

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4515319号
(P4515319)

(45) 発行日 平成22年7月28日(2010.7.28)

(24) 登録日 平成22年5月21日(2010.5.21)

(51) Int.Cl.

F 1

G06F 13/00 (2006.01)
G06F 9/50 (2006.01)G06F 13/00 357Z
G06F 9/46 465Z

請求項の数 10 (全 29 頁)

(21) 出願番号 特願2005-129182 (P2005-129182)
 (22) 出願日 平成17年4月27日 (2005.4.27)
 (65) 公開番号 特開2006-309383 (P2006-309383A)
 (43) 公開日 平成18年11月9日 (2006.11.9)
 審査請求日 平成20年2月28日 (2008.2.28)

(出願人による申告) 平成16年度、独立行政法人情報通信研究機構「テラビット級スーパーネットワークの研究開発」委託研究、産業再生法第30条の適用を受ける特許出願

(73) 特許権者 000005108
 株式会社日立製作所
 東京都千代田区丸の内一丁目6番6号
 (74) 代理人 100075513
 弁理士 後藤 政喜
 (74) 代理人 100084537
 弁理士 松田 嘉夫
 (74) 代理人 100114236
 弁理士 藤井 正弘
 (72) 発明者 東村 邦彦
 東京都国分寺市東恋ヶ窪一丁目280番地
 株式会社日立製作所 中央研究所内
 (72) 発明者 片岡 幹雄
 東京都国分寺市東恋ヶ窪一丁目280番地
 株式会社日立製作所 中央研究所内
 最終頁に続く

(54) 【発明の名称】コンピュータシステム

(57) 【特許請求の範囲】

【請求項1】

複数のサーバとパケット転送装置を備えるコンピュータシステムにおいて、前記各サーバは前記パケット転送装置と接続され、前記パケット転送装置は、ネットワークを介してクライアント端末と接続され、前記パケット転送装置は、前記クライアント端末から要求されたコンテンツの可変長の識別子から計算された固定長のハッシュ値を含むコンテンツ要求を受信し、前記受信したコンテンツ要求に含まれるハッシュ値を抽出し、前記抽出したハッシュ値に基づいて前記コンテンツ要求の転送先を前記複数のサーバのうちいずれのサーバにするかを決定することを特徴とするコンピュータシステム。

【請求項2】

前記クライアント端末は、要求に係るコンテンツの可変長の識別子から固定長のハッシュ値を計算し、前記計算されたハッシュ値をコンテンツ要求に格納し、前記コンテンツ要求をパケット転送装置に送信することを特徴とする請求項1に記載のコンピュータシステム。

【請求項3】

前記クライアント端末からのコンテンツ要求を受信するプロキシサーバを備え、前記プロキシサーバは、前記パケット転送装置及び前記クライアント端末と接続され、前記クライアント端末は、要求に係るコンテンツの可変長の識別子を含むコンテンツ要求を前記プロキシサーバに送信し、

10

20

前記プロキシサーバは、受信したコンテンツ要求から、要求に係るコンテンツの可変長の識別子を抽出し、前記抽出された識別子から固定長のハッシュ値を計算し、前記計算されたハッシュ値をコンテンツ要求に格納し、前記コンテンツ要求をパケット転送装置に送信することを特徴とする請求項1に記載のコンピュータシステム。

【請求項4】

前記プロキシサーバは、

前記クライアント端末からコンテンツ要求を受信するインターフェース部と、

当該要求に係るコンテンツの可変長の識別子から固定長のハッシュ値を計算し、

前記計算されたハッシュ値をコンテンツ要求に格納するプロセッサと、を備えることを特徴とする請求項3に記載のコンピュータシステム。 10

【請求項5】

前記クライアント端末は、

要求に係るコンテンツの可変長の識別子から固定長のハッシュ値を計算し、

前記計算されたハッシュ値を要求パケット内に書き込むプロセッサと、を備えることを特徴とする請求項1に記載のコンピュータシステム。

【請求項6】

前記パケット転送装置は、

前記コンテンツ要求に含まれる固定長のハッシュ値を抽出するプロセッサと、

前記コンテンツの可変長の識別子に対応するハッシュ値と前記コンテンツ要求の転送先との対応を含む転送先選択情報を記憶した記憶部と、を備えることを特徴とする請求項1に記載のコンピュータシステム。 20

【請求項7】

前記プロセッサは、

前記抽出されたハッシュ値及び前記転送先選択情報によって、前記コンテンツ要求の転送先を決定し、

前記決定された転送先にパケットを転送することを特徴とする請求項6に記載のコンピュータシステム。

【請求項8】

前記パケット転送装置は、

前記コンテンツ要求に代えて、前記クライアント端末から要求されたコンテンツの可変長の識別子から計算された固定長のハッシュ値を含むセッション確立要求を受信し、 30

前記受信したセッション確立要求に含まれるハッシュ値を抽出し、前記抽出したハッシュ値に基づいて前記セッション確立要求の転送先を前記複数のサーバのうちいずれのサーバにするかを決定することを特徴とする請求項1に記載のコンピュータシステム。

【請求項9】

パケット転送装置、オリジンサーバ、キャッシュサーバ及び制御サーバを含んだコンピュータシステムであって、

前記オリジンサーバは、クライアント端末にコンテンツを提供し、

前記キャッシュサーバは、前記オリジンサーバから提供されたコンテンツを記憶し、

前記制御サーバは、前記キャッシュサーバに記憶されたコンテンツを管理し、 40

前記パケット転送装置は、前記クライアント端末から要求されたコンテンツの可変長の識別子から計算された固定長のハッシュ値を含むコンテンツ要求を受信し、前記受信したコンテンツ要求に含まれるハッシュ値に基づいて、当該コンテンツ要求を、前記オリジンサーバ又は前記制御サーバへ振分けることを特徴とするコンピュータシステム。

【請求項10】

前記制御サーバは、前記キャッシュサーバが記憶したコンテンツの情報を記憶するキャッシュ所在情報を備えることを特徴とする請求項9に記載のコンピュータシステム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、パケット転送装置を備えるコンピュータシステムに関し、特に、URLスイッチを用いた転送制御に関する。

【背景技術】

【0002】

インターネット上で提供される各種サービスは、ストリーム型の通信プロトコルであるTCP上で実現されることが多い（非特許文献1参照）。例えば、World-Wide Webは、TCP上のプロトコルであるHypertext Transport Protocol（HTTP、非特許文献2参照）を使用している。

【0003】

HTTPプロトコルでは、以下の手順でサーバからコンテンツを取得する。まず、クライアント端末はサーバに対してTCPコネクションを確立する。ストリームの先頭には要求するコンテンツのURLが含まれる。サーバは、このURLを解釈することによってユーザが要求するコンテンツを認識し、認識したURLに対応するコンテンツをクライアント端末に返信する。10

【0004】

大規模なシステムになると、サーバは多数のユーザからの要求を処理する必要がある。この多量の処理に対応するため、複数のサーバに分散させて処理をすることがある。このとき、サーバとクライアント端末の間の通信経路上に負荷分散装置を設け、その負荷分散装置でユーザからの処理を振り分ける。負荷分散装置（URLスイッチ）は、クライアント端末が要求するURLを認識して転送先を決める。20

【0005】

URLスイッチの動作は以下の通りである。

【0006】

クライアント端末がサーバに対して接続するTCPコネクションを、URLスイッチが代理して終端する。URLスイッチは、クライアント端末が送信したコンテンツのURLを抽出する。そして、抽出したURLに対応した転送先を決定する。

【0007】

例えば、負荷分散のためのURLスイッチであれば、複数のサーバの中から当該コンテンツを格納しているサーバを選択する。URLスイッチは、選択されたサーバとの間でコネクションを設定し、クライアント端末からのリクエストを転送する。以後、URLスイッチはクライアント端末とサーバの間で交換されるパケットを中継する。30

【0008】

この動作によって、クライアント端末からのリクエストがURLスイッチにより適切なサーバに転送される。

【非特許文献1】IETF, RFC793, "Transmission Control Protocol", 1981年9月

【非特許文献2】IETF, RFC2616, "Hypertext Transfer Protocol -- HTTP/1.1", 1999年6月

【非特許文献3】Z. Genova and K. Christensen, "Using signatures to improve URL routing.", In proceedings of IEEE International Performance, Computing, and Communications Conference, 2002年4月, p. 45 - 5240

【発明の開示】

【発明が解決しようとする課題】

【0009】

URLスイッチの動作の第1の特徴は、受信した要求からURLを抽出することである。また、第2の特徴は、抽出したURLをキーとしたテーブル検索（又は、何らかの演算）によって転送先を決定することである。URLスイッチの転送先判別処理を高速化するためには、前述した二つの処理を高速化する必要がある。

【0010】

従来の方法では、URLスイッチがリクエストからURLを抽出し、URLをキーとして転送先を選択している。クライアント端末からのリクエストに含まれるURLの長さは50

一定しない可変長であり、一般的には 50 バイト程度である。この URL を抽出し、この URL に基づいてテーブルを検索（又は、演算）するためには、大きな処理能力、記憶容量が必要となる。すなわち、この方法では、高速かつ経済的な処理の実現が困難である。

【0011】

また、URL スイッチによって保持されるテーブルを小さくするために、テーブルの検索キーに URL のハッシュ値を用いる方法が提案されている（非特許文献 3 参照。）。この方法によると、URL スイッチは要求パケットから URL を抽出し、抽出した URL からハッシュ値を計算し、ハッシュ値をキーとしてテーブルを検索している。

【0012】

本発明は、転送先決定処理を複数の装置で分割して処理することによって、URL スイッチの計算量及び記憶容量を削減することを目的とする。10

【課題を解決するための手段】

【0013】

本発明は、複数のサーバとパケット転送装置を備えるコンピュータシステムにおいて、前記各サーバは前記パケット転送装置と接続され、前記パケット転送装置は、ネットワークを介してクライアント端末と接続され、前記パケット転送装置は、前記クライアント端末から要求されたコンテンツの可変長の識別子から計算された固定長のハッシュ値を含むコンテンツ要求を受信し、前記受信したコンテンツ要求に含まれるハッシュ値を抽出し、前記抽出したハッシュ値に基づいて前記コンテンツ要求の転送先を前記複数のサーバのうちいずれのサーバにするかを決定する。20

【0014】

すなわち、本発明のコンピュータシステムの一実施形態では、クライアント端末が要求に係るコンテンツの URL から固定長の値を生成し、その値を要求の中に含める。そして、URL スイッチは、要求に埋め込まれた固定長の値を抽出し、当該値に基づいて要求に対する処理を決定する。

【発明の効果】

【0015】

本発明によると、URL スイッチで行う処理が簡略化されることによって、URL スイッチの性能を向上させることができる。また、URL スイッチの処理を高速化することによって、URL スイッチを使用するアプリケーションが高速化し、ユーザに対してより高い品質のサービスを提供することができる。30

【発明を実施するための最良の形態】

【0016】

本明細書では、本発明を Web サーバの負荷分散に適用した実施の形態を説明する。

【0017】

Web サーバの負荷分散では、同一サイトに対するコンテンツ要求を複数のサーバに分散して処理することで、1台のサーバあたりの負荷を削減し、応答時間を短縮する。各サーバはコンテンツの提供を分担するが、本発明の実施の形態では、ハッシュ値から担当するサーバを決定するものとする。

【0018】

従来の URL スイッチを含むコンピュータシステムでは、URL スイッチがコンテンツ要求パケットに含まれる URL を抽出し、抽出した URL のハッシュ値を計算した上で、コンテンツ要求の転送先を決定する。これに対し、本発明の実施の形態では、ハッシュ値の計算をクライアント端末等に分散し、URL スイッチでの URL 抽出処理を簡略化することができる。40

【0019】

ハッシュ値を計算するために使用するハッシュ関数は多様である。例えば、MD5 は任意の長さのデータから 128 ビットのハッシュ値を計算するものであり、インターネット上のデータの認証などに広く用いられている。MD5 は、IETF, RFC1321, "The MD5 Message-Digest Algorithm" に記載されている。50

【0020】

また、C R C 3 2 は、任意の長さのデータから 3 2 ビットのハッシュ値を計算するものであり、主にデータの誤り検出に用いられている。C R C 3 2 は、ISO, IS3309, "ISO Information Processing Systems-Data Communication High-Level Data Link Control Procedure-Frame Structure" に記載されている。

【0021】

(実施形態 1)

図 1 は、第 1 の実施の形態のコンピュータシステムの構成を示すブロック図である。

【0022】

第 1 の実施の形態のコンピュータシステムは、クライアント端末 1 0 1 - 1 ~ 1 0 1 - n 10 、ネットワーク 1 0 2 及びサーバ群 1 0 4 によって構成されている。

【0023】

サーバ群 1 0 4 には、複数の W e b サーバ 1 0 4 - 1 ~ 1 0 4 - m 及び U R L スイッチ 1 0 5 を含む。W e b サーバ 1 0 4 - 1 等は、クライアント端末 1 0 1 - 1 等にコンテンツを提供するコンピュータである。U R L スイッチ 1 0 5 は、クライアント端末 1 0 1 - 1 等から送信されたコンテンツ要求 (U R L) を解析し、コンテンツ要求を W e b サーバ 1 0 4 - 1 ~ 1 0 4 - m に振り分ける負荷分散装置である。

【0024】

ネットワーク 1 0 2 は、クライアント端末 1 0 1 - 1 等とサーバ群 1 0 4 とを接続し、例えば、T C P / I P プロトコルによって通信可能である。 20

【0025】

クライアント端末 1 0 1 - 1 等は、W e b サーバ 1 0 4 - 1 等に、H T T P プロトコルによってコンテンツを要求するコンピュータである。なお、クライアント端末 1 0 1 - n のように、ネットワーク 1 0 2 との間にプロキシサーバ 1 0 3 を設けてもよい。

【0026】

なお、本実施の形態では後述するように、H T T P プロトコルによるコンテンツ要求に U R L のハッシュ値を含めている。U R L スイッチ 1 0 5 は、ハッシュ値に基づいてコンテンツ要求の転送先を決定する。

【0027】

図 2 は、第 1 の実施の形態のクライアント端末 1 0 1 - 1 の構成を示すブロック図である。なお、図 2 には、クライアント端末 1 0 1 - 1 の構成を示すが、他のクライアント端末 1 0 1 - 2 ~ 1 0 1 - n も同じ構成を有する。 30

【0028】

クライアント端末 1 0 4 - 1 は、ネットワークインターフェース 2 0 1 、C P U 2 0 2 、メモリ 2 0 3 及び二次記憶装置 2 0 4 を備えるコンピュータである。これらのクライアント端末 1 0 4 - 1 内の各構成は、相互に内部バス 2 0 5 で接続されている。

【0029】

ネットワークインターフェース 2 0 1 は、ネットワーク 1 0 2 に接続されている。よって、クライアント端末 1 0 1 - 1 は、ネットワーク 1 0 2 を介して U R L スイッチ 1 0 5 との間でデータや制御信号を送受信する。なお、一つのネットワークインターフェースを図示するが、2 以上のネットワークインターフェースを備えてよい。 40

【0030】

C P U 2 0 2 は、メモリ 2 0 3 に記憶されている各種プログラムを実行し、クライアント端末 1 0 1 - 1 を制御する。

【0031】

メモリ 2 0 3 には、処理プログラムが記憶される。処理プログラムは、ハッシュ値計算処理サブプログラム 2 0 6 及びコンテンツ要求サブプログラム 2 0 7 を含む。

【0032】

ハッシュ値計算処理サブプログラム 2 0 6 は、コンテンツ要求に係る U R L について固定長のハッシュ値を計算する。例えば、C R C 3 2 によって、ハッシュ値を計算する。コ 50

ンテンツ要求サブプログラム 207 は、ハッシュ値計算処理サブプログラム 206 によって計算されたハッシュ値を含んだコンテンツ要求を生成する。

【0033】

なお、クライアント端末 101-n のように、Web サーバ 104-1 等への経路の途中に存在するプロキシサーバ 103 においてハッシュ値を算出する場合は、ハッシュ値計算サブプログラム 206 を備えなくてよい。

【0034】

二次記憶装置 204 は、クライアント端末 101-1 の電源遮断時も記憶内容を保持するハードディスク装置である。

【0035】

図 3 は、第 1 の実施の形態のプロキシサーバ 103 の構成を示すブロック図である。

10

【0036】

プロキシサーバ 103 は、ネットワークインターフェース 301-1、301-2、CPU 302、メモリ 303 及び二次記憶装置 304 を備えるコンピュータである。これらの 103 内の各構成は、相互に内部バス 305 で接続されている。

【0037】

ネットワークインターフェース 301-1 は、ネットワーク 102 に接続されている。また、ネットワークインターフェース 301-2 は、クライアント端末 101-n に接続されている。

【0038】

CPU 302 は、メモリ 303 に記憶されている各種プログラムを実行し、プロキシサーバ 103 を制御する。

20

【0039】

メモリ 303 には、処理プログラムが記憶される。処理プログラムは、ハッシュ値計算処理サブプログラム 306 及びコンテンツ要求サブプログラム 307 を含む。

【0040】

ハッシュ値計算処理サブプログラム 306 は、クライアント端末 101-n から送信されたコンテンツ要求に係る URL に基づいて固定長のハッシュ値を計算する。コンテンツ要求サブプログラム 307 は、ハッシュ値計算処理サブプログラム 306 によって計算されたハッシュ値を含んだコンテンツ要求を生成する。

30

【0041】

二次記憶装置 304 は、プロキシサーバ 103 の電源遮断時も記憶内容を保持するハードディスク装置である。

【0042】

図 4 は、第 1 の実施の形態の URL スイッチ 105 の構成を示すブロック図である。

【0043】

URL スイッチ 105 は、ネットワークインターフェース 401-1 ~ 401-m+1、パケット処理装置 402 及びメモリ 403 を備えるコンピュータである。これらの URL スイッチ 105 内の各構成は、相互に内部バス 404 で接続されている。

【0044】

ネットワークインターフェース 401-1 は、ネットワーク 102 に接続されている。また、ネットワークインターフェース 401-2 ~ 401-m+1 は Web サーバ 104-1 ~ 104-m に接続されている。

40

【0045】

パケット処理装置 402 は、入力されたパケットを所定の宛先に転送するネットワークプロセッサである。

【0046】

メモリ 403 には、処理プログラムが記憶される。処理プログラムは、ハッシュ値抽出処理サブプログラム 405 及び転送先決定処理サブプログラム 406 を含む。メモリ 403 には、URL スイッチ 105 が送受信するパケットが一時的に格納されるバッファが設

50

けられている。また、メモリ 403 には、転送先選択テーブル 407 が記憶される。

【0047】

ハッシュ値抽出処理サブプログラム 405 は、クライアント端末 101-n 等から送信されたコンテンツ要求からハッシュ値を抽出する。転送先決定処理サブプログラム 406 は、ハッシュ値抽出処理サブプログラム 405 によって抽出されたハッシュ値をキーとして、転送先選択テーブル 407 を検索し、コンテンツ要求の転送先を決定する。

【0048】

転送先選択テーブル 407 は、ハッシュ値と転送先との対応を格納している。

【0049】

図 5 は、第 1 の実施の形態の転送先選択テーブル 407 の構成の説明図である。

10

【0050】

転送先選択テーブル 407 は、URL のハッシュ値 4071 と、転送先の Web サーバの IP アドレス 4072 とを含む。

【0051】

URL のハッシュ値 4071 は、コンテンツ要求に係る URL から計算された固定長のハッシュ値である。転送先 4072 は、当該ハッシュ値の計算の元となった URL に対応する Web サーバの IP アドレスである。すなわち、当該ハッシュ値を含むコンテンツ要求の転送先となる Web サーバの IP アドレスである。

【0052】

URL スイッチ 105 は、クライアント端末 101-1 等から受信したコンテンツ要求に含まれるハッシュ値をキーとして、転送先選択テーブル 407 を検索し、転送先の Web サーバを特定する。

20

【0053】

図 6 は、第 1 の実施の形態のコンテンツ要求パケットの説明図である。

【0054】

クライアント端末 101-1 からのコンテンツ要求パケットには、X-Hash ヘッダが設けられている。X-Hash ヘッダには、この URL に対するハッシュ値（例えば、0123ABCD）が格納される。

【0055】

このコンテンツ要求パケットを受信した URL スイッチ 105 は、X-Hash ヘッダからハッシュ値を抽出し、コンテンツ要求パケットの転送先を決定する。

30

【0056】

次に、コンテンツ要求の手順について説明する。

【0057】

図 7 は、第 1 の実施の形態のクライアント端末 101 のコンテンツ要求処理のフローチャートである。

【0058】

クライアント端末 101 は、コンテンツを要求する URL が決定すると、ハッシュ値計算処理サブプログラム 306 によって、URL のハッシュ値を計算する（ステップ 701）。

40

【0059】

次に、計算したハッシュ値をコンテンツ要求のヘッダ（X-Hash ヘッダ）に格納して、コンテンツ要求パケットを作成する（ステップ 702）。

【0060】

次に、Web サーバ 104-1 等との間でセッションを確立する（ステップ 703）。具体的には、クライアント端末 101-1 は URL スイッチ 105 と接続し、URL スイッチ 105 との間でセッションを確立する。

【0061】

そして、ステップ 702 で作成したコンテンツ要求パケットを URL スイッチ 105 に対して送信する（ステップ 704）。その後、URL スイッチ 105 経由で Web サーバ

50

104 - 1等からコンテンツを受信する（ステップ705）。

【0062】

図8は、第1の実施の形態のURLスイッチ105のセッション中継処理のフローチャートである。このセッション中継処理は、クライアント端末101-1等からのコンテンツ要求を受信すると実行される。

【0063】

URLスイッチ105は、クライアント端末101からのコンテンツ要求パケットを受信すると、当該コンテンツ要求パケット内のX-Hashヘッダの有無を判定する（ステップ801）。

【0064】

その結果、X-Hashヘッダが存在する場合は、X-Hashヘッダからハッシュ値を抽出する（ステップ802）。このハッシュ値抽出処理は、URLを抽出するより簡単な処理である。なぜなら、可変長の文字列（URL）を抽出する処理より固定長の値（ハッシュ値）を抽出する処理の方が単純だからである。

【0065】

次に、抽出されたハッシュ値をキーとして転送先検索テーブル407を検索し（ステップ803）、該当するエントリが存在するか否かを判定する（ステップ804）。

【0066】

その結果、該当エントリが発見された場合は、検索されたエントリのWebサーバ104-1を転送先に決定し、決定されたWebサーバ104-1との間でセッションを確立する（ステップ805）。

【0067】

その後、クライアント端末101-1からのコンテンツ要求を転送先Webサーバ104-1に送信する（ステップ806）。

【0068】

そして、クライアント端末101-1とWebサーバ104-1との間のセッションが切断されるまでそのセッションを中継する（ステップ807）。

【0069】

ステップ801でX-Hashヘッダが存在しない場合、又は、ステップ804で該当するエントリが発見されなかった場合は、予め定められたWebサーバを転送先に決定し、決定されたサーバとの間でセッションを確立する（ステップ808）。その後、クライアント端末からのコンテンツ要求を転送先サーバに送信する（ステップ806）。

【0070】

X-Hashヘッダが存在しないコンテンツ要求が転送されたWebサーバは、クライアント端末101-1から送信されたURLからハッシュ値を求め、適切なWebサーバにコンテンツ要求を転送する。この場合は、URLスイッチとは別に、Webサーバも図5に示した転送先選択テーブル407を有しており、この転送先選択テーブル407に基づいて適切なWebサーバを検索する。これによって、X-Hashヘッダを埋め込むことができないクライアント端末101-1からコンテンツ要求があった場合でも、正しく負荷を分散できる。なお、コンテンツ要求が転送されたWebサーバが、通常の手順で、URLを解析し、適切なWebサーバにコンテンツ要求を転送してもよい。

【0071】

このように、URLスイッチ105がセッションを中継することによって、クライアント端末101からのコンテンツ要求を適切なサーバ104に転送する。

【0072】

図9は、第1の実施の形態のコンテンツ転送のシーケンス図である。

【0073】

クライアント端末101-1は、URLスイッチ105との間のセッションを確立するためにTCPの3-wayハンドシェイクを行う（図7のステップ703）。

【0074】

10

20

30

40

50

具体的には、クライアント端末101-1は、URLスイッチ105に対してSYNパケットを送信して、クライアント端末101-1とURLスイッチ105との間のセッション確立を要求する。URLスイッチ105は、SYNパケットを受信すると、セッションの確立が可能であれば、SYN, ACKパケットを送信する。クライアント端末101-1は、SYN, ACKパケットを受信すると、ACKパケットを送信する。このパケットの交換によって、クライアント端末101-1とURLスイッチ105との間のセッションが確立する。

【0075】

そして、クライアント端末101-1は、コンテンツ要求パケットを送信する（図7のステップ704）。URLスイッチ105は、コンテンツ要求パケットを受信すると、受信したパケットからハッシュ値を抽出し（図8のステップ802）、転送先を決定する（図8のステップ803）。

【0076】

そして、決定された転送先Webサーバ104-1とURLスイッチ105との間のセッションを確立する（ステップ805）。その後、セッションを確立したWebサーバ104-1に対し、クライアント端末101-1からのコンテンツ要求パケットを送信する（図8のステップ806）。

【0077】

Webサーバ104-1は、コンテンツ要求パケットを受信すると、OKパケットを送信する。URLスイッチ105は、Webサーバ104-1が送信したOKパケットを、クライアント端末101-1に転送する。また、Webサーバ104-1は、コンテンツを送信する。

【0078】

その後、Webサーバ104-1は、コンテンツの送信が終了すると、FINパケットを送信して、セッションの終了を要求する。URLスイッチは、Webサーバ104-1が送信したFINパケットを、クライアント端末101-1に転送する。

【0079】

クライアント端末101-1は、FINパケットを受信すると、Webサーバ104-1との間のセッションを終了し、FIN, ACKパケットを送信する。Webサーバ104-1は、FIN, ACKパケットを受信すると、ACKパケットを送信する。このパケットの交換によって、クライアント端末101-1とWebサーバ104-1との間のセッションが終了する。

【0080】

以上説明したように、第1の実施の形態では、クライアント端末がURLからハッシュ値を求めるこことによって、URLスイッチにおける識別子抽出、計算処理を単純化している。これにより、URLスイッチに実装が必要なプログラム量及び価格を削減することができ、高速な負荷分散を実現することができる。

【0081】

すなわち、第1の実施の形態では、クライアントからのコンテンツ要求の転送先決定処理のうち、URL抽出処理を簡略化する。このため、クライアント端末においてURLから固定長のハッシュ値を生成し、その値をコンテンツ要求に含める。これによって、URLスイッチが抽出すべき値が可変長の長い文字列から、固定長の短い文字列になる。これによってURLスイッチによる処理を軽減することができる。

【0082】

また、URLスイッチによる転送先決定処理に使用されるデータがURLからハッシュ値になることによって、データ量が約50バイトから数バイトに減る。これによって、URLスイッチの記憶容量を削減することができる。

【0083】

（実施形態2）

次に、本発明を、分散キャッシュを備えるコンピュータシステムに適用した例を説明す

10

20

30

40

50

る。

【0084】

キャッシュサーバは、クライアント端末とコンテンツサーバ（オリジンサーバ）との間に設けられる。キャッシュサーバは、クライアント端末からのコンテンツ要求に対してサーバを代理して返信することによって、クライアント端末への応答時間の短縮、コンテンツサーバの負荷低減、及びキャッシュサーバとコンテンツサーバとの間のトラフィックの削減をしている。

【0085】

キャッシュサーバは、より多数のクライアント端末を収容することによって、複数のクライアント端末から要求される共通のコンテンツの記憶量を増やし、キャッシュのヒット率を向上させる。

10

【0086】

しかし、多数のクライアント端末からのコンテンツ要求がキャッシュサーバに集中するため、キャッシュサーバの処理性能がボトルネックとなる。これを解決するため、複数のキャッシュサーバでクライアント端末を分担し、かつキャッシュされたコンテンツを共有することによって、コンテンツの記憶量の増大と、サーバのボトルネック解消の両者を実現する。

【0087】

このような分散型キャッシュとして、2つの方法が考えられる。

【0088】

20

第1の方法は、分散された複数のキャッシュサーバ同士が直接通信し、お互いのキャッシュしているコンテンツの情報を交換するものである。この方法は、Internet Cache Protocolとして、RFC2186、"Internet Cache Protocol (ICP)" に詳しく記載されている。

【0089】

第2の方法は、キャッシュされたコンテンツの情報を制御サーバにて集中的に管理し、各キャッシュサーバは制御サーバの指示に従ってコンテンツを取得するものである。このような分散キャッシュ制御方法は、例えば、特開2005-10970号公報に記載されている。

【0090】

第1の方法は第2の方法と比較して、多数のキャッシュサーバが互いに通信する。すなわち、第1の方法は、キャッシュサーバ間の通信が、サーバ台数の2乗に比例して増加する。よって、第1の方法はスケーラビリティが低い。

30

【0091】

一方、第2の方法では、キャッシュの有無に関する問い合わせが制御サーバに集中する。この解決のために、制御サーバの処理を、URLスイッチを利用した専用機器に移し、制御サーバをオフロードする。このため、ネットワークプロセッサなどのパケット処理に特化した機器を用い、パケット処理を効率化する必要がある。

【0092】

しかし、パケット処理に特化した機器では、パケットに対して行える処理は限られている。例えば、特定の位置に存在する固定長の情報は高速に抽出できるが、URLのような可変長の文字列を高速に扱うことはできない。

40

【0093】

これを解決するため、第2の実施の形態では、キャッシュサーバから制御サーバへのコンテンツ要求パケット中にURLのハッシュ値を埋め込む。これによって、パケット処理に特化した機器においてもURLの解析と同等の処理を高速に行うことができる。

【0094】

第2の実施の形態においても、実施例1と同様に、URLのハッシュ計算にCRC32を利用する。

【0095】

図10は、第2の実施の形態のコンピュータシステムの構成を示すブロック図である。

50

【0096】

第1の実施の形態のコンピュータシステムは、キャッシュサーバ1001-1～1001-2、クライアント端末1002-1～1002-4、ネットワーク1003-1～1003-2、URLスイッチ1004、オリジンサーバ1005-1～1005-2、ネットワーク1006、ネットワーク1007及びキャッシュ制御サーバ1008によって構成されている。

【0097】

クライアント端末1002-1等は、オリジンサーバ1005-1等に、HTTPプロトコルによってコンテンツを要求するコンピュータである。クライアント端末1002-1等は、前述した第1の実施の形態のクライアント端末101-1と同じ構成を有する(図2参照)。
10

【0098】

キャッシュサーバ1001-1、1001-2は、クライアント端末1002-1等がオリジンサーバ1005-1等から取得したコンテンツを記憶する。

【0099】

URLスイッチ1004は、クライアント端末1002-1等からのコンテンツ要求パケットを解析して、キャッシュが記憶されているか否かを判定し、コンテンツ要求を振り分ける負荷分散装置である。第2の実施の形態のコンピュータシステムでは、URLのハッシュ値がコンテンツ要求パケットに含まれているので、URLスイッチ1004は、このURLのハッシュ値を用いてキャッシュの有無を判定する。
20

【0100】

オリジンサーバ1005-1及び1005-2は、はクライアント端末1002-1等にコンテンツを提供するコンピュータである。

【0101】

各キャッシュサーバ1001-1、1001-2は、一つ以上のクライアント端末1002-1等とネットワーク1003-1等を介して接続されている。URLスイッチ1004は、キャッシュサーバ1001-1等とURLスイッチ1004とはネットワーク1006を介して接続されている。また、URLスイッチ1004とオリジンサーバ1005-1等とはネットワーク1007を介して接続されている。URLスイッチ1004には、キャッシュ制御サーバ1008が接続されている。これらのネットワーク1003-1、1003-2、1006及び1007は、例えば、TCP/IPプロトコルによって通信可能である。
30

【0102】

図11は、第2の実施の形態のキャッシュサーバ1001-1の構成を示すブロック図である。なお、図11には、キャッシュサーバ1001-1の構成を示すが、他のキャッシュサーバ1001-2も同じ構成を有する。

【0103】

キャッシュサーバ1001-1は、インターフェース1101-1及び1101-2、1102、メモリ1103及び二次記憶装置1104を備えるコンピュータ装置である。これらのキャッシュサーバ1001-1内の各構成は相互に内部バス1105で接続されている。
40

【0104】

インターフェース1101-1は、クライアント端末側のネットワーク1003に接続されている。インターフェース1101-2は、他のキャッシュサーバおよびURLスイッチに接続されている。

【0105】

CPU1102は、メモリ1103に記憶されている各種プログラムを実行し、キャッシュサーバを制御する。

【0106】

メモリ1103には、処理プログラムが記憶される。処理プログラムは、キャッシュ有
50

無判定処理サブプログラム 1106、ハッシュ値計算処理サブプログラム 1107、コンテンツ要求処理サブプログラム 1108 及びコンテンツ登録処理サブプログラム 1109 を含む。

【0107】

キャッシュ有無判定処理サブプログラム 1106 は、二次記憶装置 1104 内に格納されたコンテンツデータベース 1110 を検索し、ユーザが要求するコンテンツを当該キャッシュサーバが格納しているか否かを判定する。

【0108】

ハッシュ値計算処理サブプログラム 1107 は、クライアント端末 1002-1 等から送信されたコンテンツ要求に係る URL から固定長のハッシュ値を計算する。

10

【0109】

コンテンツ要求処理サブプログラム 1108 は、クライアント端末 1002-1 等から要求されたコンテンツが、当該キャッシュサーバに格納されていないときに、キャッシュ制御サーバ 1008 にコンテンツを要求する。

【0110】

コンテンツ登録処理サブプログラム 1109 は、URL スイッチ 1004 を経由してオリジンサーバ 1005-1 等から、又は直接他のキャッシュサーバ 1001-2 からコンテンツを取得する。そして、その内容を当該キャッシュのコンテンツデータベース 1110 内に格納する。

20

【0111】

二次記憶装置 1104 は、キャッシュサーバ 1001-1 の電源遮断時も記憶内容を保持するハードディスク装置であり、コンテンツデータベース 1110 が格納されている。

【0112】

コンテンツデータベース 1110 は、コンテンツと当該コンテンツの URL との対応を格納している。

【0113】

図 12 は、第 2 の実施の形態のコンテンツデータベース 1110 の構成の説明図である。

【0114】

コンテンツデータベース 1110 は、URL 11101 と、コンテンツ 11102 とを含む。

30

【0115】

URL 11101 は、当該コンテンツがオリジンサーバに格納されている場所を示す URL である。コンテンツ 11102 は、コンテンツのデータそのものである。

【0116】

キャッシュサーバ 1001-1 は、クライアント端末 1002-1 等から受信したコンテンツ要求に含まれる URL をキーとして、コンテンツデータベース 1110 を検索し、当該キャッシュサーバにコンテンツが格納されているかを判定する。

【0117】

図 13 は、第 2 の実施の形態の URL スイッチ 1004 の構成を示すブロック図である。

40

【0118】

URL スイッチ 1004 は、ネットワークインターフェース 1201-1 ~ 1201-3、パケット処理装置 1202 及びメモリ 1203 を備えるコンピュータである。これらの各構成は、相互に内部バス 1204 で接続されている。

【0119】

ネットワークインターフェース 1201-1 は、ネットワーク 1006 に接続されている。ネットワークインターフェース 1201-2 は、キャッシュ制御サーバ 1008 に接続されている。また、ネットワークインターフェース 1201-3 はネットワーク 1007 に接続されている。

50

【 0 1 2 0 】

パケット処理装置 1202 は、入力されたパケットを所定の宛先に転送するネットワークプロセッサである。

【 0 1 2 1 】

メモリ 1203 には、処理プログラムが記憶される。処理プログラムは、ハッシュ値抽出処理サブプログラム 1205、転送先決定処理サブプログラム 1206 及びハッシュ値登録処理サブプログラム 1207 を含む。メモリ 1203 には、URLスイッチ 1004 が送受信するパケットが一時的に格納されるバッファが設けられている。また、メモリ 1203 には、転送先選択テーブル 1208 が記憶される。

【 0 1 2 2 】

ハッシュ値抽出処理サブプログラム 1205 は、クライアント端末 1002-1 等から送信されたコンテンツ要求からハッシュ値を抽出する。転送先決定処理サブプログラム 1206 は、ハッシュ値抽出処理サブプログラム 1205 によって抽出されたハッシュ値をキーとして、転送先選択テーブル 1208 を検索し、コンテンツ要求の転送先を決定する。ハッシュ値登録処理サブプログラム 1207 は、転送先選択テーブル 1208 へのハッシュ値の登録要求及び削除要求を受け付ける。

【 0 1 2 3 】

転送先選択テーブル 1208 は、ハッシュ値と転送先との対応を格納する。

【 0 1 2 4 】

図 14 は、第 2 の実施の形態の転送先選択テーブル 1208 の構成の説明図である。

20

【 0 1 2 5 】

転送先選択テーブル 1208 は、システム内のキャッシュサーバに格納される全てのキャッシュのデータを含む。各データは、URL のハッシュ値 12081 と、転送先のキャッシュ制御サーバの IP アドレス 12082 とを含む。

【 0 1 2 6 】

URL のハッシュ値 12081 は、コンテンツ要求に係る URL から計算された固定長のハッシュ値である。転送先 12082 は、当該ハッシュ値の計算の元となった URL に対応するキャッシュを格納するキャッシュサーバを制御するキャッシュ制御サーバの IP アドレスである。すなわち、転送先 12082 は、当該ハッシュ値が含まれるコンテンツ要求の転送先となるキャッシュ制御サーバの IP アドレスである。

30

【 0 1 2 7 】

URL スイッチ 1004 は、キャッシュサーバ 1001-1 等から受信したコンテンツ要求に含まれるハッシュ値をキーとして、転送先選択テーブル 1208 を検索し、転送先のキャッシュ制御サーバを特定する。

【 0 1 2 8 】

図 15 は、第 2 の実施の形態のキャッシュ制御サーバ 1008 の構成を示すブロック図である。

【 0 1 2 9 】

キャッシュ制御サーバ 1008 は、ネットワークインターフェース 1301、CPU 1302、メモリ 1303 及び二次記憶装置 1304 を備えるコンピュータである。これらの各構成は、相互に内部バス 1306 で接続されている。

40

【 0 1 3 0 】

ネットワークインターフェース 1301 は、URL スイッチ 1004 に接続されている。

【 0 1 3 1 】

CPU 1302 は、メモリ 1303 に記憶されている各種プログラムを実行し、制御サーバを制御する。

【 0 1 3 2 】

メモリ 1303 には、処理プログラムが記憶される。処理プログラムは、キャッシュ有無判定処理サブプログラム 1307、コンテンツ転送指示処理サブプログラム 1308 及

50

びURL登録処理サブプログラム1309を含む。

【0133】

キャッシュ有無判定処理サブプログラム1307は、URLスイッチから送信されたコンテンツ要求パケットからURLを抽出し、キャッシュ所在データベース1305を検索して、キャッシュを保有するキャッシュサーバを決定する。

【0134】

コンテンツ転送指示処理サブプログラム1308は、キャッシュを保有するキャッシュサーバに、コンテンツの転送を指示する。

【0135】

URL登録処理サブプログラム1309は、キャッシュサーバが取得したコンテンツの情報を、キャッシュ所在データベースに登録する。 10

【0136】

二次記憶装置1304は、プロキシサーバ103の電源遮断時も記憶内容を保持するハードディスク装置であり、キャッシュ所在データベース1305が格納されている。

【0137】

キャッシュ所在DB1305は、URLのハッシュ値とコンテンツ要求の転送先との対応を格納している。

【0138】

図16は、第2の実施の形態のキャッシュ所在データベース1305の構成の説明図である。 20

【0139】

キャッシュ所在データベース1305は、URL13051と、キャッシュを格納するキャッシュサーバのIPアドレス13052とを含む。

【0140】

URL13051は、当該コンテンツがオリジンサーバに格納されている場所を示すURLである。キャッシュサーバ13052は、当該コンテンツが格納されているキャッシュサーバのIPアドレスである。キャッシュ制御サーバ1008は、キャッシュサーバ1001-1等から受信したコンテンツ要求に含まれるURLをキーとして、キャッシュ所在データベース1305を検索し、どのキャッシュサーバにコンテンツが格納されているかを判定する。 30

【0141】

図17、図18は、第2の実施の形態のコンテンツ要求パケットの説明図である。

【0142】

図17は、クライアント端末1002からのコンテンツ要求パケットを示す。コンテンツ要求パケットには、要求するコンテンツのURLが含まれている。

【0143】

キャッシュサーバ1001-1は、クライアント端末1002からのコンテンツ要求パケットを受信すると、受信したパケットからURLを抽出する。具体的には、Refererヘッダーから"http://www.example.com/"を抽出し、GETヘッダーから"index.html"を抽出する。そして、これらを合成し、URL"http://www.example.com/index.html"を得る。 40

【0144】

図18は、キャッシュサーバ1001-1からのコンテンツ要求パケットを示す。

【0145】

キャッシュサーバ1001-1からのコンテンツ要求パケットは、クライアント端末1002-1からのコンテンツ要求パケット(図17)に、X-Hashヘッダが付加されている。X-Hashヘッダには、このURLに対するハッシュ値(例えば、0123ABCD)が格納される。

【0146】

このコンテンツ要求パケットを受信したURLスイッチ105は、X-Hashヘッダからハッシュ値を抽出し、コンテンツ要求パケットの転送先を決定する。 50

【0147】

また、キャッシュ制御サーバ1008は、キャッシュサーバ1001-1等からのコンテンツ要求パケットを受信すると、受信したパケットからURLを抽出し、"http://www.example.com/index.html"を得る。

【0148】

次に、コンテンツ要求の手順について説明する。

【0149】

図19は、第2の実施の形態のキャッシュサーバ1001-1のコンテンツ要求処理のフローチャートである。

【0150】

まず、キャッシュサーバ1001は、クライアント端末1002-1からのコンテンツ要求パケット(図17)を受信すると、コンテンツ要求パケットに含まれているコンテンツのURLを抽出する(ステップ1501)。そして、抽出したURLをキーにして、コンテンツデータベース1110を検索し(ステップ1502)、該当するコンテンツが存在するか否かを判定する(ステップ1503)。

【0151】

その結果、クライアント端末1002-1が要求するコンテンツがコンテンツデータベース1110に存在する場合、コンテンツデータベース1110からコンテンツを取得し、取得したコンテンツをクライアント端末1002-1に送信する(ステップ1504)。

【0152】

一方、当該コンテンツがコンテンツデータベース1110に存在しない場合、キャッシュサーバ1001は、ハッシュ値計算処理サブプログラム1107によって、要求のコンテンツのURLからハッシュ値を計算する(ステップ1505)。そして、計算したハッシュ値をコンテンツ要求のヘッダ(X-Hashヘッダ)に格納したコンテンツ要求パケットを作成し、作成したコンテンツ要求パケットを、オリジンサーバ1005-1に送信する(ステップ1506)。このコンテンツ要求はURLスイッチ1004を経由する。

【0153】

このコンテンツ要求に対して、2通りの返信が存在する。第一の返信は、オリジンサーバ1005-1から転送されるコンテンツである。この場合、キャッシュサーバ1001-1はオリジンサーバ1005-1からコンテンツを受信する(ステップ1508)。

【0154】

一方、第二の返信は、他のキャッシュサーバ1001-2から転送されるコンテンツである。この場合、要求元キャッシュサーバ1001-1は、キャッシュ制御サーバ1008から他のキャッシュサーバ1001-2からコンテンツが転送される旨を受信し、その後、他のキャッシュサーバ1001-2からコンテンツを受信する(ステップ1511)。

【0155】

要求元キャッシュサーバ1001-1は、コンテンツを受信した後、受信したコンテンツをクライアント端末1002-1に転送する(ステップ1509)。その後、コンテンツデータベース1110にコンテンツを登録する(ステップ1510)。そして、当該キャッシュにコンテンツが追加されたことをキャッシュ制御サーバ1008に送信する(ステップ1512)。

【0156】

図20は、第2の実施の形態のURLスイッチ1004のコンテンツ要求処理のフローチャートである。

【0157】

URLスイッチ1004は、キャッシュサーバ1001-1等とオリジンサーバ1005-1等との間に設けられている。よって、URLスイッチ1004は、キャッシュサーバ1001-1からオリジンサーバ1005へのコンテンツ要求パケット(図18)を捕

10

20

30

40

50

捉することができる。

【0158】

URLスイッチ1004は、キャッシングサーバ1001からオリジンサーバ1005へのコンテンツ要求パケットを捕捉すると、捕捉したパケットからURLのハッシュ値を抽出する（ステップ1601）。そして、抽出したハッシュ値をキーとして、転送先選択テーブル1208を検索し（ステップ1602）、該当するエントリが存在するか否かを判定する（ステップ1603）。

【0159】

その結果、転送先選択テーブルに対応するエントリが存在した場合、いずれかのキャッシングサーバに当該コンテンツが存在する。よって、検索されたエントリに記載された制御サーバを転送先と決定し、決定された制御サーバにコンテンツ要求パケットを転送する（ステップ1604）。

10

【0160】

一方、対応するエントリが存在しない場合、いずれのキャッシングサーバにも当該コンテンツが存在しない。よって、URLスイッチはコンテンツ要求パケットをオリジンサーバ1005-1へ転送する（ステップ1605）。キャッシングサーバ1001から送信されたコンテンツ要求パケットの宛先はオリジンサーバ1005-1なので、URLスイッチ1004はオリジンサーバ1005-1のIPアドレスを取得する必要はない。

【0161】

図21は、第2の実施の形態のキャッシング制御サーバ1008のコンテンツ要求処理のフローチャートである。

20

【0162】

制御サーバ1008は、URLスイッチ1004から転送されたコンテンツ要求パケットを受信すると、コンテンツ要求パケットからコンテンツのURLを抽出する（ステップ1701）。そして、抽出したURLをキーとして、キャッシング所在データベース1305を検索し（ステップ1702）、該当するエントリが存在するか否かを判定する（ステップ1703）。

【0163】

当該URLに対するエントリが存在した場合、他のキャッシングサーバに当該コンテンツが存在する。よって、要求元キャッシングサーバ1001-1に対して、他のキャッシングサーバ1001-2からコンテンツが転送される旨を返信する。そして、検索されたエントリに含まれるキャッシングサーバ1001-2を選択し、当該URLのコンテンツを要求元のキャッシングサーバ1001-1に対して転送する指示をする（ステップ1704）。

30

【0164】

一方、当該URLに対するエントリが存在しなかった場合、URLスイッチ1004によるキャッシングの有無判定が間違っていたことになる。URLスイッチ1004では、ハッシュ値に基づいてキャッシングの有無を判定しているので、ハッシュ値の衝突が生じうる。

【0165】

このため、キャッシング所在データベース1305にエントリが存在しなかった場合は、キャッシング制御サーバ1008はオリジンサーバ1005-1にコンテンツ要求パケットを転送し、コンテンツを取得する（ステップ1705）。

40

【0166】

そして、オリジンサーバ1005-1から得られたコンテンツを要求元キャッシングサーバ1001-1に返信する（ステップ1706）。そして、要求元キャッシングサーバ1001-1は、コンテンツをクライアント端末1002-1に転送し、当該コンテンツをキャッシングに格納する。

【0167】

図22、図23は、第2の実施の形態のコンテンツ転送のシーケンス図である。

【0168】

50

図22は、キャッシュサーバから要求されたコンテンツが、いずれのキャッシュサーバにも存在しない場合のシーケンスを示す。

【0169】

キャッシュサーバ1001-1は、クライアント端末からコンテンツ要求パケットを受信すると、URLスイッチ1004との間のセッションを確立するためにTCPの3-wayハンドシェイクを行う。

【0170】

具体的には、キャッシュサーバ1001-1は、URLスイッチ1004に対してSYNパケットを送信して、キャッシュサーバ1001-1とURLスイッチ1004との間のセッション確立を要求する。URLスイッチ1004は、SYNパケットを受信すると、セッションの確立が可能であれば、SYN,ACKパケットを送信する。キャッシュサーバ1001-1は、SYN,ACKパケットを受信すると、ACKパケットを送信する。このパケットの交換によって、キャッシュサーバ1001-1とURLスイッチ1004との間のセッションが確立する。

【0171】

そして、キャッシュサーバ1001-1は、クライアント端末から受信したコンテンツ要求パケットを受信すると、受信したパケットにURLのX-Hash値を追加して、URLスイッチ1004に送信する(図19のステップ1506)。URLスイッチ1004は、コンテンツ要求パケットを受信すると、受信したパケットからハッシュ値を抽出し(図20のステップ1601)、転送先を決定する(図20のステップ1602)。

10

20

【0172】

図22に示す場合は、転送先選択テーブル1208に該当するエントリが存在しないので(要求されたコンテンツがキャッシュサーバに存在しないので)、オリジンサーバ1005-11005-1にコンテンツ要求パケットを転送する。このため、URLスイッチ1004とオリジンサーバ1005-1との間のセッションを確立する。その後、URLスイッチ1004は、セッションを確立したオリジンサーバ1005-1に対し、クライアント端末からのコンテンツ要求パケットを送信する(図20のステップ1605)。

【0173】

オリジンサーバ1005-1は、コンテンツ要求パケットを受信すると、OKパケットを送信する。URLスイッチ1004は、オリジンサーバ1005-1が送信したOKパケットを、キャッシュサーバ1001-1に転送する。また、オリジンサーバ1005-1は、コンテンツを送信する。

30

【0174】

その後、オリジンサーバ1005-1は、コンテンツの送信が終了すると、FINパケットを送信する。URLスイッチ1004は、オリジンサーバ1005-1が送信したFINパケットを、キャッシュサーバ1001-1に転送する。

【0175】

キャッシュサーバ1001-1は、FINパケットを受信すると、オリジンサーバ1005-1との間のセッションを終了し、FIN,ACKパケットを送信する。オリジンサーバ1005-1は、FIN,ACKパケットを受信すると、ACKパケットを送信する。このパケットの交換によって、キャッシュサーバ1001-1とオリジンサーバ1005-1との間のセッションが終了する。

40

【0176】

その後、キャッシュサーバ1001-1は、転送されたコンテンツのURLをキャッシュ制御サーバ1008に送る。キャッシュ制御サーバ1008は、受信したURLを、キャッシュ所在データベース1305に追加する。

【0177】

その後、キャッシュ制御サーバ1008は、受信したURLのハッシュ値を計算し、計算したハッシュ値をURLスイッチ1004に送る。URLスイッチ1004は、転送されたハッシュ値を、当該ハッシュ値を計算したキャッシュ制御サーバ1008と対応付け

50

て、転送先選択テーブル 1208 に追加する。

【0178】

図 23 は、他のキャッシュサーバが要求のコンテンツを格納していた場合のシーケンスを示す。

【0179】

要求元キャッシュサーバ 1001-1 は、クライアント端末からコンテンツ要求パケットを受信すると、URLスイッチ 1004 との間のセッションを確立するために、TCP の 3-way ハンドシェイクを行う。この手順は図 22 に示したものと同様である。

【0180】

そして、要求元キャッシュサーバ 1001-1 は、クライアント端末から受信したコンテンツ要求パケットを受信すると、受信したパケットに URL の X-Hash 値を追加して、URL スイッチ 1004 に送信する（図 19 のステップ 1506）。URL スイッチ 1004 は、コンテンツ要求パケットを受信すると、受信したパケットからハッシュ値を抽出し（図 20 のステップ 1601）、転送先を決定する（図 20 のステップ 1602）。

【0181】

図 23 に示す場合は、転送先選択テーブル 1208 に該当するエントリが存在するので（要求されたコンテンツがキャッシュサーバに存在するので）、他のキャッシュサーバ 1001-2 からコンテンツを取得する。このため、URL スイッチ 1004 とキャッシュ制御サーバ 1008 との間のセッションを確立する。その後、URL スイッチ 1004 は、キャッシュ制御サーバ 1008 に対し、クライアント端末からのコンテンツ要求パケットを送信する（図 20 のステップ 1604）。

【0182】

キャッシュ制御サーバ 1008 は、コンテンツ要求パケットを受信すると、キャッシュ所在データベース 1305 を検索して、要求コンテンツを格納するキャッシュサーバ 1001-2 を特定する（図 21 のステップ 1702）。そして、特定されたキャッシュサーバ 1001-2 にコンテンツの転送を指示する（図 21 のステップ 1704）。

【0183】

要求コンテンツを格納するキャッシュサーバ 1001-2 は、要求元キャッシュサーバ 1001-1 にコンテンツを送信する。

【0184】

その後、コンテンツの送信が終了すると、キャッシュ制御サーバ 1008 は、FIN パケットを送信する。URL スイッチ 1004 は、キャッシュ制御サーバ 1008 が送信した FIN パケットを、キャッシュサーバ 1001-1 に転送する。

【0185】

キャッシュサーバ 1001-1 は、FIN パケットを受信すると、キャッシュ制御サーバ 1008 との間のセッションを終了し、FIN, ACK パケットを送信する。キャッシュ制御サーバ 1008 は、FIN, ACK パケットを受信すると、ACK パケットを送信する。このパケットの交換によって、キャッシュサーバ 1001-1 とキャッシュ制御サーバ 1008 との間のセッションが終了する。

【0186】

その後、キャッシュサーバ 1001-1 は、転送されたコンテンツの URL をキャッシュ制御サーバ 1008 に送る。キャッシュ制御サーバ 1008 は、受信した URL を、キャッシュ所在データベース 1305 に追加する。

【0187】

以上説明したように、第 2 の実施の形態では、URL スイッチ 1004 において URL のハッシュ値を用いて転送先を判定している。よって、従来、制御サーバにかかっていた負荷を URL スイッチ 1004 にオフロードすることができる。

【0188】

また、キャッシュサーバ 1001-1 で URL のハッシュ値を求め、URL スイッチ 1004 が当該ハッシュ値を用いてキャッシュ有無を判定するので、URL スイッチ 100

10

20

30

40

50

4の処理負荷を軽減することができる。また、第1の実施の形態と同様に、URLスイッチ1004が、URLの抽出に代わってハッシュ値を抽出することによって、URLスイッチ1004における処理を単純化することができる。

【0189】

(実施形態3)

第1及び第2の実施の形態では、サーバとクライアント端末間のセッションを確立した後に交換されるパケットにハッシュ値を埋め込んだ。以下に説明する第3の実施の形態では、TCPセッション確立時にハッシュ値を埋め込んだパケットを使用する。

【0190】

図24は、第3の実施の形態のTCPのヘッダフォーマットの説明図である。

10

【0191】

TCPにおけるセッションの確立時に送信されるパケットは、TCPヘッダの制御フィールドにセッション確立要求であることを示すSYNフラグが設定されている。第3の実施の形態では、SYNパケットのTCPオプション列の中に、要求するコンテンツのURLのハッシュ値を埋め込む。

【0192】

このため、第3の実施の形態を実現するためには、クライアント端末に実装されている既存のTCP STACKを改造する必要がある。よって、第1及び第2の実施の形態の方法と、第3の実施の形態の方法とは、利用形態によって使い分けることが望ましい。

【0193】

20

なお、第3の実施の形態は、第1の実施の形態のコンピュータシステム(図1)及び第2の実施の形態のコンピュータシステム(図10)のいずれにも適用することができる。

【0194】

次に、第3の実施の形態を、第1の実施の形態のコンピュータシステム(図1)に適用した場合の、コンテンツ要求の手順について説明する。

【0195】

図25は、第3の実施の形態のクライアント端末101-1等のコンテンツ要求処理のフローチャートである。

【0196】

クライアント端末101-1は、コンテンツを要求するURLが決定すると、URLのハッシュ値を計算する(ステップ711)。

30

【0197】

次に、計算したハッシュ値を、TCP SYNパケットのTCPオプション列に格納して、SYNパケットを作成する(ステップ712)。

【0198】

次に、クライアント端末101-1は、URLのハッシュ値が埋め込まれたSYNパケットを送信し、Webサーバ104-1等との間でセッションを確立する(ステップ713)。

【0199】

その後、コンテンツ要求パケットをURLスイッチ105に対して送信する(ステップ714)。その後、URLスイッチ105経由でサーバ104-1からコンテンツを受信する(ステップ715)。

40

【0200】

図26は、第3の実施の形態のURLスイッチ105のセッション中継処理のフローチャートである。このセッション中継処理は、クライアント端末101-1からのTCP SYNパケットを受信すると実行される。

【0201】

URLスイッチ105は、クライアント端末101-1からのSYNパケットを受信すると、当該SYNパケット内にハッシュ値が埋め込まれているか否かを判定する(ステップ811)。

50

【0202】

その結果、S Y N パケットにハッシュ値が存在する場合は、S Y N パケットからハッシュ値を抽出する（ステップ812）。

【0203】

次に、抽出されたハッシュ値をキーとして転送先検索テーブル407を検索し（ステップ813）、該当するエントリが存在するか否かを判定する（ステップ814）。

【0204】

その結果、該当エントリが発見された場合は、U R Lスイッチ105は検索されたエントリのW e b サーバ104-1を転送先に決定する（ステップ815）。

【0205】

その後、クライアント端末101-1からのS Y N パケットを転送先サーバに送信して、クライアント端末101-1とW e b サーバ104-1との間のセッションを確立する（ステップ816）。

【0206】

そして、クライアント端末101-1とW e b サーバ104-1との間のセッションが切断されるまでそのセッションを中継する（ステップ817）。

【0207】

ステップ811でS Y N パケットにハッシュ値が存在しない場合、又は、ステップ814で該当するエントリが発見されなかった場合は、予め定められたW e b サーバ104-1を転送先に決定する（ステップ818）。その後、クライアント端末101-1からのS Y N パケットを転送先サーバに送信して、クライアント端末101-1とW e b サーバ104-1との間のセッションを確立する（ステップ816）。

【0208】

図27は、第3の実施の形態のコンテンツ転送のシーケンス図である。

【0209】

クライアント端末101-1は、W e b サーバ104-1との間のセッションを確立するためにT C P の3-w a y ハンドシェイクを行う。

【0210】

具体的には、クライアント端末101-1は、W e b サーバ104-1に対して、U R Lのハッシュ値を埋め込んだS Y N パケットを送信する（図25のステップ713）。U R Lスイッチ105は、クライアント端末101-1からW e b サーバ104-1に対するS Y N パケットを捕捉する。

【0211】

そして、U R Lスイッチ105は、捕捉したS Y N パケットからハッシュ値を抽出し（図26のステップ812）、S Y N パケットの転送先を決定する（図26のステップ813～815、818）。そして、決定された転送先サーバにS Y N パケットを送信して（図26のステップ816）、クライアント端末101-1とU R Lスイッチ105との間のセッション確立を要求する。

【0212】

U R Lスイッチ105は、以後、クライアント端末101-1交換されるパケットを中継する。

【0213】

W e b サーバ104-1は、S Y N パケットを受信すると、セッションの確立が可能であれば、S Y N , A C K パケットを送信する。クライアント端末101-1は、S Y N , A C K パケットを受信すると、A C K パケットを送信する。このパケットの交換によって、クライアント端末101-1とU R Lスイッチ105との間のセッションが確立する。

【0214】

そして、クライアント端末101-1は、コンテンツ要求パケットを送信する。U R Lスイッチ105は、セッションを確立したW e b サーバ104-1に、受信したコンテンツ要求パケットを転送する。

10

20

30

40

50

【0215】

Webサーバ104-1は、コンテンツ要求パケットを受信すると、OKパケットを送信する。URLスイッチ105は、Webサーバ104-1が送信したOKパケットを、クライアント端末101-1に転送する。また、Webサーバ104-1は、コンテンツを送信する。

【0216】

その後、Webサーバ104-1は、コンテンツの送信が終了すると、FINパケットを送信する。URLスイッチ105は、Webサーバ104-1が送信したFINパケットを、クライアント端末101-1に転送する。

【0217】

クライアント端末101-1は、FINパケットを受信すると、Webサーバ104-1との間のセッションを終了し、FIN, ACKパケットを送信する。Webサーバ104-1は、FIN, ACKパケットを受信すると、ACKパケットを送信する。このパケットの交換によって、クライアント端末101-1とWebサーバ104-1との間のセッションが終了する。

【0218】

以上説明したように、第3の実施の形態では、SYNパケットのTCPオプション列の中に、要求するコンテンツのURLのハッシュ値を埋め込む。よって、URLスイッチ105は、TCPオプション列中のハッシュ値によって、セッション確立処理の時点でストリームの転送先のサーバを決定することができる。

【0219】

また、TCPのヘッダの解析は、TCPストリーム内のデータに含まれるハッシュ値の抽出よりさらに単純である。よって、URLスイッチ105における処理をさらに簡略にすることができる。

【図面の簡単な説明】

【0220】

【図1】第1の実施の形態のコンピュータシステムのブロック図である。

【図2】第1の実施の形態のクライアント端末のブロック図である。

【図3】第1の実施の形態のプロキシサーバのブロック図である。

【図4】第1の実施の形態のURLスイッチのブロック図である。

【図5】第1の実施の形態の転送先選択テーブルの説明図である。

【図6】第1の実施の形態のコンテンツ要求パケットの説明図である。

【図7】第1の実施の形態のクライアント端末のコンテンツ要求処理のフローチャートである。

【図8】第1の実施の形態のURLスイッチのセッション中継処理のフローチャートである。

【図9】第1の実施の形態のコンテンツ転送のシーケンス図である。

【図10】第2の実施の形態のコンピュータシステムのブロック図である。

【図11】第2の実施の形態のキャッシュサーバのブロック図である。

【図12】第2の実施の形態のコンテンツデータベースの説明図である。

【図13】第2の実施の形態のURLスイッチのブロック図である。

【図14】第2の実施の形態の転送先選択テーブルの説明図である。

【図15】第2の実施の形態のキャッシュ制御サーバのブロック図である。

【図16】第2の実施の形態のキャッシュ所在データベースの説明図である。

【図17】第2の実施の形態のクライアント端末からのコンテンツ要求パケットの説明図である。

【図18】第2の実施の形態のキャッシュサーバからのコンテンツ要求パケットの説明図である。

【図19】第2の実施の形態のキャッシュサーバのコンテンツ要求処理のフローチャートである。

10

20

30

40

50

【図20】第2の実施の形態のURLスイッチのコンテンツ要求処理のフローチャートである。

【図21】第2の実施の形態のキャッシュ制御サーバのコンテンツ要求処理のフローチャートである。

【図22】第2の実施の形態のコンテンツ転送のシーケンス図である。

【図23】第2の実施の形態のコンテンツ転送のシーケンス図である。

【図24】第3の実施の形態のTCPのヘッダフォーマットの説明図である。

【図25】第3の実施の形態のクライアント端末のコンテンツ要求処理のフローチャートである。

【図26】第3の実施の形態のURLスイッチのセッション中継処理のフローチャートである。 10

【図27】第3の実施の形態のコンテンツ転送のシーケンス図である。

【符号の説明】

【0221】

101 クライアント端末

102 ネットワーク

103 プロキシ

104 Webサーバ

105 URLスイッチ

1001 キャッシュサーバ

1002 Webクライアント端末

1003、1006、1007 ネットワーク

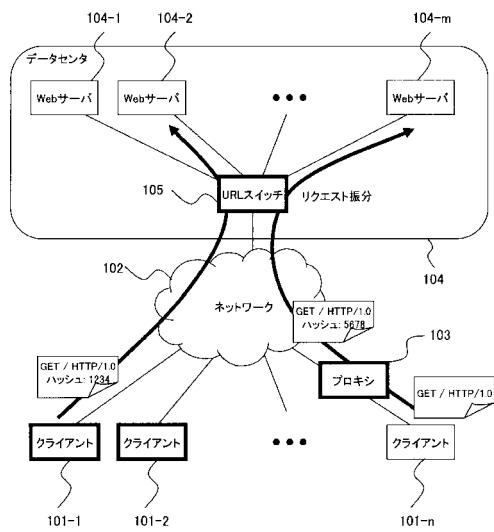
1004 URLスイッチ

1005 オリジンサーバ

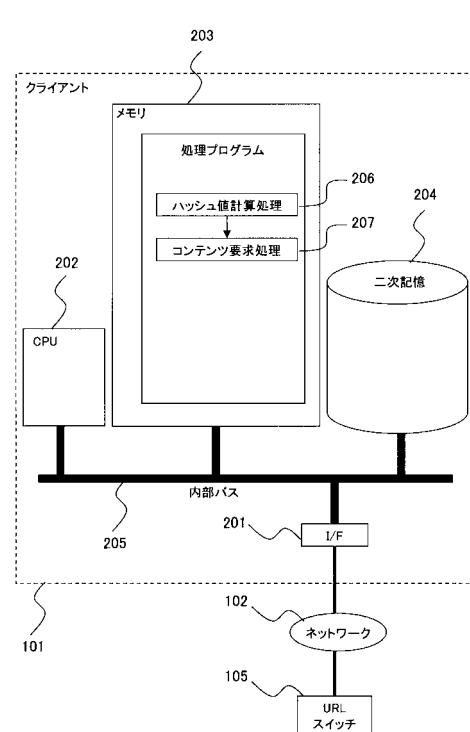
10

20

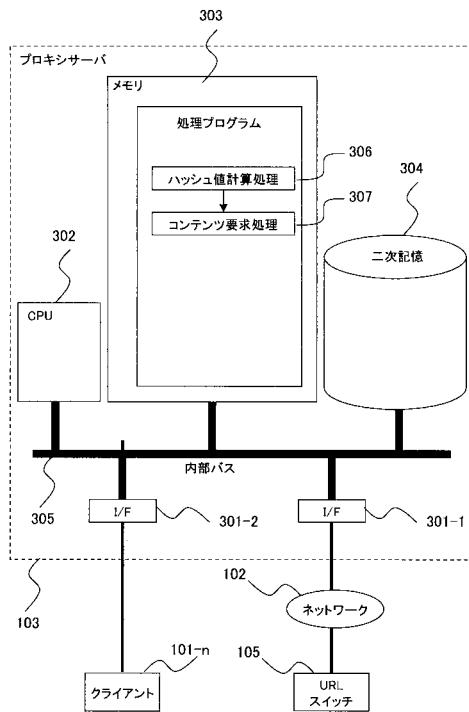
【図1】



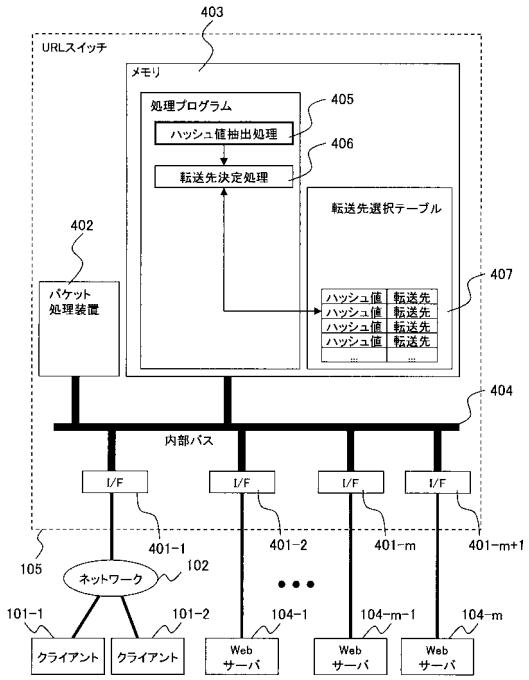
【図2】



【図3】



【図4】



【図5】

This diagram shows a "転送先選択テーブル" (Forwarding Destination Selection Table) (407) with the identifier 4071 above it and 4072 below it. The table consists of columns for "ハッシュ値" (Hash Value) and "転送先" (Forwarding Destination). The entries are:

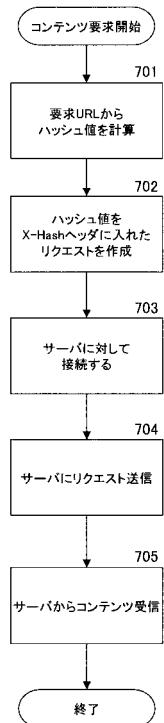
| ハッシュ値 | 転送先 |
|------------|---------------------------|
| 0x12345678 | Webサーバ1 (IP: 192.168.0.1) |
| 0x23456789 | Webサーバ1 (IP: 192.168.0.1) |
| 0x3456789A | Webサーバ2 (IP: 192.168.0.2) |
| ... | ... |

The table is labeled "転送先選択テーブル" at the bottom.

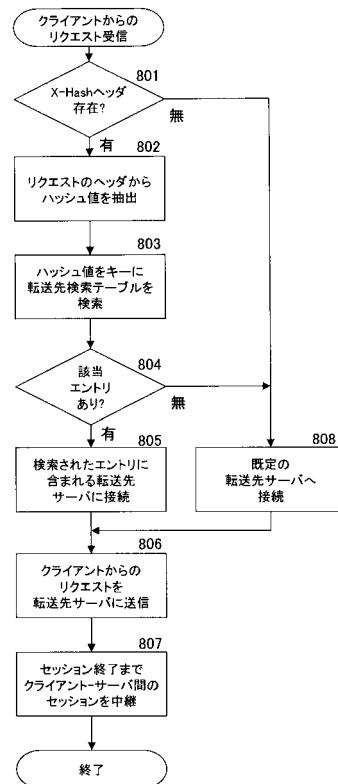
【図6】



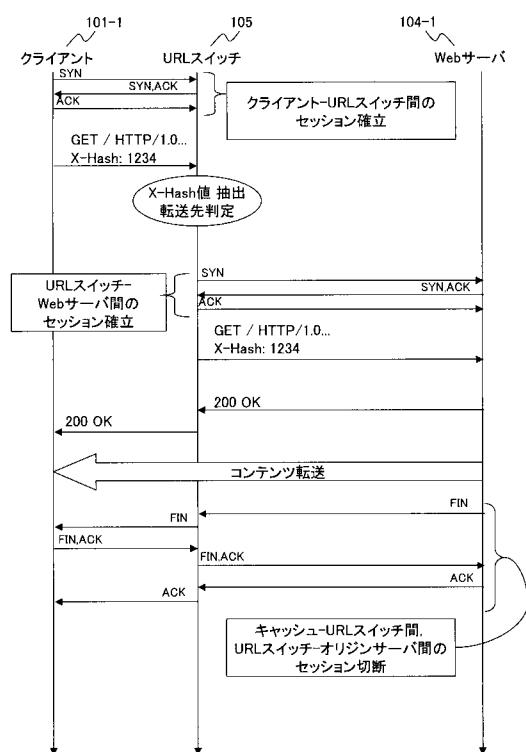
【図7】



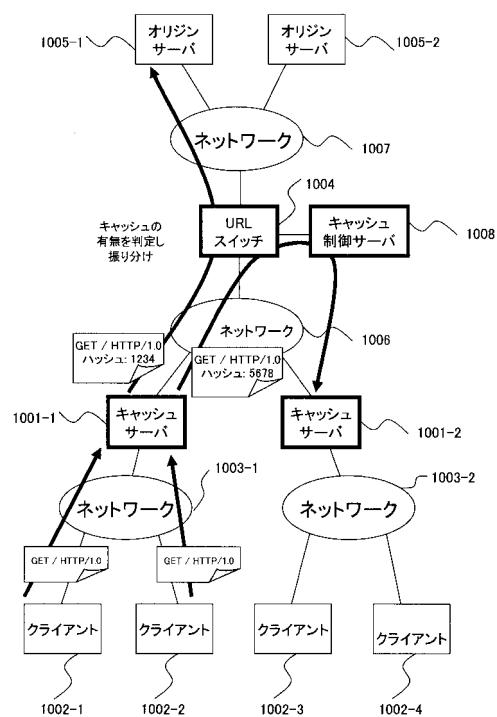
【図8】



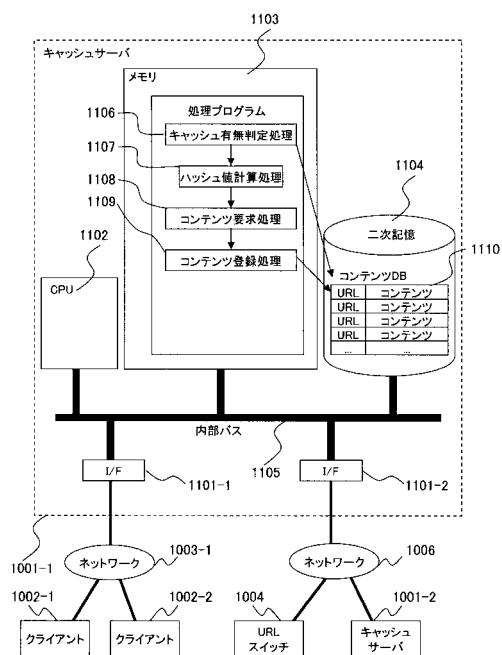
【図9】



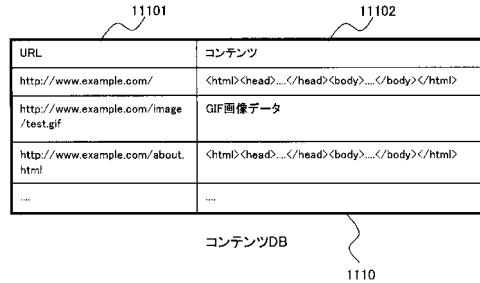
【図10】



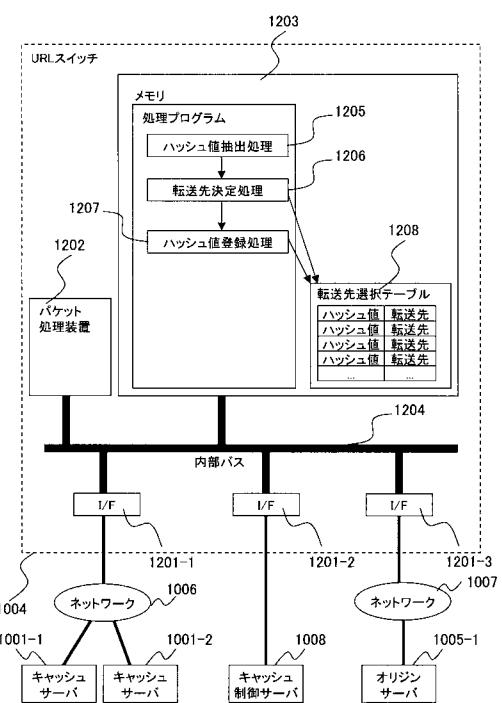
【図11】



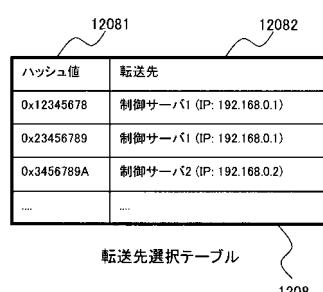
【図12】



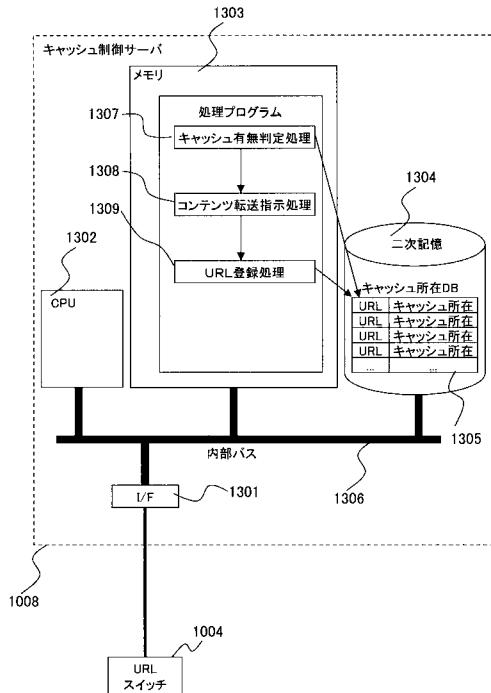
【図13】



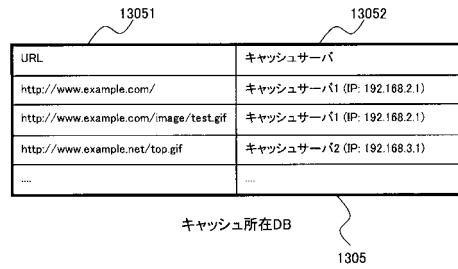
【図14】



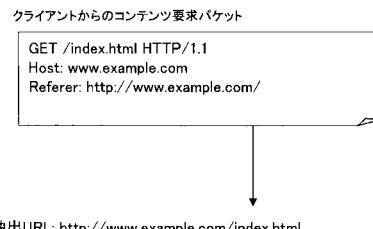
【図15】



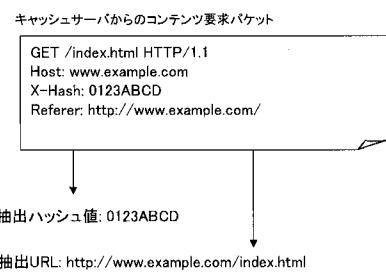
【図16】



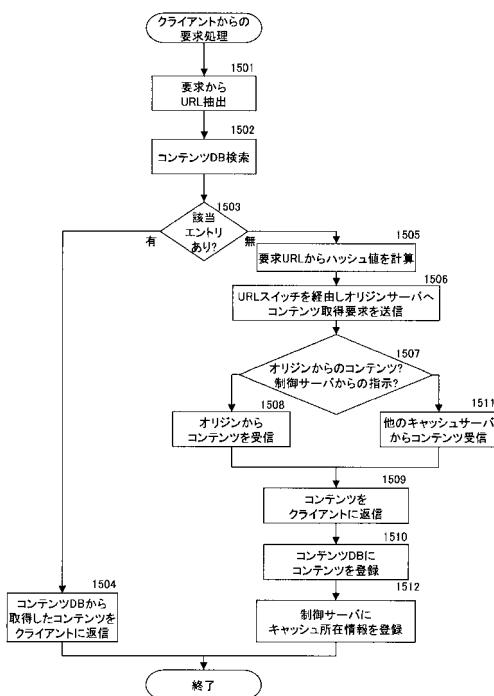
【図17】



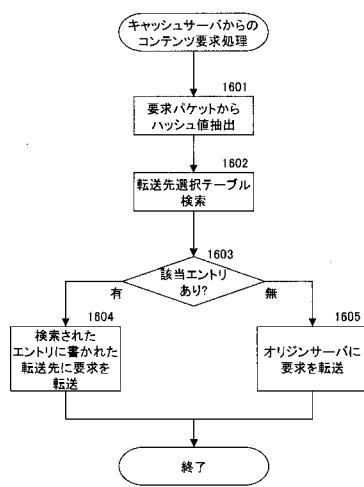
【図18】



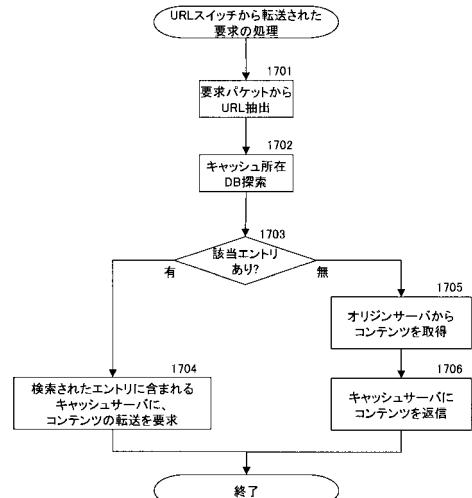
【図19】



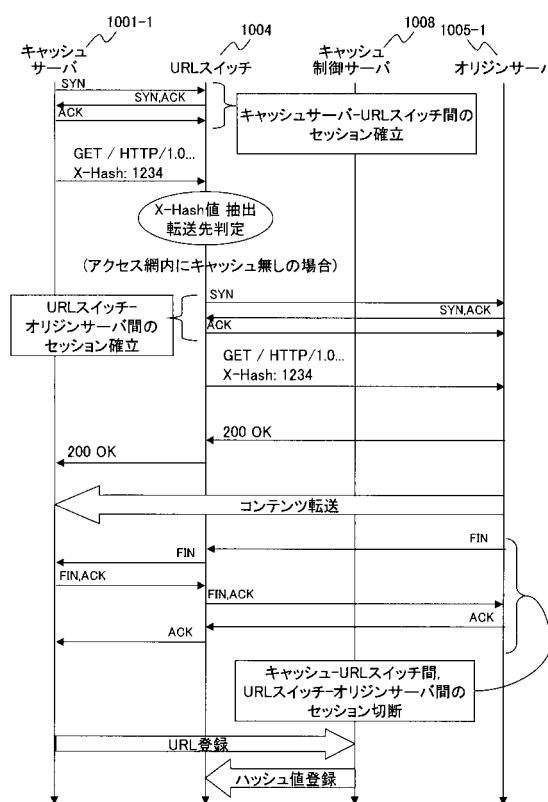
【 四 20 】



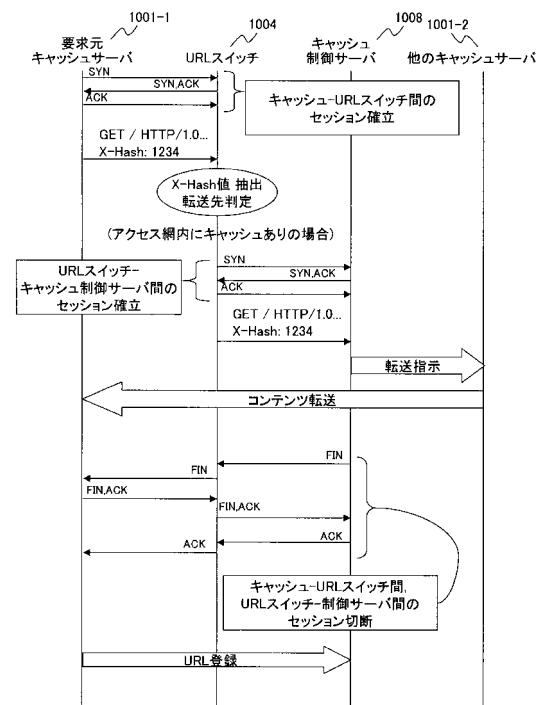
【 図 2-1 】



【习题 2.2】



〔 図 2 3 〕



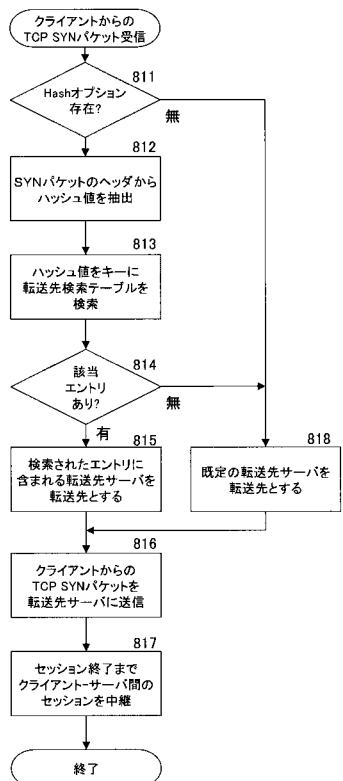
【図24】



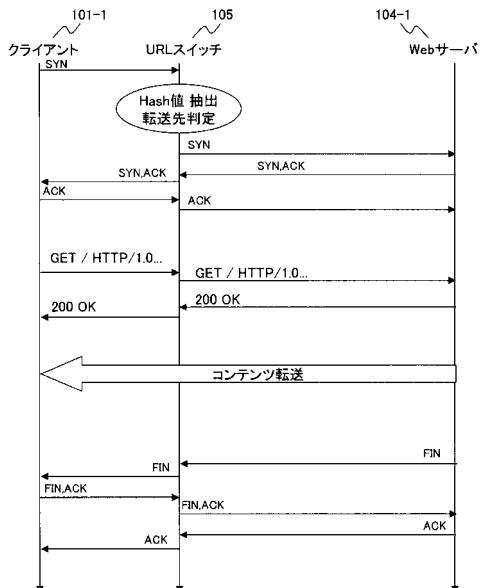
【図25】



【図26】



【図27】



フロントページの続き

(72)発明者 山本 淳二

東京都国分寺市東恋ヶ窪一丁目280番地 株式会社日立製作所 中央研究所内

(72)発明者 鈴木 敏明

東京都国分寺市東恋ヶ窪一丁目280番地 株式会社日立製作所 中央研究所内

審査官 千本 潤介

(56)参考文献 特開2004-118522(JP,A)

Z.Genova, K.J.Christensen , Using Signatures to mprove URL Routing , In proceedings of I
EEE International Performance, Computing, and Communications Conference ,米国 , IEEE ,
2002年 4月 , P45-52

(58)調査した分野(Int.Cl. , DB名)

G 06 F 13 / 00

G 06 F 9 / 50