

(11) Número de Publicação: **PT 2176771 E**

(51) Classificação Internacional:
G06F 13/10 (2009.01)

(12) FASCÍCULO DE PATENTE DE INVENÇÃO

<p>(22) Data de pedido: 2009.02.10</p> <p>(30) Prioridade(s): 2008.02.14 US 31038</p> <p>(43) Data de publicação do pedido: 2010.04.21</p> <p>(45) Data e BPI da concessão: 2010.12.22 015/2011</p>	<p>(73) Titular(es): INTERNATIONAL BUSINESS MACHINES CORPORATION NEW ORCHARD ROAD ARMONK, NY 10504 US</p> <p>(72) Inventor(es): DANIEL CASPER US MARK BENDYK US JOHN FLANAGAN US CATHERINE HUANG US MATTHEW KALOS US</p> <p>(74) Mandatário: ELSA MARIA MARTINS BARREIROS AMARAL CANHÃO PT RUA DO PATROCÍNIO 94 1399-019 LISBOA</p>
--	--

(54) Epígrafe: **PROVISÃO DE ENDEREÇAMENTO INDIRECTO DE DADOS NUM SISTEMA DE PROCESSAMENTO DE ENTRADA/SAÍDA EM QUE A LISTA DE ENDEREÇOS INDIRECTOS DE DADOS É NÃO-CONTIGUA**

(57) Resumo:

SISTEMAS, MÉTODOS E PRODUTOS DE PROGRAMA DE COMPUTADOR PARA PROPORCIONAR ENDEREÇAMENTO DE DADOS INDIRECTO NUM SUBSISTEMA DE I/O DE UM SISTEMA DE PROCESSAMENTO DE I/O. O PRODUTO DE PROGRAMA DE COMPUTADOR INCLUI UM MEIO DE ARMAZENAMENTO TANGÍVEL LEGÍVEL POR UM CIRCUITO DE PROCESSAMENTO E ARMAZENANDO INSTRUÇÕES PARA EXECUÇÃO PELO CIRCUITO DE PROCESSAMENTO PARA EXECUTAR UM MÉTODO. O MÉTODO INCLUI RECEBER UMA PALAVRA DE CONTROLO PARA UMA OPERAÇÃO DE I/O. A PALAVRA DE CONTROLO INCLUI UM ENDEREÇO DE DADOS INDIRECTO PARA DADOS ASSOCIADOS COM A OPERAÇÃO DE I/O. O ENDEREÇO DE DADOS INDIRECTO INCLUI UM LOCAL INICIAL DE UMA LISTA DE ENDEREÇOS DE ARMAZENAMENTO QUE ESPECIFICAM OS DADOS COLECTIVAMENTE, A LISTA ABRANGE DOIS OU MAIS LOCAIS DE ARMAZENAMENTO NÃO-CONTÍGUOS. OS DADOS SÃO REUNIDOS EM RESPOSTA À LISTA. OS DADOS REUNIDOS SÃO TRANSMITIDOS PARA UMA UNIDADE DE CONTROLO NO SISTEMA DE PROCESSAMENTO DE I/O.

DESCRIÇÃO

"PROVISÃO DE ENDEREÇAMENTO INDIRECTO DE DADOS NUM SISTEMA DE PROCESSAMENTO DE ENTRADA/SAÍDA EM QUE A LISTA DE ENDEREÇOS INDIRECTOS DE DADOS É NÃO-CONTÍGUA"

CAMPO DA INVENÇÃO

A presente revelação refere-se, em geral, a um processamento de entrada/saída e, em particular, a proporcionar uma lista de endereçamento indirecto de dados não-contígua num subsistema I/O de um sistema de processamento I/O.

DESCRIÇÃO DOS ANTECEDENTES DA TÉCNICA

As operações de entrada/saída (I/O) são utilizadas para transferir dados entre a memória e dispositivos de I/O de um sistema de processamento I/O. Especificamente, os dados são escritos, a partir de uma memória, num ou mais dispositivos de I/O e os dados são extraídos de um ou mais dispositivos de I/O para a memória executando operações de I/O.

Para facilitar o processamento de operações de I/O, é empregue um subsistema de I/O do sistema de processamento I/O. O subsistema de I/O está acoplado à memória principal e aos dispositivos de I/O do sistema de processamento I/O e direcciona o fluxo de informação entre a memória e os dispositivos I/O. Um exemplo de um subsistema I/O é um subsistema de canal. O

subsistema de canal utiliza percursos de canal como meios de comunicação. Cada percurso de canal inclui um canal acoplado a uma unidade de controlo, estando a unidade de controlo, além disso, acoplada a um ou mais dispositivos de I/O.

O subsistema de canal pode empregar palavras de comando de canal (CCW) para transferir dados entre os dispositivos de I/O e a memória. Uma CCW especifica o comando a ser executado. Para comandos iniciando certas operações de I/O, a CCW designa a área de memória associada com a operação, a acção a ser tomada sempre que uma transferência de ou a partir da área é completada e outras opções.

Durante o processamento de I/O, um canal obtém uma lista de CCW a partir da memória. O canal analisa cada comando a partir da lista de CCW e encaminha vários comandos, cada comando como uma entidade própria, para uma unidade de controlo acoplada ao canal. A unidade de controlo processa, então, os comandos. O canal segue o estado de cada comando e controla o momento de envio do próximo conjunto de comandos para a unidade de controlo para processamento. O canal assegura que cada comando seja enviado para a unidade de controlo como uma entidade própria. Além disso, o canal infere certas informações associadas com o processamento da resposta a partir da unidade de controlo para cada comando.

Executar o processamento I/O à base de CCW pode envolver uma grande quantidade de processamento adicional para o subsistema de canal, uma vez que os canais analisam as CCW, seguem a informação de estado e reagem às respostas das unidades de controlo. Portanto, pode ser benéfico desviar muita da carga de processamento, associada com a interpretação e gestão das CCW

e da informação de estado do subsistema de canal, para as unidades de controlo. A simplificação do papel dos canais na comunicação entre as unidades de controlo e um sistema operativo no sistema de processamento I/O pode aumentar a capacidade de comunicação, uma vez que é executada menos troca de informação. Contudo, alterar as sequências de comando, bem como os papéis do subsistema de canal e das unidades de controlo, permite que a quantidade de dados que é transferida numa única operação seja superior a um megabyte. Esta é a máxima quantidade de dados que pode ser transferida com uma lista contínua de endereços de dados indirectos de transporte quando o tamanho de página do sistema é 4 k bytes. Presentemente, uma Palavra de Comando de Canal (CCW) existente não pode conter mais do que 64 k bytes numa única operação de I/O devido à limitação do campo de contagem de dois bytes na CCW. A Palavra de Controlo de Transporte (TCW) resolveu esse problema ao aumentar a contagem de bytes para quatro bytes na TCW, mas, depois surge a limitação seguinte de um megabyte devido à lista de endereços de dados indirectos de transporte (TIDAL) ter de estar contida numa página com 4 k bytes o que apenas permite 256 entradas na lista de endereços.

O documento US 5584039 divulga um método para proporcionar o endereçamento de dados indirecto num subsistema de I/O.

BREVE SUMÁRIO DA INVENÇÃO

A invenção proporciona um método como reivindicado na reivindicação 1 e o sistema e o programa de computador correspondentes.

BREVE DESCRIÇÃO DOS DESENHOS

O assunto que é considerado como a invenção é particularmente salientado e reivindicado distintamente nas reivindicações na conclusão da descrição. Os objectivos anteriores e outros objectivos, características e vantagens da invenção são evidentes a partir da descrição detalhada seguinte, considerada em conjunção com os desenhos anexos, nos quais:

A Fig. 1 mostra uma forma de realização de um sistema de processamento de I/O, incorporando e utilizando um ou mais aspectos da presente invenção;

A Fig. 2A mostra um exemplo de uma palavra de comando de canal da técnica anterior;

A Fig. 2B mostra um exemplo de um programa de canal de uma palavra de comando de canal da técnica anterior;

A Fig. 3 mostra uma forma de realização de um protocolo de ligação da técnica anterior, utilizado na comunicação entre um canal e uma unidade de controlo para executar o programa de canal de uma palavra de comando de canal da Fig. 2B;

A Fig. 4 mostra uma forma de realização de um programa de canal de uma palavra de controlo de transporte, de acordo com um aspecto da presente invenção;

A Fig. 5 mostra uma forma de realização de um protocolo de ligação utilizado para comunicar entre um canal e uma unidade de controlo para executar o programa de canal de uma palavra de controlo de transporte da Fig. 4, de acordo

com um aspecto da presente invenção;

A Fig. 6 mostra uma forma de realização de um protocolo de ligação da técnica anterior utilizado para comunicar entre um canal e uma unidade de controlo de modo a processar quatro comandos de escrita de um programa de canal de uma palavra de controlo de transporte;

A Fig. 7 mostra uma forma de realização de um protocolo de ligação utilizado para comunicar entre um canal e uma unidade de controlo para processar os quatro comandos de escrita de um programa de canal de uma palavra de controlo de transporte, de acordo com um aspecto da presente invenção;

A Fig. 8 mostra uma forma de realização de uma unidade de controlo e de um subsistema de canal, de acordo com um aspecto da presente invenção;

A Fig. 9 mostra uma forma de realização de uma palavra de endereço de dados indirecto do modo de transporte (TIDAW), de acordo com um aspecto da presente invenção;

A Fig. 10 mostra uma forma de realização de um processo para proporcionar o endereçamento de dados indirectos não-contíguos num subsistema I/O; e

A Fig. 11 mostra uma forma de realização de um artigo manufacturado incorporando um ou mais aspectos da presente invenção.

A descrição detalhada explica as formas de realização preferidas da invenção, conjuntamente com vantagens e características, a título de exemplo e recorrendo aos desenhos.

DESCRIÇÃO DETALHADA DA INVENÇÃO

De acordo com um aspecto da presente invenção, o processamento de entrada/saída (I/O) é facilitado. O processamento I/O é facilitado num exemplo, com a redução de comunicações entre componentes de um sistema de processamento I/O utilizado para executar o processamento de I/O. Por exemplo, é reduzido o número de trocas e sequências entre um adaptador de comunicações de I/O, tal como um canal, e uma unidade de controlo. Isto é conseguido enviando uma pluralidade de comandos do adaptador de comunicações I/O para a unidade de controlo como uma entidade única para execução pela unidade de controlo e pelo envio, pela unidade de controlo, dos dados resultantes dos comandos, se existirem, como uma entidade única.

A pluralidade de comandos é incluída num bloco, referido, daqui em diante, como bloco de controlo de comando de transporte (TCCB), cujo endereço é especificado numa palavra de controlo de transporte (TCW). A TCW é enviada por um sistema operativo (OS), ou outra aplicação, para o adaptador de comunicações de I/O que, por sua vez, encaminha o TCCB numa mensagem de comando para a unidade de controlo para processamento. A unidade de controlo processa cada comando que não possua um seguimento do estado relativo a esses comandos individuais por parte do adaptador de comunicações de I/O. A pluralidade de comandos é também referida como programa de canal, que é analisado e executado na unidade de controlo e não no adaptador de comunicações de I/O.

Numa forma de realização exemplificativa, a TCW para uma operação de I/O inclui apontadores indicando a localização dos dados de controlo (e. g., o TCCB) e os dados de utilizador associados com a operação de I/O. Numa forma de realização exemplificativa, os apontadores referem-se a um endereço de dados indirecto, aqui referido como lista de endereços de dados indirectos de modo de transporte (TIDAL). A TIDAL inclui uma lista de endereços onde os dados estão localizados; estes endereços são aqui referidos como palavras de endereço de dados indirecto de modo de transporte (TIDAW). No projecto de sistema actual com páginas de 4 k, a TIDAL não pode ser maior do que quatro kilobytes (4k), limitando, assim, o número de TIDAW contínuas que podem ser incluídas numa TIDAL única a 256 (cada TIDAL tem 16 bytes). Isto limita a quantidade total de dados que pode ser transmitida por uma operação de I/O a um megabyte, assumindo páginas de 4k, cada TIDAL tem 16 bytes e que a TIDAL tem de estar contida numa página de 4k. Numa forma de realização exemplificativa da presente invenção, esta limitação é removida permitindo que um endereço TIDAW indique a localização de início das próximas TIDAW noutra localização de armazenamento que pode estar numa página diferente. Um indicador na TIDAW é utilizado para indicar se o endereço da TIDAW inclui um endereço de dados ou um endereço da continuação da lista TIDAW. Desta forma, as TIDAW constituindo uma TIDAL única podem estar localizadas em locais de armazenamento não-contíguos e, portanto, podem ser incluídas num única lista TIDAL mais de 256 TIDAW. Isto permite que mais dados sejam transferidos numa única operação de I/O. Por exemplo, numa forma de realização exemplificativa, onde o campo de contagem na TCW é de quatro bytes, podem ser transferidos cerca de quatro gigabytes (quatro gigabytes menos um byte) numa única operação de I/O.

Um exemplo de um sistema de processamento de I/O incorporando e utilizando um ou mais aspectos da presente invenção é descrito em referência à Fig. 1. O sistema 100 de processamento de I/O inclui um sistema 101 hospedeiro, que inclui, além disso, por exemplo, uma memória 102 principal, uma ou mais unidades 104 de processamento centrais (CPU), um elemento 106 de controlo de armazenamento e um subsistema 108 de canal. O sistema 101 hospedeiro pode ser um sistema de computação de larga escala, tal como um computador principal ou um servidor. O sistema 100 de processamento de I/O inclui também uma ou mais unidades 110 de controlo e um ou mais dispositivos 112 de I/O, cada um dos quais é descrito em seguida.

A memória 102 principal armazena dados e programas, que podem ser introduzidos a partir dos dispositivos 112 de I/O. Por exemplo, a memória 102 principal pode incluir um ou mais sistemas 103 operativos (OS) que são executados por um ou mais dos CPU 104. Por exemplo, uma CPU 104 pode executar um sistema 103 operativo Linux® e um sistema 103 operativo z/OS® como diferentes instâncias de máquina virtual. A memória 102 principal é endereçável directamente e assegura um processamento de dados de alta velocidade pelas CPU 104 e pelo subsistema 108 de canal.

A CPU 104 é o centro de controlo do sistema 100 de processamento de I/O. Contém meios de sequenciação e processamento para a execução de instruções, acção de interrupção, funções de sincronização, carregamento do programa inicial e outras funções relacionadas com o equipamento. A CPU 104 está ligada ao elemento 106 de controlo de armazenamento através de uma ligação 114, tal como um barramento bidireccional ou unidireccional.

O elemento 106 de controlo de armazenamento está ligado à memória 102 principal através de uma ligação 116, tal com um barramento; às CPU 104 através da ligação 114; e ao subsistema 108 de canal através de uma ligação 118. O elemento 106 de controlo de armazenamento controla, por exemplo, a colocação numa fila e a execução dos pedidos feitos pela CPU 104 e subsistema 108 de canal.

Numa forma de realização exemplificativa, o subsistema 108 de canal proporciona uma interface de comunicação entre o sistema 101 hospedeiro e as unidades 110 de controlo. O subsistema 108 de canal está ligado ao elemento 106 de controlo de armazenamento, como descrito anteriormente, e a cada unidade 110 de controlo através de uma ligação 120, tal como uma ligação série. A ligação 120 pode ser implementada como uma ligação óptica, empregando guias de onda monomodo ou multimodo num tecido de Canal em Fibra. O subsistema 108 de canal dirige o fluxo de informação entre os dispositivos 112 de I/O e a memória 102 principal. Liberta as CPU 104 da tarefa de comunicar directamente com os dispositivos 112 de I/O e permite que o processamento de dados prossiga paralelamente com o processamento de I/O. O subsistema 108 de canal utiliza um ou mais percursos 122 de canal como as ligações de comunicação na gestão do fluxo de informação para ou dos dispositivos 112 de I/O. Como parte do processamento de I/O, o subsistema 108 de canal também executa as funções de gestão de percursos para testar a disponibilidade de percursos de canal, seleccionando um percurso 122 de canal disponível e iniciando a execução da operação com os dispositivos 112 de I/O.

Cada percurso 122 de canal inclui um canal 124 (os canais 124 estão localizados dentro do subsistema 108 de canal, num

exemplo, como mostrado na Fig. 1), uma ou mais unidades 110 de controlo e uma ou mais ligações 120. Noutro exemplo, é também possível ter um ou mais comutadores dinâmicos (não representados) como parte do percurso 122 de canal. Um comutador dinâmico está ligado a um canal 124 e a uma unidade 110 de controlo e proporciona a capacidade de interligar fisicamente quaisquer duas ligações que estejam ligadas ao comutador. Noutro exemplo, é também possível ter múltiplos sistemas e, portanto, múltiplos subsistemas de canal (não representados) ligados à unidade 110 de controlo.

Também existem sub-canais (não mostrados) localizados dentro do subsistema 108 de canal. Um sub-canal é proporcionado para e dedicado a cada dispositivo 112 de I/O acessível a um programa através do subsistema 108 de canal. Um sub-canal (e. g., uma estrutura de dados, tal como uma tabela) proporciona a aparência lógica de um dispositivo para o programa. Cada sub-canal proporciona informação relativa ao dispositivo 112 I/O associado e a sua ligação ao subsistema 108 de canal. O sub-canal também proporciona informação relativa às operações de I/O e outras funções envolvendo o dispositivo 112 de I/O associado. O sub-canal é o meio pelo qual o subsistema 108 de canal proporciona informação à cerca dos dispositivos 112 de I/O associados às CPU 104, que obtêm esta informação executando instruções de I/O.

O subsistema 108 de canal está ligado a uma ou mais unidades 110 de controlo. Cada unidade 110 de controlo proporciona a lógica para operar e controlar um ou mais dispositivos 112 de I/O e adapta, através da utilização de meios comuns, as características de cada dispositivo 112 de I/O ao interface de ligação proporcionado pelo canal 124. As meios

comuns proporcionam, para a execução de operações de I/O, indicações relativas o estado do dispositivo 112 de I/O e da unidade 110 de controlo, controlo dos dados de temporização sobre o percurso 122 de canal e certos níveis de controlo do dispositivo 112 de I/O.

Cada unidade 110 de controlo está ligada através de uma ligação 126 (e. g., um barramento) a um ou mais dispositivos 112 I/O. Os dispositivos 112 I/O recebem informação ou armazenam informação na memória 102 principal e/ou outra memória. Exemplos de dispositivos 112 de I/O incluem leitores de cartões e furadores, unidades de banda magnética, dispositivos de armazenamento de acesso directo, monitores, teclados, impressoras, dispositivos apontadores, dispositivos de teleprocessamento, controladores de comunicação e equipamento baseado em sensores, para nomear alguns.

Um ou mais dos componentes anteriores do sistema de processamento 100 de I/O são, além disso, descritos em "IBM® z/Architecture Principles of Operation," Publicação Nº SA22-7832-05, 6^a Edição, Abril 2007, na Patente U.S. Nº 5461721 intitulada "System For Transferring Data Between I/O Devices And Main Or Expanded Storage Under Dynamic Control Of Independent Indirect Access Words (IDAWS)," Cormier et al., publicado em 24 de Outubro de 1995; e na Patente U.S. Nº 5526484 intitulada "Method And System For Pipelining The Processing Of Channel Command Words," Casper et al, publicado em 11 de Junho de 1996. IBM é uma marca registada da International Business Machines Corporation, Armonk, Nova Iorque, EUA. Outros nomes utilizados aqui podem ser marcas registadas, marcas ou nomes de produtos da International Business Machines Corporation ou outras empresas.

Numa forma de realização, são utilizadas palavras de comando de canal (CCW) para transferir dados entre os dispositivos 112 de I/O e a memória 102. Uma CCW especifica o comando a ser executado e inclui outros campos para processamento de controlo. Um exemplo de uma CCW é descrito em referência à Fig. 2A. Uma CCW 200 inclui, por exemplo, um código 202 de comando especificando o comando a ser executado (e. g., ler, ler para trás, controlo, amostrar e escrever); uma pluralidade de bandeiras 204 utilizadas para controlar a operação de I/O; para comandos que especificam a transferência de dados, um campo 206 de contagem que especifica o número de bytes na área de armazenamento designada pelo CCW para ser transferida; e um endereço 208 de dados que aponta para uma localização na memória principal que inclui dados, quando o endereçamento directo é utilizado, ou para uma lista (e. g., uma lista contínua) de palavras de endereço de dados indirecto modificadas (MIDAW) para serem processadas, quando o endereçamento de dados indirecto modificado é utilizado. O endereçamento indirecto modificado é descrito, além disso, no pedido de Patente Publicada U.S. Número 20080043563, intitulada "Flexibly Controlling The Transfer Of Data Between Input/Output Devices And Memory," Brice et al., apresentada a 15 de Agosto de 2006.

Uma ou mais CCW dispostas para execução sequencial formam um programa de canal, também referido aqui por programa de canal CCW. O programa de canal CCW é instalado por, por exemplo, um sistema operativo, ou outro software. O software instala as CCW e obtem os endereços de memória atribuídos ao programa de canal. Um exemplo de um programa de canal CCW é descrito em referência à Fig. 2B. Um programa 210 de canal CCW inclui, por exemplo, uma CCW de definição de extensão 212 que tem um apontador 214 para

uma localização de memória para dados de definição de extensão 216 para ser utilizado com o comando de definição de extensão. Neste exemplo, uma transferência no canal 218 (TIC) segue-se ao comando de definição de extensão que refere o programa de canal para outra área de memória (e. g., uma área de aplicação) que inclui um ou mais CCW, de forma a localizar o registo 217 que tem um ponteiro 219 para localizar os dados de registo 220, e um ou mais CCW 221 de leitura. Cada CCW 221 de leitura tem um ponteiro 222 para uma área 224 de dados. A área de dados inclui um endereço para aceder directamente aos dados ou uma lista de palavras de endereços de dados (e. g., MIDAW ou IDAW) para aceder aos dados indirectamente. Além disso, o programa 210 de canal CCW inclui uma área pré-determinada no subsistema de canal definida pelo endereço do dispositivo chamado o sub-canal para o estado 226 resultante da execução do programa de canal CCW.

O processamento de um programa de canal CCW é descrito em referência à Fig. 3, bem como em referência à Fig. 2B. Em particular, a Fig. 3 mostra um exemplo das várias trocas e sequências que ocorrem entre um canal e uma unidade de controlo quando um programa de canal CCW está a ser executado. O protocolo de ligação utilizado para a comunicação é o FICON (Conectividade por Fibra), neste exemplo. A informação relativa ao FICON é descrita em "Fibre Channel Single Byte Command Code Sets-3 Mapping Protocol (FC-SB-3), T11/Projecto 1357-D/Rev. 1.6, INCITS (Março de 2003).

Em referência à Fig. 3, um canal 300 inicia uma troca com uma unidade 302 de controlo e envia um comando de definição de extensão e os dados 304 a este associados para a unidade 302 de controlo. O comando é buscado a partir da CCW de definição de extensão 212 (Fig. 2B) e os dados são obtidos a partir da área

de dados de definição de extensão 216. O canal 300 utiliza o TIC 218 para localizar o CCW de localização de registo e a CCW de escrita. Obtém o comando de localização de registo 306 (Fig. 3) a partir da CCW de localização de registo 217 (Fig. 2B) e obtém os dados a partir dos dados de localização de registo 220. O comando 308 de escrita e dados (Fig. 3) são obtidos a partir da CCW de escrita 221 e área 224 de dados (Fig. 2B). Cada um destes é enviado para a unidade 302 de controlo.

A unidade 302 de controlo inicia uma troca 310 com o canal 300, em resposta à troca aberta do canal 300. Isto pode ocorrer antes ou depois do comando 306 de localização e dados e/ou do comando 308 e dados de escrita. Juntamente com a troca aberta, é encaminhada uma resposta (CMR) para o canal 300. A CMR proporciona uma indicação para o canal 300 que a unidade 302 de controlo está activa e em operação.

A unidade 302 de controlo proporciona o estado ao canal 300 e fecha a troca 312. Em resposta a isto, o canal 300 armazena os dados, examina o estado e fecha a troca 314, que indica à unidade 302 de controlo que o estado foi recebido.

O processamento do programa de canal CCW anterior para escrever 4k de dados requer que duas trocas sejam abertas e fechadas e seis sequências. O número total de trocas e sequências entre o canal e a unidade de controlo é reduzido através do colapsar de múltiplos comandos do programa de canal num TCCB. O canal, e. g. o canal 124 da Fig. 1, utiliza uma TCW para identificar a localização do TCCB, bem como as localizações para aceder a e armazenar o estado e os dados associados com a execução do programa de canal. A TCW é interpretada pelo canal e não é enviada ou vista pela unidade de controlo.

Um exemplo de um programa de canal para escrever 4k de dados, como na Fig. 2B, mas que inclui um TCCB, em vez de CCW individuais separadas, é descrito em referência à Fig. 4. Como mostrado, um programa 400 de canal, referido aqui como um programa de canal TCW, inclui uma TCW 402 especificado uma localização na memória de um TCCB 404, bem como uma localização na memória de uma área 406 de dados ou uma TIDAL 410 (*i. e.*, uma lista de palavras de endereço de dados indirecto de modo de transporte (TIDAW), semelhante às MIDAWs) que aponta para a área 406 de dados, e uma área 408 de estado. A TIDAW 412 zero, a TIDAW 414 um e a TIDAW 416 dois (colectivamente TIDAW 412-416) podem referenciar diferentes localizações na área 406 de dados para adquirir ou armazenar dados. As TIDAW 412-416 na TIDAL 410 podem estar localizadas sequencialmente na memória ou localizadas de forma não-contígua relativamente umas às outras. Apesar de apenas três TIDAW 412-416 estarem representadas na TIDAL 410, deve entender-se que qualquer número de TIDAW pode ser incluído na TIDAL 410.

O processamento de um programa de canal TCW é descrito em referência à Fig. 5. O protocolo de ligação utilizado para estas comunicações é, por exemplo, o Protocolo de Canal de Fibra (FCP). Em particular, são utilizadas 3 fases de protocolo de ligação FCP, permitindo que sejam utilizados adaptadores de barramento hospedeiro que suportam FCP para executar transferências de dados controladas por CCW. O FCP e as suas fases são descritas, além disso, em "Information Technology - Fibre Channel Protocol for SCSI, Terceira Versão (FCP-3)," T10 Projecto 1560-D, Revisão 4, 13 de Setembro de 2005.

Em referência à Fig. 5, um canal 500 inicia uma troca com uma unidade 502 de controlo e envia o TCCB 504 para a unidade

502 de controlo. Num exemplo, o TCCB 504 é a iniciativa de sequência são transferidos para a unidade 502 de controlo num comando FCP, referido como unidade de informação (IU) FCP_CMND ou um comando de transporte IU. A unidade 502 de controlo transmite uma mensagem de pronto para transferir (XFER_RDY) IU 510 para o canal 500 quando está pronta para receber dados para os comandos de escrita recebidos no TCCB 504. Em resposta a receber a XFER_RDY IU 510, o canal 500 transfere os dados 506 para a unidade 502 de controlo, através de, por exemplo, uma FCP_Data IU. A unidade 502 de controlo executa os múltiplos comandos do TCCB 504 (e. g., comando de definição de extensão, comando de localização de registo, comando de escrita como palavras de controlo de dispositivo (DCW)) e escreve os dados 506 recebidos a partir do canal 500. Também proporciona o estado e termina a troca 508. Num exemplo, o estado final é enviado numa trama de estado FCP que tem um bit activo no, por exemplo, byte 10 ou 11 da carga de um FCP_RSP IU, também referido como resposta de transporte IU. A carga FCP_RSP IU pode ser utilizada para transportar o estado de terminação do FICON juntamente com informação de estado adicional.

A forma de realização do protocolo de ligação mostrada na Fig. 5 é utilizada quando a XFER_RDY está activa. Na forma de realização mostrada na Fig. 5, o canal 500 não pode enviar os dados 506 para a unidade 502 de controlo até que tal seja pedido pela unidade 502 de controlo através da XFER_RDY IU 510. Numa forma de realização exemplificativa alternativa, a XFER_RDY é desactivada e a unidade de controlo não transmite uma XFER_RDY IU 510 para o canal 500. Assim, o canal 500 não tem de esperar que a unidade 502 de controlo peça os dados 506 antes de enviar os dados 506. Esta forma de realização alternativa, na qual a XFER_RDY é desactivada pode ser utilizada quando o canal 500 e a

unidade 502 de controlo estão localizados em locais geográficos muito distantes entre si (e. g., mais de vinte quilómetros, mais de cinquenta quilómetros) para melhorar o desempenho. A não ser que especificado doutra forma, a discussão neste documento assume que a XFER_RDY está activa.

Noutro exemplo, para escrever 4k de dados de cliente, o canal 500 utiliza as fases do protocolo de ligação FCP, como se segue:

1. Transferir o TCCB no FCP_CMND IU e a iniciativa de sequência para a unidade 502 de controlo.
2. Esperar por uma XFER_RDY IU indicando que a unidade de controlo está pronta para receber os dados.
3. Transferir a IU de dados, e a iniciativa de sequência para a unidade 502 de controlo.
4. O estado final é enviado numa trama de estado FCP que tem um bit activo no, por exemplo, byte 10 ou 11 da carga de um FCP_RSP IU. O campo FCP_RSP_INFO ou campo de amostragem é utilizado para transportar o estado de terminação do FICON juntamente com informação de estado adicional.

Executando o programa de canal TCW da Fig. 4, exite apenas uma troca iniciada e terminada (ver também a Fig. 5), em vez de duas trocas para o programa de canal CCW da Fig. 2B (ver também a Fig. 3). Além disso, para o programa de canal TCW, existem quatro sequências de comunicação (ver Figs. 4-5), em comparação com as seis sequências para o programa de canal CCW (ver Figs. 2B-3).

O número de trocas e sequências permanece o mesmo para um programa de canal TCW, mesmo se comandos adicionais tiverem sido adicionados ao programa. Comparem-se, por exemplo, as comunicações do programa de canal CCW da Fig. 6 com as comunicações do programa de canal TCW da Fig. 7. No programa de canal CCW da Fig. 6, cada um dos comandos (e. g., o comando 600 de definição de extensão, o comando 601 de localização de registo, o comando 602 e dados de escrita, o comando 604 e dados de escrita, o comando 606 e dados de localização de registo, o comando 608 e dados de escrita, e o comando 620 e dados de escrita) é enviado em sequências separadas a partir do canal 610 para a unidade 612 de controlo. Este programa de canal CCW requer o início e a terminação de duas trocas (e. g., trocas iniciadas 622, 624 e trocas terminadas 626, 628) e dez sequências de comunicação. Isto é comparado com as quatro sequências e uma troca para o programa de canal TCW da Fig. 7, que consegue efectuar a mesma tarefa que o programa de canal CCW da Fig. 6.

Como mostrado na Fig. 7, um canal 700 inicia uma troca com uma unidade 702 de controlo e envia um TCCB 704 para a unidade 702 de controlo. O TCCB 704 inclui o comando de definição de extensão, os dois comandos de localização de registo e os quatro comandos de leitura em DCW, como descrito anteriormente. De forma semelhante ao exemplo mostrado na Fig. 5, a unidade 702 de controlo pode utilizar uma XFER_RDY IU 710 para notificar o canal 700 que está pronto a receber dados, desde que o suporte à XFER_RDY não esteja desactivado. O canal 700 transmite 16k de dados 706 para a unidade de controlo 702 numa única sequência após a recepção da XFER_RDY IU 710. O canal 700 insere um CRC a cada 4k dos 16k dos dados 706 na sequência. A inserção de um CRC a cada 4k permite à unidade 702 de controlo verificar os 16k de

dados incrementalmente, em vez de armazenar todos os 16k para verificação antes de completar os comandos de escrita no TCCB 704. Adicionalmente, a unidade 702 de controlo proporciona o estado para o canal 700 e termina a troca 708. Assim, o programa de canal TCW da Fig. 7 requer muito menos comunicações para transferir a mesma quantidade de dados do que o programa de canal CCW da Fig. 6, apesar de suportar a verificação de dados incremental através da inserção de CRC múltiplos no fluxo de dados de saída a partir do canal 700.

Considerando agora a Fig. 8, uma forma de realização do canal 124 no subsistema 108 de canal e a unidade 110 de controlo da Fig. 1 que suporta a execução do programa de canal TCW é mostrada em maior detalhe. A unidade 110 de controlo inclui a lógica 802 de controlo CU para analisar e processar mensagens de comando contendo um TCCB, tal como o TCCB 704 da Fig. 7, bem como dados recebidos a partir do canal 124 através da ligação 120. A lógica de controlo CU 802 pode extraír DCW e dados de controlo a partir do TCCB recebido na unidade 110 de controlo para controlar um dispositivo, por exemplo, o dispositivo 112 de I/O através da ligação 126. A lógica 802 de controlo CU envia comandos de dispositivo e dados para o dispositivo 112 de I/O, e recebe informação de estado e outra informação de retorno a partir do dispositivo 112 de I/O. Quando a lógica 802 de controlo recebe dados, tais como o primeiro limite de bloco de verificação dos 16 k bytes de dados 706 da Fig. 7, a lógica 802 de controlo CU escreve estes dados recebidos no buffer 804 de dados para armazenamento temporário, até que o CRC recebido para o limite de bloco de verificação seja verificado, depois, os dados podem ser enviados para o dispositivo 112 de I/O. Esta operação continua para cada limite de bloco de verificação até que a operação de I/O esteja concluída.

A unidade 110 de controlo pode, além disso, incluir outras listas de dados ou elementos de memória (não mostrados) para armazenar mensagens adicionais ou informação de estado associada com comunicações entre o canal 124 e o dispositivo 112 de I/O.

O canal 124 no subsistema 108 de canal inclui elementos para suportar a comunicação com a unidade 110 de controlo. Por exemplo, o canal 124 pode incluir lógica 806 de controlo CHN que faz a interface com a lógica 812 de recolha de dados CHN. A lógica 812 de recolha de dados é descrita, em seguida, recorrendo à Fig. 10. Numa forma de realização exemplificativa, a lógica 806 de controlo CHN controla a comunicação entre o subsistema 108 de canal e a unidade 110 de controlo. A lógica 806 de controlo CHN pode interagir directamente com a lógica 802 de controlo CU através da ligação 120 para enviar comandos e receber respostas, tais como comandos de transporte e IU de resposta. Em alternativa, as interfaces de transferências de mensagens e/ou buffers (não mostrados) adicionais podem ser colocados entre a lógica 806 de controlo CHN e a lógica 802 de controlo CU. Os registos 814 de subsistema CHN podem incluir valores fixos que proporcionam informação de configuração e estado, bem como informação de estado dinâmica, actualizados quando os comandos são transportados e respostas são recebidas. Os registos de subsistema 814 CHN podem ser registos de hardware dedicados e/ou registos virtuais estabelecidos utilizando o mapeamento de memória.

Numa forma de realização, os registos do subsistema 814 CHN incluem a TIDAL 410 e as TIDAW 412-416 da Fig. 4 como registos mapeados na memória.

Um exemplo de uma TIDAW 900 é mostrado na Fig. 9. A TIDAW 900 proporciona endereçamento indirecto aos dados utilizados num programa de canal TCW, tal como as TIDAW 412-416 da Fig. 4. A TIDAW 900 inclui o indicador 902, uma contagem 904 e um endereço 906. Cada campo (*i. e.*, indicadores 902, contagem 904 e endereço 906) no formato da TIDAW 900 é atribuído a um byte de endereço particular para suportar a análise dos campos. Apesar de um arranjo dos campos dentro da TIDAW 900 ser mostrado na Fig. 9, deve entender-se que a ordem dos campos pode ser rearranjada para obter ordenações alternativas.

Numa forma de realização exemplificativa, os indicadores 902 incluem um último indicador TIDAW e um indicador transporte-transferência no canal (T-TIC), em adição a outros indicadores. O último indicador TIDAW indica que a TIDAW associada é a última TIDAW numa TIDAL, de forma consistente com a definição de MIDAW. Quando a contagem atinge zero com o último conjunto de indicadores TIDAW, a transferência de dados para a operação associada de I/O está completa. O indicador T-TIC indica se o conteúdo do endereço 906 inclui dados ou o endereço da próxima TIDAW na TIDAL. Numa forma de realização exemplificativa, quando o indicador T-TIC é estabelecido, então, o endereço 906 na TIDAW é o endereço da próxima TIDAW na TIDAL. Da mesma forma, o endereço 906 pode ser utilizado para aceder a uma TIDAW num local de armazenamento não-contígua à TIDAW actual. Assim, a lista de TIDAW pode conter mais de 256 entradas, excedendo portanto o número máximo de palavras de endereço de dados indirecto permitido nas implementações actuais. Numa forma de realização exemplificativa, quando o indicador T-TIC é estabelecido, o endereço 906 deve ter os quatro bits de ordem mais baixa colocados a zero porque a TIDAW de 16 bytes deve estar numa fronteira de endereços de 16 bytes.

Quando o indicador T-TIC não está estabelecido, então, o endereço 906 na TIDAW é o endereço de uma porção de dados que constitui os dados a serem reunidos para a operação de I/O. A quantidade dos dados é indicada no campo 904 de contagem. Se o último indicador TIDAW e o indicador T-TIC não estiverem estabelecidos, então a próxima TIDAW está situada na próxima localização de armazenamento (e. g., está contígua à TIDAW actual).

Em referência agora à Fig. 10, será agora descrito um processo 1000 para reunir dados para uma operação de I/O de acordo com formas de realização exemplificativas, e em referência ao sistema de processamento 100 de I/O da Fig. 1. No bloco 1002, o subsistema 108 de canal recebe uma palavra de controlo para uma operação de I/O (e. g., uma TCW). A palavra de controlo inclui um endereço de dados indirecto que aponta para o endereço inicial de uma lista de endereços de armazenamento (e. g., uma TIDAL possuindo múltiplas TIDAW) que reúnem os dados utilizados pela operação de I/O. Em formas de realização exemplificativa, os dados são dados de cliente (e. g., dados de entrada ou de saída da operação de I/O). Em formas de realização alternativas exemplo, os dados são dados de controlo (e. g., um TCCB). Numa forma de realização exemplificativa, a lista de endereços de armazenamento abrange duas ou mais locais de armazenamento não-contíguas.

No bloco 1004, os dados são reunidos por instruções localizadas no subsistema 108 de canal. A reunião é baseada no conteúdo da lista. Numa forma de realização exemplificativa, cada entrada da lista (e. g., cada TIDAW) inclui um endereço de armazenamento bem como um indicador T-TIC para indicar se o endereço de armazenamento é a localização de uma porção dos

dados ou se o endereço de armazenamento aponta para a localização doutra porção da lista contendo mais endereços de armazenamento. Desta maneira, uma única TIDAL pode estar contida em múltiplas páginas removendo quaisquer restrições sobre o comprimento da TIDAL. Quando o indicador T-TIC indica que o endereço de armazenamento é a localização de uma porção de dados (e. g., o indicador T-TIC não está estabelecido), então são acedidos os dados situados na localização de armazenamento e adicionados aos dados. Esta adição de novos dados pode ser executada de qualquer forma conhecida na técnica, tal como a fusão com os dados já reunidos, juntar aos dados já reunidos, etc.

Quando o indicador T-TIC indica que o endereço de armazenamento actual é a localização de outra porção da lista, então, o processamento continua acedendo a uma nova TIDAW localizada na localização de armazenamento especificada. Os dados situados na localização de armazenamento especificada na nova TIDAW são acedidos e adicionados aos dados. Numa forma de realização exemplificativa, a TIDAW inclui uma contagem 904 para especificar a quantidade de dados que deve ser lida (ou escrita) a partir de cada endereço de armazenamento. Quando uma TIDAW com o último indicador TIDAW é estabelecida a reunião está completada. Em caso contrário, a próxima TIDAW é acedida e a reunião de dados continua.

No bloco 1006, os dados são transmitidos para uma unidade 110 de controlo pelo subsistema 108 de canal.

Efeitos técnicos de formas de realização exemplificativa incluem proporcionar o endereçamento de dados indirectos não-contíguos num subsistema de I/O. Permitir que múltiplas

páginas de 4 k bytes contêm a lista de endereços de armazenamento (TIDAL) permite que o número de TIDAW seja ilimitado, e portanto mais dados podem estar associados com uma única operação de I/O. Grandes transferências de dados podem reduzir o conjunto de dados adicionais na comunicação evitando a troca de sinalização adicional e outros atrasos associados com múltiplas mensagens mais pequenas.

Os exemplos seguintes mostram formas nas quais formas de realização exemplificativa podem ser utilizadas para executar mais do que o movimento de grandes blocos de dados.

Formas de realização exemplificativa podem ser utilizadas para ajudar na pré-fixação de um programa de canal por um sistema operativo. Por exemplo, quando um programa de canal é passado para um sistema operativo, o sistema operativo pode modificar o programa de canal para adicionar, substituir, ou modificar CCW ou DCW de comando. Numa forma de realização exemplificativa, isto implica que um TCCB TIDAL seja criado onde os primeiros TIDAW apontam para uma área de armazenamento contendo a primeira parte do TCCB modificado. A segunda TIDAW aponta para o restante do TCCB não-modificado.

Formas de realização exemplificativa também podem ser utilizadas para ajudar na pré-fixação de programas de canal por um hipervisor. Se o hipervisor necessitar de modificar o programa de canal, este necessita de criar uma TIDAL com mais uma entrada do que aquela passada pelo visitante. Se a TIDAL passada pelo visitante já tiver o tamanho máximo (256 TIDAW), então o hipervisor pode fazer as suas próprias modificações criando um TCCB TIDAL que tem duas TIDAW, uma das quais é a T-TIC para a primeira TIDAW da TIDAL passada pelo visitante.

As T-TIC da TIDAL podem também facilitar os requisitos de armazenamento para sistemas operativos. Se o programa de canal contendo um TCCB TIDAL necessita de ser modificado, ocupa menos espaço criar uma TIDAW e uma T-TIC do que o armazenamento atribuído para o TCCB TIDAL do chamador mais uma TIDAW extra.

Como descrito anteriormente, as formas de realização podem ser realizadas na forma de processos implementados em computador e aparelhos para pôr em prática esses processos. Em formas de realização exemplificativa, a invenção é realizada num código de programa em computador executado por um ou mais elementos de rede. As formas de realização incluem um produto de programa de computador 1100 como mostrado na Fig. 11 num meio 1102 utilizável em computador com lógica 1104 de código de programa contendo instruções realizadas em meios tangíveis como um artigo manufacturado. Artigos manufacturados exemplo para um meio utilizável em computador 1202 podem incluir disquetes maleáveis, CD-ROM, discos rígidos, drives flash com bus série universal (USB), ou qualquer outro meio de armazenamento legível por computador, nos quais, quando a lógica 1104 de código de programa em computador é carregada e executada por um computador, o computador se torna um aparelho para praticar a invenção. As formas de realização incluem a lógica 1104 de código de programação em computador, por exemplo, seja armazenado num meio de armazenamento, carregado em e/ou executado por um computador, ou transmitido sobre algum meio de transmissão, tal como sobre linhas ou cabos eléctricos, através de fibras ópticas, ou através de radiação electromagnética, nos quais, quando a lógica 1104 de código de programa em computador é carregada e executada por um computador, o computador se torna um aparelho para praticar a invenção. Quando implementado num microprocessador genérico, os segmentos da lógica 1104 de código

de programa em computador configuram o microprocessador para criar circuitos lógicos específicos.

Lisboa, 14 Janeiro de 2011

REIVINDICAÇÕES

1. Método para proporcionar endereçamento de dados indirecto num subsistema de I/O de um sistema (100) de processamento, compreendendo o método:

receber (1102) uma palavra de controlo para uma operação de I/O no subsistema de I/O, incluindo a palavra de controlo um endereço de dados indirecto para dados associados com a operação de I/O, incluindo o endereço de dados indirecto um local inicial de uma lista de endereços de armazenamento que especificam os dados colectivamente, abrangendo a lista dois ou mais locais de armazenamento não-contíguos;

reunir (1004) os dados em resposta à lista; e

transmitir (1006) os dados reunidos para uma unidade de controlo no sistema de processamento de I/O;

em que a lista inclui ainda um primeiro indicador associado com cada endereço de armazenamento na lista, possuindo o primeiro indicador um primeiro valor para indicar que o endereço de armazenamento associado se refere a uma porção dos dados e possuindo o primeiro indicador um segundo valor para indicar que o endereço de armazenamento associado se refere a um local de armazenamento no qual endereços de armazenamento adicionais na lista estão localizados;

em que, para cada endereço de armazenamento na lista, a reunião inclui:

aceder ao endereço de armazenamento e ao primeiro indicador associado com o endereço de armazenamento;

adicionar conteúdos a partir do endereço de armazenamento aos dados em resposta ao facto de o primeiro indicador possuir o primeiro valor;

caracterizado por:

se aceder a um segundo endereço de armazenamento e a um segundo indicador localizado no endereço de armazenamento em resposta ao facto de o primeiro indicador possuir o segundo valor e adicionar conteúdos a partir do segundo endereço de armazenamento aos dados em resposta ao facto de o segundo indicador possuir o primeiro valor.

2. Método da reivindicação 1, em que um segundo endereço de armazenamento na lista está localizado contiguamente a um primeiro endereço de armazenamento na lista quando o primeiro indicador possui o primeiro valor e o segundo endereço de armazenamento na lista está localizado não-contiguamente ao primeiro endereço de armazenamento quando o primeiro indicador possui o segundo valor.
3. Método da reivindicação 1, em que a lista inclui ainda um campo de contagem associado com cada endereço de armazenamento na lista, indicando o campo de contagem um número de bytes a ser lido do endereço de armazenamento e a reunião depender dos campos de contagem e dos endereços de armazenamento.

4. Método da reivindicação 1, em que os dados incluem dados de controlo.
5. Método da reivindicação 1, em que os dados incluem dados de cliente.
6. Método da reivindicação 1, em que a palavra de controlo é uma palavra de controlo de transporte (TCW), a lista de endereços de armazenamento é uma lista de endereços de dados indirectos de transporte (TIDAL) e cada endereço de armazenamento é uma palavra de endereço de dados indirecto de transporte (TIDAW).
7. Método da reivindicação 1, em que o subsistema de I/O é um subsistema de canal.
8. Método da reivindicação 1, em que a palavra de controlo inclui um indicador para indicar que esta inclui um endereço de dados indirecto.
9. Sistema compreendendo meios adaptados para executar todos os passos do método de acordo com qualquer reivindicação anterior.
10. Programa de computador compreendendo instruções para executar todos os passos do método de acordo com qualquer reivindicação de método anterior, quando o referido programa de computador é executado num sistema informático.

Lisboa, 14 de Janeiro de 2011

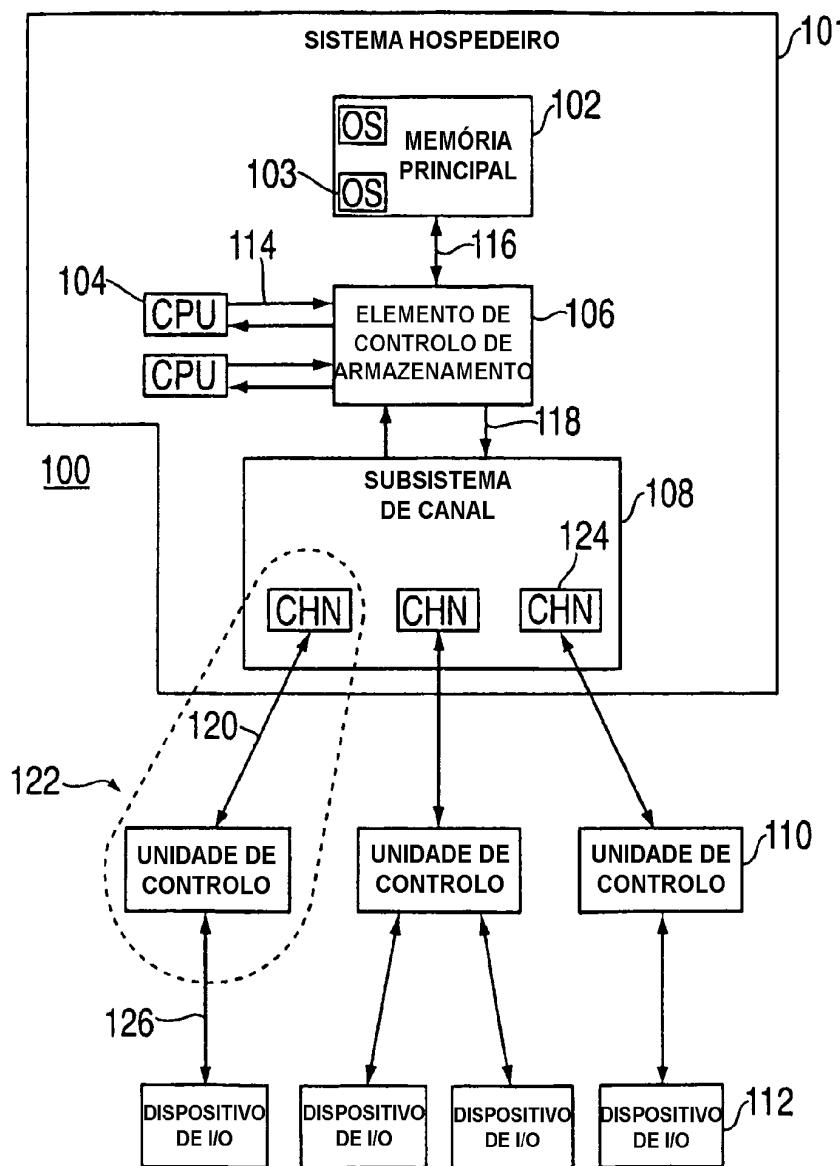


FIG. 1

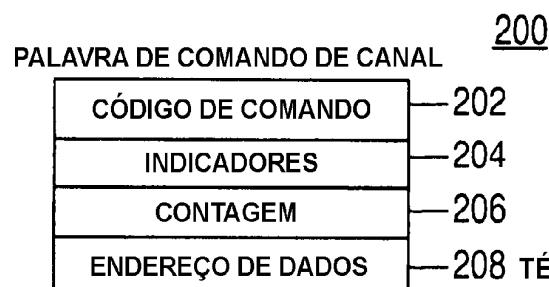


FIG. 2A

TÉCNICA ANTERIOR

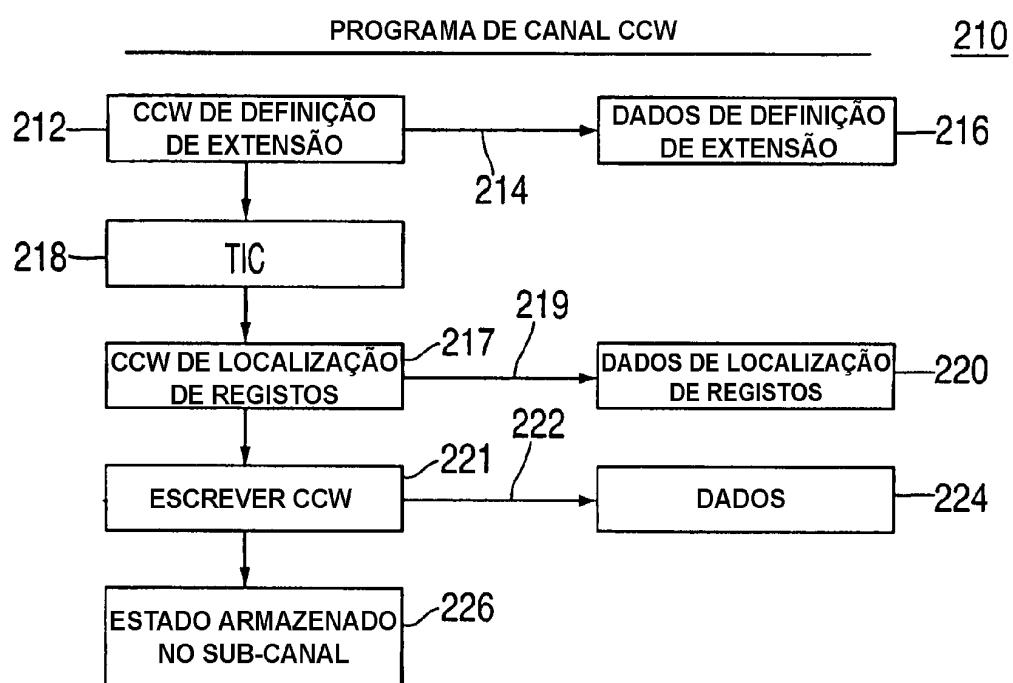


FIG. 2B
TÉCNICA ANTERIOR

**PROTOCOLO DE LIGAÇÃO PARA ESCRITA DE 4K
PROGRAMA DE CANAL CCW**

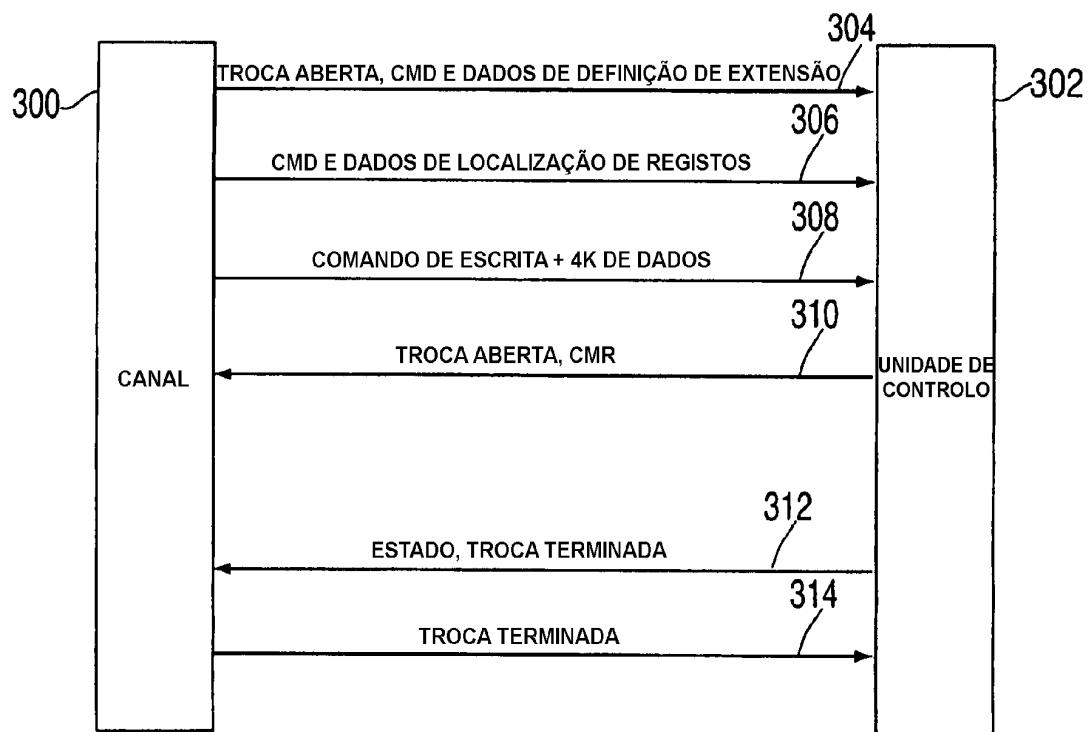
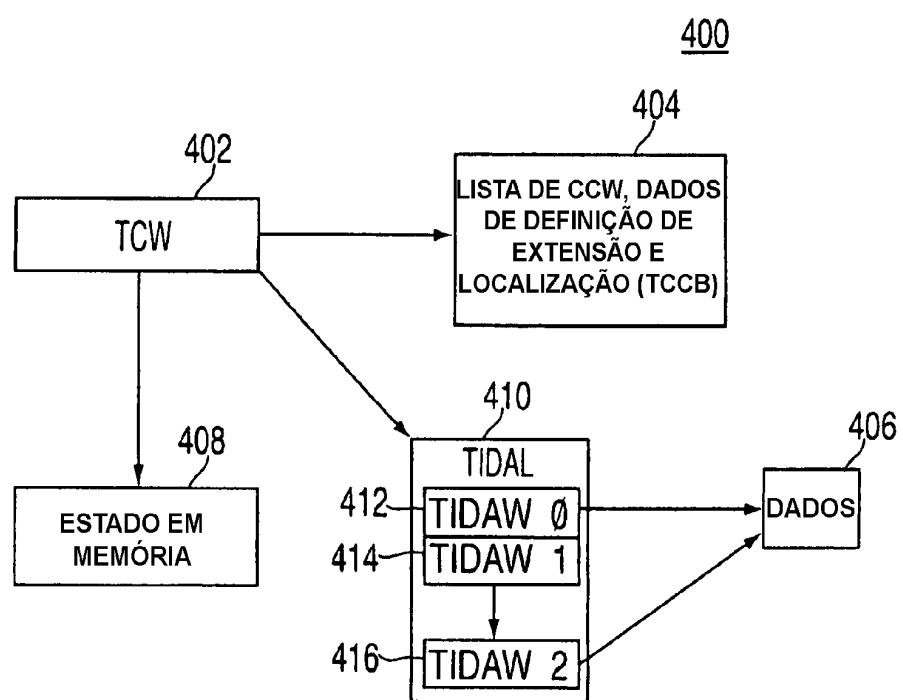


FIG. 3
TÉCNICA ANTERIOR

PROGRAMA DE CANAL TCW

**FIG. 4**

PROTOCOLO DE LIGAÇÃO PARA ESCRITA DE 4K
PROGRAMA DE CANAL TCW

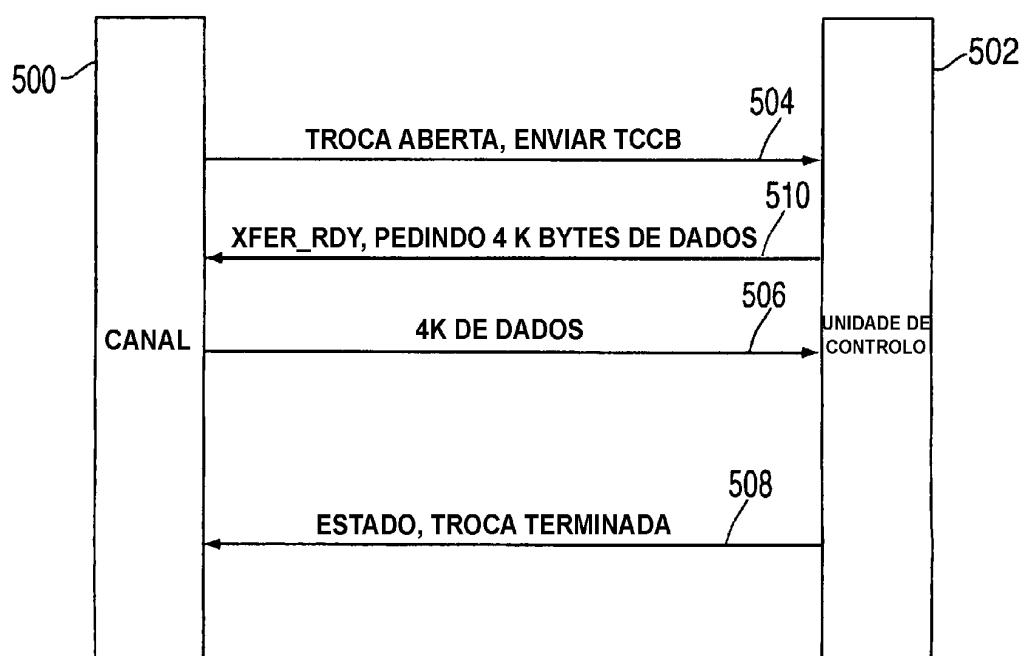


FIG. 5

PROGRAMA DE CANAL CCW

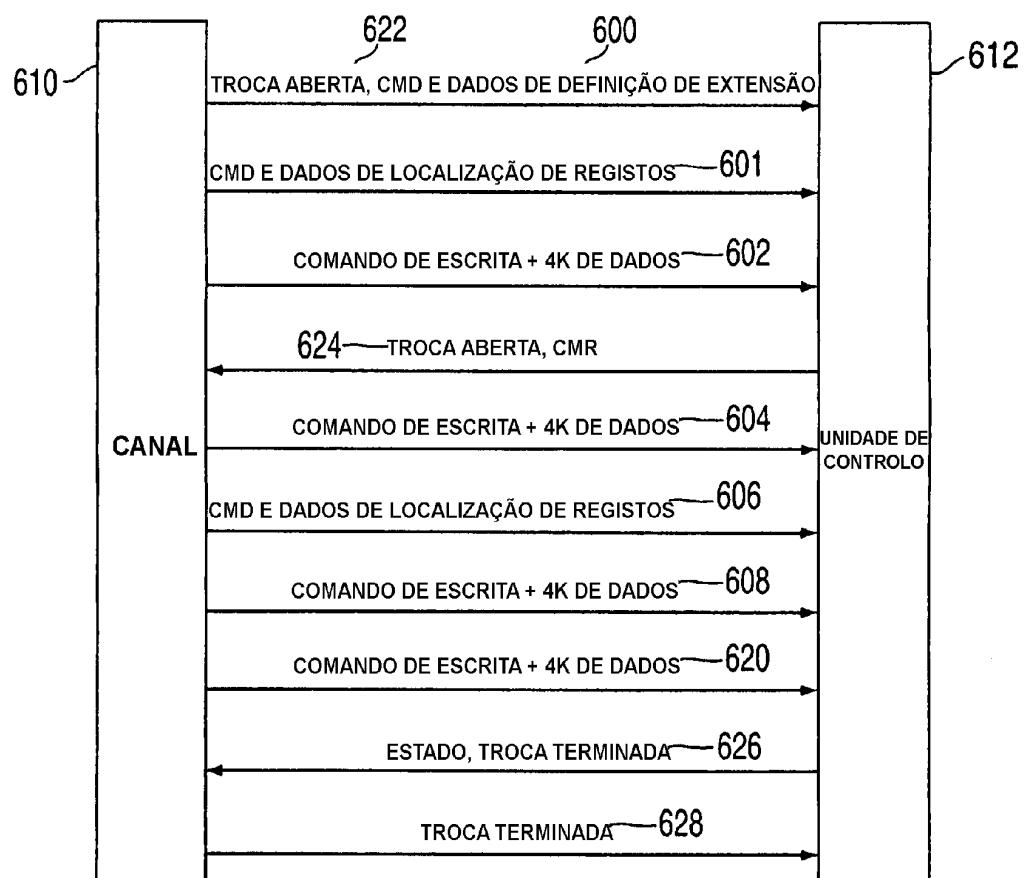


FIG. 6
TÉCNICA ANTERIOR

PROGRAMA DE CANAL TCW

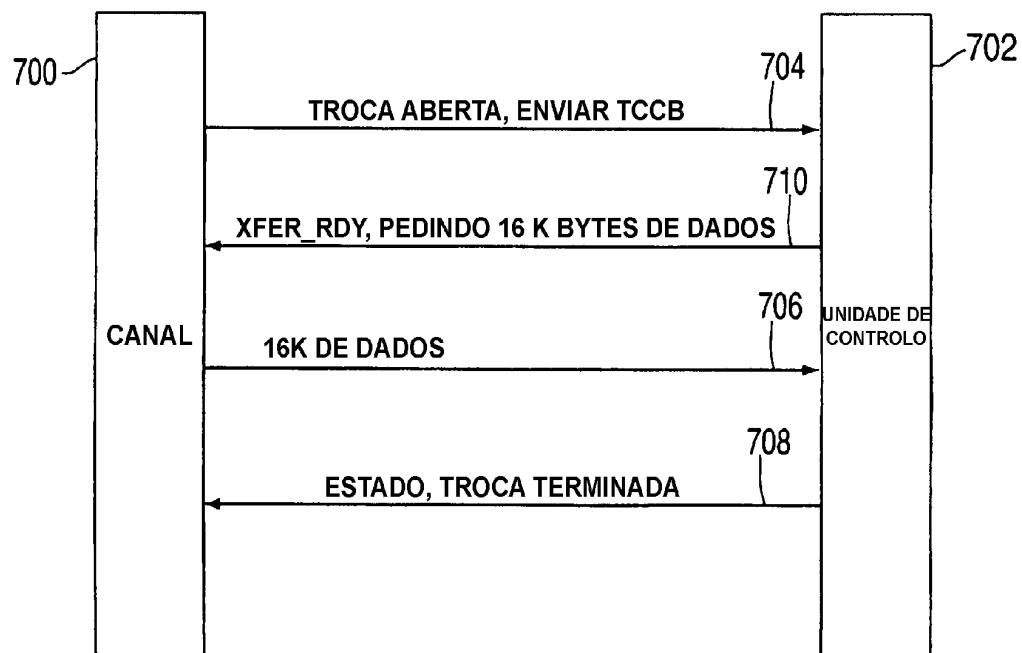


FIG. 7

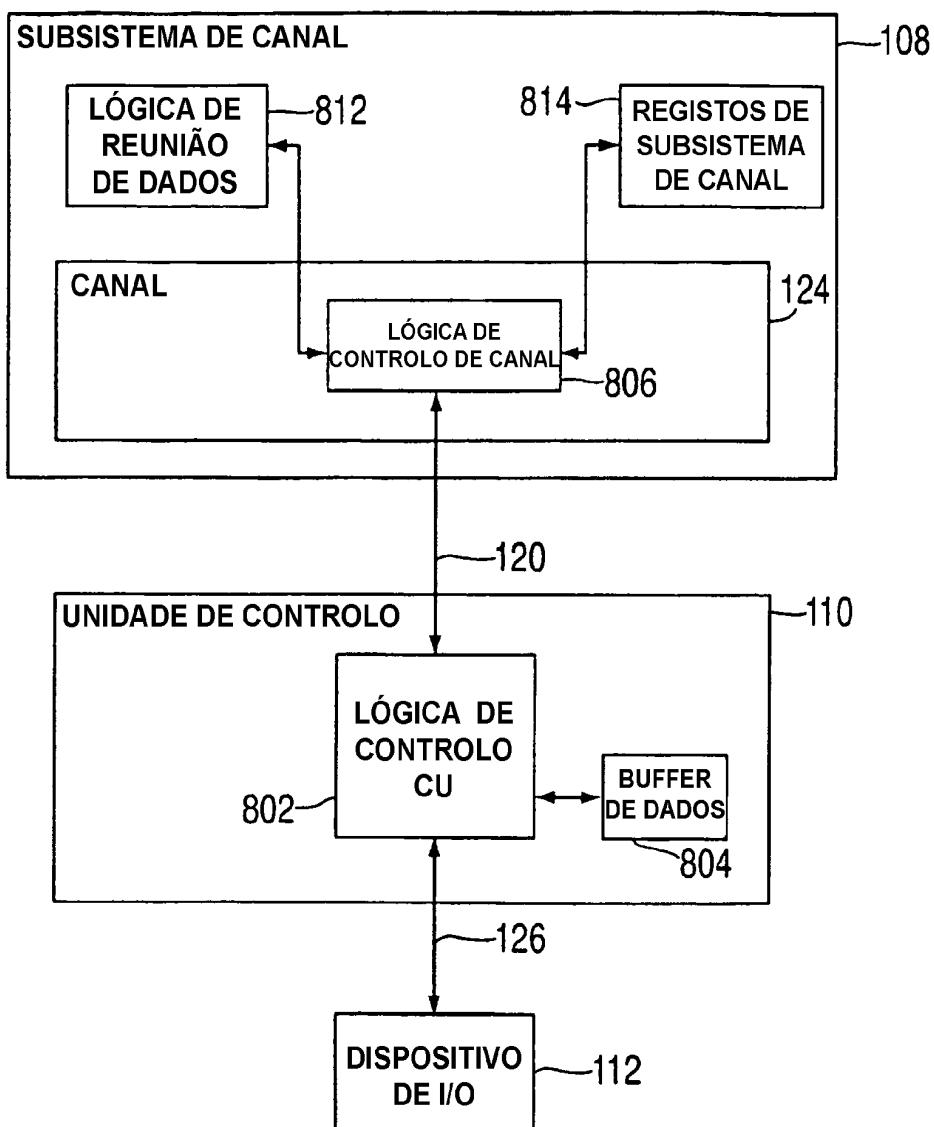


FIG. 8

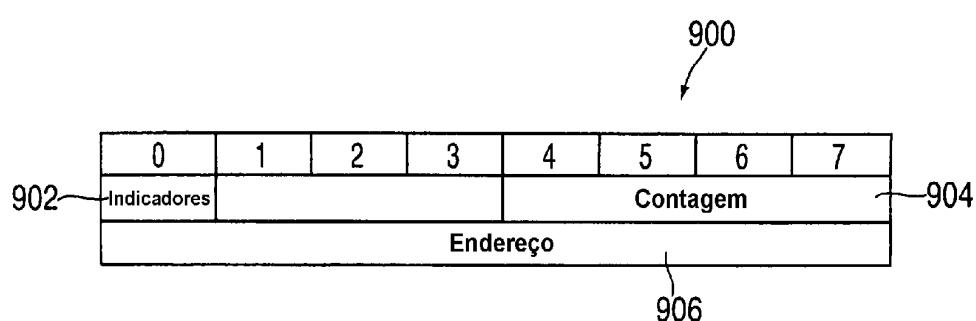


FIG. 9

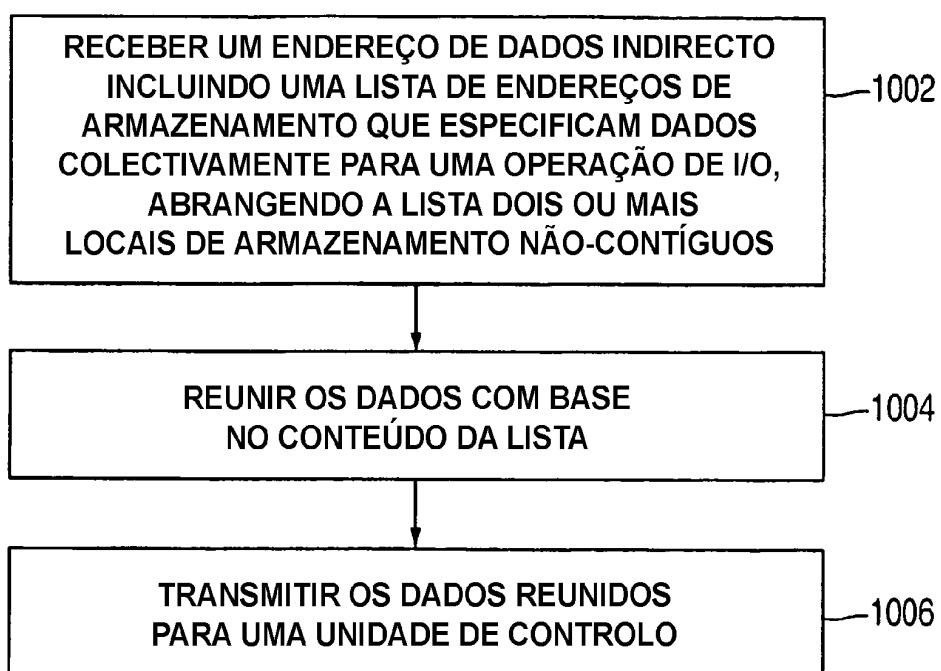
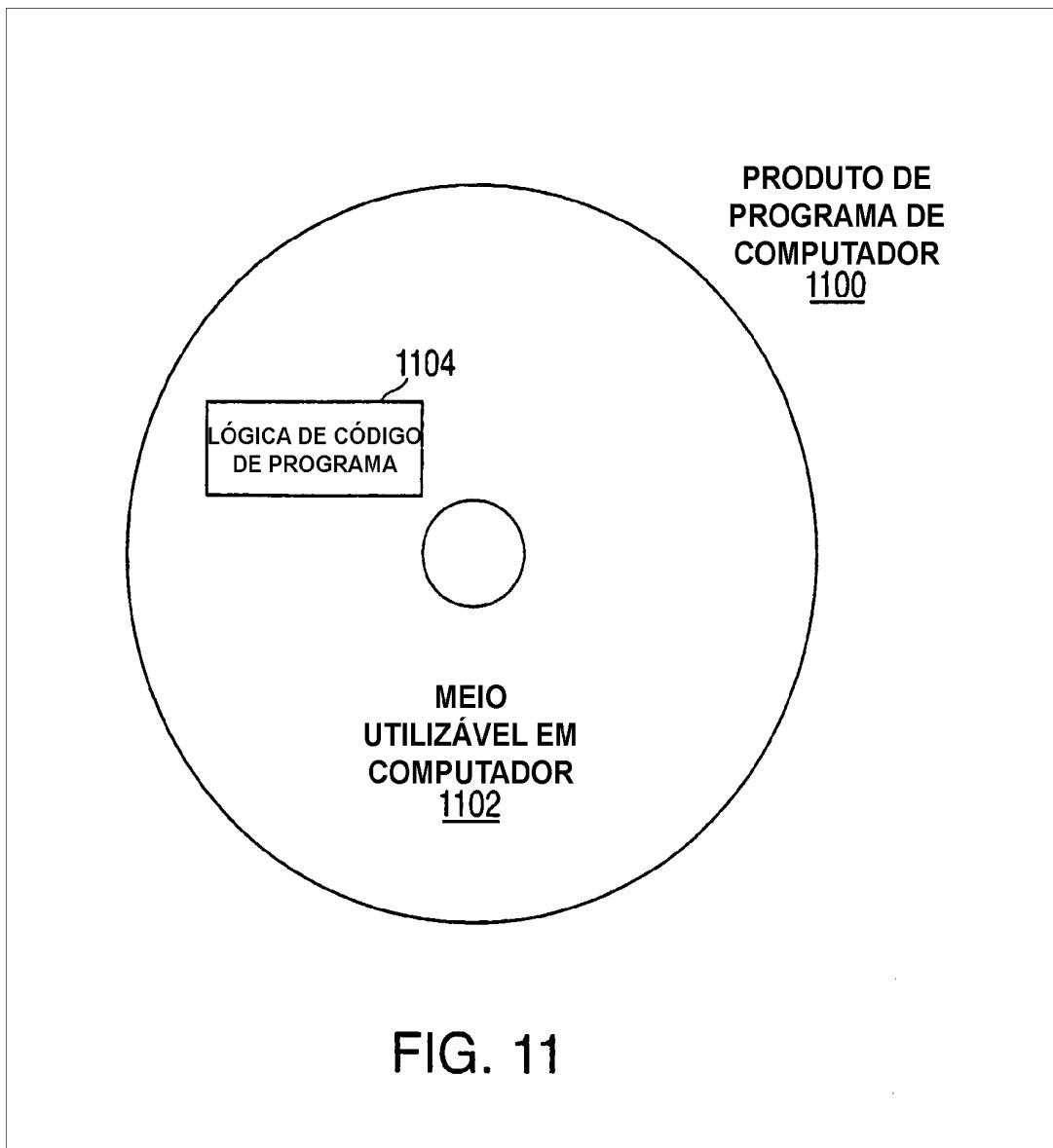


FIG. 10



RESUMO

"PROVISÃO DE ENDEREÇAMENTO INDIRECTO DE DADOS NUM SISTEMA DE PROCESSAMENTO DE ENTRADA/SAÍDA EM QUE A LISTA DE ENDEREÇOS INDIRECTOS DE DADOS É NÃO-CONTÍGUA"

Sistemas, métodos e produtos de programa de computador para proporcionar endereçamento de dados indirecto num subsistema de I/O de um sistema de processamento de I/O. O produto de programa de computador inclui um meio de armazenamento tangível legível por um circuito de processamento e armazenando instruções para execução pelo circuito de processamento para executar um método. O método inclui receber uma palavra de controlo para uma operação de I/O. A palavra de controlo inclui um endereço de dados indirecto para dados associados com a operação de I/O. O endereço de dados indirecto inclui um local inicial de uma lista de endereços de armazenamento que especificam os dados colectivamente, a lista abrange dois ou mais locais de armazenamento não-contíguos. Os dados são reunidos em resposta à lista. Os dados reunidos são transmitidos para uma unidade de controlo no sistema de processamento de I/O.