

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号
特許第4738406号
(P4738406)

(45) 発行日 平成23年8月3日 (2011.8.3)

(24) 登録日 平成23年5月13日 (2011.5.13)

(51) Int. Cl.

F I

GO 6 F 11/30 (2006.01)

GO 6 F 11/34 (2006.01)

GO 6 F 11/30 E

GO 6 F 11/34 S

請求項の数 9 (全 16 頁)

(21) 出願番号	特願2007-506778 (P2007-506778)	(73) 特許権者	390009531
(86) (22) 出願日	平成17年4月6日 (2005.4.6)		インターナショナル・ビジネス・マシーンズ・コーポレーション
(65) 公表番号	特表2007-532993 (P2007-532993A)		INTERNATIONAL BUSINESS MACHINES CORPORATION
(43) 公表日	平成19年11月15日 (2007.11.15)		アメリカ合衆国10504 ニューヨーク州 アーモンク ニュー オーチャードロード
(86) 国際出願番号	PCT/EP2005/051531		
(87) 国際公開番号	W02005/098626	(74) 代理人	100108501
(87) 国際公開日	平成17年10月20日 (2005.10.20)		弁理士 上野 剛史
審査請求日	平成20年4月1日 (2008.4.1)	(74) 代理人	100112690
(31) 優先権主張番号	10/821,053		弁理士 太佐 種一
(32) 優先日	平成16年4月8日 (2004.4.8)	(74) 代理人	100091568
(33) 優先権主張国	米国 (US)		弁理士 市位 嘉宏
		最終頁に続く	

(54) 【発明の名称】 トランザクションを識別し、トランザクションをサポートする容量を管理する方法

(57) 【特許請求の範囲】

【請求項 1】

第1のサーバが複数の第2のサーバに要求を転送するようにされたネットワーク上でサーバ群内のサーバについて相対的負荷を監視する方法であって、
前記複数の第2のサーバの各々について、アドレス及びそれぞれのカウンタを含む第1の表を準備するステップと、
キャッシュ・ダンプを形成するために、アドレス解決プロトコル・キャッシュのコピーをストレージ域にダンプするステップと、
前記複数の第2のサーバのうちのどれが前記キャッシュ・ダンプに存在するのかを前記キャッシュ・ダンプから決定するステップと、
存在する前記第2のサーバの各々について、それぞれのカウンタを増分するステップと、
前記決定するステップと前記増分するステップとが完了した後で、前記それぞれのカウンタを互いに比較して、前記第2のサーバ間のロード不均衡が閾値より大きいかどうかを判断するステップと、
前記ロード不均衡が前記閾値より大きい場合には、前記ロード不均衡に関するメッセージを送信するステップと、を含む方法。

【請求項 2】

前記ダンプするステップと前記決定するステップとが繰り返し実行され、前記アドレス解決プロトコル・キャッシュが、前記ダンプするステップの各々の繰り返しの後でクリアされる、請求項1に記載の方法。

【請求項 3】

前記準備するステップは、最初に実行された後は、サーバが前記複数の第 2 のサーバに加えられたか又はこれから除去された場合にのみ実行される、請求項 1 に記載の方法。

【請求項 4】

定期的に、前記カウンタの値をログ・ファイルに書き込むステップと、前記カウンタをクリアするステップとをさらに含む、請求項 1 に記載の方法。

【請求項 5】

インターネットに接続されており、アドレス解決プロトコル・キャッシュを有する第 1 のサーバと、

前記第 1 のサーバとネットワークを形成するように接続された複数の第 2 のサーバであって、前記第 1 のサーバが、インターネットから受信した複数の要求を前記複数の第 2 のサーバのそれぞれのサーバに転送する、複数の第 2 のサーバと、

前記第 1 のサーバ上で稼動するプログラムであって、前記複数の第 2 のサーバの各々について、アドレス及びそれぞれのカウンタを含む第 1 の表を準備するステップと、前記アドレス解決プロセス・キャッシュをキャッシュ・ダンプにダンプするステップと、前記キャッシュ・ダンプ内の各々のアドレスについて、前記キャッシュ・ダンプ内の前記アドレスが前記第 1 の表内のアドレスと一致する場合には、それぞれのカウンタを増分するステップと、前記複数のカウンタを互いに比較して、前記第 2 のサーバ間のロード不均衡が閾値より大きいかどうかを判断するステップと、前記ロード不均衡が前記閾値より大きい場合には、前記ロード不均衡に関するメッセージをオペレータ・コンソールに送信するステップとを実行するプログラムと、を含むネットワーク。

【請求項 6】

前記プログラムが前記ステップを繰り返し実行し、前記アドレス解決プロトコル・キャッシュが各前記ダンプするステップの後でクリアされる、請求項 5 に記載のネットワーク。

【請求項 7】

前記準備するステップは、最初に実行された後は、サーバが前記複数の第 2 のサーバに加えられたか又はこれから除去された場合にのみ実行される、請求項 5 に記載のネットワーク。

【請求項 8】

機械可読フォーマットのコンピュータ・プログラムであって、複数のサーバの各々について、アドレス及びそれぞれの第 1 のカウンタを含む第 1 の表を準備する手順と、

キャッシュ・ダンプを形成するために、アドレス解決プロトコル・キャッシュをダンプする手順と、

前記キャッシュ・ダンプ内の各々のアドレスと一致する前記第 1 の表内の各々のアドレスについて、前記第 1 の表内のそれぞれのカウンタを増分する手順と、

前記それぞれの第 1 のカウンタを互いに比較して、前記第 2 のサーバ間のロード不均衡が閾値より大きいかどうかを判断する手順と、

前記ロード不均衡が前記閾値より大きい場合には、前記ロード不均衡に関するメッセージをオペレータ・コンソールに送信する手順と、

をコンピュータに実行させるためのコンピュータ・プログラム。

【請求項 9】

前記プログラムが、前記第 2 のサーバの各々のサーバが所与の期間にわたりアクティブであった時間の相対量の報告を提供する手順をさらにコンピュータに実行させる、請求項 8 に記載のコンピュータ・プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本出願は、一般に、適時かつ経済的な方法で、コンピュータ・サービスを管理すること

10

20

30

40

50

に関する。より具体的には、本出願は、サーバの1つ又はそれ以上のものが、期待される負荷を負担していない場合を検出するために、サーバ群に対するトラフィックを監視することに関する。

【背景技術】

【0002】

どのような種類のコンピュータ・システムを構築する場合にも、リソースを望ましい結果に適合させることが重要である。割り当てられたコンピューティング・リソースが少なすぎると、鈍い応答、顧客の不満、及び、一般に不良な結果がもたらされることになる。しかしながら、多すぎるリソースは無駄になり、事業収益に影響する。従って、アプリケーションとそれぞれのリソースの実際の能力との間の適合性を良くすることが、システムの財政上の成功に重要である。このことは、コンピュータ・リソースが顧客に与えられて共有され、採算性がこれらのリソースの最善の使用に大きく依存するプロビジョニング・システムにおいて、特に当て嵌まる。しかしながら、コンピュータ・システムの構築における最善の努力にもかかわらず、リソースの能力は、実際のリアルタイムの状況で必ずしも十分に利用されているわけではなく、見分けるのが困難な形で設計の構成とは異なることがある。例えば、評判の良いウェブサイトは、アクセスのために単一のアドレスを公開するが、このアドレスを使用するのは、実際の対話を管理する多数のトランザクション・サーバの1つに要求を転送するサーバに対してのみである。既存のプログラムは、要求を種々のトランザクション・サーバに配信するが、これらのプログラムは、必要な情報のすべてを持たないか又は問題を修正するための能力がないという状況が生じることがある。

【0003】

一例を示すことができる。ニュース・サービスは、速報についてコンタクトを取ることができるウェブ・アドレスを提供する。このアドレスは、着信要求を、通常の負荷に対処できる5から10のサーバ群に向ける単一のウェブ・サーバを指す。しかしながら、ツイントワーへのテロリスト攻撃といった大災害の間は、ニュース・サービスには要求が殺到する可能性がある。情報技術サービスを非常に迅速に提供することに特化したプロビジョニング・サービスの使用は、ニュース・サービスが、予想外の負荷を取り扱うための付加的なサーバと連動し、新しい要求は付加的なサーバに方向転換されるようにすることを可能にする。しかしながら、プロトコルによれば、第1に、所与のユーザからの要求を取り扱うサーバは、トランザクション全体を通して、その要求を取り扱い続けるため、より良いサービスを提供することができる他のサーバがあるにもかかわらず、多数のユーザが元のサーバに残り、限られた帯域幅を得るために競い合うことになる。既存のシステムは、あるサイトにアクセスしているユーザの数を追跡する能力、又は、元のサーバ及び最近割り当てられたサーバの両方が、等しく利用されているかどうかをチェックする能力がないことがある。

【0004】

使用状況の他の面から見ると、大容量の大型サーバは、その能力のすべてを利用していないことがあるが、このことは、使用状況を追跡するのに利用可能なデータがないため、システム管理ソフトウェアによって、明確に認識されていないという可能性がある。過去には、この情報は、障害分析中に、プロジェクトに関する事後分析で発見されることが最も多かった。

【発明の開示】

【発明が解決しようとする課題】

【0005】

従って、サーバのおおよその使用状況を追跡して、不均衡が示され修正されるようにすることができる機構を有することが望ましい。

【課題を解決するための手段】

【0006】

本発明の第1の態様は、第1のサーバが複数の第2のサーバに要求を転送するようにされたネットワーク上でサーバ群内のサーバについて相対的負荷を監視するための方法を提

10

20

30

40

50

供する。この方法は、複数の第2のサーバの各々について、アドレス及びそれぞれのカウンタを含む第1の表を準備するステップと、キャッシュ・ダンプを形成するために、アドレス解決プロトコル・キャッシュのコピーをストレージ域にダンプするステップと、複数の第2のサーバのうちのどれがキャッシュ・ダンプに存在するのかをキャッシュ・ダンプから決定するステップと、存在する第2のサーバの各々について、それぞれのカウンタを増分するステップと、を含む。決定するステップと増分するステップとが完了した後で、この方法は、複数のそれぞれのカウンタを互いに比較して、第2のサーバ間のロード不均衡が閾値より大きいかどうかを判断する。ロード不均衡が閾値より大きい場合には、ロード不均衡に関するメッセージが送信される。

【0007】

10

本発明の第2の態様は、インターネットに接続されており、アドレス解決プロトコル・キャッシュを有する第1のサーバと、第1のサーバ上で稼動するプログラムと、第1のサーバとネットワークを形成するように接続された複数の第2のサーバと、を含むネットワークを提供する。第1のサーバは、インターネットから受信した複数の要求を、複数の第2のサーバのそれぞれのサーバに転送する。プログラムは、

【0008】

複数の第2のサーバの各々について、アドレス及びそれぞれのカウンタを含む第1の表を準備するステップと、

【0009】

アドレス解決プロセス・キャッシュをキャッシュ・ダンプにダンプするステップと、

20

【0010】

キャッシュ・ダンプ内の各々のアドレスについて、キャッシュ・ダンプ内のアドレスが第1の表内のアドレスと一致する場合には、それぞれのカウンタを増分するステップと、

【0011】

複数のカウンタを互いに比較して、第2のサーバ間のロード不均衡が閾値より大きいかどうかを判断するステップと、

【0012】

ロード不均衡が閾値よりも大きい場合には、ロード不均衡に関するメッセージを送信するステップと、
を実行する。

30

【0013】

本発明の1つの態様は、サーバ間の不均衡を検出して、修正を行うことができるように通知を提供するための方法を含む監視プログラムを提供する。本発明の一実施形態は、アドレス解決プロトコル（ARP）キャッシュを読み取って、どのサーバ・アドレスが存在するかを決定し、次いで、ARPキャッシュをクリアする。アドレス解決プロトコルは、それがスヌープするトラフィックと新しい要求から、キャッシュを再設定する。この照会及びクリア操作を数秒毎に実行することにより、監視プログラムは、種々のサーバの活動レベルのピクチャを形成することができる。顕著な不均衡が存在する場合には、プログラムは、メッセージを制御コンソールに送信して、その問題が知られ、解決することができるようにする。監視プログラムは、その結果を定期的にログ・ファイルに書き込んで、必要に応じて、種々の期間についての報告が生成されることを可能にすることができる。この機構により、不均衡を迅速に検出できるだけでなく、期待されるリソースの使用パターンをその実際の使用と比較することにより、十分に使用されていない状態を見つけることもできる。

40

【0014】

本発明によるコンピュータ・プログラム製品は、プログラムが上述の方法を行うように実行されるデータ処理装置の動作を制御するための一組の命令を実装する。

【発明を実施するための最良の形態】

【0015】

本発明の特徴と信じられる新規なフィーチャは、添付の特許請求の範囲に示される。し

50

かしながら本発明自体、好ましい使用様式、並びに本発明の更なる目的及び利点は、添付の図面と併せて、以下の例示的な実施形態の詳細な説明を参照して読むことにより、最も良く理解されるであろう。

【0016】

ここで図面を参照すると、図1は、インターネット120を通じて顧客にサービスを提供するプロビジョニング・システムの使用を示している。サービス・プロバイダ124は、種々の被管理サービス128を与える。これらのサービスは、以下に限定されるものではないが、サーバ、端末、ディスク・スペース、及び広範囲のソフトウェアを含むことができる。サービス128を望む顧客122は、インターネット120を通じて、又は電話といった他の手段（図示せず）を用いて、プロバイダ124にコンタクトすることができる。次いで、プロビジョニングは、自動プロビジョニング・システム126により管理される。契約上の取り決めは、種々の形態をとることができる。例えば、顧客は、サービス・プロバイダ124と契約を結んで、サーバ及び端末を顧客のサイトに設置し、顧客122が人員追加するときには、プロバイダ124が端末を追加するようにすることができる。ソフトウェアのアップグレードは、要求が受信された時に、自動プロビジョニング・システムにより、自動的に顧客のコンピュータにダウンロードされるようにすることができる。第2の顧客は、異なるプロジェクトが開始したり終了したりするため、コンピューティングのニーズが時間と共に大幅に変化する事業を有するとすることができる。この顧客は、必要に応じて、自分の場所に基本コンピュータ機器を持つが、プロジェクトの有効期間中に、サービス・プロバイダからのサーバ、ストレージ、及びソフトウェアを利用する契約をすることができる。こうした第3者プロビジョニング・システムは、引用によりここに組み入れる、「第3者リソース・プロビジョニング管理のための方法及びシステム」と題する米国特許出願第20020169876号に説明されている。

【0017】

図2は、図1のプロビジョニング・システム126によるプロビジョニングが可能で、本発明を実施することができるデータ処理システムのネットワークの図形表示を示す。ネットワーク・データ処理システム100は、ネットワーク102に接続されており、これは、ネットワーク・データ処理システム100内で互いに接続された種々のデバイスとコンピュータとの間に通信リンクを提供するのに使用される。

【0018】

図示例においては、サーバ104は、ネットワーク102及びインターネット120に接続されている。さらに、ストレージ・ユニット106とサーバ108、110、112、114及び116がネットワーク102に接続されている。図示例においては、サーバ104のアドレスは、インターネット上のユーザのためのコンタクト・ポイントとして提供される。しかしながら、サーバ104は、次いで、要求を他のサーバ108ないし116に転送する。サーバ104は、サーバ間でロード・バランシングを実行して、ロードをできるだけ均等に拡げるようにするプログラムを有するが、ロード・バランシングは、「厄介」なものであり、要求がサーバ110といった所与のサーバに送信されると、そのユーザからのすべての要求がそのセッション中、同じサーバに向けられることになる。負荷が重い間は、プロビジョニング・システム126は、付加的なサーバを提供して作業負荷を処理することができる。ネットワーク・データ処理システム100は、図示されていない他のデバイスを含むことができる。図示例においては、ネットワーク・データ処理システム100は、サーバ104によりインターネット120に接続される。インターネット120は、互いに通信を行うために、伝送制御プロトコル／インターネット・プロトコル（TCP／IP）のプロトコル・スイートを用いるネットワーク及びゲートウェイの世界的な集合を表す。インターネットの中心は、データ及びメッセージを送る何千もの商業、政府、教育及びその他のコンピュータ・システムで構成される、主ノード又はホスト・コンピュータ間的高速データ通信回線のバックボーンである。ネットワーク・データ処理システム100は、例として、イントラネット、ローカル・エリア・ネットワーク（LAN）、又は広域ネットワーク（WAN）といった複数の異なる種類のネットワークとして実

装することができる。図1は、一例であり、本発明の構造的な制限を意図するものではない。

【0019】

図3を参照すると、サーバ104ないし116の1つ、又はプロビジョニング・システム126の一部であるサーバとして実装することができるデータ処理システムのブロック図が、本発明の好ましい実施形態により示されている。データ処理システム200は、システム・バス206に接続された複数のプロセッサ202及び204を含む対称型マルチプロセッサ(SMP)システムとすることができる。代替的には、単一プロセッサ・システムを採用してもよい。さらに、システム・バス206には、ローカル・メモリ209に対するインターフェースを提供するメモリ・コントローラ/キャッシュ208が接続される。I/Oバス・ブリッジ210は、システム・バス206に接続されて、I/Oバス212に対するインターフェースを提供する。メモリ・コントローラ/キャッシュ208とI/Oバス・ブリッジ210は、図示のように統合することができる。

10

【0020】

I/Oバス212に接続された周辺機器相互接続(PCI)バス・ブリッジ214は、PCIローカル・バス216に対するインターフェースを提供する。複数のモデムをPCIローカル・バス216に接続することができる。典型的なPCIバスの実装は、4つのPCI拡張スロット又はアドイン・コネクタをサポートする。図2のクライアント108ないし112への通信リンクは、アドイン・コネクタを通してPCIローカル・バス216に接続されたモデム218及びネットワーク・アダプタ220により提供することができる。

20

【0021】

付加的なPCIバス・ブリッジ222及び224は、付加的なPCIローカル・バス226及び228に対するインターフェースを提供し、ここから付加的なモデム又はネットワーク・アダプタをサポートすることができる。このようにして、データ処理システム200は、複数のネットワーク・コンピュータへの接続を可能にする。メモリ・マップ・グラフィックス・アダプタ230及びハードディスク232もまた、図示のように、直接又は間接的に、I/Oバス212に接続することができる。

【0022】

当業者であれば、図3に示すハードウェアは異なってもよいことを理解するであろう。例えば、光学ディスク・ドライブ及び同様なものといった他の周辺デバイスもまた、図示されるハードウェアに加えて、又はその代わりに使用することができる。図示例は、本発明に関して構造的な制限を示唆することを意味するものではない。

30

【0023】

図3に示すデータ処理システムは、例えば、拡張対話式エグゼクティブ(AIX(R))オペレーティング・システム又はLinux(R)オペレーティング・システムを稼動する、ニューヨーク州アームック所在のインターナショナル・ビジネス・マシーンス・コーポレーションの製品であるIBM(R) eServer(R) pSeries(R)システムとすることができる。

【0024】

通信するためには、すべてのネットワーク・デバイスは、固有の、製造者により定義されたメディア・アクセス制御(MAC)アドレスを有し、これにより識別されることができる。MACアドレスは6バイトを含み、各対がコロンで分離された、12個の16進数字として書かれる。例示的なMACアドレスは、0A:4B:00:00:07:08となる。一方、インターネット・プロトコル(IP)アドレスは、デバイスとインターネットとの間の接続のためのアドレスである。IPアドレスは領域によって分けられ、インターネット・サービス・プロバイダ(ISP)といった、その領域内のローカルな登録機構により割り当てられる。IPアドレスは、アドレスが属するクラスによって定まるフォーマットの4バイトである。10000000 00001011 00000011 00011111として表される2進数アドレスをもつ例示的なIPアドレスは、128.

40

50

11. 3. 31と書かれる。図2のネットワーク102といった1つのネットワーク内では、ホストはそのMACアドレスにより知られるが、TCP/IPといった高レベルのプロトコルでは、IPアドレスといったシンボリック・アドレスによりホストを指定する。

【0025】

コンピュータが、所与のIPアドレスにメッセージを送信することを望む場合には、このIPアドレスを、低レベルのデバイスが理解できる物理ハードウェア・アドレス(MAC)に変換しなくてはならない。この変換は、この目的のために、ARPキャッシュと呼ばれるルックアップ表を維持するアドレス解決プロトコル(ARP)により提供される。ARPが、そのキャッシュ内にMACアドレスをまだもっていない場合には、この情報の要求はネットワークに配信される。

10

【0026】

図4を参照すると、ARP要求メッセージ300のフォーマットが示されている。メッセージの最初の2バイトは、「1」がイーサネットを示す、ハードウェア・アドレス・タイプ302である。次の2バイトは、0x0800がIPを示す、プロトコル・アドレス・タイプ304である。1バイトのハードウェア・アドレス長フィールド306及び1バイトのプロトコル・アドレス長フィールド308が次の2バイトであり、次いで「1」が要求を示し、「2」が応答を示す、2バイトの動作コード310が続く。次の6バイトは、送信側のMACアドレス312であり、次いで、4バイトの送信側のIPアドレス314、6バイトのターゲットMACアドレス316、及び4バイトのターゲットIPアドレス318が続く。

20

【0027】

ARP要求を送信するのに先立って、コンピュータは、それ自体のMACアドレス312及びIPアドレス314を書き込み、ターゲットIPアドレス316を書き込み、動作コード310を「1」に設定し、ターゲットMACアドレスをゼロに設定する。この要求は、次いで、配信され、同時に、要求が送信されたという表示をもって、ターゲットIPアドレスがARPキャッシュに加えられる。送信側が応答を受信した場合には、受信側のMACアドレスを、対応するIPアドレスと共にARPキャッシュに格納し、同時に、その入力項目が完了したとしてマークする。要求処理の相手方においては、ターゲット機械がそれ自体のアドレスを認識し、最初に、該ターゲットが、このソースから付加的なコンタクトを受信する可能性があるため、送信側のアドレスをそれ自体のキャッシュに保存し、次に、メッセージに応答して、それ自体のMACアドレスを提供する、という2つの動作により要求に応答する。アドレスをそれ自体のものとして認識しない他のコンピュータは、配信メッセージを破棄する。

30

【0028】

ARP要求が、個々のネットワーク内で応答を引き出さない場合には、要求は他のネットワーク・ハブに配信されて、望ましい接続を探し当てる。この場合、物理ハードウェア・アドレスは、ルート情報を加えた形で戻される。ARPキャッシュは、情報が再度アクセスされたかどうかにかかわらず依存することもあるし、しないこともある、所定の時間の長さについてのコンタクト情報を維持する。所定の間隔、例えば10分が経過すると、アドレスは削除されて、より最近のアドレスのためにスペースが利用可能になる。ARP入力項目がタイムアウトするまでの時間制限、並びに、この時間制限を更新できるかどうかは、使用されているオペレーティング・システムにより管理される。

40

【0029】

図5は、簡略化された例示的なARPキャッシュ400を示す。このシステムにおいては、アドレス解決プロトコルは、要求を通して受信したアドレスを保存するだけでなく、通過する通信ストリームをスヌープして、キャッシュ内にまだ存在しないアドレスがあれば、それを追加する。ARPキャッシュの正確なフォーマットは、システムごとに異なっていてよく、この図に示すキャッシュは、本発明に必要な情報だけを含む。キャッシュ400には、各々のアクティブなアドレスについての記録があり、少なくともサーバ又はクライアントのIPアドレス410、対応するMACアドレス415、及び時間フィールド

50

ド 4 2 0 を示す。時間フィールド 4 2 0 は、規定の期間が経過した場合、或いは、アドレスが規定の期間においてアクティブではなかった場合に、いつ項目を表から除去すべきかを判断するのに用いられる。例示的なキャッシュ 4 0 0 は、存続期間 (T T L) フィールドを使用する。このシステムでは、項目がキャッシュに加えられた時に、2 1 6 0 秒の時間値がこのフィールドに挿入される。この値は毎秒 1 ずつ減り、ゼロに到達した場合には、スペースは、別のアドレスによる使用のために開放される。しかしながら、関連アドレスが要求される又はスヌープされる度に、T T L フィールドは開始値である 2 1 6 0 にリセットされる。このようにして、項目は、そのアドレスがアクティブである限り、2 1 6 0 秒の間、キャッシュ内に残る。システムは、さらに、逆の発想を使用して、入力に際してカウンタをゼロに設定し、次いで、限界に到達してスペースが開放されるまで、カウンタを毎秒増分させることができる。カウンタをゼロにリセットすることができる場合には、項目は、通常の活動を有する限り、キャッシュ内に残るが、カウンタをリセットすることができない場合には、項目は一定の長さの時間しかキャッシュ内に残らない。

10

【0030】

本発明のプログラムが使用される場合には、プログラムは定期的に A R P キャッシュをクリアする。クリアが頻繁に実行される場合には、キャッシュは、最後にクリアされて以来アクティブであるアドレスだけを含む。本発明のプログラムは、最初にインストールされた場合には、典型的には、キャッシュをチェックし、およそ毎秒に一回それをクリアする。関連のサーバがどれだけ速くキャッシュに再追加されるかを監視することにより、プログラムは、サーバ活動の概略を求めることができる。

20

【0031】

キャッシュ・ダンプがチェックされるに伴い、情報は図 7 に示す表に格納される。表 6 0 0 は、サーバ・アドレス 6 0 2 のフィールドと、サーバに関連するカウンタ 6 0 4 のフィールドとを含む。サーバ・アドレス 6 0 2 がキャッシュ・ダンプにおいて見出される度に、それぞれのカウンタ 6 0 4 が増分される。

【0032】

ここで図 6 を参照すると、本発明のプログラムの例示的な実施形態による、複数のサーバの使用レベルを比較するのに必要な動作を見ることができる。プログラムは、ネットワーク (例えばネットワーク 1 0 2) とより大きなインターネットとの間のインターフェースとしての機能を果たすゲートウェイ・サーバ (例えば 1 0 4) 上にインストールされる。このネットワークに送られた全パケットは、サーバ 1 0 4 と関連するルータを通過して、A R P キャッシュは、サーバ 1 0 8 ないし 1 1 6 の使用に関する情報を含むようになる。例えば、以前に説明したニュース・サービスは、公開されたウェブ・アドレスに配置されたサーバ上に、本発明のプログラムをインストールする。

30

【0033】

革新的なプログラムにアクセス可能な、関心のあるサーバ、例えば、サーバ 1 0 8 ないし 1 1 6 のアドレスのリストが維持される。一時的な条件のためにシステムが付加的なサーバを加えた場合には、それらのサーバのアドレスがリストに加えられる。プログラムが開始するに伴い、追跡されるべきサーバのアドレスが、項目 6 0 2 として表 6 0 0 に入力される (ステップ 5 1 0)。次いで、サーバに関連するカウンタ 6 0 4 がクリアされてゼロになる (ステップ 5 1 5)。プログラムは、A R P キャッシュの内容を作業域にダンプし、キャッシュ内の項目をクリアすることにより開始する (ステップ 5 2 0)。このように、アドレス解決プロトコルは、革新的なプログラムがサーバの活動をチェックしている間、動作し続けることができる。次いで、追跡プログラムが、取得されたばかりの A R P ダンプ内の最初の項目を調べる (ステップ 5 2 5)。そこに格納されたアドレスが、追跡されているサーバの 1 つのアドレスかどうかに関しての判断がなされる (ステップ 5 3 0)。答えが「イエス」である場合には、サーバのカウンタ 6 0 4 は 1 だけ増分され (ステップ 5 3 5)、そうでない場合には、このステップは省かれる。監視プログラムは、次いで、ダンプ内の最後の項目を調べているかどうかを判断する (ステップ 5 4 0)。調べていない場合には、追跡プログラムはダンプ内の次の項目に進み (ステップ 5 4 5)、次いで

40

50

、ステップ530に戻り、キャッシュ・ダンプの残りをチェックする。

【0034】

キャッシュの終わりに到達した場合には、追跡プログラムは、表600で収集した情報を分析することができる。プログラムがARPキャッシュをダンプし、分析したのが初めてであれば、例えば、上述の例における最後の2160秒といった、キャッシュにより設定された期間内でサーバがアクティブであったかどうかを知らせることしかできない。しかしながら、監視プログラムがキャッシュを繰り返しダンプし、クリアするに伴い、全サーバがたいてい又は常にキャッシュ内に現れているかどうかの知識を構築することができる。サーバが、時々ARPキャッシュ・ダンプに現れなくても驚くことはないが、これが頻繁に起こる場合、特にオペレータが利用可能な他のインジケータが、サーバがビジー状態であると示す場合には、一般に、問題の現れである。

10

【0035】

監視プログラムは、種々のサーバのカウンタを比較して、それらの差が閾値よりも大きいかどうかを調べる（ステップ550）。答えが「イエス」である場合には、メッセージが、収集された関連データのコピーと併せて、制御コンソールに送信される（ステップ555）。従って、このプログラム自体が問題を修正することはないが、疑わしい条件が存在することをオペレータに知らせる。サーバ上のロードが閾値よりも大きく異なることがない場合には、メッセージは省かれる。

【0036】

累計は、定期的にログ・ファイルに書き込まれ、クリアされることができる。ロギングの期間は、数分から数時間又は数日まで異なることができる。間隔がどのように決められても、プログラムは、この期間が経過したかどうかをチェックする（ステップ560）。期間が経過していた場合には、ログ記録又はログ報告が書き込まれ、カウンタ604がクリアされる（ステップ565）。ARPキャッシュのダンプ及びクリアに戻る前に、プログラムは、サーバの数に変更されたかどうかを判断する（ステップ570）。一般に、この答えは「ノー」であり、プログラムは再びキャッシュをダンプし、クリアする（ステップ520）。変化する条件が、サーバの数を増加させる又は減少させる原因となった場合には、プログラムはステップ510に戻り、予想されるサーバすべてのアドレスが表600に記録され、次いで、チェックが再び始まる。

20

【0037】

すべてのコンピュータ・アプリケーションにおけるように、このプログラムでは、十分な情報を提供することと、作業の追跡にあまりに時間がかかり、実際の作業を終わらせる時間がないこと、との間にトレードオフがある。監視プログラムは、連続的で、密集したキャッシュのダンプを使用することにより、最も効率よく動作する。しかしながら、アクティブなプロセスにおける信頼性が増加するに伴い、モニタは、例えば、15分毎又は1時間毎に1分ずつ、その動作を省略することができる。このようにして、プログラムは、常時稼動することなく、必要なデータを取得することができる。もちろん、例えば、他のシステム・モニタが潜在的なシステム問題があることを示した場合には、監視プログラムは、オペレータの入力により連続稼動に戻ることができる。

30

【0038】

ここで、サーバ間の不均衡の検出について、図8を参照して説明する。この例では、IPアドレス0000.0c07.bc01をもつ単一のルータが、IPアドレス192.168.12.174、192.168.129.191、192.168.135.167、及び192.168.135.136における、4つの異なるサーバに照会を送信する。以下は、幾つかの一連のキャッシュ・ダンプの一部を表す表、及び、関連する走査の反復後にもたらされる表を表す。このシステムでは、キャッシュは、ゼロから上向きに増分される年数フィールドを使用することに注目されたい。

40

【表 1】

プロトコル	I P アドレス	年 数	MAC	タイプ	インターフェース
インターネット	192.168.12.174	6	0000.8645.4b12	AR PA	Vlan12
インターネット	192.168.143.168	1	0002.55c2.0699	AR PA	Vlan143
インターネット	192.168.129.191	5	0000.864a.2ccd	AR PA	Vlan129
インターネット	192.168.137.29	7	0004.bce4.c226	AR PA	Vlan137
インターネット	192.168.135.167	2	0000.864a.3501	AR PA	Vlan135
インターネット	192.168.142.174	0	0003.47a7.c7ef	AR PA	Vlan142
インターネット	192.168.129.131	9	0001.0384.0268	AR PA	Vlan129
インターネット	192.168.135.136	35	0002.5513.4d84	AR PA	Vlan135
インターネット	192.168.135.172	36	0002.8b2a.b65c	AR PA	Vlan135

10

20

【0039】

キャッシュ・ダンプ、時間 = 0

【表 2】

192.168.12.174	1
192.168.129.191	1
192.168.135.167	1
192.168.135.136	1

30

【0040】

1 回目の反復後の表

【0041】

予想されるサーバのすべてが、この最初の走査において見出された。1 秒後に、ダンプが再び生じる。

【表 3】

プロトコル	I P アドレス	年 数	M A C	タイプ	インターフェース
インターネット	192.168.129.191	0	0000.864a.2ccd	ARP A	Vlan129
インターネット	192.168.143.168	0	0002.55c2.0699	ARP A	Vlan143
インターネット	192.168.135.167	0	0000.864a.3501	ARP A	Vlan135
インターネット	192.168.12.174	0	0000.8645.4b12	ARP A	Vlan12
インターネット	192.168.129.131	0	0001.0384.0268	ARP A	Vlan129
インターネット	192.168.142.174	0	0003.47a7.c7ef	ARP A	Vlan142
インターネット	192.168.137.29	0	0004.bce4.c226	ARP A	Vlan137
インターネット	192.168.135.172	0	0002.8b2a.b65c	ARP A	Vlan135

10

20

【0042】

キャッシュ・ダンプ、時間 = 1

【表 4】

192.168.12.174	2
192.168.129.191	2
192.168.135.167	2
192.168.135.136	1

30

【0043】

2 回目の反復後の表

【0044】

今度は、3つのサーバがダンプ内で探索されたが、‘136サーバは見出されなかった。1秒後に、ダンプが再び生じる。

【表 5】

プロトコル	I P アドレス	年 数	M A C	タイプ	インターフェース
インターネット	192. 168. 129. 191	0	0000. 864a. 2ccd	ARP A	Vlan129
インターネット	192. 168. 12. 174	0	0000. 8645. 4b12	ARP A	Vlan12
インターネット	192. 168. 135. 167	0	0000. 864a. 3501	ARP A	Vlan135
インターネット	192. 168. 136. 146	0	0010. b48e. 3be5	ARP A	Vlan136
インターネット	192. 168. 129. 131	0	0001. 0384. 0268	ARP A	Vlan129
インターネット	192. 168. 142. 174	0	0003. 47a7. c7ef	ARP A	Vlan142
インターネット	192. 168. 137. 29	0	0004. bce4. c226	ARP A	Vlan137
インターネット	192. 168. 135. 172	0	0002. 8b2a. b65c	ARP A	Vlan135

10

20

【 0 0 4 5 】

キャッシュ・ダンプ、時間 = 2

【表 6】

192. 168. 12. 174	3
192. 168. 129. 191	3
192. 168. 135. 167	3
192. 168. 135. 136	1

30

【 0 0 4 6 】

3 回目の反復後の表

【 0 0 4 7 】

ここでも、‘ 1 3 6 サーバがキャッシュ・ダンプ内に見出されなかった。これは、問題の検出とすることができるが、閾値に到達するまでは、そのように認識されない。閾値が 3 5 秒に設定されていると仮定し、時間を飛び越して進めることにする。ここで、3 6 回目の反復におけるキャッシュ・ダンプを示す。

【表 7】

プロトコル	I P アドレス	年 数	M A C	タイ プ	インターフェース
インターネット	192.168.137.29	0	0004.bce4.c226	ARP A	Vlan137
インターネット	192.168.12.174	0	0000.8645.4b12	ARP A	Vlan12
インターネット	192.168.146.17	0	0800.5bfc.9eb5	ARP A	Vlan430
インターネット	192.168.135.167	0	0000.864a.3501	ARP A	Vlan135
インターネット	192.168.129.191	0	0000.864a.2ccd	ARP A	Vlan129
インターネット	192.168.129.131	0	0001.0384.0268	ARP A	Vlan129
インターネット	192.168.142.174	0	0003.47a7.c7ef	ARP A	Vlan142
インターネット	192.168.135.172	0	0002.8b2a.b65c	ARP A	Vlan135

10

20

【0048】

キャッシュ・ダンプ、時間＝35

【表 8】

192.168.12.174	36
192.168.129.191	36
192.168.135.167	36
192.168.135.136	1

30

【0049】

36回目の反復後の表

【0050】

この反復において、サーバ「136」の出現と他のサーバのそれとの間の差が、閾値35に到達する。この時点で、オペレータは明白な問題を知らされる。

【0051】

この方法及びシステムは、サーバ群の活動を監視し、報告して、変更が必要な時にオペレータが対応することを可能にする方法を提供するものであるが、これは、以前は明らかではなかった。

40

【0052】

本発明は、完全に機能するデータ処理システムの内容で説明されているが、重要な点は、当業者であれば、本発明のプロセスは、コンピュータ可読媒体の命令形態及び種々の形態で配信することができ、しかも、本発明は、配信を実行するのに実際に使用される信号担持媒体の特定の種類に関係なく、等しく適用されるものと理解するであろうということである。

【0053】

コンピュータ可読媒体の例は、フレキシブル・ディスク、ハード・ディスク・ドライブ、RAM、CD-ROM、DVD-ROMといった書き込み可能型媒体、並びにデジタル

50

及びアナログ通信リンクといった伝送型媒体、例えば、無線周波伝送及び光波伝送といった伝送形態を用いた有線又は無線通信リンクを含む。コンピュータ可読媒体は、特定のデータ処理システムにおける実際の使用のためにデコードされる、コード化フォーマットの形態をとることができる。

【0054】

本発明の説明は、例示及び説明の目的のために提示されたものであり、網羅的であること又は開示された形態に発明を限定することを意図するものではない。多数の修正及び変形態様が当業者には明白であろう。実施形態は、本発明の原理、実用的な適用例を最も良く説明し、他の当業者が、想定される特定の使用に適した種々の修正をもつ種々の実施形態について本発明を理解することを可能にするように選択され、説明されたものである。

【図面の簡単な説明】

【0055】

【図1】インターネットに接続された、プロビジョニング・システムの図形表示を示す。

【図2】プロビジョニング・システムにより管理されるリソースの一部となることができ、本発明の監視プログラムを実装することができるサーバ群を示す。

【図3】サーバとして実装することができるデータ処理システムのブロック図である。

【図4】例示的なARP要求記録を示す。

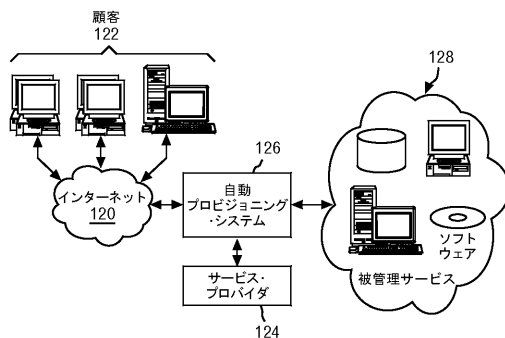
【図5】本発明の例示的な実施形態による例示的なARPキャッシュの一部を示す。

【図6】本発明の例示的な実施形態によるサーバの使用を監視する監視プログラムの動作のフローチャートを示す。

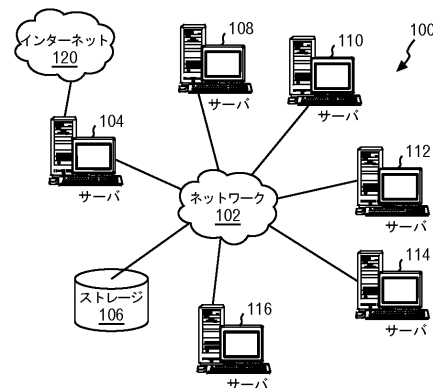
【図7】本発明の例示的な実施形態によるそれぞれのサーバの活動を追跡するのに使用される表を示す。

【図8】問題がどのように検出され得るかを例示する、ネットワーク及び関連するIPアドレスを示す。

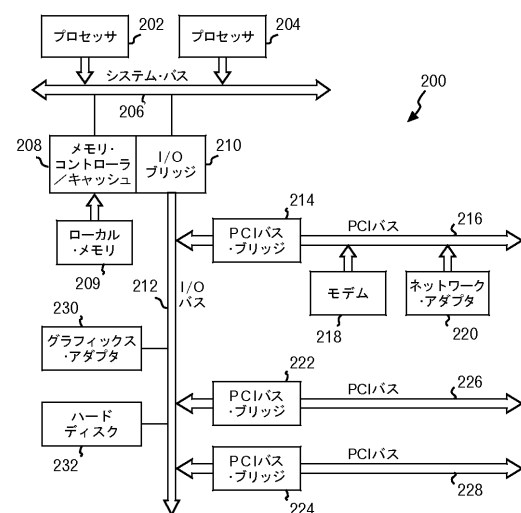
【図1】



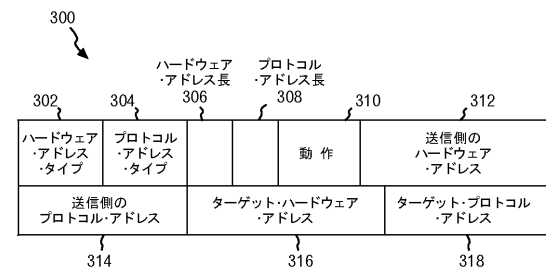
【図2】



【図3】



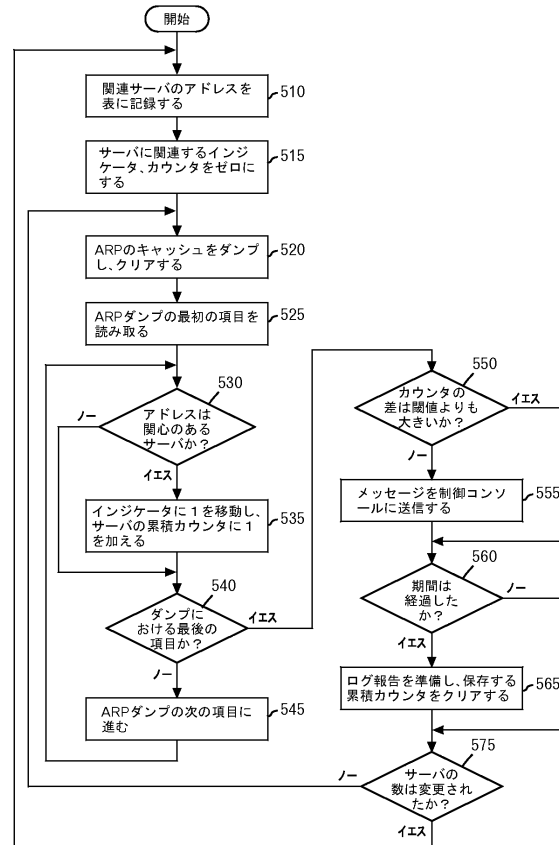
【図4】



【図 5】

410 IPアドレス	415 MACアドレス	420 存続期間
192.168.24.138	00:c0:4f:2f:96:1f	2159
192.168.231.44	00:b0:d0:56:44:fb	2142
192.168.174.3	00:0b:db:d5:80:c6	2081
192.168.174.62	00:00:cb:df:70:2f	2057
192.168.150.237	00:b0:d0:59:b8:2f	2160
192.168.232.42	00:c0:4f:4c:ff:5b	1999
192.168.170.83	00:b0:d0:14:cb:6b	1965
192.168.158.35	00:06:5b:b2:8b:bf	1956
192.168.244.150	00:b0:d0:59:40:b5	1890
192.168.154.23	00:10:b4:b7:06:0c	1700
192.168.244.251	00:b0:d0:58:29:cf	1115
192.168.213.187	00:06:5b:b2:9c:69	977
192.168.49.18	00:c0:4f:4f:d4:62	930
192.168.244.184	00:b0:d0:b7:78:7f	422
⋮	⋮	⋮

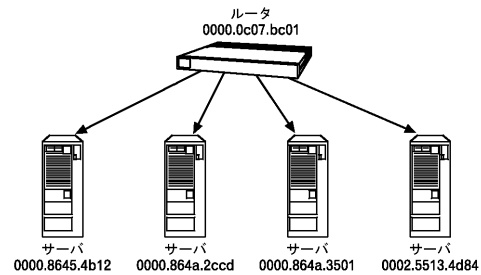
【図 6】



【図 7】

602 サーバ・ アドレス	604 カウンタ
サーバ・ アドレス	カウンタ
⋮	⋮
サーバ・ アドレス	カウンタ

【図 8】



フロントページの続き

(74)代理人 100086243

弁理士 坂口 博

(72)発明者 チルドレス、ロンダ

アメリカ合衆国 78717 テキサス州 オースチン ウェスト・ポトシ・コーブ 8515

(72)発明者 クムール、デービッド、ブルース

アメリカ合衆国 78737 テキサス州 オースチン アパルーサ・ラン 8934

(72)発明者 ペネル、ニール、レイモンド

アメリカ合衆国 78612 テキサス州 シーダー・クリーク コロラド・コーブ 106

審査官 井出 和水

(56)参考文献 特表2002-540492 (JP, A)

特開平11-143804 (JP, A)

米国特許第5664106 (US, A)

国際公開第97/10543 (WO, A1)

特開2003-152729 (JP, A)

特開2002-016622 (JP, A)

特開平10-303961 (JP, A)

特開2002-033747 (JP, A)

(58)調査した分野(Int.Cl., DB名)

G06F 11/30

G06F 11/34

H04L 12/24 - H04L 12/28