



República Federativa do Brasil
Ministério do Desenvolvimento, Indústria
e Comércio Exterior
Instituto Nacional de Propriedade Industrial

(21) PI0615816-1 A2



* B R P I 0 6 1 5 8 1 6 A 2 *

(22) Data de Depósito: 11/09/2006
(43) Data da Publicação: 24/05/2011
(RPI 2107)

(51) Int.Cl.:
H04L 12/28 2006.01
H04L 29/10 2006.01
G06F 15/16 2006.01

(54) Título: COMUNICAÇÕES TOLERANTE A FALTA
EM REDES ROTEADAS

(30) Prioridade Unionista: 16/12/2005 US 11/275.185,
12/09/2005 US 60/716.122, 12/09/2005 US 60/716.122, 16/12/2005
US 11/275.185

(73) Titular(es): MICROSOFT CORPORATION

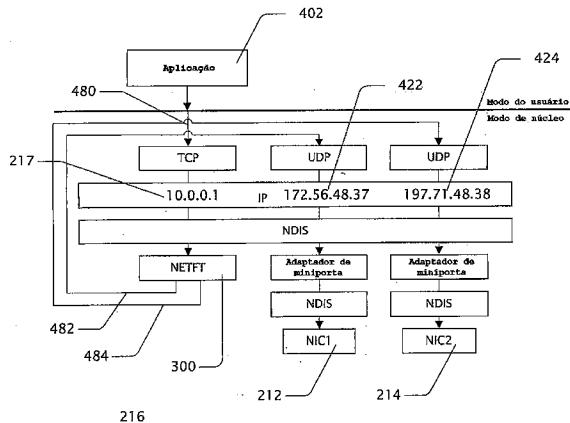
(72) Inventor(es): DAVID A. DION, MICHAEL T. MASSA, RUDOLF
OPAVSKY

(74) Procurador(es): ALEXANDRE FERREIRA

(86) Pedido Internacional: PCT US2006035497 de 11/09/2006

(87) Publicação Internacional: WO 2007/033179de 22/03/2007

(57) Resumo: COMUNICAÇÕES TOLERANTE A FALTA EM
REDES ROTEADAS. Um método para prover comunicações de rede
tolerante à falta entre uma pluralidade de nós para uma aplicação,
incluindo prover uma pluralidade de caminhos de comunicações
iniciais sobre uma pluralidade de redes unidas entre a pluralidade de
nós, receber um pacote de dados em um nó de envio da aplicação, o
nó de envio sendo um da pluralidade de nós, o pacote de dados sendo
endereçado pela aplicação para um endereço em um da pluralidade de
nós e selecionar um primeiro caminho selecionado para o pacote de
dados dentre a pluralidade de caminhos de comunicações iniciais onde
o primeiro caminho selecionado é um caminho preferido.



"COMUNICAÇÕES TOLERANTE A FALTA EM REDES ROTEADAS"

ANTECEDENTES

Em um ambiente de rede de computador, múltiplos nós podem se comunicar através de uma rede. Caso a rede experimente uma falha, a comunicação entre os nós pode ser interrompida.

SUMÁRIO

O seguinte apresenta um resumo simplificado da revelação a fim de prover um entendimento básico para o leitor. Esse resumo não é uma visão geral extensiva da revelação e ele não identifica elementos essenciais ou críticos da invenção ou delineia o escopo da invenção. Sua única finalidade é apresentar alguns conceitos revelados aqui em uma forma simplificada como um prenúncio para a descrição mais detalhada que é apresentada a seguir.

Os exemplos seguintes provêm comunicação de rede de computador tolerante a falta via arquiteturas de pilha de rede únicas requerendo mínima consideração pelo software de aplicação que opera nos nós de rede.

Muitos dos aspectos auxiliares serão mais facilmente verificados à medida que eles se tornam melhor entendidos por referência à descrição detalhada seguinte considerada em conjunto com os desenhos acompanhantes.

DESCRIÇÃO DOS DESENHOS

A presente descrição será entendida melhor a partir da descrição detalhada seguinte lida à luz dos desenhos acompanhantes, nos quais:

A figura 1 é um diagrama de blocos mostrando uma

arquitetura de pilha de rede exemplar.

A figura 2 é um diagrama de blocos mostrando um ambiente de computação de rede incluindo dois nós exemplares acoplados via duas redes.

5 A figura 3 é um diagrama de blocos mostrando um acionador de comunicações tolerante à falta exemplar, NETFT.

A figura 4 é um diagrama de blocos mostrando uma arquitetura de comunicações tolerante à falta exemplar incluindo NETFT e uma aplicação.

10 A figura 5 é um diagrama de fluxo mostrando os dados fluindo através de um ambiente de comunicações tolerante à falta incluindo um nó de fonte e um nó de destino acoplado via Trajetória A através da rede 1 e Trajetória B através da rede 2.

15 A figura 6 é um diagrama de fluxo mostrando os dados fluindo através do ambiente de comunicações tolerante à falta mostrado na figura 5 com a adição de várias falhas de comunicações possíveis.

20 A figura 7 é um diagrama de blocos mostrando um outro exemplo de um acionador de comunicações tolerante à falta, NETFT.

A figura 8 é um diagrama de blocos mostrando uma arquitetura de comunicações tolerante à falta exemplar incluindo NETFT e uma aplicação.

25 A figura 9 é um diagrama de fluxo mostrando os dados fluindo através de um ambiente de comunicações tolerante à falta incluindo um nó de fonte e um nó de destino acoplados via Trajetória A através da rede 1 e Trajetória B atra-

vés da rede 2.

A figura 10 é um diagrama de fluxo mostrando os dados fluindo através do ambiente de comunicações tolerante à falta mostrado na figura 9 com a adição de várias falhas de comunicações possíveis.

A figura 11 é um diagrama de blocos mostrando um ambiente de computação exemplar no qual a tecnologia descrita acima pode ser realizada.

Numerais de referência semelhantes são usados para indicar partes semelhantes nos desenhos acompanhantes.

DESCRÍÇÃO DETALHADA

A descrição detalhada provida abaixo em conjunto com os desenhos anexos é planejada como uma descrição dos presentes exemplos e não é planejada para representar as únicas formas nas quais os presentes exemplos podem ser construídos ou utilizados. A descrição apresenta as funções dos exemplos e a seqüência de etapas para construir e operar os exemplos. Entretanto, as mesmas funções e seqüências ou equivalentes podem ser executadas por exemplos diferentes.

Embora os presentes exemplos sejam descritos e ilustrados aqui como sendo realizados em um sistema de computação e rede, o sistema descrito é provido como um exemplo e não uma limitação. Como aqueles versados na técnica verificarão, os presentes exemplos são adequados para aplicação em uma variedade de tipos diferentes de sistemas de computação e rede.

A figura 1 é um diagrama de blocos mostrando uma arquitetura de pilha de rede exemplar 100. Uma pilha de rede

("pilha") geralmente se junta com aplicações de software via interfaces de pilha de rede e/ou outras interfaces para prover funcionalidade de comunicações de rede para aplicações. Uma aplicação é tipicamente dita estar em (ou unida a) o 5 "topo" da pilha. Uma rede é tipicamente dita estar em (ou unida a) a "base" da pilha. Vários elementos de uma pilha de rede podem ser citados como em ou perto do topo ou base da pilha, ou mais alto ou mais baixo na pilha em relação um ao outro. Por exemplo, na figura 1, o acionador de protocolo 10 130 está mais alto na pilha do que a NIC 180 que é mostrada na base da pilha nessa figura particular. Várias representações de uma pilha de rede podem ou não incluir alguns elementos de pilha ou podem agrupar, ordenar ou nomear os elementos em várias maneiras, dependendo da finalidade ou foco 15 da representação, como entendido por aqueles versados na técnica.

O termo "acionador" como usado aqui se refere a um programa de controle ou semelhante que possibilita que um nó opere com um dispositivo particular, tais como uma impressora, placa de interface de rede ou outro subsistema de computador, ou opere com um ou mais programas tais como pilhas de rede, acionadores de protocolo e/ou outro software de computador ou firmware ou semelhantes. Por exemplo, um acionador de protocolo tipicamente opera com uma pilha de rede.

25 Uma aplicação pode passar um pacote de dados para uma pilha destinada para uma aplicação operando em um outro nó. Nesse caso, é dito que os dados fluem "para baixo" da pilha e são enviados para fora através de uma rede. Os dados

recebidos por um nó são ditos fluir "para cima" da pilha até que eles alcançam a aplicação destinada. Tais sistemas de rede são bem conhecidos para aqueles versados na técnica.

Em um exemplo, uma pilha é baseada na especificação da interface do acionador de rede ("NDIS") que define uma interface de programação de aplicação ("API") padrão para placas de interface de rede ("NICs"), tal como NIC 180, e abstrai o hardware da rede dos acionadores de rede. NDIS também especifica uma interface padrão entre acionadores de rede em camadas, dessa maneira abstraindo acionadores de nível inferior que gerenciam hardware, tal como um acionador de miniporta, dos acionadores de nível superior, tal como acionadores de protocolo. Múltiplos acionadores de protocolo que se conformam com NDIS podem coexistir em um único nó.

Também, se um nó inclui múltiplas NICs, talvez porque ele está conectado em mais do que uma rede, a NDIS roteia o tráfego da rede para a NIC apropriada via seu acionador associado como indicado pelo tráfego. Uma ilustração da NDIS é mostrada na figura 1. Outros padrões de pilha de rede, tecnologias e/ou arquiteturas, tal como a interface de ligação de dados aberta ("ODI"), a interface do provedor da ligação de dados ("DLPI"), a interface do acionador uniforme ("UDI") ou outras tecnologias, podem ser usados com os exemplos seguintes também com modificações apropriadas, como seria entendido por aqueles versados na técnica. Como uma questão de conveniência, a terminologia NDIS e NDIS é usada com exemplos por toda essa descrição, mas outros padrões, tecnologias e/ou arquiteturas podem ser usados em todos esses exem-

plos com modificações apropriadas, a menos que de outra forma mencionado.

Como mostrado na figura 1, acoplado na NIC 180 via NDIS 120 está o acionador de miniporta 160. Um acionador de miniporta tipicamente interage com NDIS via uma interface de miniporta da NDIS 162. O acionador de miniporta 160 pode ser associado com a NIC 180 e pode gerenciar suas operações, incluindo enviar e receber dados através da NIC. O acionador de miniporta 160 tipicamente faz interface com acionadores de nível superior, tais como acionador intermediário 140 e acionador de protocolo 130. Um acionador de miniporta é considerado um acionador de NIC. Miniportas de NIC geralmente executam essas operações específicas do hardware necessárias para gerenciar uma NIC particular com funcionalidade comum ou independente da NIC provido pela NDIS. Um nó pode incluir múltiplas NICs com cada NIC geralmente tendo um acionador de NIC associado. Alguns exemplos nessa descrição descrevem o uso de acionadores de miniportas, mas, como será entendido por aqueles versados na técnica, qualquer tipo de acionador de NIC ou semelhante pode ser usado nesses exemplos, a menos que de outra forma mencionado.

O acionador de protocolo ou transporte 130 se junta na NDIS 120 via uma interface de protocolo da NDIS 134. Acionadores de protocolo ou acionadores de protocolo de transporte geralmente provêem a funcionalidade para criar, enviar e receber pacotes de dados que são enviados de um nó para outro através da pilha de rede e através de uma rede. Como conhecido para aqueles versados na técnica, um protoco-

lo de transporte de entrega confiável ou garantido comum pode ser TCP/IP (protocolo de controle de transmissão / protocolo da Internet). UDP (protocolo de datagrama do usuário) através de IP pode ser um protocolo de entrega comum não confiável ou não garantido. TCP, UDP e/ou outros protocolos, tal como IPX/SPX (troca de pacotes pela Internet / troca de pacote seqüenciada) podem ser usados com os exemplos seguintes a menos que de outra forma mencionado.

Os acionadores intermediários da NDIS ("IM") 140 são mostrados entre os acionadores de protocolo 130 e as miniportas da NIC da NDIS 160 na figura 1. Para acionadores de protocolo os acionadores IM aparecam ser miniportas da NDIS enquanto para acionadores da NIC eles se parecem com acionadores de protocolo. Os pacotes de dados fluindo para cima ou para baixo da pilha de rede passam através do acionador IM 140 que pode ignorar, inspecionar, filtrar, enviar, redirecionar e/ou modificar os pacotes de dados. Um acionador intermediário 140 pode também ser conhecido como um acionador de filtro.

A figura 2 é um diagrama de blocos mostrando um ambiente de computação de rede 200 incluindo dois nós exemplares 210 e 260 acoplados via duas redes 202 e 282. Os nós 210 e 260, cada um, podem ser computadores pessoais ("PCs"), computadores clientes, servidores, hospedeiros, laptops, dispositivos portáteis, dispositivos eletrônicos do consumidor ou qualquer um de vários outros tipos de dispositivos de computação ou processamento, máquinas ou sistemas. Um exemplo não limitador de um tipo de sistema de computação é des-

crito em detalhes abaixo com relação à figura 11. Os círculos 212, 214, 262 e 264 representam NICs associadas com seus nós respectivos. Um exemplo não limitador de um tipo de NIC é também descrito abaixo com relação à figura 11 como o adaptador de rede 1113.

Como usado aqui, o termo nó se refere a qualquer sistema de computação, dispositivo ou processo que é unicamente endereçável, ou de outra forma unicamente identificável, em uma rede (por exemplo, rede 202) e que é operável para se comunicar com outros nós na rede. Por exemplo, e sem limitação, um nó pode ser um computador pessoal, um computador servidor, um dispositivo de mão ou laptop, um dispositivo de mesa digitalizadora, um sistema de multiprocessador, um sistema com base em microprocessador, um conversor de sinal de freqüência, um dispositivo eletrônico do consumidor, um PC de rede, um minicomputador, um computador de grande porte ou semelhantes. Um exemplo não limitador de um nó 210, na forma de um sistema de computação, é apresentado abaixo com relação à figura 11.

As redes 202 e 282 podem ser a mesma rede, podem existir nas mesmas subredes ou diferentes, podem ser lógica ou fisicamente acopladas ou isoladas uma da outra, podem usar tecnologias de rede similares ou diferentes, etc. Em particular, as redes 202 e 282 podem ser redes roteadas, isto é, redes incluindo roteadores que enviam pacotes de protocolo roteável. Protocolos roteáveis são tipicamente considerados protocolos de comunicações usados para rotear dados de uma rede para outra. Um exemplo de um protocolo roteável

é TCP/IP. O envio de um pacote de dados em um modo roteável implica usar um protocolo de transporte roteável para formar e/ou enviar o pacote de dados. Aqueles versados na técnica estarão familiarizados com protocolos roteáveis e topologias, sistemas e arquiteturas de rede de roteamento.

Em um exemplo, as redes 202 e 282 podem ser independentes uma da outra tal que se existe um problema ou falha com uma rede é improvável afetar o estado operacional da outra. Em outros exemplos, três ou mais redes podem ser usadas. Nos exemplos onde maiores graus de tolerância à falta são desejados, um maior número de redes junto com a conectividade associada de nós para essas redes, incluindo um número similar de NICs instaladas em um nó, pode ser utilizado.

A NIC 212, associada com o nó 210, é mostrada com um endereço exemplar de 172.56.48.37 e é acoplada na rede 1 202. A NIC 214, também associada com o nó 210, é mostrada com um endereço exemplar de 197.71.48.38 e é acoplada na rede 2 282. A NIC 262, associada com o nó 260, é mostrada com um endereço exemplar de 172.56.48.38 e é também acoplada na rede 1 202. A NIC 264, também associada com o nó 260, é mostrada com um endereço exemplar de 197.71.48.39 e é também acoplada na rede 2 282. Esses endereços podem, na prática, ser endereços IPv4 ou IPv6 ou semelhantes ou qualquer outro tipo de endereço de rede tipicamente relacionado com o protocolo sendo usado.

Cada nó pode incluir uma ou mais NICs. As setas 201 e 203, também mostradas na figura 11 como seta 1114, representam uma primeira rota ou caminho de comunicações

("Trajetória A") sobre a rede 1 202 entre os nós 210 e 260. As setas 282 e 283 representam uma segunda rota ou caminho de comunicações ("Trajetória B") sobre a rede 2 282 entre os nós 210 e 260. Na prática, podem existir um ou mais caminhos sobre uma ou mais redes entre os dois ou mais nós no ambiente 200. O termo "caminho" como usado aqui é definido como uma rota de comunicações, ou ligação de comunicações, entre nós em uma rede. Uma tal rota ou ligação pode ser dinâmica em que a rota exata entre os nós pode mudar com o tempo.

Os blocos 216 e 266 representam uma aplicação e uma pilha de rede, incluindo um acionador de comunicações tolerante à falta ("FT"), provido em cada um dos nós 210 e 260. O acionador FT do bloco 216 é mostrado com um endereço exemplar de 10.0.0.1 e o acionador FT do bloco 266 é mostrado com um endereço exemplar de 10.0.0.2. Esses endereços são tipicamente considerados endereços virtuais. Esses endereços podem ser endereços IPv4 ou IPv6 ou semelhantes ou qualquer outro tipo de endereço de rede ou de comunicações. Acionadores FT podem ou não ter endereços virtuais como mostrado nos vários exemplos abaixo.

Uma pilha de rede tolerante à falta é uma pilha de rede incluindo um acionador FT, tal como NETFT descrito abaixo em conjunto com a figura 3, ou semelhante. Um acionador FT, tal como NETFT, operando em combinação com uma pilha de rede tipicamente permite que os nós se comuniquem via uma ou mais trajetórias de comunicações, tais como Trajetória A e Trajetória B, sobre uma ou mais redes. Caso qualquer uma dessas trajetórias de comunicações falhe, os nós podem con-

tinuar a comunicação dado pelo menos um caminho operacional.

Uma tal falha de caminho pode resultar da falha de uma NIC ou falha de qualquer elemento de um caminho, incluindo conexões, cabeamento ou outros meios de comunicações (incluindo 5 radiofreqüência ("RF") ou infravermelho ("IR") e semelhantes), roteadores, bocas de conexão, chaves, barreiras de proteção, provedores de serviço da Internet ("ISPs"), falha de força para qualquer nó, dispositivo ou sistema da rede ou semelhante.

10 Em um exemplo, uma falha de comunicações pode resultar em um evento conecte e use ("PnP"). Um evento PnP pode indicar a remoção de uma NIC do seu nó ou em uma mudança de sentido dos meios. Uma desconexão do sentido dos meios, por exemplo, tipicamente resulta de uma falha que faz com 15 que a NIC perca o sinal ou portadora nos meios de rede, tais como um cabo de rede, ligação RF ou IR ou semelhante. Uma desconexão do sentido dos meios pode ser causada pela desconexão do cabo de rede ou portadora da NIC ou desligamento da outra extremidade do cabo (uma boca de conexão ou chave, por 20 exemplo). Uma conexão do sentido dos meios é tipicamente o oposto, tais como a reconexão do cabo, religamento na boca de conexão ou chave ou semelhantes. Esses tipos de eventos, também conhecidos como eventos de conectividade, são geralmente eventos locais já que eles ocorrem em ou estão próximos 25 do próprio nó. Tais eventos de conectividade local tipicamente resultam em uma indicação de evento, tal como um evento PnP ou semelhante, em um nó.

Em um outro exemplo, uma falha de comunicações po-

de ser detectada pelo uso de pacotes de pulsos enviados entre os nós. Falha de um tal pacote de pulsos pode indicar falha de um caminho entre nós. Pacotes de pulsos tendem a ser marcados tal que o acionador FT pode detectá-los com a 5 recepção e removê-los para o fluxo de pacote sendo passado para cima da pilha de rede. Em um exemplo, pacotes de pulsos podem ser realizados usando protocolo de controle de rota ("RCP") pela formação dos pacotes de RCP. Tais pacotes de pulsos podem ser usados para validar o estado operacional de 10 extremidade a extremidade de um caminho. Isto é, pelo envio de um pacote de pulsos do nó 210 através da Trajetória A para o nó 260 e pelo nó 210 recebendo uma resposta para o pacote de pulsos enviado do nó 260, é geralmente considerado que a Trajetória A está operacional de extremidade a extremidade. Caso o pulso falhe (nenhuma resposta de pulso recebida em resposta ao pulso enviado), uma tal falha pode indicar que a Trajetória A não está operacional, talvez devido à 15 falha de algum elemento da rede 202 tais como um roteador, chave, conexão ou semelhante, ou devido à falha do próprio nó alvo. Em particular, o nó 210 pode ter uma NIC operacional 212 e sentido dos meios válido, indicando que ele está 20 apropriadamente conectado na rede, mas ainda pode detectar uma falha de pulso devido a alguma falha de rede ou sistema para baixo da linha.

25 A figura 3 é um diagrama de blocos mostrando um acionador de comunicações tolerante à falta exemplar, NETFT

300. O NETFT 300 pode ser realizado como um acionador de miniporta da NDIS (figura 1, 160) para uso com uma pilha de rede da NDIS e para prover comunicações de rede entre nós tolerantes às falhas de caminho. Isto é, as comunicações entre dois ou mais nós podem continuar quando cada um está usando um NETFT a despeito da falha de qualquer componente no caminho contanto que pelo menos um caminho permaneça operacional.

Em um exemplo, a realização do acionador FT como um acionador de miniporta da NDIS provê pelo menos dois benefícios. Primeiro, porque um tal acionador FT geralmente se acomoda abaixo de quaisquer acionadores de protocolo na pilha, a confiabilidade do protocolo tende a ser provida por qualquer acionador de protocolo confiável de nível superior que geralmente não é afetado pela adição da tolerância à falta no nível da ligação provida por um acionador FT. Por exemplo, quando usando um acionador FT em combinação com um acionador de protocolo tal como um acionador de TCP/IP, o acionador FT tipicamente detectará caminhos falhos e roteará os pacotes de dados sobre caminhos operacionais de extremidade a extremidade independente de qualquer acionador de protocolo. Caso qualquer perda de pacote ocorra devido à mudança de caminhos, o acionador do protocolo TCP/IP, que geralmente se acomoda acima do acionador FT na pilha, tende a detectar tais perdas e executar quaisquer operações de nova tentativa ou reenvio para garantir que o protocolo confiável seja bem-sucedido na entrega do pacote.

Um segundo benefício de colocar o acionador FT a-

baixo do acionador de protocolo na pilha é que tipicamente nenhuma degradação da capacidade de roteamento do protocolo é introduzida. Quando assim configurado, qualquer operação de envelopamento que um acionador FT executa em um pacote de dados pode utilizar um protocolo roteável, tais como TCP ou UDP, assim garantindo que tais dados sejam roteáveis, além de ser tolerantes à falta ao nível da ligação. "Envelopar de maneira roteável" um pacote de dados é envelopar um pacote de dados usando um protocolo roteável.

NETFT, como uma parte de uma pilha de rede, geralmente se junta em uma aplicação de software via NDIS ou outras interfaces da pilha de rede. Uma tal união geralmente possibilita que as aplicações enviem e recebam pacotes de dados através de redes unidas na base da pilha. Em um exemplo, as aplicações tendem a usar um endereço virtual como o endereço de fonte para seus pacotes de dados, esse endereço virtual sendo conhecido para NETFT e mapeado e comunicado para outros nós na rede como descrito abaixo. Como mostrado na figura 3, o NETFT inclui um adaptador de miniporta 302 (também conhecido como um elemento de processamento), uma base de dados de roteamento 304 e um ou mais adaptadores de monitor de rota 306 e adaptadores de tunelamento 308.

O adaptador de tunelamento 308 tipicamente representa uma NIC no nó local (ou, em alguns casos, uma NIC virtual) e mantém um soquete usado para envelopar os pacotes para NETFT no nó alvo. Existe tipicamente um adaptador de tunelamento 308 associado com cada NIC no nó local com cada NIC sendo unida em uma rede provendo um caminho para um ou-

tro nó. Cada rede pode ou não ser isolada de qualquer outra rede. Um adaptador de tunelamento 308 é tipicamente associado com um acionador de protocolo de envelopamento e envelopa os pacotes de dados através de um protocolo de envelopamento 5 para e de sua NIC associada via interfaces da NDIS. Um exemplo de um protocolo de envelopamento é UDP. Alternativamente, outros protocolos, tais como TCP, IPX ou SPX, podem ser usados para envelopamento. Um adaptador de tunelamento 308 pode se tornar inativo caso a NIC associada ou a conexão dos 10 meios se torne inativa.

Uma base de dados de roteamento 304, quando realizada em NETFT, é tipicamente uma estrutura de dados simples, que pode ficar localizada na memória do sistema, que inclui entradas mapeando um endereço virtual para um ou mais caminhos para um NETFT similar em um outro nó. Em um exemplo, os 15 mapeamentos são representados por adaptadores de monitor de rota tal como adaptador do monitor de rota 306 que são tipicamente associados com um adaptador de tunelamento tal como adaptador de tunelamento 308. De forma geral, uma base de dados de roteamento tal como a base de dados de roteamento 304 incluirá um conjunto de adaptadores de rota para cada adaptador de tunelamento, cada adaptador de rota sendo associado com um nó alvo diferente alcançável através do caminho 20 associado com o adaptador de tunelamento. Quando usando 25 TCP/IP, por exemplo, a base de dados pode mapear um endereço virtual de destino para um endereço físico de um nó remoto específico.

Uma base de dados de roteamento 304 pode também incluir informação de prioridade para cada caminho. Tal in-

formação de prioridade pode ser usada para indicar um caminho preferido ou primário para um outro nó e/ou pode incluir informação sobre a velocidade do caminho ou outras características. Um caminho preferido é o caminho calculado pelo NETFT para ser usado através de outros caminhos possíveis, quando possível, com base na informação de prioridade e/ou estado do caminho. A informação de prioridade pode alternativamente indicar um algoritmo de equilíbrio de carga de roteamento para fazer uso de múltiplos caminhos para um nó alvo para equilibrar a carga do tráfego através dos caminhos, ou possibilitar algum outro esquema de priorização do caminho.

A tabela de mapeamento da base de dados da tabela de roteamento exemplar 304 é mostrada na tabela 1.

Tipo de endereço de destino	Endereço	Prioridade
Virtual	10.0.0.2	--
Físico: Trajetória A	172.56.48.38	1
Físico: Trajetória B	197.71.48.39	2

Tabela 1

Com referência à tabela 1 e figura 2, a tabela 1 mostra uma tabela de mapeamento exemplar como seria usada pelo NETFT operando no nó 216. A tabela 1 mostra o endereço de destino virtual 10.0.0.2, o endereço virtual como mapeado para o nó 266, mapeado para o endereço físico 172.56.48.38 associado com a Trajetória A para o nó 266 e endereço físico 197.71.48.39 associado com a Trajetória B

para o nó 266. A Trajetória A é mostrada com primeira prioridade e a Trajetória B com segunda prioridade. A tabela 1 é provida como um exemplo e não é planejada para ser limitadora.

5 Quando enviando dados do nó 216 para o nó 266, uma tal tabela de mapeamento é tipicamente usada para envelopar um pacote destinado para o endereço de destino virtual 10.0.0.2 enviando o pacote via um protocolo de envelopamento, tal como UDP, para o endereço de destino físico 10 172.56.48.38, assim envelopando o pacote do nó 216 sobre a Trajetória A para o nó 266. Uma tal tabela de mapeamento pode ser criada na base de dados de roteamento (figura 3, 304) para cada conjunto de caminhos estabelecidos entre dois nós. 15 Uma tal tabela de mapeamento pode ser implementada em várias formas, usar vários esquemas de prioridade e/ou armazenar outra informação incluindo estado operacional do caminho. A estrutura da tabela de mapeamento, número de caminhos, formatos de endereço, etc. mostrados na tabela 1 são providos como exemplos e não são planejados para ser limitadores.

20 O endereço virtual do nó local, endereços virtuais do nó remoto e informação de prioridade e outra do caminho são tipicamente providos para nós por um mecanismo fora da banda e passados para NETFT via suas interfaces de NDIS. Esse mecanismo fora da banda pode ser tão simples como um administrador de sistemas usando uma aplicação de gerenciamento para especificar a informação ou ele pode ser um sistema automatizado ou semelhante. Tais mecanismos fora da banda 25 são bem conhecidos para aqueles versados na técnica.

Como mostrado na figura 3, o adaptador de miniporta 302 (também conhecido como o elemento de processamento do acionador) tipicamente analisa um pacote de dados fluindo para baixo da pilha da rede, examina o endereço virtual de destino do pacote e usa informação da base de dados de roteamento 304 para determinar qual adaptador de tunelamento 308 através do qual envelopar o pacote de dados. Os pacotes que chegam, ou pacotes de dados fluindo para cima da pilha, são enviados para cima da pilha em direção ao seu endereço virtual de destino, o protocolo de envelopamento tendo previamente removido os cabeçalhos do pacote de envelopamento. Em particular, o adaptador de tunelamento 308 inspeciona os pacotes que chegam e envia pacotes de pulsos para um adaptador de monitor de rota 306 e envia outros pacotes para cima da pilha via um adaptador de miniporta 302. Aspectos de envelopamento dos pacotes de dados usando um protocolo de envelopamento e como os cabeçalhos do protocolo são adicionados e removidos pelos acionadores de protocolo são bem conhecidos para aqueles versados na técnica.

O adaptador do monitor de rota 306 tipicamente representa um nó remoto acessível através de um caminho específico identificado por um adaptador de tunelamento associado. O adaptador do monitor de rota 306 tipicamente proverá um endereço físico para o nó remoto, o endereço físico também correspondendo com um caminho específico para o nó remoto. Esse endereço físico é tipicamente usado para mapeamentos em uma base de dados de roteamento 304. Existe tipicamente um adaptador de monitor de rota para cada caminho dis-

tinto para um nó remoto, cada adaptador do monitor de rota sendo associado com um adaptador de tunelamento representando um caminho. Em um exemplo, com referência de volta à figura 2, o nó 210 é mostrado unido no nó 260 através de dois caminhos, um através da rede 1 202 ("Trajetória A") e o outro através da rede 2 282 ("Trajetória B"). O NETFT operando no nó 210 pode incluir um primeiro adaptador de monitor de rota ("RMA-A") provendo o endereço físico 172.56.48.38 do nó remoto 260 associado com sua NIC 262. O RMA-A pode ser associado com um primeiro adaptador de tunelamento ("TA-A") no nó 210 que pode ser associado com a Trajetória A. O NETFT no nó 210 pode também incluir um segundo adaptador de monitor de rota ("RMA-B") provendo o segundo endereço físico 197.71.48.39 do nó remoto 260 associado com sua NIC 264. O RMA-B pode ser associado com um segundo adaptador de tunelamento ("TA-B") no nó 210 que pode ser associado com a Trajetória B.

Com referência à figura 3, o adaptador do monitor de rota 306 tipicamente monitora o estado de um caminho para um nó remoto e indica um caminho falho ou não operacional na base de dados de roteamento 304. A monitoração tipicamente inclui receber quaisquer indicações de evento e/ou observar quaisquer falhas de pulso e atualizar a base de dados 304 adequadamente. Em um exemplo, um evento indicando a falha de uma NIC ou conexão de meios pode resultar na desativação do adaptador de tunelamento 308. Em um outro exemplo, uma falha de pulso pode resultar na desativação do adaptador do monitor de rota 306 associado com o nó remoto específico para o qual o pulso falhou.

A figura 4 é um diagrama de blocos mostrando uma arquitetura de comunicações tolerante à falta exemplar 216 incluindo NETFT 300 e uma aplicação 402. Nesse exemplo, a aplicação 402 envia pacotes de dados para o NETFT 300 via a pilha usando um endereço de fonte virtual 217 e um endereço de destino virtual representando o nó de destino. Tais pacotes de dados de saída fluem via a trajetória 480 da aplicação e através da pilha da rede para o acionador 300. O acionador 300 tipicamente determina quais dos caminhos possíveis 10 cada pacote deve adotar, geralmente usando informação de prioridade e informação do estado operacional do caminho armazenado na base de dados do roteamento e envelopa o pacote para o nó alvo sobre o caminho selecionado usando o endereço de fonte físico apropriado 422 ou 424.

15 A aplicação 402 pode enviar um pacote de dados através do NETFT 300 via o protocolo TCP, como mostrado na figura 4. Alternativamente, UDP ou qualquer outro protocolo pode ser usado. Também, como mostrado, o NETFT 300 pode usar o protocolo UDP para envelopar pacotes para o nó alvo. Alternativamente, TCP ou qualquer outro protocolo pode ser usado para envelopamento. Além do que, exemplos alternados 20 podem não fazer uso de adaptadores de miniporta ou acionadores da NDIS, mas podem usar outros mecanismos ou arquiteturas para executar funções similares. Finalmente, os vários elementos da pilha da rede e semelhantes podem operar em um modo de usuário ou um modo de núcleo, como mostrado ou de outra forma, ou nos sistemas com ou sem modos de operação 25 equivalentes.

A figura 5 é um diagrama de fluxo mostrando os dados fluindo através de um ambiente de comunicações tolerante à falta 500 incluindo um nó de fonte 216 e um nó de destino 266 acoplados via Trajetória A sobre a rede 1 202 e Trajetória B sobre a rede 2 282. Nesse ambiente exemplar 500, os dados são mostrados sendo enviados da aplicação operando no nó 216 para a aplicação atendendo no endereço virtual de destino no nó 266. Os pacotes de dados fluem para baixo da pilha da rede operando no nó 216 usando o protocolo TCP para 10 o NETFT como mostrado pela trajetória 501. Assumindo, como mostrado, que a Trajetória A é o caminho selecionado, o NETFT mapeia os pacotes de dados do endereço virtual de fonte sendo usado pela aplicação para a Trajetória A e envelopa os dados através do protocolo UDP usando o endereço de destino físico da Trajetória A para o nó alvo 266, para fora da 15 NIC do nó 216 como também mostrado pela trajetória 501 e para a rede 1 202 via a ligação 201. Os dados então fluem através da rede 1 202, sobre a ligação 203 e para o nó 266, fluindo para cima da pilha de rede operando no nó 266 como 20 mostrado pela trajetória 503. Os dados então fluem através do acionador de protocolo UDP, o mesmo protocolo que foi usado no lado de envio como o protocolo de envelopamento, onde os cabeçalhos do protocolo UDP são retirados dos pacotes de dados que são então passados para o NETFT operando no nó 266. O NETFT então envia os pacotes de dados para cima da pilha para a aplicação que está atendendo no endereço virtual de destino. As respostas tendem a fluir na ordem inversa.

A figura 6 é um diagrama de fluxo mostrando os da-

dos fluindo através do ambiente de comunicações tolerante à falta 500 mostrado na figura 5 com a adição de várias falhas de comunicações possíveis 610, 612, 620, 622 e 630. Outras falhas de comunicações são também possíveis. A falha 610 indica uma falha da NIC 1 operando no nó de envio 216. Uma tal falha pode ocorrer caso a NIC 1 seja removida do nó, caso o acionador da NIC 1 falhe, caso a própria NIC 1 falhe ou semelhantes. A falha pode ser detectada pelo NETFT via uma indicação de evento, tal como um evento PnP ou semelhante e/ou uma falha de pulso. Em uma tal situação, a Trajetória A é tipicamente considerada como tendo falhado e o NETFT selecionará um caminho operacional alternado de extremidade a extremidade. Um caminho operacional de extremidade a extremidade é tipicamente um caminho que pode entregar com sucesso os dados do nó de fonte e aplicação inteiramente para o nó de destino e aplicação.

A falha 620 indica uma falha dos meios de rede acoplando com a NIC 1 do nó 216. Essa falha pode ser devido a um cabo sendo desconectado da NIC 1, do cabo ficando desconectado de algum dispositivo da rede 1 202, do dispositivo no qual o cabo está conectado no lado da rede ser desenergizado ou falhar ou semelhante. Esse tipo de falha pode também ser detectada pelo NETFT via uma indicação de evento, tal como um evento PnP ou semelhante e/ou uma falha de pulso e um caminho alternado selecionado.

A falha 630 indica uma falha de algum tipo dentro da rede 202 resultando nos pacotes de dados falhando em alcançar o nó de destino 266. Nesse caso de falha, o nó de en-

vio 216 pode ainda estar acoplado na rede 202 com uma indicação apropriada de sentido de meios, porém a Trajetória A ficou interrompida mais para baixo da rede. Dada uma tal falha, o NETFT operando no nó de envio 216 pode não detectar a 5 falha via uma indicação de evento se as indicações locais mostram a conectividade na rede 202 como boa, mas podem detectar a falha via a falha do pulso da Trajetória A.

A falha 622 da ligação 203 e a falha 612 da NIC 1 operando no nó de recepção 266 tendem a ser similares às falhas 10 correspondentes mostradas para o nó 216. Mas essas falhas, não sendo locais ao nó 216 podem não ser detectadas via indicações de evento, mas podem ser detectadas via falha do pulso.

Qualquer uma dessas falhas, e outras falhas, podem 15 ser detectadas pelo NETFT operando no nó 216 e resultam nele selecionando um caminho operacional alternado de extremidade a extremidade, tal como Trajetória B sobre a rede 2 282. Nesse exemplo, como mostrado na figura 6, o NETFT envelopa os dados para baixo da trajetória alternada 681 e sobre a 20 rede 2 282 para o nó receptor 266. Caso a condição de falha seja corrigida e o estado operacional de extremidade a extremidade restaurado na Trajetória A, o NETFT operando no nó de envio 216 pode detectar a recuperação e novamente fazer uso da Trajetória A. Além do que, quaisquer respostas do nó 25 266 de volta para o nó 216 podem ser envelopadas em um modo tolerante à falta similar pelo NETFT.

A figura 7 é um diagrama de blocos mostrando um outro exemplo de um acionador de comunicações tolerante à

falta, NETFT 700. Esse exemplo é similar ao exemplo mostrado na figura 3, mas inclui variações como descrito abaixo. Nesse exemplo, uma aplicação de software pode não precisar usar endereços virtuais. No lugar disso, uma aplicação pode usar 5 um endereço de destino físico para endereçar pacotes de dados para o nó alvo.

O adaptador de protocolo 710 geralmente se junta no adaptador de miniporta 702 (também conhecido como o elemento de processamento do acionador) e em um adaptador de 10 miniporta da NIC (não mostrado). Existe tipicamente um adaptador de protocolo para cada NIC instalada no nó, cada adaptador de protocolo sendo associado com uma NIC via seu adaptador de NIC. Como cada adaptador de protocolo é associado com uma NIC, ele é também associado com o caminho acoplado 15 na NIC. O adaptador de protocolo 710 é operável para aceitar pacotes de dados de uma aplicação via o elemento de processamento 702 e passar os pacotes de dados para a NIC associada sem a necessidade de envelopamento.

O elemento de processamento 702 tipicamente analisa 20 um pacote de dados fluindo para baixo da pilha da rede, examina o endereço de destino físico do pacote e usa a informação da base de dados de roteamento 704 para determinar se o pacote pode ser enviado através de um adaptador de protocolo 710 ou precisa ser envelopado através de um adaptador 25 de tunelamento 308 para o nó alvo. De forma geral, se o caminho indicado pelo endereço do destino físico é operacional de extremidade a extremidade, o pacote de dados será enviado

através desse caminho. De outra forma, um caminho alternado pode ser selecionado sobre o qual o pacote pode ser envelopado.

Nesse exemplo, a base de dados de roteamento 704 mantém mapeamentos de endereços de destino físico e caminhos, junto com prioridade e outras informações como descrito acima. Uma tabela de mapeamento da base de dados de roteamento exemplar 704 é mostrada na tabela 2.

Tipo de endereço de destino	Endereço	Prioridade
Físico: Trajetória A	172.56.48.38	1
Físico: Trajetória B	197.71.48.39	2

Tabela 2

10 Com referência à tabela 2 e figura 2, a tabela 2 mostra uma tabela de mapeamento exemplar como seria usada pelo NETFT operando no nó 216. A tabela 2 mostra um mapeamento incluindo o endereço de destino físico 172.56.48.38 associado com a Trajetória A para o nó 266 e o endereço de destino físico 197.71.48.39 associado com a Trajetória B para o nó 266. A Trajetória A é mostrada com primeira prioridade e a Trajetória B com segunda prioridade.

20 Quando enviando dados do nó 216 para o nó 266, uma tal tabela de mapeamento é tipicamente usada no envio (ou envelopamento se necessário) de um pacote de dados sendo enviado para o endereço de destino físico 172.56.48.38 do nó 266. Se o caminho associado com o endereço de destino origi-

nal está operacional, o pacote de dados tende a ser enviado para o nó de destino sem envelopamento. Se esse caminho não está disponível, então o pacote de dados é enviado através do caminho alternado para o endereço de destino físico 5 197.71.48.39 do nó 266 através do envelopamento. Outros aspectos do NETFT 700 são geralmente similares a esses do NETFT como descrito para a figura 3.

A figura 8 é um diagrama de blocos mostrando uma arquitetura de comunicações tolerante à falta exemplar 216 incluindo o NETFT 700 e uma aplicação 402. Nesse exemplo, a aplicação 402 envia pacotes de dados para o NETFT 700 via a pilha usando um endereço de fonte físico e um endereço de destino físico 801 representando o nó de destino. Tais pacotes de dados de saída fluem via trajetória 880 da aplicação 10 e através da pilha de rede para o acionador 700. O acionador 15 700 tipicamente determina quais dos caminhos possíveis cada pacote deve adotar, geralmente usando a informação de prioridade e a informação do estado operacional do caminho armazenadas na base de dados de roteamento e envia o pacote para 20 o nó alvo através do caminho indicado pelo endereço de destino físico original ou, se esse caminho não é a operação de extremidade a extremidade, envelopa o pacote através de um caminho alternado como indicado nesse exemplo pela rota 882 25 e NIC 2 892.

A aplicação 402 pode enviar um pacote de dados através do NETFT 700 via o protocolo TCP, como mostrado na figura 8. Alternativamente, UDP ou qualquer outro protocolo pode ser usado. Também, como mostrado, o NETFT 700 pode usar

o protocolo UDP para envelopar pacotes para o nó alvo. Alternativamente, TCP ou qualquer outro protocolo pode ser utilizado para envelopamento. Além do que, outros exemplos podem não fazer uso de acionadores da NDIS, mas podem usar 5 outros mecanismos ou arquiteturas para executar funções similares. Finalmente, os vários elementos da pilha de rede e semelhantes podem operar em um modo de usuário ou um modo de núcleo, como mostrado ou de outra forma, ou em sistemas com ou sem modos de operação equivalentes.

10 A figura 9 é um diagrama de fluxo mostrando os dados fluindo através de um ambiente de comunicações tolerante à falta 900 incluindo um nó de fonte 816 e um nó de destino 966 acoplados via Trajetória A através da rede 1 202 e Trajetória B através da rede 2 282. Nesse ambiente exemplar 15 900, os dados são mostrados sendo enviados da aplicação operando no nó 216 para a aplicação atendendo no endereço físico de destino no nó 266. Os pacotes de dados fluem para baixo da pilha da rede operando no nó 216 usando o protocolo TCP para o NETFT como mostrado pela trajetória 901. Assumindo, como mostrado, que a Trajetória A é o caminho selecionado, o NETFT envia os pacotes de dados usando o endereço de destino físico provido pela aplicação via a NIC 1 do nó 216 sobre a Trajetória A e rede 1 202 via a ligação 201. Os dados então fluem através da rede 1 202, sobre a ligação 203 e 20 para o nó 966, fluindo para cima da pilha de rede operando no nó 966 como mostrado pela trajetória 903. Os dados então fluem através do NETFT e do acionador de protocolo (um acionador de protocolo para o mesmo protocolo que foi usado no

lado de envio como o protocolo de envio) e para cima para a aplicação. As respostas tendem a fluir na ordem inversa.

A figura 10 é um diagrama de fluxo mostrando os dados fluindo através do ambiente de comunicações tolerante à falta 900 mostrado na figura 9 com a adição de várias falhas de comunicações possíveis 1010, 1012, 1020, 1022 e 1030. Outras falhas de comunicações são também possíveis. A falha 1010 indica uma falha da NIC 1 operando no nó de envio 816. Uma tal falha pode ocorrer caso a NIC 1 seja removida do nó, seu acionador de NIC falhe, a própria NIC falhe ou semelhantes. A falha pode ser detectada pelo NETFT via uma indicação de evento, tal como um evento PnP ou semelhante e/ou uma falha de pulso. Em uma tal situação, a Trajetória A é tipicamente considerada como tendo falhado e o NETFT selecionará um caminho operacional alternado de extremidade a extremidade.

A falha 1020 indica uma falha dos meios de rede se unindo com a NIC 1 do nó 816. Essa falha pode ser devido a um cabo sendo desconectado da NIC 1, do cabo ficar desconectado de algum dispositivo da rede 1 202, do dispositivo em que o cabo está conectado no lado da rede ser desligado ou falhar ou semelhantes. Esse tipo de falha pode também ser detectada pelo NETFT via uma indicação de evento, tal como um evento PnP ou semelhante e/ou uma falha de pulso e um caminho alternado selecionado.

A falha 1030 indica uma falha de algum tipo dentro da rede 202 resultando nos pacotes de dados falhando em alcançar o nó de destino 966. Nesse caso de falha, o nó de en-

vio 816 pode ainda estar unido na rede 202 com uma indicação de sentido de meios apropriada, porém a Trajetória A ficou interrompida mais para baixo da rede. Dada uma tal falha, o NETFT operando no nó de envio 816 pode não detectar a falha via uma indicação de evento, tal como um evento PnP ou semelhante, se as indicações locais mostram a conectividade na rede 202 como boa, mas pode detectar a falha através da falha de pulso da Trajetória A.

A falha 1022 da ligação 203 e a falha 1012 da NIC 10 1 operando no nó receptor 966 tendem a ser similares às falhas correspondentes mostradas para o nó 816. Mas essas falhas, não sendo locais ao nó 816 podem não ser detectadas via indicações de evento, mas podem ser detectadas via falha do pulso.

15 Qualquer uma dessas falhas e outras falhas podem ser detectadas pelo NETFT operando no nó 816 e resultam nele selecionando um caminho operacional alternado de extremidade a extremidade, tal como a Trajetória B sobre a rede 2 282. Nesse exemplo, como mostrado na figura 10, o NETFT envelopa 20 os dados para baixo da trajetória alternada 1081 e sobre a rede 2 282 para o nó receptor 966. Caso a condição de falha seja corrigida e o estado operacional de extremidade a extremidade restaurado na Trajetória A, o NETFT operando no nó de envio 816 pode detectar a recuperação e novamente fazer 25 uso da Trajetória A. Além disso, quaisquer respostas do nó 966 de volta para o nó 816 podem ser enviadas ou envelopadas, dependendo do estado operacional da Trajetória A e Trajetória B, em um modo similar tolerante à falta por seu

NETFT.

A figura 11 é um diagrama de blocos mostrando um ambiente de computação exemplar 1100 no qual a tecnologia descrita acima pode ser realizada. Um ambiente de computação adequado pode ser realizado com numerosos sistemas de uso especial ou uso geral. Exemplos de sistemas bem conhecidos podem incluir, mas não são limitados a, computadores pessoais ("PC"), dispositivos de mão ou laptops, sistemas com base em microprocessador, sistemas de multiprocessador, servidores, estações de trabalho, dispositivos eletrônicos de consumidor, conversores de sinal de freqüência e semelhantes.

O ambiente de computação 1100 geralmente inclui um sistema de computação de uso geral na forma de um dispositivo de computação 1100 unido em vários dispositivos periféricos 1102, 1103, 1104 e semelhantes. O sistema 1100 pode se unir a vários dispositivos de entrada 1103, incluindo teclados e dispositivos de indicação, tais como um mouse ou trackball, via uma ou mais interfaces de I/O 1112. Os componentes do dispositivo de computação 1101 podem incluir um ou mais processadores (incluindo unidades de processamento central ("CPU"), unidades de processamento gráfico ("GPU"), microprocessadores ("uP") e semelhantes) 1107, memória do sistema 1109 e um barramento do sistema 1108 que tipicamente se une aos vários componentes. O processador 1107 tipicamente processa ou executa várias instruções executáveis pelo computador para controlar a operação do dispositivo de computação 1101 e para se comunicar com outros dispositivos eletrô-

nicos e/ou de computação, sistemas ou ambientes (não mostrados) via várias conexões de comunicações tal como uma conexão de rede 1114 ou semelhante. O barramento do sistema 1108 representa qualquer número de vários tipos de estruturas de barramento, incluindo um barramento de memória ou controlador de memória, um barramento periférico, um barramento serial, uma porta gráfica acelerada, um processador ou barramento local usando qualquer uma de uma variedade de arquiteturas de barramento e semelhantes.

A memória do sistema 1109 pode incluir meios legíveis por computador na forma de memória volátil, tal como memória de acesso aleatório ("RAM") e/ou memória não volátil, tais como memória somente de leitura ("ROM") ou memória flash ("FLASH"). Um sistema básico de entrada/saída ("BIOS") pode ser armazenado em não volátil ou semelhante. A memória do sistema 1109 tipicamente armazena dados, instruções executáveis pelo computador e/ou módulos do programa compreendendo instruções executáveis pelo computador que são imediatamente acessíveis para e/ou atualmente operadas por um ou mais dos processadores 1107.

Os dispositivos de armazenamento em massa 1104 e 1110 podem ser unidos no dispositivo de computação 1101 ou incorporados no dispositivo de computação 1100 via acoplamento no barramento do sistema. Tais dispositivos de armazenamento em massa 1104 e 1110 podem incluir uma unidade de disco magnético que lê dê e/ou escreve em um disco magnético não volátil removível (por exemplo, um "disco flexível") 1105 e/ou uma unidade de disco ótico que lê dê e/ou escreve

em um disco ótico não volátil tais como um CD-ROM, DVD ROM 1106. Alternativamente, um dispositivo de armazenamento em massa, tal como disco rígido 1110, pode incluir meio de armazenamento não removível. Outros dispositivos de armazenamento em massa podem incluir placas de memória, bastões de memória, dispositivos de armazenamento em fita e semelhantes.

Qualquer número de programas de computador, arquivos, estruturas de dados e semelhantes pode ser armazenado no disco rígido 1110, outros dispositivos de armazenamento 1104, 1105, 1106 e memória do sistema 1109 (tipicamente limitado para o espaço disponível) incluindo, por meio de exemplo, sistemas operacionais, programas de aplicação, arquivos de dados, estruturas de diretório e instruções executáveis pelo computador.

Dispositivos de saída, tal como dispositivo de exibição 1102, podem ser unidos no dispositivo de computação 1101 via uma interface, tal como adaptador de vídeo 1111. Outros tipos de dispositivos de saída podem incluir impressoras, saídas de áudio, dispositivos táteis ou outros mecanismos de saída sensores ou semelhantes. Dispositivos de saída podem habilitar o dispositivo de computação 1101 a interagir com operadores humanos ou outras máquinas ou sistemas. Um usuário pode fazer interface com o ambiente de computação 1100 via qualquer número de dispositivos de entrada diferentes 1103 tais como um teclado, mouse, barra de direção, base de jogos, porta de dados e semelhantes. Esses e outros dispositivos de entrada podem ser unidos no processador 1107

via interfaces de entrada/saída 1112 que podem ser unidas no barramento do sistema 1108 e podem ser unidas por outras interfaces e estruturas de barramento, tais como uma porta paralela, porta de jogos, barramento serial universal ("USB"),
5 fire wire, porta de infravermelho e semelhantes.

O dispositivo de computação 1101 pode operar em um ambiente de rede via conexões de comunicações em um ou mais dispositivos de computação remotos através de uma ou mais redes locais ("LAN"), redes remotas ("WAN"), redes da área
10 de armazenamento ("SAN"), a Internet, ligações de rádio, ligações óticas e semelhantes. O dispositivo de computação 1101 pode ser acoplado em uma rede via adaptador de rede 1113 ou semelhante ou, alternativamente, via um modem, ligação de linha de assinante digital ("DSL"), ligação de rede
15 digital de serviços integrados ("ISDN"), ligação da Internet, ligação sem fio ou semelhantes.

A conexão de comunicações 1114, tal como uma conexão de rede, tipicamente provê um acoplamento nos meios de comunicações, tal como uma rede. Meios de comunicações tipicamente provêem instruções legíveis por computador e executáveis por computador, estruturas de dados, arquivos, módulos do programa e outros dados usando um sinal de dados modulado, tal como uma onda portadora ou outro mecanismo de transporte. O termo "sinal de dados modulado" tipicamente
20 significa um sinal que tem uma ou mais de suas características ajustadas ou alteradas em uma tal maneira a fim de codificar a informação no sinal. Por meio de exemplo, e não limitação, meios de comunicações podem incluir meios ligados
25

por fiação, tal como uma rede ligada por fiação ou conexão ligada por fiação direta ou semelhante, e meios sem fio, tais como acústico, de radiofrequência, de infravermelho ou outros mecanismos de comunicações sem fio.

5 Aqueles versados na técnica verificarão que dispositivos de armazenamento utilizados para prover instruções legíveis por computador e executáveis por computador e dados podem ser distribuídos através de uma rede. Por exemplo, um computador remoto ou dispositivo de armazenamento pode armazenar instruções legíveis pelo computador e executáveis pelo computador na forma de aplicações de software e dados. Um computador local pode acessar o computador remoto ou dispositivo de armazenamento via a rede e transferir parte ou toda uma aplicação de software ou dados e pode executar quaisquer instruções executáveis pelo computador. Alternativamente, o computador local pode transferir pedaços do software ou dados quando necessário ou processar de maneira distributiva o software executando algumas das instruções no computador local e algumas nos computadores remotos e/ou dispositivos.

10 20

 Aqueles versados na técnica também verificarão que, pela utilização de técnicas convencionais, todas ou porções das instruções executáveis pelo computador do software podem ser executadas por um circuito eletrônico dedicado tal como um processador de sinal digital ("DSP"), arranjo lógico programável ("PLA"), circuitos discretos e semelhantes. O termo "aparelho eletrônico" pode incluir dispositivos de computação ou dispositivos eletrônicos do consu-

25

midor compreendendo qualquer software, firmware ou semelhante, ou dispositivos eletrônicos ou circuitos não compreendendo software, firmware ou semelhantes.

O termo "firmware" tipicamente se refere a instruções executáveis, código ou dados mantidos em um dispositivo eletrônico tal como uma ROM. O termo "software" se refere geralmente a instruções executáveis, código, dados, aplicações, programas ou semelhantes mantidos em ou em qualquer forma dos meios legíveis por computador. O termo "meios legíveis por computador" tipicamente se refere à memória do sistema, dispositivos de armazenamento e seus meios associados, meios de comunicações e semelhantes.

REIVINDICAÇÕES

1. Método para prover uma aplicação com comunicações de rede tolerante à falta entre uma pluralidade de nós, **CARACTERIZADO** pelo fato de que compreende:

5 prover uma pluralidade de caminhos de comunicações iniciais via uma pluralidade de redes unidas na pluralidade de nós,

10 receber um pacote de dados em um nó de envio da aplicação, o nó de envio sendo um da pluralidade de nós, o pacote de dados sendo endereçado pela aplicação para um endereço em um da pluralidade de nós, e

15 selecionar um primeiro caminho selecionado para o pacote de dados dentre a pluralidade de caminhos de comunicações iniciais, onde o primeiro caminho selecionado é um caminho preferido.

2. Método, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que também compreende:

20 detectar um evento de conectividade local associado com um da pluralidade de caminhos de comunicações iniciais, e

indicar se ou não um caminho da pluralidade de caminhos de comunicações iniciais está operacional com base no evento de conectividade local.

3. Método, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que também compreende:

enviar um pacote de pulsos roteável através de um da pluralidade de caminhos de comunicações iniciais,

monitorar uma resposta para o pacote de pulsos ro-

teável para determinar um estado operacional de extremidade a extremidade do um de pluralidade de caminhos de comunicações iniciais, e

indicar se ou não um da pluralidade de caminhos de
5 comunicações iniciais está operacional de extremidade a ex-
tremidade com base na monitoração da resposta.

4. Método, de acordo com a reivindicação 1,
CARACTERIZADO pelo fato de que o endereço é um endereço de
protocolo da Internet versão 4 ou um endereço de protocolo
10 da Internet versão 6.

5. Método, de acordo com a reivindicação 1,
CARACTERIZADO pelo fato de que o pacote de dados recebido da
aplicação é um pacote do protocolo de controle de transmis-
são ou um pacote do protocolo de datagrama do usuário.

15 6. Método, de acordo com a reivindicação 1,
CARACTERIZADO pelo fato de que as instruções executáveis por
computador para executar o método são armazenadas em um meio
legível por computador.

7. Método, de acordo com a reivindicação 1,
20 **CARACTERIZADO** pelo fato de que o endereço é um endereço vir-
tual.

8. Método, de acordo com a reivindicação 7,
CARACTERIZADO pelo fato de que também compreende tunelar de
maneira roteável o pacote de dados sobre o primeiro caminho
25 selecionado sem requerer que a aplicação fique ciente de
qual da pluralidade de caminhos de comunicações iniciais é o
caminho selecionado.

9. Método, de acordo com a reivindicação 8,

CARACTERIZADO pelo fato de que o pacote de dados tunelado de maneira roteável sobre o primeiro caminho selecionado é um pacote do protocolo de controle de transmissão ou um pacote do protocolo de datagrama do usuário.

5 10. Método, de acordo com a reivindicação 7,

CARACTERIZADO pelo fato de que também compreende:

detectar uma falha do primeiro caminho selecionado,

10 selecionar um segundo caminho selecionado dentre a pluralidade de caminhos de comunicações iniciais onde o segundo caminho selecionado é tanto um caminho preferido quanto está operacional de extremidade a extremidade, e

15 tunelar de maneira roteável o pacote de dados sobre o segundo caminho selecionado sem requerer que a aplicação fique ciente de qual do um ou mais caminhos de comunicações iniciais é o segundo caminho selecionado.

20 11. Método, de acordo com a reivindicação 1,
CARACTERIZADO pelo fato de que o primeiro caminho selecionado é selecionado com base em um endereço de destino físico incluído com o pacote de dados.

25 12. Método, de acordo com a reivindicação 11,
CARACTERIZADO pelo fato de que também compreende enviar o pacote de dados sobre o primeiro caminho selecionado sem requerer que a aplicação fique ciente de qual da pluralidade de caminhos de comunicações iniciais é o primeiro caminho selecionado.

13. Método, de acordo com a reivindicação 11,
CARACTERIZADO pelo fato de que também compreende:

detectar uma falha do primeiro caminho selecionado,

selecionar um segundo caminho selecionado da pluralidade de caminhos de comunicações iniciais onde o segundo caminho selecionado é tanto um caminho preferido quanto está operacional de extremidade a extremidade, e

10 tunelar de maneira roteável o pacote de dados sobre o segundo caminho selecionado sem requerer que a aplicação fique ciente de qual de um ou mais caminhos de comunicações iniciais é o segundo caminho selecionado.

14. Método, de acordo com a reivindicação 13, **CARACTERIZADO** pelo fato de que o pacote de dados tunelado de maneira roteável sobre o segundo caminho selecionado é um pacote do protocolo de controle de transmissão ou um pacote 15 do protocolo de datagrama do usuário.

15. Método para prover uma aplicação com comunicações de rede tolerante à falta entre uma pluralidade de nós, **CARACTERIZADO** pelo fato de que compreende:

20 prover uma pluralidade de caminhos de comunicações iniciais via uma pluralidade de redes unidas na pluralidade de nós,

25 receber um pacote de dados em um nó receptor, o nó receptor incluindo uma pilha de rede tolerante à falta e sendo um da pluralidade de nós, o pacote de dados sendo destinado para a aplicação,

determinar se o pacote de dados foi tunelado de maneira roteável, e

se o pacote de dados foi tunelado de maneira rote-

ável, enviar o pacote de dados para cima da pilha da rede tolerante à falta.

16. Método, de acordo com a reivindicação 15, **CARACTERIZADO** pelo fato de que também comprehende:

5 receber um pacote de pulsos roteável no nó receptor sobre um da pluralidade de caminhos de comunicações iniciais, e

10 responder com uma resposta para o pacote de pulsos roteável para indicar um estado operacional de extremidade a extremidade de um da pluralidade de caminhos de comunicações iniciais.

15 17. Método, de acordo com a reivindicação 15, **CARACTERIZADO** pelo fato de que o pacote de dados é um pacote do protocolo de controle de transmissão ou um pacote do protocolo de datagrama do usuário.

18. Método, de acordo com a reivindicação 15, **CARACTERIZADO** pelo fato de que as instruções executáveis pelo computador para executar o método são armazenadas em um meio legível por computador.

20 19. Sistema para prover uma aplicação com comunicações de rede tolerante à falta entre uma pluralidade de nós, **CARACTERIZADO** pelo fato de que comprehende:

25 um primeiro acionador tolerante à falta acoplado em uma primeira pilha da rede e operando em um primeiro nó, o primeiro nó sendo um da pluralidade de nós,

um segundo acionador tolerante à falta acoplado em uma segunda pilha de rede e operando em um segundo nó, o segundo nó sendo um da pluralidade de nós, e

o primeiro acionador tolerante à falta e o segundo acionador tolerante à falta sendo unidos via uma pluralidade de caminhos de comunicações iniciais através de uma pluralidade de redes.

5 20. Sistema, de acordo com a reivindicação 19, **CARACTERIZADO** pelo fato de que o primeiro acionador tolerante à falta compreende:

um elemento de processamento unido na aplicação via a primeira pilha de rede,

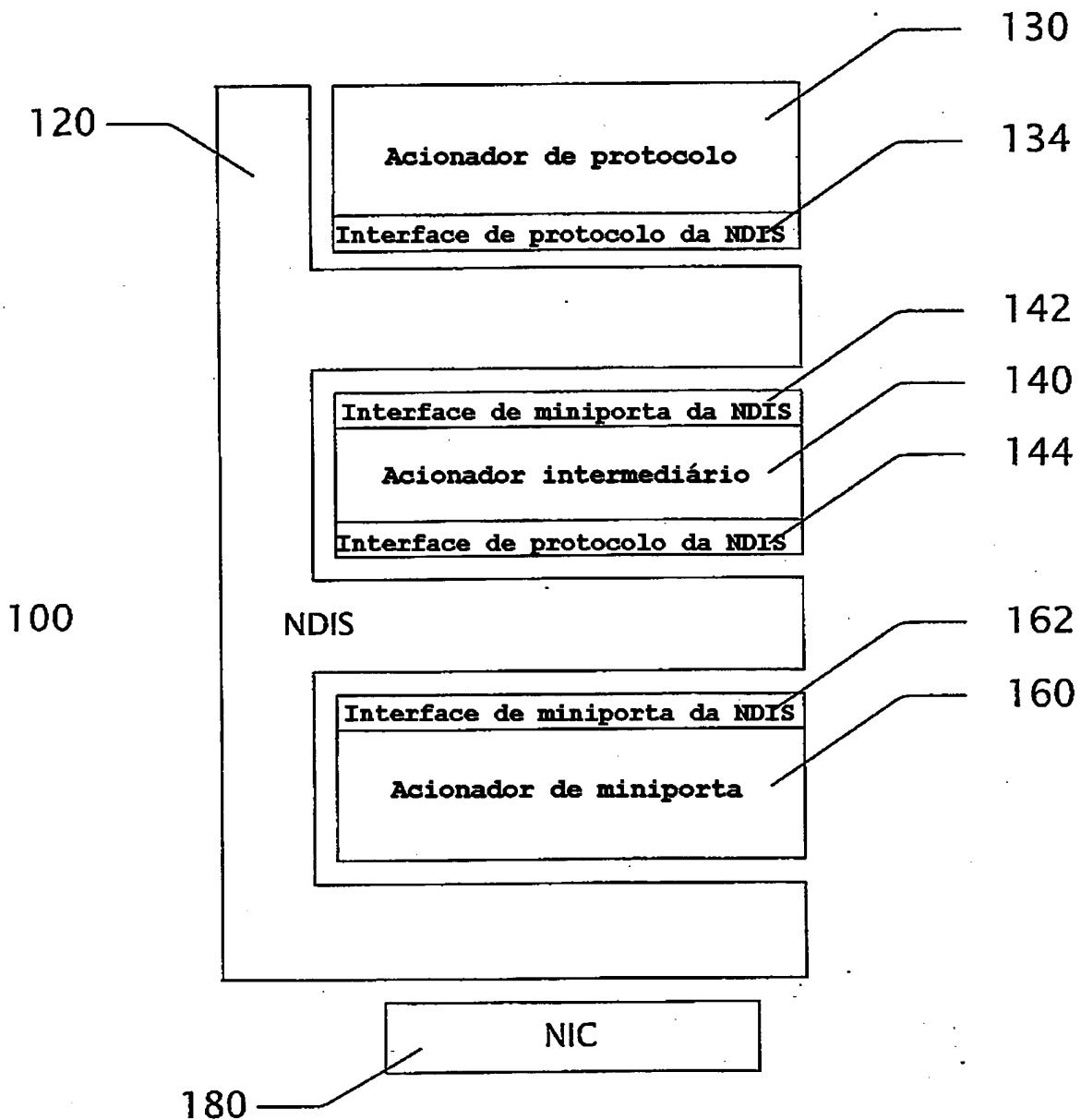
10 uma base de dados de roteamento unida no elemento de processamento incluindo:

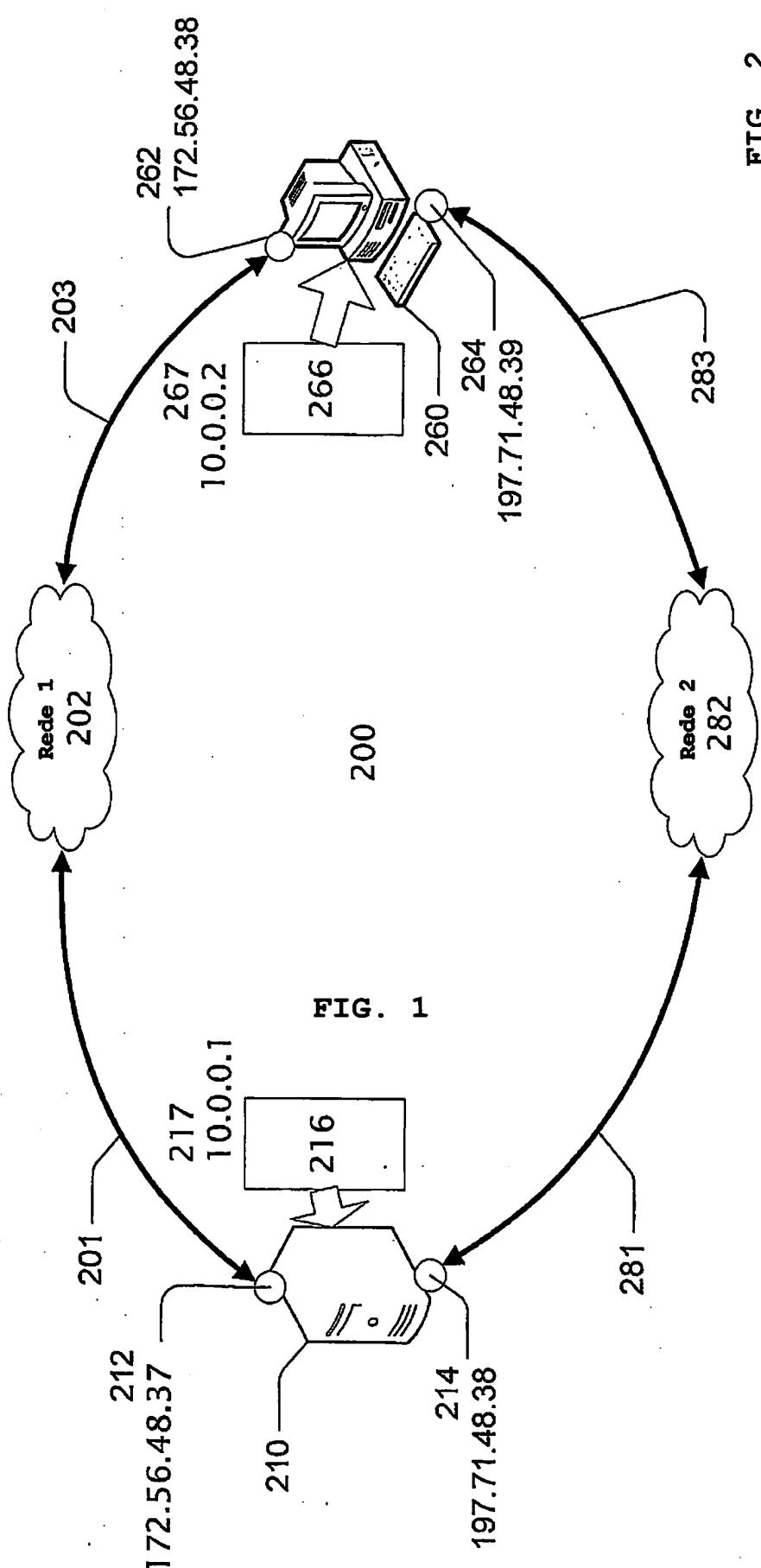
uma entrada representando um caminho para o segundo nó incluindo um endereço físico do segundo nó, o caminho sendo um da pluralidade de caminhos de comunicações iniciais, e

uma indicação de um estado operacional de extremidade a extremidade do caminho para o segundo nó,

20 um adaptador de protocolo unido no elemento de processamento e unido em uma da pluralidade de redes via a primeira pilha de rede, a uma da pluralidade de redes sendo associada com o caminho para o segundo nó, e

um adaptador de tunelamento associado com o caminho para o segundo nó, unido no elemento de processamento e unido em uma da pluralidade de redes via a primeira pilha de rede.





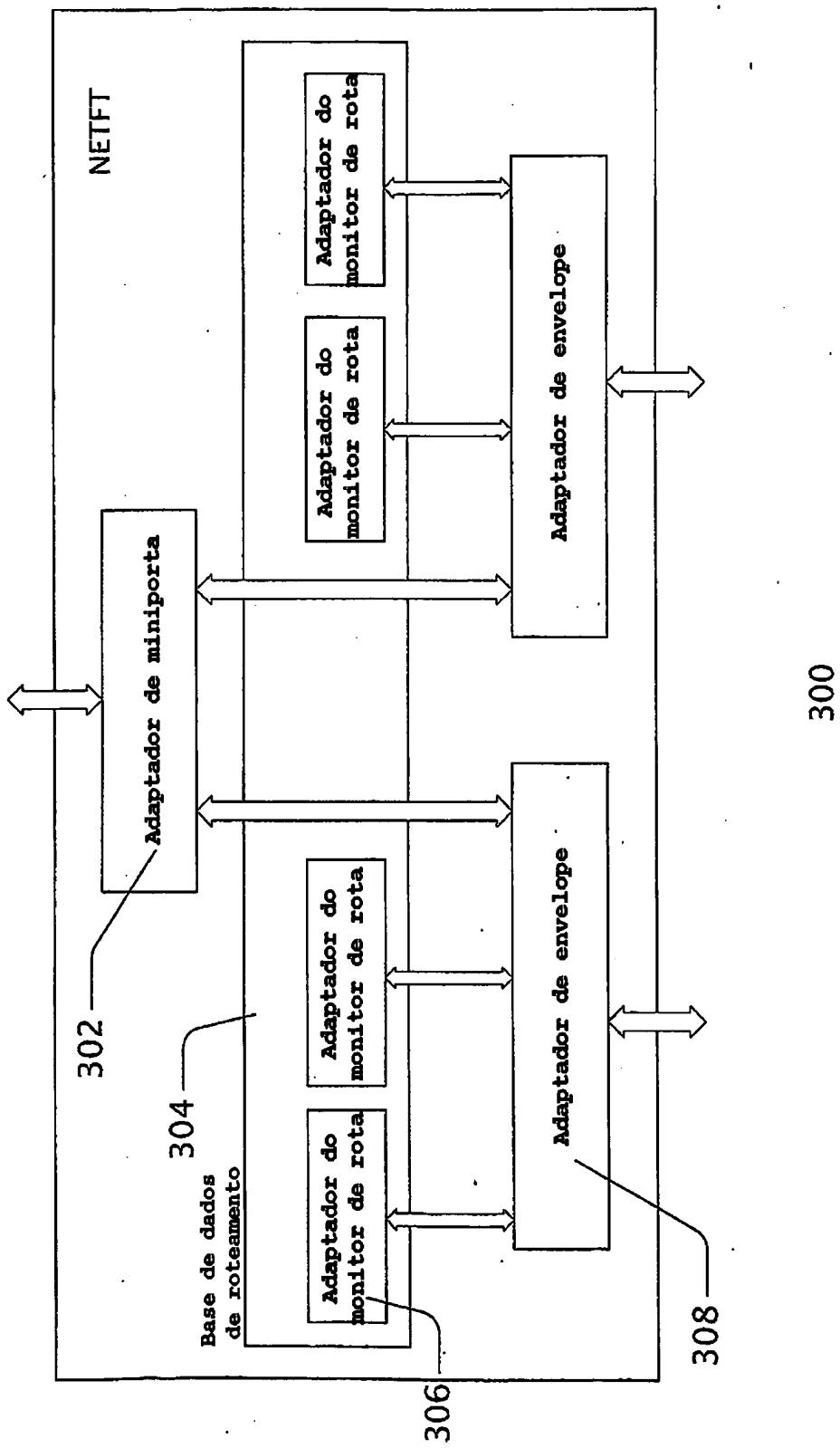


FIG. 3

300

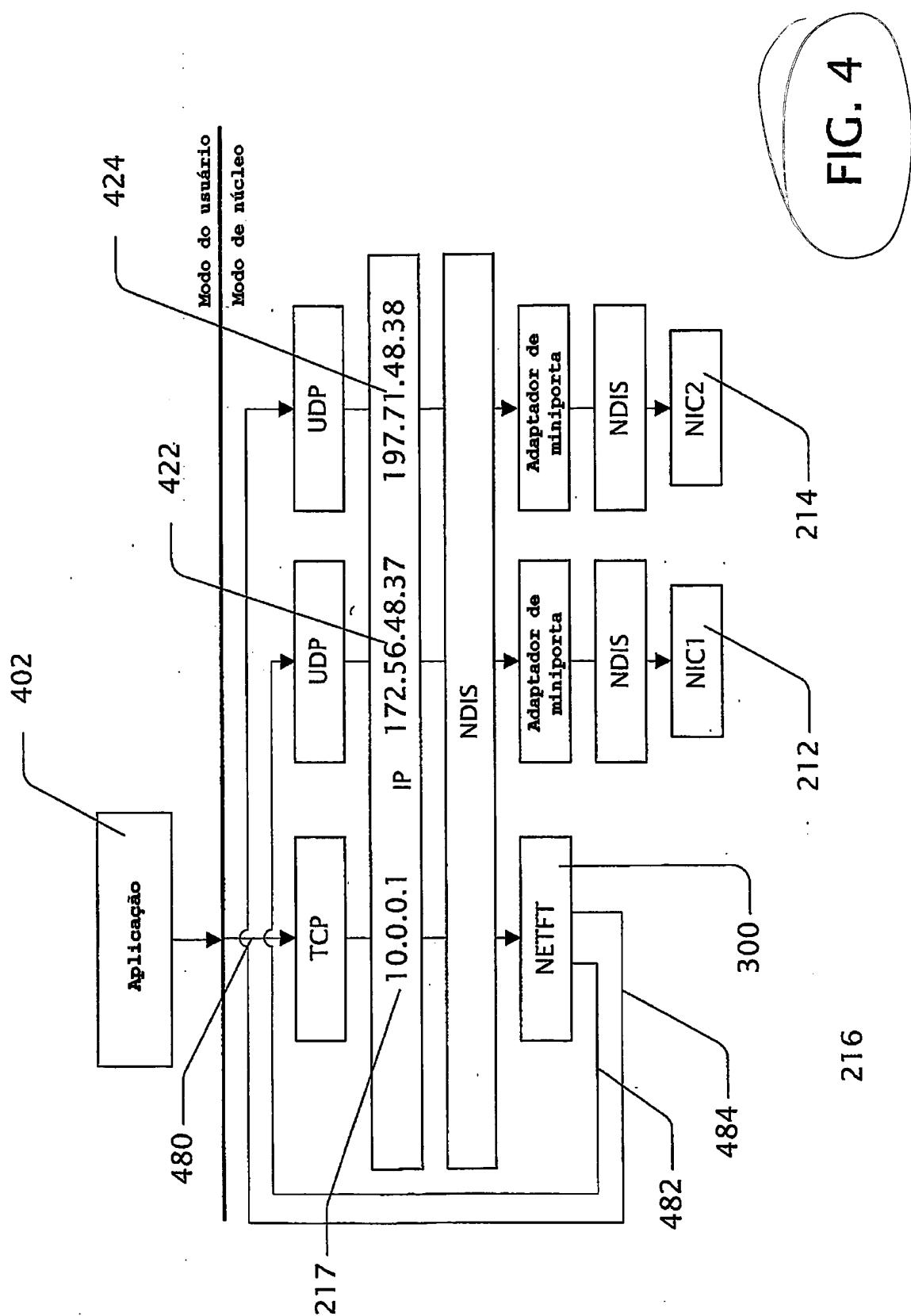
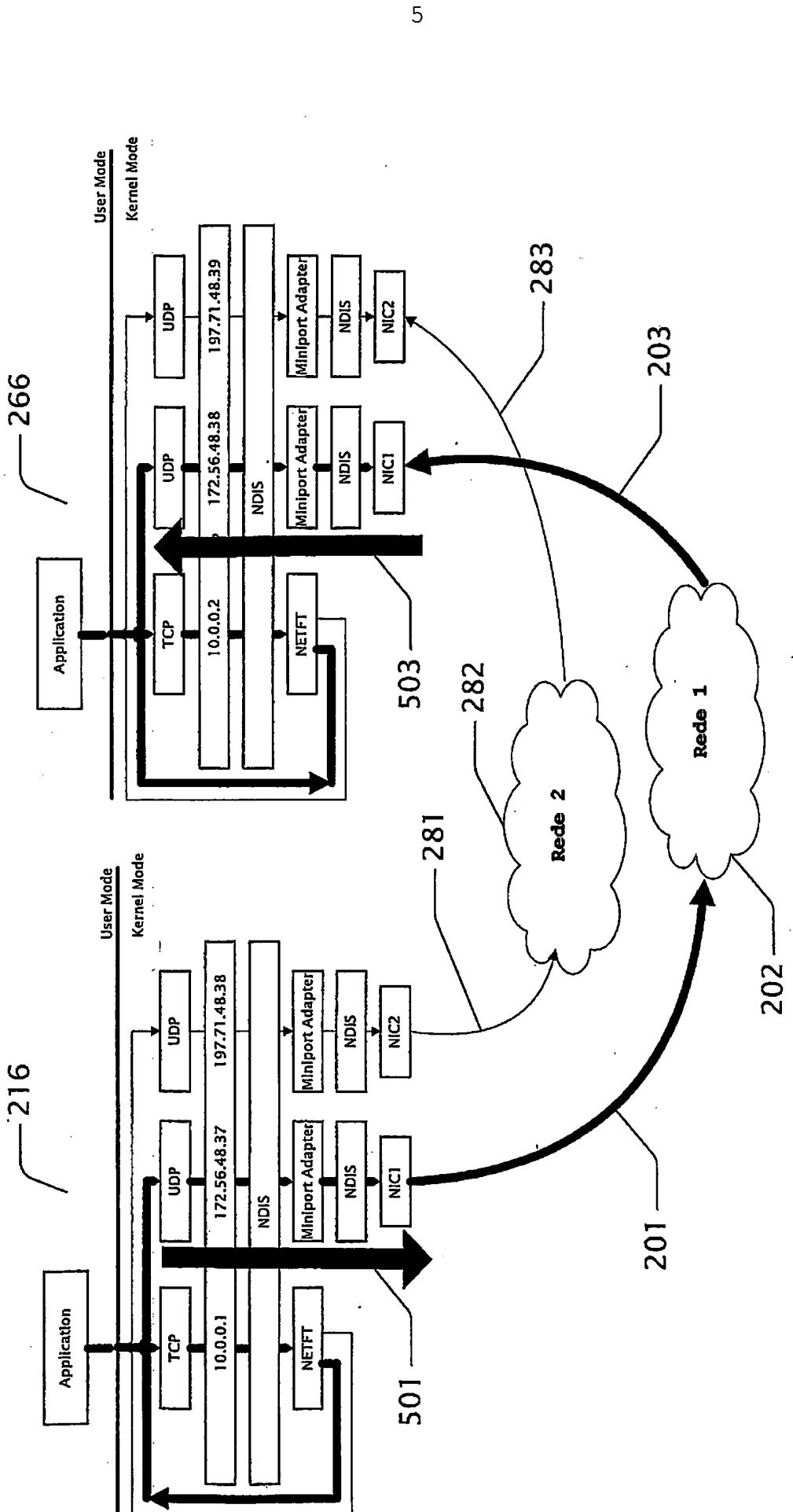


FIG. 4

216

FIG. 5

500



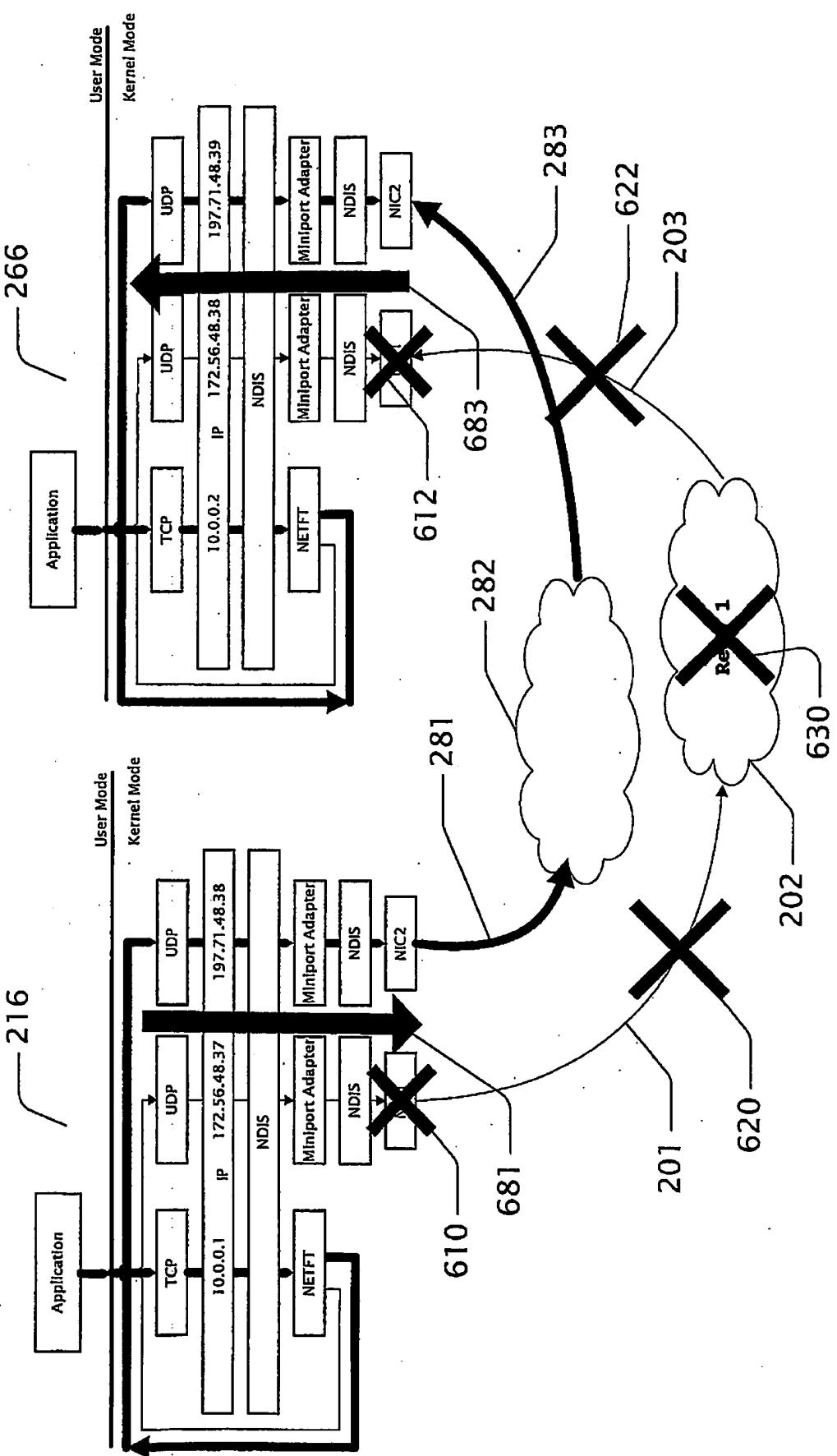
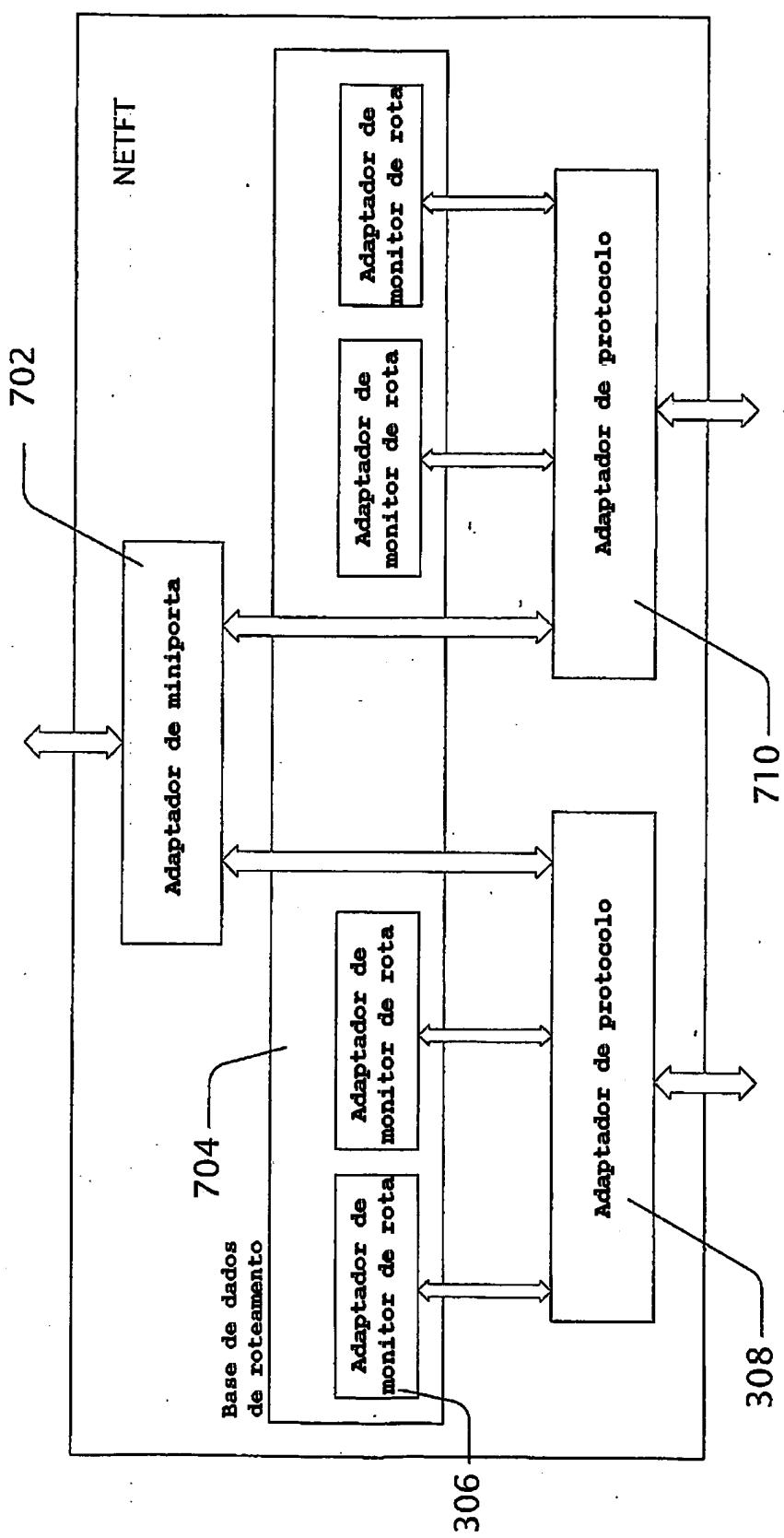


FIG. 6

60

FIG. 7



700

FIG. 8

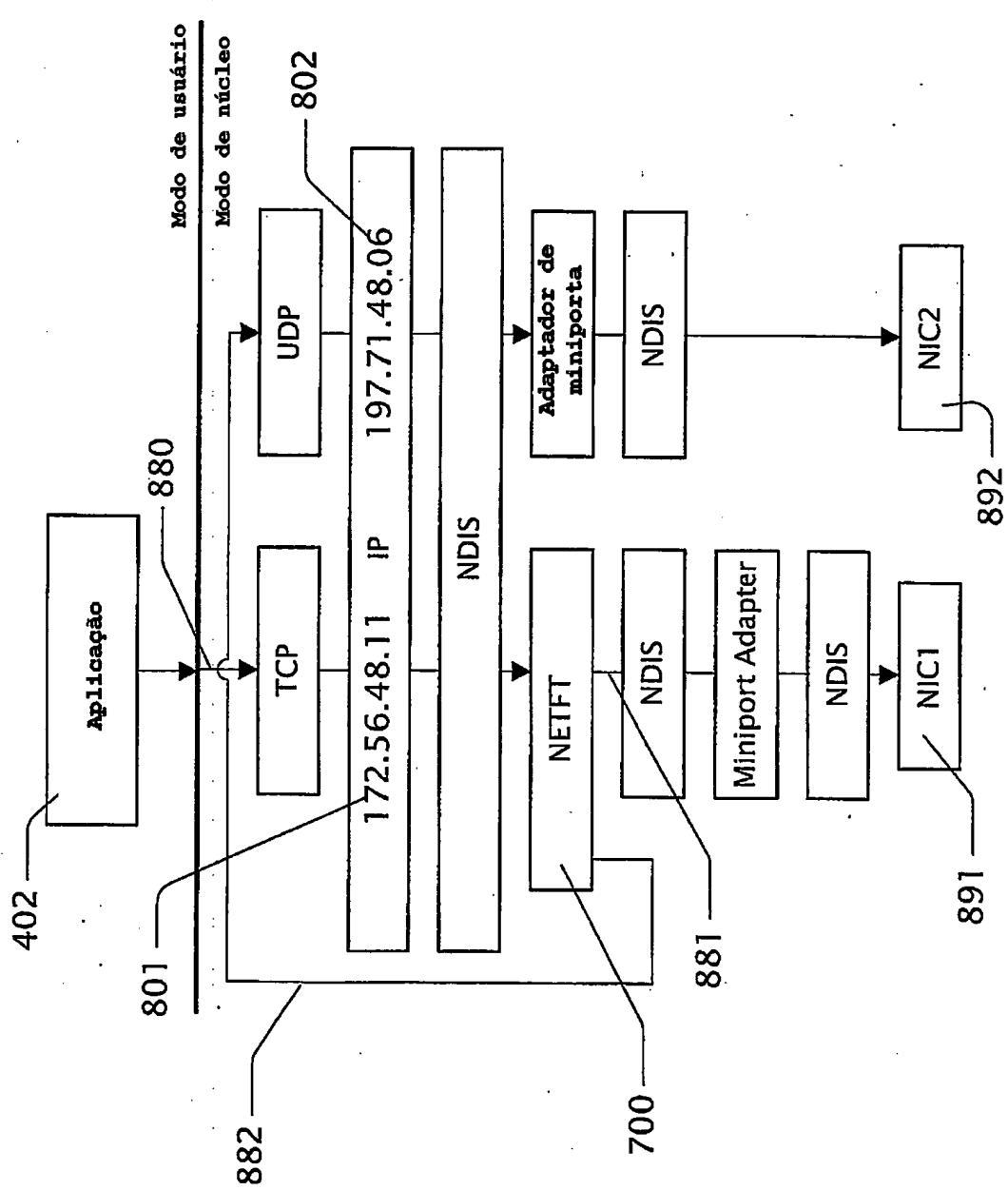
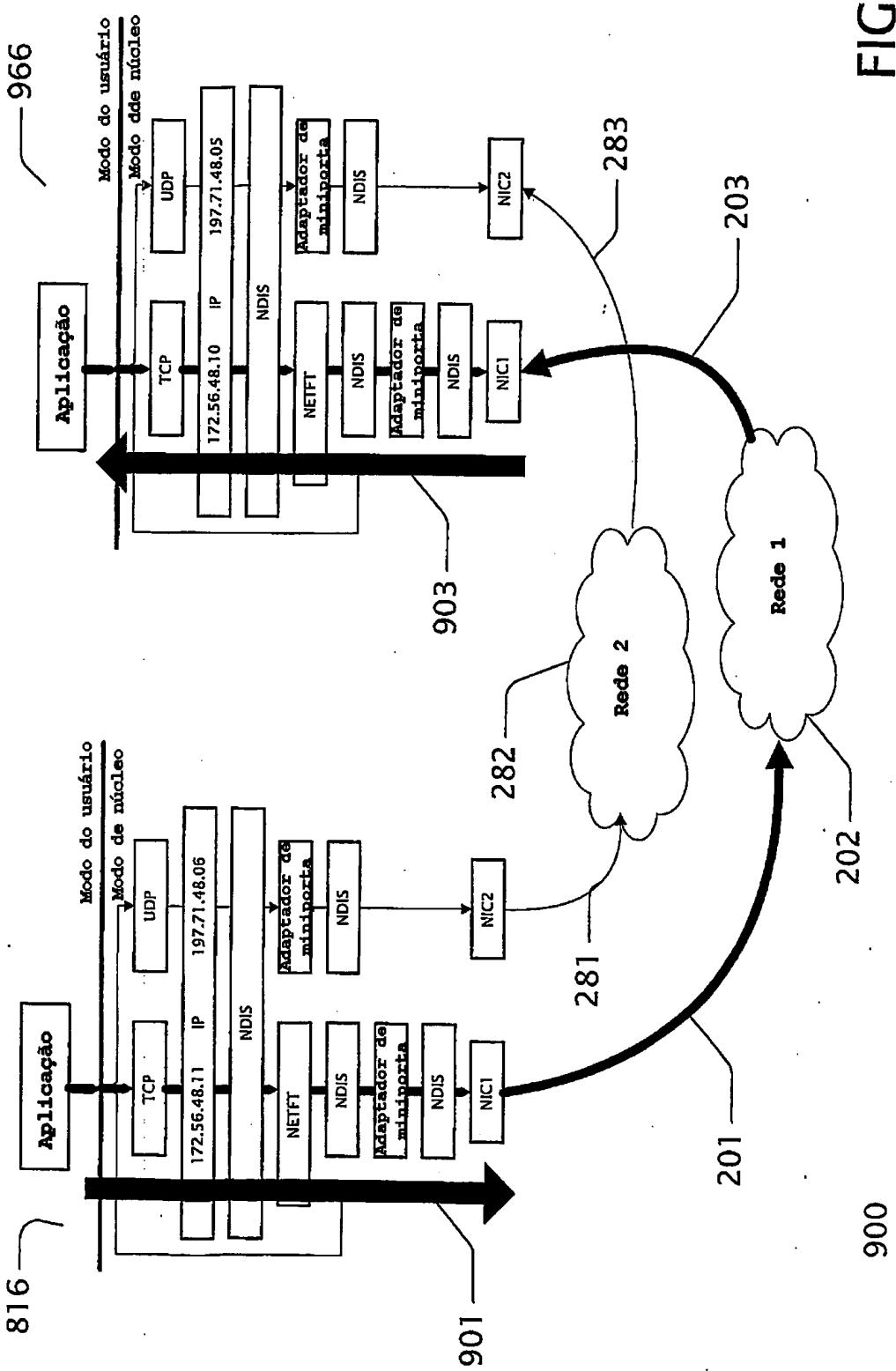


FIG. 9



900

202

203

283

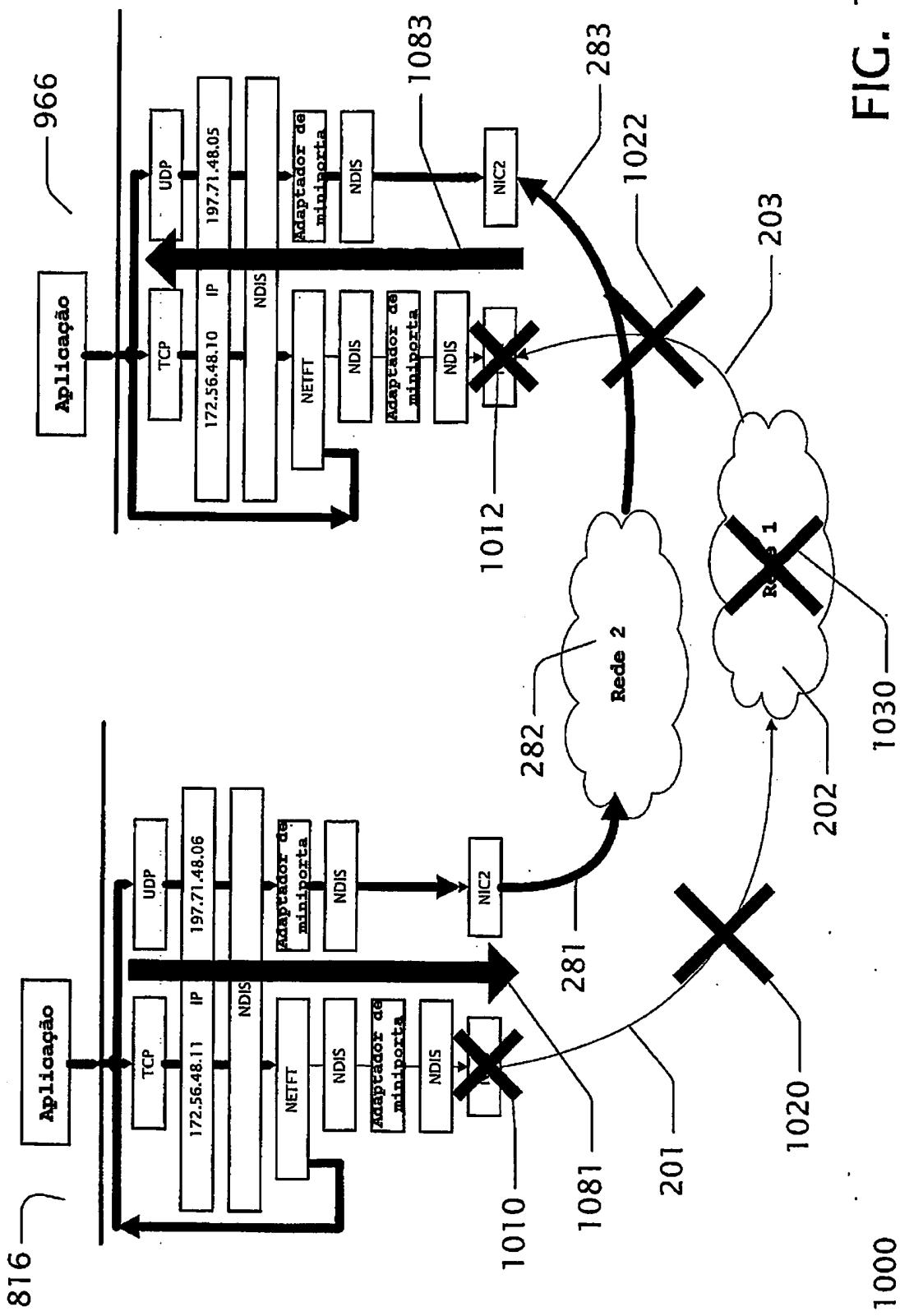
282

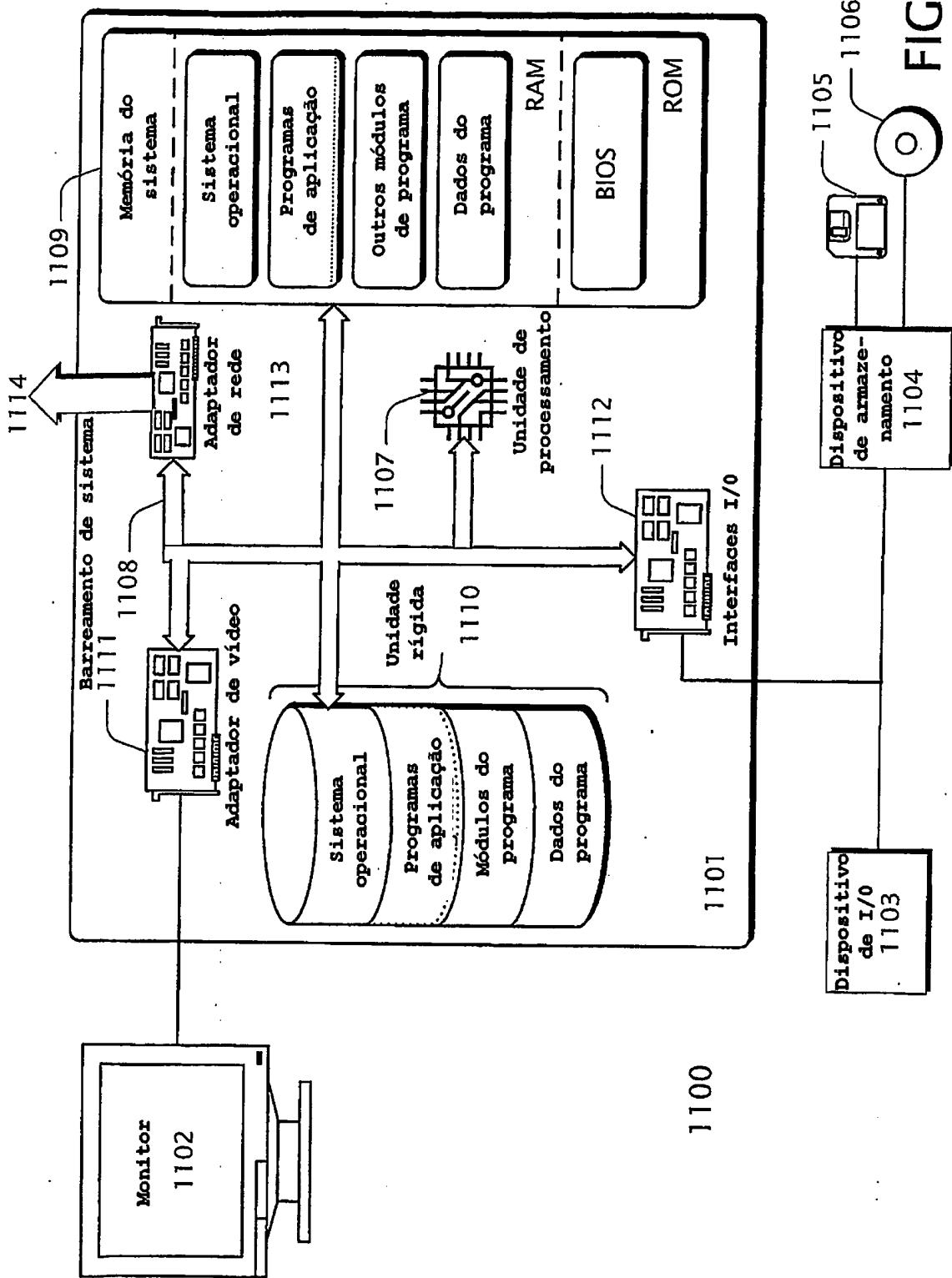
281

901

903

966





10615816-1

RESUMO

"COMUNICAÇÕES TOLERANTE A FALTA EM REDES ROTEADAS"

Um método para prover comunicações de rede tolerante à falta entre uma pluralidade de nós para uma aplicação, incluindo prover uma pluralidade de caminhos de comunicações iniciais sobre uma pluralidade de redes unidas entre a pluralidade de nós, receber um pacote de dados em um nó de envio da aplicação, o nó de envio sendo um da pluralidade de nós, o pacote de dados sendo endereçado pela aplicação para um endereço em um da pluralidade de nós e selecionar um primeiro caminho selecionado para o pacote de dados dentre a pluralidade de caminhos de comunicações iniciais onde o primeiro caminho selecionado é um caminho preferido.