



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

(11) PI 0912211-7 B1



(22) Data do Depósito: 05/05/2009

(45) Data de Concessão: 07/01/2020

(54) Título: MÉTODO E ARRANJO PARA GERENCIAMENTO DE OPERAÇÕES DE CENTRO DE PROCESSAMENTO DE DADOS

(51) Int.Cl.: G06F 1/20; G06F 1/329; G06F 9/50; H05K 7/20; H04L 29/08; (...).

(52) CPC: G06F 1/206; G06F 1/329; G06F 9/5027; H05K 7/20836; G06F 9/505; (...).

(30) Prioridade Unionista: 04/05/2009 US 12/435,401; 05/05/2008 US 61/050,425; 05/05/2008 US 61/050,429; 05/05/2008 US 61/050,420.

(73) Titular(es): SIEMENS INDUSTRY, INC..

(72) Inventor(es): WILLIAM THOMAS PIENTA; PORNSAK SONGKAKUL.

(86) Pedido PCT: PCT US2009002762 de 05/05/2009

(87) Publicação PCT: WO 2009/137026 de 12/11/2009

(85) Data do Início da Fase Nacional: 05/11/2010

(57) Resumo: ARRANJO PARA GERENCIAMENTO DE OPERAÇÕES DE CENTRO DE PROCESSAMENTO DE DADOS PARA AUMENTAR A EFICIÊNCIA DA REFRIGERAÇÃO A presente invenção refere-se a um método que inclui uma etapa de obter informação de característica de eficiência para cada um de uma pluralidade de unidades condicionadoras de ar em uma localização que contém uma pluralidade de computadores servidores. O método também inclui empregar uma ou mais unidades de processamento para alocar uma ou mais tarefas de processamento a um de uma pluralidade de computadores servidores baseado na informação de característica de eficiência.

Relatório Descritivo da Patente de Invenção para "**MÉTODO E ARRANJO PARA GERENCIAMENTO DE OPERAÇÕES DE CENTRO DE PROCESSAMENTO DE DADOS**".

CAMPO DA INVENÇÃO

[001] A presente invenção refere-se, em geral, a centros de processamento de dados, e mais particularmente, ao gerenciamento das operações de um centro de processamento de dados.

ANTECEDENTES DA INVENÇÃO

[002] Centros de processamento de dados são partes de prédios ou instalações nos quais uma grande quantidade de computadores servidores é instalada. O acúmulo denso de computadores servidores resulta na geração de uma grande quantidade de calor em uma área localizada. O centro de processamento de dados tem que ser refrigerado de uma maneira confiável a fim de evitar o desligamento do, ou danos ao hardware dos computadores servidores. O desligamento de computadores servidores devido à sobrecarga de calor pode causar perda econômica significativa.

[003] Conseqüentemente, unidades de refrigeração especializadas têm sido desenvolvidas para implementação diretamente em centros de processamento de dados. Estas unidades de refrigeração especializadas são algumas vezes conhecidas na técnica como unidades de condicionamento de ar de sala de computador ("CRACs") ou unidades de manuseio de ar de sala de computador. Nesta revelação, será entendido que a unidade de condicionamento de ar ou CRAC engloba qualquer dispositivo usado para realizar refrigeração em um centro de processamento de dados. Os CRACs têm sido empregados como um resultado do fato de que os sistemas HVAC ordinários de prédios não são configurados otimamente para manusear o calor concentrado gerado com os centros de processamento de dados. Deste modo, Os CRACs são frequentemente usados em conjunto com, mas adicional-

mente, às unidades de refrigeração comuns de um prédio empregadas para sistemas de conforto humano.

[004] Muitos CRACs têm controles simples embutidos que ajustam saída da unidade com bases em fatores tais como temperatura do ar ambiente. Em alguns casos, os CRACs têm controladores que interagem com o sistema de automação do prédio que controla ou inclui o sistema HVAC do prédio, dentre outras coisas.

[005] Ao mesmo tempo em que os CRACs fornecem uma solução para a necessidade por potência de refrigeração melhorada dentro de um centro de processamento de dados que têm diversos computadores servidores, a energia consumida pelos CRACs é significativa. Portanto, é desejável reduzir o consumo de energia associado com a refrigeração adequada dos centros de processamento de dados.

SUMÁRIO

[006] A presente invenção endereça as necessidades identificadas acima, bem como outras, fornecendo um método e arranjo que aloca tarefas de processamento para servidores com base na localização dos servidores com respeito às unidades de condicionamento de ar. Além disso, modalidades da invenção alocam tarefas de processamento para servidores com bases nas características de eficiência das unidades de condicionamento de ar dentro do centro de processamento de dados. Como resultado, o calor gerado pela execução de tarefas de processamento é distribuído vantajosamente para as unidades de condicionamento de ar com base em suas eficiências.

[007] Uma primeira modalidade é um método que inclui uma etapa de obter informação de característica de eficiência para cada um de uma pluralidade de unidades de condicionamento em uma localização que contém uma pluralidade de computadores servidores. O método também inclui empregar uma ou mais unidades de processamento para alocar uma ou mais tarefas de processamento a um da pluralidade

de computadores servidores com base na informação de característica de eficiência.

[008] Em algumas modalidades, uma primeira requisição de processamento é alocada a um primeiro de uma pluralidade de processadores com base em uma proximidade térmica do primeiro da pluralidade de computadores a uma unidade selecionada da pluralidade de unidades de condicionamento de ar. Adicionalmente, a unidade selecionada da pluralidade de unidades de condicionamento de ar é selecionada com base na informação de característica de eficiência.

[009] Uma segunda modalidade é um arranjo que inclui um sistema de gerenciamento de computador servidor que tem uma memória e um circuito de processamento. O circuito de processamento é configurado para coordenar o uso de uma pluralidade de computadores servidores. A memória armazena informação de característica de eficiência para uma pluralidade de unidades de condicionamento de ar em uma localização que contém uma pluralidade de computadores servidores. O circuito de processamento é configurado adicionalmente para alocar uma ou mais tarefas de processamento a um de uma pluralidade de computadores servidores com base na informação de característica de eficiência.

[0010] As vantagens e características descritas acima ficarão mais facilmente evidentes para os indivíduos com conhecimentos comuns na técnica através de referência a descrição detalhada a seguir em conjunto com os desenhos em anexo.

BREVE DESCRIÇÃO DAS FIGURAS

[0011] A figura 1 mostra um diagrama de bloco esquemático de um arranjo ilustrativo de acordo com uma primeira modalidade da invenção implementada para coordenar o processamento de aplicações em um centro de processamento de dados ilustrativo;

[0012] A figura 2 mostra um conjunto de curvas de característica

de eficiência ilustrativas para unidades de condicionamento de ar no centro de processamento de dados mostrado na figura 1;

[0013] A figura 3 mostra um conjunto ilustrativo de operações que podem ser executadas de acordo com a presente invenção;

[0014] A figura 4 mostra em detalhes adicionais uma modalidade ilustrativa de pelo menos uma das operações da figura 3; e

[0015] A figura 5 mostra em detalhes adicionais uma segunda modalidade de pelo menos uma das operações da figura 3.

DESCRIÇÃO DETALHADA

[0016] A figura 1 mostra um arranjo 100 de acordo com uma modalidade ilustrativa da invenção. O arranjo 100 é mostrado usado em conjunto com um centro de processamento de dados 102 que inclui uma pluralidade de computadores servidores 104₁, 104₂... 104₁₈ e uma pluralidade de unidades de condicionamento de ar 106₁, 106₂, 106₃, e 106₄. O arranjo 100 inclui um sistema de gerenciamento de computador servidor 108 que tem, dentre outras coisas, uma memória 110 e um circuito de processamento 112. Nesta modalidade, o arranjo 100 inclui adicionalmente um elemento BAS 120, que é conectado comunicativamente ao circuito de processamento 112.

[0017] Cada um dos computadores servidores 104₁ a 104₁₈ é parte de um conjunto de computadores que fornece serviços de processamento de aplicação para pelo menos um, e tipicamente uma grande quantidade de computadores clientes não mostrados. Os computadores servidores 104₁ a 104₁₈ são tipicamente dispostos em bastidores e espalhados por todo o espaço do centro de processamento de dados 102. Por exemplo, como mostrado na figura 1, os computadores servidores 104₁, 104₂, 104₃, 104₄ e 104₅, podem ser agrupados em um primeiro bastidor, os computadores servidores 104₆, 104₇, 104₈, e 104₉, podem ser agrupados em um segundo bastidor, os computadores servidores 104₁₀, 104₁₁, 104₁₂, 104₁₃ e 104₁₄, podem ser agrupados em

um terceiro bastidor, e os computadores servidores 104₁₅, 104₁₆, 104₁₇, e 104₁₈, podem ser agrupados em um quarto bastidor.

[0018] Cada uma das unidades de condicionamento de ar 106₁, 106₂, 106₃ e 106₄, é uma unidade de condicionador de ar de sala de computador ou de manuseio de ar de sala de computador, referenciados coletivamente como CRACs, ou qualquer unidade de condicionamento de ar que possa ser adequadamente empregada para refrigerar especificamente um centro de processamento de dados ou outra área que seja uma grande geradora de calor, e exija refrigeração significativa localizada. Estes dispositivos são bem conhecidos na técnica. Nesta modalidade, cada uma das unidades de condicionamento de ar (ACUs) 106₁, 106₂, 106₃ e 106₄ é operativamente acoplada ao elemento BAS 120 de modo que o elemento BAS 120 pode realizar pelo menos alguma medida de controle sobre as operações das ACUs 106_n. Por exemplo, se uma ACU 106_n tem um sensor e controle de temperatura independente, o elemento BAS 120 é conectado operacionalmente para se sobrepôr ao controle local de ligar/desligar, e ou para fornecer um ponto de regulação para a ACU 106_n. Outras ACUs podem ser configuradas para controle mais extensivo através de um controlador de campo externo do sistema de automação do prédio. Em qualquer dos casos, o elemento BAS 120 preferencialmente é operável conectado para fornecer gerenciamento geral e/ou controle de cada uma das ACUs 106₁, 106₂, 106₃, e 106₄.

[0019] O elemento BAS 120 é um ou mais dispositivos que são configurados para comunicar com, e operar com, um sistema de automação de prédio tal como um sistema HVAC ou semelhante. Estes sistemas são conhecidos na técnica e podem ter uma arquitetura geral do sistema APOGEE® disponibilizado pela Siemens Building Technologies Inc. O elemento BAS 120 inclui pelo menos um circuito de processamento 140 e uma memória 142. O elemento BAS 120 pode ade-

quadamente tomar a forma de uma estação de trabalho supervisora em um BAS tal como a estação INSIGHT® disponibilizada pela Siemens Building Technologies Inc., de Buffalo Grove, Illinois. Na alternativa, o elemento BAS 120 pode adequadamente ser um controlador de campo configurável, tal como o controlador de campo PXC Modular, também disponibilizado pela Siemens Building Technologies Inc. Em geral, o circuito de processamento 140 é configurado através de outros circuitos para comunicar dados BAS (tal como pontos de ajuste, valores de sensor e comandos) com outros dispositivos BAS tais como outros controladores, ou mesmo com sensores e atuadores. O elemento BAS 120 pode adicionalmente incluir dispositivos de I/O especiais digitais ou analógicos visto que pode ser necessário se comunicar com elementos de controle das ACUs 106₁, 106₂, 106₃, e 106₄. Nesta modalidade, o elemento BAS 120 é adicionalmente operável conectado para comunicar informação com o sistema de gerenciamento de computador servidor 108, e particularmente o circuito de processamento 112. Para este fim, é fornecida uma interface de dados adequada entre o elemento BAS 120, que é configurado para um sistema BAS, e o sistema de gerenciamento de computador servidor 108, o qual tipicamente não é estabelecido para comunicação com um sistema BAS.

[0020] O sistema de gerenciamento de computador servidor 108 é um sistema de computação que é geralmente configurado para coordenar o uso da pluralidade de computadores servidores 104₁, 104₂. Estes dispositivos são geralmente conhecidos. Para coordenar a utilização do servidor, o circuito de processamento 112 do sistema de gerenciamento de computador servidor 108 executa software de virtualização 114. Software de virtualização 114, como é conhecido na técnica, é software que, quando executado por um processador de computador, gerencia a alocação de processos de aplicação entre uma pluralidade de computadores servidores, tal como em um centro de proces-

samento de dados.

[0021] De acordo com esta modalidade da presente invenção, o circuito de processamento 112 é configurado adicionalmente para alocar processos de aplicação entre os computadores servidores 104₁, 104₂, etc. com base nas características de eficiência das ACUs 106₁, 106₂, 106₃ e 106₄. Como será discutido abaixo em detalhes, o circuito de processamento 112 aloca processos de aplicação entre os computadores servidores 104₁ a 104₁₈ de modo que o calor resultante gerado pelo processamento das aplicações seja distribuído entre as ACUs 106₁, 106₂, 106₃ e 106₄ de uma maneira que leva em conta as eficiências relativas das unidades para diferentes níveis de carga térmica.

[0022] Para este fim, a memória 110 armazena informação de característica de eficiência para a pluralidade de ACUs 106₁, 106₂, 106₃ e 106₄ no centro de processamento de dados 102. A informação de eficiência característica pode compreender adequadamente curvas de performance de carga parcial das ACUs 106₁, 106₂, 106₃ e 106₄. Por exemplo, a figura 2 mostra gráficos ilustrativos 202, 204, 206 e 208, em que cada gráfico mostra curvas de eficiência 212, 214, 216 e 218 para várias cargas. O eixo-x representa a saída de refrigeração, e o eixo-y representa o consumo de energia.

[0023] As curvas de eficiência 212, 214, 216 e 218 podem ser geradas em uma pluralidade de modos, em que alguns dependem do tipo de unidade de condicionamento de ar. Em geral, as curvas podem ser geradas pela identificação do consumo de energia para uma pluralidade de valores de saída de refrigeração. As curvas 212, 214, 216 e 218 podem ser geradas a partir dos pontos de dados resultantes através de interpolação ou técnicas de ajuste de curva.

[0024] Para obter pontos de dados, a saída de refrigeração é geralmente medida como a mudança na temperatura do ar (ΔT) entre a entrada e saída da ACU, multiplicada pelo fluxo de ar através da uni-

dade. O consumo de energia associado pode ser medido de várias formas dependendo do tipo de unidade. Se a unidade é inteiramente independente (por exemplo, contém um compressor e refrigerante), o consumo de energia elétrica da ACU fornece uma boa medida do consumo de energia. Se, entretanto, o dispositivo depende de refrigerante fornecido e/ou bombeado a partir de outra fonte, tal como água gelada, então o consumo de energia pode ser medido como o fluxo de refrigerante multiplicado pela troca na temperatura do refrigerante entre a entrada e saída da ACU. O consumo de energia elétrica do ventilador da unidade também tem que ser adicionado. Obviamente, esta medida de consumo de energia pode ser feita de forma mais precisa ao levar em conta a perda de transmissão do refrigerante da e para a ACU, bem como outras perdas.

[0025] Em todo caso, depois de serem obtidos vários pontos de dados de saída de refrigeração contra consumo de energia para cada unidade de condicionamento de ar 106₁, 106₂, 106₃, 106₄, as curvas de eficiência 212, 214, 216, e 218 são geradas com o uso de interpolação em linha reta ou pode ser usado ajuste de curvas. Em todo caso, dados representativos das curvas 212, 214, 216, ou 218 são armazenados na memória 110.

[0026] Com referência novamente a figura 1, como discutido acima, o sistema de gerenciamento de computador servidor 108 é configurado para alocar (através do software de virtualização) uma ou mais tarefas de processamento a um da pluralidade de computadores servidores 104₁ - 104₁₈ com base em parte na informação de característica de eficiência das unidades de condicionamento de ar 106₁, 106₂, 106₃, 106₄. Em uma primeira operação ilustrativa, as aplicações são direcionadas para computadores servidores 104_n, 104_o, etc. com base na proximidade térmica dos servidores 104_n, 104_o, etc. para as unidades de condicionamento de ar 106₁ a 106₄ e as características de eficiên-

cia armazenadas das unidades de condicionamento de ar 106₁ a 106₄.

[0027] Será entendido que a expressão “proximidade térmica” como usada neste documento significa uma extensão a qual dois itens são acoplados termicamente, por exemplo, quanto calor é transferível entre eles. Por exemplo, se um computador servidor é adjacente à entrada da uma ACU de modo que a unidade recebe e remove todo o calor gerado pelo computador servidor, então eles estão termicamente acoplados proximamente, e assim tem um alto grau de “proximidade térmica”. Enquanto a proximidade térmica é frequentemente um resultado da proximidade física, é possível para um computador servidor ser fisicamente próximo, mas não ser muito fortemente acoplado termicamente a uma ACU particular. A proximidade térmica relativa entre dispositivos de refrigeração tais como ACUs e equipamento gerador de calor em um ambiente tal como um centro de processamento de dados pode prontamente ser determinada ou pelo menos estimada por versados na técnica.

[0028] Em todo caso, quando uma aplicação é atribuída a um computador servidor 104_n a execução da aplicação faz com que o computador servidor 104_n gere energia térmica. O circuito de processamento 112 aloca as tarefas de processamento de modo que o calor gerado pelos computadores servidores 104 quando executam as tarefas de processamento é distribuído eficientemente entre as várias ACUs 106₁, 106₂, 106₃ e 106₄.

[0029] A título de exemplo, considere uma situação na qual 100 aplicações têm que ser alocadas aos computadores servidores 104₁ a 104₁₈. Na técnica anterior, uma forma de alocar as aplicações pode ser simplesmente alocar uma quantidade substancialmente igual de aplicações para cada um dos processadores, de modo que, neste exemplo, cada um dos computadores servidores 104₁ a 104₁₈ também pode ser levado em conta. Em todo caso, o método da técnica anterior pode

incluir distribuição mais ou menos igual da carga de processamento entre os computadores servidores.

[0030] Entretanto, pode ser o caso de que esta alocação não resulte no uso eficiente das ACUS 106₁, 106₂, 106₃, 106₄. Para este fim, assume-se que a ACU 106₁ geralmente é responsável por refrigerar (isto é, é termicamente próxima a) os computadores servidores 104₁ a 104₅, a ACU 106₂ geralmente é responsável por refrigerar os computadores servidores 104₆ a 104₉, a ACU 106₃ geralmente é responsável por refrigerar os computadores servidores 104₁₀ a 104₁₄, a ACU 106₄ geralmente é responsável por refrigerar os computadores servidores 104₁₅ a 104₁₈. Pode haver o caso em que a carga térmica provoque uma alocação mais ou menos igual das aplicações entre os processadores 104₁ a 104₁₈ e não resulte nas ACUs 106₁, 106₂, 106₃, 106₄ operando em seus níveis mais eficientes possíveis. De fato, mesmo assumindo que o calor gerado pelos computadores servidores 104₁ a 104₁₈ é distribuído igualmente entre as ACUs 106₁, 106₂, 106₃, 106₄, as ACUs 106₁, 106₂, 106₃, 106₄ como um grupo podem não estar operando a seus níveis de eficiência mais altos. Em particular, cada ACU 106_n pode ter um nível diferente no qual a mesma opera de forma mais eficiente, e pode ser mais ou menos eficiente que as outras unidades em vários níveis de saída. Isto é evidenciado pela variabilidade das curvas 212, 214, 216 e 218. Deste modo, de acordo com a presente invenção, o circuito de processamento 112 atribui as aplicações aos computadores servidores 104₁ a 104₁₈ de modo que o calor é alocado de acordo com os níveis de operação de eficiência combinada das ACUs 106₁, 106₂, 106₃, 106₄.

[0031] Em uma ilustração simples, assume-se que em certo nível de carga térmica *médio*, a primeira ACU 106₁ pode estar operando a um nível muito mais eficiente do que uma segunda ACU 106₂. Neste caso, deve ser vantajoso colocar mais carga de processamento nos

computadores servidores 104₁ a 104₅ porque os mesmos estão termicamente mais próximos à primeira ACU mais eficiente 106₁, e colocar menos carga de processamento nos computadores servidores 104₆ a 104₉ porque os mesmos estão termicamente próximos a ACU menos eficiente 106₂. Entretanto, ao mesmo tempo, colocar muita carga de processamento nos processadores termicamente próximos à primeira ACU 106₁ e muito pouca nos computadores servidores termicamente próximos a segunda ACU 106₂ também pode resultar em eficiência reduzida. Assim o circuito de processamento 112 informação de eficiência tal como as curvas 212, 214, 216 e 218 para distribuir a carga térmica para as ACUs 106₁, 106₂, 106₃, 106₄ de uma maneira eficiente. A carga térmica é distribuída para uma ACU particular 106_m ao atribuir as tarefas de processamento aos computadores servidores 104_n que são termicamente próximos a (ou seja, termicamente acoplados fortemente a) aquela ACU 106_m.

[0032] O circuito de processamento 112 é, por consequência, configurado para determinar a alocação de pelo menos alguns processos com base nas características de eficiência das quatro ACUs 106₁, 106₂, 106₃, 106₄. Em geral, esta determinação é adicionalmente dependente da proximidade térmica dos computadores servidores 104₁ a 104₁₈ às ACUs 106₁, 106₂, 106₃, 106₄.

[0033] A figura 3 mostra uma modalidade ilustrativa das operações gerais do circuito de processamento 112 da figura 1. Na etapa 305, o circuito de processamento 112 obtém características de eficiência para as ACUs 106₁, 106₂, 106₃, 106₄. Como discutido acima, as características de eficiência podem adequadamente compreender as curvas de carga parcial 212, 214, 216 e 218 de saída de refrigeração como uma função do consumo de energia, como ilustrado nos gráficos 202, 204, 206 e 208 da figura 2. Tipicamente, o circuito de processamento 112 obtém esta informação através de comunicações, ou pela entrada do

usuário, como resultado de um procedimento de estabelecimento para o centro de processamento de dados 102. O circuito de processamento 112 armazena a informação de eficiência na memória 110. Em algumas modalidades, as características de eficiência das ACUs 106₁, 106₂, 106₃, 106₄ podem ser armazenadas em uma memória do elemento BAS 120, ou outro computador.

[0034] Na etapa 310, o circuito de processamento 112 determina como alocar a carga de processamento com base nas características das ACUs 106₁, 106₂, 106₃, 106₄. Para este fim, o circuito de processamento 112 também tem informação que correlaciona os computadores servidores 104₁ a 104₁₈ às ACUs 106₁, 106₂, 106₃, 106₄. Por exemplo, cada computador servidor 104_n é identificado como tendo uma ACU 106_m mais próxima termicamente. Termicamente mais próxima, significa que a carga térmica gerada pelo computador servidor 104_n é suportada primariamente por aquela ACU 106_m, ou em outras palavras, o computador servidor 104_n é mais próximo termicamente à ACU 106_m.

[0035] Em uma modalidade mais sofisticada, cada computador servidor 104_n pode ter uma *função* de impacto térmico que consiste de um polinômio ponderado, em que cada termo representa a capacidade produtora de carga térmica de uma das ACUs 106₁, 106₂, 106₃, 106₄. Esta modalidade reconhece que a carga térmica gerada por um computador servidor 104_n pode em alguns casos ser manuseada, pelo menos em parte, por mais do que uma das ACUs 106₁, 106₂, 106₃, 106₄. Esta função de impacto térmico para um computador servidor 104_n é expressa como:

$$H_{total,n} = I_{1061,n}H_{total,n} + I_{1062,n}H_{total,n} + I_{1063,n}H_{total,n} + I_{1064,n}H_{total,n}$$

[0036] onde $I_{1061,n}$, $I_{1062,n}$, $I_{1063,n}$, $I_{1064,n}$, são fatores de ponderação que totalizam 1, e $H_{total,n}$ é o calor total gerado pelo computador servidor 104_n. Os fatores de ponderação representam a proporção do calor

gerado pelo computador servidor 104_n que é carregada para cada uma das ACUs 106_1 , 106_2 , 106_3 , 106_4 . Por exemplo, o computador servidor 104_{15} pode ter fatores de ponderação $I_{106_1,15} = 0,02$, $I_{106_2,15} = 0,12$, $I_{106_3,15} = 0,01$, e $I_{106_4,15} = 0,85$, enquanto o computador servidor 104_{12} pode ter fatores de ponderação $I_{106_1,12} = 0,02$, $I_{106_2,12} = 0,00$, $I_{106_3,12} = 0,9$, e $I_{106_4,12} = 0,08$.

[0037] O circuito de processamento 112 então usa a informação de correlação computador servidor/ACU para distribuir as tarefas de processamento de modo que a carga térmica provoque a combinação das ACUs 106_1 , 106_2 , 106_3 , 106_4 a operar em níveis de carga de melhor (ou muito boa) eficiência coletivamente.

[0038] Em um exemplo muito simples da etapa 310, o circuito de processamento 112 pode executar um algoritmo que inicia com um a primeira alocação proposta de tarefas, estima o calor gerado deste modo, e então identifica a carga nas unidades de condicionamento 106_1 , 106_2 , 106_3 , e 106_4 com o uso da informação de correlação de proximidade térmica servidor/ACU. Uma vez que a carga nas ACUs 106_1 , 106_2 , 106_3 , 106_4 é estimada, o circuito de processamento 112 usa a informação de eficiência armazenada (por exemplo, 212, 214, 216 e 218) para calcular a eficiência global (ou valor de consumo de energia) para a primeira alocação proposta de tarefas. Depois disso, o circuito de processamento 112 gera uma *segunda* alocação proposta de tarefas que é ligeiramente (ou significativamente) diferente da primeira alocação proposta de tarefas. O circuito de processamento 112 usa o mesmo procedimento para determinar a eficiência global da segunda alocação proposta de tarefas. O circuito de processamento 112 repete este processo para alocações propostas de tarefas adicionais. O circuito de processamento 112 também pode usar qualquer quantidade de algoritmos convergentes para resolver uma melhor (ou muito boa) alocação proposta de tarefas com base nas eficiências globais das vá-

rias alocações propostas de tarefas.

[0039] Entretanto, será avaliado que a alocação de tarefas para obter carregamento térmico eficiente das ACUs 106₁, 106₂, 106₃, 106₄ pode ser executada de outras formas. Algumas outras formas são discutidas abaixo em conexão com as figuras 4 e 5.

[0040] Com referência novamente a figura 3, uma vez que o circuito de processamento 112 determina a alocação das tarefas de computação na etapa 310, o circuito de processamento 112 na etapa 315 provoca a alocação de tarefas de processamento de acordo com a determinação. As operações da etapa 315 podem ser executadas com o uso de capacidades conhecidas de sistemas de gerenciamento de computador servidor. Depois da etapa 315, os computadores servidores 104₁ a 104₁₈ podem executar adequadamente a tarefa de processamento que foi atribuída a eles.

[0041] Em uma etapa opcional 320, o circuito de processamento 112 adicionalmente comunica informação de carga para o controle BAS 120. A informação de carga identifica tanto a carga de processamento global nos servidores 104₁ a 104₁₈, como qualquer mudança na carga dos servidores 104₁ a 104₁₈, ou qualquer estimativa de carga térmica nas ACUs 106₁, 106₂, 106₃, 106₄. O controle BAS 120 em alguns casos pode usar esta informação para tomar decisões proativas sobre a refrigeração térmica, ou para controlar aspectos das ACUs 106₁, 106₂, 106₃, 106₄ para acomodar qualquer aumento ou redução necessários na saída de refrigeração.

[0042] A figura 4 mostra um exemplo de determinação de alocação de tarefa que pode ser usado como etapa 310 na figura 3. As operações da figura 4 descrevem uma determinação de alocação de tarefa que ocorre de forma incremental durante operação em andamento do centro de processamento de dados. Mais especificamente, as operações da figura 4 podem ser usadas quando nova tarefa ou um novo

conjunto de tarefas de processamento é requisitado aos computadores servidores 104₁ a 104₁₈. Enquanto as operações da figura 4 são descritas com respeito a uma única requisição de aplicação ou processamento adicional, será avaliado que as etapas da figura 4 podem ser usadas adequadamente com um grupo de requisições adicionais de processamento.

[0043] Em alguns casos, o circuito de processamento 112 pode determinar proativamente onde as próximas n aplicações devem ser atribuídas antes de as requisições serem realmente recebidas.

[0044] Na etapa 405, o circuito de processamento 112 do sistema de gerenciamento de computador servidor 108 recebe uma requisição de processamento adicional. Para este fim, de vez em quando, os computadores clientes que tem acesso ao cento de processamento de dados 102 fornecem requisições para tarefas de aplicações serem executadas por um ou mais computadores servidores 104₁ a 104₁₈. Assumindo que a requisição de processamento adicional é de natureza não trivial, o circuito de processamento 112 avança para a etapa 410 para determinar uma alocação para a tarefa.

[0045] Na etapa 410, o circuito de processamento 112 determina qual das ACUs 106_m deve manusear carga térmica adicional com melhor eficiência. Esta determinação pode ser executada através da identificação da carga térmica corrente em cada uma das ACUs 106₁, 106₂, 106₃, 106₄, e usa as características de eficiência para identificar a ACU 106_m que requer a menor quantidade de consumo de energia incremental adicional para fornecer a refrigeração adicional.

[0046] Por exemplo, com referência a figura 2, assume-se que uma carga térmica existente nas 106₁, 106₂, 106₃, 106₄ é mostrada como ponto A em cada um dos gráficos 212, 214, 216 e 218. Pode ser visto que a inclinação aumenta como uma função da ineficiência nas curvas 212, 214, 216 e 218. Em outras palavras o aclave aumenta por-

que é requerida mais energia para obter os ganhos incrementais na refrigeração. Assumindo-se que um carregamento térmico corrente como indicado pelos pontos A nos gráficos 202, 204, 206 e 208, o circuito de processamento 112 na etapa 410 da figura 4 deve determinar que a ACU 106₂ deve manusear melhor a carga térmica adicional que resulta do processamento adicional, comparada com as ACUs 106₁, 106₃ e 106₄. O circuito de processamento 112 deve fazer esta determinação porque a inclinação da curva 214 no ponto A é o mínimo entre aqueles das curvas 212, 214, 216 e 218.

[0047] Com referência novamente a figura 4, depois da etapa 410, o circuito de processamento 112 executa a etapa 415. Na etapa 415, o circuito de processamento 112 faz com que a tarefa de processamento adicional seja alocada ou atribuída para um dos computadores servidores que é termicamente próximo a (isto é, é refrigerado principalmente por) ACU 106_m que foi identificada na etapa 410. Na operação ilustrativa da etapa 410 discutida acima em conexão com a figura 2, o circuito de processamento 112 na etapa 415 deve fazer com que a atribuição da tarefa adicional para um dos computadores servidores (por exemplo, servidores 104₆ a 104₉) que são termicamente próximos ou termicamente mais próximos a ACU 106₂.

[0048] Como discutido acima, o circuito de processamento 112 pode realizar a etapa 410 (e etapa 415) proativamente, em antecipação a uma nova requisição. Neste caso, o resultado das etapas 410 e 415 é identificar onde as próximas *n* tarefas serão atribuídas uma vez que as mesmas sejam recebidas. O circuito de processamento 112 armazena a informação gerada e atribui adequadamente as requisições de tarefas de processamento que chegam.

[0049] Em geral, a operação de atribuição de tarefa de processamento da figura 4, que é com base na eficiência incremental pode resultar na adição de novas tarefas em uma maneira eficiente em ener-

gia. Entretanto, as operações da figura 4 podem não resultar na melhor eficiência global. Por exemplo, as operações da figura 4, não levam em conta necessariamente as tarefas de processamento terminadas em uma forma de maximizar a eficiência. As tarefas terminadas podem reduzir a carga térmica no centro de processamento de dados. Ao mesmo tempo em que a carga térmica reduzida devido ao término de tarefas é pelo menos levada na etapa 410 (por considerar o carregamento térmico atual das ACUs 106₁, 106₂, 106₃, 106₄), o estado corrente do carregamento térmico nas ACUs 106₁, 106₂, 106₃, 106₄ pode ser ineficiente. A ineficiência pode ser devida ao fato de que os terminos de tarefas podem alterar drasticamente o carregamento térmico de forma que a mera atribuição incremental de tarefas adicionais pode não compensar adequadamente.

[0050] Adicionalmente, as operações da figura 4 presumem que todas (ou pelo menos um dado conjunto) as ACUs estejam correntemente ligadas, e possam ser carregadas variavelmente para obter um resultado ótimo. Em muitos casos, entretanto, a operação de menos do que todas as ACUs pode ser suficiente para manusear a carga térmica. A seleção inteligente de quais ACUs devem ser ativadas para manusear certos níveis de carga pode aumentar a eficiência no consumo de energia.

[0051] Para endereçar as questões com respeito às operações da figura 4, a figura 5 mostra um conjunto de operações que pode ser usado para alocar e/ou realocar tarefas de aplicação *existentes* para obter eficiência de refrigeração térmica. Para este fim, as operações da figura 5 determinam quais das ACUs 106₁, 106₂, 106₃, 106₄ devem ser ativadas para vários níveis de carga térmica global no centro de processamento de dados 102. Uma vez que é determinada uma melhor ou próxima a melhor configuração, os processos de aplicação podem ser alocados aos computadores servidores 104₁ a 104₁₈ adequa-

damente.

[0052] As etapas da figura 5 podem, mas não têm que, ser executadas em um sistema de gerenciamento de computador servidor 108 que também emprega as etapas da figura 4. As operações da figura 5 podem ser usadas como parte da etapa 310 da figura 3. Para este fim, as operações descritas abaixo usam a modelagem de carga para determinar a alocação mais (ou aproximadamente a mais) eficiente da carga térmica (e conseqüentemente carga de processamento) no centro de processamento de dados 102. O resultado das operações da figura 5 identifica uma melhor configuração de ACUs para ser acionada para uma dada carga térmica TL pra o centro de processamento de dados 102 integralmente. O sistema de gerenciamento de computador servidor 108 pode em seguida atribuir ou reatribuir tarefas de processamento de computador entre os computadores servidores 104_1 a 104_{18} de acordo com a configuração de ACU identificada.

[0053] Assim, em um exemplo, o circuito de processamento 112 pode primeiro obter a partir do elemento BAS 129 a carga térmica corrente TL no centro de processamento de dados, com base no carregamento das ACUs 106_1 , 106_2 , 106_3 , 106_4 . O circuito de processamento 112 pode, em seguida, executar as operações da figura 5 para determinar se uma alocação *melhor* das tarefas de processamento existentes pode obter um carregamento mais eficiente das ACUs 106_1 , 106_2 , 106_3 , 106_4 .

[0054] Alternativamente, a unidade de processamento 112 pode predeterminar as melhores configurações de ACU para uma pluralidade de cargas térmicas dadas, e em seguida armazenar os resultados. Deste modo, a unidade de processamento 112 pode ter armazenado na memória 110 uma pluralidade de pontos de dados da melhor (ou próximo a melhor) alocação de carga térmica entre as ACUs 106_1 , 106_2 , 106_3 , 106_4 para uma pluralidade de níveis de carga térmica do

centro de processamento de dados. Em seguida, de vez em quando, o circuito de processamento 112 obtém a carga térmica corrente TL e identifica o melhor modelo de carga térmica com base nos pontos de dados armazenados.

[0055] Como uma consequência, as operações da figura 5 podem ser geradas antes das operações reais do centro de processamento de dados, de modo que os resultados são armazenados para uso futuro. Alternativamente, as operações da figura 5, podem ser geradas em tempo quase real, em resposta ao recebimento da carga térmica calculada real TL .

[0056] Na etapa 505, o circuito de processamento 112 seleciona uma configuração de ACUs para ser ativado para uma dada carga térmica global TL para o centro de processamento de dados 102. Na etapa 510, o circuito de processamento 112 determina uma relação de carga da configuração de ACU selecionada, LR_{comb} , pela divisão do valor da TL pela capacidade de carregamento total da configuração de ACU selecionada, $MAXI_{comb}$. Em outras palavras,

$$LR_{comb} = TL/MAXI_{comb}$$

[0057] Na etapa 515, o circuito de processamento 112 determina se a configuração selecionada é uma combinação válida com base valor da relação de carga LR_{comb} . Por exemplo, se a configuração não pode resfriar adequadamente o centro de processamento de dados 102 sob as condições da carga térmica TL , então a configuração de ACU selecionada não é válida. Na modalidade corrente, o circuito de processamento de dados 112 executa a etapa 515 determinando se a relação de carga calculada LR_{comb} está entre 0,3 e 1,0. Em caso afirmativo, em seguida o circuito de processamento 112 avança para a etapa 520. Em caso negativo, então a configuração de ACU não é válida e o circuito de processamento 112 avança para a etapa 525, discutida adicionalmente abaixo.

[0058] Para chegar à etapa 520, foi determinado que a configuração/combinção de ACU selecionada é válida. Na etapa 520, o circuito de processamento 112 determina o consumo de energia para a combinação de ACU selecionada, PD_{comb} . Para este fim, o circuito de processamento 112 pode primeiro determinar, com base nas características de eficiência para as ACUs 106_1 , 106_2 , 106_3 , 106_4 (por exemplo, curvas 212, 214, 216 e 218 da figura 2), a melhor alocação da carga térmica TL entre a combinação de ACUs selecionada. Com o uso desta melhor alocação da carga TL entre a combinação de ACU selecionada, o circuito de processamento 112 em seguida calcula o consumo de energia global para cada uma das ACUs na combinação de ACU selecionada.

[0059] Depois disso, na etapa 530, o circuito de processamento 112 determina se o consumo de energia para a combinação selecionada é menor do que o consumo de energia corrente mais baixo para todas as combinações de ACU previamente avaliadas. Em outras palavras, o circuito de processamento 112 determina se $PD_{comb} < PD_{melhorcomb}$ em que $PD_{melhorcomb}$ é o consumo de energia da melhor combinação de ACUs calculada até agora, para a TL corrente. Se a resposta na etapa 530 é negativa, então o circuito de processamento 112 avança para a etapa 525, discutida adicionalmente abaixo. Se a resposta na etapa 530 é positiva, entretanto, então o circuito de processamento 112 na etapa 535 armazena a combinação de ACUs selecionada como a melhor combinação corrente para a carga TL , e determina $PD_{melhorcomb} = PD_{comb}$. O circuito de processamento 112 em seguida avança para a etapa 525.

[0060] Na etapa 525, a avaliação da configuração/combinção de ACU selecionada para o nível de carga TL do centro de processamento de dados é completada. Na etapa 525 o circuito de processamento 112 determina se existe alguma outra combinação de ACUs a ser ava-

liada. Em caso afirmativo, então o circuito de processamento 112 retorna a etapa 505 para selecionar outra configuração de ACUs. Em caso negativo, então a avaliação está completa, e o circuito de processamento 112 na etapa 535 armazena a combinação de ACU associada com a $PD_{melhorcomb}$ como a combinação de ACU apropriada para o nível de carga TL do centro de processamento de dados.

[0061] O circuito de processamento 112 pode subsequentemente alocar processos de aplicação de modo que a carga térmica TL será suportada pela combinação de ACUs que corresponde a $PD_{melhorcomb}$. Adicionalmente, como na etapa 320 da figura 3, o circuito de processamento 112 informa o elemento BAS 120 da combinação de ACUs que deve ser ativado com base na carga térmica corrente TL .

[0062] Em outra modalidade, um conjunto de operações pode ser realizado de vez em quando para determinar se uma realocação dos processos de aplicação entre os computadores servidores 104₁ a 104₁₈ pode permitir que uma das ACUs 106_n que operam correntemente seja desligada. O processo envolve calcular a capacidade de refrigeração máxima (a uma eficiência máxima ou aceitável) de cada unidade ACU ligada. A capacidade máxima total das ACUs ligadas é comparado, em seguida, à carga térmica corrente atual para determinar a capacidade de reserva das ACUs ligadas. Se a carga da ACU menos eficiente ligada está abaixo da capacidade de reserva para um dado período de tempo, então a ACU menos eficiente ligada é desligada, e as tarefas de processamento são realocadas entre os computadores servidores 104₁ a 104₁₈ adequadamente.

[0063] Será avaliado que as modalidades descritas acima sejam meramente ilustrativas e que os versados na técnica podem prontamente imaginar suas próprias implementações e modificações que incorporem os princípios da presente invenção e abrangidas pelo espírito e escopo da mesma.

REIVINDICAÇÕES

1. Método para gerenciamento de operações de centro de processamento de dados, caracterizado pelo fato de que inclui as seguintes etapas

a) obter informação de característica de eficiência para cada um de uma pluralidade de unidades condicionadoras de ar (106₁ - 106₁₈) em uma localização (102) que contém uma pluralidade de computadores servidores (104₁ - 104₁₈); e

b) empregar uma ou mais unidades de processamento (108) para alocar uma ou mais tarefas de processamento a um de uma pluralidade de computadores servidores (104₁ - 104₁₈) com base na informação de característica de eficiência e para

i) identificar para um nível de carga térmica total de cada um da dita pluralidade de computadores servidores (104₁ - 104₁₈) uma combinação de unidades de condicionamento de ar (106₁ - 106₁₈) capaz de refrigerar o nível de carga térmica total utilizando menos energia do que uma pluralidade de outras combinações de unidades de condicionamento de ar (106₁ - 106₁₈) capazes de refrigerar a carga térmica total;

ii) alocar tarefas de processamento a computadores servidores (104₁ - 104₁₈) com base em sua proximidade térmica à combinação identificada de unidades de condicionamento de ar (106₁ - 106₁₈).

2. Método de acordo com a reivindicação 1, caracterizado pelo fato de que a informação de característica de eficiência é representativa de uma curva de performance de eficiência para uma amplitude da saída de refrigeração variável.

3. Método de acordo com a reivindicação 2, caracterizado pelo fato de que a etapa b) compreende adicionalmente empregar as uma ou mais unidades de processamento (108) para alocar uma primeira requisição de processamento a um primeiro da pluralidade de

computadores (104₁ - 104₁₈) com base em uma proximidade térmica do primeiro da pluralidade de computadores (104₁ - 104₁₈) a uma unidade selecionada da pluralidade de unidades condicionadoras de ar (106₁ - 106₁₈).

4. Método de acordo com a reivindicação 3, caracterizado pelo fato de que as uma ou mais unidades de processamento (108) identifica a unidade selecionada da pluralidade de unidades condicionadoras de ar (106₁ - 106₁₈) com base na informação de característica de eficiência.

5. Método de acordo com a reivindicação 1, caracterizado pelo fato de que adicionalmente compreende:

empregar as uma ou mais unidades de processamento (108) para alocar uma ou mais tarefas de processamento a um primeiro da pluralidade de computadores servidores (104₁ - 104₁₈) com base em parte em uma distância do primeiro da pluralidade de computadores servidores (104₁ - 104₁₈) de uma primeira unidade de condicionamento de ar (106₁ - 106₁₈), e fazer com que a primeira unidade de condicionamento de ar (106₁ - 106₁₈) reduza uma saída de refrigeração da mesma.

6. Método de acordo com a reivindicação 5, caracterizado pelo fato de que adicionalmente compreende fazer com que a primeira unidade de condicionamento de ar (106₁ - 106₁₈) seja desligada.

7. Método de acordo com a reivindicação 5, caracterizado pelo fato de que o primeiro da pluralidade de computadores servidores (104₁ - 104₁₈) é mais distante que pelo menos um outro computador servidor (104₁ - 104₁₈) da primeira unidade de condicionamento de ar (106₁ - 106₁₈).

8. Método de acordo com a reivindicação 1, caracterizado pelo fato de que adicionalmente compreende desligar unidades de condicionamento de ar (106₁ - 106₁₈) fora da combinação identificada

de unidades de condicionamento de ar (106₁ - 106₁₈) em resposta à conclusão da subetapa ii).

9. Arranjo para gerenciamento de operações de centro de processamento de dados, caracterizado pelo fato de que compreende um sistema de gerenciamento de servidor de computador (108) que tem um circuito de processamento e uma memória, o circuito de processamento configurado para coordenar o uso de uma pluralidade de computadores servidores (104₁ - 104₁₈), o circuito de processamento executando software de virtualização configurado para gerenciar processamento de aplicação na pluralidade de computadores servidores (104₁ - 104₁₈);

a memória que armazena informação de característica de eficiência para uma pluralidade de unidades de condicionamento de ar (106₁ - 106₁₈) em uma localização (102) que contém uma pluralidade de computadores servidores (104₁ - 104₁₈);

sendo que o circuito de processamento é adicionalmente configurado para alocar uma ou mais tarefas de processamento a um da pluralidade de computadores servidores (104₁ - 104₁₈) com base na informação de característica de eficiência e para

i) identificar para um nível de carga térmica total de cada um da dita pluralidade de computadores servidores (104₁ - 104₁₈) uma combinação de unidades de condicionamento de ar (106₁ - 106₁₈) capaz de refrigerar o nível de carga térmica total utilizando menos energia do que uma pluralidade de outras combinações de unidades de condicionamento de ar (106₁ - 106₁₈) capazes de refrigerar a carga térmica total;

ii) alocar tarefas de processamento a computadores servidores (104₁ - 104₁₈) com base em sua proximidade térmica à combinação identificada de unidades de condicionamento de ar (106₁ - 106₁₈).

10. Arranjo de acordo com a reivindicação 9, caracterizado

pelo fato de que a informação de característica de eficiência é representativa de uma curva de performance de eficiência para uma amplitude da saída de refrigeração variável.

11. Arranjo de acordo com a reivindicação 10, caracterizado pelo fato de que o circuito de processamento é configurado adicionalmente para alocar uma primeira requisição de processamento a um primeiro da pluralidade de computadores (104₁ - 104₁₈) com base em uma proximidade térmica do primeiro da pluralidade de computadores (104₁ - 104₁₈) a uma unidade selecionada da pluralidade de unidades condicionadoras de ar (106₁ - 106₁₈).

12. Arranjo de acordo com a reivindicação 11, caracterizado pelo fato de que o circuito de processamento é configurado adicionalmente para identificar a unidade selecionada da pluralidade de unidades condicionadoras de ar (106₁ - 106₁₈) com base na informação de característica de eficiência.

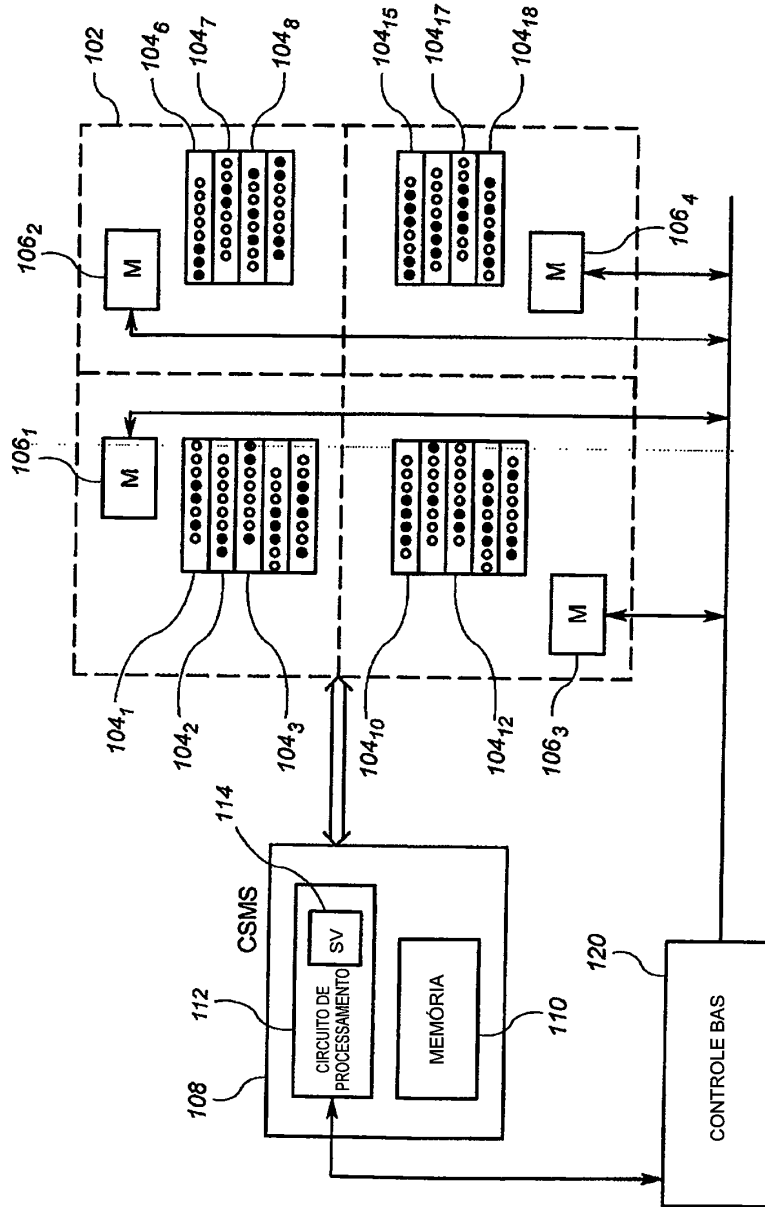


FIG. 1

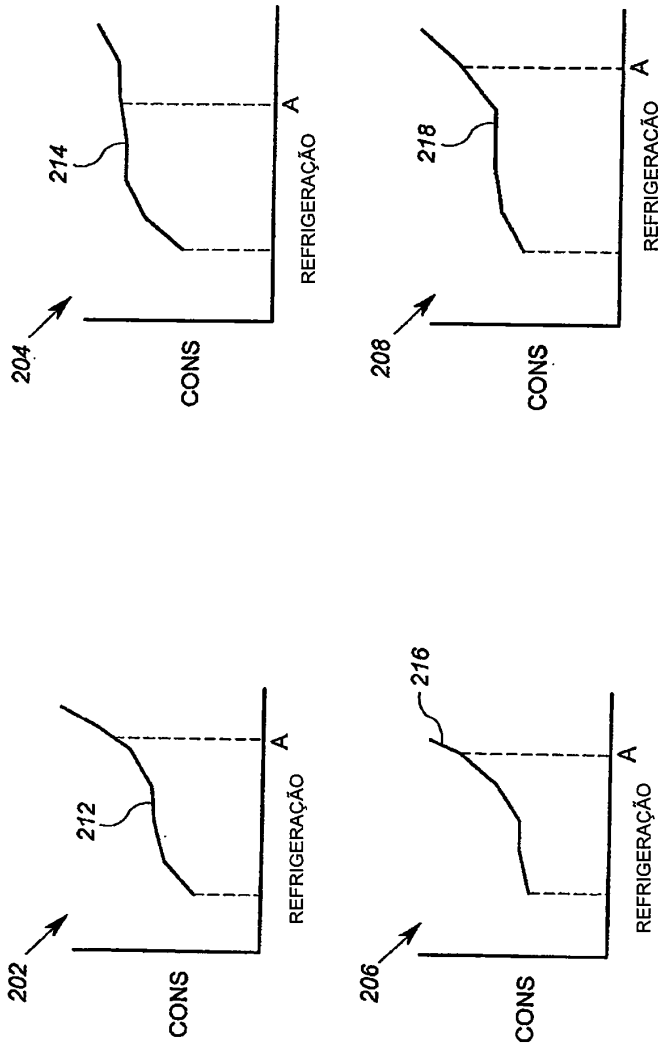


FIG. 2

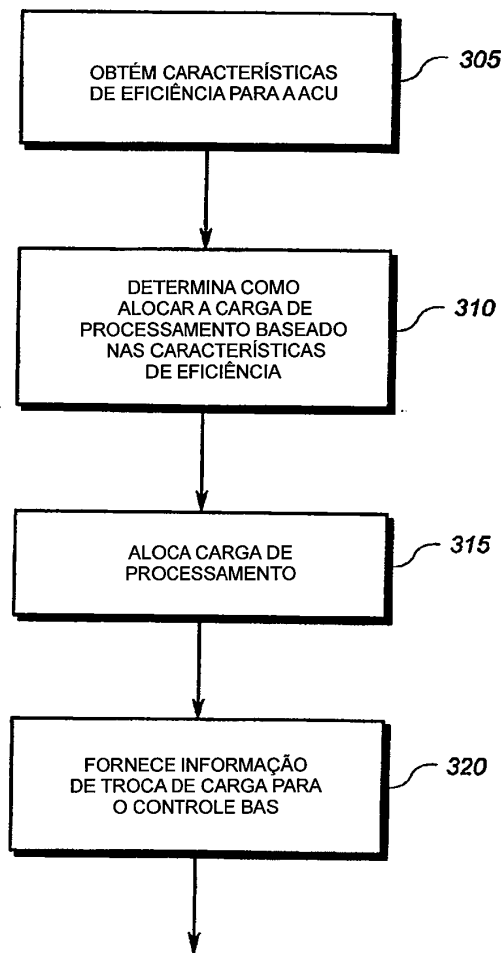


FIG. 3

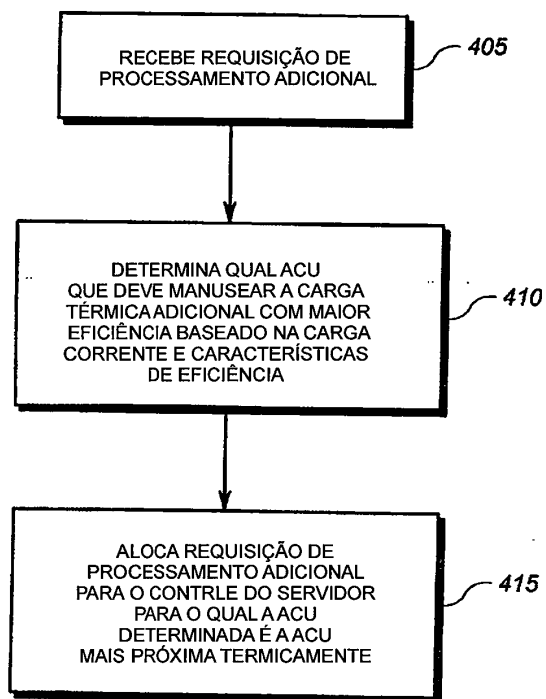


FIG. 4

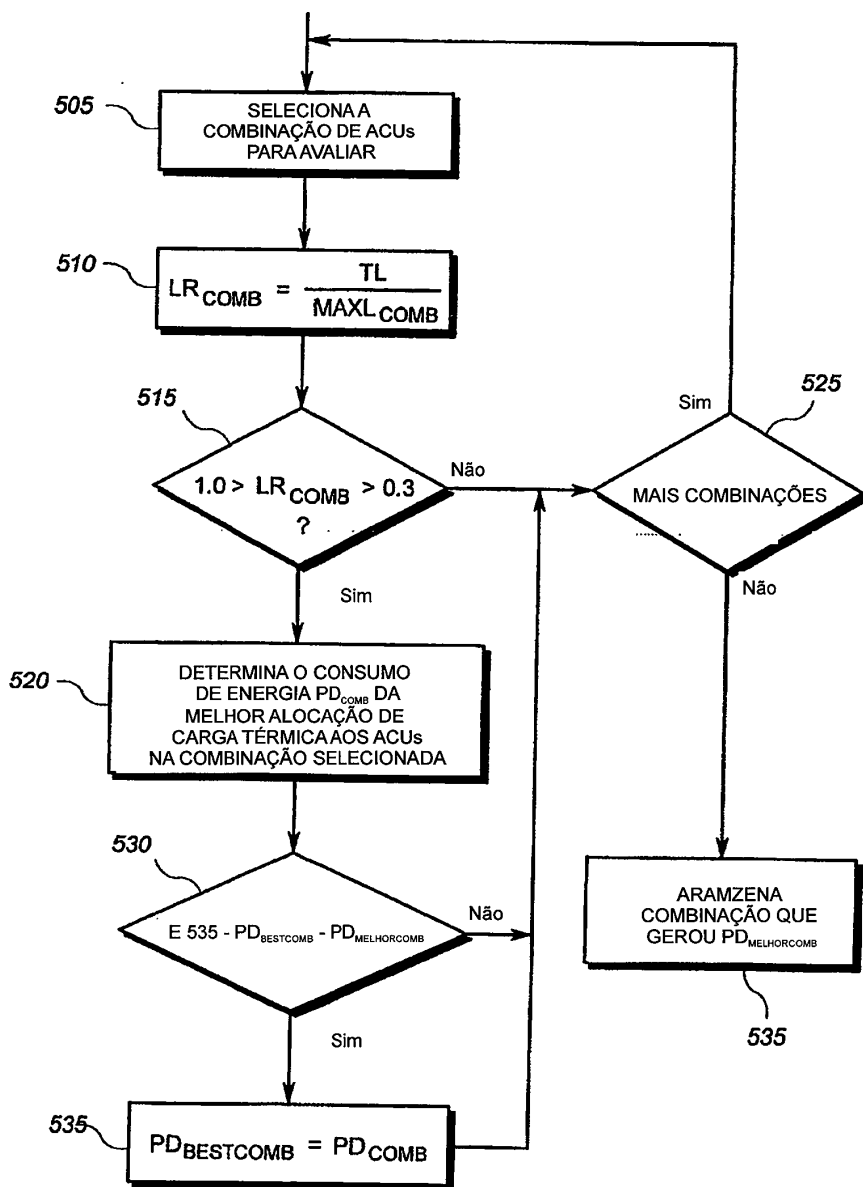


FIG. 5