

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4145449号
(P4145449)

(45) 発行日 平成20年9月3日(2008.9.3)

(24) 登録日 平成20年6月27日(2008.6.27)

(51) Int.Cl.		F I			
G 0 6 F	13/00	(2006.01)	G 0 6 F	13/00	3 5 7 Z
G 0 6 F	12/00	(2006.01)	G 0 6 F	12/00	5 4 6 A

請求項の数 17 (全 14 頁)

(21) 出願番号	特願平11-341724	(73) 特許権者	596092698
(22) 出願日	平成11年12月1日(1999.12.1)		ルーセント テクノロジーズ インコーポ レーテッド
(65) 公開番号	特開2000-194643(P2000-194643A)		アメリカ合衆国、07974-0636
(43) 公開日	平成12年7月14日(2000.7.14)		ニュージャージー、マレイ ヒル、マウン テン アヴェニュー 600
審査請求日	平成13年3月30日(2001.3.30)	(74) 代理人	100064447
審判番号	不服2005-1068(P2005-1068/J1)		弁理士 岡部 正夫
審判請求日	平成17年1月18日(2005.1.18)	(74) 代理人	100085176
(31) 優先権主張番号	09/201752		弁理士 加藤 伸晃
(32) 優先日	平成10年12月1日(1998.12.1)	(74) 代理人	100106703
(33) 優先権主張国	米国(US)		弁理士 産形 和央
		(74) 代理人	100096943
			弁理士 臼井 伸一

最終頁に続く

(54) 【発明の名称】 ウェブ資源を持続的に格納するための方法および装置

(57) 【特許請求の範囲】

【請求項1】

複数のバージョンを有する電子ドキュメントを持続的に格納するための方法であって、前記電子ドキュメントの前記複数のバージョンの一つを要請するリクエストであって、該リクエストされたバージョンの生成時間を示すタイムスタンプを含むリクエストを受信するステップ；

前記電子ドキュメントの前記リクエストされたバージョンを修正することで、前記電子ドキュメントの前記リクエストされたバージョンに埋め込まれたハイパーリンクを前記電子ドキュメントの前記リクエストされたバージョンのタイムスタンプを含むように更新するステップ；および

前記リクエストされた前記電子ドキュメントをクライアントに送信するステップを含むことを特徴とする方法。

【請求項2】

さらに、前記電子ドキュメントの前のバージョンが実質的に同一である場合、前記電子ドキュメントの前のバージョンへのポインタを格納するステップを含むことを特徴とする請求項1の方法。

【請求項3】

前記電子ドキュメントを識別するアドレスが前記タイムスタンプを含むことを特徴とする請求項1の方法。

【請求項4】

前記アドレスが均一資源ロケータ（“URL”）であることを特徴とする請求項3の方法。

【請求項5】

前記均一資源ロケータ（“URL”）が前記タイムスタンプを識別するための関連するリクエストヘッダを有することを特徴とする請求項4の方法。

【請求項6】

さらに、前記リクエストされたタイムスタンプを持つ前記電子ドキュメントのバージョンが存在しない場合、前記リクエストされたタイムスタンプの前の最も最近のタイムスタンプを持つ前記電子ドキュメントのバージョンを送信するステップを含むことを特徴とする請求項1の方法。

10

【請求項7】

前記リクエストがブラウザを用いて指定されることを特徴とする請求項1の方法。

【請求項8】

前記タイムスタンプが相対タイムスタンプであることを特徴とする請求項1の方法。

【請求項9】

複数のバージョンを持つ電子ドキュメントを格納するためのシステムであって、前記電子ドキュメントの前記複数のバージョンを電子ドキュメントのアーカイブ内に格納するためのメモリ；および

前記メモリに結合されたプロセッサを備え、このプロセッサが：

前記電子ドキュメントの前記複数のバージョンの一つを要請するリクエストであって、該リクエストされたバージョンの生成時間を示すタイムスタンプを含む、リクエストを受信し；

20

前記電子ドキュメントの前記リクエストされたバージョンを修正することで、前記電子ドキュメントの前記リクエストされたバージョンに埋込まれたハイパーリンクを前記電子ドキュメントの前記リクエストされたバージョンのタイムスタンプを含むように更新し；

前記リクエストされた前記電子ドキュメントをクライアントに送信する；
 ように構成されることを特徴とするシステム。

【請求項10】

前記プロセッサが、さらに、前記電子ドキュメントの前のバージョンが実質的に同一である場合、前記電子ドキュメントの前のバージョンへのポインタを格納するように構成されることを特徴とする請求項9のシステム。

30

【請求項11】

前記電子ドキュメントを識別するアドレスが前記タイムスタンプを含むことを特徴とする請求項9のシステム。

【請求項12】

前記アドレスが均一資源ロケータ（“URL”）であることを特徴とする請求項11のシステム。

【請求項13】

前記均一資源ロケータ（“URL”）が前記タイムスタンプを識別するための関連するリクエストヘッダを有することを特徴とする請求項12のシステム。

40

【請求項14】

前記リクエストがブラウザを用いて指定されることを特徴とする請求項9のシステム。

【請求項15】

前記プロセッサが、さらに、前記リクエストされたタイムスタンプを持つ前記電子ドキュメントのバージョンが存在しない場合、前記リクエストされたタイムスタンプの前の最も最近のタイムスタンプを持つ前記電子ドキュメントのバージョンを送信するように構成されることを特徴とする請求項9のシステム。

【請求項16】

前記タイムスタンプが相対タイムスタンプであることを特徴とする請求項9のシステム。

50

【請求項 17】

電子ドキュメントにアクセスするための装置であって、

その上に実現されたコンピュータにて読出可能なプログラムコード手段を持つコンピュータにて読出可能な媒体を備え、このコンピュータにて読出可能なプログラムコード手段が、コンピュータに対して：

前記電子ドキュメントの前記複数のバージョンの一つを要請するリクエストであって、該リクエストされたバージョンの生成時間を示すタイムスタンプを含む、リクエストを受信し；

前記電子ドキュメントの前記リクエストされたバージョンを修正することで、前記電子ドキュメントの前記リクエストされたバージョンに埋込まれたハイパーリンクを前記電子ドキュメントの前記リクエストされたバージョンのタイムスタンプを含むように更新し；

前記リクエストされた前記電子ドキュメントをクライアントに送信する；
ように指令するものであるコンピュータ読出可能媒体を備えることを特徴とする装置。

【発明の詳細な説明】**【0001】****【関連する特許出願】**

本発明は、本発明と同時に出願され、本発明の譲受人に譲渡された、以下の合衆国特許出願、つまり：“A Method And Apparatus For Persistent Access to Web Resources (ウェブ資源に持続的にアクセスするための方法および装置)” (Attorney Docket Number Ong 9)、 “A Method And Apparatus For Persistent Access to Web Resources Using Relative Time-Stamps (ウェブ資源に相対タイムスタンプを用いて持続的にアクセスするための方法および装置)” (Attorney Docket Number Ong 10)、および “A Method And Apparatus For Resolving Domain Names of Persistent Web Resources (持続的ウェブ資源のドメイン名を解決するための方法および装置)” (Attorney Docket Number Ong 11) とも関連するために、これらについても参照されたい。

【0002】**【発明の属する技術分野】**

本発明は、インターネット資源へのアクセス技法、より詳細には、インターネット資源へのパシスト(持続的)なアクセスを確保するための方法および装置に関する。

【0003】**【従来の技術】**

World Wide Web (“ウェブ(Web)”)は、莫大な量の情報を提出および配布するためのダイナミック(動的)な方法を提供する。ウェブ上の情報には、インターネットに接続されており、例えば、Netscape Communications Corporation of Mountain View, CA. から市販される Netscape Navigator Communicator™ などのブラウザを所有する人なら誰でもアクセスすることができる。ウェブは多数の媒体オプションをユーザに提供する。ウェブは、当初は、主としてパーソナルコンピュータを介してのアクセスに限られていたが、最近では、普遍的にますます多様なパーソナル電子デバイスから利用(アクセス)できるようになりつつある。加えて、ディスプレイ技術も絶えず進歩しており、ウェブは、終局的には、従来のペーパーベースの媒体に完全に取って代ることも予測される。

【0004】

ペーパーベースの媒体は、通常は、関連するタイムスタンプを持ち、過去のある任意の時点において提出された情報を簡単に調べることができる。例えば、新聞の記事は、その新聞の特定の出版日を指定することができる。信賴できるリファレンス(参照資料)として引用することができる。ただし、ウェブの内容(コンテンツ)は、ダイナミックな(動的な/可変的な)性格を持ち、このため、ウェブのドキュメントは、通常は、信賴できる参照源とはならない。つまり、現時点では、ウェブの内容は、将来の任意の時点で同一の形式にて存在し、これに、将来も引き続いて、同一の均一資源ロケータ(Uniform Resource Locator, “URL”)にて、持続的にアドレス(アクセス)できるという保証はない。

つまり、幾つかのウェブサイトは、幾らかのアーカイブ(長期保存)ウェブドキュメント

10

20

30

40

50

へのアクセスを提供するが、ただし、ユーザは、これら過去のウェブドキュメントに、たとえ可能であったとしても、一貫した予測可能なやり方でアクセスすることはできない。

【0005】

非営利のコンピュータライブラリサービスおよびリサーチ機関であるオンラインコンピュータライブラリセンタ(Online Computer Library Center, Inc., "OCLC")は、一般インターネット資源に対するインターネットアドレスおよび別名を管理するためのOCLC PURL(持続的均一資源ロケータ"Persistent Uniform Resource Locator")と呼ばれるソフトウェアツールを提供する。技術均一資源ロケータ(Persistent Uniform Resource Locator, PURL)は、インターネット資源に、高信頼に、長期間に渡って、最小の保守にて、アクセスできることを保証するために、インターネット資源に対するフレキシブルなネーミング(名前の指定)および名前分析サービスを提供する。概略的には、OCLC PURLは、ウェブ資源をどこに保存すべきかについての助言をインターネットユーザに与える。ただし、上述のように、インターネットは、絶えず、拡張および変化を続けている。そして、いったん、URLが変更されると、そのURLを用いての以前の全ての参照(アクセス)は無効となり、ユーザは、もはや、そのインターネット資源にはアクセスできなくなる。しかも、これら変更の管理は、しばしば、非常に煩雑となる。

【0006】

【発明が解決しようとする課題】

均一資源ロケータ(USR)は、インターネット資源の位置(ロケーション)を直接にポイントするが、持続的均一資源ロケータ(USR)は、PURLを実際のURLに翻訳する中間のアドレス解決(翻訳)サービスをポイントする。いったん、ウェブ資源がOCLCに登録され、PURLを割り当てられると、そのウェブ資源はPURLを介してアクセスすることが可能になる。PURLは、ある資源に、その資源の位置が変わった場合でも、パシステント(持続的)な名前を割り当てる。こうして、ウェブ資源および他の資源内で参照されるPURLは、時間がたっても持続的に有効となり、ウェブ資源が移動される度にその参照を更新する必要はなくなる(PURLは持続的なものとなり、ウェブ資源が移動された場合でも、同一のPURLを用いて同一のウェブ資源および他の資源の内容にアクセスすることが可能になる)。ただし、OCLCによって維持されるPURL"転送(forwarding)"アドレスは、常に、更新された状態に維持することを必要とされる。換言すれば、現状では、ドキュメントを移動する度に、OCLCに、そのドキュメントの新たなアドレスを通知することが要求される。

【0007】

【課題を解決するための手段】

本発明は、一般的には、ウェブ資源へのパシステント(持続的)なアクセスを提供するための方法および装置に関する。本発明の一面によると、ウェブ資源を識別するための"URL"が、オプションとして、タイムスタンプを含むように強化される。このタイムスタンプは、"URL"内に、任意の適当な形式にて指定することができる。ウェブブラウザおよびウェブサーバが開示されるが、これらは、タイムスタンプパラメータを収容し、ユーザが、明確な標的(ターゲット)時間にて、任意のウェブアドレスを参照することを可能にする。こうして、本発明は、ウェブ(資源)を、ペーパーベースの媒体と殆ど同様な、編成された、信頼できる参照源とすることを可能にする。

【0008】

開示されるウェブブラウザは、オプションとして、所望の日時の指定を容易にする機構を備えることも、あるいは、ユーザが、手操作にて、タイムスタンプを、URLに、ブラウザの"ロケーション(Location)"ウインドウに指定されるように付加できるようにすることもできる。一つの実施例によると、パシステント(持続的)ウェブサーバは、(1)タイムスタンプを含むURLを受信し、(2)アーカイブからリクエストされたウェブページに対応するベースウェブページを検索し、(3)ベースウェブページを修正することで、埋め込まれたハイパーリンクをリクエストされたウェブページと同一のタイムスタンプを含むように更新し、(4)更新されたページをリクエストされたページとしてクライア

10

20

30

40

50

ントに送り返す。パシステント（持続的）ウェブサーバは、抽出されたURLを選択されたタイムスタンプフォーマットに従って翻訳する。加えて、パシステントウェブサーバは、ウェブ資源の全てのバージョンを格納するためのパシステント（持続的）アーカイブを備え、ウェブユーザはこれら全てのバージョンに持続的にアクセスすることが可能になる。本発明は、任意のウェブ資源へのタイムスタンプを持つ参照が所望の資料を参照することを保証する。こうして、ウェブ上で歴史的な（過去の）調査を行なっている人が、任意の時間期間に対する（任意の時間期間に対して有効な）任意の情報を検索できるようにする。

【0009】

【発明の実施の形態】

図1は、インターネットあるいはWorld Wide Web（“ウェブ（Web）”）環境130を通じて一つあるいは複数のパシステント（持続的）ウェブサーバ140、150からの情報にアクセスする本発明によるウェブブラウザ100を示す。本発明は、テキスト、オーディオ、ビデオ、アニメーションドキュメントなどを含むウェブ資源、すなわち電子ドキュメントへのパシステントなアクセスを提供する（過去のドキュメントにも持続的にアクセスできるようにする）。本発明の一つの実施例においては、ウェブ資源を識別する均一資源ロケータ（“USR”）が、タイムスタンプを含むように強化される。ウェブブラウザ100およびパシステントウェブサーバ140、150は、追加のタイムスタンプパラメータを備え、こうして、ユーザは、特定の目標日付を持つ任意のウェブアドレスを参照することが可能となる（過去の任意の特定の日付のウェブ資源にアクセスすることが可能になる）。例えば、“http://cnn.com?time=2+2+1997”なるURLは、1997年2月2日付けのCNNホームページのウェブ内容（コンテンツ）を指定する。このようにして、ウェブは、編成された、しかも、ペーパーベースの媒体と同程度に信頼できる参照源となり得る。

【0010】

一例としてのタイムスタンプのフォーマット

タイムスタンプは、“URL”内に任意の適当なフォーマットにて含めることができる（この方法は当業者においては明白であると思われる）。例えば、www.Lucent.comのウェブページの1998年2月2日の時点での内容を参照する（にアクセスする）ためには、URLは、以下のように表される：

http://www.lucent.com?time=2+2+1998

もう一つのバリエーションとして、追加の時間粒度（より詳細な時間の指定）を、URL内に特定の日の時刻を含めることで、指定することもできる。例えば、www.Lucent.comのウェブページの1998年2月2日1:23.p.mの時点での内容にアクセスするためには、URLは、以下のように表される：

http://www.lucent.com?time=13+23+2+2+1998

特に改めて指定されない限り、時間帯は、ユーザの省略時の時間帯であるものと想定される。上述の一例としてのタイムスタンプのフォーマットは、Common Gateway Interface（CGI）による探索指数に従う。勿論、月、日、年（あるいは他の時間単位）は、任意の順番で表すことができる。URLがタイムスタンプを持たない場合は、省略時の値は、最も最新のバージョンとなる。

【0011】

加えて、相対タイムスタンプを用いることもできる。例えば、www.Lucent.comのウェブページの昨日の時点での内容にアクセスするためには、URLは、以下のように表される：

http://www.lucent.com?time=yesterday

さらに、埋め込まれたハイパーリンクが相対タイムスタンプを含む場合は、その相対タイムスタンプは現在のウェブページに基づくこととなる。つまり、現在のウェブページが：

http://www.lucent.com?time=1998_1_10

なる形式のURLを持ち、さらに、このページが：

http://www.lucent.com?time=yesterday

10

20

30

40

50

なる形式の埋め込まれたハイパーリンクを持つ場合、ブラウザおよびサーバは、そのURLを：

http://www.lucent.com?time=1998_1_9

なる形式を持つものと解釈および翻訳することとなる。

【0012】

他の相対タイムスタンプとして、現在のウェブページの時刻からの時間オフセットを含めることも、例えば、指定された時間期間を足したり、引いたりすることも考えられる。この方法では、例えば、“+10D”は、現時点で見られるウェブページの日付に10日を加えた時刻（日付）を示すこととなる。

【0013】

あるドキュメントより（現ドキュメントのタイムスタンプを起点として）前にあるいは後にアーカイブ（保存）されたバージョンを参照する（にアクセスする）ために、“next_archive”、あるいは“previous_archive”なるラベルを用いてURLを表すこともできる。この場合は、サーバは、アーカイブを通じて（内を）所望のドキュメントが見つかるまで探索することとなる。例えば、現在見ているドキュメントが：

http://www.lucent.com/doc.html?rttime=1997_5_12

なる形式のURLを持ち、同一のドキュメントの7/2/97、6/1/97、4/1/97および3/1/97なる年月日にアーカイブ（保存）された異なるバージョンが存在する場合は、現ドキュメント内に以下のリンクを用いることで、これら記録の一つを参照（一つにアクセス）することができる：

doc.html?rttime=next_archive

このような相対アーカイブタイムスタンプを用いると、同ドキュメントの異なるバージョン間での移動がより効率的になる。

【0014】

ここで、“リクエスト・ヘッダスキーム（request-header scheme）”と呼ばれるもう一つの実現においては、タイムスタンプは、HTTPリクエストヘッダの一つとして、例えば：

Time-Stamp:June 9,1998

【0015】

なる形式にて表される。

ここで、“専用文字スキーム（special character scheme）”と呼ばれるもう一つの実現においては、URL内に含まれるタイムスタンプを示すために、専用文字が以下のように用いられる：

<http://www.cnnfn.com;time=2+2+1998&timezone=server>

【0016】

ウェブサーバ

ウェブブラウザ100は、従来のブラウザ、例えば、Microsoft Internet Explorer™あるいはNetscape Navigator™として実現し、これを本発明の特徴および機能を組み込むように修正することが考えられる。後に詳細に説明するように、ウェブブラウザ100は、単に、ユーザが所望の日時を指定できるように、新たなオプション選択パネルを組み込むことのみを必要とされる。実際、従来のウェブブラウザ100を利用し、ユーザが、手操作にて、タイムスタンプをURLに、ブラウザ100の“Location”ウィンドウに従って付加するようにすることもできる。

【0017】

一つの実現においては、ユーザは、タイムスタンプをターンオンあるいはオフするオプションを有する。タイムスタンプが起動されている場合は、ブラウザ100は、URLに必要な変更を加えたうえで、URLをウェブ130に送る。対応するウェブサーバ140、150がタイムスタンプを認識できるという保証はなく、このため、サーバ140、150によって送り返されるドキュメントに含まれる埋め込まれたハイパーリンクがタイムスタンプを含まないことも考えられる。このため、このような状況においては、ウェブブラ

10

20

30

40

50

ウエブブラウザ 100 が、タイムスタンプオプションが起動されている場合は、ユーザがそのハイパーリンクをクリックした際に、埋め込まれているハイパーリンクと関連する URL を自動的に変換し、適当なタイムスタンプを付加するようにすることが考えられる。このため、ウエブブラウザ 100 は、URL を選択されたタイムスタンプフォーマットに従って変換できる必要がある。リクエスト・ヘッダスキームを用いる実現では、ブラウザ 100 を、専用のリクエストヘッダ（例えば、“Time-Stamp:June9,1998”）を送信できるように修正する必要がある。加えて、HTML を、特定のタイムスタンプを持つ任意の埋め込まれたハイパーリンクに対して、新たなタイムスタンプタグを含むように修正する必要がある。例えば：

```
<A HREF=" www.lucent.com">Lucent Web Site</A>
```

10

なる形式を持つハイパーリンクに対しては、HTML を、1998年2月2日なるタイムスタンプを示すように、以下のように修正する必要がある：

```
<A HREF=" www.lucent.com"><TIMESTAMP TIME="2+2+1998 "
```

```
TIMEZONE=server><TIMESTAMP>Lucent Web Site</A>
```

【0018】

パシステント（持続的）ウェブサーバ

パシステント（持続的）ウェブサーバ 140、150 は、従来のハードウェアおよびソフトウェアとして実現し、以下に説明する機能および動作を遂行するように修正することが考えられる。より具体的には、パシステントウェブサーバ 140、150 は、（1）タイムスタンプを含む URL を受信すること、（2）タイムスタンプを抽出すること、（3）アーカイブから正確なページを検索すること、（4）ベースウェブページを修正することで、埋め込まれたハイパーリンクをリクエストされたウェブページと同一のタイムスタンプを含むように更新すること、および（5）要求されたページをクライアントに送り返すことができる必要がある。さらに、パシステントウェブサーバ 140、150 は、抽出された URL を、選択されたタイムスタンプフォーマットに従って分析できる必要がある。本発明の一つの実施例においては、リクエストされた時間に対応するウェブ資源のバージョンが存在しない場合は、そのウェブ資源が以前にアーカイブ（保存）されたバージョンから変更されていないものと想定され、リクエストされた時間より前の最も最近の時間を持つウェブ資源のバージョンが送り返される。

20

【0019】

加えて、パシステントウェブサーバ 140、150 は、それらが扱ったウェブの過去の全ての情報を保存できる必要がある。このため、図 1 に示すように、各パシステントウェブサーバ、例えば、パシステントウェブサーバ 140、150 は、それぞれ、ウェブ資源の全てのバージョンを、それらがウェブユーザによって持続的に利用できるように、格納するためのパシステント（持続的）アーカイブ 145、155 を備える。これらパシステントアーカイブ 145、155 は、どのようなメモリデバイスを用いて実現することもできるが、ただし、好ましくは、CD-ROM、CR-R、WORM、DVD-ROM などの持続的な（消去不能な）メモリデバイスが用いられる。

30

【0020】

パシステントウェブサーバ 140、150 にて日付を持つ URL をサポートするためには、パシステントウェブサーバ 140、150 は、それらの全内容（コンテンツ）を、それらがいつでも取り出せるように、年代順に格納する必要がある。図 2 に示す一つの実現においては、パシステントアーカイブ 145、155 は、ウェブサイトの全内容を、永久メモリデバイスに、ある種の年代順ダイレクトリ構造に従って格納する。図 2 は、ウェブサイトの内容を年代順に配列するダイレクトリ構造 200 を示す。このダイレクトリ構造 200 内のおのおのの葉、例えば、葉 210 は、ある日付（ここでは、1998年6月26日）を持つ URL に対応する。例えば：

40

```
http://www.nytimes.com?time=24+2+1998
```

なる日付を持つ URL は、概念的に：

```
http://www.nytimes.com/archive/1998/2/24/
```

50

と等しい。

【 0 0 2 1 】

勿論、ウェブサイトの全内容を格納することは、メモリの利用の観点からは効率が悪い。多くのウェブページは、あったとしても、毎日、少ししか変更されない。このため、単にアーカイブから冗長を除去することで、メモリ効率を大幅に向上させることができる。いったん冗長が除去されれば、正規のウェブサイトメモリに追加されるべきメモリ要件は、通常は、それほど大きくはならない。

【 0 0 2 2 】

図 3 は、パシスレントアーカイブ 1 4 5、1 5 5 の冗長を低減するためのアーカイブプロセス 3 0 0 を示す。このアルゴリズムにおいて言及される全てのファイルあるいはサブダイレクトリは、図 2 の一例としてのダイレクトリ構造 2 0 0 のアーカイブサブダイレクトリ 2 2 0 の下に来ることに注意する。図 3 に示すように、アーカイブプロセス 3 0 0 は、最初に、ステップ 3 1 0 において、各サブダイレクトリ、例えば、サブダイレクトリ A に対して、それより以前に作成され、サブダイレクトリ A と同一の内容を持つサブダイレクトリ B が存在するか否かを決定するためのテストを遂行する。ステップ 3 1 0 において、それより以前に作成され、サブダイレクトリ A と同一の内容を持つサブダイレクトリ B は存在しないことが決定された場合は、サブダイレクトリレベルでは、パシスレントアーカイブ 1 4 5、1 5 5 の冗長を低減することはできず、プログラム制御は、ステップ 3 3 0 に進む。

【 0 0 2 3 】

他方、ステップ 3 1 0 において、それより以前に作成され、サブダイレクトリ A と同一の内容を持つサブダイレクトリ B が存在することが決定された場合は、ステップ 3 2 0 において、サブダイレクトリ A が、サブダイレクトリ B をポイントする別名（代替入口点 / 代替名）となる。例えば、図 4 A に示すように、ウェブサイトの現在（今日）の内容が昨日の内容と同一である場合は、今日に対して昨日のサブダイレクトリをポイントする別名が生成され、同様に、図 4 B に示すように、今月の内容が先月の内容と同一である場合は、今月に対して先月のサブダイレクトリをポイントする別名が生成される。

【 0 0 2 4 】

その後、ステップ 3 3 0 において、各ファイル、例えば、ファイル A に対して、それより以前に作成され、ファイル A と同一の内容を持つファイル B が存在するか否かを決定するためのテストが遂行される。ステップ 3 3 0 において、それより以前に作成され、ファイル A と同一の内容を持つファイル B は存在しないことが決定された場合は、ファイルのレベルでパシスレントアーカイブ 1 4 5、1 5 5 の冗長を低減することはできない。このため、プログラム制御は、ステップ 3 5 0 において終結する。

【 0 0 2 5 】

他方、ステップ 3 3 0 において、それより以前に作成され、ファイル A と同一の内容を持つファイル B が存在することが決定された場合は、ステップ 3 4 0 において、ファイル A がファイル B をポイントする別名となる。その後、プログラム制御は、ステップ 3 5 0 において終結する。

【 0 0 2 6 】

アーカイブプロセス 3 0 0 は、これが多数の（一致する）ファイルあるいはダイレクトリを探索する必要があるために、現実的でなくなることも考えられる。つまり、このプロセスの実行時間は、アーカイブ内のエンティティの数と共に指数関数的に増大することとなる。当業者においては明白なように、これに対する多くの次善の解決策が可能である。一つの極めて単純な解決策においては、単に、今日アーカイブしたいものが、最も最近に追加されたアーカイブ（例えば、昨日の内容）に対してチェックされる（と比較される）。殆どのウェブサイトは、以前にアーカイブされたものと僅かしか違わないために、このアプローチはかなり妥当なものとなる。このアプローチは、周知のファイルシステムの増分バックアップと類似する。

【 0 0 2 7 】

ウェブサーバがパシステント（持続的）ウェブサーバでない場合でも、その影響を最小限に押さえることが要求される。このために、一つの実現においては、リクエストがタイムスタンプを含むが、あるウェブサーバがそのタイムスタンプを認識できない場合、そのウェブサーバは、リクエストされたウェブ資源の最も最近のバージョンを送り返すようにされる。

【 0 0 2 8 】

パシステントアーカイブのメモリ要件を低減するためのもう一つの方法として、ウェブサーバを、正しいアーカイブされたデータを探索できるという観点で、よりスマートなものにすることが考えられる。例えば、あるウェブ資源のパシステントな（持続的な）記録は、そのウェブ資源の前に保存されているバージョンと幾らか異なるバージョンのみに制限することが考えられる。例えば、あるアーカイブがウェブ資源の次のような5つの異なるバージョン、つまり、6/4/1996、6/12/1996、3/23/1997、2/1/1998、2/3/1998を含み、リクエストされた日付がこれらアーカイブされているどのバージョンとも等しくない場合、ウェブサーバは、そのリクエストされた日付は、それより前の最も最近の日付を持つバージョンと同一であるものと想定することが考えられる。加えて、あるディレクトリ内に専用のシンボルリンク（あるいは、MacOS上の別名、あるいは、MS Windows上のショートカット）を用いることで、現在のディレクトリの下に見つけることができないファイルあるいはディレクトリを探索すべき場所を表し、こうして、該当するディレクトリの下には、変更された部分のみを格納することも考えられる。変更されていない全てのデータは、このような専用のリンクの連鎖を通じて参照することができる。

【 0 0 2 9 】

パシステント（持続的）DNSサーバ

パシステント（持続的）ドメイン名サーバ（domain name server、DNS）は、従来のハードウェアおよびソフトウェアとして実現し、以下に説明する機能および動作を遂行するように修正することが考えられる。従来のDNSサーバは、任意のドメイン名参照を、それらがDNSデータベース内に存在しない場合、拒否する。本発明による日付を持つURLの一つの長所は、これを（URLを）、過去のウェブ資源を参照（にアクセス）するためにも用いることができることである。例えば、従来の方法では、会社Aが会社Bに併合された場合、“www.A.com”を通じて参照（アクセス）される全てのウェブページはもはや無効となる。つまり、ユーザが会社Aに関するドキュメントにアクセスすることを希望する場合、ユーザは、全ての参照を会社Bのウェブサイトのどこかに変更することが必要となる。

【 0 0 3 0 】

ただし、会社Aの過去の情報は、DNSサーバが、会社Aの名前での参照（アクセス）を拒否することなく、代わりに、会社Aの過去の情報がどこに存在するかを知っているアーカイブサービス会社に相談（照会）するようになれば、アクセスすることが可能となる。DNSサーバ自身も、名前をIPアドレスに解決（翻訳）するプロセスをより迅速にするための幾らかの過去のデータを格納する。図5A～図5Bは、DNSサーバデータベースに格納されるデータの一例を、それぞれ、会社AとBが併合する前と後について示す。この方法では、図5Bに示すように、ユーザが併合の後に、www.A.comについて知りたいと希望した場合でも、DNSサーバは、ユーザのリクエストを会社Bと関連する新たなIPアドレスに再ダイレクトするための十分な情報を持つ。データベース内にリストされている日付は、対応するドメイン名が有効である期間を表す。図面の例では、“www.A.com. 2/2/1999”なる日付を持つドメイン名参照は無効であり、“www.A.com. 2/2/1992”なる日付の場合は有効であることが示される。

【 0 0 3 1 】

図6は、本発明によるDNSサーバプロセス600を示す。図6に示すように、DNSサーバプロセス600は、最初に、ステップ610において、ドメイン名リクエストを受信する。次に、ステップ620において、ドメイン名リクエストが日付を持つか否かを決定するためのテストが遂行される。ステップ620において、ドメイン名リクエストが日付

10

20

30

40

50

を持たないことが決定された場合は、ステップ 6 3 0 において、通常の名前探索結果が送り返される。

【 0 0 3 2 】

反対に、ステップ 6 2 0 において、ドメイン名リクエストが日付を持つことが決定された場合は、DNSサーバプロセス 6 0 0 は、ステップ 6 4 0 において、DNSデータベースを、そのドメイン名を求めて、その日付の制約内で探索する。次に、ステップ 6 5 0 において、その日付を持つドメイン名が見つかったか否か決定するためのテストが遂行される。ステップ 6 5 0 において、目的の日付を持つドメイン名が見つからなかったことが決定された場合は、DNSサーバは、ステップ 6 6 0 において、さらなる探索のために、アーカイブサービス会社に相談し、その後、プログラム制御は、ステップ 6 7 0 に進む。

10

【 0 0 3 3 】

一方、ステップ 6 5 0 において、目的の日付を持つドメイン名が見つかったことが決定された場合は、ステップ 6 7 0 において、探索結果が送り返される。このとき、再リダイレクトの場合は、その指標も送り返される。その後、プログラム制御は終結する。

【 0 0 3 4 】

ドメイン名がDNSサーバによって解決（翻訳）された後、ウェブブラウザ 1 0 0 は、DNSサーバから送り返されたものに従ってリクエストをウェブサーバ 1 4 0、1 5 0 に送信する必要がある。例えば、ユーザからのリクエストが、“http://www.A.com?time=2+2+1992”なる形式のURLである場合、ブラウザ 1 0 0 は、DNSサーバにドメイン名解決リクエストを“www.A.com 2/2/1992”なる形式で送信する。説明の例では、会社 A は、現在は会社 B の一部であるために、（図 5 B に示すように）結果として、“123.2.3.222 redirect”が送り返される。こうして、ウェブブラウザ 1 0 0 は、現在、サーバのIPアドレスを入手しており、それが、再ダイレクトアドレスであることも知っている。こうして、ウェブブラウザ 1 0 0 は、結果として、会社 B のウェブサーバ 1 4 0、1 5 0 にリクエストを“http://123.2.3.222?http://www.A.com?time=2+2+1992”なる形式にて送信する。会社 B のウェブサーバ 1 4 0、1 5 0 は、正しい（該当する）情報を得るために、どのようにすれば、会社 A のこの古いアドレスを適当な位置にマッピングすることができるか知っている。

20

【 0 0 3 5 】

この新たなDNSの一つの副次的な長所は、幾つかの名前は、それらが一旦過去のものとなったら、再使用できることである。例えば、会社 A という名前を持つもう一つの会社が、所定の期間の後に、DNSデータベースを以下のエントリにて更新することで、www.A.comなるドメイン名を用いることが可能となる：

30

【表 1】

URL	開始日	終日日	IP アドレス	再ダイレクト 先URL
WWW.A.COM	13/2/1980	1/5/1998	123.2.3.222	WWW.B.COM
WWW.A.COM	1/1/2000	PRESENT	234.2.2.12	--
WWW.B.COM	23/1/1985	PRESENT	123.2.3.222	--

40

【 0 0 3 6 】

このようにして、ドメイン名を、無駄にすることなく、再利用することが可能となる。

【 0 0 3 7 】

ダイナミックなウェブ内容へのパシステントなアクセス

こうして、ウェブは、ダイナミックな（動的に変化する）内容（コンテンツ）で一杯となる。これら内容（コンテンツ）には、WebCamなどからのリアルタイムビデオや、WebCast

50

イベントなどからのオーディオストリームや、Java、Javascript、あるいはActive-X起動のウェブページなどが考えられる。アプリケーションによっては、一連のダイナミックな内容を保存することは必要とされない、若しくは、要求されず、加えて、ウェブページの全ての広告部分は、保存することは必要とされない、若しくは、要求されない。

【0038】

例えば、銀行や証券会社の取引データなどの年代順のデータでは、記録の一部分のみを、与えられたタイムスタンプの制約内で抽出することができる。例えば、ある日付における残高をチェックするためには、ブラウザ100は、リクエストを以下のような形式にて送信する：

http://bank.com/?id=1234&time=2+3+1998

サーバ140、150は、単に、データを1998年3月2日まで検索あるいは再計算し、結果を送り返すことのみを必要とされる。このようなアプリケーション環境においては、全ての取引はもともとタイムスタンプを持つために、サービスにこの機能を追加することは極めて簡単である。

【0039】

リアルタイムな内容（コンテンツ）に対しては、タイムスタンプを付加することの唯一の制約は、メモリ要件である。保存すべき情報の量に対して多くのメモリ空間が利用できる場合は、ウェブサイトの管理者は、全てのリアルタイム内容を保存することを選択し、メモリ空間が十分でない場合は、それらの幾らかを、例えば、1日分、1週間分、あるいは1年分のデータを保存することを選択することが考えられる。

【0040】

ダイナミックに生成される宣伝に対しては、ウェブサイトの管理者は、（特別な理由により）古い広告を‘再提示（reshow）’することが適当であるか、あるいは、‘実際の（real）’アーカイブされているウェブ内容とは特に関係ない古い広告を新たな最新のコマercialと置換すべきか否かを決定することを要求される。

【0041】

ウェブページ上でダイナミックに実行されるスクリプトについては、2つの可能な状況が考えられる。そのスクリプトが時間とともに変化しない場合は、そのスクリプトは安全に保存することができるが、そのスクリプトの挙動が時間とともに変化する場合、例えば、クロックやカレンダーディスプレイである場合は、そのスクリプトは、日付を持つURLを扱うように修正することを必要とされ、Java、JavascriptあるいはActive-X標準に適切な修正を施すことが必要となる。CGIアプリケーションによって動的に生成されるウェブページについても同様である。

【0042】

以上、本発明について説明したが、説明の実施例およびバリエーションは単に本発明の原理を解説するためのものであり、当業者においては理解できるように、本発明の範囲および精神から逸脱することなく様々な修正が可能である。

【図面の簡単な説明】

【図1】本発明によるインターネット/World Wide Web（“ウェブ（Wed）”）環境を示す図である。

【図2】ウェブサイトの内容（コンテンツ）を年代順に配列するダイレクトリ構造を示す図である。

【図3】図1のパシステントアーカイブの冗長を低減するためにウェブサイト（サーバ）によって実現される一例としてのアーカイブプロセスの流れ図である。

【図4A】図3のアーカイブプロセスによって遂行される以前にアーカイブされた内容と同一のウェブサイトの内容に別名を付ける様子を示す図である。

【図4B】図3のアーカイブプロセスによって遂行される以前にアーカイブされた内容と同一のウェブサイトの内容に別名を付ける様子を示す図である。

【図5A】本発明によるDNSサーバのデータベースのサンプルテーブルを示す図である。

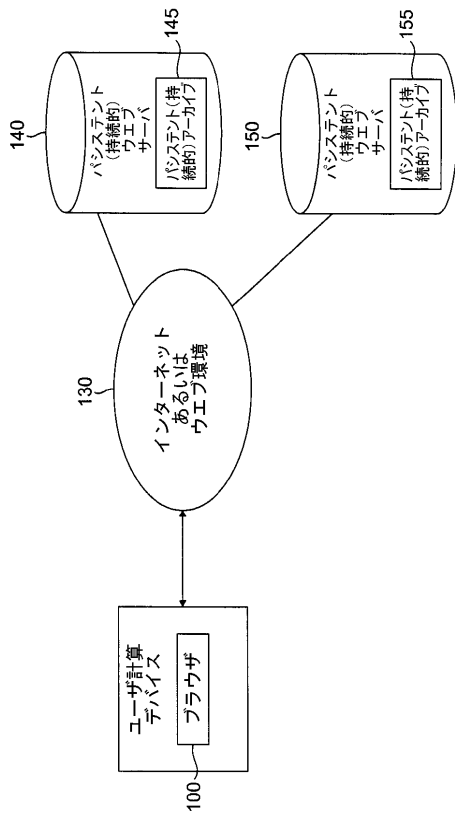
【図5B】本発明によるDNSサーバのデータベースのサンプルテーブルを示す図である。

【図6】タイムスタンプを持つURLに対応するIPアドレスを決定するためにDNSサーバによって遂行される一例としてのDNSサーバプロセスの流れ図である。

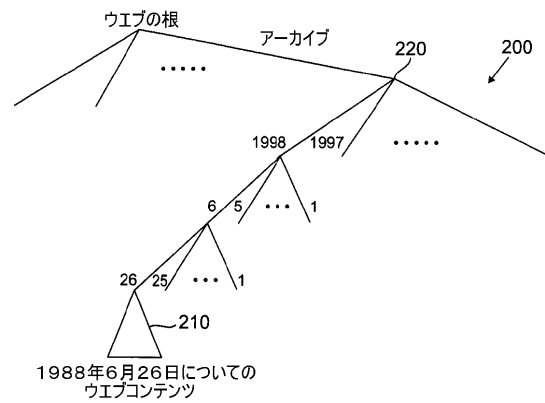
【符号の説明】

- 100 ウェブブラウザ
- 130 インターネット/World Wide Web (“ウェブ(Web)”)環境
- 140、150 パシステントウェブサーバ
- 145、155 パシステントアーカイブ
- 200 年代順ダイレクトリ構造

【図1】



【図2】



フロントページの続き

(74)代理人 100101498

弁理士 越智 隆夫

(74)代理人 100096688

弁理士 本宮 照久

(74)代理人 100104352

弁理士 朝日 伸光

(72)発明者 ピン - ウェン オン

アメリカ合衆国 07748 ニュージャージー, ミドルタウン, ローレル アヴェニュー 430

合議体

審判長 大日方 和幸

審判官 清水 稔

審判官 桑原 清

(56)参考文献 特開平9 - 297703 (JP, A)

Mike Burner 著, 「全世界のWebを保存する」, Internetworking
株式会社アスキー, 1997年8月1日発行, 第3巻第8号, p. 24 - 30

(58)調査した分野(Int.Cl., DB名)

G06F 12/00

G06F 13/00