectado ao dispositivo A, o link 2 está conectado ao dispositivo B, o link 3 está conectado ao dispositivo C e o link 4 está conectado ao dispositivo D. Da mesma forma, para outras topologias, cada link da porta de interconexão direta 102 deve ser conectado ao dispositivo vizinho apropriado exigido por essa topologia.

[0049] Em uma modalidade adicional, conforme mostrado na figura 5, o dispositivo de gateway 50 também pode ser usado para ligar em ponte, alternar ou rotear entre várias interconexões diretas. Neste caso, as portas de interconexão direta 102 no dispositivo de gateway 50 serão divididas entre as diferentes interconexões diretas (mostradas na Figura 5 como A e B). Todos os dispositivos na interconexão direta A seriam então alcançáveis a partir da interconexão direta A

através do dispositivo de gateway 50 e vice-versa. No exemplo fornecido, o tráfego de um dispositivo na interconexão direta A destinado a um dispositivo na interconexão direta B atravessaria primeiro a interconexão direta A para a porta de interconexão direta 102 mostrada como A4 no dispositivo de gateway 50. O dispositivo de gateway 50 encaminharia esse tráfego através a porta de interconexão direta 102 mostrada como B2, onde seria então encaminhado através da Interconexão direta B para o nó de destino.

[0050] Ainda em outra modalidade da presente invenção, os gateways poderiam ser usados como comutadores de acesso e a interconexão direta formaria a espinha dorsal (veja a Figura 6). Nesse caso, todos ou a maioria dos nós na interconexão direta seriam compostos de portas de gateway e todos ou a maioria dos outros dispositivos seriam conectados diretamente às portas de rede tradicionais 100 do dispositivo de gateway 50. No exemplo fornecido na Figura 6, o tráfego do grupo de dispositivos A destinado ao grupo de dispositivos B seria encaminhado primeiro para o dispositivo de gateway 50A. A função de encaminhamento do gateway 50A reconheceria que o dispositivo de destino é alcançável através da interconexão direta e encaminharia o tráfego através de uma das portas 102 de interconexão direta do dispositivo de gateway 50A. A interconexão direta encaminharia o tráfego para uma das portas de interconexão direta do gateway dispositivo 50B. A função de encaminhamento do gateway 50B reconheceria que o dispositivo de destino é alcançável através de uma de suas portas de rede padrão e encaminharia o tráfego adequadamente.

[0051] A fim de maximizar a eficiência do tráfego de E / S, a Figura 7 fornece um exemplo de como, em uma única implantação de interconexão direta, as portas de interconexão direta 102 podem ser escolhidas para minimizar a distância média (número de saltos) de cada dispositivo de gateway 50 na interconexão direta à porta do gateway

mais próxima. Em um toro 4x4 2D com os nós numerados conforme a Figura 7, fazer com que o gateway atue como nós 1,6,11 e 16 minimizaria a distância aos outros nós na interconexão direta. No entanto, um versado na técnica também entenderia que, se um subconjunto dos nós tende a gerar uma quantidade maior de E / S do que a média, a implantação dos nós de gateway pode ser influenciada para estar mais próxima desses nós de E / S mais altos. São conhecidos vários algoritmos que podem ser usados para determinar a localização ideal para esses nós de E / S (consulte, por exemplo, Bae, M., Bose, B. : "Resource Placement in Torus- Based Networks" em: Proc. IEEE Internacional Parallel Processing Symposium, pp. 327-331. IEEE Computer Society Press, Los Alamitos (1996); Dillow, David A. et al.; "I / O Congestion Avoidance via Routing and Object Placement" United States: N. p., 2011 Print. Proceedings of the Cray User Group conference (CUG 2011), Fairbanks, AK, USA; Almohammad, B., Bose, B. : "Resource Placements in 2D Tori" in: Proc. IPPS 1998 Proceedings of the 12th International Parallel Processing Symposium on International Parallel Processing Symposium, p. 431, IEEE Computer Society, Washington, DC (1998); Dillow, David A. et al.; "Enhancing I/O Throughput via Efficient Routing and Placement for Large-scale Parallel File Systems" in: Conference Proceedings of the IEEE International Performance, Computing and Communications Conference (IPCCC), 2011 IEEE 30th International; Ezel, M. et al.: "I/O Router Placement and Fine-Grained Routing on Titan to Support Spider II" in: "Proceedings of the Cray User Group Conference (CUG), May 2014; e Babatunde, A. et al.: "I/O Node Placement for Performance and Reliability in Torus Networks", Department of Computer Science, Texas AM University, Jan 2006).

[0052] Em uma modalidade preferida, as portas de interconexão direta 102 atuarão como portas de interconexão direta padrão e enca-

minham autonomamente o tráfego restante na interconexão direta (isto é, FLITs de encaminhamento). Elas também recombina FLITs em pacotes de rede para o tráfego destinado a dispositivos que não estão na interconexão direta (consulte a Publicação do Pedido de Patente PCT N° WO 2015/120539 A1 para obter um método ideal para rotear pacotes em uma rede de interconexão direta distribuída). O dispositivo de gateway 50, preferencialmente, também tem a capacidade de transmitir / receber pacotes de rede para / de cada uma das portas de rede tradicional 100 e portas de interconexão direta 102, e também ser capaz de interpretar e encaminhar esse tráfego com base na camada 2, 3 ou como pelo acima.

[0053] Em uma modalidade preferida, as interfaces de gerenciamento de rede para o norte padrão seriam expostas (por exemplo, CLI, OpenFlow, SNMP, REST, etc.) para permitir que um sistema de gerenciamento de rede gerencie o dispositivo de gateway 50.

[0054] Em uma modalidade, quando várias portas de gateway são conectadas à mesma interconexão direta, todos os pacotes de um determinado fluxo devem preferencialmente sair na mesma porta de gateway para ajudar a garantir a entrega de pacotes em ordem.

[0055] O dispositivo de gateway 50 deve, preferencialmente, ser configurado para agregar tabelas de encaminhamento MAC entre as portas de interconexão direta 102 conectadas à mesma interconexão direta (ou seja, quando uma porta de interconexão direta aprende de um endereço/nó _id tuple VLAN / MAC, essa tupla deve preferencialmente ser compartilhada com as outras portas de interconexão direta 102 conectadas à mesma interconexão direta).

[0056] Em uma modalidade preferida, quando uma solicitação de ARP é recebida em uma ou mais das portas de interconexão direta 102 conectadas à mesma interconexão direta, a decisão de qual porta de interconexão direta deve responder deve ser coordenada pelo ga-

teway 50 para garantir que apenas uma única resposta é transmitida (veja, por exemplo, na Figura 8). Isso pode ser feito, por exemplo, escolhendo a porta de interconexão direta mais próxima da origem da solicitação do ARP, a primeira porta de interconexão direta a receber o ARP, através de um esquema de seleção round-robin, através de um hash de uma parte da solicitação de ARP (endereço de IP, MAC de origem, etc.) ou outro algoritmo conhecido dos versados na técnica. A Figura 8b mostra uma modalidade deste mecanismo de coordenação, em que a função de coordenação é distribuída entre os ASICs de interconexão direta 104 e cada ASIC de interconexão direta é conectado ao barramento de coordenação 112 usado para comunicar informações e decisões de coordenação. Em outra modalidade, esta função de coordenação pode ser centralizada em um ASIC de coordenação dedicado e cada ASIC de interconexão direta é conectado ao ASIC de coordenação através de um barramento de coordenação dedicado ou compartilhado 112.

[0057] Quando mais de um gateway está conectado ao mesmo toro, os dispositivos de gateway 50 devem, preferencialmente, coordenar seus conhecimentos sobre a topologia do toro e a resposta às solicitações de ARP de maneira semelhante ao caso do gateway único discutido acima (consulte a Figura 9 para a árvore lógica nessa coordenação). Em uma modalidade, os gateways poderiam se descobrir através de um protocolo de transmissão. Em outra modalidade, os gateways podem ser configurados para conhecer a localização dos outros gateways. Ainda em outra modalidade, um algoritmo de consenso como o Raft poderia ser usado para garantir a coordenação dos gateways e a consistência das associações tráfego / gateway / porta (ver, por exemplo, Ongaro, D., Ousterhout, J. : "In Search of an Understandable Consensus Algorithm (Extended Version)" in: 2014 USENIX Annual Technical Conference, June 19, 2014, Philadelphia, PA; Woos, D.

et al.: "Planning for Change in a Formal Verification of the Raft Consensus Protocol" in: Certified Programs and Proofs (CPP), jan. 2016).

[0058] [58] Em geral, sempre que um nó do toro desejar se comunicar com um recurso acessível através de um ou mais dispositivos de gateway 50, o (s) gateway (s) devem preferencialmente coordenar qual porta de gateway é escolhida para fornecer acesso a esse recurso de maneira semelhante ao exemplo do ARP descrito acima. Exemplos disso incluem tráfego anycast, broadcast e multicast, protocolos de descoberta de nós e serviços e descoberta de vizinhos IPv6.

[0059] Como uma consideração adicional, é importante observar que, em muitos casos, o roteamento não mínimo é usado dentro de uma interconexão direta. Como o (s) gateway (s) em uma interconexão direta possui portas de gateway em vários locais da topologia, é possível que o tráfego destinado a uma porta de gateway atravessasse primeiro uma das outras portas de gateway. Portanto, seria preferível aumentar a eficiência fazendo com que uma única porta de gateway processasse o tráfego, em vez de permitir que o tráfego atravessasse para uma porta de gateway mais distante. Um exemplo disso é fornecido na Figura 10, em que o nó 5 está enviando tráfego para o nó 1, mas devido ao roteamento não mínimo, o pacote atravessará o nó 6. Em uma modalidade, para melhorar a eficiência, o gateway com a porta localizada no nó 6 reconhecerá que o pacote está destinado a outra porta de gateway e processará o pacote como se estivesse destinado ao nó 6 em vez de encaminhá-lo através da interconexão direta via nó 2.

[0060] Como observado acima, o Direct Interconnect ASIC 104 fornece conectividade entre o Switch ASIC 106 e uma interconexão direta. A fim de garantir que uma pessoa habilitada na técnica possa fabricar e trabalhar um dispositivo de gateway de rede da presente invenção, as Figuras 11 e 12 descrevem a operação de uma modalidade preferida do Direct Interconnect ASIC 104. Esta operação também é

aplicável, com as modificações necessárias a outros ASICs dentro do escopo desta invenção, incluindo ASICs 105.107 e 108, conforme apropriado.

[0061] Como é bem conhecido na técnica, o Switch ASIC 106 transmite e recebe quadros Ethernet. A Figura 11 descreve como o Ethernet Frame 200, gerado pelo Switch ASIC 106, é processado pelo Direct Interconnect ASIC 104. O Banco de Dados de Endereço MAC 201 contém uma lista de endereços MAC de Ethernet e os nós de interconexão direta associados a cada endereço MAC de Ethernet. O Ethernet Frame 200 recebido do Switch ASIC 106 é examinado e os endereços MAC de origem e destino são recuperados de seu cabeçalho. Esses endereços MAC são usados como índices para o banco de dados de endereços MAC 201. O endereço MAC de origem do Ethernet Frame 301 é combinado com o número do nó da corrente para criar ou atualizar uma associação entre esse número de nó e o endereço MAC de origem no banco de dados de endereços MAC 201.

[0062] Se o endereço MAC de destino do Ethernet Frame 200 estiver no banco de dados de endereços MAC 201, o número de nó 202 associado a esse endereço MAC será recuperado do banco de dados de endereços Mac 201. O número de nó 202 será usado como um índice na rota de origem O banco de dados 206 e a rota de origem 203 associados ao Nó Número 202 são recuperados do Banco de Dados de Rota de Origem 206. Como é bem conhecido na técnica, os bancos de dados de rotas de origem contêm uma lista de destinos de rede e um ou mais caminhos através da rede para alcançar cada destino. Um banco de dados de rota de origem pode ser preenchido manualmente ou pode depender de algoritmos de descoberta de topologia automatizada e de rota bem conhecidos. O Ethernet Frame 200 é então convertido em FLITs 204. Um FLIT é um tipo de quadro especializado usado em interconexões diretas e pode ser de tamanho fixo ou variá-

vel. Em uma modalidade preferida, os FLITs terão um tamanho fixo. Em outra modalidade, FLITs terão um tamanho variável dentro de um tamanho mínimo e máximo. Ainda em outra modalidade, os FLITs serão dimensionados de modo que o Ethernet Frame 200 se encaixe exatamente na carga útil do FLIT.

[0063] Se o Ethernet Frame 200 for maior que a carga útil de um único FLIT, vários FLITs 204 serão criados. Se o Ethernet Frame 200 couber na carga útil de um único FLIT, um único FLIT será criado. A rota de origem 203 é então inserida no cabeçalho do primeiro dos FLITs 204 juntamente com o número de nó do nó atual. FLITs 204 são então transmitidos a partir da porta de saída 205 especificada na rota de origem 203.

[0064] Se o endereço MAC de destino do Ethernet Frame 200 não estiver no banco de dados de endereços MAC 201 ou se o endereço MAC de destino do Ethernet Frame 200 indicar que é um pacote Ethernet de transmissão, então, o Ethernet Frame 200 será convertido em FLITs 204 como no caso descrito acima, embora uma rota de origem não seja incluída. Uma vez criados os FLITs 204, um flag no cabeçalho do primeiro FLIT é configurado para indicar que esses FLITs devem ser transmitidos para todos os nós na interconexão direta. Um valor de time-to-live (TTL) também é definido no cabeçalho do primeiro FLIT. O TTL determina o número máximo de vezes que os FLITs transmitidos podem ser encaminhados através da interconexão direta. Em uma modalidade, os quadros de Ethernet anycast e multicast são tratados como se fossem quadros de transmissão, como acima.

[0065] A Figura 12 descreve como os FLITs 204 são processados pelo Direct Interconnect ASIC 104 em uma modalidade preferida do Direct Interconnect ASIC 104. Os FLITs 204 do Direct Interconnect 301 são recebidos pelo Direct Interconnect ASIC 104 e o cabeçalho do primeiro destes FLITs é examinado para ver se o flag de transmissão

está definido. Se o flag de transmissão não estiver definido, a Source Route (Rota de Origem) 203 é recuperada do primeiro FLIT e é determinado se a rota de origem indica que o nó atual é o nó de destino dos FLITs 204. Em outra modalidade, os FLITs contêm o número de nó do nó de destino e é o número do nó usado para determinar se o nó atual é o nó de destino para os FLITs 204. Se o nó atual é o nó de destino dos FLITs 204, os FLITs 204 são combinados para formar o Ethernet Frame 301. O endereço MAC de origem do Ethernet Frame 301 é combinado com o número do nó no cabeçalho do primeiro FLIT para criar ou atualizar uma associação entre o referido número de nó e o endereço MAC de origem no banco de dados de endereços MAC 201. O Ethernet Frame 301 é então transmitido para o Switch ASIC 106.

[0066] Se for determinado que este não é o nó de destino, a rota de origem é usada para determinar a porta de saída 302 para os FLITs 204. Os FLITs 204 são então transmitidos pela porta de saída 302.

[0067] Se o flag de transmissão é definido no primeiro cabeçalho FLIT, os FLITs 204 são combinados para formar o Ethernet Frame 301. O endereço MAC de origem do Ethernet Frame 301 é combinado com o número do nó no cabeçalho do primeiro FLIT para criar ou atualizar uma associação entre o referido número de nó e o endereço MAC de origem no banco de dados de endereços MAC 301. O Ethernet Frame 301 é transmitido para o Switch ASIC 106.

[0068] O TTL no cabeçalho do primeiro FLIT é então decrementado em um. Se o TTL agora é igual a zero, os FLITs 204 são descartados. Se o TTL for maior que zero, os FLITs 204 serão transmitidos para todas as portas de saída, exceto a porta de entrada da qual os FLITs 204 foram originalmente recebidos.

[0069] Em outras modalidades do Direct Interconnect ASIC 104, o roteamento de origem não pode ser usado. Em uma modalidade, o endereço MAC de destino do Ethernet Frame 200 será usado por cada

nó para executar uma pesquisa de rota do próximo salto local. Em outra modalidade, as informações do nó de destino no cabeçalho FLIT serão usadas por cada nó para executar uma pesquisa de rota local do próximo salto.

[0070] Será óbvio para os versados na técnica que outras modalidades do Direct Interconnect 104 e do Switch ASIC 106 podem ser projetadas para funcionar com protocolos diferentes da Ethernet. Em uma modalidade, esses elementos serão projetados para funcionar com Gen-Z. Nesse caso, a Direct Interconnect 104 esperaria receber pacotes Gen-Z Core64 em vez de Ethernet Frames. Em vez de endereços MAC Ethernet, os GCIDs Gen-Z (IDs de Componentes Globais) seriam usados e associados aos números de nós de interconexão direta.

Embora modalidades específicas da invenção tenham sido descritas, será evidente para um versado na técnica que variações e modificações nas modalidades podem ser feitas dentro do escopo das reivindicações a seguir.

REIVINDICAÇÕES

1. Dispositivo de gateway de rede dedicado (50), o qual é capaz de ligar em ponte, alternar ou rotear tráfego de rede entre redes tradicionais e de interconexão direta, caracterizado pelo fato de compreender:

- um primeiro conjunto de pelo menos uma porta de rede tradicional (100) com um único link por porta e

- um segundo conjunto de pelo menos uma porta de interconexão direta (102) com pelo menos dois links por porta;

em que o primeiro conjunto de portas de rede tradicional (100) está conectado a um dos computadores e dispositivos que formam uma rede tradicional; e

em que o segundo conjunto de pelo menos uma porta de interconexão direta (102) está conectado a uma rede de interconexão direta.

2. Dispositivo (50), de acordo com a reivindicação 1, caracterizado pelo fato de o primeiro conjunto de portas de rede tradicional (100) compreender pelo menos um dos conectores SFP+, QSFP e QSFP+.

3. Dispositivo (50), de acordo com a reivindicação 1, caracterizado pelo fato de o segundo conjunto de pelo menos uma porta de interconexão direta (102) compreender pelo menos um dos conectores MXC, MTP e MTO.

4. Dispositivo (50), de acordo com a reivindicação 1, caracterizado pelo fato de as portas de rede tradicional (100) estarem conectadas a pelo menos uma dentre uma porta de computador e uma porta de roteador na rede tradicional.

5. Dispositivo (50), de acordo com a reivindicação 1, caracterizado pelo fato de o segundo conjunto de portas de interconexão direta (102) ser conectado a um patch panel / hub passivo (60) usado

na implementação da rede de interconexão direta

6. Dispositivo (50), de acordo com a reivindicação 1, caracterizado pelo fato de o segundo conjunto de portas de interconexão direta (102) ter cada uma delas conectada a pelo menos um de seus próprios circuitos integrados específicos para aplicativos de interconexão direta dedicada (ASIC) 104, matriz de portas programáveis em campo (FPGA), processador de uso geral, processador de rede e dispositivo capaz de atuar como um nó em uma rede de interconexão direta.

7. Dispositivo (50), de acordo com a reivindicação 1, caracterizado pelo fato de o segundo conjunto de portas de interconexão direta (102) ter cada uma conectada a pelo menos um dos circuitos integrados específicos para aplicativos (ASICs) compartilhados, matrizes de portas programáveis em campo (FPGAs), processadores de uso geral, processadores de rede e dispositivos capazes de atuar como um nó em uma rede de interconexão direta.

8. Dispositivo (50), de acordo com a reivindicação 1, caracterizado pelo fato de pelo menos uma das funções de ligação em ponte, comutação e roteamento ser realizada por pelo menos uma das chaves de rede ASIC, matriz de portas programáveis em campo (FPGA), processador de uso geral, processador de rede e um dispositivo capaz de executar o encaminhamento de tráfego de rede.

9. Dispositivo (50), de acordo com a reivindicação 1, caracterizado pelo fato de pelo menos uma das funções de ligação em ponte, comutação e roteamento ser realizada por pelo menos um dentre um controlador de rede ASIC, a matriz de portas programáveis em campo (FPGA), um processador de uso geral, um processador de rede e um dispositivo capaz de executar o encaminhamento de tráfego de rede.

10. Dispositivo (50), de acordo com a reivindicação 9, ca-

racterizado pelo fato de pelo menos um dos ASICs, FPGAs, processadores de uso geral, processadores de rede e nós ser capaz de atuar como um nó de interconexão direta com tráfego destinado / originado localmente enviado por uma linha de interface de rede tradicional.

11. Dispositivo (50), de acordo com a reivindicação 1, caracterizado pelo fato de as portas de interconexão direta tomarem o lugar de um dispositivo dentro da rede de interconexão direta.

12. Dispositivo (50), de acordo com a reivindicação 1, caracterizado pelo fato de o segundo conjunto de portas de interconexão direta ser conectado a várias redes de interconexão direta.

13. Dispositivo de gateway de rede dedicado (50) capaz de, pelo menos, um dentre ligação em ponte, comutação e roteamento de tráfego de rede entre redes tradicional e de interconexão direta, caracterizado pelo fato de compreender:

- uma primeira porta que é uma porta de interconexão direta configurada para ser conectada a uma rede de interconexão direta;

e

- uma segunda porta que é uma porta de rede padrão configurada para ser conectada a pelo menos um dos comutadores e dispositivos que formam uma rede tradicional.

14. Dispositivo (50), de acordo com a reivindicação 13, caracterizado pelo fato de a primeira porta compreender pelo menos um dentre um conector MXC, MTP e MTO.

15. Dispositivo (50), de acordo com a reivindicação 13, caracterizado pelo fato de a segunda porta compreender pelo menos um dentre um conector SFP+, QSFP e QSFP+.

16. Dispositivo (50), de acordo com a reivindicação 13, caracterizado pelo fato de a primeira porta estar conectada a um patch panel / hub passivo usado na implementação da rede de interconexão direta.

17. Dispositivo (50), de acordo com a reivindicação 13, caracterizado pelo fato de a segunda porta estar conectada a pelo menos uma das portas de comutador e roteador na rede tradicional.

18. Dispositivo de gateway de rede dedicado (50) para pelo menos um dentre ligação em ponte e roteamento de tráfego de rede entre uma rede tradicional e uma rede de interconexão direta, caracterizado pelo fato de compreender:

- um primeiro conjunto de portas de rede tradicional (100) com um único link por porta, sendo essas portas conectadas a dispositivos finais que formam uma primeira rede tradicional; e

- um segundo conjunto de portas de interconexão direta (102) com dois ou mais links por porta, sendo essas portas conectadas à rede de interconexão direta,

- em que a referida rede de interconexão direta atua como uma estrutura que permite ao tráfego da rede rotear do dispositivo de gateway de rede dedicado (50) para outro dispositivo de gateway de rede dedicado (50), o referido outro dispositivo de gateway de rede dedicado (50) que compreende:

- um primeiro conjunto de portas de rede tradicional (100) com um único link por porta, sendo essas portas conectadas a dispositivos finais que formam uma segunda rede tradicional; e

- um segundo conjunto de portas de interconexão direta (102) com dois ou mais links por porta, sendo essas portas conectadas à rede de interconexão direta.

19. Dispositivo de gateway de rede dedicado (50) para pelo menos um dentre ligação em ponte, comutação e roteamento de tráfego de rede entre a rede tradicional e as redes de interconexão direta, caracterizado pelo fato de compreender:

- um primeiro conjunto de pelo menos uma porta de rede tradicional (100) com um único link por porta, sendo essas portas co-

nectadas a pelo menos um dos comutadores e dispositivos que formam uma rede tradicional;

- um segundo conjunto de pelo menos uma porta de interconexão direta (102) com um ou mais links por porta, estando essas portas conectadas a uma rede de interconexão direta; e

uma pluralidade de portas de interconexão direta que são logicamente associadas para atuar como um único nó de interconexão direta.

20. Método implementado por computador de pelo menos um dentre ligação em ponte, comutação e roteamento de tráfego de rede entre uma rede tradicional e uma rede de interconexão direta, caracterizado pelo fato de compreender as etapas de:

- conectar o dispositivo de gateway dedicado (50), como definido na reivindicação 1, à rede tradicional e à rede de interconexão direta, o referido dispositivo de gateway dedicado (50) atuando como um ou mais nós na rede de interconexão direta; e

- encaminhar o tráfego de rede por meio do dispositivo de gateway (50) entre a rede tradicional e a rede de interconexão direta com base nos cabeçalhos ou no conteúdo do tráfego de rede.

21. Método implementado por computador para coordenar qual dispositivo de gateway (50) com capacidade de fornecer acesso a um recurso localizado em uma rede tradicional quando o referido recurso estiver acessível por mais de um dispositivo de gateway (50), caracterizado pelo fato de compreender as etapas de:

- receber de um tráfego ARP, broadcast, multicast ou anycast em uma porta de interconexão direta, em que o referido tráfego está solicitando acesso ao recurso localizado na rede tradicional e em que a referida porta de interconexão direta está vinculada por um ou mais saltos a mais de um dispositivo de gateway (50), cada um dos quais é capaz de fornecer acesso ao recurso;

- calcular uma porta ideal do dispositivo de gateway dentre os mais de um dispositivo de gateway que deve fornecer acesso ao recurso;

- criar uma associação entre o tráfego, o nó de interconexão direta e a porta ideal calculada do dispositivo de gateway; e

- comunicar a associação com cada um dos mais de um dispositivo de gateway para garantir que a porta ideal calculada do dispositivo de gateway forneça acesso ao recurso.

22. Método implementado por computador, de acordo com a reivindicação 21, caracterizado pelo fato de a etapa de calcular a porta ideal do dispositivo de gateway que deve fornecer acesso ao recurso compreender a determinação de qual das mais de uma porta do dispositivo de gateway está mais próxima da porta de interconexão direta.

23. Método implementado por computador, de acordo com a reivindicação 21, caracterizado pelo fato de que a etapa de calcular a porta ideal do dispositivo de gateway que deve fornecer acesso ao recurso compreende empregar um algoritmo de consenso para garantir a consistência do tráfego.

24. Método implementado por computador, de acordo com a reivindicação 21, caracterizado pelo fato de que a etapa de comunicação da associação ser realizada por um barramento de coordenação dedicado ou compartilhado.

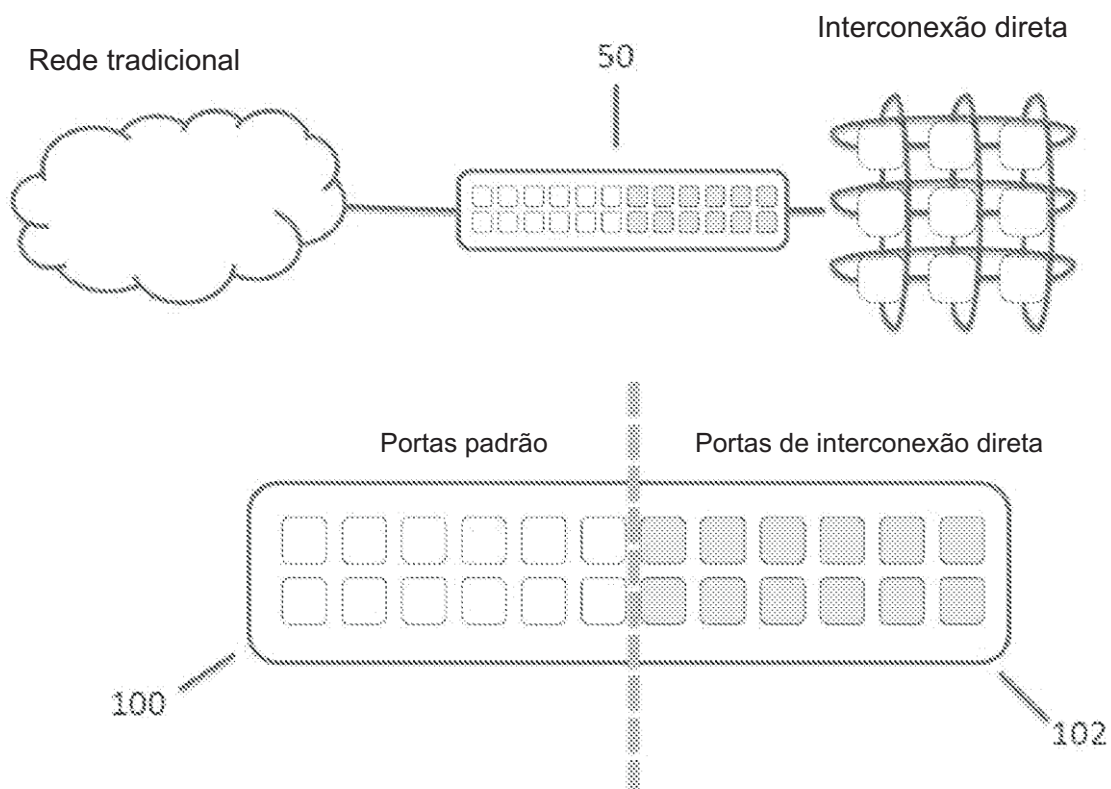


FIG. 1

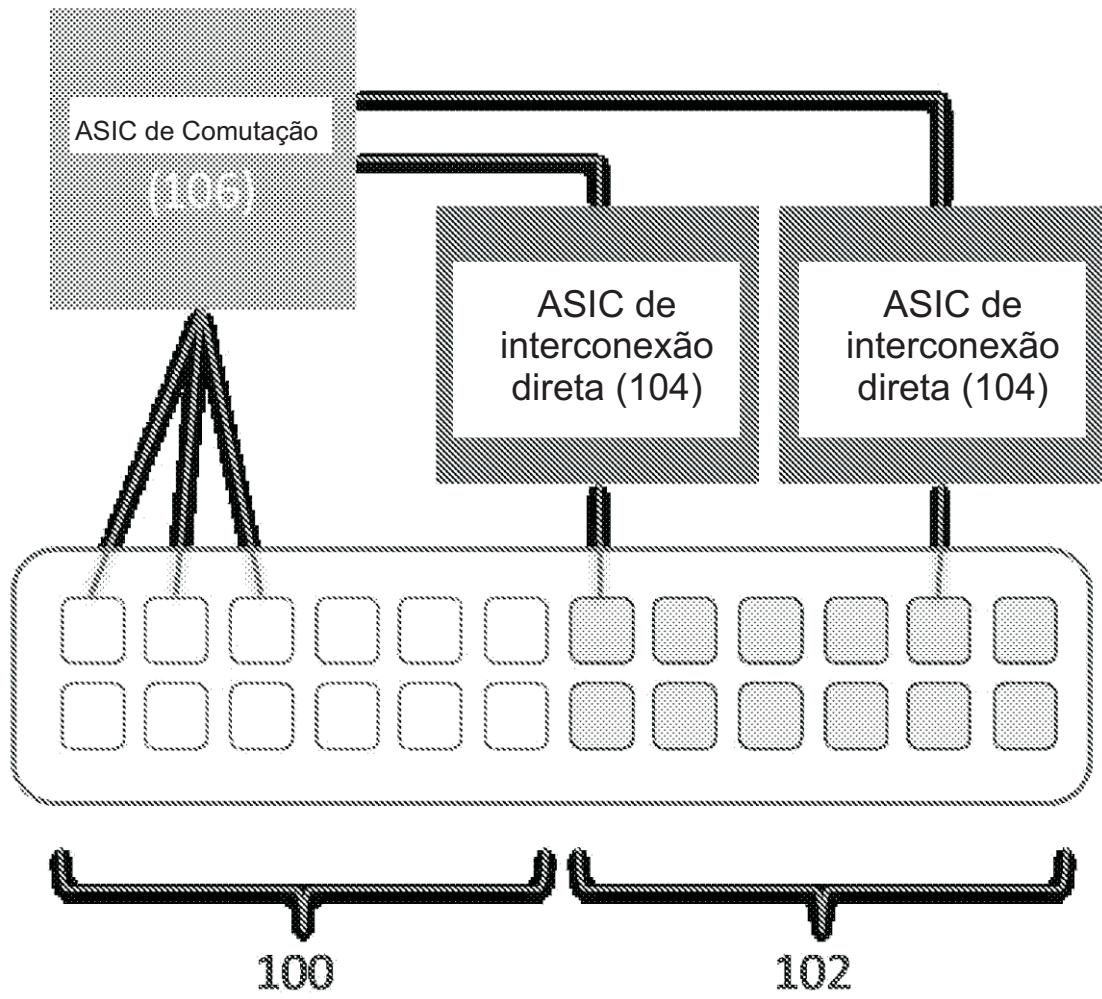


FIG. 2

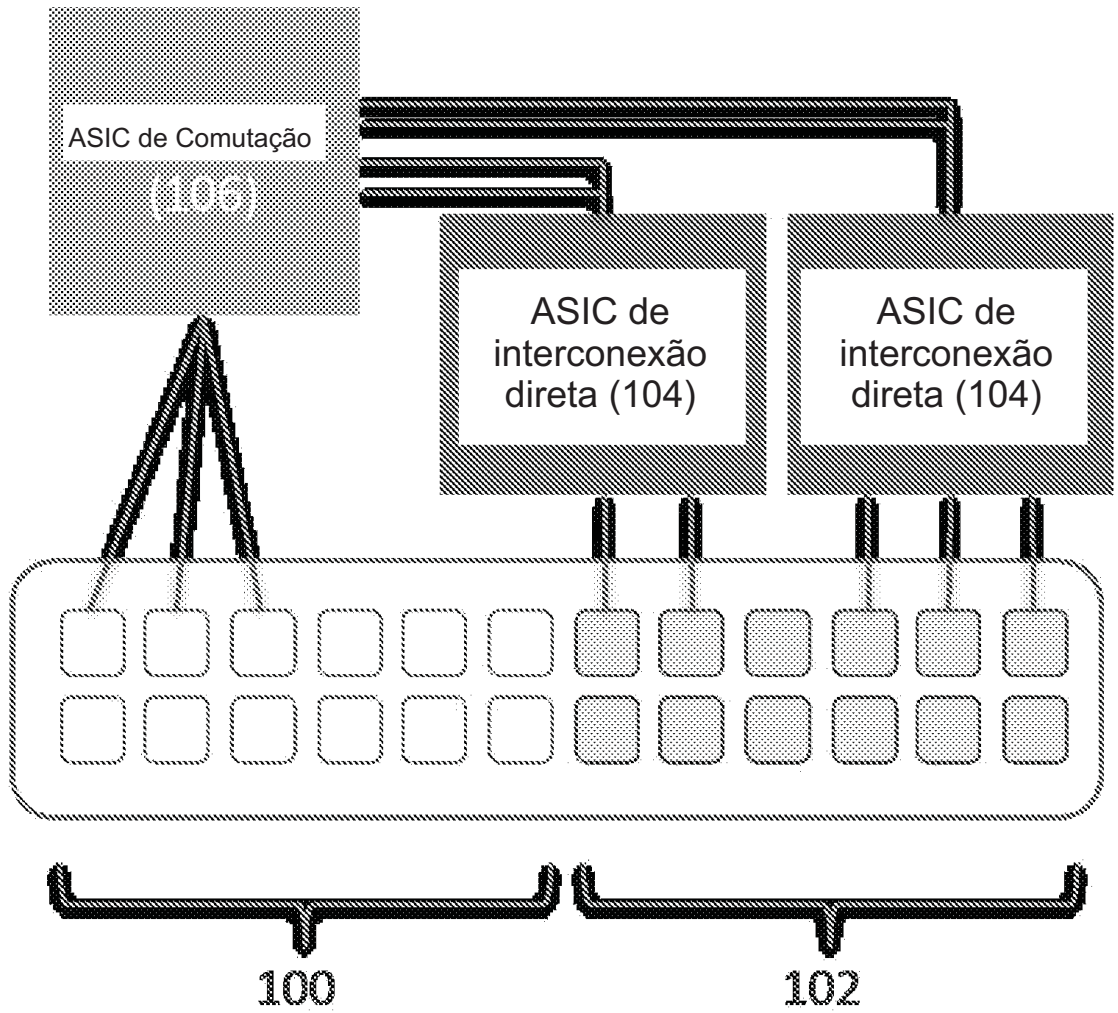


FIG. 2b

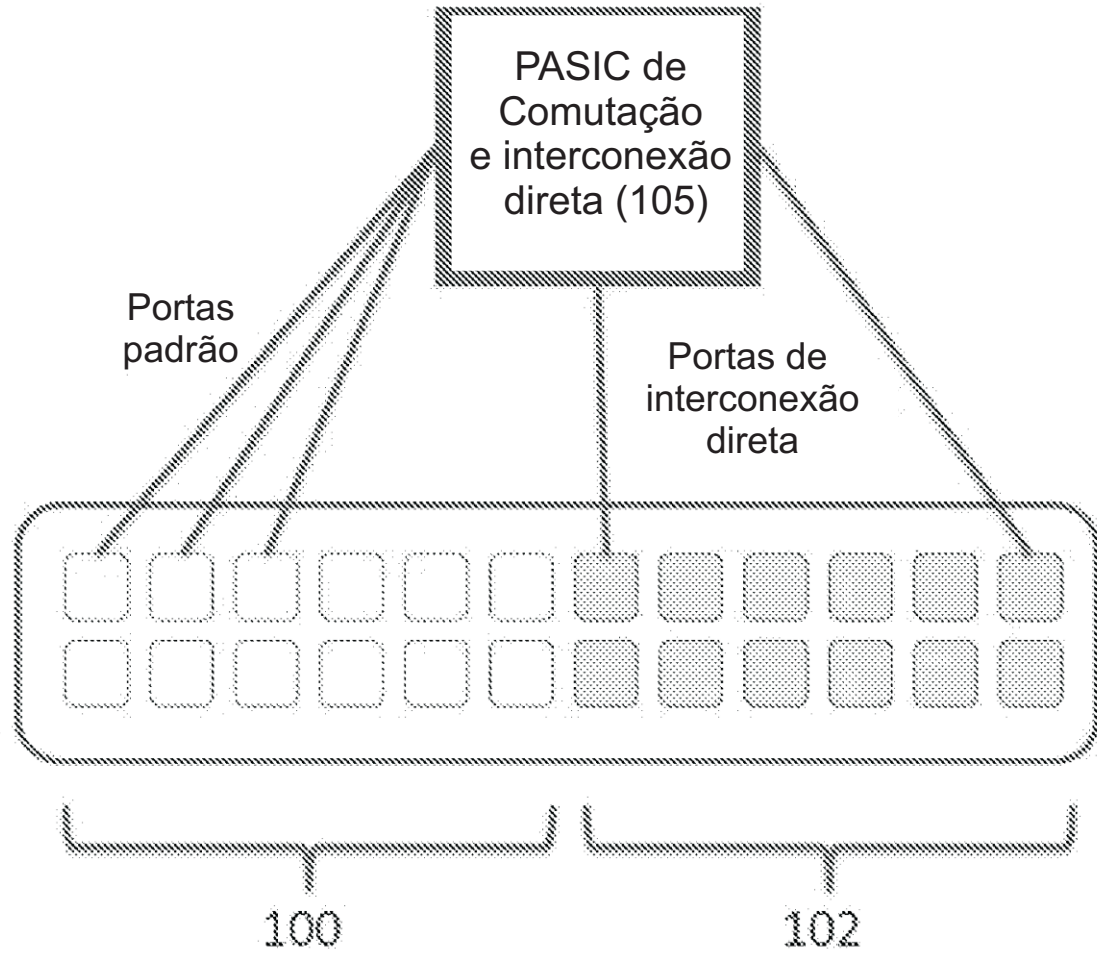


FIG. 2c

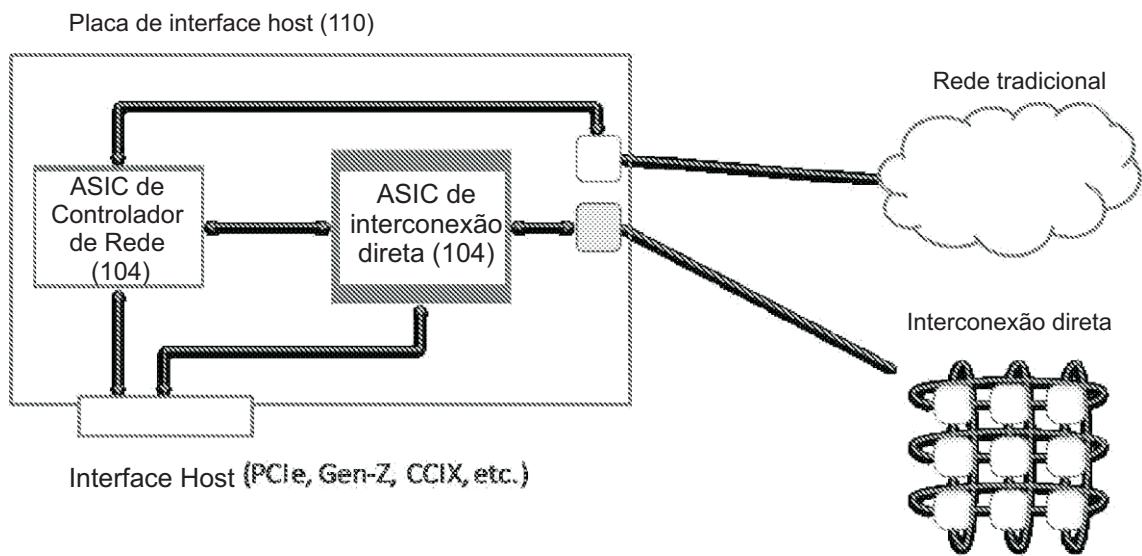


FIG. 2d

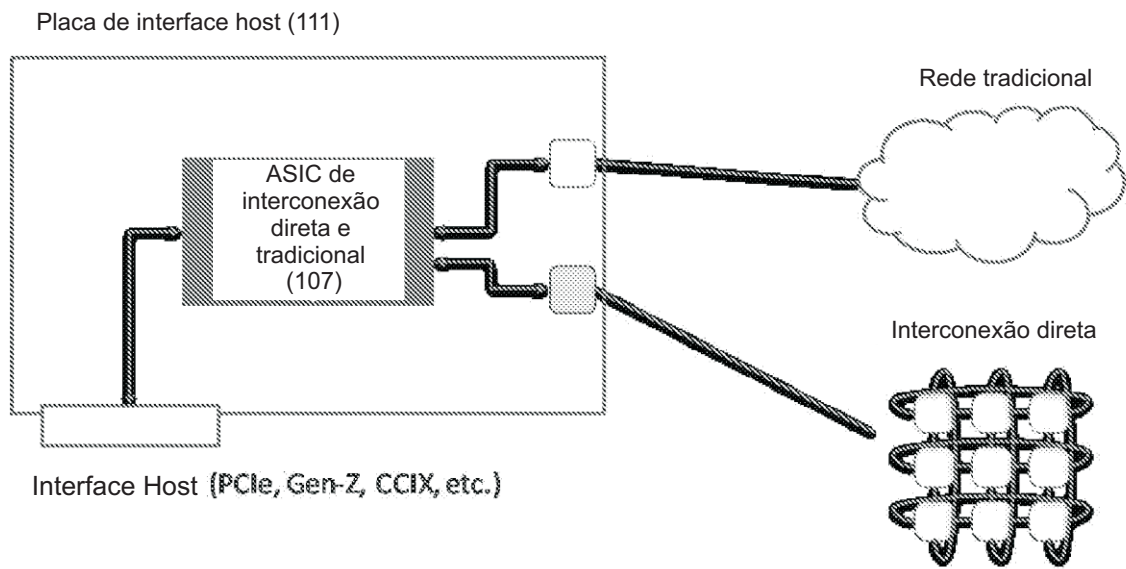


FIG. 2e

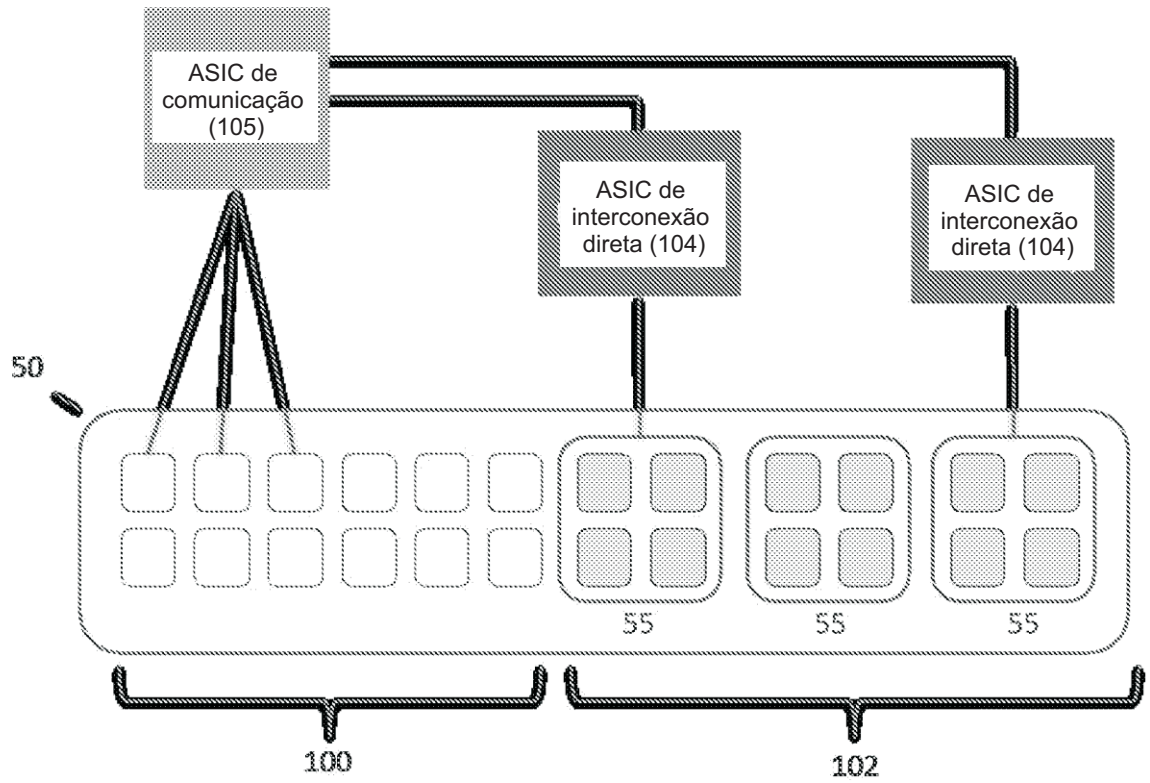


FIG. 2f

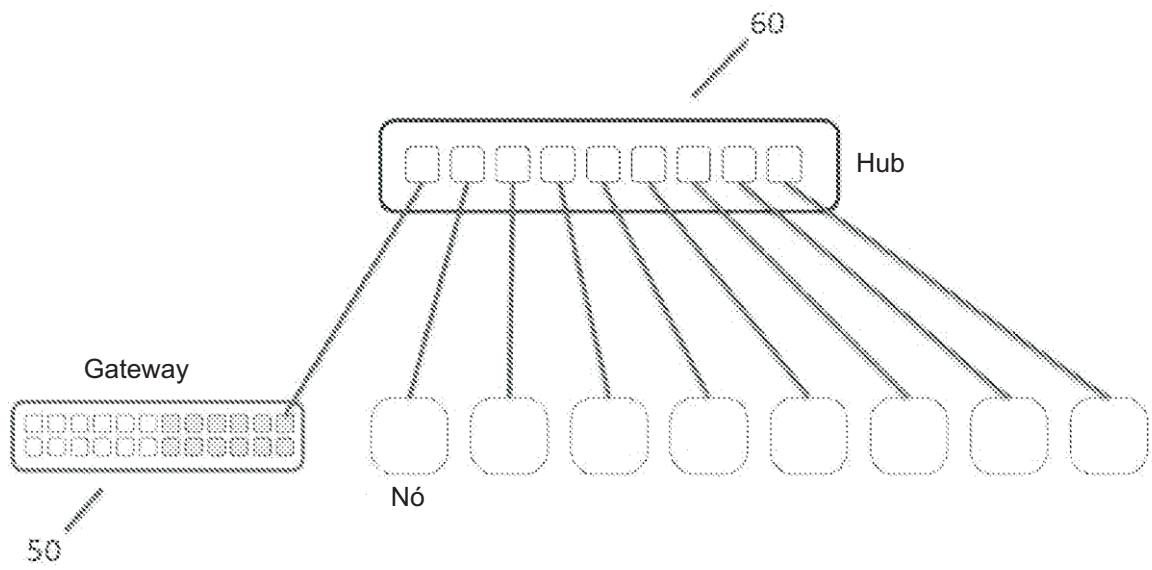


FIG. 3

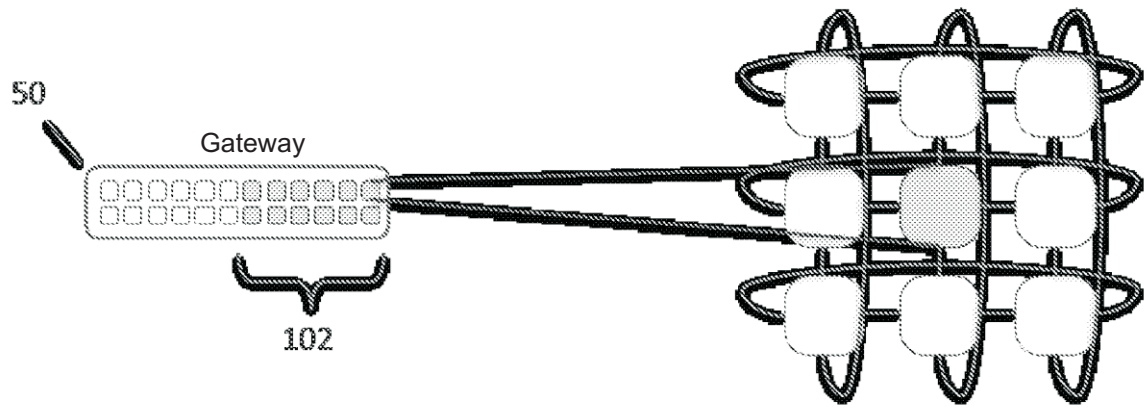


FIG. 4

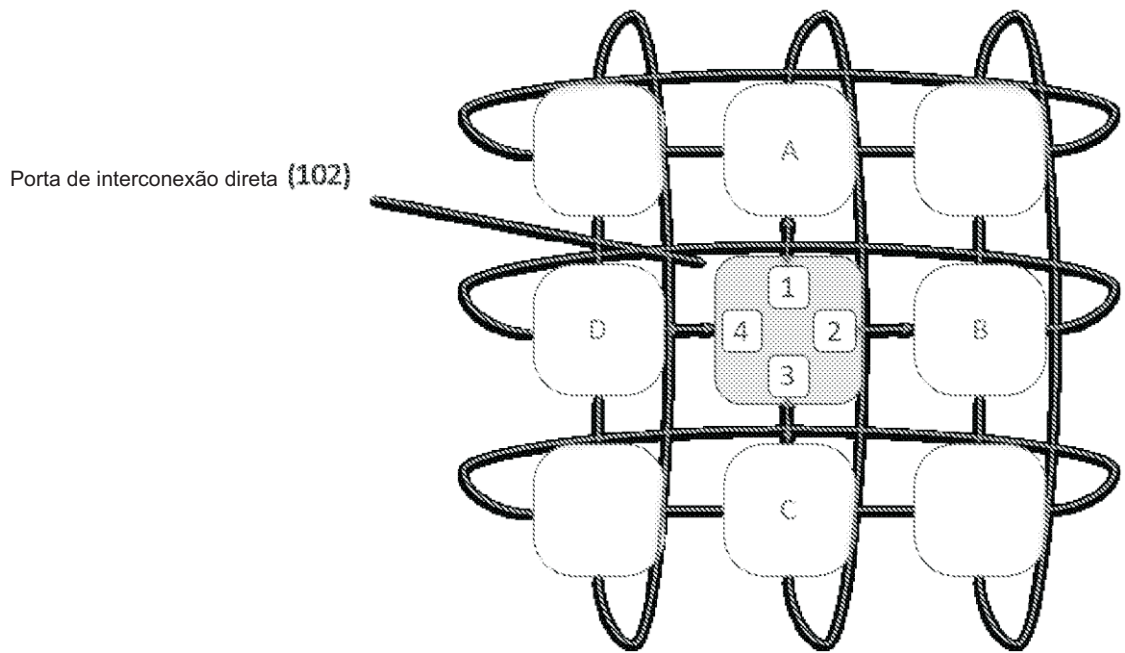


FIG. 4b

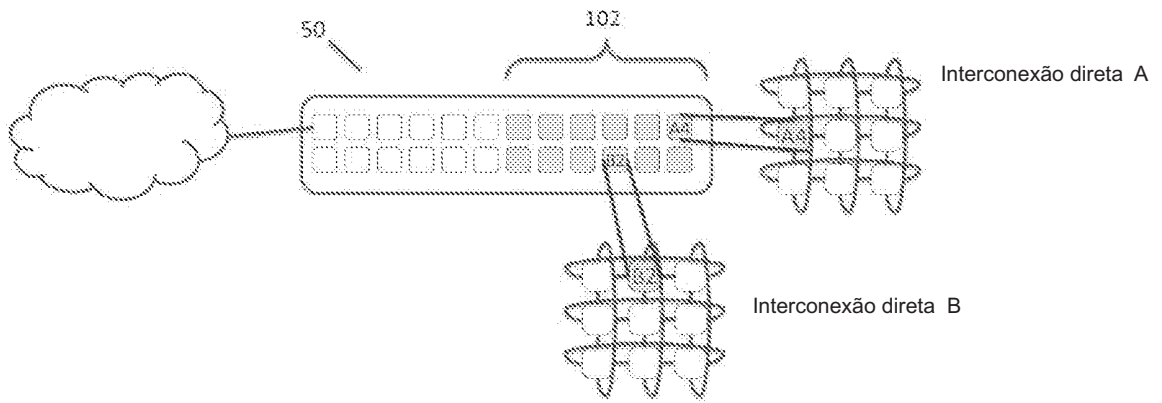


FIG. 5

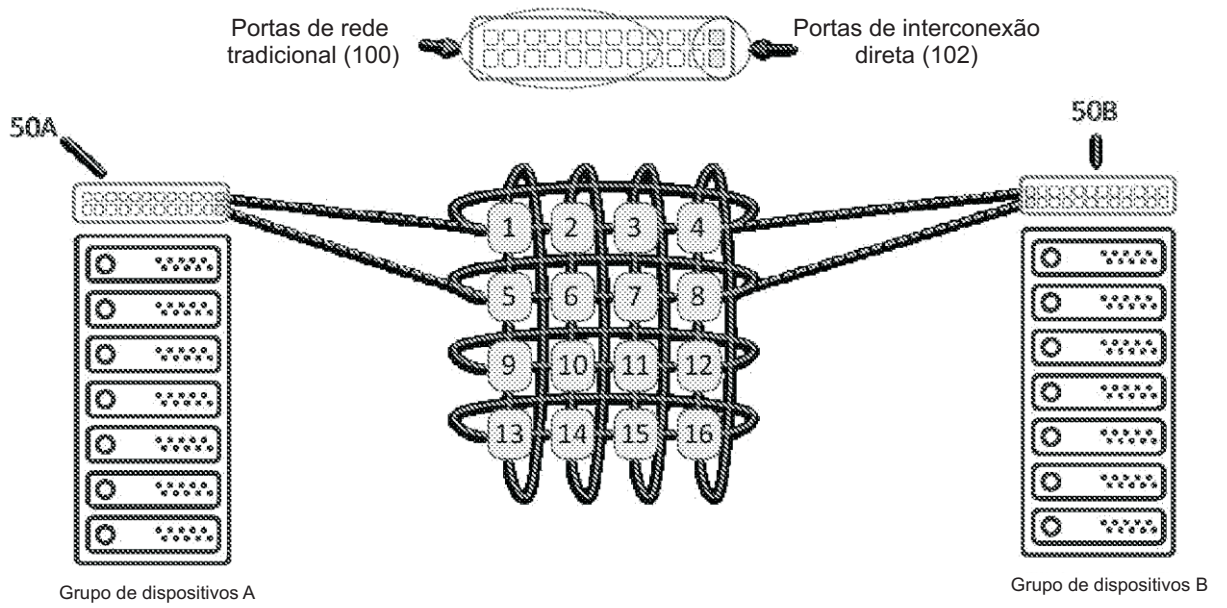


FIG. 6

Rede comutada tradicional

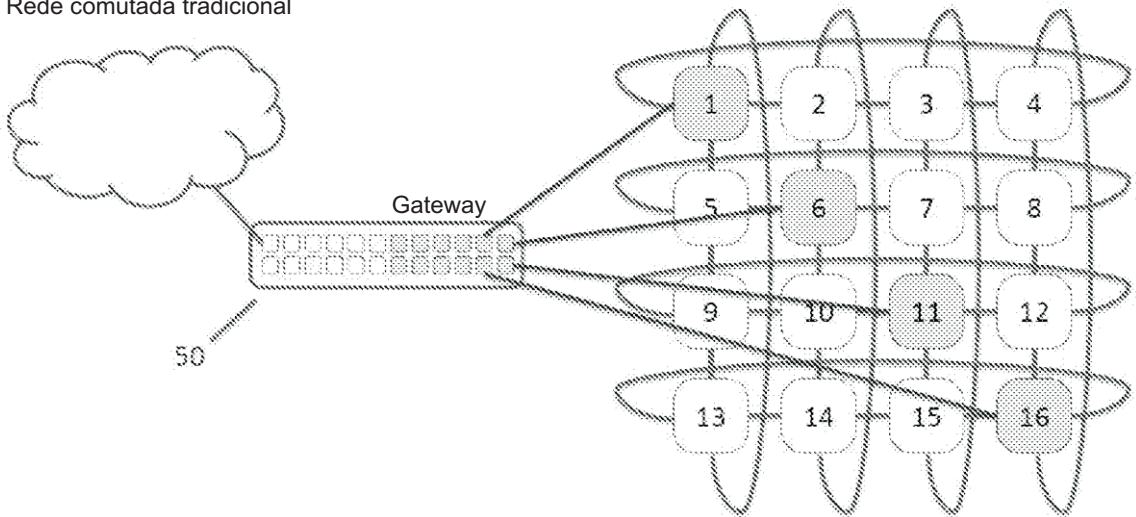


FIG. 7

- Nós 1, 6, 11 e 16 são portas no mesmo dispositivo de gateway
- Nó 5 envia pacote de broadcast
- Nós 1, 6, 11 e 16 recebem pacote de broadcast
- Gateway determina que os nós 1 e 6 são os nós mais próximos do nó 5
- Gateway escolhe o nó 1 como nó ótimo
- O nó 1 está associado com esse tráfego e lidará com tráfego futuro relacionado com o pacote de broadcast

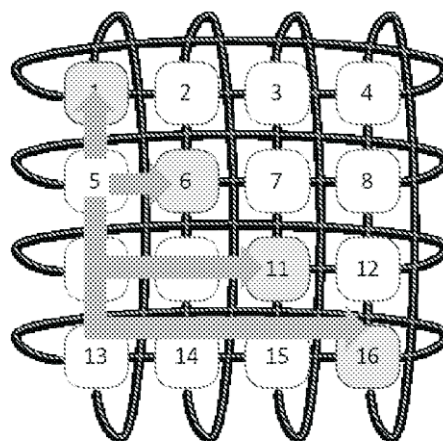


FIG. 8

ASIC de interconexão direta

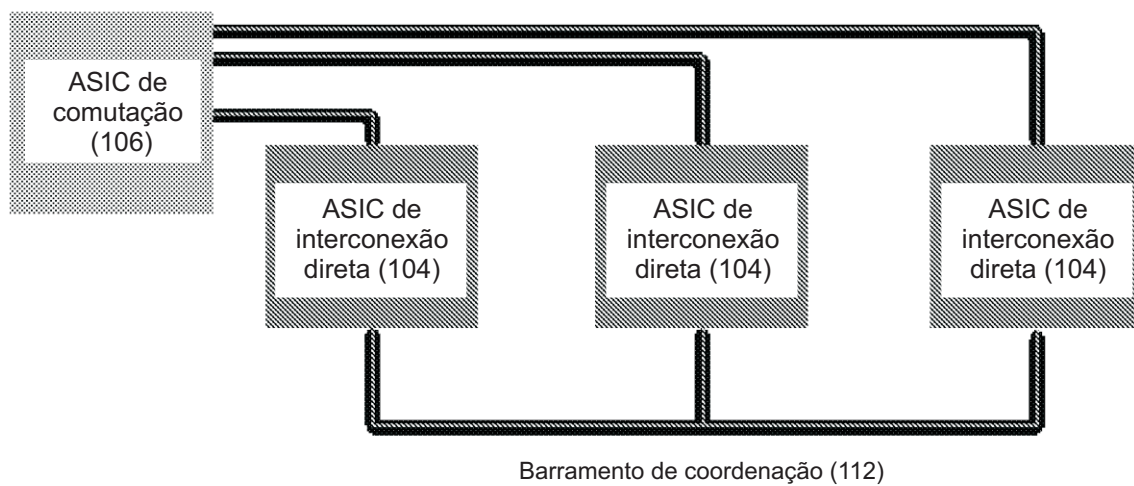


FIG. 8b

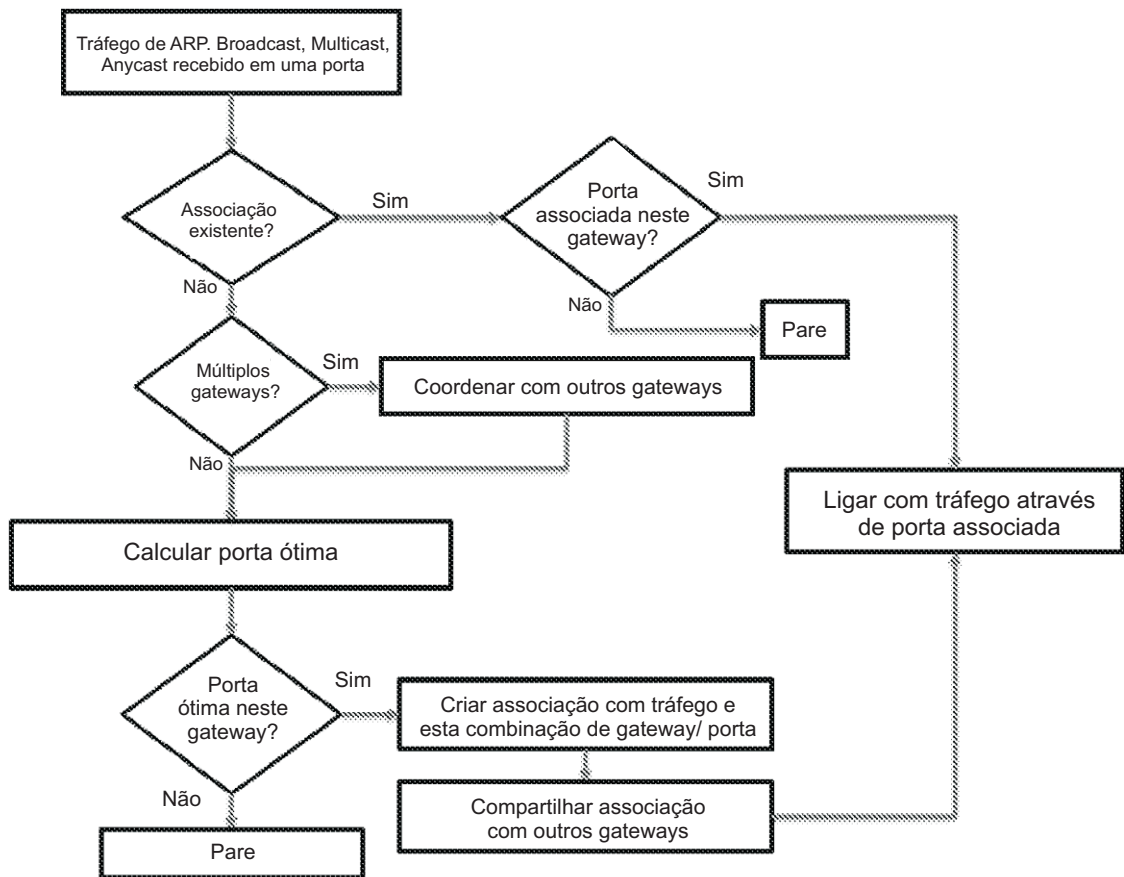


FIG. 9

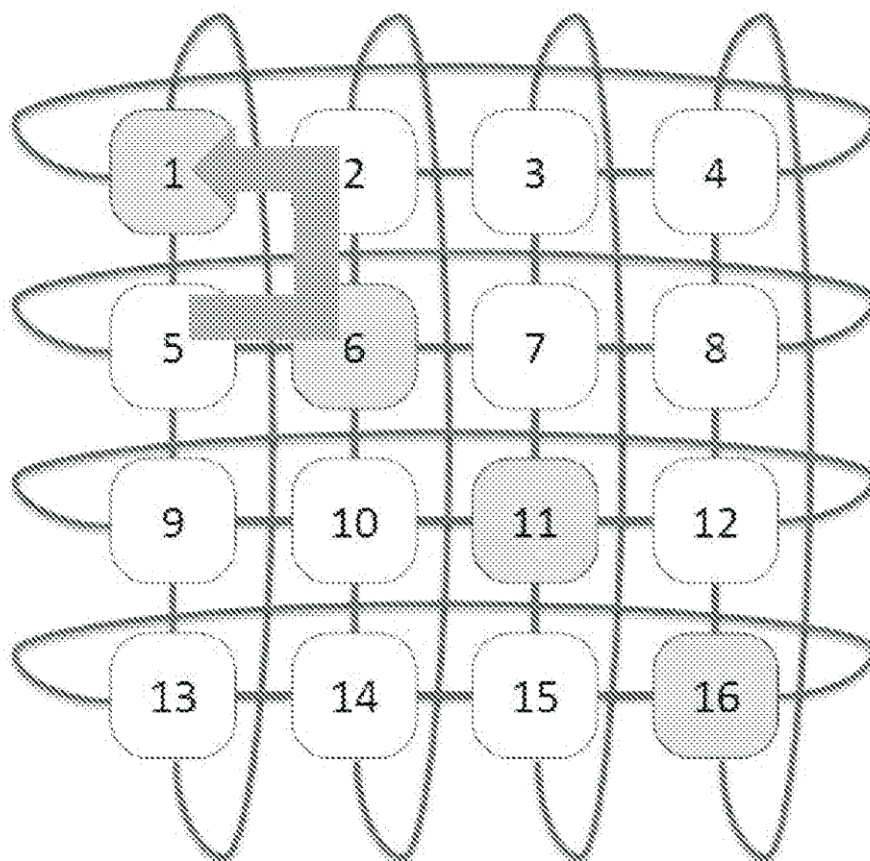


FIG. 10

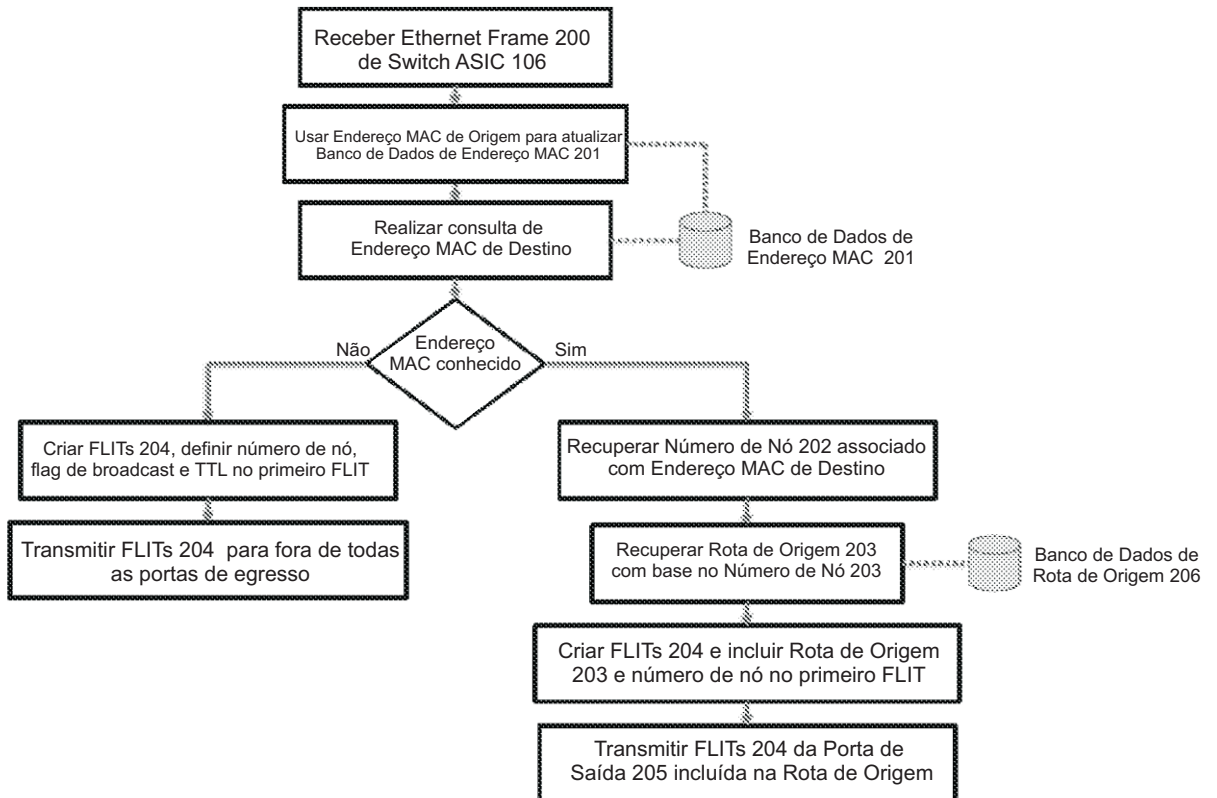


FIG. 11

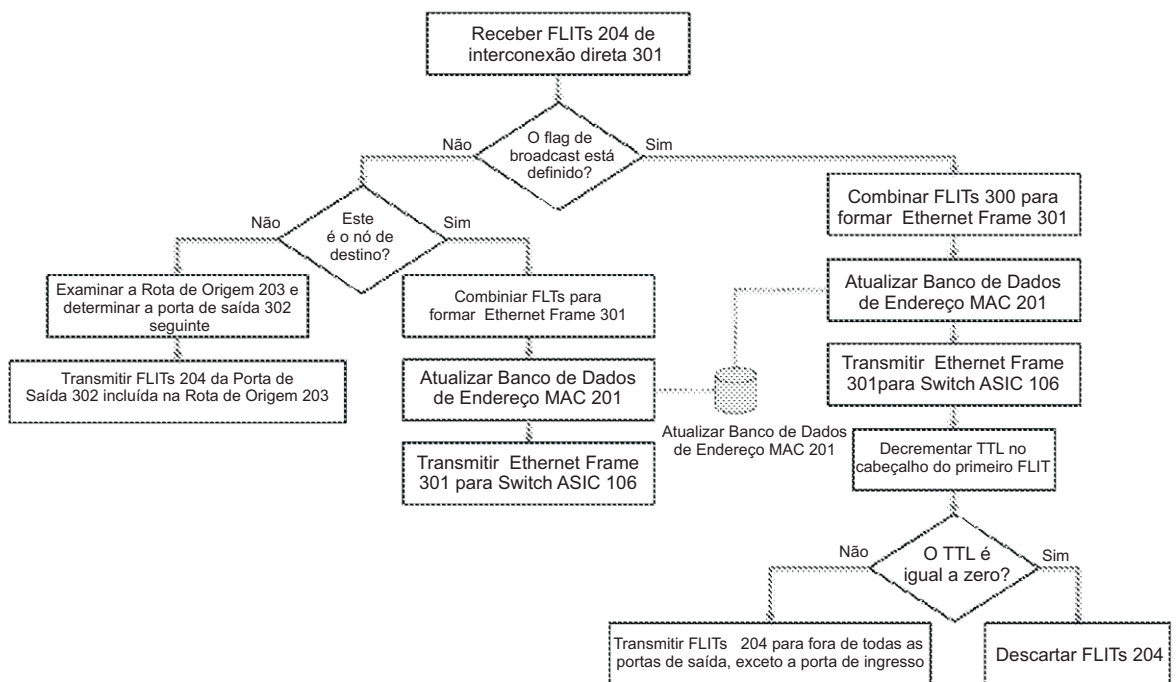


FIG. 12

RESUMO

Patente de Invenção: **"GATEWAY DE INTERCONEXÃO DIRETA"**.

Um dispositivo de gateway de rede dedicado que é capaz de ligar em ponte, alternar ou rotear o tráfego de rede entre a rede tradicional e as redes de interconexão direta, compreendendo: um primeiro conjunto de uma ou mais portas de rede tradicionais com um único link por porta, geralmente compreendendo um ou mais conectores SFP + QSFP e QSFP +, tais portas sendo conectadas a comutadores ou dispositivos que formam uma rede tradicional; e um segundo conjunto de uma ou mais portas de interconexão direta com um alto número de links por porta (dois ou mais) geralmente compreendendo um ou mais conectores MXC, MTP e MTO, tais portas sendo conectadas a uma rede de interconexão direta.