

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第5264961号  
(P5264961)

(45) 発行日 平成25年8月14日(2013.8.14)

(24) 登録日 平成25年5月10日(2013.5.10)

(51) Int.Cl. F I  
**G06F 12/00 (2006.01)** G O 6 F 12/00 5 4 6 A  
 G O 6 F 12/00 5 1 3 J

請求項の数 9 (全 22 頁)

(21) 出願番号	特願2011-114971 (P2011-114971)	(73) 特許権者	596004587
(22) 出願日	平成23年5月23日 (2011.5.23)		マサチューセツ インスティテュート オブ テクノロジー
(62) 分割の表示	特願2008-269979 (P2008-269979) の分割		アメリカ合衆国 02139 マサチューセツ州, ケンブリッジ, マサチューセツ アベニュー 77
原出願日	平成20年10月20日 (2008.10.20)	(74) 代理人	100066692
(65) 公開番号	特開2011-204252 (P2011-204252A)		弁理士 浅村 皓
(43) 公開日	平成23年10月13日 (2011.10.13)	(74) 代理人	100072040
審査請求日	平成23年6月22日 (2011.6.22)		弁理士 浅村 肇
(31) 優先権主張番号	60/092, 710	(74) 代理人	100094673
(32) 優先日	平成10年7月14日 (1998.7.14)		弁理士 林 拓三
(33) 優先権主張国	米国 (US)	(74) 代理人	100091339
(31) 優先権主張番号	09/314, 863		弁理士 清水 邦明
(32) 優先日	平成11年5月19日 (1999.5.19)		
(33) 優先権主張国	米国 (US)		

最終頁に続く

(54) 【発明の名称】 組込型コンテンツ分散ゴースト・サーバを用いたグローバル文書ホスティング・システム

(57) 【特許請求の範囲】

【請求項 1】

コンテンツ配信ネットワーク (CDN) に提供することを認められたドメイン名サービス (DNS) によって、

異なるネットワーク位置にある第 1 又は第 2 のクライアント・ローカル名サーバのそれぞれから、同じ仮想サーバ・ホスト名を識別する第 1 及び第 2 のドメイン名クエリーを受信するステップであって、

前記第 1 のドメイン名クエリーは、第 1 のコンテンツ・オブジェクトに対する第 1 の最終使用者からの要求に関し、

前記第 2 のドメイン名クエリーは、第 2 のコンテンツ・オブジェクトに対する第 2 の最終使用者からの要求に関し、

前記第 1 及び第 2 のコンテンツ・オブジェクトは、どちらも同じ仮想サーバ・ホスト名に関し、同じコンテンツ・オブジェクト又は異なるコンテンツ・オブジェクトの一方である前記受信ステップと、

前記仮想サーバ・ホスト名を含む前記第 1 のドメイン名クエリーに回答し、前記第 1 のクライアント・ローカル名サーバのインターネット・プロトコル (IP) アドレスに基づいて、

( i ) コンテンツ・サーバの第 1 の部分集合に関する 1 以上の IP アドレスの第 1 の組を決定するステップと、

( i i ) 前記第 1 のクライアント・ローカル名サーバに、前記 1 以上の IP アドレスの

10

20

第 1 の組を返すステップと、

( i i i ) その後、前記第 1 のコンテンツ・オブジェクトに対する第 1 の最終使用者からの新たな要求を受信するステップと、

( i V ) 前記第 1 の最終使用者からの新たな要求に回答して、前記第 1 の部分集合内のコンテンツ・サーバから前記第 1 のコンテンツ・オブジェクトを返すステップと、

前記仮想サーバ・ホスト名を含む前記第 2 のドメイン名クエリーに回答し、前記第 2 のクライアント・ローカル名サーバの IP アドレスに基づいて、

( i ) コンテンツ・サーバの第 2 の部分集合に関する 1 以上の IP アドレスの第 2 の組を決定するステップと、

( i i ) 前記第 2 のクライアント・ローカル名サーバに、前記 1 以上の IP アドレスの第 2 の組を返すステップと、

( i i i ) その後、前記第 2 のコンテンツ・オブジェクトに対する第 2 の最終使用者からの新たな要求を受信するステップと、

( i V ) 前記第 2 の最終使用者からの新たな要求に回答して、前記第 2 の部分集合内のコンテンツ・サーバから前記第 2 のコンテンツ・オブジェクトを返すステップと

を含むコンテンツを配信する方法であって、

コンテンツ・サーバの前記第 1 及び第 2 の部分集合は異なる、前記方法。

**【請求項 2】**

請求項 1 記載の方法において、

コンテンツ・サーバの前記第 1 又は第 2 の部分集合のそれぞれに関する前記 1 以上の IP アドレスは、前記第 1 又は第 2 のクライアント・ローカル名サーバのそれぞれに近く、過負荷とならずに、前記仮想サーバ・ホスト名に関するコンテンツ・オブジェクトをホストすると思われるコンテンツ・サーバの IP アドレスを含む前記配信方法。

**【請求項 3】**

請求項 1 記載の方法において、前記仮想サーバ・ホスト名は、ジェネリック最上位レベルドメイン名を、前記 DNS を実行するサービス・プロバイダ及びコンテンツ・サーバの前記部分集合をユニークに識別する識別子と共に含む、前記方法。

**【請求項 4】**

請求項 1 記載の方法において、前記 DNS は、コンテンツ・サーバの前記第 1 の部分集合と同じ場所にある第 1 の名サーバと、コンテンツ・サーバの前記第 2 の部分集合と同じ場所にある第 2 の名サーバとを備える、前記方法。

**【請求項 5】**

請求項 4 記載の方法において、前記第 1 及び第 2 の名サーバは、少なくとも第 1 及び第 2 レベルを有する DNS 階層の第 2 レベルにある、前記方法。

**【請求項 6】**

請求項 5 記載の方法において、前記 DNS 階層の前記第 1 レベルは、前記第 1 又は第 2 のドメイン名クエリーの受信に回答して、前記第 2 レベルにある前記第 1 又は第 2 の名サーバの IP アドレスを返す少なくとも一つの他の名サーバを含む、前記方法。

**【請求項 7】**

請求項 1 記載の方法において、前記仮想サーバ・ホスト名は、前記コンテンツ配信ネットワーク ( CDN ) を指し示し、コンテンツ・オブジェクトに関する情報を符号化する値を含む、前記方法。

**【請求項 8】**

請求項 1 記載の方法において、前記コンテンツ・オブジェクトは、マークアップ言語ページ、マークアップ言語ページに組み込まれたオブジェクト、及びメディア・ストリームのうちの一つである、前記方法。

**【請求項 9】**

請求項 1 記載の方法において、前記コンテンツ・オブジェクトは、テキスト、グラフィックス、画像、ダウンロード、オーディオ又はビデオのうちの一つである、前記方法。

**【発明の詳細な説明】**

10

20

30

40

50

## 【技術分野】

## 【0001】

本発明は一般的にコンピュータ・ネットワーク内での情報検索 (retrieval) に関する。更に詳細には、本発明はインターネット・サービス・プロバイダ (ISP) およびインターネット・コンテンツ・プロバイダの問題を解決するための、インターネット上でのコンテンツのホスティングおよび配信の新規な方法に関する。

## 【0002】

(関連技術の説明)

ワールド・ワイド・ウェブ (WWW) はインターネットのマルチメディア情報検索システムである。ウェブ環境の中で、クライアント・マシンはハイパー伝送プロトコル (HTTP: Hypertext Transfer Protocol) を使用してウェブ・サーバへのトランザクションを処理する、これはファイル (例えば、テキスト、グラフィックス、画像、音響、ビデオ等) への使用者のアクセスを提供する既知のアプリケーション・プロトコルであり、ハイパーテキスト・マークアップ言語 (HTML) として知られている標準ページ記述言語を使用する。HTML は基本文書形式を提供し、開発者が別のサーバ及びファイルへの「リンク」を指定することを可能とする。インターネットの範疇において、サーバへのネットワーク経路は、ネットワーク接続を定義する特別シンタックスを有する、いわゆるユニフォーム・リソース・ロケータ (URL: Uniform Resource Locator) で識別される。HTML 互換ブラウザ (例えば、ネットスケープ・ナビゲータ (Netscape Navigator) またはマイクロソフト・インターネット・エクスプローラ (Microsoft Internet Explorer)) をクライアント・マシンで使用することは、URL 経由でのリンクの指定を含む。それに応答してクライアントはリンク内で識別されたサーバへ要求を行い、その返事として文書または HTML に基づいてフォーマットされたその他のオブジェクトを受信する。ウェブ・サーバで保持されている文書の集合をウェブ・サイトと呼ぶ場合もある。

## 【0003】

ウェブ・サーバに関する従来技術ではそのコンテンツを別のサーバへ反映保存 (mirror) することが良く知られている。実際、現在のところコンテンツ・プロバイダがそのコンテンツを読者により近いところに配置するための唯一の方法は、そのウェブ・サイトのコピーを国内的にまた国際的に異なる場所にあるウェブ・ホスティング所に設置されている機械の上に構築することである。ウェブ・サイトのこれらのコピーはミラー・サイトとして知られている。不幸にしてミラー・サイトはコンテンツ・プロバイダに対して不必要な経済的また運営上の重荷を追わせ、それらは規模による経済効果をもたらさない。経済的には、1つの一次サイトと1つのミラー・サイトを具備するコンテンツ・プロバイダの全体コストは一次サイト単一のコストの2倍以上である。この追加コストは2つの要因による: (1) コンテンツ・プロバイダは各々のミラー・サイトに関して別々のホスティング設備を契約しなければならない、そして(2) コンテンツ・プロバイダはミラー・サイトの同期を維持するための追加のオーバーヘッド費用を負担しなければならない。

## 【0004】

ミラーリングに関する問題点を解決する努力の中で、例えばシスコ (Cisco)、レゾネイテ (Resonate)、ブライツ・タイガー (Bright Tiger)、F5 ラブズ (F5 Labs) およびアルテオン (Alteon) の様な企業は、ミラー・サイトの同期と負荷のバランスを保たせるソフトウェア並びにハードウェアを開発している。これらの機械はコンテンツ・プロバイダの助けになるが、内在している規模拡張可能性 (scalability) の問題解決には失敗している。コンテンツ・プロバイダはミラーリングに関するコストを負担する意志はあるが、この技術はそれ自体、数ウェブ (10 未満) の規模を超えない。

## 【0005】

これらの経済的および規模拡張可能性の問題に加えて、ミラーリングはまた運用面での問題も残している。ミラー・サイトを使用するコンテンツ・プロバイダは遠隔地に物理的な空間を賃借りして管理しなければならないだけでなく、そのサイトを同期させかつ負荷バランスを取るためのソフトウェアおよびハードウェアを購入して保守しなければなら

10

20

30

40

50

ない。現行の解決方法では、コンテンツ・プロバイダは複数のウェブ・サイトを保守するために必要な人員、技術およびその他の項目を供給する必要がある。纏めると、ミラーリングはコンテンツ・プロバイダに対してコンテンツを制作するという彼ら本来のビジネスに関係しない機能に対して経済的およびその他の資源を浪費させている。

【 0 0 0 6 】

更に、コンテンツ・プロバイダはまた彼らのコンテンツの管理を維持し続けたいと希望している。今日、ISPの中には、コンテンツ・プロバイダと最終使用者との間のリンクを中断できるキャッシュ・ハードウェアを設置しているところもある。この様なキャッシングの効果は破滅的な結果をコンテンツ・プロバイダにもたらし、それらには(1)コンテンツ・プロバイダが彼らのウェブ・ページへの正確なヒット数を得ることを妨げる(10 それによって広告主からの収入が減少する)、(2)コンテンツ・プロバイダがコンテンツを仕立て上げて特定の視聴者へ宣伝することを妨げる(これはコンテンツ・プロバイダのウェブ・ページの有効性を著しく制限する)、(3)時期遅れの情報をそのクライアントに提供することになる(これは最終使用者をイライラさせ怒らせることになる)というような事が含まれる。

【 0 0 0 7 】

当業分野において、使用者がインターネット・コンテンツを更に効率的に(すなわち不必要にネットワーク資源に負担を掛けることなく)獲得することを可能とし、同様にコンテンツ・プロバイダが彼らのコンテンツに対する管理を維持できるような、集中化されていないホスティング解法を提供する大きな必要性が残されている。20

【 0 0 0 8 】

本発明はこれらおよびその他の従来技術に関連する問題を解決する。

【 0 0 0 9 】

(発明の簡単な概要)

本発明の一般的な目的は、ウェブ・コンテンツを効率的に、効果的にかつ信頼性を持って最終使用者に提供するように設計された有機的で、大きな耐故障性を有する構造基板を形成する、多数の広域的に配備されたインターネット・サーバを含むコンピュータ・ネットワークを提供することである。

【 0 0 1 0 】

本発明の別の更に一般的な目的は、ウェブ基盤コンテンツを配信するための基本的に新規でより良い方法を提供することである。本発明によるアーキテクチャは、好適に集中化された制御を具備しない分散型サーバの大規模ネットワーク上でコンテンツを知的に転送しかつ複写するための方法を提供する。30

【 0 0 1 1 】

本発明の別の目的は、コンテンツを使用者の近くに移動させるネットワーク・アーキテクチャを提供することである。本発明によるアーキテクチャはウェブ・サイトが多数の視聴者の開発を、関連するトラフィックを取り扱うための大きな基盤構造を構築するとを気遣わずに行えるようにしている。

【 0 0 1 2 】

本発明の更に別の目的は、分散型ウェブ・コンテンツ用の故障耐性のあるネットワークを提供することである。このネットワーク・アーキテクチャはより豊かなウェブ・ページの配信を高速化するために使用され、コンテンツ・プロバイダが多数の視聴者に対して信頼性高くまた経済的に、好適に最終使用者の近くに配置されているサーバからサービスを提供することが出来るようにしている。40

【 0 0 1 3 】

本発明の更に別の特徴は、コンテンツ・プロバイダとその最終使用者との直接的な関係に影響を与えることなく、大規模なネットワーク上でコンテンツを配信し管理出来ることである。

【 0 0 1 4 】

本発明の更に別の特徴はインターネット用の分散型拡張可能基盤構造を提供することで50

あり、これはウェブ・コンテンツを配信する負担をコンテンツ・プロバイダから好適に、例えば世界レベルで展開されている数百箇所のホスティング・サーバ・ネットワークにシフトさせる。

【0015】

一般的に本発明はネットワーク・アーキテクチャであり、これは真の全世界的規模でのホスティングを支援する。本発明による枠組み構造はコンテンツ・プロバイダがその最も人気のあるコンテンツを世界中の無制限のポイントに複写可能とする。別の特徴として、或る1つの地理的な場所で複写された実際のコンテンツがその場所での視聴者に特定的に調整される。更に、コンテンツはそれが要求された場所へ、コンテンツ・プロバイダ部分での努力またはオーバーヘッドを伴うことなく自動的に送られる。

10

【0016】

従って本発明の更に一般的な目的は、コンテンツ・プロバイダが彼らのコンテンツの制御を維持することを可能とする、全世界的ホスティング枠組み構造を提供することである。

【0017】

本発明のホスティング枠組み構造は、分散方式で動作する1組のサーバを含む。サービス対象である実際のコンテンツは好適に1組のホスティング・サーバ(時にゴースト・サーバと呼ばれる)上でサポートされる。このコンテンツはHTMLページ・オブジェクトを含み、これは従来通りコンテンツ・プロバイダ・サイトからサービスされる。しかしながら本発明によれば、ウェブ・ページの基本HTMLドキュメントはコンテンツ・プロバイダのサイトからサービスされるが、そのページに対する1つまたは複数の組み込まれたオブジェクトはホスティング・サーバからサービスされ、好適にクライアントのマシンに最も近いホスティング・サーバからサービスされる。基本HTMLドキュメントをコンテンツ・プロバイダのサイトからサービスすることにより、コンテンツ・プロバイダはそのコンテンツに対する管理を維持する。

20

【0018】

指定された組み込みオブジェクトのサービスを行うためにどのホスティング・サーバを使用するかは決定は、ホスティング枠組み構造内のその他の資源によって実施される。特にこの枠組み構造は第2組のサーバ(またはサーバ資源)を含み、これらは最上位レベルドメイン・ネーム・サービス(DNS)を提供するように構成されている。加えて、この枠組み構造はまた第3組のサーバ(またはサーバ資源)を含み、これらは低レベルDNS機能を提供するように構成されている。クライアント・マシンが指定されたウェブ・ページに対するHTTP要求をウェブ・サイトに発すると、先に説明したように基本HTMLドキュメントがウェブ・サイトからサービスされる。そのページに対する組み込みオブジェクトは、好適に最上位および低レベルDNSサーバで識別された特定のホスティング・サーバからサービスされる。使用する適切なホスティング・サーバの位置決めをするために、最上位レベルDNSサーバは、その組み込みオブジェクトに対する要求に応答するための指定された低レベルDNSサーバを識別するために、ネットワーク内での使用者の場所を決定する。続いて最上位レベルDNSサーバはその要求を識別された低レベルDNSサーバへ回送し、これは続いてその要求を、オブジェクトをそのクライアントに返すサービスを行う指定されたホスティング・サーバ用のIPアドレスに解読する。

30

40

【0019】

更に一般的に、いくつかのレベルから構成されたDNSサーバの階層を有することが可能である(場合によっては好ましい)。階層の下に行くに従って、最適な地域に近づく。

【0020】

本発明の更に別の特徴は1つの手段であり、これによってコンテンツはサーバの集合を通して分配および複写する事が可能であり、メモリの使用がいくつかの制約を条件として最適化され、この制約とは要求を満足させるために全てのオブジェクトの十分な数のコピーが存在し、オブジェクトのコピーは分散されていてどのサーバも過負荷とならず、これらのコピーは時間と共に同一サーバ上に配置される傾向にあり、またそれらのコピーはそ

50

れらを要求するクライアントに近い地域に配置されるということである。従って、この枠組み構造の中で動作しているサーバはコンテンツ・データベース内の全てのコピーを保持するわけではない。むしろ、指定されたサーバは最小量のデータのコピーを保持し、全システムが要求された水準のサービスを提供するようにしている。本発明のこの特徴は、全てのものを全ての場所にキャッシュする技法や、またはオブジェクトを予め指定された場所のみにキャッシュする技法に比較して更に効率的なホスティング技法を可能とする。

#### 【0021】

全世界的ホスティングの枠組み構造は、動作の各レベルにおいて故障耐性を有する。特に、最上位レベルDNSサーバは、組み込みオブジェクトに対する要求をサービスするためにクライアントによって使用されるはずの低レベルDNSサーバのリストを返す。同様に、各ホスティング・サーバは好適にバディーサーバを含み、これはその関連するホスティング・サーバが故障状態の際にホスティング責任を引き受けるために使用される。

10

#### 【0022】

本発明によれば、ホスティング・サーバの組の間での負荷バランスは組み込みオブジェクト要求を分配するための新奇な技術によって一部実現されている。特に、各組み込みオブジェクトURLは好適に仮想サーバ・ホスト名をそのURLの中にプリペンド（追加:prepending）することで修正変更される。更に一般的には仮想サーバ・ホスト名がURLの中に挿入される。好適に、この仮想サーバ・ホスト名は1つの値（シリアル番号と呼ばれる場合もある）を含み、この値は指定されたハッシュ関数をURLに適用したり、またはそのオブジェクトに関する情報をその値の中に符号化する事によって生成される。この機能は組み込みオブジェクトを仮想サーバ・ホスト名の指定された組の中にランダムに分散させるように働く。加えて、組み込みオブジェクトに対する指定された指紋値が、指定されたハッシュ関数を組み込みオブジェクト自体に適用して生成される。この指定された値はその組み込みオブジェクトが既に修正されているか否かを識別する指紋として働く。好適にそれらの値（すなわち、仮想サーバ・ホスト名および指紋用に）を生成するために使用される機能は、指定されたウェブ・ページに対してオフライン過程で適用される。従って、そのページに対するHTTP要求が受信されると、その基本HTMLドキュメントがそのウェブ・サイトでサービスされ、またそのページの組み込みオブジェクトのある部分がその要求を発したクライアント・マシンに近い（最も近い必要は無いが）ホスティング・サーバからサービスされる。

20

30

#### 【0023】

以上で本発明の関連する目的並びに機能の概要を説明した。これらの目的は本発明の更に突出した機能およびアプリケーションのいくつかを単に図示するものと考えられるべきである。多くのその他の有益な結果を、開示された発明を異なる方法で適用したり、または本発明を記述されるように修正変更することによって得ることが出来る。従って、その他の目的および本発明の更に完全な理解は、好適な実施例の以下の詳細な説明を参照する事により得られるであろう。

#### 【0024】

本発明およびその特長を更に完全に理解するために、添付図と関連してなされる以下の詳細な説明を参照する。

40

#### 【図面の簡単な説明】

#### 【0025】

【図1】図1はその中で本発明が実現されている代表的なシステムである。

【図2】図2は基本ドキュメントおよび組み込みオブジェクトの組を図示するマークアップ言語ドキュメントの簡略化された表現である。

【図3】図3は本発明に基づく全世界ホスティング・システムの高レベル図表である。

【図4】図4はウェブ・ページを本発明で使用される修正された組み込みオブジェクトURLに処理するための方法を図示する簡略化された流れ図である。

【図5】図5はウェブ・ページに対するHTTP要求に、本発明がいかに応答するかを図示する簡略化された状態図である。

50

## 【 0 0 2 6 】

( 好適な実施例の詳細な説明 )

既知のインターネット・クライアント・サーバ・システムが図 1 に図示されるように実現されている。クライアント・マシン 1 0 がウェブ・サーバ 1 2 にネットワーク 1 4 を介して接続されている。図示する目的で、ネットワーク 1 4 はインターネット、イントラネット、イクストラネットまたは任意のその他の既知のネットワークである。ウェブ・サーバ 1 2 はクライアントでアクセス可能な複数のサーバの 1 つであって、クライアントの 1 つはマシン 1 0 で図示されている。代表的なクライアント・マシンはブラウザ 1 6 を含み、これはネットワークのサーバにアクセスするために使用される既知のソフトウェア・ツールである。ウェブ・サーバは複数のファイル ( 集散的に「ウェブ」サイトと呼ぶ ) をハイパーテキスト・ドキュメントおよびオブジェクトの形式でサポートする。インターネットの世界ではサーバへのネットワーク・パスはいわゆるユニフォーム・リソース・ロケータ ( URL ) で識別される。

代表的なウェブ・サーバ 1 2 はプロセッサ 1 8、オペレーティング・システム 2 0、およびネットスケープ・エンタプライズ・サーバ ( Netscape Enterprise Server ) の様なウェブ・サーバ・プログラム 2 2 を含むコンピュータである。サーバ 1 2 はまた管理および運営用のグラフィカル・ユーザ・インタフェース ( GUI )、およびアプリケーション・プログラミング・インタフェース ( API ) サポートする表示器を含み、このアプリケーション・プログラミング・インタフェースはアプリケーション開発者に対して、その中枢機能の拡張そして / またはカスタマイズをコモン・ゲートウェイ・インタフェース ( CGI ) プログラム、プラグイン、サーブレット ( Servlets )、アクティブ・サーバ・ページ ( Active Server Pages )、サーバ・サイド・インクルード ( SSI: Server Side Include ) 機能または相当機能を含むソフトウェア・プログラムを通して行えるようにしている。

## 【 0 0 2 7 】

代表的ウェブ・クライアントは x86-、PowerPC<sup>(R)</sup> または RISC 型のパーソナル・コンピュータであり、これは IBM(R) OS/2(R) または Microsoft Windows ( 登録商標 ) '95 の様なオペレーティング・システムを含み、これは Netscape Navigator 4.0 ( またはそれ以上 ) の様な、Java ( 登録商標 ) Virtual Machine ( JVM ) を有しアプリケーション・プラグイン ( application plugins ) またはヘルパー・アプリケーション ( helper application ) をサポートするウェブ・ブラウザを含む。クライアントはまたノート型コンピュータ、携帯型コンピュータ装置 ( 例えば PDA )、インターネット・アプライアンス ( Internet appliance ) またはその他のコンピュータ・ネットワークに接続可能な装置であっても構わない。

## 【 0 0 2 8 】

図 2 に示されるように、典型的なウェブ・ページはマークアップ言語 ( 例えば HTML ) 主または基本ドキュメント 2 8、並びに多数の組み込みオブジェクト ( 例えば画像、音声、ビデオなど ) 3 0 を含む。従って、典型的なページでは、2 0 またはそれ以上の組み込み画像またはオブジェクトは非常に一般的である。これらの画像の各々はウェブ内で独立なオブジェクトであり、個別にリトリブ ( または変更する事が認められている ) されるものである。従ってウェブ・クライアントの一般的な行動様式は、基本 HTML ドキュメントをフェッチし、続いて直ちに組み込みオブジェクトをフェッチする事であり、これらは典型的に ( 常にというわけでは無いが ) 同一サーバ上に配置されている。本発明によれば、好適にマークアップ言語基本ドキュメント 2 8 はウェブ・サーバ ( すなわちコンテンツ・プロバイダ・サイト ) からサービスされる一方で、指定された数 ( またはおそらく全て ) の組み込みオブジェクトはその他のサーバからサービスされる。見て分かるように好適に指定された組み込みオブジェクトはクライアント・マシンに近く過負荷では無く、また要求されたファイルの最新版を既に持っていると思われるサーバ ( ウェブ・サーバ自体以外 ) からサービスされる。

## 【 0 0 2 9 】

次に図 3 を参照すると、この動作は本発明のホスティング・システムで実現される。見て分かるように、ホスティング・システム 3 5 は 1 組の広範に配備された複数のサーバ (

10

20

30

40

50

またはサーバ資源)を含み、これは大規模、故障耐性を有する基盤構造を形成し、これはウェブ・コンテンツを効率的、効果的または信頼性高く最終使用者にサービスするように設計されている。これらのサーバは世界的に配備されていても、また任意の希望する地理的地域に広がっていても良い。理解されるように、ホスティング・システムはその様なコンテンツを知的に回送しかつ複製するための分散型アーキテクチャを提供する。この目的のために、全世界的ホスティング・システム35は3つの基本型サーバ(またはサーバ資源):ホスティング・サーバ(場合によってはゴーストと呼ばれる)36、最上位レベルDNSサーバ38、および低レベルDNSサーバ40を含む。図示はされていないが、DNS階層の中で別のレベルも存在するはずである。これに代わって、最上位レベルと低レベルサーバの機能を結合する単一DNSレベルであっても構わない。この図示された実施例では、発明の枠組み構造35はインターネット・サービス・プロバイダ(ISP)で配備されているが、これは本発明の制約ではない。本発明の全世界的ホスティング枠組み構造35を配備する、1つまたは複数のISPは好適に多数のマシンを有し、これらはそれらのネットワーク上のゴースト・サーバ構成要素36および低レベルDNS構成要素40の両方を実行する。これらのマシンはネットワークを通して分散されており;好適にそれらはネットワーク交換ポイント42およびネットワーク・アクセス・ポイント44の周りに集中されているが、これは必要事項では無い。加えて、ISPは好適に最上位レベルDNS38で動作する少数のマシンを有し、これはまたネットワーク上に分散されている。

#### 【0030】

制約を意味しないが、好適に枠組み構造35の中で使用されている指定されたサーバは、プロセッサ、オペレーティング・システム(例えば、Linux,UNIX(登録商標),Windows(登録商標)NT,等)、ウェブ・サーバ・アプリケーション、及び本発明で使用される1組のアプリケーション処理ルーチンを含む。これらの処理ルーチンはプロセッサで実行される1組の命令としてソフトウェアで好便に実現されており、以下に更に詳細に説明する種々の処理工程または方法ステップを実施する。これらのサーバは好適にネットワークの先端(例えば、ポイント・オブ・プレゼンス、またはPOSS)に配置されている。

#### 【0031】

ネットワーク内の何処にホスティング・サーバを設置するかを決定するにはいくつかの要因がある。従って、例えばサーバ位置は好適に需要主導ネットワーク・マップで決定され、これはプロバイダ(例えばISP)がトラフィック要求を監視することを可能とする。トラフィック・パターンを研究する事により、ISPは与えられたトラフィック・プロフィールに対してサーバ位置を最適化する。

#### 【0032】

本発明によれば、指定されたウェブ・ページ(基本HTMLドキュメントおよび1組の組み込みオブジェクトを含む)は分散方式でサービスされる。従って、好適に基本HTMLドキュメントは通常そのページを管理するコンテンツ・プロバイダからサービスされる。組み込みオブジェクトまたはそのいくつかのサブセットは優先的にホスティング・サーバ36からサービスされ、特に最初にそのウェブ・ページに対する要求を發したクライアント・マシンに近い、指定されたホスティング・サーバ36からサービスされる。加えて、ホスティング・サーバ間で好適に負荷がバランスされ、クライアントがそのページを完成するそのオブジェクトを要求した際に、指定された組み込みオブジェクトがそのクライアント近くの指定されたホスティング・サーバから効率的にサービスされるようにしている。

#### 【0033】

この方法でページ・コンテンツをサービスするために、組み込みオブジェクトに関連するURLが修正変更される。良く知られているように、1つのページ内でサービスされる各々の組み込みオブジェクトはそれ自身のURLを有する。典型的にURLは1つのホスト名を有し、これはこのオブジェクトが従来同様に、すなわち本発明を参照することなくサービスされるコンテンツ・プロバイダのサイトを識別する。本発明によれば、組み込みオブジェクトURLは最初に、好適にオフライン状態でそのURLが全世界的ホスティング

10

20

30

40

50



グ・サーバでサービスされるべく調整されるように修正変更される。オブジェクトURLを修正変更するための好適な方法を図示する流れ図が図4に図示されている。

【0034】

処理ルーチンはステップ50で開始され、指定されたページ内の全ての組み込みオブジェクトの処理が完了したか否かが判定される。全て完了している場合は、処理ルーチンは終了する。しかしながらそうで無い場合、処理ルーチンは次の組み込みオブジェクトをステップ52で獲得する。ステップ54で、仮想サーバ・ホスト名が指定された組み込みオブジェクトに対してURLの中にプリペンドされる。この仮想サーバ・ホスト名は1つの値(例えば、数字)を含み、これは例えば指定されたハッシュ関数をそのURLに対して適用して生成される。良く知られているように、ハッシュ関数は任意長のビット・ストリングを入力として取り、固定長のビット・ストリング(ハッシュ値)を出力として生成する。この様な関数は2つの条件を満足する:(1)同一ハッシュ値を生成する2つの異なる入力を見つけることは不可能である、また(2)入力とそのハッシュ値を与えられた場合、同じハッシュ値を具備した異なる入力を見つけることは不可能である。ステップ54で、その組み込みオブジェクトに対するURLが値xx,xxxにハッシュされ、これは続いて仮想サーバ・ホスト名の中に含まれる。このステップはランダムにそのオブジェクトを指定された仮想サーバ・ホスト名に分配する。

10

【0035】

本発明は、先に説明したハッシュ関数を適用することによる仮想サーバ・ホスト名の生成に限定するものではない。代替りのまた好適な実施例として、仮想サーバ・ホスト名が下記のように生成される。仮の見本的なホスト名 a1234.g.akamai tech\_net を考える。1234 という値は、シリアル番号と呼ばれる場合もあり、好適にそのオブジェクトに関する情報、例えばそのサイズ(大きいかまたは小さいか)、その期待される人気、そのオブジェクトが生成された日時、ウェブ・サイトの識別子、オブジェクトの型式(例えば、動画または静止画)、およびおそらくは指定された乱数関数で生成されたいくつかの線数ビットを含む。もちろん、指定されたシリアル番号がその様な情報を全て符号化する必要もまたその様な構成要素の多くを符号化する必要も無い。実際、最も簡単な場合、このシリアル番号は単一整数であっても構わない。とにかく、情報は1つのシリアル番号の中に好適な方法で符号化される。従って、例えば第1ビットはサイズを表すために使用され、第2ビットは人気を表すために使用され、その他のビットの組は日時を表すために使用される等々。ハッシングの例のところでは先に説明したように、シリアル番号はまた負荷バランス用に、また或る型式のトラヒックを或る型式のサーバに向けるためにも使用される。典型的に同一ページ上のほとんどのURLは同一シリアル番号を有して、ページ毎に必要な識別名(DN: distinguish name)アクセス数を最小化する。この要求はオブジェクトが大きくなるに伴って重要さは少なくなる。

20

30

【0036】

従って、本発明によれば仮想サーバ・ホスト名が指定された組み込みオブジェクトに対してURLの中にプリペンドされ、このホスト名は1つの値(またはシリアル番号)を含み、これは指定された関数をそのURLまたはオブジェクトに適用して生成される。この関数はハッシュ関数、符号化関数等である。

40

【0037】

流れ図に戻ると、処理ルーチンは次にステップ56に続き、指定された値をオブジェクトのURLに含める。好適に、指定された値は指定されたハッシュ関数を組み込みオブジェクトに適用して生成される。このステップはそのオブジェクトのユニークな指紋を生成し、これはそのオブジェクトが修正変更されているか否かを判定するために有用である。その後、この処理ルーチンはステップ50に戻って繰り返される。

【0038】

以上の背景の下に、本発明の全世界的ホスティング枠組み構造を特定の例を参照して次に説明する。特に、ボストンのクライアント・マシンの使用者が、通常アトランタで管理されているコンテンツ・プロバイダのウェブ・ページを要求すると仮定する。図示を目的

50

として、コンテンツ・プロバイダがネットワーク内で全世界的ホスティング・アーキテクチャを使用していると仮定する、このネットワークは全世界的、国際的、国内、地域、ローカルまたは私的なものであって構わない。図5はシステムの種々の構成要素とクライアントからの要求が如何にして処理されるかを示している。先に説明したようにこの動作は制約を目的としているものではない。

**【0039】**

ステップ1：ブラウザが要求をプロバイダのウェブ・サイトに送る（項目1）。アトラクタのコンテンツ・プロバイダ・サイトは全世界的ホスティング枠組み構造が実施されていない場合と同様にしてこの要求を受信する。違いは何がプロバイダ・サイトから返却されるかである。通常のページを返却する代わりに、本発明に基づいてこのウェブ・サイトは図4の流れ図に図示されている方法に基づき修正変更された組み込みオブジェクトURLを具備したページを返却する。先に説明したように、これらのURLは以下のように変更される：

**【0040】**

100,000個の仮想ゴースト・サーバが存在すると仮定する、もっともネットワーク上には比較的少数（例えば100）しか物理的には存在しないかも知れないが。これらの仮想ゴースト・サーバまたは仮想ゴーストはホスト名で識別される：ghostxxx.ghosting.com ここでxxxは0から99,999の間の数で置換される。コンテンツ・プロバイダのウェブ・サイトが新たな情報で更新された後、コンテンツ・プロバイダ・サイト上で実行するスクリプトが実行されて、組み込みURLを書き換える。好適に、組み込みURL名は0と99,999の間の数にハッシュされる、もっともこの範囲は本発明を制限するものではない。次に組み込みURLがその数を具備する仮想ゴーストを参照するように切り換えられる。例えば、下記はプロバイダのサイトからの組み込みURLである：

```
<IMG SRC = http://www.provider.com/TECH/images/space.story.gif>
```

**【0041】**

このURLで参照されたオブジェクトに対するシリアル番号が数1467の場合、好適にこのURLは次のように書き換えられる：

```
<IMG SRC = http://ghost1467.ghosting.akamai.com/www.provider.com/TECH/images/space.story.gif>
```

**【0042】**

この方法でシリアル番号を使用することにより、組み込みURLが100,000個の仮想ゴースト・サーバ名に対してほぼ均等に分散される。プロバイダ・サイトは個別の好みに応じて、スクリーン上の種々のオブジェクトを再配置してそのページを個人化出来ることに注意されたい。更に、プロバイダはまた広告をダイナミックに挿入して、どれだけの人が各々の広告を見たかを計数する事が可能である。

**【0043】**

好適な実施例に基づけば、組み込みURLに対して追加の修正変更がなされて、全世界的ホスティング・システムがサービスする情報が陳腐化しないように保証している。先に説明したように、好適に組み込みURLに含まれるデータのハッシュもまた組み込みURL自身の中に挿入されている。すなわち、各々の組み込みURLはそれが指し示すデータの指紋を含んでいる。含まれる情報が変化すると、指紋も変化し、これによって使用者が古いデータを参照する事が防止される。

**【0044】**

第2ハッシュはビットのストリームを入力として取り込み、そのストリームの指紋と呼ばれることもある出力を生成する。指紋の重要な属性は、2つの異なるストリームはほぼ確実に2つの異なる指紋を生成することである。その様なハッシュの例はMD2およびMD3関数であるが、例えば単純なチェックサムの様な更に透過性の良い方法も使用できる。具体化するために、ハッシュの出力を128ビット署名と仮定する。この署名は1つの数字として解釈され、続いて組み込みURLの中に挿入される。例えば、プロバイダ・ウェブ・サイトからの画像space.story.gif内のデータのハッシュが数28765の場合、修正

10

20

30

40

50

変更された組み込みURLは実際下記のように見えるはずである：

<IMG

SRC=http://ghost1467.ghosting.akamai.com/28765/www.provider.com/TECH/images/space.story.gif">

【 0 0 4 5 】

ページが変更される度に、好適に各々の組み込みURLに対するハッシュが再計算され、必要に応じてそのURLが書き換えられる。URLのいずれかのデータが変化する場合、例えば新たなまた異なる画像がspace.story.gifの名前で挿入される場合、そのデータのハッシュは異なるのでURL自身が別のものとなる。この手法は元のページが更新された結果、陳腐化されたデータをこのシステムがサービスすることを防止する。

10

【 0 0 4 6 】

例えば、画像space.story.gifがコンテンツ・プロバイダのサーバ上で更に最新の版に置き換えられたと仮定する。画像のデータが変化するので、URLのハッシュも同様に变化する。従って、新たな組み込みURLは新たな数が指紋に挿入されない限り同一に見える。更新の後にそのページを要求した全ての使用者は、新たな画像を指し示すページを受信する。古い画像は決して参照されることはなく、更に更新された情報の代わりに間違っ

て返却されることは出来ない。

【 0 0 4 7 】

纏めると、好適に2つのハッシュ操作が存在し、コンテンツ・プロバイダのページを修正変更するために行われる。第1に、ハッシングはドメイン名を仮想ゴースト名に変換するためのシリアル番号が選択される処理工程の、1つの構成要素である。理解されるように、この第1変換はクライアントを全世界ホスティング・システムに転送して組み込みURLをリトリブさせるように働く。次に組み込みURLで指し示されたデータのハッシュが計算されそのURLの中に挿入される。この第2変換は陳腐化および時期遅れのコンテンツがゴースト・サーバからサービスされることを防止する。好適に、これら2つの変換はオフラインで実施されるので性能上の隘路となる可能性は無い。

20

【 0 0 4 8 】

一般化すると好適なURL技法は次の通りである。図示されているドメイン www.domainname.com/frontpage.jpg は次のように変換される：

xxx.yy.zzzz.net/aaaa/ www.domainname.com/frontpage.jpg,

30

此处で

xxxx=シリアル番号フィールド

yy =低レベルDNSフィールド

zzzz=最上位レベルDNSフィールド

aaaa=その他の情報（例えば、指紋）フィールド

DNS階層の別のレベルが使用されている場合、別の低レベルDNSフィールド、例えば xxx.y<sub>1</sub>y<sub>2</sub>.zzz.net/aaaa/...が存在するはずである。

【 0 0 4 9 】

ステップ2：初期ページをコンテンツ・プロバイダ・サイトから受信した後、ブラウザはそのページを表示するために組み込みURLをロードする必要がある。これを実施するための第1ステップは、変更されたホスト名を解明するために使用者マシン（または使用者ISPでの）DNSサーバに接触することである、この場合は：

40

ghost1467.ghosting.akamai.com である。理解されるように、本発明の全世界的ホスティング・アーキテクチャはDNSシステムを、その名前がクライアントに近くまたそのページを既に持っていると思われるゴーストの1つに解明されるように取り扱う。これがどのように行われるかを理解するために、以下にクライアントから発せられたDNSクエリー（query）の進行を説明する。

【 0 0 5 0 】

ステップ3：先に説明したように、好適に2つの型式のDNSサーバが本発明のシステム内に存在する：最上位レベルと低レベルである。ghosting.com用の最上位レベルDNS

50

サーバ38は、.comドメインの通常のDNSサーバとは異なる特別の機能を有する。最上位レベルDNSサーバ38は適切な制御処理ルーチンを含み、これはネットワーク内の何処に使用者が配置されているかを決定し、続いてその使用者を近くのakamai.com(すなわち、低レベルDNS)サーバ40に転送するために使用される。.comドメインと同様に、akamai.comは好適にネットワークを通して分散されている多数の最上位レベルDNSサーバ38を、故障耐性を高めるために有している。従って指定された最上位レベルDNSサーバ38は使用者をインターネット(指定された組み込みオブジェクトに対する要求を満足するために使用されるホスティング・サーバ36の集合を有する)内の地域に転送し、一方低レベルDNSサーバ40(識別された地域内の)はその集合内で特定のホスティング・サーバを識別し、ここからオブジェクトが実際にサービスされる。

10

**【0051】**

更に一般的に先に触れたように、DNS処理はいくつかのレベルの処理を含むことが可能であり、その各々はクライアントをより良いゴースト・サーバに向けるように働く。ゴースト・サーバ名はまた更に多くのフィールドを持つことが可能である。例えば、“a123.g.akamaitech.net”を“a123.ghost.akamai.com.”の代わりに使用することが出来る。唯ひとつのDNSレベルのみが使用されている場合、代表URLは“a123.akamai.com.”である。

**【0052】**

その他の技術を使用することも可能であるが、使用者のネットワーク内での場所は、要求を行ったクライアント・マシンのIPアドレスを調べることによって好適に演繹される。この例において、DNSサーバは使用者のマシン上で動作しているが、これは要求事項ではない。使用者が例えばISP DNSサーバを使用している場合、処理ルーチンはその使用者がこのサーバの近く(インターネットの意味で)に配置されていると仮定する。これに代わって、使用者の場所またはIPアドレスを最上位レベルDNSに送られた要求の中に直接符号化することが出来る。ネットワーク内のIPアドレスの物理的位置を決定するために、最上位DNSサーバはネットワーク・マップを構築し、これは続いて関連する位置を識別するために使用される。

20

**【0053】**

従って、例えば最上位DNSにa1234.g.akamatech.netに関する解釈要求が来た際に、最上位DNSは要求者の返却アドレスを調べて、ネットワーク・マップに従い、そのアドレスに基づいて応答を作り出す。この例で、a1234はシリアル番号、gは低レベルDNSを参照するフィールド、そしてakamaitechは最上位レベルDNSを参照する。このネットワーク・マップは好適に全てのインターネット・プロトコル(IP)ブロックのリストを含み、各々のIPブロックに対して、マップはその要求を送信する場所を決定する。このマップは好適にネットワーク条件およびトラヒックに基づいて絶えず更新される。

30

**【0054】**

ネットワーク内の何処で要求が発せられたかを決定した後、最上位DNSサーバはそのDNS要求をネットワーク内でその使用者に近い低レベルDNSサーバに転送する。要求を転送する能力はDNSシステム内の標準機能である。加えて、この転送はそのローカル低レベルDNSが停止している場合、接続されるバックアップ・サーバが存在するような方法で実施できる。

40

**【0055】**

好適に、これらの最上位レベルDNSのghosting.comドメインに関する転送上のTTL(time to live)スタンプは長めに設定されている。これにより使用者のDNSサーバそして/またはISPのDNSサーバでのDNSキャッシングが最上位レベルDNSサーバが過負荷状態となることを防止できる。使用者マシンまたはISPのDNSサーバ内のghosting.akamai.comに対するTTLが経過すると、最上位レベル・サーバに接続され、ローカル低レベルghosting.akamai.comDNSサーバに対する新たな転送が、新たなTTLスタンプと共に返却される。このシステムは現在の集中化されたホスティング解釈の中で行われているものに比較して、桁違いに多量の最上位レベルDNS検索を行わしめるもので

50

無いことに注意されたい。これは、最上位レベル転送のTTLが高く設定されているためであり、従ってほとんど多数の利用者は彼らのローカルDNSで、近くの低レベルghosting.akamai.comDNSサーバに直接導かれるからである。

【0056】

更に、最上位レベルDNSサーバに対する故障耐性は、一般的な.comドメインで行われているのと同様にDNSによって自動的に提供される。低レベルDNSサーバに対する故障耐性は、好適に単一サーバのみの代わりに可能性のある低レベルDNSサーバのリストを返却することで提供される。低レベルDNSサーバの1つが停止していても、未だ利用者はリスト上に載っていて動作している1つに接続することが可能である。

【0057】

故障耐性はまた「オーバフロー制御」機構を介して処理することも可能であり、ここでクライアントはそのオブジェクトにサービスを行う十分な容量を有することが知られている、地域内の低レベルDNSに転送される。この代替手法は特定の地域からの大量の要求が存在するとき、または1つの地域内の容量が縮退している様な状況で非常に有用である。一般的に、クライアントが地域に転送される方法は、どの地域も過負荷とならないという制約の下で、複数のクライアントが経験する総待ち時間を最小とすようになされる。地域容量の制約の下で総待ち時間を最小化することは、好適に最小コスト多重利益フロー・アルゴリズム(min-cost multicommodity flow algorithm)を使用して実現できる。

【0058】

ステップ4：この時点で、利用者は近隣のghosting.comDNSサーバ38のアドレスを有する。利用者のローカルDNSサーバは近隣の低レベルDNSサーバ40に接続し、名前ghost1467.ghosting.akamai.comへの変換を要求する。ローカルDNSサーバは、利用者に近く、過負荷状態で無く、また要求されたデータを既に持っていると考えられる、ネットワーク上のゴースト・サーバ36の1つのIPアドレスを返却する責任がある。

【0059】

仮想ゴースト名を実際のゴーストにマッピングするための基本機構はハッシングである。1つの好適な技術はいわゆる無矛盾ハッシングであり、米国明細書No.09/042,228、1998年3月13日登録、米国明細書No.09/088,825、1998年6月2日、各々の名称は「複数の資源の中で要求を分配するための方法並びに装置」、マサチューセッツ工科大学所有、に記述されており、この明細書は此処に挙げることで参照されている。無矛盾ハッシュ関数はマシン故障およびクラッシュの際にシステムを強固なものとする。これはまたシステムをしなやかに、ほとんどの項目の配置を変更することなく、またこのシステムに関する完全な情報を持つことなく成長させることを可能とする。

【0060】

本発明によれば、仮想ゴースト名は対照表を用いて実ゴースト・アドレスにハッシュされ、ここで対照表はネットワーク状態およびトラフィックに基づいて、負荷バランスと故障耐性が保証されるように、断続的に更新される。好適に、解読の表は各々のシリアル番号に対して生成される。例えば、シリアル番号1はゴースト2および5に解読される、シリアル番号2はゴースト3に解読され、シリアル番号3はゴースト2, 3, 4に解読される等々である。目的はどのゴーストもその容量を超えない解を定義し、全ての解の中で全てのゴーストの総数が最小化されるようにする事である。これはシステムが各々の地域で利用できるメモリの最大利益を確実に享受出来るようになされる。この点は既存の負荷バランス手法に較べて大きな特長であり、従来手法は全てのものを全ての場所にキャッシュしたり、または負荷に関係なく或る場所の或るオブジェクトのみをキャッシュする傾向にある。一般的に、負荷が短い時間の間に余り大きく変化しない場合は、解が一貫して時間と共に変化しない様に割り当てられることが望ましい。この手法はまた好適に、ゴーストがどれだけ利用者に近い、またその時点でそのゴーストにどれだけ重い負荷が掛かっているかも考慮している。

【0061】

同一の仮想ゴーストは好適に、その利用者がネットワーク内の何処に配置されているか

10

20

30

40

50

によって異なる実ゴースト・アドレスに変換されることも注意されたい。例えば、ゴースト・サーバ18.98.0.17が米国内に配置されており、またゴースト・サーバ132.68.1.28はイスラエル国内に配置されていると仮定する。ボストンから発せられた ghost1487.ghosting.akamai.com に対するDNS要求は、18.98.0.17に解読され、一方テル・アビブから発せられた要求は132.68.1.28に解読されるはずである。

#### 【0062】

2つのレベルのDNSサーバは種々のゴースト・サーバを監視してそれらの負荷を考慮に入れる一方、仮想ゴースト名を実アドレスに変換する。これはゴースト上および低レベルDNSサーバ上で動作するソフトウェア処理ルーチンで取り扱われる。1つ実施例において、負荷情報は1つの地域内のサーバ間で循環され、それらが各々のシリアル番号に対して解を計算できるようにしている。解を計算するための1つのアルゴリズムは下記のように働く。サーバは最初に各々のシリアル番号に対して計画された負荷（使用者要求の数に基づいて）を計算する。次にこのシリアル番号は負荷が増える順に処理される。各々のシリアル番号に対して、要求されるサーバのランダムな優先順位リストが無矛盾ハッシング法を用いて割り当てられる。次に各々のシリアル番号が優先順位リストからサーバの最小初期セグメントに解かれ、どのサーバも過負荷とならないようにされる。例えば、シリアル番号に対する優先順位リストが2, 5, 3, 1, 6の場合、最初にそのシリアル番号に対する負荷をゴースト2にマッピングすることを試みるように試行される。これがゴースト2を過負荷にする場合、その負荷はゴースト2および5の両方に割り当てられる。これによってこれらのサーバのいずれかに過大な負荷が掛かる場合、その負荷はゴースト2, 3, および5といった具合に割り当てられる。1つのサーバ上での計画された負荷はそのサーバを含む全ての解を調べ、またそのシリアル番号からそのサーバに送られると予想される負荷の量を加えることで計算出来る。この解を生成する方法は、繰り返し形式で使用される際に最も効果的であり、そこでは割り当ては全てのシリアル番号が全てのゴーストにマッピングされているデフォルト状態から開始される。先の処理手順に基づいて解テーブルを洗練することにより、負荷は最小の繰り返し量を使用してバランスされる（これにより1つの地域内の利用可能なメモリが最大限節約される）。

#### 【0063】

これら低レベルDNS変換に関するTTLは短く設定されていて、ゴーストの1つに重い負荷が検出された際に迅速に応答出来るようにしている。このTTLは1つのパラメータであって、ゴースト上の高負荷に対するタイムリーな応答と低レベルDNSサーバ上で誘発された負荷との間のバランスを保つために、システムで調整することが出来る。しかしながら、低レベルDNS変換に対するTTLが1-2分に設定されていたとしても、実際に低レベルDNS対照をしなければならない使用者はほとんど居ないことに注意されたい。ほとんどの使用者は彼らのマシンまたは彼らのISP上にキャッシュされているDNS変換を見るはずである。従って、ほとんどの使用者は彼らのローカルDNSサーバから直接、彼らが望むデータを有する近隣のゴーストへ行く。実際に低レベルDNS対照を行うそれらの使用者には非常に短い待ち時間が追加される、しかしながらこの待ち時間はほとんどのデータを近隣からリトリブする長所を鑑みると短い。

#### 【0064】

先に示したように、低レベルDNSサーバに対する故障耐性は、最上位レベルDNSに対して単一のサーバ・アドレスの代わりに、可能な複数の低レベルDNSサーバのリストを返却することにより提供される。使用者のDNSシステムはこのリストをキャッシュ（標準DNSシステムの一部）、リスト上の最初の1つが何らかの理由で停止している場合にはリスト上の別のサーバの1つに接続する。低レベルDNSサーバはDNSの標準機能を使用して、ゴースト・サーバに対して特別水準の故障耐性を提供する。名前が変換される際に、単一名を返却する代わりに、名前のリストが返却される。何らかの理由でゴーストに対する一次故障耐性方法（バディー・システム（Buddy system）として知られていて、後ほど説明する）が故障している場合、クライアントのブラウザはリスト上のその他のゴーストの1つに接続するはずである。

10

20

30

40

50

## 【 0 0 6 5 】

ステップ 5 : 次にブラウザはa123.ghosting.akamai.com/.../www.provider.com/TECH/images/space.story.gif と名付けられたオブジェクトに対して近隣のゴーストから要求を行う。元のサーバの名前 (www.provider.com) が好適に URL の一部として含まれていることに注意されたい。ゴースト上で動作するソフトウェアはページ名を元のホスト名および実ページ名に分解する。ファイルのコピーが既にゴースト上に格納されている場合は、そのデータは直ちに戻される。しかしながら、データのコピーがゴースト上に存在しない場合は、コピーが元サーバまたは別のゴースト・サーバからリトリブされる。ゴーストは元サーバが誰であるかを、その名前がブラウザからそのゴーストに送られてきた URL の中に符号化されているので知っていることに注意されたい。1 度コピーがリトリブされると、これは使用者に戻され、また好適にそれはまたゴースト上に要求に対して返答するために格納される。

10

## 【 0 0 6 6 】

追加の保証として、その使用者がそのサーバに本当に近いかを確認することが好適であろう。これはそのファイルに対する要求に応答する前に、クライアントの IP アドレスを調べて行われる。これはクライアントの DNS サーバがクライアントから遠く離れている様な希な場合に有用である。その様な場合、ゴースト・サーバはその使用者をより近いサーバ (または、そのクライアントにより近いサーバと解釈されるに違いない別の仮想アドレス) に転送することが出来る。その転送が仮想サーバに対する場合、更に別の転送が行われるのを防止するために、目印しが付けられなければならない。好適な実施例では、転送は大きなオブジェクトに対してのみ実施される : 従って、要求されたオブジェクトが或る総サイズを超えているかを確認するための確認が、転送を行う前になされる。

20

## 【 0 0 6 7 】

長いダウンロード性能は、変化するネットワーク条件に基づいて、クライアントが接続されているサーバをダイナミックに変更することにより改善できる。これはオーディオおよびビデオをダウンロードする際 (ここでは接続は長くなりがちで品質が特に重要である) に特に有用である。この様な場合、使用者は流れ途中の別のサーバに向けられることが可能である。クライアントの転送を行うための制御構造は先に説明したものと類似しているが、クライアントのブラウザまたはメディア再生機の中に設置されるソフトウェアを含むことも出来る。このソフトウェアはクライアントの接続の性能およびおそらくはネットワークの状態も同様に監視する。クライアントの接続がサーバを変更することにより改善出来ると思われる場合は、システムはクライアントをその接続の残り部分に関して新たなサーバに向ける。

30

## 【 0 0 6 8 】

ゴーストに対する故障耐性は、相棒 (バディー) ・システムで提供され、此処では各々のゴーストが指定されたバディー・ゴーストを有する。1 つのゴーストが停止すると、その相棒 (バディー) がその仕事 (および IP アドレス) を引継いで、サービスが中断されないようにする。このシステムの別の特徴はバディー・ゴーストが故障に対して待機時間を持たないことである。代わりに、全てのマシンは常に活動状態であり、故障が発生すると、その負荷はバディーによって肩代わりされ、低レベル DNS システムにより別の動作中ゴーストとバランスが取られる。バディー・システムの別の特徴は、故障耐性が長い時間待つことなく提供されることである。

40

## 【 0 0 6 9 】

グローバル・ホスティング・システムの更に別の安全機能として、或るオブジェクトに対する総トラフィックを定められた制限内に保持するためにゲート機構が使用できる。ゲート機構の 1 つの実施例は次のように動作する。1 つのオブジェクトに対する要求の数が、或る指定された閾値を超えると、そのサーバはそのオブジェクトをサービスしないように選択できる。これはそのオブジェクトが非常に大きな場合、非常に有効であろう。代わりに、そのクライアントはそのクライアントに後ほど再接続するように依頼するもっと小さなオブジェクトでサービスを受けることが出来る。または、そのクライアントを転送する

50

事も可能である。ゲートを実現するための別の方法は、そのクライアントに対して「切符」を提供することであり、これはそのクライアントが予め定められた時間後にそのオブジェクトを受信できるようにするものである。この方法において、ゴースト・サーバはそのオブジェクトのサービスを行う前にその切符上の時刻を確認する必要がある。

【0070】

本発明のグローバル・ホスティング・システム手法はグローバルISPまたは地域ISPの集成体の方法であって、ホスティング売り上げを生み出すために彼らのネットワーク基盤構造を強化し、ネットワーク帯域幅を節約するためのものである。本発明のグローバル・ホスティング手法を提供するISPはコンテンツ・プロバイダに対してコンテンツをかれらの使用者に対してISPネットワーク上の最も近いポイントから配信する機能を提供し、従って高速で信頼性の高いアクセスを保証する。保証されたウェブ・サイト性能は全てのウェブを使用するビジネスに重要であり、グローバル・ホスティングはこの要求を満足するサービスを作り出すことを可能とする。

10

【0071】

本発明に基づくグローバル・ホスティングはまたISPがコンテンツがネットワークをどの様にまた何処を移動するかを制御することを可能とする。グローバル・ホスティング・サーバはISPネットワークの端点（例えば多数のネットワーク交換機およびアクセス・ポイント）に設定することが出来る。これはISPがそれが受け持っているサイトに対して、コンテンツを直接ネットワーク交換ポイントおよびアクセス・ポイントの中にサービスすることを可能とする。高価な基幹リンクがコンテンツ・プロバイダのサイトからネットワーク交換機およびアクセス・ポイントへ、冗長なトラフィックを搬送する必要はもはや無くなる。代わりに、コンテンツはISPのネットワークから、高価なネットワーク資源をその他のトラフィックに開放して、直接サービスされる。

20

【0072】

グローバル・ホスティングはネットワーク・トラフィックを低減するが、これはまたグローバルISPが急速に拡大しているホスティング市場の一部を獲得する手段であり、この市場拡大は現在毎年10億ドルと見込まれている。

【0073】

グローバル・ホスティング解法はまた多くの利益をコンテンツ・プロバイダにもたらし、特に彼らのウェブ・サイトの性能を国内および国際的に改善する効果的かつ経済的な解決策を提供する。本発明によるホスティング・ソフトウェアはコンテンツ・プロバイダに対して高速でかつ信頼性の高いインターネット・アクセスを、コンテンツを彼らの加入者に対してISPネットワーク上の最も近いポイントから配信する手段を提供することにより保証する。加えて、更に別の利点が更に詳細に以下に説明されており、グローバル・ホスティング解法はまたネットワーク・トラフィックを削減する重要な利点をも提供する。

30

【0074】

ひとたびグローバル・ホスティング・サーバがISPネットワークの周辺機器（すなわち、多数のネットワーク交換機およびアクセス・ポイント）にインストールされると、コンテンツはネットワーク交換機およびアクセス・ポイントの中に直接サービスされる。コンテンツを直接ISPのネットワークから効率的に配信する結果、本発明はウェブ・サイト性能を事実上改善する。現行のコンテンツ配信システムと比較して、本発明のグローバル・ホスティング解法は高価な基幹リンクを使用して冗長なトラフィックをコンテンツ・プロバイダのウェブ・サイトからネットワーク交換機およびアクセス・ポイントへ搬送する必要が無い。

40

【0075】

本発明のグローバル・ホスティング手法で提供される特有な特長を以下に示す：

【0076】

1. コンテンツ・プロバイダの運営費削減：

ほとんどの競合する解法はコンテンツ・プロバイダに対して彼らのコンテンツを担当す

50



る各々のウェブ・サイトにサーバを購入することを要求する。その結果、コンテンツ・プロバイダは世界中で異なるISPと個別の契約を結ばなければならない。加えて、コンテンツ・プロバイダは一般的にこれらの遠隔場所でコンテンツの複写とサーバの保守の責任を負う。

#### 【0077】

本発明では、ISPがグローバル・ホスティングの機能のほとんどの一次責任を負う。コンテンツ・プロバイダは好適に彼らの単一サーバのみを保守する。このサーバ上のコンテンツはソフトウェアによりそれらがアクセスされる場所へ自動的に複写される。プロバイダ（またはそれについてはISP）による介入または計画は不要である。コンテンツ・プロバイダにはグローバル・ネットワーク上の全てのサーバに対する瞬時のアクセスが提供される；コンテンツが複写されるべき場所の選択または遠隔地にサーバを追加購入する必要は無い。

10

#### 【0078】

##### 2. インテリジェントかつ効率的データ複写：

ほとんどの競合する解法はコンテンツ・プロバイダに対して彼らのコンテンツを1つの商業的ホスティング・サイトのサーバ上に複写するか、または彼らのコンテンツを地理的に遠距離にあるサーバ上にミラー化するように要求する。どちらの方法も効率的ではない。前者の解法では、コンテンツは未だインターネット上の単一場所に配置されている（従って、ほとんどの使用者から遠く離れている）。後者の場合には、ウェブ・サイトの全てのコンテンツが遠隔サーバにコピーされるが、コンテンツの小さな部分であっても実際、遠隔地に配置される必要がある。たとえ安価なメモリを使用しても、この様なミラーリングに関連する追加コストは複数のサイトにミラー化するには膨大となり、これはほとんどの使用者が未だミラー・サイトから遠く離れていることを意味する。ミラーリングはまたコンテンツ・プロバイダに対して、全てのサイトが整合して最新であることを保証しなければならないという別の欠点を有し、これはたとえ数サイトであっても軽視出来ない仕事である。

20

#### 【0079】

本発明では、コンテンツは自動的にグローバル・サーバ・ネットワークに、インテリジェントかつ効率的なやり方で複写される。コンテンツはそれらが必要とされる場所にものみ複写される。更に、コンテンツが変更されると、新たなコピーが好適にネットワークを通して自動的に複写される。

30

#### 【0080】

##### 3. 自動的コンテンツ管理：

多くの既存の解法は、コンテンツ配信、コンテンツ複写および異なるサーバ間での負荷バランスに対して能動的な管理を必要とする。特に、コンテンツを何処で担当するかの決定は手動で行われなければならない、またデータを複写する処理工程は、中央化されたバッチ方式で処理される。これに対して、本発明の特徴は受動管理である。複写は要求ベースのプル方式で、コンテンツが本当にそれを必要としている所にものみ送られるように、好適になされる。更に、処理工程は好適に全自動化されている；ISPはコンテンツが如何にしてまた何処に複写されるべきかを悩む必要は無い、そして/またはコンテンツ・プロバイダも同様である。

40

#### 【0081】

##### 4. 無制限な、経済的拡張性：

競合する解法は拡張性が低く、数個のサイトまでである。例えば、ミラーリングに基づく解法は典型的に最大3または4サイトを接続して使用される。規模拡大の障壁には、全サイトを複写する費用、全てのノードでの計算資源を複写するコスト、およびコンテンツ・プロバイダが彼らのサーバ上で使用する多数の種類のソフトウェア・パッケージを管理する複雑さが含まれる。

#### 【0082】

本発明のユニークなシステム・アーキテクチャは数百、数千または百万個のノードまで

50

拡張可能である。ホスティング・ネットワーク内のサーバは異常状態になったり故障する可能性があるが、システムの全体機能は影響されない。グローバル・ホスティング枠組み構造は資源の有効利用を行う；サーバおよびクライアント・ソフトウェアは全てのノードで複写される必要は無いが、それはホスティング・サーバのみが各ノードで動作しているからである。加えて、グローバル・ホスティング・サーバは高い信頼性を要求されない、標準の簡単なハードウェア上で動作するように設計されている。

【 0 0 8 3 】

5 . 突然の負荷集中に対する防護 :

競合する解法はコンテンツ・プロバイダに対して、予期していない突然の負荷集中に対する防護策を提供していない。ミラーリングおよび関連する負荷バランス解法は、コンテンツ・プロバイダが負荷をサーバの集合の中で分配することを可能とするが、サーバの集合容量はピーク時要求を処理するのに十分でなければならない。これはコンテンツ・プロバイダが実際の平均負荷の代わりに予想されるピーク負荷に見合ったレベルの資源を購入し保守しなければならないことを意味する。インターネットの高度に変化が激しくまたは予想できない性質を考えると、この様な解法は高価で資源の無駄遣いである。

10

【 0 0 8 4 】

本発明のホスティング・アーキテクチャは、ISPがコンテンツ・プロバイダに対して突然の負荷集中保証を提供するためにホスティング・サーバの単一ネットワークを利用することを可能とする。すなわち、プロバイダ・サイト上での予期せぬ高い負荷に対して自動的に適合しサポートする保証である。ISPは多くのプロバイダを同一グローバル・ネットワーク上で集合しているので、資源は更に効率的に使用される。

20

【 0 0 8 5 】

6 . 実質的帯域幅節約 :

競合する解法は、ISPまたはコンテンツ・プロバイダに対して実質的な帯域幅節約は提供しない。ミラーリングの使用を通して、あるリンク(すなわち、ニューヨークとロスアンジェルス間)に関して帯域幅を節約することが可能である。しかしながら、グローバル・ホスティング無しでは、コンテンツに対するほとんどの要求はインターネットを通過する必要があり、従って帯域幅コストを増加させる。本発明のホスティング枠組み構造は、それら自身の基幹を有するISPに対する事実上の基幹帯域幅を節約する。コンテンツはネットワークを通して配信され、ネットワーク交換ポイントの隣に配置することができるので、ISPおよびコンテンツ・プロバイダは共に基幹料金がほとんどのコンテンツ要求に課せられないため、事実上の節約を経験する。

30

【 0 0 8 6 】

7 . グローバル・ネットワークへの瞬時アクセス :

競合する解法はコンテンツ・プロバイダに対して、コンテンツを担当、そして/または複写するサイトの小さな集合を手動で選択するように要求する。たとえISPが広範に広がっている場所に多数のホスティング・サイトを有していたとしても、特別に選択された(および対価が支払われている)これらのサイトのみが、そのコンテンツ・プロバイダに対するコンテンツを担当するために使用される。

【 0 0 8 7 】

これに対して、本発明のグローバル・ホスティング解法は、ISPが彼らのクライアントに対してサーバのグローバル・ネットワークへの瞬時のアクセスを提供することを可能とする。グローバル・ネットワークへの瞬時のアクセスを提供するために、コンテンツは好適に絶え間なくまたダイナミックにネットワーク内を移動される。例えば、コンテンツ・プロバイダがアジア地区の顧客に興味のあるコンテンツを追加した場合、コンテンツ・プロバイダはそのコンテンツがアジア地区に配置されているサーバに自動的に移動されることを保証される。加えて、グローバル・ホスティング枠組み構造は、コンテンツが最終使用者の非常に近くに(企業市場の場合使用者の建物の様な近くに)移動されることを可能とする。

40

【 0 0 8 8 】

50

#### 8. グローバルISPおよび集成体への設計：

ほとんどの競合する解法はコンテンツ・プロバイダで購入され管理されるように設計されており、彼らの多くは既に単一サーバを管理ための統括管理および運用作業に対して絶え間なく挑戦し消耗している。本発明のホスティング手法はグローバルISPで配備され、コンテンツ・プロバイダに提供出来る新たなサービスを提供する。このサービスの特徴は、コンテンツ・プロバイダに要求される運営および管理を最小とし、従ってコンテンツ・プロバイダがユニークなコンテンツを制作するというその中枢ビジネスに集中出来るようにする。

#### 【0089】

##### 9. 固有データベースおよび秘密情報の効果的管理：

多くの競合する解法は、コンテンツ・プロバイダに対して彼ら固有のデータベースを地理的に遠く離れた複数のサイトへ複写することを要求する。その結果、コンテンツ・プロバイダはその固有のまた通常秘密であるデータベースに対する効果的な管理を失ってしまう。これらの問題を取り除くために、本発明のグローバル・ホスティング解法はコンテンツ・プロバイダが彼らのデータベースに対する完全な管理を維持することを保証する。先に説明したように、コンテンツに対する最初の要求はコンテンツ・プロバイダの中央ウェブ・サイトに向けられ、これは続いて効果的かつ管理されたデータベース・アクセスを実施する。好適に、ページ要求に対する高帯域幅、固定部分はグローバル・ホスティング・ネットワークからリトリブされる。

#### 【0090】

##### 10. コンテンツ・プロバイダのソフトウェアとの互換性：

多くの競合する解法はコンテンツ・プロバイダに対して特定セットのサーバとデータベースを使用することを要求する。これらの特に、一定しない要求はコンテンツ・プロバイダが最も効率的に最新技術を使用する能力に制約を与え、コンテンツ・プロバイダの既存基盤構造をお金を掛けて変更する必要を迫る場合もある。これらの問題を取り除くことにより、本発明のグローバル・ホスティング・アーキテクチャは効果的にコンテンツ・プロバイダとISPとの間のインタフェースを行い、これはコンテンツ・プロバイダで使用されているシステムまたはサーバに対して何らの前提も必要としない。更に、コンテンツ・プロバイダのシステムは本発明のアーキテクチャを修正変更したりまたは中断することなく、更新、変更または完全な置き換えを行うことが可能である。

#### 【0091】

11. ダイナミック・コンテンツ、個人化された広告またはEコマース、と陳腐化されていないコンテンツとの間の相互干渉の排除：

多くの競合する解法（例えば、全てのコンテンツの単純なキャッシング）はダイナミック・コンテンツ、個人化された広告およびEコマースと相互干渉し、使用者に対して陳腐なコンテンツをサービスする可能性がある。その他のソフトウェア会社はこれらの問題を部分的に取り除く試みを続けているが（例えば、全てのキャッシュされたコピーに対するヒット計数の保存）、これらの解法の各々は機能の部分的または完全に失わせかねない（例えば、広告を個人化させる機能）。これに対して、グローバル・ホスティング解法はダイナミック・コンテンツ、個人化された広告またはEコマースと相互干渉することは無く、それはこれらの作業の各々が好適にコンテンツ・プロバイダの中央サーバで処理されるためである。

#### 【0092】

##### 12. グローバル・ネットワーク対応設計：

グローバル・ホスティング・アーキテクチャは高度な拡張性を有し、従って全世界規模ネットワーク基盤上に配備できる。

#### 【0093】

グローバル・ホスティング・アーキテクチャの各々の構成要素に関して先に説明した機能は、好適にプロセッサの中で実行可能なソフトウェアで実現される、すなわち、コンピュータの随意アクセスメモリ内に存在するコード・モジュール内の命令またはプログラム

10

20

30

40

50

・コードの組で実現される。コンピュータから要求されるまで、命令の組は別のコンピュータ・メモリ、例えばハードディスク・ドライブの中、または光磁気ディスク（CD ROM内での使用される場合）またはフロッピー（登録商標）・ディスク（フロッピー（登録商標）・ディスク・ドライブ内で使用される場合）の様な携帯型メモリの中に格納されていたり、またはインターネットまたはその他のコンピュータ・ネットワークを經由してダウンロードされる。

【0094】

加えて、説明された種々の方法は汎用コンピュータの中で選択的に起動したりソフトウェアで再構成することで簡便に実現できるが、通常の技術を有する当業者は、その様な方法がハードウェア、ファームウェア、または更に要求された方法のステップを実行するように構築された特殊装置で実現できることを理解するであろう。

10

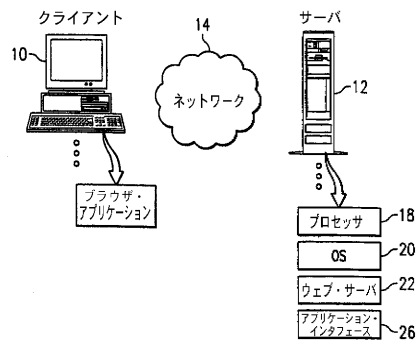
【0095】

更に、ここで使用されるように、ウェブ「クライアント」は広く解釈されるべきであり、インターネットのような、全ての既知のまたは今後開発されるであろうコンピュータ・ネットワークに直接または間接的に接続されていたりまたは接続可能な、全てのコンピュータまたは構成要素を意味する。用語ウェブ「サーバ」は広く解釈されるべきであり、コンピュータ、コンピュータ・プラットフォーム、コンピュータまたはプラットフォームの付属物、またはそれらの全ての構成要素を意味する。もちろん「クライアント」は広く解釈されるべきであり、ファイルを要求または獲得する人を意味し、また「サーバ」はファイルをダウンロードするエンティティーである。

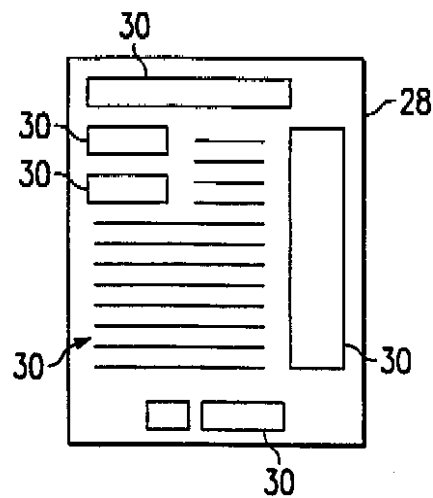
20

此処まで我々の発明を説明してきたが、我々が新たに請求し要望する特許内容は添付の特許請求の項に記載されている。

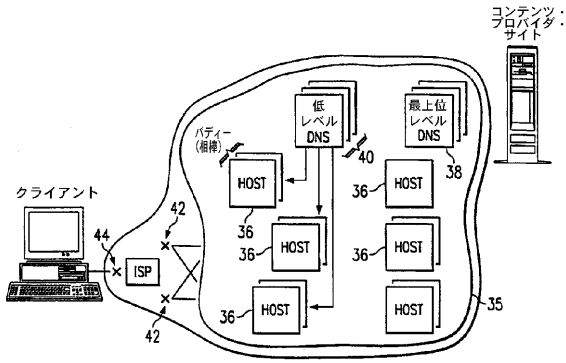
【図1】



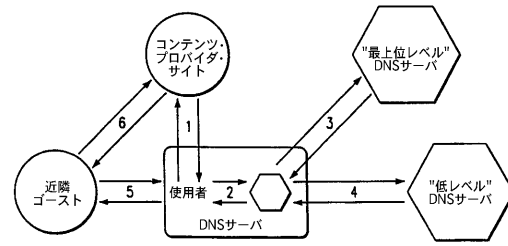
【図2】



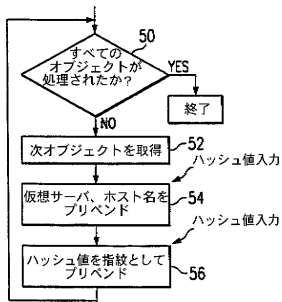
【図3】



【図5】



【図4】



---

フロントページの続き

- (72)発明者 レイトン、エフ、トムソン  
アメリカ合衆国 マサチューセッツ、ニュートンビル、チャールズデン パーク 15
- (72)発明者 ルイン、ダニエル、エム  
アメリカ合衆国 マサチューセッツ、ケンブリッジ、バッサー ストリート 292、ナンバー  
ディ6

審査官 田川 泰宏

- (56)参考文献 特開平08-044643(JP,A)  
国際公開第97/029423(WO,A1)

- (58)調査した分野(Int.Cl., DB名)  
G06F 12/00