(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
17 April 2003 (17.04.2003)

PCT

(10) International Publication Number
WO 03/032201 A1

(51) International Patent Classification[7]: G06F 17/30

(21) International Application Number: PCT/US02/31907

(22) International Filing Date: 7 October 2002 (07.10.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/974,665 9 October 2001 (09.10.2001) US
09/974,664 9 October 2001 (09.10.2001) US

(71) Applicant: **WILDBLUE COMMUNICATIONS, INC.** [US/US]; 7600 E. Orchard, Suite 360N, Englewood, CO 80111 (US).

(72) Inventors: **WEAVER, Jeffrey, Charles**; 8057 S. Leyden Street, Englewood, CO 80112 (US). **SUNDELIN, Andrew, Wilson**; 1405 S. Gaylord Street, Denver, CO 80210 (US). **MOORE, Thomas, Evans**; 10 Lynn Road, Englewood, CO 80110 (US).

(74) Agent: **STACY, Wayne, O.**; Cooley Godward LLP, Patent Group, 11951 Freedom Drive, One Freedom Square-Reston Town Center, Reston, VA 20190-5601 (US).

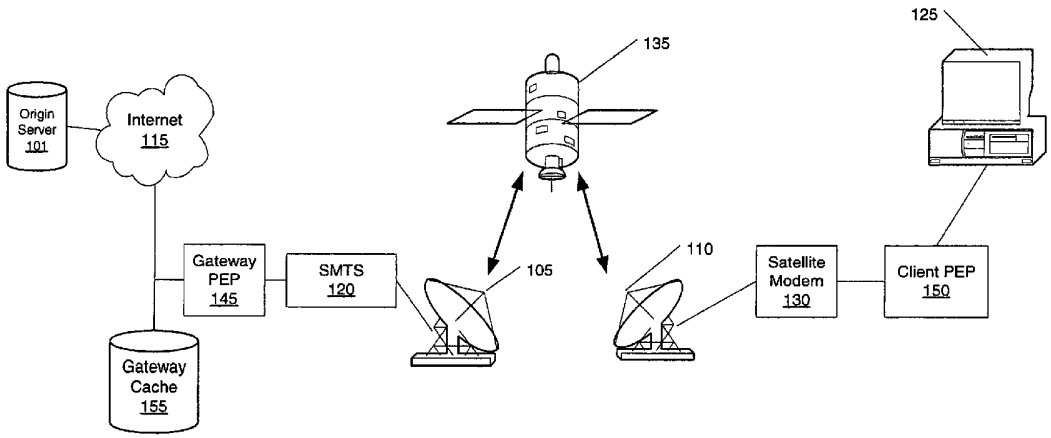
(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: PERFORMANCE ENHANCING PROXY FOR HIGH LATENCY DATA



(57) Abstract: A system and method for reducing the impact of high latency links for interactive applications (Fig. 2) is described. In one embodiment, a user, connected to the Internet by a satellite network, requests a web page from an origin server (101, Fig. 2). The origin server generates an index page associated with the requested web page and passes that index page to a gateway PEP (performance enhancing proxy) (145, Fig. 2), which parses the index page to identify the objects identified therein. Next, the gateway PEP generates the requests to prefetch those identified objects and passes those requests to the origin server. These objects are then retrieved and pushed to the user (150, Fig. 2). To avoid repetitive requests, the fetch requests normally generated by a web browser are blocked by a client PEP.

WO 03/032201 A1

PERFORMANCE ENHANCING PROXY FOR HIGH LATENCY DATA

FIELD OF THE INVENTION

The present invention relates to systems and methods for delivery of data. In particular, but not by way of limitation, the present invention relates to a system and method for reducing the impact of high latency links over satellite and wireless systems.

BACKGROUND OF THE INVENTION

With the increasing demand for high-speed access to the Internet, Internet service providers and Internet users are increasingly turning to satellite networks. These satellite networks have the potential to offer significant coverage improvements over traditional, broadband terrestrial networks. Additionally, satellite networks allow for connections to the Internet to be established without the expense of laying an extensive network of terrestrial lines.

Even with all of the advantages offered by satellite networks, they suffer from significant drawbacks that limit their commercial appeal to Internet users. In particular, users accessing the Internet over a satellite network can experience lengthy delays. Although portions of this delay are inherent to satellite networks because of the large distances that signals must travel, much of the delay is caused by the use of inefficient protocols to transmit data over the satellite link.

Traditional data transmission protocols were designed for conventional terrestrial networks with relatively uniform low latencies between components. These traditional protocols, however, are extremely inefficient for satellite and wireless networks. For example, traditional protocols do not account for the fact that the latency between a gateway satellite dish 105 and a user satellite dish 110 is extremely high while the latency between the Internet 115 and, for example, a SMTS (Satellite Modem Termination System) 120 is relatively low. (See FIGURE 1). Despite the drawbacks to using these traditional protocols over satellite systems, terrestrial Internet service providers and system designers continue to incorporate them into their technology. Moreover, because satellite Internet users want access to the same Internet content as terrestrial Internet users, satellite Internet users necessarily have to interface with standard Internet protocols designed for terrestrial use. The result of the continued use

of these inefficient protocols has resulted in a commercial product that is often less than satisfactory and has slowed the adoption of an otherwise beneficial technology.

Some companies have recently undertaken an effort to maximize the efficiency of TCP/IP, one of the traditional data transmission protocols, with regard to satellite networks. For example, Flash Networks has developed a TCP/IP enhancement. This enhancement is described in PCT patent application WO9847166, entitled "*Data Communication Protocol.*" Although enhancements to TCP/IP somewhat improve the performance of satellite networks, TCP/IP enhancements alone are not sufficient to make Internet service via a satellite network a viable consumer product. Accordingly, further enhancements to transmission protocols are needed if Internet service provision through satellite networks is to become widespread.

Recent efforts to improve the performance of satellite networks have been directed toward enhancing HTTP, another of the traditional data transmission protocols. These efforts, however, have been generally unsatisfactory. Accordingly, a system and method are needed to meaningfully improve the performance of HTTP with regard to satellite networks. In particular, a system and method are needed to reduce the latency and thus increase the effective speed that a user experiences when accessing the Internet through a satellite network.

SUMMARY OF THE INVENTION

In one innovative aspect of the present invention, a system and method for reducing the impact of high latency links, whether satellite links or wireless links, are disclosed. Briefly, this embodiment of the present invention reduces the impact of high latency satellite and wireless links by performing predictive prefetching and pushing of data objects needed by a user. Other embodiments are also described herein.

In one embodiment, for example, a user--through a web browser--initially requests a web page. This request is passed to a client PEP (performance enhancing proxy) that is often incorporated with the user's satellite modem. The client PEP then passes this request through the satellite network to the gateway PEP, which then passes the request to an origin server. Responsive to receiving the request, the origin server generates an index page (also called a container page) associated with the requested web page and passes that index page to the gateway PEP, which parses the index page to identify the objects listed therein. Next, the gateway PEP generates the requests to

fetch, i.e., prefetch, those identified objects and passes those requests to the origin server. The origin PEP can also pass the index page and a list of those objects being prefetched to the client PEP. The client PEP can then pass at least the index page to the user.

As with the standard Internet, once the user receives the index page, it generates a fetch request for each--or at least some of--the objects identified therein. In a typical satellite-based system, these fetch requests would normally be passed through the satellite network to the origin server. However, in one embodiment of the present invention, the fetch requests, or at least some portion thereof, can be blocked, i.e., not passed to the origin server, by the client PEP because the client PEP knows that certain fetch requests are already being filled through the requests previously passed from the gateway PEP to the origin server.

In one embodiment, the gateway PEP assumes that all embedded objects are needed by the client PEP. In alternate embodiments, however, the gateway PEP can be made aware of the prior activities of the client PEP and prefetch or push only those objects not already cached by the client PEP. In either of these embodiments, however, when the origin server returns prefetched objects, these objects are passed through the satellite network to the client PEP and on to the user.

From the user's perspective, the objects were received as a result of the fetch requests that the user originated. In reality, however, the objects could have been requested long before the user's browser even generated those fetch requests. In some cases, the objects may be in the process of being transmitted to the client PEP even before the user receives the index page. Thus, the present invention can, among other things, minimize the impact of the high latency between the gateway satellite dish and the local satellite dish is minimized.

BRIEF DESCRIPTION OF THE DRAWINGS

Various objects and advantages and a more complete understanding of the present invention are apparent and more readily appreciated by reference to the following Detailed Description and to the appended claims when taken in conjunction with the accompanying Drawings wherein:

FIGURE 1 is a present satellite network system connecting a user with the Internet;

FIGURE 2 is a satellite network system constructed in accordance with one embodiment of the present invention;

FIGURE 3 is an alternate embodiment of a satellite network system constructed in accordance with one embodiment of the present invention;

FIGURE 4 is an embodiment of a wireless network system constructed in accordance with the principles of the present invention;

FIGURE 5 is a data flow diagram illustrating the data flow between the components of the system shown in FIGURE 2;

FIGURE 6 is a data flow diagram illustrating an alternate data flow between the components of the system shown in FIGURE 2; and

FIGURE 7 is a data flow diagram illustrating a bandwidth optimization technique in accordance with one embodiment of the present invention.

DETAILED DESCRIPTION

Although the present invention is open to various modifications and alternative constructions, a preferred exemplary embodiment that is shown in the drawings is described herein in detail. It is to be understood, however, that there is no intention to limit the invention to the particular forms disclosed. One skilled in the art can recognize that there are numerous modifications, equivalents and alternative constructions that fall within the spirit and scope of the invention as expressed in the claims.

Referring now to FIGURE 1, there is illustrated a present satellite network system 100 connecting a user 125 with the Internet 115. In this embodiment, the user 125 is connected to a satellite modem 130 and a local satellite dish 110. The user 125 receives data--in some cases both Internet and TV--from a gateway satellite dish 105 via a geostationary satellite 135. (Collectively the local satellite dish 110, the geostationary satellite 135 and the gateway satellite dish 105 are referred to as the "satellite system.") The data that the gateway satellite dish 105 transmits to the user 125 can originate from the Internet 115 and origin server 101 or a broadcast system (not shown). The SMTS (Satellite Modem Termination System) 120 provides the necessary intelligence to bridge between the Internet 115 and the satellite system. This general configuration shown in FIGURE 1 is similar to the system used for modern

satellite TV systems. In fact, Hughes™ Network Systems offers a combination satellite TV system and Internet access (called "DirecDuo"™) using this general system.

Although the system described above has tremendous downloading capacity, this capacity is not readily accessible to a highly interactive activity such as web browsing. In fact, this downloading capacity presently is best used by non-interactive or low interactive activities such as streaming media or downloading large files. Those applications that require continued and frequent interaction between the user 125 and a device on the gateway side of the satellite system are often slow (1) because of the high latency between the gateway satellite dish 105 and the local satellite dish 110 and (2) because of the numerous roundtrip transmissions between the user 125 and the Internet 115 needed to successfully transfer data. Accordingly, the system shown in FIGURE 1 is somewhat undesirable for many Internet applications and other highly interactive applications.

Referring now to FIGURE 2, there is illustrated a satellite network system 140 constructed in accordance with the principles of the present invention. As with the system shown in FIGURE 1, this system includes a gateway satellite dish 105, a local satellite dish 110, and a geostationary satellite 135. Unlike the system shown in FIGURE 1, however, this embodiment of the present invention includes a gateway PEP (Performance Enhancing Proxy) 145, a client PEP 150 and a gateway cache 155. Although the client PEP 150 is shown separately from the satellite modem 130 and the user 125, in other embodiments, the client PEP 150 could be integrated into the satellite modem 130 and/or into the user's computer. For example, the client PEP 150 could be a browser plug-in or otherwise incorporated with the user's browser application.

The gateway PEP 145 and the client PEP 150 can replace or supplement traditional data transmission protocols, e.g., TCP/IP and HTTP, with a transport protocol that is optimized for a satellite network. For example, the protocol between the Internet 115 and the gateway PEP 145 could be TCP/IP and the protocol between the client PEP 150 and the user 125 could be TCP/IP. The protocol between the gateway satellite dish 105 and the local satellite dish 150, however, could be a satellite transport protocol instead of standard TCP/IP. In other words, the PEP can speak standard Internet protocols on one interface and satellite protocols on a different interface. Alternatively, the satellite transport protocol could replace traditional HTTP

with an enhanced protocol. As those of skill in the art can understand, by replacing the traditional transmission protocols with a protocol optimized for a satellite network, interactive applications, e.g., web browsing, can better utilize the downloading capacity of a satellite network and thereby make interactive applications faster, more efficient and commercially viable.

Figure 3 illustrates an alternate embodiment of a satellite network system 160 constructed in accordance with the principles of the present invention. In this embodiment, a client-side cache 165 and distribution service 170 are disposed intermediate the local satellite dish 110 and the individual users 125. The distribution service 170 could be a MMDS (Multichannel Multipoint Distribution System), a LMDS (Local Multipoint Distribution System), or any other type of distribution service.

The operation of the system illustrated in FIGURE 3 is similar to the operation of the system shown in FIGURE 2. The client PEP 150, however, can be modified to interact with multiple users 125 rather than a single user. Additionally, the client PEP 150 could be modified to take advantage of a client-side cache 165, which could be resident in the same device as the PEP, thereby offering the PEP the advantages of local storage.

Referring now to FIGURE 4, there is illustrated an embodiment of a wireless network system 175 constructed in accordance with the principles of the present invention. In this embodiment, the gateway PEP 145 is disposed between the Internet 115 and a wireless broadcast device 180. The gateway PEP 145 could enhance or modify the transmission protocols used by the wireless broadcast device 180 and thereby provide optimizations similar to those achieved in the previously-described satellite network systems.

Notably, the user 125 could be any type of wireless-adapted device ranging from a PDA to a cell phone to a wireless-connected mobile computer. In the preferred embodiment, the user 125 would incorporate a client PEP or at least portions of a PEP. Moreover, the user could include storage for storing data used by the client PEP.

The operation of the present invention is best illustrated by example. Although the following examples involve an Internet application, one skilled in the art can recognize that the present invention can be applied to other types of interactive

applications. Moreover, the following examples recite the execution of events in a particular order. The order of these events is exemplary only, and the present invention is not necessarily limited to the recited order.

FIGURE 5 illustrates the data flow between the components of the system shown in FIGURE 2. In this embodiment, the user, through a browser, sends a request to the client PEP (step 185). For example, the user could issue a "get cnn.com" request. The client PEP would then pass this request--through the satellite system--to the gateway PEP (step 190), which would then pass the request to an Internet-connected (or other network) origin server (step 195). Next, the origin server would return an index page to the gateway PEP (step 200). The gateway PEP would then parse the index page (step 205) to identify the objects within the page, generate a prefetch list (step 205), and attempt to prefetch the objects included in the prefetch list (step 230). Although the exact timing is not necessarily relevant, the gateway PEP next sends the index page and a list of the objects being prefetched to the client PEP (step 210). The client PEP passes the index page to the user (step 215) and in certain embodiments, the client PEP also passes the list of objects being prefetched to the user (step 215).

Assuming that the user includes a standard commercial browser, upon receiving the index page, the user generates GET requests for the objects identified in the index page and passes those requests to the client PEP (step 220). Because the client PEP has been told by the gateway PEP which objects that the gateway PEP is prefetching on its behalf (via the object list), the client PEP knows what objects may already be in transit from the gateway PEP. If the client PEP receives GET requests from the user's browser for objects on the object list, the client PEP blocks the requests (step 225), i.e., does not forward the requests, because it is already anticipating delivery of these objects. The user, however, does not necessarily know that the fetch requests have been blocked.

Concurrently, or generally concurrently, with certain ones of the previously described steps, the gateway PEP requests the index page objects from the origin server (step 230). The origin server sends the requested objects to the gateway PEP (step 235), and the gateway PEP passes those objects to the client PEP (step 240). Finally, the retrieved objects can be passed to the user (step 245) so that the user can assemble the objects and display the requested web page.

As can be appreciated by those of skill in the art, the above-described method presents a significant improvement over existing technology. In particular, the above-described method minimizes the impact of high latency links in satellite and wireless networks by, for example, minimizing the number of round trip interacts necessary to retrieve data. Instead of requiring the user to actually wait for an index page and then request the objects identified therein, this implementation may anticipate the needs of the user, fetch those objects before the user actually requests them and push them to the user/client PEP. Because the high latency link is generally higher latency than typical Internet latency to the origin server, these objects can often be pushed before the index page has even arrived at the end user. Thus, objects needed by a user may be in transit to the user before the user even requests them. If the objects are not needed, the client PEP can merely discard them or store them locally for future use.

Another distinct advantage of this embodiment of the present invention involves its use of existing technology. For example, one embodiment of the present invention incorporates existing browser technology. In other words, the user's browser does not necessarily need to be modified to benefit from the present invention. The browser generates the same requests that it would in a normal system. However, the client PEP intercepts these requests and blocks them if appropriate. By allowing the user to use a standard browser, systems in accordance with various aspects of the present invention can increase commercial appeal because, in essence, the user need only replace his standard modem with a satellite modem and a client PEP, which as previously described, can be a single, integrated product.

The advantage of this embodiment of the invention's use of existing technology at the end points (i.e., the browser and the origin server do not have to change) can be enhanced when the addition of advance technologies such as multicast are added into the system. The gateway PEP and client PEP can utilize mulitcast between them generally without any changes to the origin server or browser. This means that a variety of business rules can be applied by the gateway PEP to determine whether to mulitcast or unicast responses to client PEPs. For example, the gateway PEP could track popular content and mulitcast content responses for these URLs. Alternatively, subscribers could indicate interest groups and when requests are made for content in those interest groups, the responses are mulitcast to the group, not just unicast to the

requestor. The advantage of this approach is that the same amount of bandwidth can be consumed by unicast traffic as by multicast traffic and, by doing this with the PEPs, the activity is transparent to the end points.

FIGURE 6 is a data flow diagram illustrating an alternate data flow between the components of FIGURE 2. Generally, steps 185-245 of FIGURE 6 are similar to the correspondingly numbered steps in FIGURE 5. In the data flow of FIGURE 6, however, certain objects were not subjected to a prefetch in step 230 and not provided to the user in step 245. Thus, in this embodiment, the client PEP passes requests for these non-prefetched objects to the gateway PEP (step 250). The gateway PEP then passes the requests to the origin server (step 255). The origin server can then retrieve the objects corresponding to the request and transmit those objects to the gateway PEP (step 260). The gateway PEP can then pass the objects to the client PEP (265), and finally, the client PEP can pass the objects to the user (step 270) for combination with the objects received in step 245.

Referring now to FIGURE 7, it is a data flow diagram illustrating a method of optimizing bandwidth in accordance with one embodiment of the present invention. As with the embodiment shown in FIGURE 5, the user initially passes a request, e.g., get cnn.com, to the client PEP (step 185). The client PEP intercepts this request and determines if the request--or a portion of the request--has been previously filled (step 280). Furthermore, the client PEP can generate a time stamp list corresponding to previously fetched and stored objects (step 280). The client PEP can pass both the request and the list of stored objects and time stamps to the gateway PEP (step 285), and the gateway PEP can pass the request and list on to the origin server (step 290).

After receiving a request and a list of previously fetched objects, the origin server can return an index page and object freshness information to the gateway PEP (step 295). The gateway PEP can then parse the index page, review the freshness information, and generate a prefetch list based on the index page and freshness information (step 300). The gateway PEP can then pass the index page, the prefetch list, and/or the freshness information to the client PEP (step 305). Next, the client PEP can pass the index page and the prefetch list on to the user (step 310). Responsive to receiving the index page and the prefetch list, the user requests the embedded objects identified in the index page (step 315). The client PEP responds to this request by

returning any fresh, locally stored data objects that correspond to the user's requests (step 315).

In parallel with certain previously described steps, the gateway PEP prefetches stale and other objects needed to fill the user's requests (step 325). These objects are returned by the origin server and passed through to the client PEP to the user (steps 330, 335, and 340). The user can then combine the objects received in step 320 with the objects received in step 340, thereby filling the request made in step 315.

In another embodiment of the present invention (not illustrated), the gateway PEP rather than the client PEP can determine whether an initial request has been previously filled. In other words, the gateway PEP can keep track of what objects have been provided to the client PEP and whether or not those objects are still valid. Otherwise, the operation of this embodiment is similar to the operation described with relation to FIGURE 7. As those of skill in the art can understand, one advantage of this embodiment is that the amount of data that needs to be transferred can be reduced.

In conclusion, the present system provides, among other things, a system and method for reducing the impact of high latency links on interactive applications. Those skilled in the art, however, can readily recognize that numerous variations and substitutions may be made in the invention, its use and its configuration to achieve substantially the same results as achieved by the embodiments described herein. Accordingly, there is no intention to limit the invention to the disclosed exemplary forms. Many variations, modifications and alternative constructions fall within the scope and spirit of the disclosed invention as expressed in the claims.

WHAT IS CLAIMED IS:

1. A method for managing a data transmission between an origin server and a client, the method comprising:

receiving, at a client performance enhancing proxy (PEP), a request for a data page, wherein the request is generated by the client;

receiving, at the client PEP, an index page generated according to the origin server, wherein the index page identifies a plurality of data objects contained in the requested data page;

receiving, at the client PEP, a list of data objects that are subject to a prefetch from the origin server;

providing the index page to the client;

receiving, at the client PEP, a fetch request generated by the client, wherein the fetch request is configured to request a first of the plurality of data objects identified by the index page;

blocking, at the client PEP, the fetch request so that the fetch request is not provided to the origin server;

receiving, at the client PEP, the first of the plurality of data objects, wherein the first of the plurality of data objects is being provided according to the prefetch; and

providing the received data object to the client;

wherein the received data object at least partially fills the fetch request generated by the client.

2. The method of claim 1, wherein receiving the request for the data page comprises:

receiving a request for a Web page.

3. The method of claim 1, further comprising:

providing the request for a data page to the origin server via a satellite network.

4. The method of claim 1, further comprising:

providing the request for a data page to the origin server via a wireless network.

5. The method of claim 1, wherein receiving an index page generated according to the origin server comprises:

receiving an index page generated by the origin server.

6. The method of claim 1, further comprising:

comparing the list of data objects with the index page to determine which of the plurality of data objects identified in the index page are subject to the prefetch.

7. The method of claim 1, further comprising:

storing a list of data objects available locally to the client PEP.

8. The method of claim 7, wherein the fetch request is a first fetch request, the method further comprising:

receiving, at the client PEP, a second fetch request generated by the client, wherein the second fetch request is configured to request a second of the plurality of data objects identified by the index page;

retrieving the second of the plurality of data objects from a cache; and

providing the second of the plurality of data objects to the client PEP;

wherein an identifier for the second of the plurality of data objects is included in the list of data objects available locally to the client.

9. The method of claim 8, further comprising;

receiving, at the client PEP, freshness information related to the second of the plurality of data objects; and

providing the second of the plurality of data objects to the client PEP responsive to the freshness information indicating that the second of the plurality of data objects is fresh.

10. A method for managing a data transmission between an origin server and a client, the method comprising:

receiving an index page identifying a plurality of data objects contained in a data page;

receiving a list of data objects identifying a first data object included in the plurality of data objects, the first data object subject to a prefetch through the origin server;

receiving a fetch request generated by the client, wherein the fetch request is configured to request the first of the plurality of data objects; and

blocking the fetch request such that the fetch request is not provided to the origin server.

11. The method of claim 10, wherein receiving the index page comprises:

receiving the index page from the origin server.

12. The method of claim 10, further comprising:

comparing the list of data objects with the index page to determine which of the plurality of data objects identified in the index page are not subject to the prefetch.

13. The method of claim 12, wherein the fetch request is a first fetch request and wherein the method further comprises:

receiving a second fetch request, wherein the second fetch request is for one of the plurality of data objects determined not to be subject to the prefetch; and

providing the second fetch request to the origin server.

14. The method of claim 13, further comprising:

receiving the data object corresponding to the second fetch request; and

providing the data object corresponding to the second fetch request to the client.

15. The method of claim 10, further comprising:

storing a list of data objects available locally to the client.

16. The method of claim 15, wherein the fetch request is a first fetch request, the method further comprising:

receiving a second fetch request generated by the client, wherein the second fetch request is configured to request a second of the plurality of data objects;

retrieving the second of the plurality of data objects from a local storage device; and

providing the second of the plurality of data objects to the client;

wherein an identifier for the second of the plurality of data objects is included in the list of data objects available locally to the client.

17. The method of claim 16, further comprising;

receiving freshness information related to the second of the plurality of data objects; and

providing the second of the plurality of data objects to the client responsive to the freshness information indicating that the second of the plurality of data objects is fresh.

18. A system method for managing a data transmission, the system comprising:

a satellite modem configured to transfer information between an origin server and a client; and

a client-side PEP coupled to the satellite modem, the client side PEP configured to intercept a data page fetch request generated by the client and to block the data page fetch request, thereby preventing the data page fetch request from being sent between the client and the origin server.

19. The system of claim 18, further comprising:

a satellite dish coupled to the satellite modem.

20. The system of claim 18, wherein the satellite modem and the client-side PEP are integrated.

21. The system of claim 18, further comprising:

a browser;

wherein the client-side PEP and the browser are integrated.

22. The system of claim 18, further comprising:

a multipoint distribution system coupled to the client-side PEP;

wherein a plurality of users are connectable to the mulitpoint distribution system.

23. The system of claim 22, wherein the multipoint distribution system comprises one of:

a multichannel, multipoint distribution system (MMDS) and a local multipoint distribution system (LMDS).

24. The system of claim 18, further comprising:

a satellite network coupled to the satellite modem, the satellite network configured to transmit data to the satellite modem.

25. The system of claim 24, wherein the satellite network comprises:

a gateway satellite dish; and

a data transmission satellite.

26. The system of claim 18, wherein the client comprises:

a browser.

27. The system of claim 18, further comprising:

a local storage device configured to store a plurality of fetched data objects.

28. A system for managing a data transmission between an origin server and a client, the system comprising:

at least a first processor device;

at least a first memory device coupled to the at least a first processor device;

a plurality of instructions stored on the at least a first memory device, the plurality of instructions configured to cause the at least a first processor to:

process a request for a data page, wherein the request is generated by the client;

process an index page generated according to the origin server, wherein the index page identifies a plurality of data objects contained in the requested data page;

process a list of data objects that are subject to a prefetch through the origin server;

provide the index page to the client;

process a fetch request generated by the client, wherein the fetch request is configured to request a first of the plurality of data objects identified by the index page;

block the fetch request such that the fetch request is not provided to the origin server;

process the first of the plurality of data objects identified by the index page, wherein the first of the plurality of data objects is being provided by the origin server according to the prefetch; and

provide the received data object to the client.

29. The system of claim 28, wherein the plurality of instructions are further configured to cause the at least a first processor to:

provide the request to the origin server via a satellite network.

30. The system of claim 28, wherein the plurality of instructions are further configured to cause the at least a first processor to:

compare the list of data objects with the index page to determine which of the plurality of data objects identified in the index page are subject to the prefetch.

31. The system of claim 30, wherein the plurality of instructions are further configured to cause the at least a first processor to:

process a second fetch request, wherein the second fetch request is for a second of the plurality of data objects subject to the prefetch; and

provide the second fetch request to the origin server.

32. The system of claim 31, wherein the plurality of instructions are further configured to cause the at least a first processor to:

process the data object corresponding to the second prefetch request; and

provide the data object corresponding to the second prefetch request to the client.

33. The system of claim 28, wherein the plurality of instructions are further configured to cause the at least a first processor to:

store a list of data objects available locally to the client.

34. The system of claim 33, wherein the fetch request is a first fetch request and wherein the plurality of instructions are further configured to cause the at least a first processor to:

process a second fetch request generated by the client, wherein the second fetch request is configured to request a second of the plurality of data objects;

retrieve the second of the plurality of data objects from a local storage device; and

provide the second of the plurality of data objects to the client;

wherein an identifier for the second of the plurality of data objects is included in the list of data objects available locally to the client.

35. The system of claim 33, wherein the plurality of instructions are further configured to cause the at least a first processor to:

receive freshness information related to the second of the plurality of data objects; and

provide the second of the plurality of data objects to the client responsive to the freshness information indicating that the second of the plurality of data objects is fresh.

36. A system for managing a data transmission between an origin server and a client, the system comprising:

means for receiving an index page identifying a plurality of data objects contained in a data page;

means for receiving a list of data objects identifying a first data object included in the plurality of data objects, the first data object subject to a prefetch through the origin server;

means for receiving a fetch request generated by the client, wherein the fetch request is configured to request the first of the plurality of data objects; and

means for blocking the fetch request such that the fetch request is not provided to the origin server.

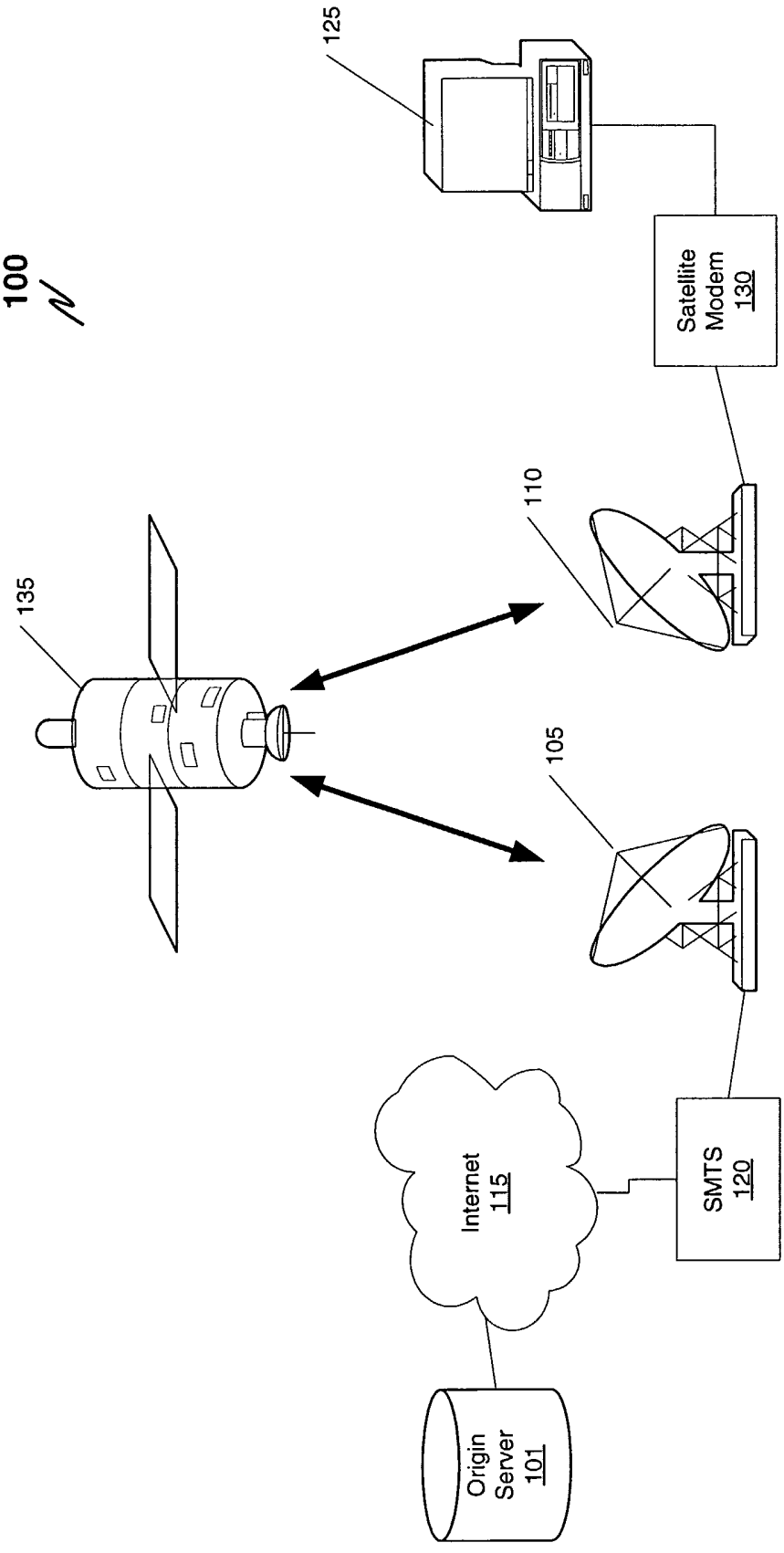


FIG. 1
(prior art)

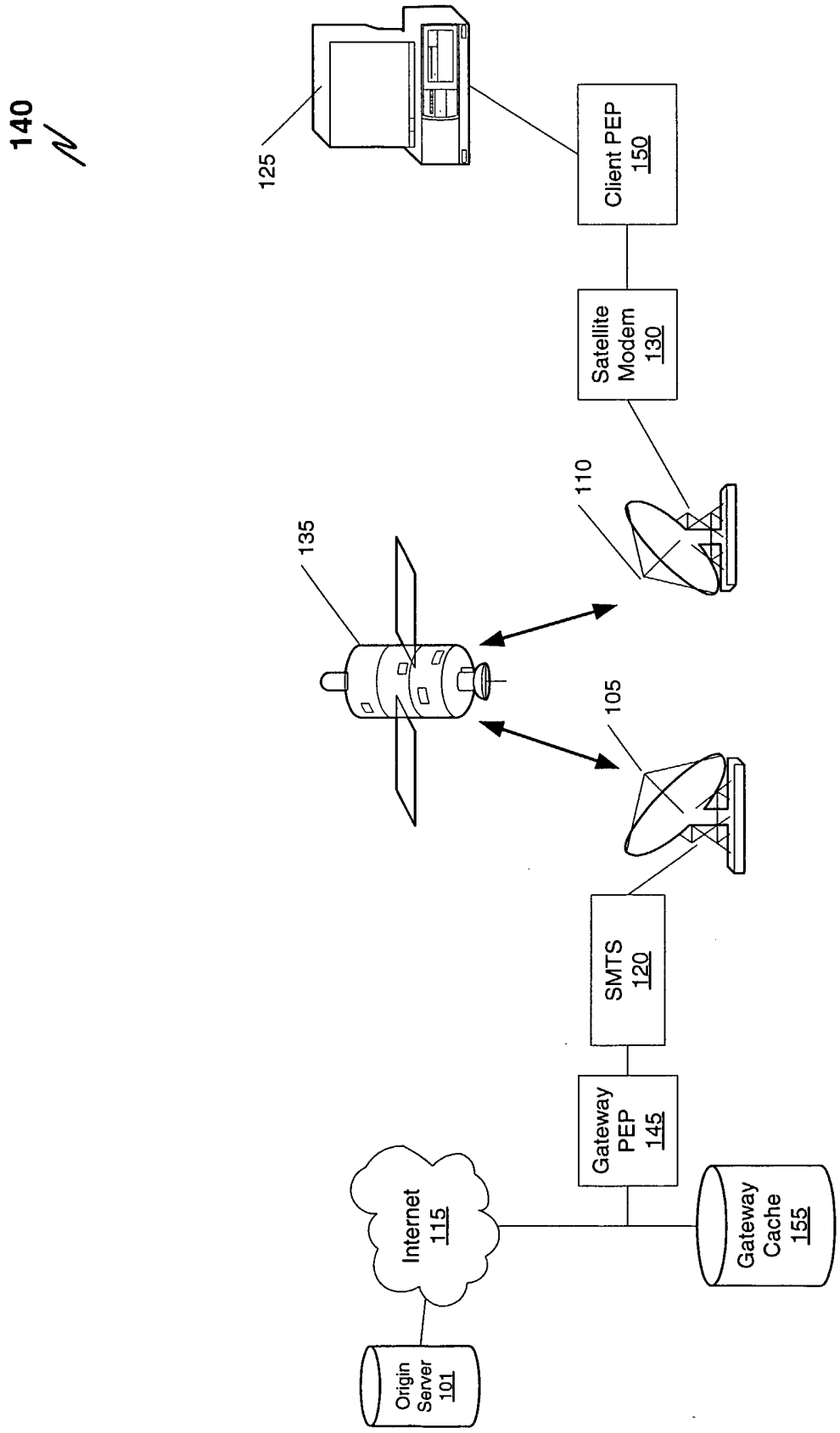

140
125
Client PEP 150
Satellite Modem 130
110
135
105
SMTS 120
Gateway PEP 145
Internet 115
Gateway Cache 155
Origin Server 101

FIG. 2

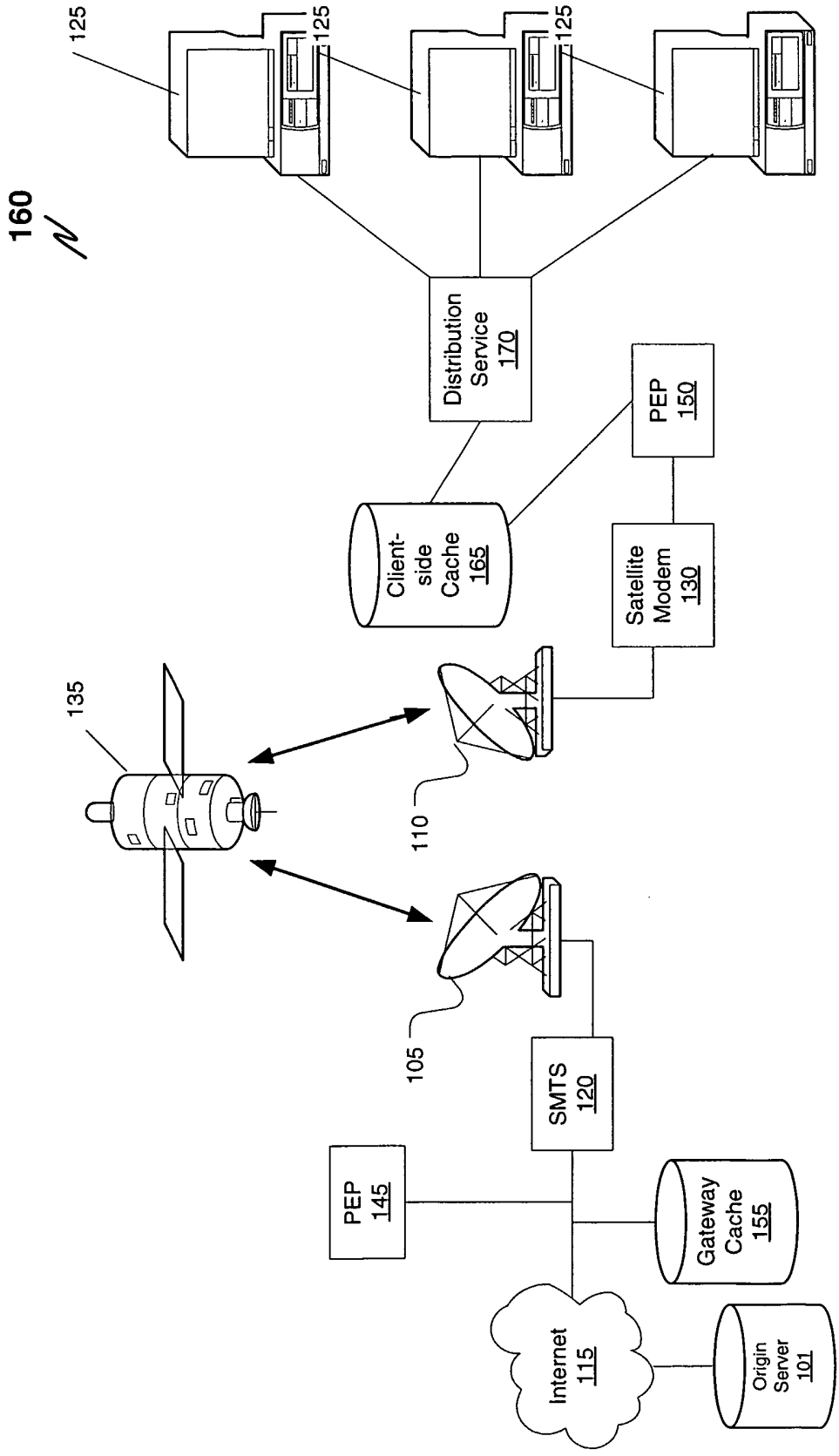

160
125
125
125
Distribution Service 170
PEP 150
Client-side Cache 165
Satellite Modem 130
135
110
105
SMTS 120
PEP 145
Gateway Cache 155
Internet 115
Origin Server 101

FIG. 3

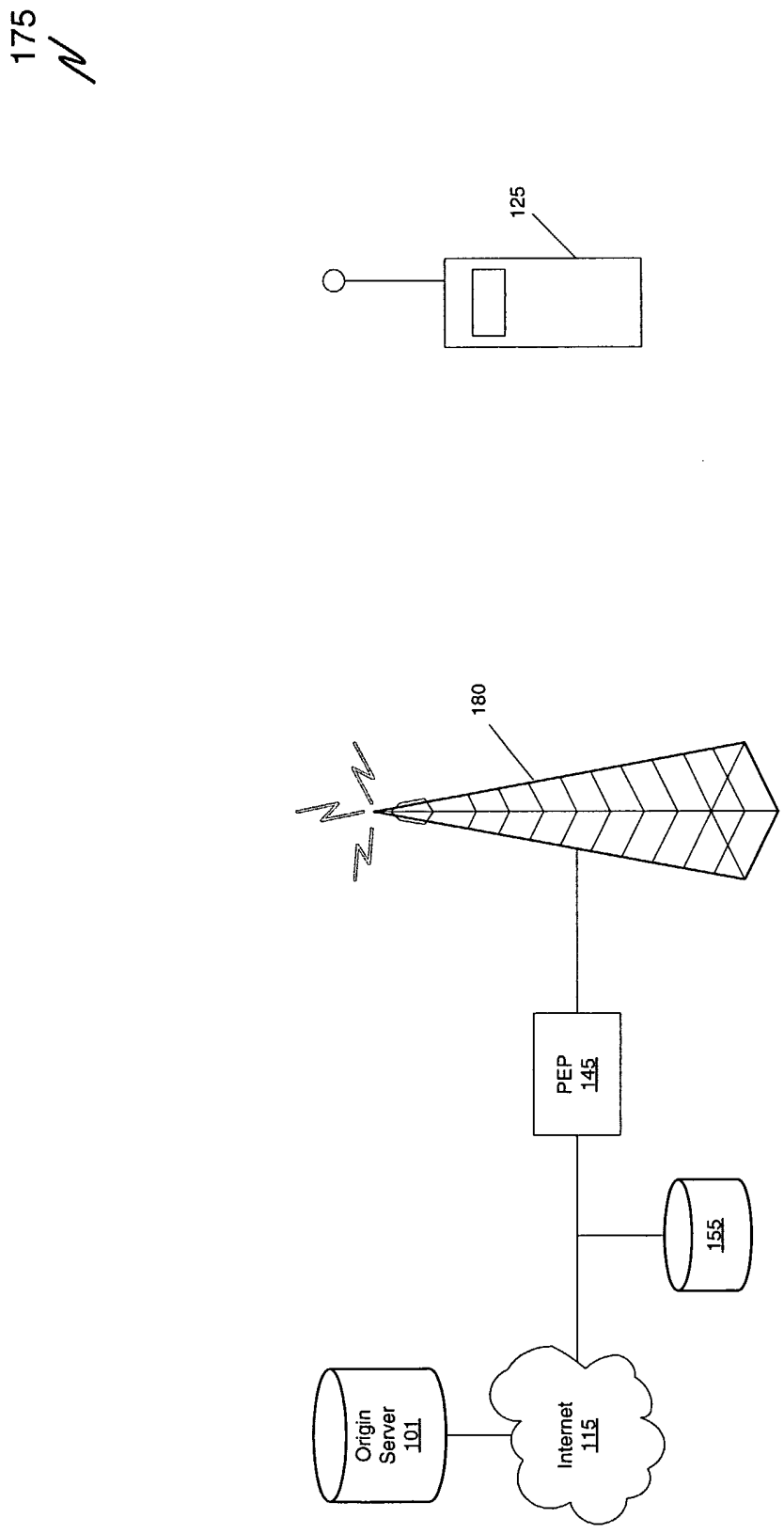


FIG. 4

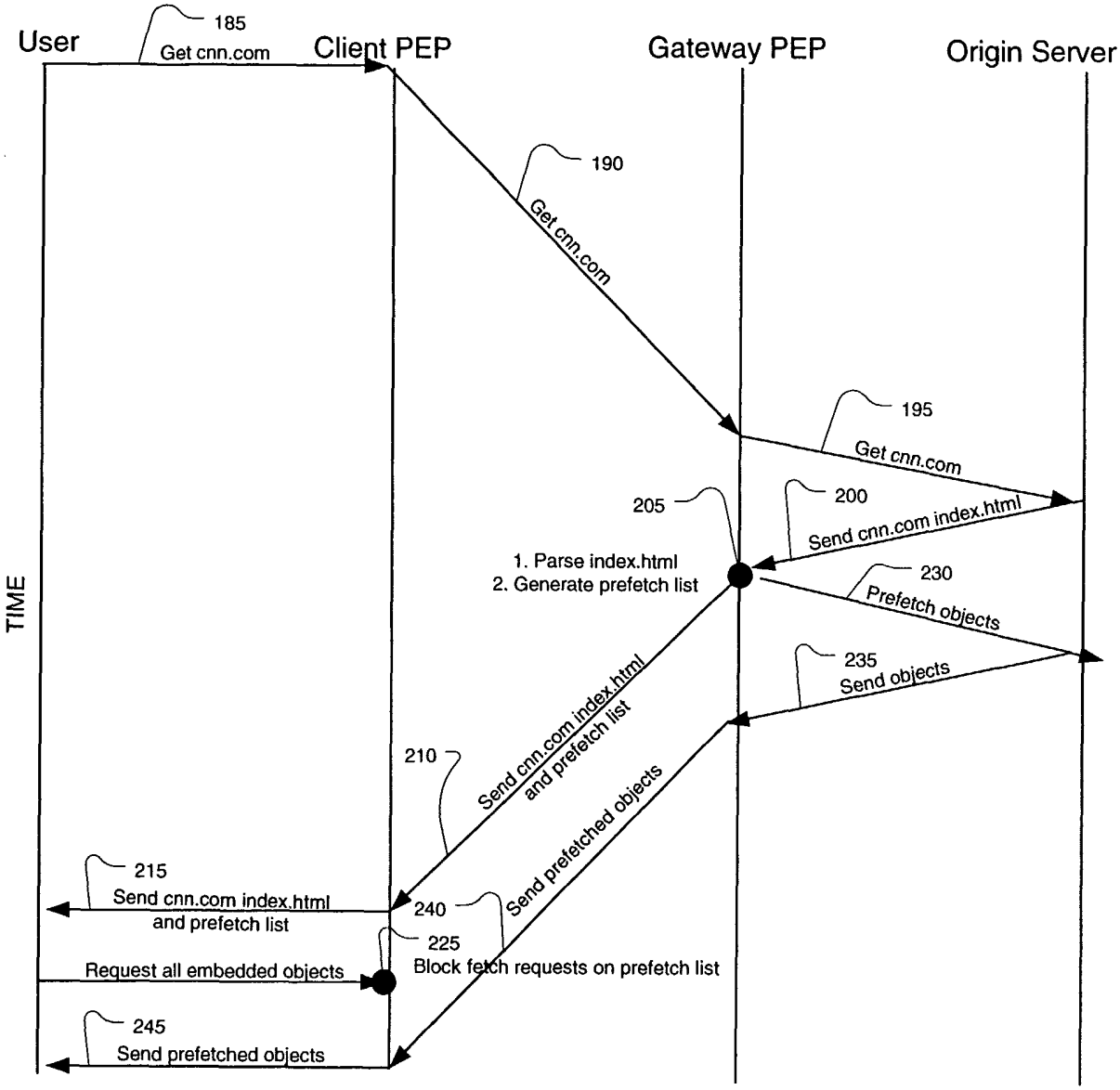

User
Client PEP
Gateway PEP
Origin Server
TIME
185
Get cnn.com
190
Get cnn.com
195
Get cnn.com
200
Send cnn.com index.html
205
1. Parse index.html
2. Generate prefetch list
230
Prefetch objects
235
Send objects
210
Send cnn.com index.html
and prefetch list
Send prefetched objects
215
Send cnn.com index.html
and prefetch list
240
225
Block fetch requests on prefetch list
Request all embedded objects
245
Send prefetched objects

FIG. 5

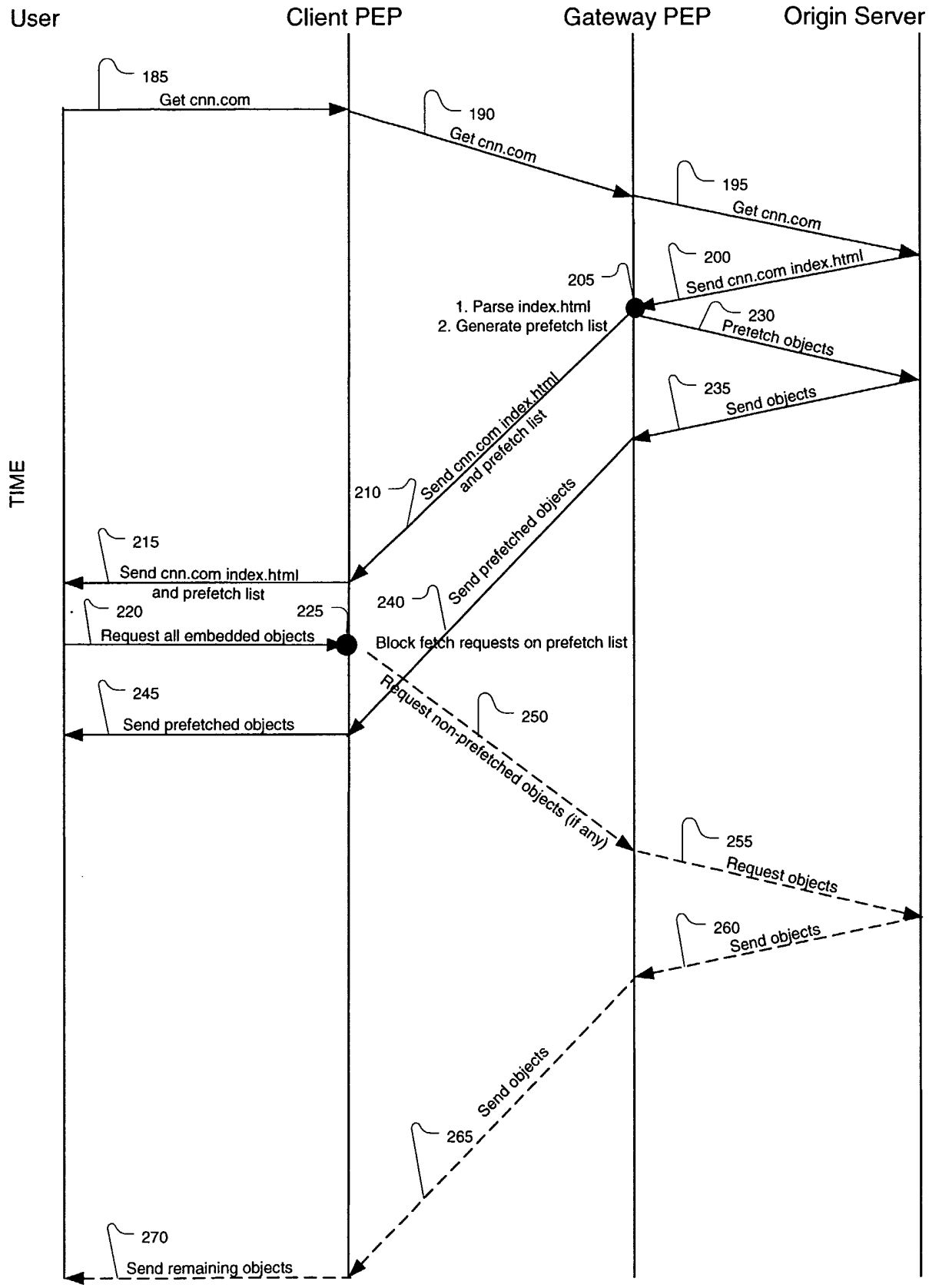

User
Client PEP
Gateway PEP
Origin Server
TIME
185
Get cnn.com
190
Get cnn.com
195
Get cnn.com
200
Send cnn.com index.html
205
1. Parse index.html
2. Generate prefetch list
230
Prefetch objects
235
Send objects
210
Send cnn.com index.html and prefetch list
Send prefetched objects
215
Send cnn.com index.html and prefetch list
220
Request all embedded objects
225
240
Block fetch requests on prefetch list
245
Send prefetched objects
250
Request non-prefetched objects (if any)
255
Request objects
260
Send objects
Send objects
265
270
Send remaining objects

FIG. 6

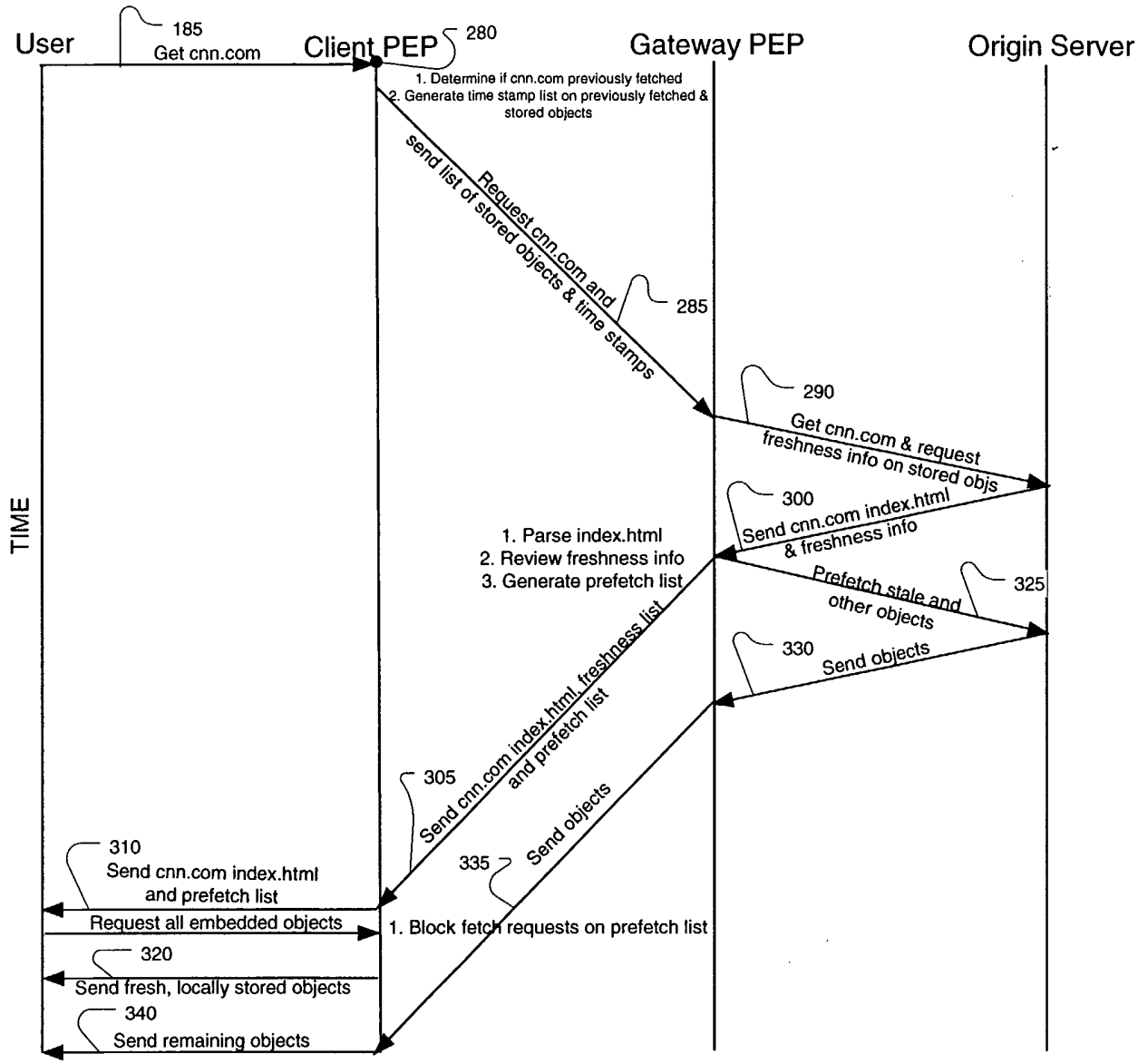

User
185
Get cnn.com
Client PEP
280
Gateway PEP
Origin Server
1. Determine if cnn.com previously fetched
2. Generate time stamp list on previously fetched & stored objects
Request cnn.com and send list of stored objects & time stamps
285
290
Get cnn.com & request freshness info on stored objs
300
Send cnn.com index.html & freshness info
1. Parse index.html
2. Review freshness info
3. Generate prefetch list
Prefetch stale and other objects
325
330
Send objects
Send cnn.com index.html, freshness list and prefetch list
305
Send objects
335
310
Send cnn.com index.html and prefetch list
Request all embedded objects
1. Block fetch requests on prefetch list
320
Send fresh, locally stored objects
340
Send remaining objects
TIME

FIG. 7

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US02/31907

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : G06F 17/30
US CL : 709/218, 219; 707/10, 501.1, 513

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 709/218, 219; 707/10, 501.1, 513

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X --- Y	US 6,282,542 B1 (CARNEAL et al) 29 August 2001 (29.08.01), Figs. 5-7; Abstract; col.6, lines 16-34; col.7, lines 16-65; col.8, lines 22-60; col.11, lines 44-61.	1-7, 10-12, 11, 18-22, 24-30, 33 and 36 --- 8-9, 13-14, 16-17, 23, 31-32 and 34-35
Y	US 6,003,087 A (HOUSEL, III et al) 14 December 1999, Figs. 2 and 8; Abstract; col.3, line 15 - col.5, line27.	8-9, 13-14, 16-17, 23, 31-32 and 34-35

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
01 December 2002 (01.12.2002)

Date of mailing of the international search report
16 DEC 2002

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231
Facsimile No. (703)305-3230

Authorized officer
Meng-Ai An
Telephone No. (703)305-3900

Form PCT/ISA/210 (second sheet) (July 1998)