



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

(21) PI 0914386-6 A2



(22) Data do Depósito: 21/10/2009

(43) Data da Publicação Nacional: 02/03/2021

(54) **Título:** CENTRO DE DADOS E MÉTODO PARA PROJETAR UM CENTRO DE DADOS

(51) **Int. Cl.:** G06F 19/00; G06F 9/44.

(30) **Prioridade Unionista:** 30/10/2008 US 12/261,250.

(71) **Depositante(es):** HEWLETT-PACKARD DEVELOPMENT COMPANY, .L.P..

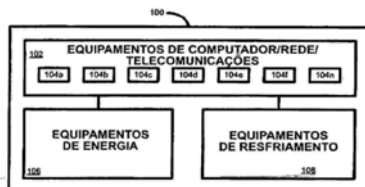
(72) **Inventor(es):** KEN JOHN HAMILTON; STEVE EINHORN; JAMES WARREN.

(86) **Pedido PCT:** PCT US2009061534 de 21/10/2009

(87) **Publicação PCT:** WO 2010/056473 de 20/05/2010

(85) **Data da Fase Nacional:** 27/04/2011

(57) **Resumo:** CENTRO DE DADOS E MÉTODO PARA PROJETAR UM CENTRO DE DADOS. De acordo com uma configuração da presente invenção, é provido um centro de dados compreendendo uma pluralidade de seções de centro de dados, cada seção tendo um nível diferente pré-definido de confiabilidade; e uma pluralidade de conjuntos de aplicações, cada conjunto de aplicações sendo povoado em uma da pluralidade de seções de centro de dados.



"CENTRO DE DADOS E MÉTODO PARA PROJETAR UM CENTRO DE DADOS"

Antecedentes

Um centro de dados é uma instalação que fornece serviços de computação para uma empresa. Um centro de dados aloja tipicamente uma variedade de equipamentos de computador e aplicações de software usadas para a provisão de serviços de computação. Os equipamentos de computador podem incluir computadores e servidores, equipamentos de rede, equipamentos de armazenagem e equipamentos de telecomunicações. Adicionalmente, equipamentos auxiliares adicionais são providos para permitir os equipamentos de computadores operarem. Tais equipamentos auxiliares podem incluir suprimentos de energia sem interrupção (UPS) e equipamentos de resfriamento.

A Associação da Indústria de Telecomunicações (TIA) TIA-942: Visão Geral de Normas de Centro de Dados e o Instituto Uptime definem um conjunto de 4 camadas de centros de dados baseados largamente em níveis de redundância. Por exemplo, os centros de dados da camada 1 oferecem a configuração mais básica, enquanto os centros de dados da camada 4 oferecem redundância total com disponibilidade de 99,995%. Não surpreendentemente, a redundância aumentada é igual a custos de capital e custos operacionais significativamente aumentados. Por meio de exemplo, até 50% de um centro de dados de camada 3 ou 4 podem ser assumidos com equipamentos de energia e resfriamento redundantes, o que se traduz em até 50% do custo de capital global do centro de dados.

Tipicamente, quando uma empresa constrói um centro de dados ela tipicamente constrói o centro de dados de camada mais alta para seu orçamento. A empresa então povoa o centro de dados com seus equipamentos de TI e povoa os equipamentos de TI com as aplicações de software da empresa.

Sumário

De acordo com um aspecto da presente invenção, é provido

um centro de dados compreendendo uma pluralidade de seções de centro de dados. Cada seção de centro de dados tem um nível diferente pré-definido de confiabilidade. Também é provida uma pluralidade de conjuntos de aplicações, cada conjunto de aplicações sendo povoado em uma da pluralidade de seções de centro de dados.

De acordo com um segundo aspecto da presente invenção, é provido um método para projetar um centro de dados. O método compreende obter detalhes de um conjunto de aplicações a serem povoadas no centro de dados. Para cada aplicação uma característica de prioridade é determinada. Baseado nas características de prioridade determinadas as aplicações são povoadas de diferentes seções de centro de dados, com cada seção de centro de dados tendo um nível diferente pré-definido de confiabilidade.

Descrição resumida

Configurações da invenção serão agora descritas, por meio de exemplos não limitantes somente, com referência aos desenhos anexos, nos quais: a-figura 1 é um diagrama de blocos mostrando um centro de dados monolítico em camadas de acordo com a técnica anterior;

A figura 2 é um diagrama de blocos mostrando um número de aplicações de software;

A figura 3A é um diagrama de fluxo delineando etapas de processamento de exemplo tomadas durante um processo de projeto de centro de dados de acordo com uma configuração da presente invenção;

A figura 3B é um diagrama de fluxo delineando etapas de processamento de exemplo tomadas durante um processo de projeto de centro de dados de acordo com uma configuração adicional da presente invenção;

A figura 4 é um diagrama de blocos mostrando um centro de dados híbrido em camadas de acordo com uma configuração da presente invenção; e

A figura 5 é um diagrama de blocos mostrando um centro de dados híbrido em camadas de acordo com configuração adicional da presente invenção.

Descrição detalhada

A figura 1 mostra um diagrama de blocos simplificado de um centro de dados monolítico em camadas 100 de acordo com a técnica anterior. O centro de dados 100 inclui

5 equipamentos de computação 102, que podem incluir computadores, servidores, equipamentos de rede e telecomunicações, nos quais são executados numerosas aplicações de software 104a a 104n. Os equipamentos 102 são energizados pelo equipamento de energia 106 e são

10 resfriados pelo equipamento de resfriamento 108. A natureza exata do equipamento de energia 106 e do equipamento de resfriamento 108 depende da classificação da camada do centro de dados 100. Por exemplo, um centro de dados de camada 4 pode ter múltiplas trajetórias de

15 distribuição de energia e resfriamento incluindo redundância $2N+1$ (isto é, 2 UPS cada um com redundância $N+1$), enquanto um centro de dados de camada 1 pode ter somente uma única trajetória para distribuição de energia e resfriamento, sem componentes redundantes.

20 Devido aos crescentes custos operacionais de operar um centro de dados, especialmente com relação à energia e resfriamento, os operadores de centro de dados estão buscando reduzir o custo de e melhorar a eficiência de

25 seus centros de dados. Atualmente isto está sendo feito aplicando soluções localizadas a energia, espaço, e resfriamento. Tais soluções localizadas incluem, por exemplo, o uso de sistemas de resfriamento mais eficientes em energia, consolidação de servidor, e subcontratação de carga de trabalho.

30 A presente invenção, entretanto, é baseada largamente na realização que eficiência significativa e economias de custos podem ser alcançadas se a natureza das aplicações pretendidas a serem operadas no centro de dados forem consideradas durante as fases de planejamento, projeto, e

35 configuração, como será explicado abaixo em maiores detalhes.

Referência será feita agora à figura 2, a qual mostra um

diagrama de blocos de um número de aplicações de software 104a a 104i que devem ser executadas ou são planejadas para serem executadas em um centro de dados. Referência adicional é feita aos diagramas de fluxo mostrados nas

5 figuras 3A e 3B. Aqueles experientes na técnica apreciarão, entretanto, que somente um número pequeno de aplicações de software são discutidas aqui por razões de clareza, e apreciarão adicionalmente que o número de aplicações de software em um centro de dados típico pode

10 correr nos milhares e além.

Na etapa 302 uma lista de aplicações de softwares a serem executados ou planejados a serem executados no centro de dados é obtida. No presente exemplo, as aplicações de software 104a a 104i estão identificadas. Estas

15 aplicações podem ser aplicações individuais ou podem ser um pacote de programas de uma ou mais aplicações.

Para cada aplicação de software 104a a 104i um impacto no negócio e nível de urgência são atribuídos (etapa 304). Neste sentido, em linha com a terminologia da Biblioteca

20 de Infraestrutura de Tecnologia ~~(TTL)~~ de informações padrões, o impacto no negócio se refere ao impacto no negócio da empresa caso esta aplicação de software não esteja disponível, devido, por exemplo, a uma falha de hardware. A urgência se refere ao atraso de tempo no qual

25 tal aplicação deve ser tornada disponível seguindo a aplicação se tornar indisponível. Por exemplo, em um ambiente bancário, uma aplicação fornecendo autorização para sacar fundos de um Caixa Automático pode ser classificada como tendo alto impacto e alta urgência,

30 enquanto uma aplicação provendo a transferência durante a noite de fundos de uma conta para outra pode ser classificada como tendo alto impacto e média urgência.

Na etapa 306 um nível de prioridade, baseado no impacto no negócio e urgência definidos é definido. A tabela 1

35 abaixo, por exemplo, mostra um mapeamento exemplar de impacto no negócio e urgência para prioridade.

Tabela 1 - Mapeamento de impacto no negócio e urgência

para prioridade

		Impacto		
		Alto	Médio	Baixo
Urgência	Alta	Crítica	Alta	Média
	Média	Alta	Média	Baixa
	Baixa	Média	Baixa	Planejar

Portanto, no presente exemplo, uma aplicação tendo uma alta urgência e alto impacto no negócio é definida como tendo uma prioridade crítica. Similarmente, uma aplicação tendo um alto impacto e média urgência é definida como tendo uma prioridade alta.

Na presente configuração as aplicações de software 104a, 104d, e 104e são determinadas a terem baixa prioridade, as aplicações 104c, 104f, e 104k como média prioridade, e as aplicações 104b, 104g, e 104i como prioridade crítica. Uma vez que a prioridade de cada aplicação de software tenha sido definida, o número e tipo de seções ou camadas de centro de dados podem ser determinados (etapa 208). Correntemente existem 4 camadas de centro de dados padrão industrial aceitas amplamente, com os centros de dados de camada 1 oferecendo os níveis de confiabilidade mais básicos, e os centros de dados de camada 4 oferecendo redundância total ou quase total com disponibilidade de 99,995%. Aqueles experientes na técnica apreciarão que diferentes números de seções ou camadas de centro de dados podem ser usados, cada um tendo um nível diferente de confiabilidade, redundância, ou outras características apropriadas.

Por exemplo, se as prioridades definidas das aplicações 104a a 104i incluírem prioridades baixa, média, e crítica, pode ser determinado inicialmente que um centro de dados compreendendo as camadas 1, 2 e 4 seja adequado. Neste caso, por exemplo, as aplicações tendo uma prioridade crítica podem ser povoadas nos equipamentos de computador em um centro de dados de Camada 4, aplicações tendo uma prioridade média em equipamentos de computador podem ser povoadas em um centro de dados de Camada 2, e aplicações tendo uma baixa prioridade podem ser povoadas

em equipamentos de computador em um centro de dados de Camada 1. Deste modo, cada aplicação é mapeada para a camada de centro de dados oferecendo um nível de confiabilidade e redundância correspondente à prioridade determinada daquela aplicação.

Na etapa 310 a capacidade de cada camada de centro de dados determinada na etapa 308 pode ser estimada. Esta estimativa pode ser baseada, por exemplo, nos requisitos de performance (tais como energia de processamento requerida, memória requerida, largura de banda de rede requerida, etc.) das aplicações intencionadas a serem povoadas em cada camada de centro de dados, um tamanho físico estimado da camada de centro de dados, e/ou uma densidade de energia estimada da camada de centro de dados.

De acordo com uma configuração adicional, um conjunto adicional de etapas, mostrado na figura 3B pode ser executado adicionalmente. As etapas adicionais visam otimizar, ou pelo melhorar, o design do centro de dados baseado em considerações financeiras.

Na etapa 312 um custo de capital estimado do centro de dados é determinado baseado em, por exemplo, o número determinado de camadas de centro de dados e sua capacidade.

Na etapa 314 as camadas de centro de dados determinadas na etapa 308 são analisadas, de uma perspectiva financeira, para determinar se qualquer consolidação das camadas pode ser conseguida. Por exemplo, em situações onde existe um grande número de aplicações de prioridade baixa e crítica, e um número baixo de aplicações de média prioridade, pode ser mais eficiente em custos projetar um centro de dados tendo uma seção de camada 1 para as aplicações de baixa prioridade e uma seção de camada 4 para as aplicações de prioridade média e crítica, ao invés de ter uma seção de camada 3 adicional só para o baixo número de aplicações de média prioridade. Isto está baseado no fato que a construção de cada seção de camada

de centro de dados tem um custo fixo mínimo associado com a mesma. Se apropriado, o design do centro de dados é racionalizado, e um novo custo estimado (etapa 316).

Na etapa 318 a capacidade de cada camada de centro de dados proposta pode ser modificada e seu efeito no custo estimado do centro de dados proposto avaliado (etapa 320).

Este processo pode ser repetido numerosas vezes, cada vez modificando características diferentes do centro de dados proposto. Deste modo, um centro de dados proposto pode ser obtido que seja inicialmente substancialmente otimizado a partir de uma perspectiva de negócios e, alternativamente, adicionalmente substancialmente otimizado de uma perspectiva financeira. Um centro de dados proposto pode incluir várias camadas de centros de dados diferentes de capacidades variadas dependendo de requisitos individuais.

As camadas de centro de dados descritas acima podem ser implementadas seja em centros de dados individuais separados fisicamente, como mostrado na figura 4, ou por um único centro de dados híbrido em camadas como mostrado na figura 5, ou em qualquer combinação ou arranjo adequado.

A figura 4 mostra um diagrama de blocos de um primeiro arranjo de centro de dados de acordo com uma configuração da presente invenção. Na figura 4, é mostrado um número de diferentes centros de dados 402 e 404. O centro de dados 402 é um centro de dados de camada 1, e aloja aplicações de baixa prioridade 104a, 104d, e 104f. O centro de dados 402 tem equipamentos de energia de camada 1 408 e equipamentos de resfriamento de camada 1 410. O centro de dados 404 é um centro de dados de camada 4 e aloja aplicações de média prioridade 104c, 104f e 104k e aplicações de prioridade crítica 104b, 104g e 104i. O centro de dados 404 tem equipamentos de energia 414 e equipamentos de resfriamento de camada 4 416. Com acesso e interconexão de rede apropriados, os centros de dados

402 e 404 provêm serviços de computação sem interrupção para empresas.

- A figura 5 mostra um centro de dados híbrido em camadas exemplar 500 projetado seguindo os métodos descritos acima. O centro de dados híbrido em camadas 500 provê diferentes seções de centro de dados provendo diferentes características de confiabilidade e redundância de diferentes camadas de centro de dados dentro de um único centro de dados físico. Por exemplo, equipamentos de rede e/ou telecomunicações 502, equipamentos de energia 504, e equipamentos de resfriamento 506 são arranjados para prover as características de confiabilidade e redundância de um centro de dados de camada 1 para as aplicações 104a, 104d, e 104e. Equipamentos de computador, rede e/ou telecomunicações 508, equipamentos de energia 510, e equipamentos de resfriamento 512 são arranjados para prover as características de confiabilidade e redundância de um centro de dados de camada 4 para as aplicações 104c, 104f, 104k, 104b, 104g, e 104i.
- Provendo um único centro de dados, economias de custos adicionais podem ser conseguidas permitindo o compartilhamento de instalações e infraestrutura comuns, tais como o compartilhamento de um recinto ou instalação física, compartilhamento de sistemas de segurança, controles de acesso, e similares.
- Baseando o design e configuração iniciais do centro de dados nas considerações comerciais, tais como a prioridade das aplicações que são executadas no centro de dados, significativas economias de custos e eficiência energética podem ser alcançadas. Por exemplo, se todas as aplicações 104a a 104i tivessem que ter sido alojadas em um único centro de dados monolítico de camada 4, custos de capital e custos operacionais significativos teriam sido desperdiçados provendo as aplicações de baixa e média prioridade com um nível de redundância e confiabilidade além e acima daquele determinado, pelo negócio, como necessário para aquelas aplicações. Em

centros de dados monolíticos existentes é estimado que até 50% das aplicações sendo executadas em tais centros de dados podem ser classificadas como não críticas para o negócio.

5 Embora as presentes configurações tenham sido descritas com referência a princípios de ITIL, aqueles experientes na técnica apreciarão que outras estruturas de trabalho de priorização de serviços comerciais, tais como ISO 20000, também podem ser usadas.

10 Em configurações adicionais, nem todas as etapas de métodos delineadas acima são executadas, ou são executadas em uma sequência diferente daquela descrita acima.

Também deve ser entendido que as técnicas da presente invenção poderiam ser implementadas usando uma variedade de tecnologias. Por exemplo, os métodos descritos aqui podem ser implementados em software sendo executado em um sistema de computador, ou implementados em hardware utilizando uma combinação de microprocessadores ou outros

20 circuitos integrados específicos para aplicação projetados especialmente, dispositivos lógicos programáveis, ou várias combinações dos mesmos. Em particular, os métodos descritos aqui podem ser

----- implementados por uma série de instruções executáveis por
25 computador residentes em um meio lido por computador. A mídia lida por computador adequada pode incluir memória volátil (p.ex., RAM) e/ou não volátil (p.ex., ROM, disco), mídia de ondas portadoras e de transmissão (p.ex., mídia de fio de cobre, cabo coaxial, fibra
30 ótica). As ondas portadoras exemplares podem assumir a forma de sinais elétricos, eletromagnéticos, ou óticos transportando correntes de dados digitais ao longo de uma rede local, uma rede acessível publicamente tal como a internet ou algum outro link de comunicação.

REIVINDICAÇÕES

1. Centro de dados, caracterizado pelo fato de compreender:
uma pluralidade de seções de centro de dados, cada seção
5 tendo um nível diferente pré-definido de confiabilidade;
e
uma pluralidade de conjuntos de aplicações, cada conjunto de aplicações sendo povoado em uma da pluralidade de seções de centro de dados.
- 10 2. Centro de dados, de acordo com a reivindicação 1, caracterizado pelo fato de cada conjunto de aplicações ter uma prioridade determinada, e adicionalmente onde cada conjunto de aplicações é povoado em uma seção de dados tendo um nível de confiabilidade correspondendo à
15 prioridade de nível determinada.
3. Centro de dados, de acordo com a reivindicação 2, caracterizado pelo fato de o nível de prioridade de cada conjunto de aplicações ser determinado baseado em um impacto no negócio e urgência determinados.
- 20 4. Centro de dados, de acordo com a reivindicação 1, caracterizado pelo fato de a capacidade de cada uma da pluralidade de seções de banco de dados ser baseada nos requisitos de performance das aplicações a serem povoadas nela.
- 25 5. Centro de dados, de acordo com a reivindicação 4, caracterizado pelo fato de o número de seções de centro de dados ser determinado em parte baseado na prioridade determinada de cada conjunto de aplicações e em parte em uma análise financeira.
- 30 6. Centro de dados, de acordo com a reivindicação 1, caracterizado pelo fato de a pluralidade de conjuntos de aplicações incluir adicionalmente aplicações planejadas a serem povoadas em uma da pluralidade de seções de centro de dados.
- 35 7. Centro de dados, de acordo com a reivindicação 1, caracterizado pelo fato de cada seção de centro de dados ser uma de um centro de dados físico independente ou uma

seção de um centro de dados dentro de um único centro de dados físico.

8. Centro de dados, de acordo com a reivindicação 7, caracterizado pelo fato de cada seção de centro de dados ser interconectada por rede.

9. Centro de dados, de acordo com a reivindicação 1, caracterizado pelo fato de cada seção de centro de dados compreender adicionalmente equipamentos de energia e resfriamento adequados para prover os níveis de confiabilidade e redundância requeridos para cada seção de centro de dados.

10. Centro de dados, de acordo com a reivindicação 9, caracterizado pelo fato de cada seção de centro de dados ser uma seção dentro de um único centro de dados físico, cada seção compartilhando elementos de infraestrutura comuns no mesmo centro de dados físico.

11. Método para projetar um centro de dados, caracterizado pelo fato de compreender:

obter detalhes de um conjunto de aplicações a serem povoadas no centro de dados;

determinar uma característica de prioridade para cada aplicação; e

determinar, baseado nas características de prioridade obtidas, uma pluralidade de seções de centro de dados nas quais as aplicações devem ser povoadas, cada seção de centro de dados tendo um nível pré-definido diferente de confiabilidade associado com a característica de prioridade determinada para cada aplicação.

12. Método, de acordo com a reivindicação 11, caracterizado pelo fato de compreender adicionalmente povoar pelo menos parte da pluralidade de aplicações em uma seção de centro de dados tendo um nível de confiabilidade correspondente ao nível determinado de prioridade de cada aplicação.

13. Método, de acordo com a reivindicação 11, caracterizado pelo fato de compreender adicionalmente determinar o nível de prioridade de cada conjunto de

aplicações baseado em um determinado impacto no negócio e urgência de cada aplicação.

14. Método, de acordo com a reivindicação 11, caracterizado pelo fato de a capacidade de cada uma da pluralidade de seções de centro de dados ser baseada nos requisitos de performance das aplicações a serem povoadas nela.

15. Método, de acordo com a reivindicação 11, caracterizado pelo fato de compreender adicionalmente determinar o número de seções de centro de dados baseado em parte na prioridade determinada de cada conjunto de aplicações e em parte em uma análise financeira.

16. Método, de acordo com a reivindicação 11, caracterizado pelo fato de compreender adicionalmente refinar a capacidade de cada uma da pluralidade de seções de centro de dados baseado em uma análise financeira.

17. Método, de acordo com a reivindicação 11, caracterizado pelo fato de compreender adicionalmente executar as etapas do método iterativamente para otimizar substancialmente o design do centro de dados.

18. Centro de dados, caracterizado pelo fato de compreender:

uma pluralidade de seções de centro de dados, cada seção tendo um nível pré-definido diferente de confiabilidade;

25 uma pluralidade de conjuntos de aplicações, cada conjunto de aplicações sendo povoado em uma da pluralidade de seções de centro de dados; e

sendo que cada conjunto de aplicações tem uma prioridade determinada, e adicionalmente sendo que cada conjunto de

30 aplicações é povoado em uma seção de dados tendo um nível de confiabilidade correspondente à prioridade de nível determinada.

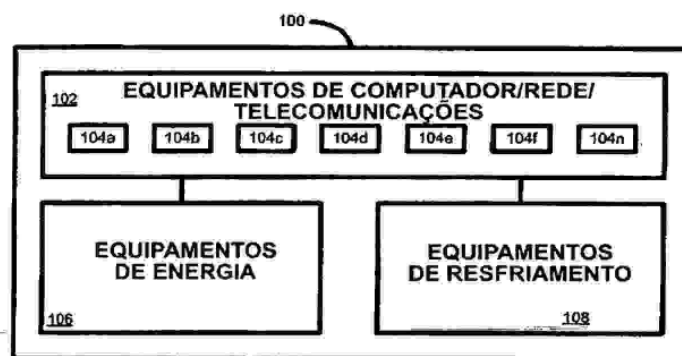


FIG.1
TÉCNICA ANTERIOR

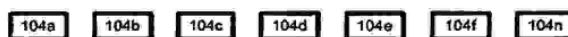


FIG.2



FIG.3A

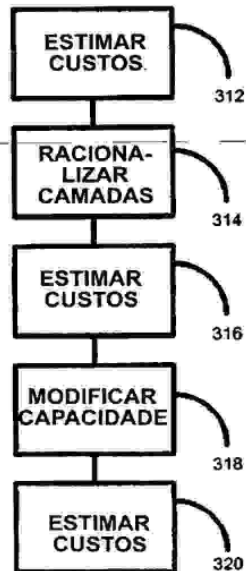


FIG.3B

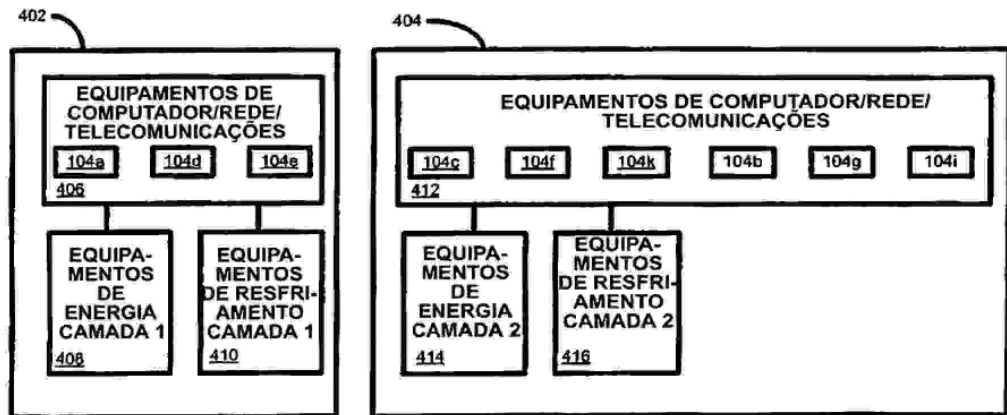


FIG. 4

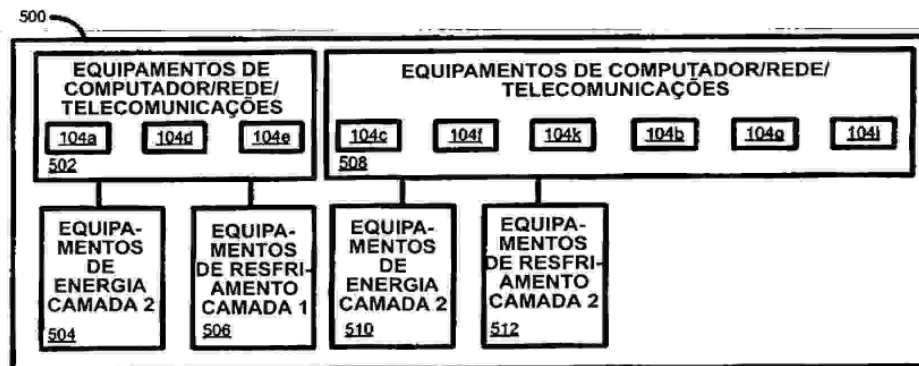


FIG. 5

RESUMO

"CENTRO DE DADOS E MÉTODO PARA PROJETAR UM CENTRO DE DADOS"

De acordo com uma configuração da presente invenção, é
5 provido um centro de dados compreendendo uma pluralidade
de seções de centro de dados, cada seção tendo um nível
diferente pré-definido de confiabilidade; e uma
pluralidade de conjuntos de aplicações, cada conjunto de
aplicações sendo povoado em uma da pluralidade de seções
10 de centro de dados.