



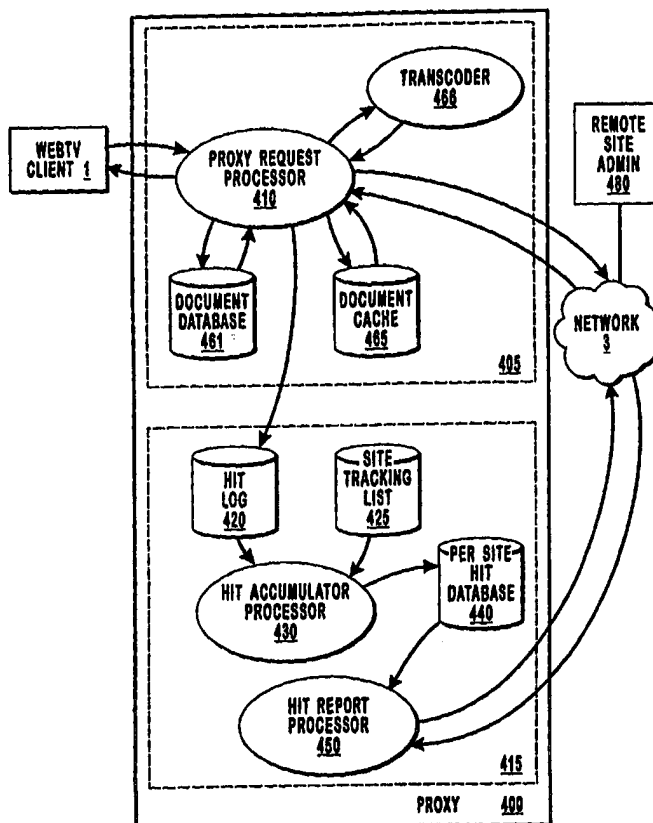
INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<p>(51) International Patent Classification <sup>6</sup> : <b>H04L</b></p>	<p><b>A2</b></p>	<p>(11) International Publication Number: <b>WO 98/45978</b> (43) International Publication Date: 15 October 1998 (15.10.98)</p>
<p>(21) International Application Number: PCT/US98/07120 (22) International Filing Date: 9 April 1998 (09.04.98) (30) Priority Data: 08/827,643 9 April 1997 (09.04.97) US (71) Applicant (for all designated States except US): WEBTV NETWORKS, INC. [US/US]; 305 Lytton Avenue, Palo Alto, CA 94301 (US). (72) Inventors; and (75) Inventors/Applicants (for US only): LOGUE, Jay, D. [-/US]; 3384 Olsen Drive, San Jose, CA 95117 (US), MIGHDOLL, Lee, S. [-/US]; 305 Lytton Avenue, Palo Alto, CA 94301 (US). (74) Agents: NYDEGGER, Rick, D. et al.; Workman, Nydegger &amp; Seeley, 1000 Eagle Gate Tower, 60 East South Temple, Salt Lake City, UT 84111 (US).</p>		<p>(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).</p> <p><b>Published</b> Without international search report and to be republished upon receipt of that report.</p>

(54) Title: METHOD AND APPARATUS FOR PROVIDING REMOTE SITE ADMINISTRATORS WITH USER HITS ON MIRRORED WEB SITES

(57) Abstract

A method and apparatus are disclosed for providing mirrored site administrators (480) with the number of hits from a proxy's document cache (465) and for dispatching document requests in a proxy (410) to more efficiently allocate the document cache space (465) within the proxy (410). A proxy (410) includes a document cache (465) storing recently requested documents, and maintains information regarding requests from the client (1) that are serviced from the proxy's document cache (465) such as the Uniform Resource Locator (URL) of the requested document and the number of cached responses. This information is provided by the proxy (410) to a remote site administrator (480). In this manner, remote site administrators (480) can more accurately track total hits. According to another aspect of the present invention, the proxy (410) implements a dispatching scheme for client requests that results in a more efficient allocation of the proxy's document cache space (465).



**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

<b>AL</b>	Albania	<b>ES</b>	Spain	<b>LS</b>	Lesotho	<b>SI</b>	Slovenia
<b>AM</b>	Armenia	<b>FI</b>	Finland	<b>LT</b>	Lithuania	<b>SK</b>	Slovakia
<b>AT</b>	Austria	<b>FR</b>	France	<b>LU</b>	Luxembourg	<b>SN</b>	Senegal
<b>AU</b>	Australia	<b>GA</b>	Gabon	<b>LV</b>	Latvia	<b>SZ</b>	Swaziland
<b>AZ</b>	Azerbaijan	<b>GB</b>	United Kingdom	<b>MC</b>	Monaco	<b>TD</b>	Chad
<b>BA</b>	Bosnia and Herzegovina	<b>GE</b>	Georgia	<b>MD</b>	Republic of Moldova	<b>TG</b>	Togo
<b>BB</b>	Barbados	<b>GH</b>	Ghana	<b>MG</b>	Madagascar	<b>TJ</b>	Tajikistan
<b>BE</b>	Belgium	<b>GN</b>	Guinea	<b>MK</b>	The former Yugoslav Republic of Macedonia	<b>TM</b>	Turkmenistan
<b>BF</b>	Burkina Faso	<b>GR</b>	Greece			<b>TR</b>	Turkey
<b>BG</b>	Bulgaria	<b>HU</b>	Hungary	<b>ML</b>	Mali	<b>TT</b>	Trinidad and Tobago
<b>BJ</b>	Benin	<b>IE</b>	Ireland	<b>MN</b>	Mongolia	<b>UA</b>	Ukraine
<b>BR</b>	Brazil	<b>IL</b>	Israel	<b>MR</b>	Mauritania	<b>UG</b>	Uganda
<b>BY</b>	Belarus	<b>IS</b>	Iceland	<b>MW</b>	Malawi	<b>US</b>	United States of America
<b>CA</b>	Canada	<b>IT</b>	Italy	<b>MX</b>	Mexico	<b>UZ</b>	Uzbekistan
<b>CF</b>	Central African Republic	<b>JP</b>	Japan	<b>NE</b>	Niger	<b>VN</b>	Viet Nam
<b>CG</b>	Congo	<b>KE</b>	Kenya	<b>NL</b>	Netherlands	<b>YU</b>	Yugoslavia
<b>CH</b>	Switzerland	<b>KG</b>	Kyrgyzstan	<b>NO</b>	Norway	<b>ZW</b>	Zimbabwe
<b>CI</b>	Côte d'Ivoire	<b>KP</b>	Democratic People's Republic of Korea	<b>NZ</b>	New Zealand		
<b>CM</b>	Cameroon		Republic of Korea	<b>PL</b>	Poland		
<b>CN</b>	China	<b>KR</b>	Republic of Korea	<b>PT</b>	Portugal		
<b>CU</b>	Cuba	<b>KZ</b>	Kazakstan	<b>RO</b>	Romania		
<b>CZ</b>	Czech Republic	<b>LC</b>	Saint Lucia	<b>RU</b>	Russian Federation		
<b>DE</b>	Germany	<b>LI</b>	Liechtenstein	<b>SD</b>	Sudan		
<b>DK</b>	Denmark	<b>LK</b>	Sri Lanka	<b>SE</b>	Sweden		
<b>EE</b>	Estonia	<b>LR</b>	Liberia	<b>SG</b>	Singapore		

## METHOD AND APPARATUS FOR PROVIDING REMOTE SITE ADMINISTRATORS WITH USER HITS ON MIRRORED WEB SITES

### FIELD OF THE INVENTION

5           The invention relates generally to the field of client-server computer networking. More specifically, the invention relates to mirroring of Web sites, notifying mirrored site administrators of hits, and allocation of the Web's content among mirroring servers based upon the Uniform Resource Locator (URL).

### BACKGROUND OF THE INVENTION

10           World Wide Web (Web) documents are commonly written in HTML (Hypertext Mark-up Language). HTML documents typically reside on Web servers and are requested by Web clients. Often delays can be introduced during Web browsing by heavy communications traffic on the Internet or slow response of a remote site, for example.

15           Providing one or more servers for mirroring Web sites located on remote servers is one means of reducing delays involved with browsing the Web. These mirroring servers, typically referred to collectively as a "proxy" or individually as "proxy servers," store frequently accessed Web sites in a local cache, thereby eliminating recurrent retrievals of commonly accessed documents. Thus, when a request for a particular Web page is

20           received from a client, the proxy server associated with the particular client looks first to its local cache to service the request rather than the remote site upon which the Web page resides. If the requested document is found locally, the request can be serviced by the proxy server and a subsequent request to the remote server for the document can be avoided. Therefore, only when a valid copy of the requested document is not in the

25           proxy's local cache would the remote server need to be accessed. In this manner, exposure to heavy communications traffic on the Internet and slow response of remote serves can be reduced.

30           While this mirroring approach is beneficial to end-users, it makes hit tracking for remote site administrators difficult. A hit is a request for a Web page, typically initiated by a user selecting a hypertext link for the Web page. The mirroring approach discussed above disrupts a remote server's ability to track the total number of requests for a given Web page because, as discussed above, some of the requests are intercepted and serviced by proxy servers. It is desirable to have an accurate count of requests for a given Web page or group of pages to track the relative popularity of a page, for example or to

35           provide feedback to advertisers whose advertisements appear on the page. Therefore, what is needed is a mechanism for tracking user hits by the proxy and a mechanism for

notifying mirrored sites, thereby allowing remote site administrators to accurately track total hits (i.e., those requests serviced from a proxy's local cache and the requests serviced by the remote server).

Another problem with the current mirroring approach is the inefficient allocation of the proxy's cache space. Currently, each client is assigned to one or more proxy servers. Therefore, the documents most recently requested by each active client will reside in the corresponding proxy server's cache. Assuming one or more clients assigned to different proxy servers have requested the same document recently, the same document might be cached in several of the proxy servers, thereby reducing the cache storage space for other frequently requested documents. Further, one or more extremely popular documents might potentially be cached in each proxy server. While redundancy of information is useful for fault tolerance, organized redundancy would be preferable. Given the foregoing, what is needed is a means of more efficiently allocating cache space within a proxy. Specifically, it would be desirable to allocate mutually exclusive portions of the Web's content to particular proxy servers.

#### SUMMARY OF THE INVENTION

A method and apparatus are described for providing mirrored site administrators with the number of hits from a proxy's document cache and for dispatching document requests in a proxy to more efficiently allocate the document cache space within the proxy. A proxy includes a document cache storing recently requested documents. The proxy is coupled to a client and to a remote server. The proxy maintains information regarding requests from the client that are serviced from the proxy's document cache. This information is provided by the proxy to a remote site administrator. In this manner, remote site administrators can more accurately track total hits (i.e., those requests serviced from a proxy's document cache plus the requests serviced by the remote server itself).

According to another aspect of the present invention, a proxy implements a dispatching scheme for client requests that results in a more efficient allocation of the proxy's document cache space. The proxy receives a document request from a client. A Uniform Resource Locator (URL) is included in the document request. The proxy forwards the request to one of a plurality of proxy servers based upon the URL.

According to another aspect of the present invention, the proxy performs a hash function on the URL that maps the URL to exactly one of the plurality of proxy servers. Advantageously, in this manner, mutually exclusive portions of the Web's content can be allocated to particular proxy servers.

### BRIEF DESCRIPTION OF THE DRAWINGS

The present invention is illustrated by way of example, and not by way of limitation, in the figures of the accompanying drawings and in which like reference numerals refer to similar elements and in which:

5           Figure 1 is a block diagram illustrating several clients connected to a proxy server in a network.

          Figure 2 is a diagram illustrating a client according to one embodiment of the present invention.

10           Figure 3 is a block diagram of a server according to one embodiment of the present invention.

          Figure 4 is a data flow diagram illustrating the interaction of proxy components according to one embodiment of the present invention.

          Figure 5A is a depiction of an exemplary site tracking list according to one embodiment of the present invention.

15           Figure 5B is a depiction of an exemplary per site hit database according to one embodiment of the present invention.

          Figure 6 is a logical view of an exemplary directory structure of a remote server.

          Figure 7 is a flow diagram illustrating a method of performing hit accumulation according to one embodiment of the present invention.

20           Figure 8 is a flow diagram illustrating a method of hit reporting according to one embodiment of the present invention.

          Figure 9 is a data flow diagram illustrating the interaction of proxy components according to another embodiment of the present invention.

25           Figure 10 is a flow diagram illustrating a method of dispatching requests to segregate the storage of documents according to one embodiment of the present invention.

### DETAILED DESCRIPTION

30           A method and apparatus are described for providing mirrored site administrators with the number of hits from a proxy's document cache and for maintaining a more efficient document caching scheme in a client-server computer network. In the following description, for purposes of explanation, numerous specific details are set forth in order to provide a thorough understanding of the present invention. It will be evident, however, to one skilled in the art that the present invention may be practiced without these specific  
35           details. Further, in other instances, well-known structures and devices are shown in block diagram.

The present invention includes various steps, which will be described below. The steps can be embodied in machine-executable instructions, which can be used to cause a general-purpose or special-purpose processor programmed with the instructions to perform the steps. Alternatively, the steps of the present invention might be performed by specific hardware components that contain hardwired logic for performing the steps, or by any combination of programmed computer components and custom hardware components.

While embodiments of the present invention will be described with respect to HTML documents, the method and apparatus described herein are equally applicable to other types of documents such as text files, images (e.g., JPEG and GIF), audio files (e.g., .WAV, .AU, and .AIFF), video files (e.g., .MOV, and .AVI), and other document types commonly found on the Web.

#### System Overview

The present invention may be included in a system, known as WebTV™, for providing a user with access to the Internet. A user of a WebTV™ client generally accesses a WebTV™ server via a direct-dial telephone (POTS, for "plain old telephone service"), ISDN (Integrated Services Digital Network), or other similar connection, in order to browse the Web, send and receive electronic mail (e-mail), and use various other WebTV™ network services. The WebTV™ network services are provided by WebTV™ servers using software residing within the WebTV™ servers in conjunction with software residing within a WebTV™ client.

Figure 1 illustrates a basic configuration of the WebTV™ network according to one embodiment. A number of WebTV™ clients 1 are coupled to a modem pool 2 via direct-dial, bi-directional data connections 29, which may be telephone (POTS, i.e., "plain old telephone service"), ISDN (Integrated Services Digital Network), or any other similar type of connection. The modem pool 2 is coupled typically through a router, such as that conventionally known in the art, to a number of remote servers 4 via a conventional network infrastructure 3, such as the Internet. The WebTV™ system also includes a WebTV™ server 5, which specifically supports the WebTV™ clients 1. The WebTV™ clients 1 each have a connection to the WebTV™ server 5 either directly or through the modem pool 2 and the Internet 3. Note that the modem pool 2 is a conventional modem pool, such as those found today throughout the world providing access to the Internet and private networks.

Note that in this description, in order to facilitate explanation the WebTV™ server 5 is generally discussed as if it were a single device, and functions provided by the

WebTV™ services are generally discussed as being performed by such single device. However, the WebTV™ server 5 may actually comprise multiple physical and logical devices connected in a distributed architecture, and the various functions discussed below which are provided by the WebTV™ services may actually be distributed among multiple WebTV™ server devices.

#### An Exemplary Client System

Figure 2 illustrates a WebTV™ client 1. The WebTV™ client 1 includes an electronics unit 10 (hereinafter referred to as "the WebTV™ box 10"), an ordinary television set 12, and a remote control 11. In an alternative embodiment of the present invention, the WebTV™ box 10 is built into the television set 12 as an integral unit. The WebTV™ box 10 includes hardware and software for providing the user with a graphical user interface, by which the user can access the WebTV™ network services, browse the Web, send e-mail, and otherwise access the Internet.

The WebTV™ client 1 uses the television set 12 as a display device. The WebTV™ box 10 is coupled to the television set 12 by a video link 6. The video link 6 is an RF (radio frequency), S-video, composite video, or other equivalent form of video link. In the preferred embodiment, the client 1 includes both a standard modem and an ISDN modem, such that the communication link 29 between the WebTV™ box 10 and the server 5 can be either a telephone (POTS) connection 29a or an ISDN connection 29b. The WebTV™ box 10 receives power through a power line 7.

Remote control 11 is operated by the user in order to control the WebTV client 1 in browsing the Web, sending e-mail, and performing other Internet-related functions. The WebTV™ box 10 receives commands from remote control 11 via an infrared (IR) communication link. In alternative embodiments, the link between the remote control 11 and the WebTV™ box 10 may be RF or any equivalent mode of transmission.

#### An Exemplary Server System

The WebTV™ server 5 generally includes one or more computer systems generally having the architecture illustrated in Figure 3. It should be noted that the illustrated architecture is only exemplary; the present invention is not constrained to this particular architecture. The illustrated architecture includes a central processing unit (CPU) 50, random access memory (RAM) 51, read-only memory (ROM) 52, a mass storage device 53, a modem 54, a network interface card (NIC) 55, and various other input/output (I/O) devices 56. Mass storage device 53 includes a magnetic, optical, or other equivalent storage medium. I/O devices 56 may include any or all of devices such

as a display monitor, keyboard, cursor control device, etc. Modem 54 is used to communicate data to and from remote servers 4 via the Internet.

As noted above, the WebTV™ server 5 may actually comprise multiple physical and logical devices connected in a distributed architecture. Accordingly, NIC 55 is used to provide data communication with other devices that are part of the WebTV™ services. Modem 54 may also be used to communicate with other devices that are part of the WebTV™ services and which are not located in close geographic proximity to the illustrated device.

### An Exemplary Proxy

Figure 4 illustrates the caching and hit accumulation features of the WebTV™ proxy 400 according to one embodiment of the present invention. In this embodiment, one or more WebTV™ servers 5 may act as a proxy 400 in providing the WebTV™ client 1 with access to the Web and other WebTV™ services. More specifically, WebTV™ server 5 functions as a "caching proxy." In this example, proxy 400 includes a proxy server 405 and a hit accumulator server 415. Client requests that are serviced from the proxy server's local document cache 465 are communicated to the hit accumulator server 415. AS will be described below, the hit accumulator server 415 maintains and organizes the data so as to provide hit tracking information to remote site administrators such as remote site administrator 480. Remote site administrator 480 may include entities such as persons authorized to gather statistical data for the remote site, persons authorized to manage and maintain the remote site the site, the remote site itself, or an automated computer system or other device configured to receive statistical data for the remote site.

In this embodiment, the proxy server 405, includes a proxy request processor 410, a document cache 465, a document database 461, and a transcoder 466. The proxy request processor 410 receives requests from the WebTV™ client 1 and sends responses to the WebTV™ client 1. The proxy request processor 410 maintains the document database 461, the document cache 465, and further determines when transcoding will be performed. The document cache 465 is used for temporary storage of Web documents such as images, text files, audio files, video files and other information which is used frequently by either WebTV™ client 1 or the proxy server 405.

When a document request is received from a client, the proxy request processor 410 determines whether to service the request from the document cache 465 by performing a search of the document cache 465. If the document is found locally, then the document may be retrieved from the document cache 465 and transferred to the client with the response. However, if the requested document is not found, then the proxy



request processor 410 requests the document from the appropriate site and upon receipt the proxy request processor 410 provides the document to the client with the response. Further, the proxy request processor 410 anticipates subsequent requests by storing the document in the document cache 465.

5           When a document is retrieved by the proxy server 405 from a remote server 4, for example, detailed information on this document may be stored in the document database 461. The stored information may subsequently be used by the proxy server 405 to speed up processing and downloading of that document in response to future requests for that document. In addition, the transcoding functions and various other functions of the  
10 WebTV™ service may be facilitated by making use of information stored in the document database 461. For example, the document database 461 may include certain historical and diagnostic information for Web pages that have been accessed by a WebTV™ client 1.

Document transcoder 466 is used to automatically revise the code of Web documents retrieved from the remote servers 4, for purposes such as: (1) correcting bugs  
15 in documents; (2) correcting undesirable effects which occur when a document is displayed by the client 1; (3) improving the efficiency of transmission of documents from the server 5 to the client 1; (4) matching hardware decompression technology within the client 1; (5) resizing images to fit on the television set 12; (6) converting documents into other formats to provide compatibility; (7) reducing latency experienced by a client 1  
20 when displaying a Web page with in-line images (images displayed in text); and, (8) altering documents to fit into smaller memory spaces.

In one embodiment, hit accumulator server 415 may act as a Web server providing a Hypertext Transport Protocol (HTTP) interface by which remote site administrators can access the accumulated hits for their sites by way of a Web browser. The hit accumulator  
25 server 415 may include a hit log 420, a hit accumulator processor 430, a site tracking list 425, a hit report processor 450, and a per site hit database 440. One method of communicating hits from a given proxy server to the hit accumulator server 415 is through a common storage device such as hit log 420. This and other methods of communicating hits will be described below. Regardless of how hits are communicated to the hit  
30 accumulator server 415, a process such as the hit accumulator processor 430 is desirable to verify the hits against a list of locations that are to be monitored. Such a list of locations may be stored in the site tracking list 425, for example. A location, in this context, refers to the location of a document. The location may be represented by a URL, a directory path, or other mechanisms for uniquely identifying a particular document. Hits that are  
35 validated by the hit accumulator processor 430 are recorded in the per site hit database 440. Thus, the per site hit database 440 will have a current count of the hits for each

location listed in the site tracking list 425. In this embodiment, the hit report processor 450 may receive requests from remote site administrators such as remote site administrator 480 for hit reports. The hit reports can be extracted from the per site hit database 440 and transmitted to the requester in an HTML report, for example.

5 While in this embodiment the proxy server 405 and the hit accumulation server 415 have been shown as separate servers, the functionality of both could be combined into one WebTV™ server 5. Additionally, the proxy 400 might be expanded to include more than one proxy server 405. When expanding the proxy 400 to include more than one proxy server 405, only one hit accumulation server 415 need be employed.

10 In alternative embodiments, hits may be communicated by a proxy server 405 to the accumulation server 415 by way of a network connection such as permanent connection through which events may be sent. Also, message passing may be employed whereby the proxy server 405 sends a message such as a datagram to the hit accumulator 415 to notify it of a document cache hit. It is appreciated that many other means of  
15 communicating information between servers are possible.

#### An Exemplary Site Tracking List

Figure 5A illustrates an exemplary site tracking list according to one embodiment of the present invention. This illustration depicts a site tracking list 435 including site  
20 tracking list records 505 for three remote sites: (1) <http://www.companyA.com/>; (2) <http://www.companyB.com/>; and (3) <http://www.companyC.com/>. In this embodiment, each site tracking list record 505 may include a list of one or more URL patterns 510.

The list of URL patterns 510 may be a list of strings identifying the initial portions (e.g., prefixes) of URLs to be tracked. In this example, the proxy 400 tracks hits for  
25 documents identified by URLs with a prefix that matches any of the URL patterns 510 specified in one of the site tracking list records 505. The hits may then be logged to a record in the per site hit database 440 corresponding to the site tracking list record 505 which contained the matching URL pattern. This form of URL pattern is useful for tracking hits for a particular grouping of Web pages beginning with the same initial  
30 sequences of characters. For example, the URLs for the Web pages of Company A might all begin with "[http://www.company\\_A.com/](http://www.company_A.com/)." Additionally, the Web pages associated with products produced by Company A might all begin with the sequence "[http://www.company\\_A.com/product/](http://www.company_A.com/product/)." Furthermore, pages related to a particular product, might all begin with the URL prefix

35 "[http://www.company\\_A.com/product/<product\\_name>](http://www.company_A.com/product/<product_name>)/" where <product\_name> identifies the particular product. To track the hits for pages relating to Company A's

Gizmo product line, therefore, the following URL pattern may be used:

"http://www.company\_A.com/product/Gizmo/." Similarly, to track the hits for all of Company A's products the following URL pattern may be used:

"http ://www.company\_A.com/product/."

5 URL patterns are not limited to prefixes, other forms of URL patterns may be used such as patterns including wild card or other special characters, or patterns in the form of standard regular expressions.

#### An Exemplary per Site Hit Database

10 Figure 5B illustrates an exemplary per site hit database according to one embodiment of the present invention. Based upon the information provided in the site tracking list 425 of Figure 5A, an exemplary per site hit database might be represented as per site hit database 440. In this example, the per site hit database 440 includes three site hit records 515 corresponding to remote sites for CompanyA, CompanyB and CompanyC.

15 In this embodiment, each site hit record 515 includes a timestamp 525. The timestamp 525 may indicate the time from which the hits have been accumulated. In this example, therefore, there have been six hits to the monitored URLs since January 16, 1997 at 10:01:58. Those of skill in the art will appreciate the timestamp 525 may represent the period of accumulation in other ways such as elapsed time since the last hit report was generated.

20 Site hit records 515 also include a remote site name 530. The remote site names 530 from front to back correspond to CompanyA, CompanyB, and CompanyC. Site hit record 515 further includes a list of hits 520. In this embodiment, the list of hits 520 includes the URLs of the documents that were requested and subsequently serviced from the proxy's local cache (e.g., document cache 465) since the time indicated by the timestamp 525. According to the site hit record 515 for CompanyA, the adl.html Web page has been requested and serviced from the proxy's local cache three times. Similarly, the sales.html and Ql.html Web pages have been provided from the proxy's cache once and twice, respectively. Based upon the accumulated hit information in a particular site hit record 515, a detailed hit report may be provided to the corresponding remote site administrator. Hit accumulation will be discussed further below.

30 Figure 6 is a logical view of an exemplary directory structure 600 that may exist on a remote server 4. This exemplary directory structure 600 illustrates the need for a flexible method of tracking the number of hits. Web pages might reside in any or all of the directories shown. In this example, the URL patterns within a site tracking list record 505 may identify a particular directory or directories in the hierarchy depicted.

35

The remote site administrator for CompanyA may want to know the number of hits in an Ads subdirectory 605 and an Events subdirectory 610. This may be due to the fact that advertising banners are shown on Web pages in these directories and the advertisers may want feedback on how many Web viewers are seeing their ads. Alternatively, the company may have its own business reasons for analyzing statistics in certain areas of their Web site. Regardless, it is apparent that simply tracking all hits for a root directory 615 on the company's server is insufficient. For example, hits would be tracked for directories in which the remote site administrator had no interest. A list of URL patterns is used to accommodate the flexibility desired. The following URL patterns may be stored in the site tracking list 425 for CompanyA to track the above-mentioned subdirectories: "http://www.companyA.com/products/Events" and "http://www.companyA.com/products/Ads/." The list of URL patterns in each site tracking list record 505 allows a remote site to enumerate specific directories, for example, in which they would like to track user hits.

The advantages of providing forms of URL patterns with wild cards becomes apparent with reference to the directory structure 600. Assume the '\*' character is a wild card. That is, it matches zero or more characters. Since, CompanyA has two subdirectories with press releases, a convenient way to track hits in both is with the following URL pattern: "http://www.companyA.com/\*press\_releases/." Without the use of a wild card, the equivalent URL patterns are as follows: "http://www.companyA.com/press\_releases/" and "http://www.companyA.com/products/press\_releases." Thus, it should be appreciated that wild cards and regular expressions provide additional efficiency and convenience in the specification of URL patterns.

#### Hit Accumulation

Figure 7 is a flow diagram illustrating a method of performing hit accumulation according to one embodiment of the present invention. In this embodiment, each site hit record 515 begins in an initial state having an indication of the remote site (e.g., the name 530) and a timestamp 525 representing the time at which hit accumulation began. Initially, the hit accumulation server 415 waits for an indication that a client request has been serviced from the proxy's local cache (step 710). For example, the hit accumulator processor 430 may determine that a new entry has been made to the hit log 420.

Upon receiving an indication that the proxy 400 has served up a cached response, the hit accumulation server 415 determines if the URL of the document retrieved from the proxy's local cache is one whose hits are to be tracked. As discussed above, not all hits

are tracked. In this embodiment, hits are tracked only for documents matching URL patterns that have been registered in a tracking list such as the site tracking list 425, discussed above. Therefore, the hit accumulator processor 430 compares the URL of the retrieved document to URL patterns 510 in each site tracking list record 505 to determine if the hit will be recorded in the per site hit database 440 (step 720). If no URL patterns 510 match the retrieved document the hit is ignored. Otherwise, if the retrieved document matches any of the URL patterns 510, then the appropriate site hit record 515 in the per site hit database 440 is updated (step 730).

Update of the site hit record 515 can be explained briefly with respect to Figure 5B. In this embodiment, the appropriate site hit record 515 is searched for an entry that matches the URL of the retrieved document. If the retrieved document's URL does not already exist in the list of hits 520 for the site hit record 515, then the URL is added and its count is set to one since this is the document's first hit. However, if the retrieved document's URL was already in the list of hits 520 (meaning it has had at least one previous hit), then only the corresponding count needs to be incremented. In this manner, each document retrieved from the proxy's local cache that matches a tracked URL pattern will have an entry in the list of hits 520 with a corresponding count indicating the number of cache hits.

### Hit Reporting

Referring now to Figure 8, a method of hit reporting according to one embodiment of the present invention is illustrated. In this embodiment of the present invention, the hit accumulator server 415, in addition to its other responsibilities, acts as a Web server providing an HTTP interface by which remote site administrators can access the accumulated hits for their respective tracked sites. The hit report processor 450 waits until a request is received from a remote site administrator (step 810). Preferably, the HTTP address on the hit accumulation server 415 can be used to identify the requester of the information. For example, the hit report for Company A, might be accessed on the hit accumulation server 415 at: "http://www.webtv.net/hits/company\_a."

To limit access to the hit reports a secure communication technology such as Secure Sockets Layer (SSL) or other available secure communication protocol can be used to keep the hit information private by providing encrypted communications across the network. Additionally, the report requests can be authenticated to assure only a particular remote server or individual can access the information (step 820).

Once a request has been received from a remote site administrator and it has been optionally authenticated, then a report can be generated from the hit data accumulated such as the list of hits 520 for the particular sit hit record 515 (step 830). In this embodiment, the report may include a list of URLs and their corresponding counts since  
5 the last report. For convenient access via the Web, the report may be formatted in an HTML format. Also, for the convenience of the remote site administrator a timestamp that identifies the starting point of the accumulation may be included in the report. The level of specificity of the URL list may be at the document level thereby allowing the remote site administrator to determine the number of hits for individual documents, however, it  
10 may also be helpful to additionally summarize the hits by directory, for example. It will be recognized that numerous other ways of formatting and arranging the hit reports are possible.

After the report has been formatted, the response containing the report is transmitted to the remote site administrator (step 840).

15 In this embodiment, before resuming the hit accumulation of Figure 7, the accumulated data in the site hit record 515 is cleared (step 850) also the timestamp 525 is reset to reflect the current time. The above steps for retrieving a report from the proxy may be periodically repeated at the convenience of the remote site administrator whenever an accurate total hit count is desired.

20 In alternative embodiments, hit reports may be provided to remote sites in a number of other ways. Hit reports need not be initiated by a request from the remote site administrator. For example, the proxy may periodically send unsolicited hit reports via e-mail, the proxy may periodically download hit updates to a device specified by the remote site administrator, or the hit reports might be transmitted to remote site administrators in  
25 the form of datagrams. In any event, the assignees of the present invention appreciate a variety of reporting mechanisms are possible.

#### Allocation of Cache Space Within a Proxy

Figure 9 is a data flow diagram illustrating the interaction of proxy components according to another embodiment of the present invention. In this embodiment, proxy 900  
30 includes a plurality of proxy servers 405 communicatively coupled to a dispatcher 910 and a hit accumulator server 415. Rather than allowing a given proxy server's cached contents to be determined based upon the requests of an associated client, the content of the Web can be distributed among proxy servers 405 by a hash algorithm executed by the  
35 dispatcher 910. The hash algorithm preferably maps a given URL to one and only one of the plurality of proxy servers 405. This can be accomplished using a portion of the output

of a secure hash algorithm such as the Message Digest 5 (MD5) hash algorithm. The hash algorithm can be thought of as a mechanism for assigning a range of URLs to each of the proxy servers 405 in the proxy 900.

In this embodiment, the dispatcher 910 receives document requests including URLs from a client such as WebTV™ client 1. Based upon the URL in the request, the dispatcher determines the proxy server 405 in which the document should be cached and forwards the client request to that proxy server 405. If the document requested by the client is not found in the proxy server's local document cache 465, then the proxy server 405 requests the document from the appropriate server (e.g., a remote server) and caches the document when it is received from the server.

If redundancy is desired, the hashed result of a URL may be used to identify a cluster of two or more proxy servers rather than a single proxy server 405. In this manner, the load required to support a popular document can be shared among a group of proxy servers.

In an alternative embodiment, a decentralized dispatching scheme can be implemented. For example, the proxy servers 405 may be arranged to form an interconnected ring configuration and the functionality of the dispatcher 910 may be incorporated into each proxy server 405. In this embodiment, the client document requests may be initially handled by one of the proxy servers 405 in the ring. If the requested document is not found in the local cache of the initial proxy server, the initial proxy server may forward the request to the appropriate proxy server based on the hashing scheme discussed above.

Figure 10 is a flow diagram illustrating a method of dispatching requests to segregate the storage of documents according to one embodiment of the present invention. While both a centralized and a decentralized request dispatching mechanism have been discussed above, the method described below is generally applicable to both. In this embodiment, initially, a document request is received from a client (step 1010).

If a centralized dispatcher such as dispatcher 910 receives the request, then based upon the URL an appropriate proxy server is determined based upon the output of the hash algorithm (step 1020).

However, in a decentralized dispatching environment, the initial proxy server receiving the client request may assume it is the appropriate proxy server and first check its local document cache 465. If the document is not present, then proxy server may perform the hash algorithm on the URL to determine which of the remaining proxy servers is appropriate for the request (step 1020).

After determining the proxy server appropriate for the client request, the request is forwarded to that proxy server (step 1030). The proxy server 405 attempt to service the request from its local document cache 465. If a cache hit occurs, then the document is immediately available from the proxy server's local document cache 465. However, if a  
5 cache miss occurs, the proxy server 405 will retrieve the document from an appropriate server and store a copy locally. In any event, the centralized or decentralized dispatching mechanism ultimately receives a response from the server (e.g., the document requested by the client) (step 1040). Finally, the response, typically in the form of an HTML document is forwarded to the client (step 1050). This method of caching documents  
10 segregates the content of the Web based upon the URL of the documents. Since each URL will map to only one proxy server 405, advantageously this approach more efficiently allocates the proxy's cache space by avoiding unnecessary redundancy.

In the foregoing specification, the invention has been described with reference to specific embodiments thereof. It will, however, be evident that various modifications and  
15 changes may be made thereto without departing from the broader spirit and scope of the invention. The specification and drawings are, accordingly, to be regarded in an illustrative rather than a restrictive sense.

What is claimed is:



1. A method of tracking hits in a proxy, the proxy including a document cache having stored therein recently requested documents, the proxy coupled to a client and to a remote server, the method comprising the steps of:

5 the proxy maintaining information regarding client requests that are serviced from the document cache; and

the proxy providing the information to a remote site administrator.

2. The method of claim 1, wherein the information includes a count representing the number of times client requests for a particular document have been serviced from the document cache.

10 3. The method of claim 2, wherein the information includes a timestamp identifying a time period to which the count corresponds.

4. The method of claim 2, the method further including the steps of:

the proxy receiving a client request for a document;

15 the proxy determining whether to service the client request from its document cache or whether to forward the client request to another server; and updating the count if the client request is serviced from the document cache.

5. The method of claim 1, wherein the remote site administrator requests the information from the proxy thereby initiating the step of providing the information to a remote site administrator.

6. The method of claim 1, wherein the step of providing the information to a remote site administrator further includes the step of the proxy transmitting unsolicited information from the proxy to the remote site administrator.

7. The method of claim 5, wherein the information provided to the remote site administrator is in the form of an Hypertext Mark-up Language (HTML) report.

8. The method of claim 5, wherein the proxy authenticates the remote site administrator's request prior to the step of providing the information to a remote site administrator.

9. A method of tracking requests for documents stored in a proxy, the proxy including one or more proxy servers, each of said one or more proxy servers including a local cache having stored therein one or more cached documents, the proxy coupled to a client and to a remote server, the method comprising the steps of:

a proxy server receiving a client request for a document;

35 the proxy recording a hit for the document if the document is available from the proxy server's local cache; and

providing an indication of the number of hits for the documents to a remote site administrator.

10. The method of claim 9, wherein the proxy further includes a hit accumulation server, the step of the proxy recording a hit for the document if the document is available from the proxy server's local cache further including the steps of:

the proxy server notifying the accumulation server that the client request was serviced from the local cache; and

the accumulation server recording the hit and a path of the document to a table if the document corresponds to one of a set of monitored Uniform Resource Locator (URL) patterns.

11. The method of claim 10, wherein the step of the proxy server notifying the accumulation server further includes the steps of:

the accumulation server monitoring a common storage device, the common storage device accessible to the one or more proxy servers;

and the proxy server logging an entry to the common storage device if the proxy server services the client request from its local cache, the entry including a URL for the document requested by the client.

12. The method of claim 11, wherein the set of monitored URL patterns represents one or more directories on a remote site for which document hits are to be tracked, the method further including the steps of:

the accumulation server detecting the entry; and

the accumulation server comparing the URL in the entry to the set of monitored URL patterns to determine whether or not to record the hit.

13. A method of tracking hits by a proxy, the proxy including one or more proxy servers and a hit accumulation server, each of said one or more proxy servers including a local cache having stored therein recently requested documents, the proxy coupled to a client and to a remote server, the method comprising the steps of:

a proxy server receiving a client request for a document;

the proxy server determining whether to service the client request from its local cache or whether to forward the client request to another server;

inserting a new log entry onto a common storage if the proxy server services the client request from its local cache, the new log entry including a location of the document;

the accumulation server detecting the new log entry by monitoring the common storage;

the accumulation server comparing the location of the document to a predetermined set of directories to determine whether or not to record the hit;

recording the hit if the location of the document matches a directory in the predetermined set of directories; and

5 providing the number of hits for a set of documents located in a first subset of directories of the predetermined set of directories to a remote site administrator.

14. The method of claim 13, wherein the common storage comprises a hit log.

10 15. The method of claim 13, wherein the location of the document comprises a Uniform Resource Locator (URL).

16. A machine-readable medium having stored thereon data representing sequences of instructions, said sequences of instructions which, when executed by a processor, cause said processor to perform the steps of:

15 maintaining information regarding client requests that are serviced from a document cache of a proxy server; and

providing the information to a remote site administrator.

17. The machine-readable medium of claim 16, wherein the information includes a count representing the number of times client requests for a particular document have been serviced from the document cache.

20 18. The machine-readable medium of claim 17, wherein the information includes a timestamp identifying a time period to which the count corresponds.

19. The machine-readable medium of claim 17, wherein said sequences of instructions further cause said processor to perform the steps of:

receiving a client request for a document;

25 determining whether to service the client request from the document cache or whether to forward the client request to another server; and

updating the count if the client request is serviced from the document cache.

20. A proxy comprising:

30 a document cache having stored therein one or more cached documents; one or more proxy servers coupled to a network for receiving and servicing client requests for documents; and

a hit accumulation server coupled to the one or more proxy servers, the hit accumulation server configured to record statistics regarding document requests serviced from the document cache and to provide the statistics to a remote site administrator.

35

21. The proxy of claim 20, wherein the one or more proxy servers each include a local cache, the one or more proxy servers are each configured to determine whether to forward a client request for a document to another server or to service the client request from its local cache, and wherein the document cache comprises the one or more local caches.

22. The proxy of claim 20, wherein the one or more proxy servers and the hit accumulation server share a common storage device for communicating document requests serviced from the document cache.

23. The proxy of claim 20, wherein the hit accumulation server is configured to provide the statistics in the form of a Hypertext Mark-up Language (HTML) report.

24. A method of dispatching document requests in a proxy, the method comprising the steps of:

receiving a document request from a client, the document request including a Uniform Resource Locator (URL); and

forwarding the request to one of a plurality of proxy servers based upon the URL.

25. The method of claim 24, wherein the step of forwarding the request to one of a plurality of proxy servers based upon the URL includes the step of performing a predetermined hash function on the URL.

26. The method of claim 25, wherein the predetermined hash function has an output, the output associating the URL with exactly one of the plurality of proxy servers.

27. The method of claim 25, wherein the predetermined hash function has an output, the output associating the URL with a group of two or more of the plurality of proxy servers.

28. A method of caching documents in a proxy, the documents being identified by Uniform Resource Locators (URLs), the method comprising the steps of:

assigning a range of URLs to each of a plurality of proxy servers in a proxy; and

caching a document in a proxy server only if the URL of the document is within the range assigned to the proxy server, thereby assuring only one copy of a particular document will reside in the proxy at any given time.

29. The method of claim 28, wherein the step of assigning a range of URLs to each of a plurality of proxy servers is accomplished with a predetermined hash algorithm.

30. A machine-readable medium having stored thereon data representing sequences of instructions, said sequences of instructions which, when executed by a processor, cause said processor to perform the steps of:

determining a proxy server of a plurality of proxy servers in which to cache a document based upon a Uniform Resource Locator (URL) associated with the document; and

storing a copy of the document in a document cache of the proxy server.

- 5           31.     The machine-readable medium of claim 30, wherein the step of determining a proxy server of a plurality of proxy servers in which to cache a document based upon a Uniform Resource Locator (URL) associated with the document further includes the step of applying a predetermined hash algorithm to the URL.

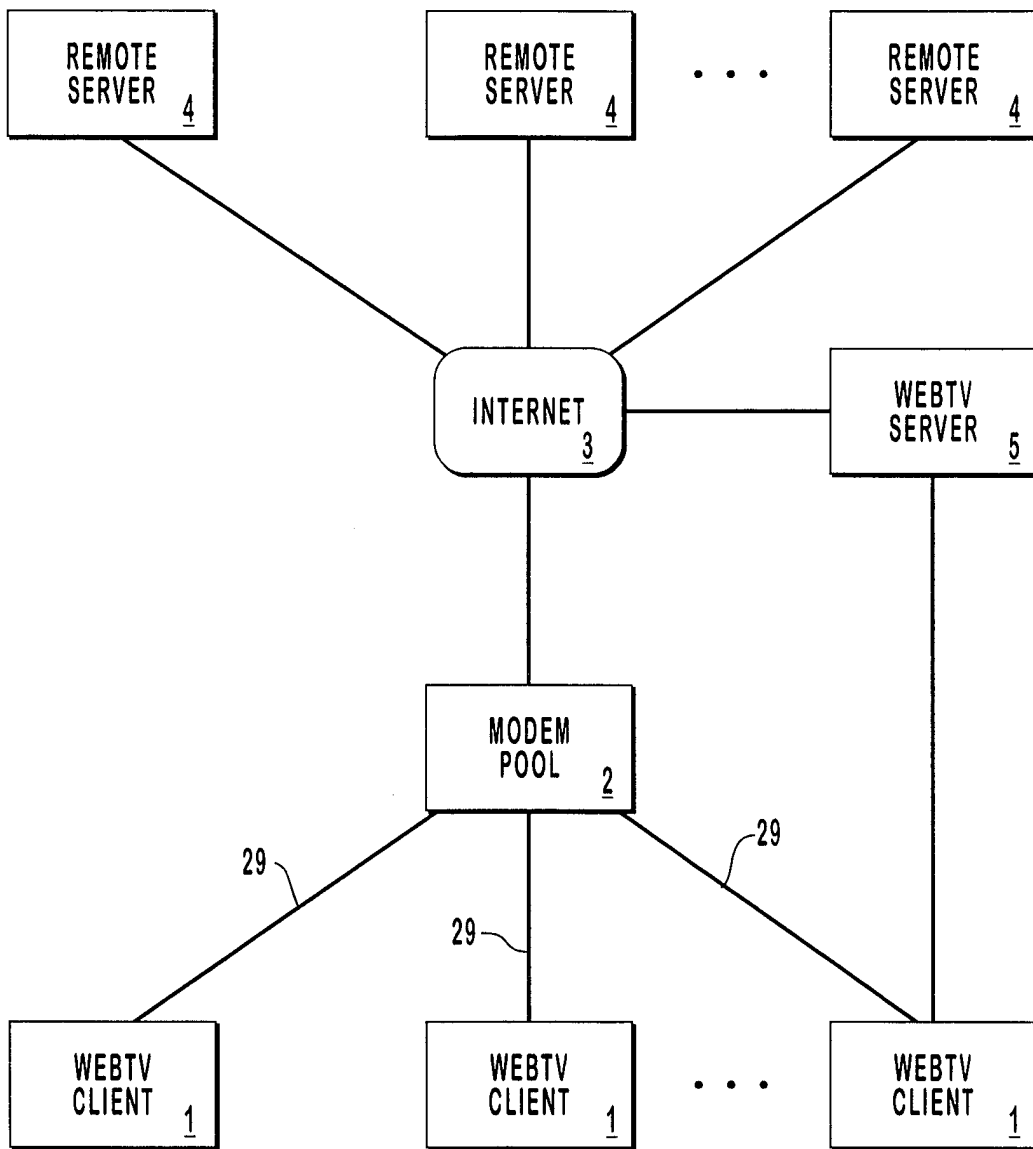


FIG. 1

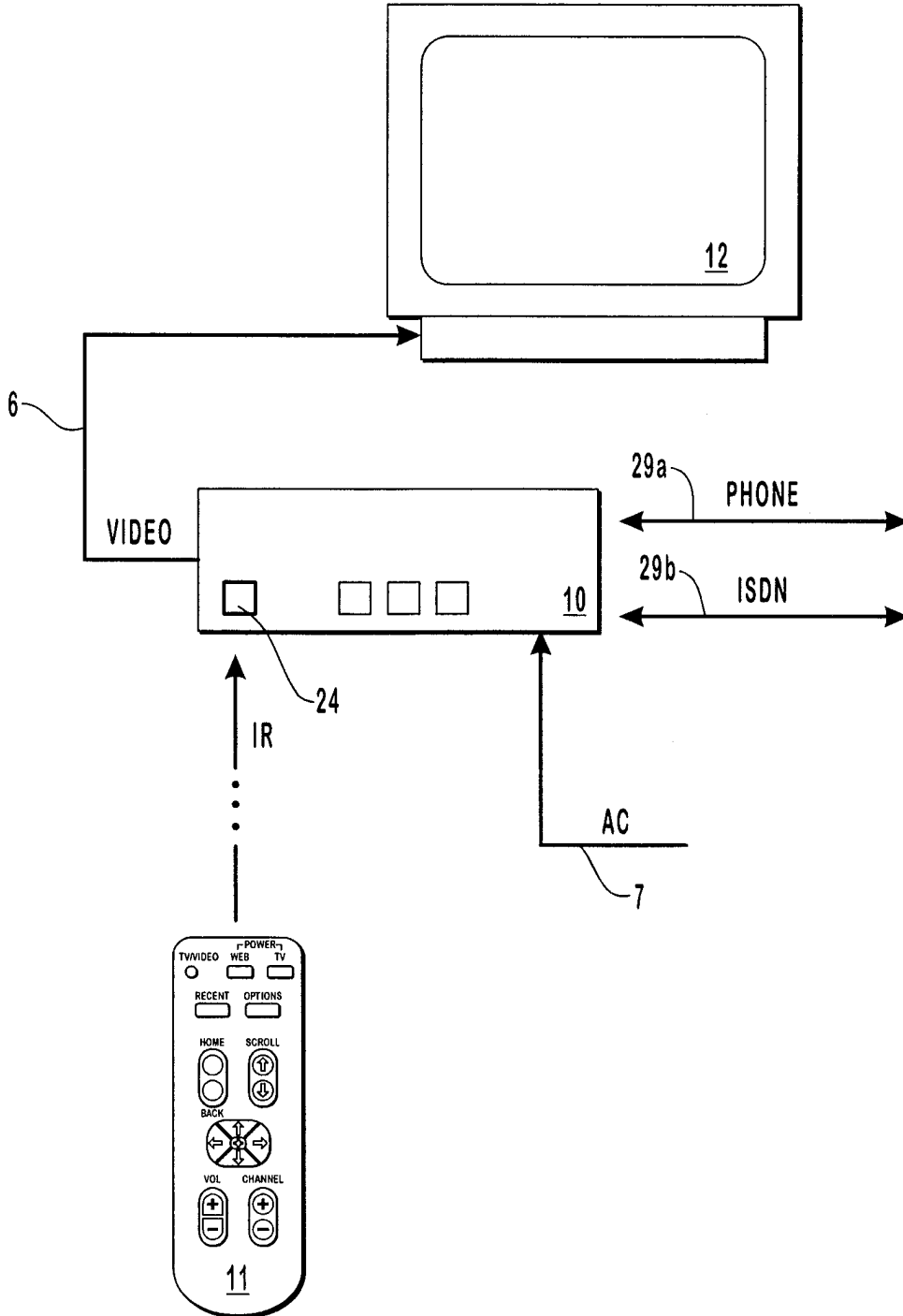


FIG. 2

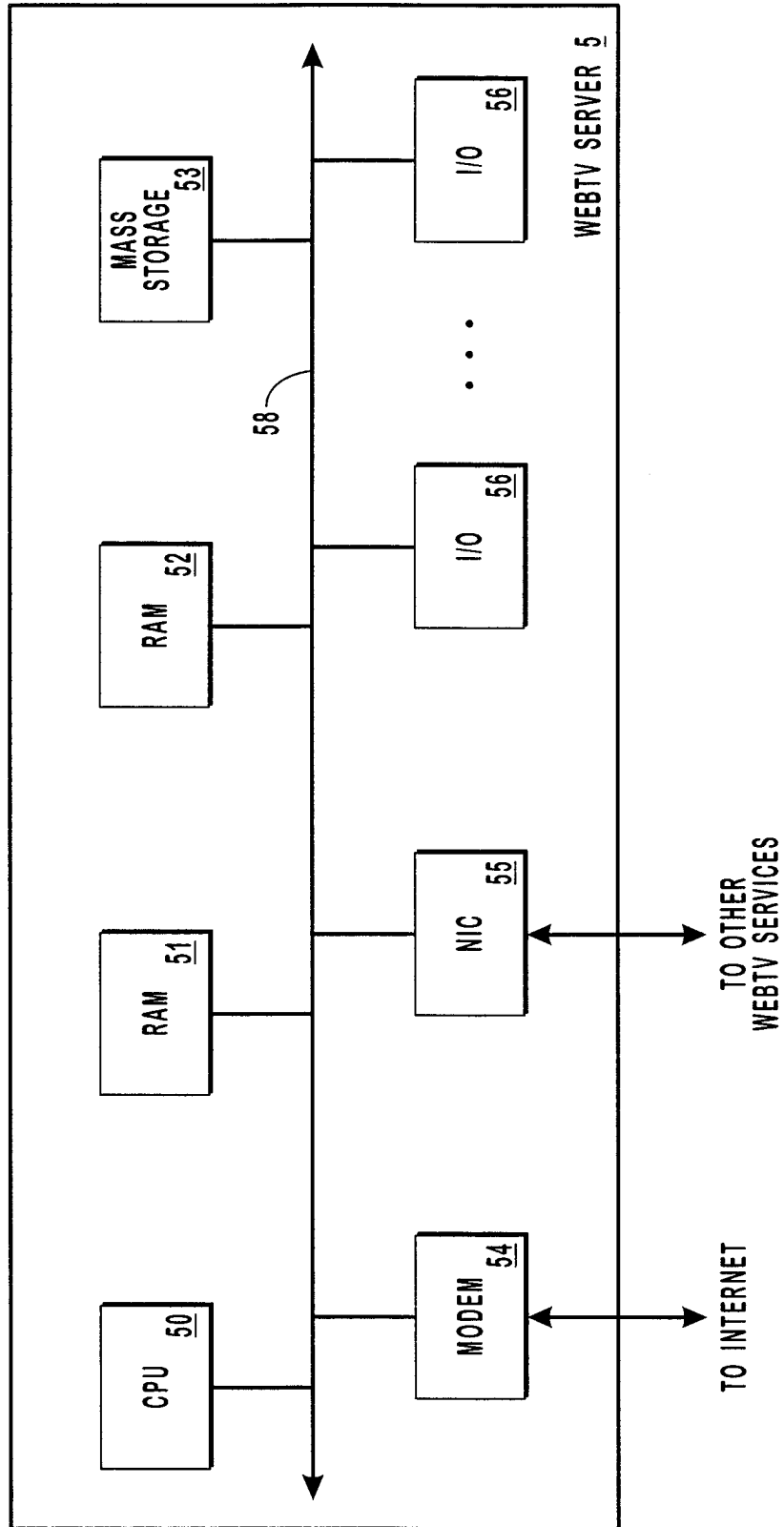


FIG. 3



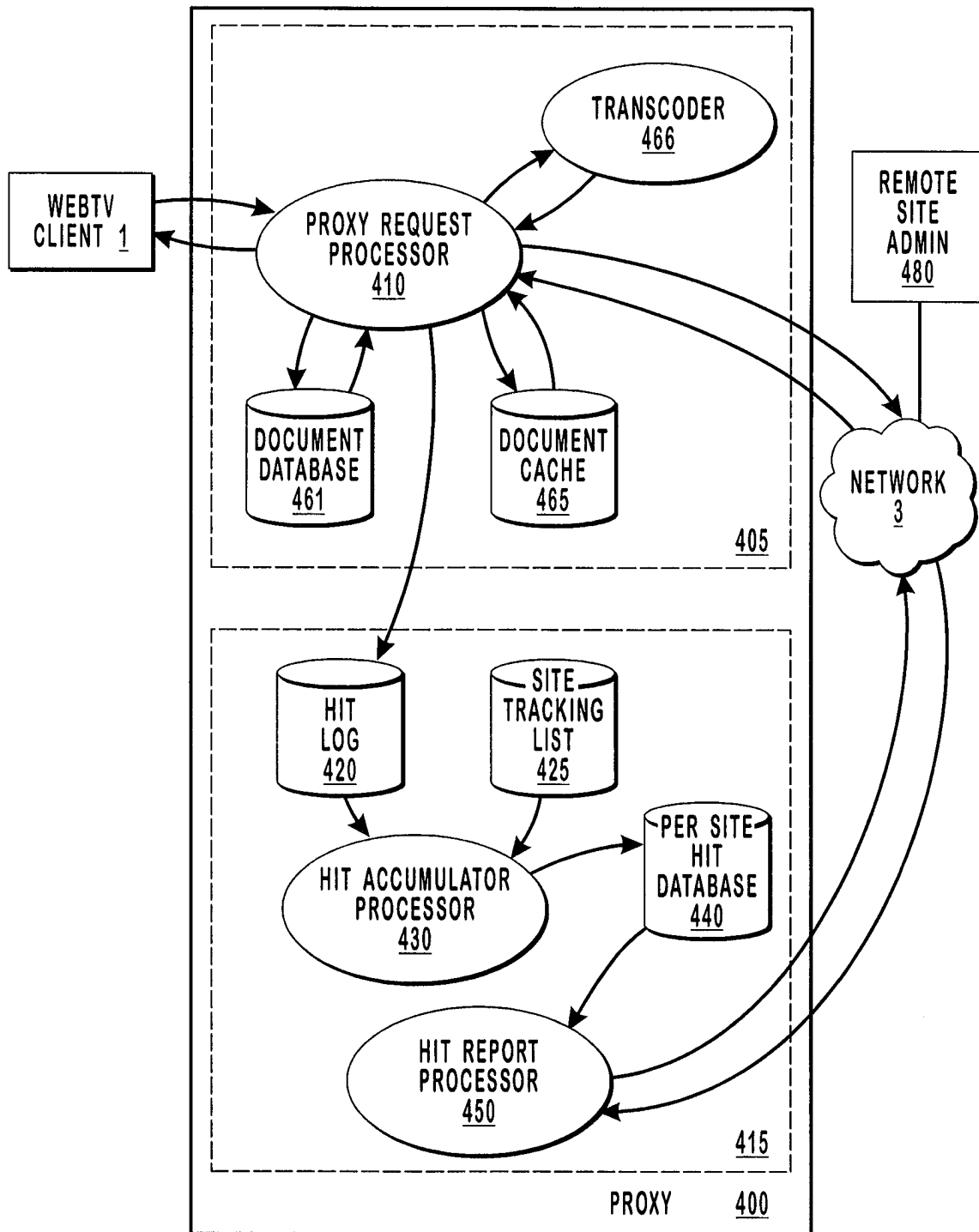


FIG. 4

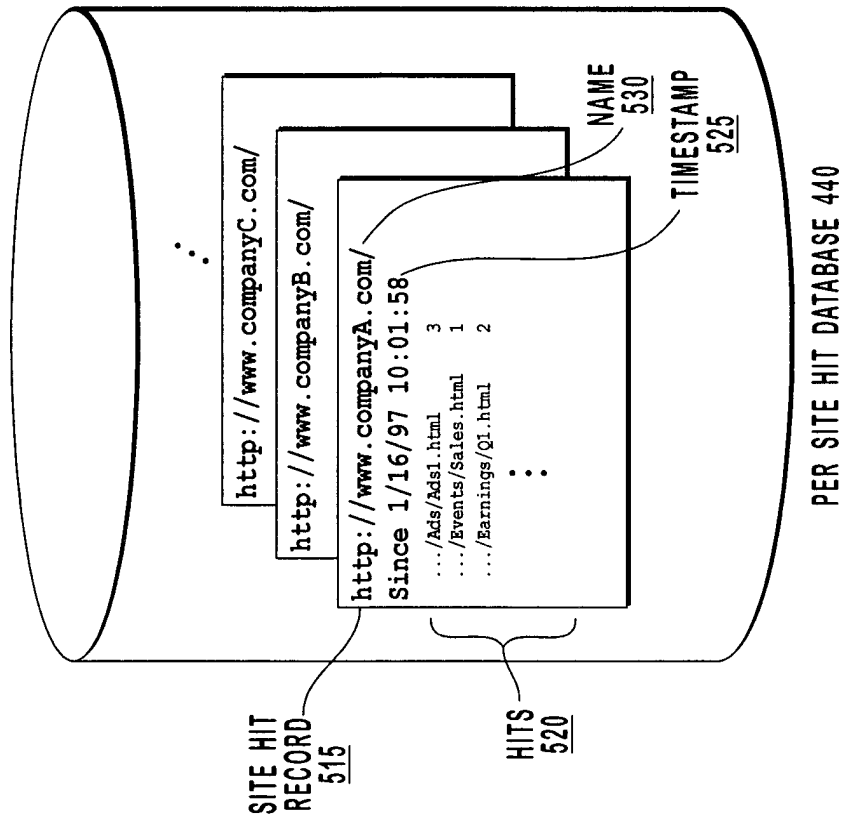


FIG. 5B

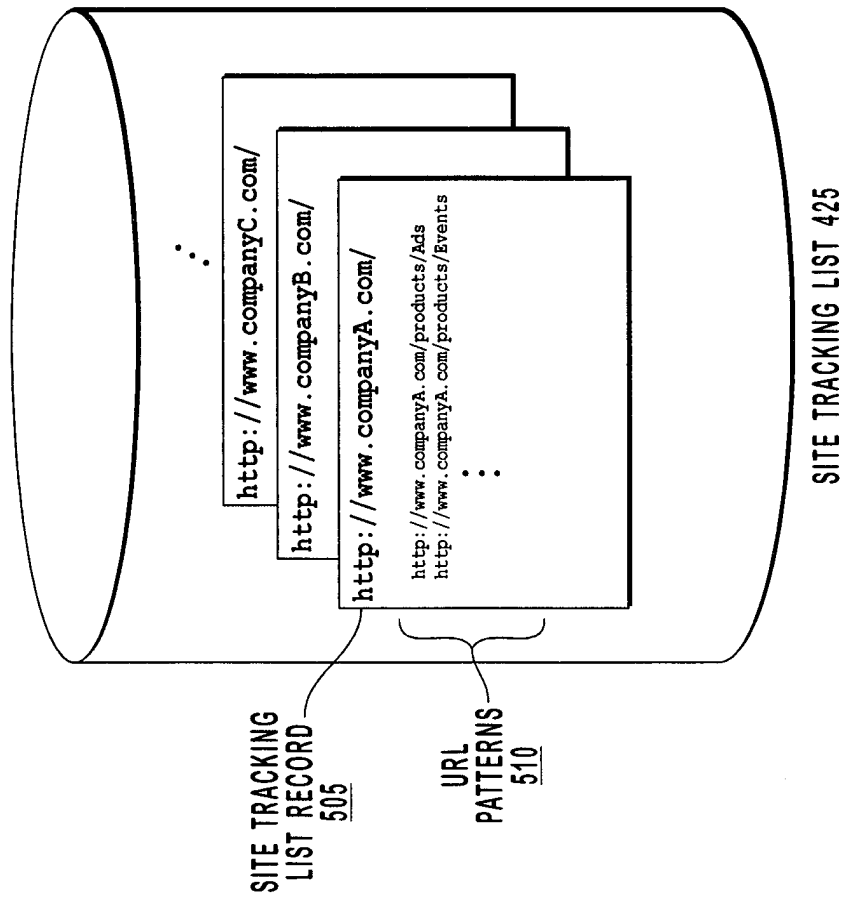


FIG. 5A

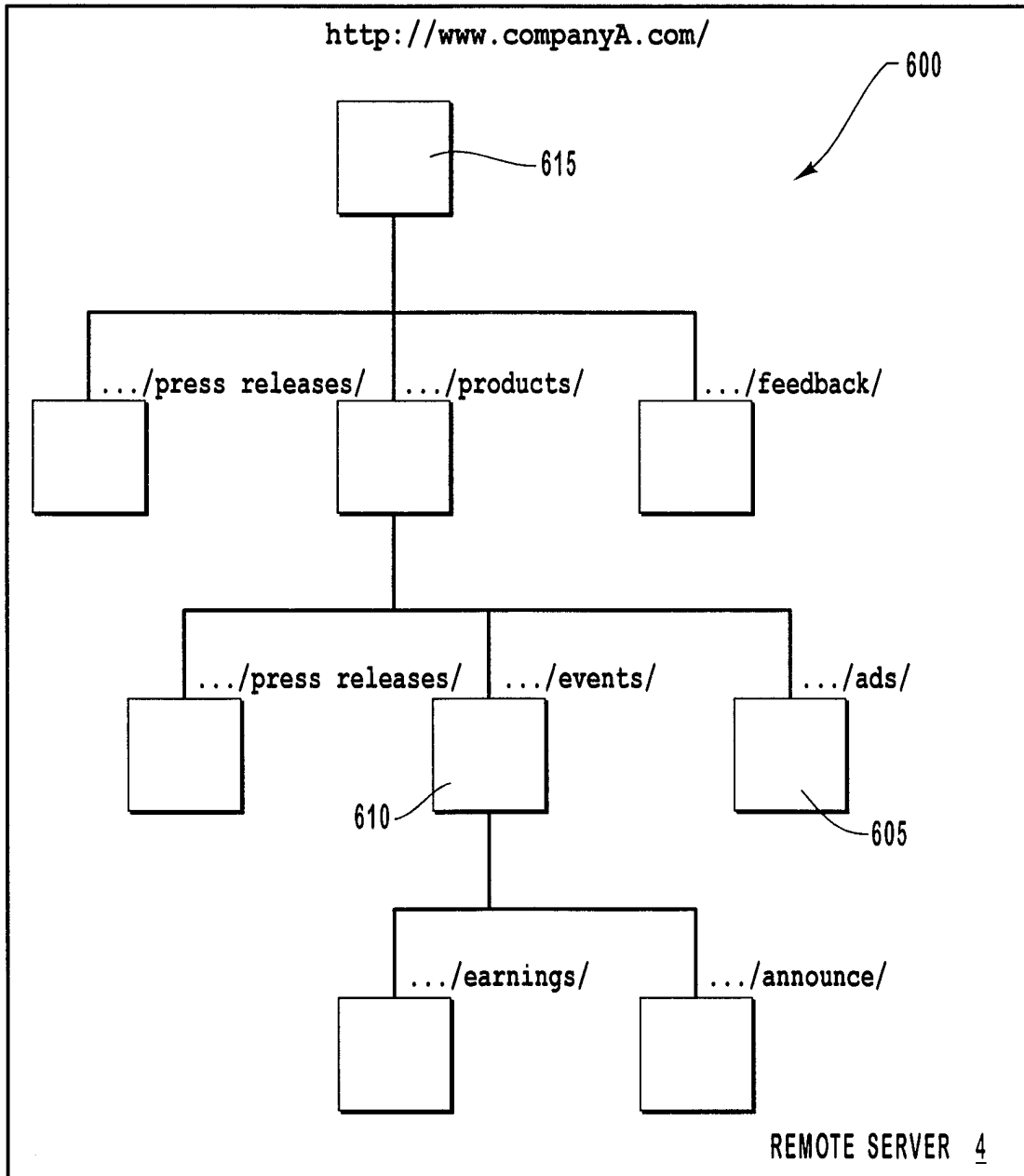


FIG. 6

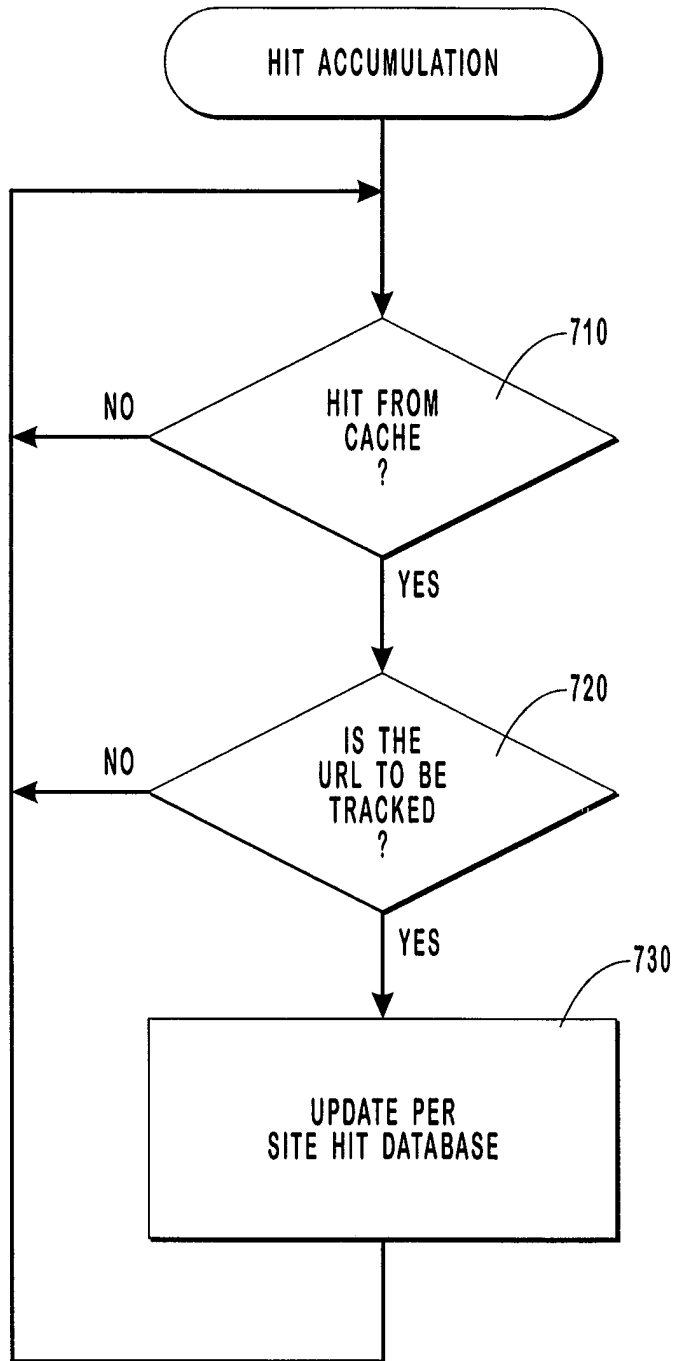


FIG. 7

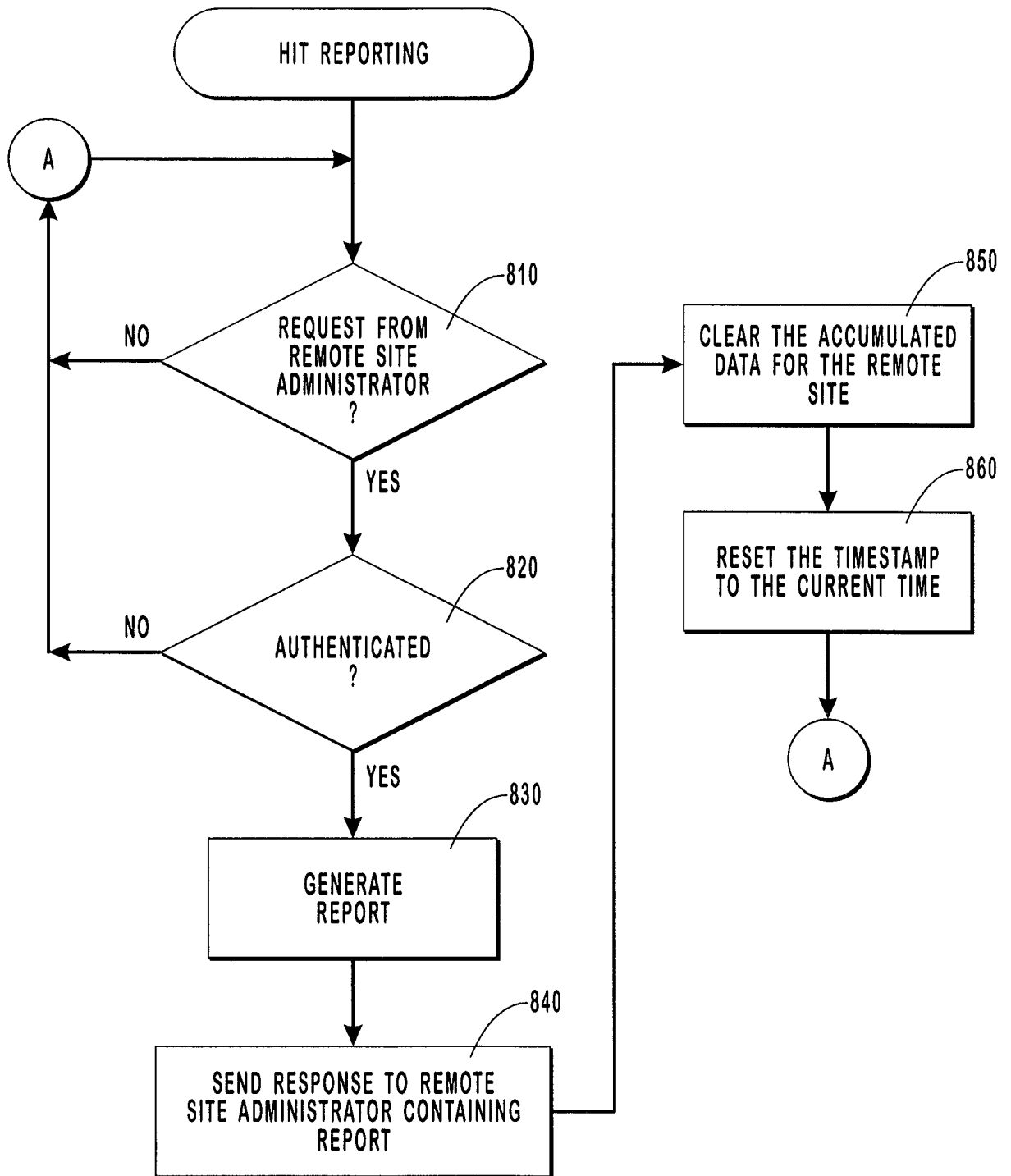


FIG. 8

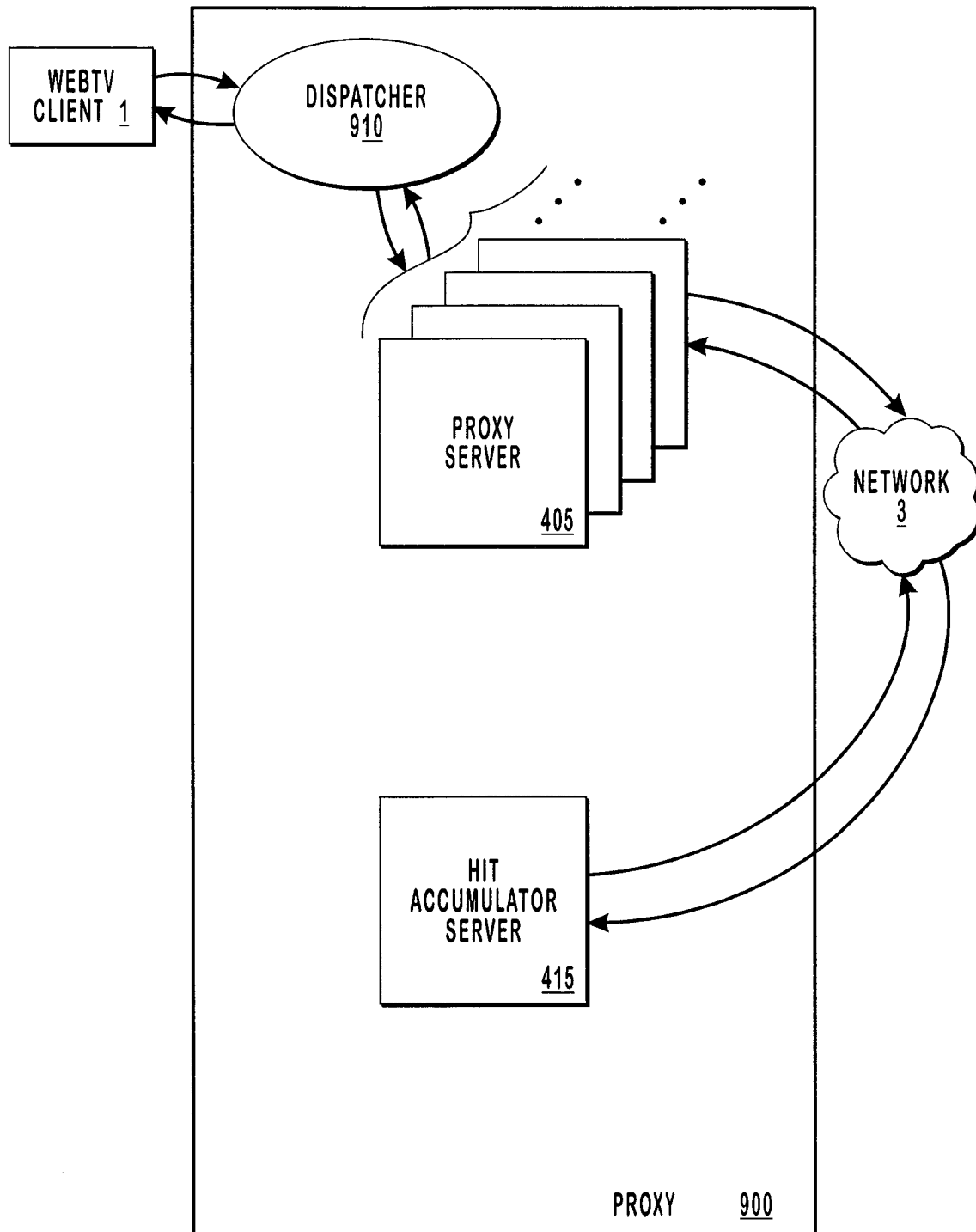


FIG. 9

10 / 10

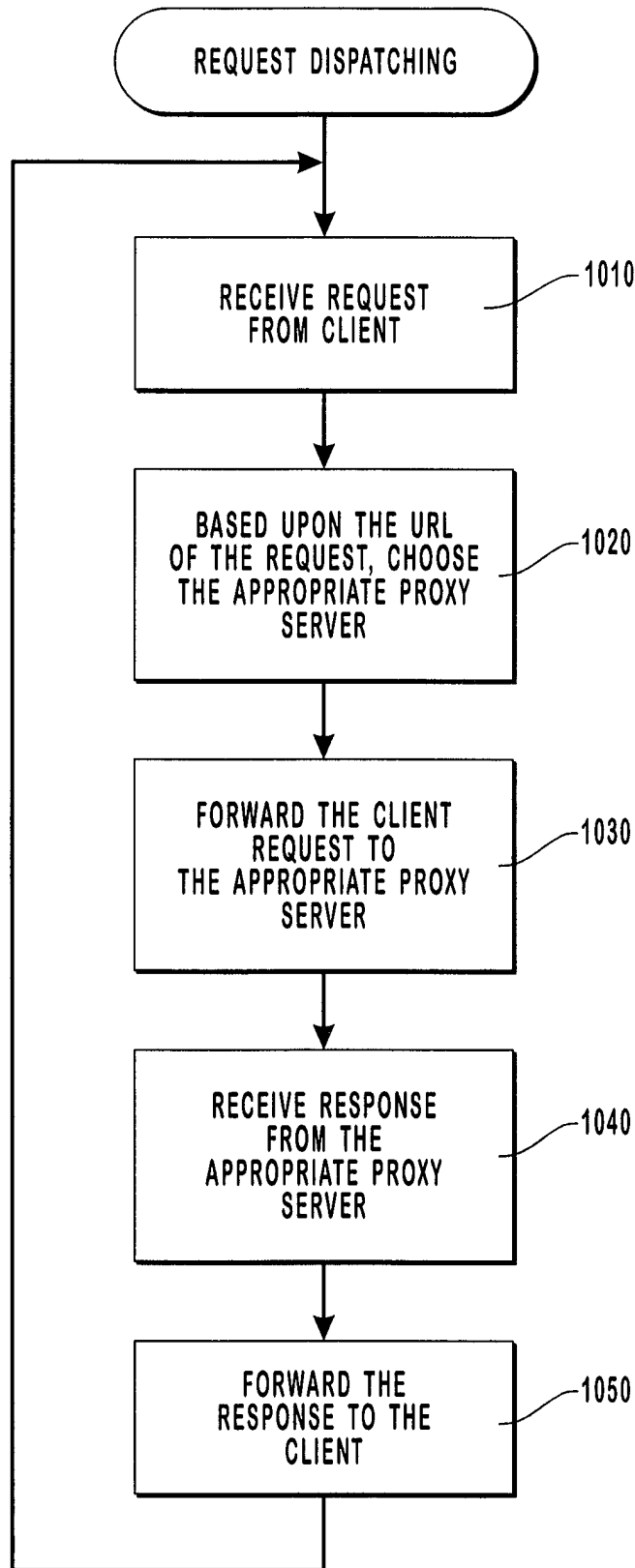


FIG. 10