

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
8 October 2009 (08.10.2009)

(10) International Publication Number
WO 2009/124011 A2

- (51) International Patent Classification:
G06F 15/16 (2006.01)
- (21) International Application Number:
PCT/US2009/038891
- (22) International Filing Date:
31 March 2009 (31.03.2009)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
12/060,213 31 March 2008 (31.03.2008) US
12/189,019 8 August 2008 (08.08.2008) US
- (71) Applicant (for all designated States except US): **AMAZON TECHNOLOGIES, INC.** [US/US]; P.O. Box 8102, Reno, NV 89507 (US).
- (72) Inventors (for all designated States except US): **SIVASUBRAMANIAN, Swaminathan**; 1200 12th Avenue South, Suite 1200, Seattle, WA 98144 (US). **RICHARDSON, David, R.**; 1200 12th Avenue South, Suite 1200, Seattle, WA 98144 (US). **MARSHALL, Bradley, E.**; 1200 12th Avenue South, Suite 1200, Seattle, WA 98144 (US). **SCOFIELD, Christopher, L.**; 1200 12th Avenue South, Suite 1200, Seattle, WA 98144 (US).
- (74) Agent: **AMAZON TECHNOLOGIES, INC.**; 8329 West Sunset Road, Suite 220, Las Vegas, NV 89113 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— without international search report and to be republished upon receipt of that report (Rule 48.2(g))

(54) Title: CONTENT MANAGEMENT

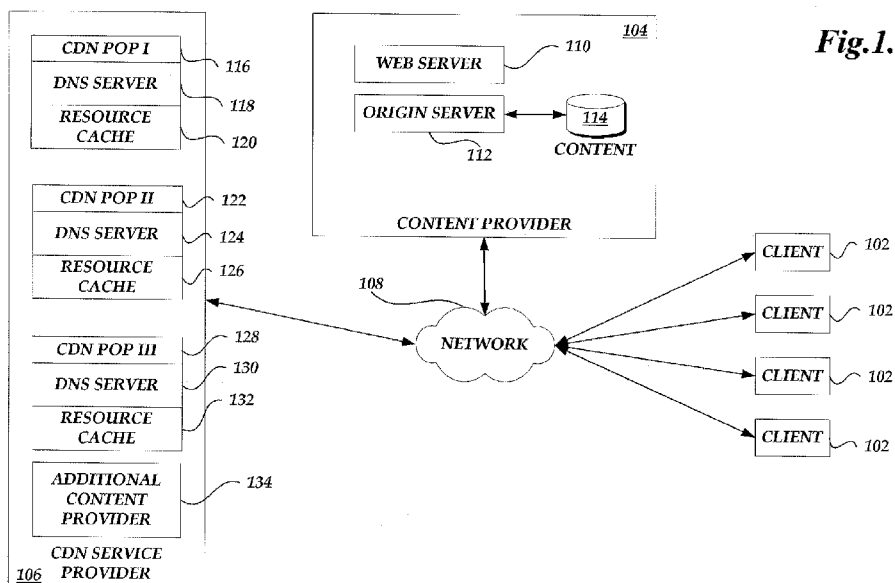


Fig. 1.

(57) Abstract: A system and method for management and processing of resource requests is provided. A content delivery network service provider determines a class associated with a set of client computing devices and monitors resources requests for the determined class. The content delivery network service provider then identifies at least one cache component for providing content to client computing devices as a function of the determined class. In other embodiments, instead of cache components, the content delivery network service provider identifies a second set of client computing devices as a function of the determined class for providing the content information.

WO 2009/124011 A2

CONTENT MANAGEMENT

BACKGROUND

[0001] Generally described, computing devices and communication networks can be utilized to exchange information. In a common application, a computing device can request content from another computing device via the communication network. For example, a user at a personal computing device can utilize a software browser application to request a Web page from a server computing device via the Internet. In such embodiments, the user computing device can be referred to as a client computing device and the server computing device can be referred to as a content provider.

[0002] Content providers are generally motivated to provide requested content to client computing devices often with consideration of efficient transmission of the requested content to the client computing device and/or consideration of a cost associated with the transmission of the content. For larger scale implementations, a content provider may receive content requests from a high volume of client computing devices which can place a strain on the content provider's computing resources. Additionally, the content requested by the client computing devices may have a number of components, which can further place additional strain on the content provider's computing resources.

[0003] With reference to an illustrative example, a requested Web page, or original content, may be associated with a number of additional resources, such as images or videos, that are to be displayed with the Web page. In one specific embodiment, the additional resources of the Web page are identified by a number of embedded resource identifiers, such as uniform resource locators ("URLs"). In turn, software on the client computing devices typically processes embedded resource identifiers to generate requests for the content. Often, the resource identifiers associated with the embedded resources reference a computing device associated with the content provider such that the client computing device would transmit the request for the additional resources to the referenced content provider computing device. Accordingly, in order to satisfy a content request, the content provider would provide client computing devices data associated with the Web page as well as the data associated with the embedded resources.

[0004] Some content providers attempt to facilitate the delivery of requested content, such as Web pages and/or resources identified in Web pages, through the utilization of a content delivery network (“CDN”) service provider. A CDN server provider typically maintains a number of computing devices in a communication network that can maintain content from various content providers. In turn, content providers can instruct, or otherwise suggest to, client computing devices to request some, or all, of the content provider’s content from the CDN service provider’s computing devices.

[0005] As with content providers, CDN service providers are also generally motivated to provide requested content to client computing devices often with consideration of efficient transmission of the requested content to the client computing device and/or consideration of a cost associated with the transmission of the content. Accordingly, CDN service providers often consider factors such as latency of delivery of requested content in order to meet service level agreements or to generally improve the quality of delivery service.

DESCRIPTION OF THE DRAWINGS

[0006] The foregoing aspects and many of the attendant advantages of this invention will become more readily appreciated as the same become better understood by reference to the following detailed description, when taken in conjunction with the accompanying drawings, wherein:

[0007] Figure 1 is a block diagram illustrative of content delivery environment including a number of client computing devices, a content provider, and a content delivery network service provider;

[0008] Figure 2 is a block diagram of the content delivery environment of Figure 1 illustrating the generation and processing of a content request from a client computing device to a content provider;

[0009] Figure 3 is a block diagram of the content delivery environment of Figure 1 illustrating the generation and processing of a DNS query corresponding to an embedded resource from a client computing device to a content delivery network service provider;

[0010] Figure 4 is a block diagram of the content delivery environment of Figure 1 illustrating the generation and processing of embedded resource requests from a client computing device to a content delivery network service provider;

[0011] Figure 5 is a block diagram of the content delivery environment of Figure 1 illustrating the processing and management of associated class information collected by a resource cache component;

[0012] Figure 6 is a block diagram of the content delivery environment of Figure 1 illustrating another embodiment of the processing and management of associated class information collected by a resource cache component;

[0013] Figure 7 is a block diagram of the content delivery environment of Figure 1 illustrating yet another embodiment of the processing and management of associated class information collected by a resource cache component;

[0014] Figure 8 is a flow diagram illustrative of a content monitoring routine implemented by a content delivery network service provider for monitoring resource requests from client computing devices; and

[0015] Figure 9 is a flow diagram illustrative a content processing routine implemented by a content delivery network service provider for managing content for delivery to a class of client computing devices.

DETAILED DESCRIPTION

[0016] Generally described, the present disclosure is directed to the management and processing of resource requests made to a content delivery network (“CDN”) service provider from client computing devices. Specifically, aspects of the disclosure will be described with regard to monitoring resource requests from a cluster of client computing devices and managing content for delivery to one or more client computing devices based on the monitored resource requests for the cluster. Further aspects of the disclosure are directed to monitoring resource requests corresponding to classes associated with client computing devices based on such monitored resource requests. Although various aspects of the disclosure will be described with regard to illustrative examples and embodiments, one skilled in the art will appreciate that the disclosed embodiments and examples should not be construed as limiting.

[0017] Figure 1 is a block diagram illustrative of content delivery environment 100 for the management and processing of content requests. As illustrated in Figure 1, the content delivery environment 100 includes a number of client computing devices 102 (generally referred to as clients) for requesting content from a content provider and/or a CDN service provider. In an illustrative embodiment, the client computing devices 102 can correspond to a wide variety of computing devices including personal computing devices, laptop computing devices, hand-held computing devices, terminal computing devices, mobile devices, wireless devices, various electronic devices and appliances and the like. In an illustrative embodiment, the client computing devices 102 include necessary hardware and software components for establishing communications over a communication network 108, such as a wide area network or local area network. For example, the client computing devices 102 may be equipped with networking equipment and browser software applications that facilitate communications via the Internet or an intranet.

[0018] Although not illustrated in Figure 1, each client computing device 102 utilizes some type of local DNS resolver component, such as a DNS Name server, that generates the DNS queries attributed to the client computing device. In one embodiment, the local DNS resolver component may be provided by an enterprise network to which the client computing device 102 belongs. In another embodiment, the local DNS resolver component may be provided by an Internet Service Provider (ISP) that provides the communication network connection to the client computing device 102.

[0019] The content delivery environment 100 can also include a content provider 104 in communication with the one or more client computing devices 102 via the communication network 108. The content provider 104 illustrated in Figure 1 corresponds to a logical association of one or more computing devices associated with a content provider. Specifically, the content provider 104 can include a web server component 110 corresponding to one or more server computing devices for obtaining and processing requests for content (such as Web pages) from the client computing devices 102. The content provider 104 can further include an origin server component 112 and associated storage component 114 corresponding to one or more computing devices for obtaining and processing requests for network resources from the CDN

service provider. One skilled in the relevant art will appreciate that the content provider 104 can be associated with various additional computing resources, such additional computing devices for administration of content and resources, DNS name servers, and the like. For example, although not illustrated in Figure 1, the content provider 104 can be associated with one or more DNS name server components that would be authoritative to resolve client computing device DNS queries corresponding to a domain of the content provider.

[0020] With continued reference to Figure 1, the content delivery environment 100 can further include a CDN service provider 106 in communication with the one or more client computing devices 102 and the content providers 104 via the communication network 108. The CDN service provider 106 illustrated in Figure 1 corresponds to a logical association of one or more computing devices associated with a CDN service provider. Specifically, the CDN service provider 106 can include a number of Point of Presence (“POP”) locations 116, 122, 128 that correspond to nodes on the communication network 108. Each POP 116, 122, 128 includes a DNS component 118, 124, 130 made up of a number of DNS server computing devices for resolving DNS queries from the client computers 102. Each POP 116, 122, 128 also includes a resource cache component 120, 126, 132 made up of a number of cache server computing devices for storing resources from content providers and transmitting various requested resources to various client computers. The DNS components 118, 124 and 130 and the resource cache components 120, 126, 132 may further include additional software and/or hardware components that facilitate communications including, but not limited, load balancing or load sharing software/hardware components.

[0021] In an illustrative embodiment, the DNS component 118, 124, 130 and resource cache component 120, 126, 132 are considered to be logically grouped, regardless of whether the components, or portions of the components, are physically separate. Additionally, although the POPs 116, 122, 128 are illustrated in Figure 1 as logically associated with the CDN service provider 106, the POPs will be geographically distributed throughout the communication network 108 in a manner to best serve various demographics of client computing devices 102. Additionally, one skilled in the relevant art will appreciate that the CDN service provider 106 can

be associated with various additional computing resources, such as an additional content provider 134, additional computing devices for administration of content and resources, and the like.

[0022] One skilled in the relevant art will appreciate that the components and configurations provided in Figure 1 are illustrative in nature. Accordingly, additional or alternative components and/or configurations, especially regarding the additional components, systems and subsystems for facilitating communications may be utilized.

[0023] With reference now to Figures 2-4, the interaction between various components of the content delivery environment 100 of Figure 1 will be illustrated. For purposes of the example, however, the illustration has been simplified such that many of the components utilized to facilitate communications are not shown. One skilled in the relevant art will appreciate that such components can be utilized and that additional interactions would accordingly occur without departing from the spirit and scope of the present disclosure.

[0024] With reference to Figure 2, a client computing device 102 generates a content request that is received and processed by the content provider 104, such as through the Web server 110. In accordance with an illustrative embodiment, the request for content can be in accordance with common network protocols, such as the hypertext transfer protocol ("HTTP"). Upon receipt of the content request, the content provider 104 identifies the appropriate responsive content. In an illustrative embodiment, the requested content can correspond to a Web page that is displayed on the client computing device 102 via the processing of information, such as hypertext markup language ("HTML"), extensible markup language ("XML"), and the like. The requested content can also include a number of embedded resource identifiers that corresponds to resource objects that should be obtained by the client computing device 102 as part of the processing of the requested content. The embedded resource identifiers can be generally referred to as resource identifiers or resource URLs.

[0025] In one embodiment, the resource URLs identify the domain of the CDN service provider 106 (e.g., "cdnprovider.com"), the same name of the resource to be requested (e.g., "resource.xxx") and the same path where the resource will be found (e.g., "path"). Additionally, the resource URL can include additional processing information (e.g., "additional information"). The resource URL would have the form of:

http://additional_information.cdnprovider.com/path/resources.xxx

[0026] In another embodiment, the information associated with the CDN service provider 106 is included with the resource URL, such as through prepending or other techniques, such that the resource URL can maintain all of the information associated with a URL previously associated with the content provider 104. In this embodiment, the resource URL would have the form of:

http://additional_information.cdnprovider.com/www.contentprovider.com/path/resource.xxx

[0027] Upon receipt of the requested content, the client computing device 102, such as through a browser software application, begins processing any of the markup code included in the content and attempts to acquire the resources identified by the embedded resource identifiers. Accordingly, the first step in acquiring the content corresponds to the issuance, by the client computing device 102 (through its local DNS resolver), of a DNS query for the Resource URL resource identifier that results in the identification of a DNS server authoritative to the “.” and the “com” portions of the translated URL. After resolving the “.” and “com” portions of the embedded URL, the client computing device 102 then issues a DNS query for the resource URL that results in the identification of a DNS server authoritative to the “.cdnprovider” portion of the embedded URL. The issuance of DNS queries corresponding to the “.” and the “com” portions of a URL are well known and have not been illustrated.

[0028] With reference now to Figure 3, in an illustrative embodiment, the successful resolution of the “.cdnprovider” portion of the resource URL identifies a network address, such as an IP address, of a DNS server associated with the CDN service provider 106. In one embodiment, the IP address can be a specific network address unique to a DNS server component of a POP. In another embodiment, the IP address can be shared by one or more POPs. In this embodiment, a further DNS query provided to the shared IP address utilizes a one-to-many network routing schema, such as anycast, such that a specific POP will receive the

request as a function of network topology. For example, in an anycast implementation, a DNS query issued by a client computing device 102 to a shared IP address will arrive at a DNS server component logically having the shortest network topology distance, often referred to as network hops, from the client computing device. The network topology distance does not necessarily correspond to geographic distance. However, in some embodiments, the network topology distance can be inferred to be the shortest network distance between a client computing device 102 and a POP. It will be appreciated by one skilled in the relevant art that a number of ways exist to determine network topology distance.

[0029] With continued reference to Figure 3, in either of the above identified embodiments (or any other embodiment), a specific DNS server in the DNS component 118 of a POP 116 receives the DNS query corresponding to the resource URL from the client computing device 102. Once one of the DNS servers in the DNS component 118 receives the request, the specific DNS server attempts to resolve the request. In one illustrative embodiment as shown in Figure 3, a specific DNS server resolves the DNS query by identifying an IP address of a resource cache component that will process the request for the requested resource. Upon selection of a resource cache component 120, 126, 132, the DNS server component 118 provides an IP address of the cache server computing device, resource cache component or load balancer/load share device associated with a resource cache component.

[0030] With reference now to Figure 4, the client computing device 102 can then utilize Internet communication protocols to request the resource from a resource cache component identified by the IP address. The resource cache component then processes the request, as will be described in greater detail below, and provides the resource to the client computing device 102. Upon receipt, the requested resource is processed by the browser application on the client computing device 102 as appropriate. It will be appreciated by one skilled in the relevant art that a selected resource cache component 120, 126, 132 can either provide the requested resource if it is available, as described above, or attempt to obtain the requested resource from another source, such as a peer cache server computing device or the origin server 112 of the content provider 104.

[0031] With continued reference to Figure 4, in conjunction with processing the resource request, the resource cache component determines a class associated with the requesting client computing device. In one embodiment, the resource cache component utilizes information received from the client request to determine the class. The class can be determined from the client directly (such as information provided by the client computing device or ISP) or indirectly (such as inferred through a client computing device's IP address). Additionally, the resource cache component can use additional information to determine the class, such as a time zone in which the requesting client computing device is located. The resource cache component associates the class information with the requested content and maintains the associated class information, for example in a log file, for sharing with the CDN service provider 106. In an illustrative embodiment, the determination of class can be statically determined based on predefined class definitions so that the association of the class can be done at the time of receiving resource requests. For example, a geographic location may be known at the time of request processing. In another embodiment, the determination of class can be dynamic and may not be determined (or determinable) at request processing. For example, behavioral classes based on observing behavior over a time window may not be determined until completion of the observation window's processing of the observed data.

[0032] In one illustrative embodiment, the class can correspond to a specific geographic region to which the client computing device belongs. In addition or alternatively, the class can correspond to a temporal domain, e.g., time zone, business hours, daytime/nighttime, holidays, and the like. In another embodiment, the class can correspond to a particular subject area, such as a news group, associated with either the requested content or the user or both. Still further the class can correspond to demographic data. Depending on the information to be collected from the requesting client computing device, the resource cache component can determine and maintain this information in a number of ways as will be appreciated by one skilled in the relevant art. Moreover, it will further be appreciated by one skilled in the relevant art that one or more computing devices at the POP where the resource request is received can perform any of the above described functions.

[0033] With reference now to Figure 5, an illustrative embodiment of how the CDN service provider 106 processes and manages the associated class information collected by the resource cache component 132 will be described. The resource cache component 132 provides the associated class information collected at the POP 128 to the CDN service provider 106. The associated class information can then be processed by the CDN service provider in a number of ways. In one embodiment, the CDN service provider can identify content information for providing to a computing device of the CDN based on the associated class information. The identified content information can include a list of frequently requested content based on the associated class information collected at the resource cache component 132. Alternatively, the identified content information can include the frequently requested content itself, as well as additional information associated with the frequently requested content such as network performance metrics, content provider service level agreements, and the like, and/or other corresponding content (for example, corresponding documents, advertisements, and the like, as will be described in further detail below with respect to Figure 7).

[0034] In addition to identifying content information, the CDN service provider 106 can identify a specific cache component 120 for providing this content information to. The identification of the specific cache component may also be based on the associated class information collected at the resource cache component 132. For example, where the class corresponds to a time zone, the CDN service provider may identify a resource cache component 120 that is authorized to service a client computing device having a different time zone. In another embodiment, the CDN service provider could simply send the identified content information to all or a subset of resource cache components of the CDN service provider and allow the cache components to determine how to manage the identified content information.

[0035] With continued reference to Figure 5, in accordance with the present illustrative embodiment, once the associated class information is processed by the CDN service provider, the identified content information is provided to the selected resource cache component 120 which in turn manages the identified content information provided by the CDN service provider 106. In one embodiment, the cache component 120 manages the identified content information by obtaining the content from a storage location in anticipation of subsequent

content requests at the cache component 120 for the identified content. In another embodiment, the cache component 120 can move the identified content from disk memory to memory of the cache component 120. In yet another embodiment, the cache component 120 can alter an expiration associated with the identified content in memory of the cache component 120 so as to ensure that the content remains available when later requested by a client computing device. In a yet further embodiment, the cache component 120 can simply verify the availability of the identified content at the cache component.

[0036] One skilled in the relevant art will appreciate that any various methodologies or algorithms may be utilized to process and manage the associated class information within the spirit and scope of the present disclosure. The following specific examples will be provided for further illustrative purposes.

[0037] In one embodiment, where the class information corresponds to content requests monitored for a set of client computing devices in a region having a first time zone, the CDN service provider processes the class information to identify a set of frequently requested resources in the first time zone. The CDN service provider then preloads content associated with the identified set of frequently requested resources at a cache component which is authorized to service a set of client computing devices in a different time zone. Accordingly, the CDN service provider is able to thereafter provide such preloaded content at lower delivery latencies. One skilled in the relevant art will further appreciate that the cache component authorized for servicing the set of client computing devices located in a different time zone can preload content at any suitable time based on processing constraints applicable to the individual cache component. In addition, it will be yet further appreciated by one skilled in the relevant art that the content may be wholly or partially preloaded at the cache component.

[0038] In another illustrative embodiment, the CDN service provider processes the associated class information to identify a list of frequently requested resources and provides this list to a set of resource cache components. Each resource cache component can then individually determine whether to preload one or more of the frequently requested resources.

[0039] In a further illustrative embodiment, where the content information is a content list derived from resource requests made by a first cluster of client computing devices

and where the content list has been provided to a cache component for servicing a second cluster of client computing devices, the CDN service provider can further monitor whether client computing devices in the second cluster subsequently request content from the list. The CDN service provider can then determine whether modifications should be made to content lists recommended by the first cluster to the second cluster based on the subsequent learned behavior of the second cluster. It will be appreciated by one skilled in the relevant art that a variety of logic may be implemented to monitor such behavior and to update recommendations based thereon.

[0040] In a still further illustrative embodiment, the CDN service provider monitors resource requests from a first cluster of client computing devices and identifies a set of content corresponding thereto. The CDN service provider can then manage the identified set of content at a resource cache component authorized to service a second cluster of client computing devices. As set forth above, there are a number of ways to manage the identified set of content, e.g., preloading the content at the resource cache component, verifying availability of content, managing cache control information related to content in memory (e.g., changing expiration data or manipulating expiration variables), and the like. It will further be appreciated by one skilled in the relevant art that the first and second clusters may be the same or different.

[0041] In yet another embodiment, the CDN service provider monitors resource requests associated with a class of computing devices and associates the resources requested with the class. The CDN service provider then, in one embodiment, preloads content at a resource cache component of the CDN as a function of the class. This function can involve a matching determination in which the CDN service provider matches a class of the client computing device with a class associated with the resources. For example, the CDN provider may match a topic of interest class, such as basketball scores, with resources that have been associated with the same class based on previous requests. Alternatively, the function can involve a correlation determination in which the CDN service provider correlates a class of the client computing device with a different class associated with the resources. For example, the CDN service provider may correlate a first geographic class, such as New York City, with another geographic class, such as San Francisco. In an illustrative embodiment, determination of matching or

correlation may be facilitated through a manual process, such as by an administrator, or based on observed behaviors.

[0042] With reference now to Figure 6, an alternative embodiment directed to how the CDN service provider 106 processes and manages the associated class information collected by the resource cache component 132 will now be described. This embodiment is virtually identical to the embodiments described above in reference to Figure 5 with the exception that instead of identifying a specific resource cache component for sending the identified content information to, a client computing device is identified. In addition, the subsequent processing functions performed by identified resource cache components in the above embodiments can now be similarly performed by identified client computing devices. In accordance with this alternative embodiment, the CDN service provider 106 includes a client component, such as a software application, for managing the identified content information at the client computing device. The CDN service provider 106 provides the client component to client computing devices desiring to participate in such a service.

[0043] With continued reference to Figure 6, in a specific illustrative embodiment, the resource cache component 132 provides the associated class information collected at the resource cache component 132 to the CDN service provider 106. In this embodiment, the associated class information includes data identifying content corresponding to requests made for a class associated with a first set of client computing devices. Upon receipt of this associated class information, the CDN service provider 106 identifies a second set of client computing devices as a function of the class of the first set of computing devices. The CDN service provider then provides content information related to the resource requests to a client computing device in the second set of client computing devices. In this illustrative embodiment, the content information includes a list of frequently requested content for the class associated with the first set of client computing devices. The client computing device in the second set of devices manages this content information (via a client side application provided by the CDN service provider) by verifying the availability of the listed content and obtaining one or more of the listed resources if not available.

[0044] With reference now to Figure 7, in one embodiment, the CDN service provider 106 may provide the class information collected from the POPs for additional processing. As previously illustrated in Figure 1, one or more additional content providers 134 may be part of the CDN service provider 106. The additional content providers 134 can correspond to additional components or services associated with the CDN service provider 106, such as advertisement engines or services, recommendation engines or services, security applications or services, and the like. Additionally, although the third party content providers 134 were previously illustrated as part of the CDN service provider 106, one skilled in the relevant art will appreciate that one or more of the third party content providers 134 may be independent.

[0045] Based on the class information provided by the CDN service provider 106 or otherwise made available by the CDN service provider 106, the third party content providers 134 can utilize the class information as inputs. For example, in one embodiment, if the third party content provider 134 corresponds to an advertisement engine or component, the class information can be processed to anticipate future advertisement display opportunities at the cache. Accordingly, the third party content provider 134 can process the input and the CDN service provider 106 can take additional steps. In another example, if the third party content provider 134 corresponds to a recommendation engine or component, the class information can be processed to anticipate additional content requests or future content request. Accordingly, the third party content provider 134 can process the input and the CDN service provider 106 can take additional steps. One skilled in the relevant art will appreciate that any various methodologies or algorithms may be utilized to process the class information within the spirit and scope of the present disclosure.

[0046] With continued reference to Figure 7, the additional content provider 134 can provide processed information back to the CDN service provider 106 for further processing. Additionally or alternatively, the additional content provider 134 can provide the processed information directly to one or more POPs, such as POP 116. In either embodiment, the additional action undertaken by the CDN service provider 106 can include notifications of additional advertisement and/or display opportunities, the identification of additional content that

can also be preloaded to a resource cache component, and/or the further management of content already stored at the resource cache component.

[0047] With reference now to Figure 8, one embodiment of a content monitoring routine 800 implemented by the CDN service provider 106 for monitoring resource requests from client computing devices will be described. One skilled in the relevant art will appreciate that actions/steps outlined for routine 800 may be implemented by one or many computing devices/components that are associated with the CDN service provider 106. Accordingly, routine 800 has been logically associated as being generally performed by the CDN service provider 106, and thus the following illustrative embodiments should not be construed as limiting.

[0048] At block 802, a resource cache component 132 receives a content request from a client computing device. Class information associated with the client computing device is then determined at block 804. In an illustrative embodiment, the determination of class information at block 804 can specifically include associating the requesting client computing device to a first cluster of other client computing devices based on a variety of criteria. In one example, the criteria correspond to a time zone in which the client computing devices are located. Continuing at block 806, the resource cache component further associates the identified class information with the requested content. Next, at block 808, the resource cache component 132 provides information to the CDN service provider 106 regarding the associated class. The routine ends at block 810.

[0049] With reference now to Figure 9, one embodiment of a content processing routine 900 for managing content for delivery to a class of client computing devices will be described. One skilled in the relevant art will appreciate that actions/steps outlined for routine 900 may be implemented by one or many computing devices/components that are associated with the CDN service provider 106. Accordingly, routine 900 has been logically associated as being performed by the CDN service provider 106.

[0050] At a block 902, the CDN service provider 106 receives information regarding associated class information, such as that provided from a resource cache component at block 810 of Figure 8. Next, at a block 904, the CDN service provider 106 processes the received class

information. As described above, a variety of logic for processing the received class information can be implemented at the CDN service provider. Continuing with the illustrative embodiment described in reference to Figure 8, the CDN service provider 106 receives information about resources requests made by client computing devices in a first time zone and processes this information by identifying content for a cache component authorized to service a client computing device in a second time zone based on the resource requests made by client computing devices in the first time zone. For example, the CDN service provider 106 may provide a list of most frequently requested content in the first time zone to the cache component which is authorized to service client computing devices in a second time zone.

[0051] Continuing at a block 906 of Figure 9, the CDN service provider manages content based on the processing performed on the associated class information at block 904. As also set forth above, a variety of logic for managing content can be implemented. In accordance with the present illustrative embodiment, the CDN service provider can manage content based on the processed information by obtaining, at the cache component in the second time zone, the identified content from a storage location or, more specifically, by preloading content from the list at the cache component in anticipation of similar requests from client computing devices in the second time zone. The routine 900 ends at block 908.

[0052] It will be appreciated by those skilled in the art and others that while processing and management functions have been described herein as being performed at various components of the CDN service provider, these functions can be distributed across one or more computing devices of the CDN service provider. In addition, the content requests and associated class information monitored at resource cache components can be maintained globally by the CDN service provider and shared with all or some subset of the cache components of the CDN service provider. Similarly, the content information identified based on processing the associated class information can be provided as a batch process to resource cache components or sent in response to a request from one cache component to another.

[0053] It will be appreciated by those skilled in the art and others that all of the functions described in this disclosure may be embodied in software executed by one or more

processors of the disclosed components and mobile communication devices. The software may be persistently stored in any type of non-volatile storage.

[0054] Conditional language, such as, among others, "can," "could," "might," or "may," unless specifically stated otherwise, or otherwise understood within the context as used, is generally intended to convey that certain embodiments include, while other embodiments do not include, certain features, elements and/or steps. Thus, such conditional language is not generally intended to imply that features, elements and/or steps are in any way required for one or more embodiments or that one or more embodiments necessarily include logic for deciding, with or without user input or prompting, whether these features, elements and/or steps are included or are to be performed in any particular embodiment.

[0055] Any process descriptions, elements, or blocks in the flow diagrams described herein and/or depicted in the attached figures should be understood as potentially representing modules, segments, or portions of code which include one or more executable instructions for implementing specific logical functions or steps in the process. Alternate implementations are included within the scope of the embodiments described herein in which elements or functions may be deleted, executed out of order from that shown or discussed, including substantially concurrently or in reverse order, depending on the functionality involved, as would be understood by those skilled in the art. It will further be appreciated that the data and/or components described above may be stored on a computer-readable medium and loaded into memory of the computing device using a drive mechanism associated with a computer readable storing the computer executable components such as a CD-ROM, DVD-ROM, or network interface further, the component and/or data can be included in a single device or distributed in any manner. Accordingly, general purpose computing devices may be configured to implement the processes, algorithms and methodology of the present disclosure with the processing and/or execution of the various data and/or components described above.

[0056] It should be emphasized that many variations and modifications may be made to the above-described embodiments, the elements of which are to be understood as being among other acceptable examples. All such modifications and variations are intended to be included herein within the scope of this disclosure and protected by the following claims.

[0056] Clause 1. A method comprising:
monitoring, at a content delivery network service provider, resource requests from a first cluster of client computing devices;
identifying a set of content corresponding to the resource requests from the first cluster;
preloading the identified set of content at a cache component authorized to service a second cluster of client computing devices.

[0057] Clause 2. The method of Clause 1, wherein the first cluster is the same as the second cluster.

[0058] Clause 3. The method of Clause 1, wherein the first cluster is different from the second cluster.

[0059] Clause 4. The method of Clause 1, wherein the first cluster of client computing devices includes client computing devices located in a first time zone.

[0060] Clause 5. The method of Clause 4, wherein the second cluster of client computing devices includes client computing devices located in a second time zone.

[0061] Clause 6. The method of Clause 1, wherein monitoring resource requests comprises monitoring resource requests from a first cluster of client computing devices in a first time zone during a time window; and wherein preloading content comprises preloading content at a cache component authorized to service the second cluster of client computing devices in a different time zone during the same time window.

[0062] Clause 7. The method of Clause 1, wherein identifying a set of content corresponding to the resource requests for the first cluster comprises identifying frequently requested content for the first cluster.

[0063] Clause 8. The method of Clause 7, wherein preloading the identified set of content comprises preloading the identified frequently requested content.

[0064] Clause 9. The method of Clause 7, wherein preloading the identified set of content comprises preloading additional content associated with the identified frequently requested content.

[0065] Clause 10. A method comprising:

obtaining a resource request from a first client computing device at a content delivery network service provider;

determining a class associated with the first client computing device;

associating the class with the requested content;

preloading the requested content at a cache component of the content delivery network service provider as a function of the associated class.

[0066] Clause 11. The method of Clause 10, wherein determining the class corresponds to observing behavior over time and determining the class based on the observed behavior.

[0067] Clause 12. The method of Clause 10, wherein preloading the requested content comprises preloading the requested content at a cache component as a function of class and as a function of time.

[0068] Clause 13. The method of Clause 10, wherein the class corresponds to a temporal domain.

[0069] Clause 14. The method of Clause 10, wherein preloading the requested content at a cache component as a function of the associated class comprises correlating a class associated with computing devices serviced by the cache component with the determined class associated with the first client computing device.

[0070] Clause 15. The method of Clause 10, wherein preloading the requested content at a cache component as a function of the associated class comprises matching a class associated with computing devices serviced by the cache component with the determined class associated with the first client computing device.

[0071] Clause 16. The method of Clause 10, wherein preloading the requested content at a cache component as a function of the associated class comprises selecting a cache component for providing content to client computing devices having a class that matches the determined class associated with the first client computing device; and preloading the requested content to the selected cache component.

[0072] Clause 17. A method comprising:

determining a class associated with a set of client computing devices;

monitoring resource requests for the determined class;
identifying at least one cache component for providing content to client computing devices as a function of the determined class;
providing content information to the identified cache component.

[0073] Clause 18. The method of Clause 17, wherein the determined class corresponds to a temporal domain.

[0074] Clause 19. The method of Clause 18, wherein the temporal domain corresponds to a first time zone, and wherein identifying at least one cache component for providing content information to client computing devices as a function of the determined class includes identifying at least one cache component for providing content information to client computing devices in a second time zone.

[0075] Clause 20. The method of Clause 17, wherein the content information comprises frequently requested content for the determined class.

[0076] Clause 21. The method of Clause 17, wherein the content information includes additional information associated with frequently requested content for the determined class.

[0077] Clause 22. The method of Clause 17, wherein the content information comprises a list of frequently requested content for the determined class.

[0078] Clause 23. The method of Clause 22 further comprising verifying availability of the listed content for delivery by the identified cache component.

[0079] Clause 24. The method of Clause 23 further comprising obtaining one or more of the listed resources if the one or more of the listed resources are not available.

[0080] Clause 25. The method of Clause 22 further comprising:
selecting content to be loaded at the identified cache component based on the list of frequently requested content;
verifying availability of the listed resources for delivery by the identified cache component; and
obtaining one or more of the listed resources if the one or more of the listed resources are not available.

[0081] Clause 26. The method of Clause 17 further comprising managing the identified content at the cache component.

[0082] Clause 27. The method of Clause 26, wherein managing the identified content at the cache component comprises moving the identified content from disk memory of the cache component to memory of the cache component.

[0083] Clause 28. The method of Clause 26, wherein managing the identified content at the cache component comprises altering cache control data associated with the identified content in memory of the cache component.

[0084] Clause 29. The method of Clause 26, wherein managing the identified content at the cache component includes obtaining the identified content from a storage location.

[0085] Clause 30. A method comprising:
determining a class associated with a set of client computing devices;
monitoring resource requests for the determined class;
identifying a second set of client computing devices as a function of the determined class;
providing content information related to the resource requests to the second set of client computing devices.

[0086] Clause 31. The method of Clause 30, wherein the determined class corresponds to a temporal domain.

[0087] Clause 32. The method of Clause 31, wherein the temporal domain corresponds to a first time zone, and wherein identifying a second set of client computing devices as a function of the determined class includes identifying a second set of client computing devices in a different time zone.

[0088] Clause 33. The method of Clause 30, wherein the content information comprises frequently requested content for the determined class.

[0089] Clause 34. The method of Clause 30, wherein the content information includes additional information associated with frequently requested content for the determined class.

[0090] Clause 35. The method of Clause 30, wherein the content information comprises a list of frequently requested content for the determined class.

[0091] Clause 36. The method of Clause 35 further comprising verifying availability of the listed resources for delivery by second set of client computing devices.

[0092] Clause 37. The method of Clause 37 further comprising obtaining one or more of the listed resources if the one or more of the listed resources are not available.

[0093] Clause 38. The method of Clause 35 further comprising:
selecting content to be loaded at the second set of client computing devices based on the list of frequently requested content;
verifying availability of the listed resources for delivery by the second set of client computing devices; and
obtaining one or more of the listed resources if the one or more of the listed resources are not available.

[0094] Clause 39. A method comprising:
monitoring resource requests at a first cache component based on observed behavior; and
providing for the availability of content associated with the observed behavior at a second cache component.

[0095] Clause 40. The method of Clause 39, wherein providing for the availability of content associated with the observed behavior at the second cache component includes moving the content from disk memory of the cache component to memory of the cache component.

[0096] Clause 41. The method of Clause 39, wherein providing for the availability of content associated with the observed behavior at the second cache component includes increasing expiration data associated with the content in memory of the cache component.

[0097] Clause 42. The method of Clause 39, providing for the availability of content associated with the observed behavior at the second cache component includes obtaining the content from a storage location.

[0098] Clause 43. A system comprising:

a first network point of presence associated with a content delivery network service provider, wherein the first network point of presence is operable to:

- obtain a resource request from a first client computing device;
- associate the first client computing device with a first cluster of client computing devices;
- monitor resource requests for the first cluster; and
- identify content for a second client computing device as a function of the monitored resources requests for the first cluster; and

a client component for managing the identified content at the second client computing device.

[0099] Clause 44. The system of Clause 43, wherein managing the identified content comprises moving the identified content from disk memory to memory of the second client computing device.

[0100] Clause 45. The system of Clause 43, wherein managing the identified content comprises increasing expiration data associated with the identified content in memory of the second client computing device.

[0101] Clause 46. The system of Clause 43, wherein managing the identified content includes obtaining the identified content from a storage location.

WHAT IS CLAIMED IS:

1. A method comprising:
 - monitoring, at a content delivery network service provider, resource requests from a first cluster of client computing devices;
 - identifying a set of content corresponding to the resource requests from the first cluster;
 - preloading the identified set of content at a cache component authorized to service a second cluster of client computing devices.
2. The method of Claim 1, wherein monitoring resource requests comprises monitoring resource requests from a first cluster of client computing devices in a first time zone during a time window; and wherein preloading content comprises preloading content at a cache component authorized to service the second cluster of client computing devices in a different time zone during the same time window.
3. The method of Claim 1, wherein identifying a set of content corresponding to the resource requests for the first cluster comprises identifying frequently requested content for the first cluster.
4. A method comprising:
 - obtaining a resource request from a first client computing device at a content delivery network service provider;
 - determining a class associated with the first client computing device;
 - associating the class with the requested content;
 - preloading the requested content at a cache component of the content delivery network service provider as a function of the associated class.
5. The method of Claim 4, wherein determining the class corresponds to observing behavior over time and determining the class based on the observed behavior.
6. The method of Claim 4, wherein preloading the requested content at a cache component as a function of the associated class comprises correlating a class associated with computing devices serviced by the cache component with the determined class associated with the first client computing device.

7. A method comprising:
 - determining a class associated with a set of client computing devices;
 - monitoring resource requests for the determined class;
 - identifying at least one cache component for providing content to client computing devices as a function of the determined class;
 - providing content information to the identified cache component.
8. The method of Claim 7, wherein the content information comprises a list of frequently requested content for the determined class.
9. The method of Claim 8, further comprising:
 - selecting content to be loaded at the identified cache component based on the list of frequently requested content;
 - verifying availability of the listed resources for delivery by the identified cache component; and
 - obtaining one or more of the listed resources if the one or more of the listed resources are not available.
10. The method of Claim 7, further comprising managing the identified content at the cache component.
11. The method of Claim 10, wherein managing the identified content at the cache component comprises altering cache control data associated with the identified content in memory of the cache component.
12. A method comprising:
 - determining a class associated with a set of client computing devices;
 - monitoring resource requests for the determined class;
 - identifying a second set of client computing devices as a function of the determined class;
 - providing content information related to the resource requests to the second set of client computing devices.
13. The method of Claim 12, wherein the content information comprises a list of frequently requested content for the determined class.

14. The method comprising:
monitoring resource requests at a first cache component based on observed behavior; and
providing for the availability of content associated with the observed behavior at a second cache component.

15. The method of Claim 13, wherein providing for the availability of content associated with the observed behavior at the second cache component includes increasing expiration data associated with the content in memory of the cache component.

16. A system comprising:
a first network point of presence associated with a content delivery network service provider, wherein the first network point of presence is operable to:
obtain a resource request from a first client computing device;
associate the first client computing device with a first cluster of client computing devices;
monitor resource requests for the first cluster; and
identify content for a second client computing device as a function of the monitored resources requests for the first cluster; and
a client component for managing the identified content at the second client computing device.

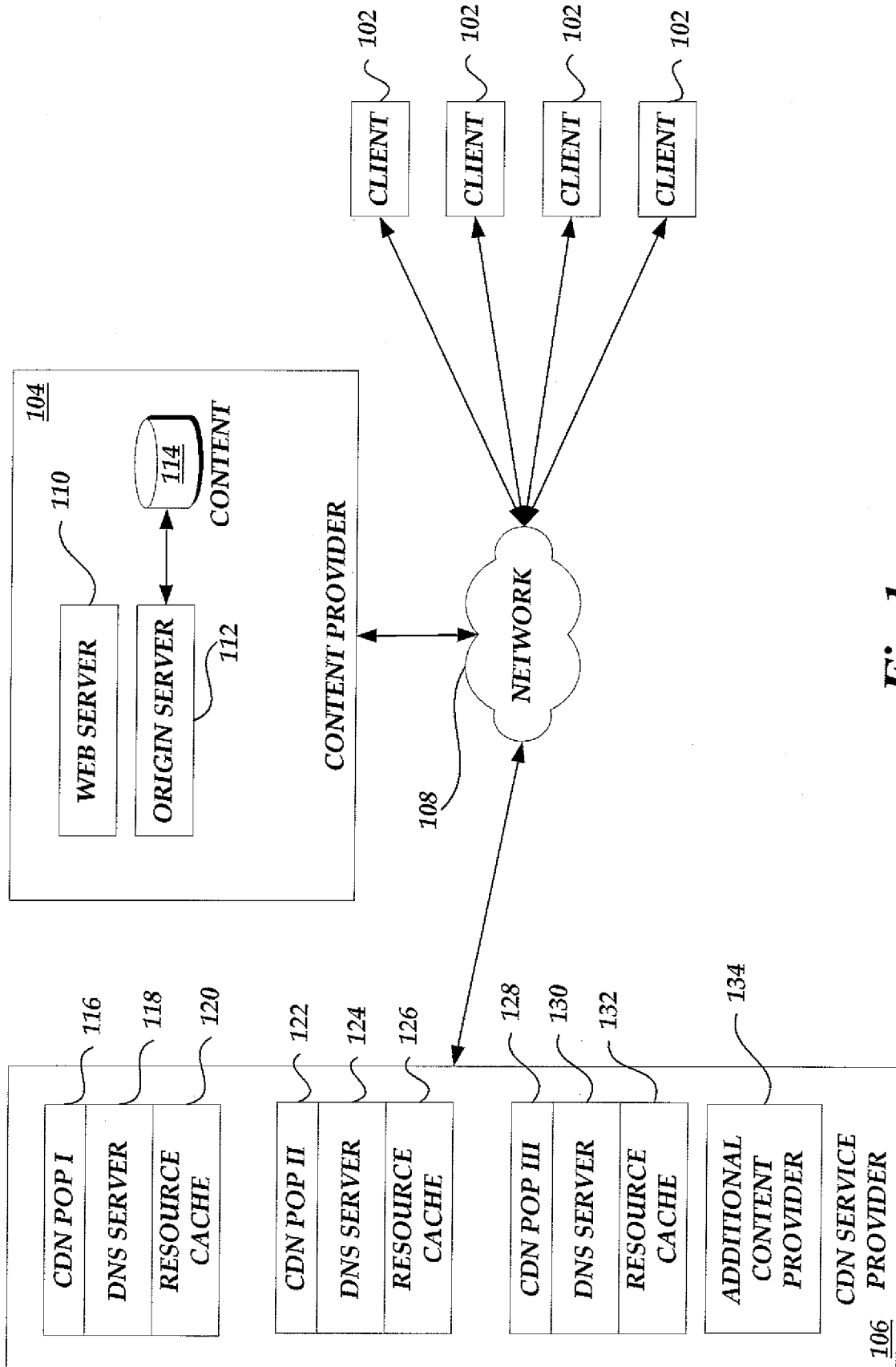


Fig. 1.

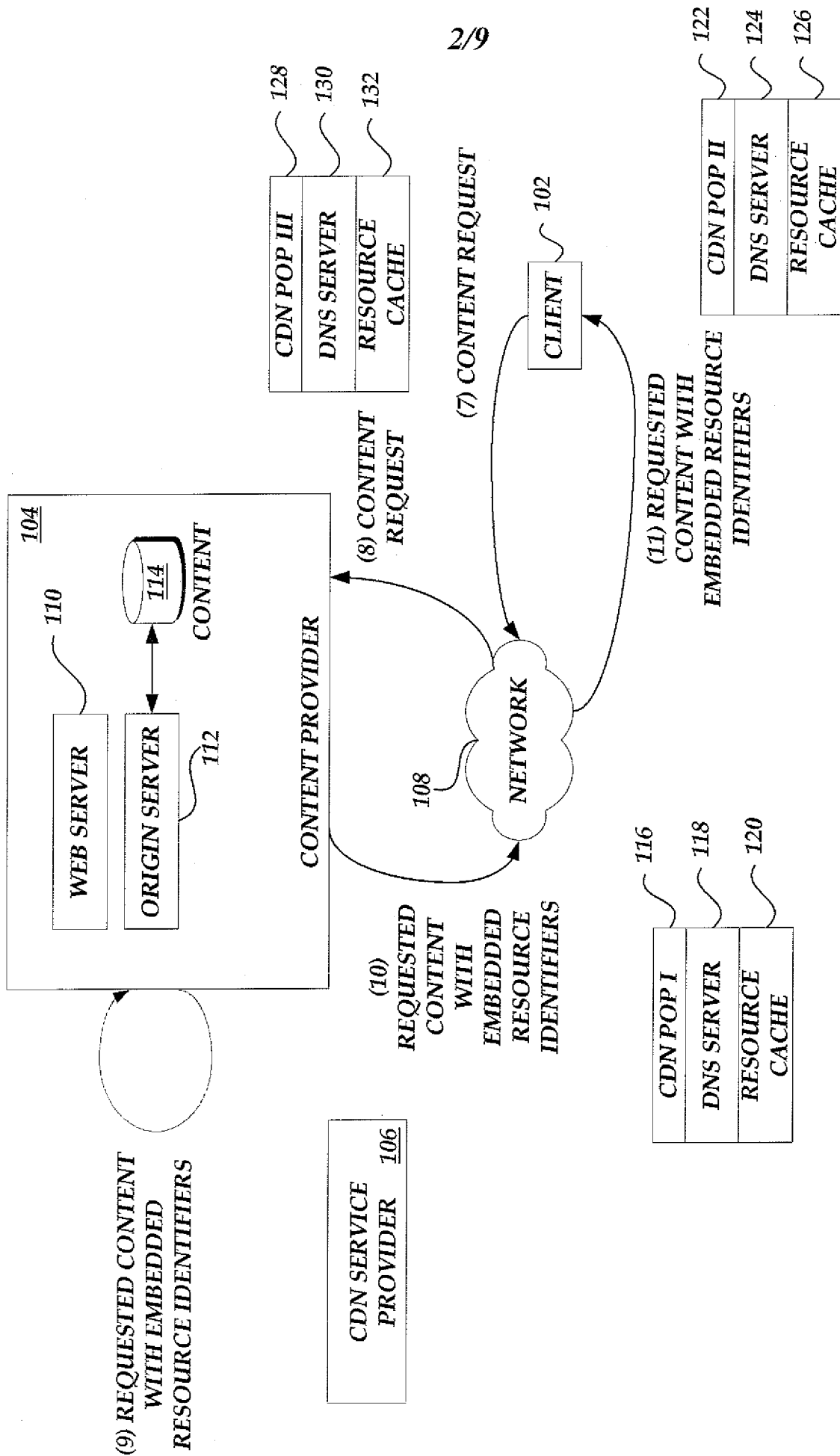


Fig. 2.

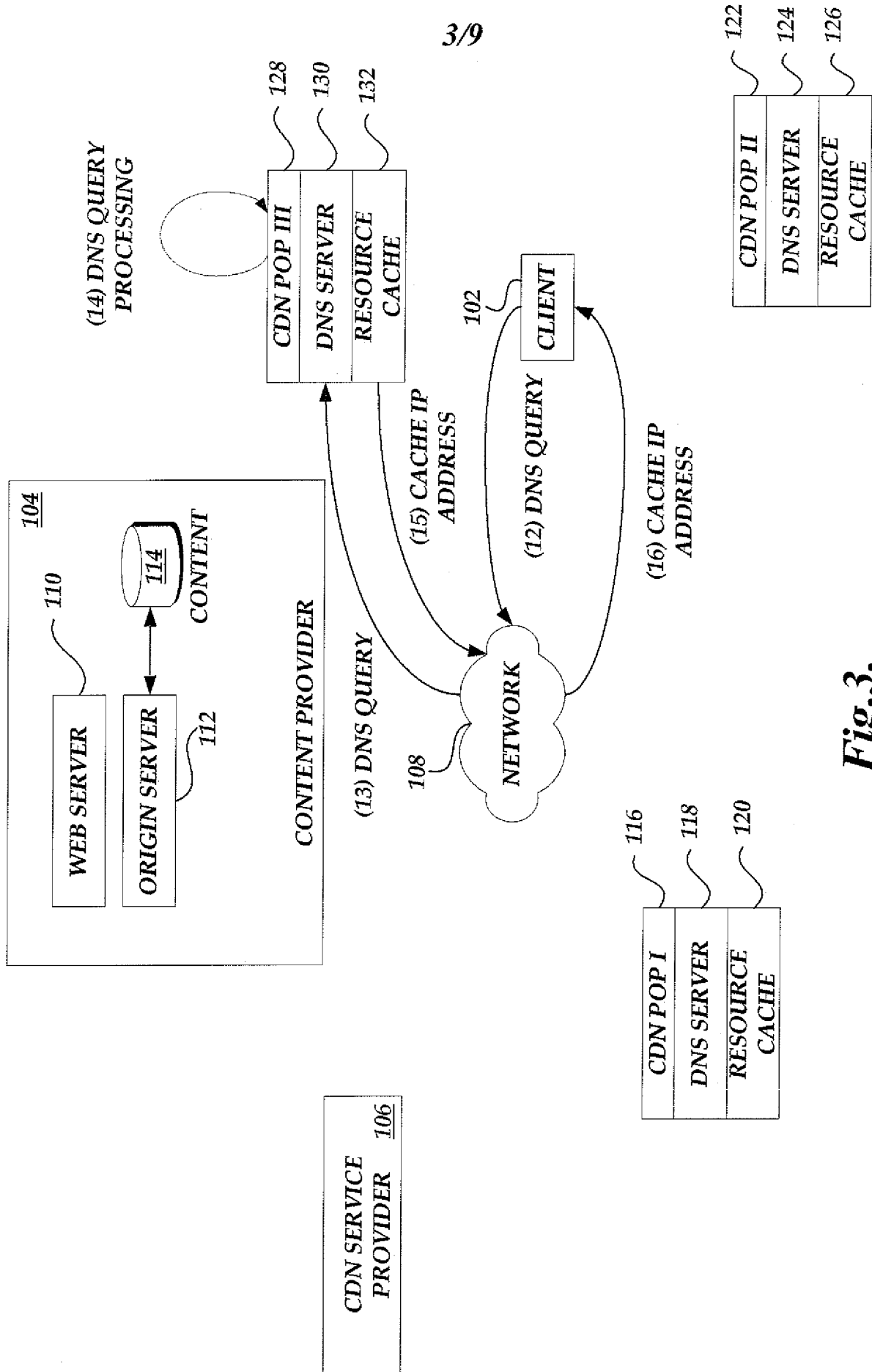


Fig. 3.

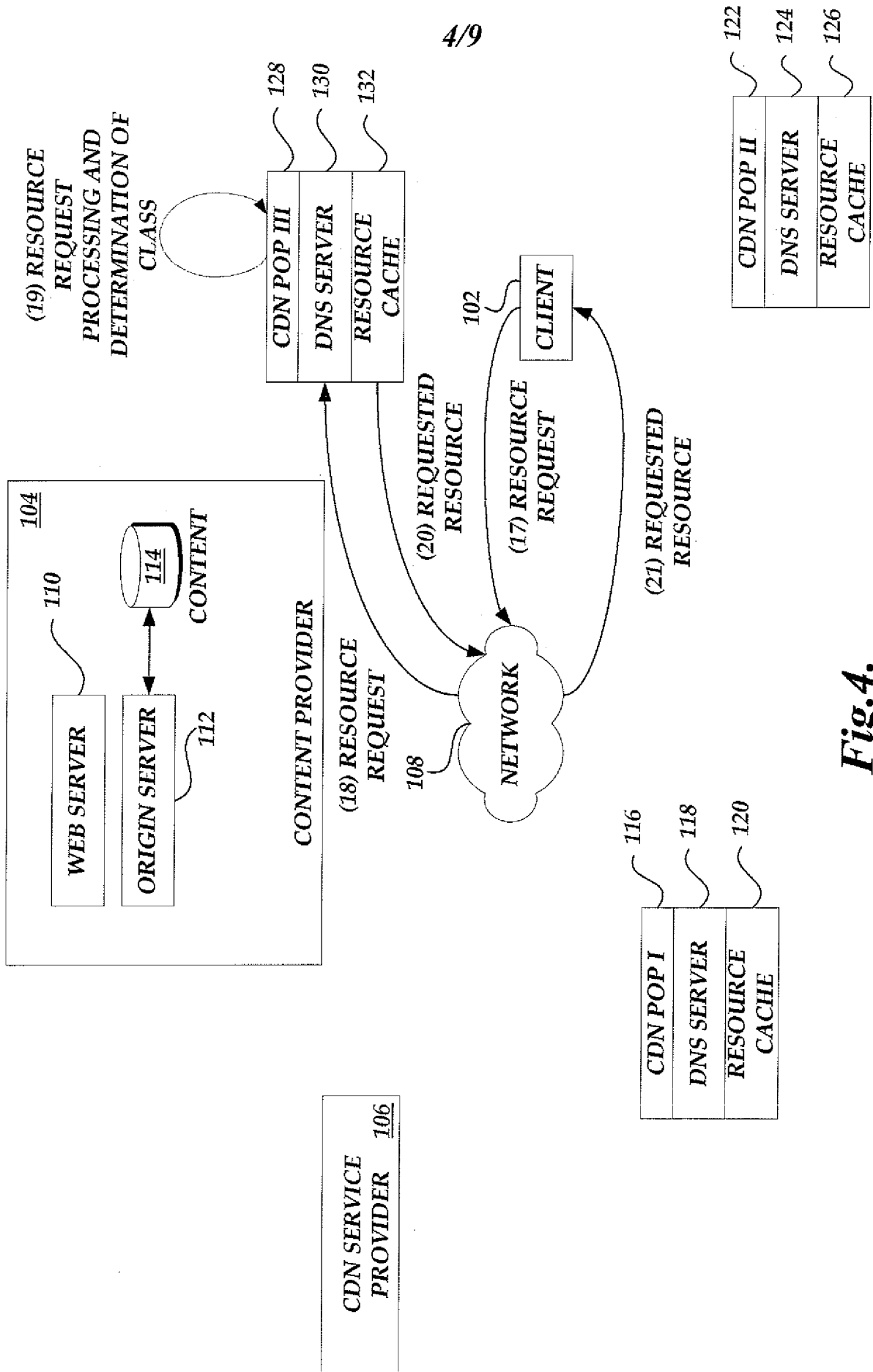


Fig.4.

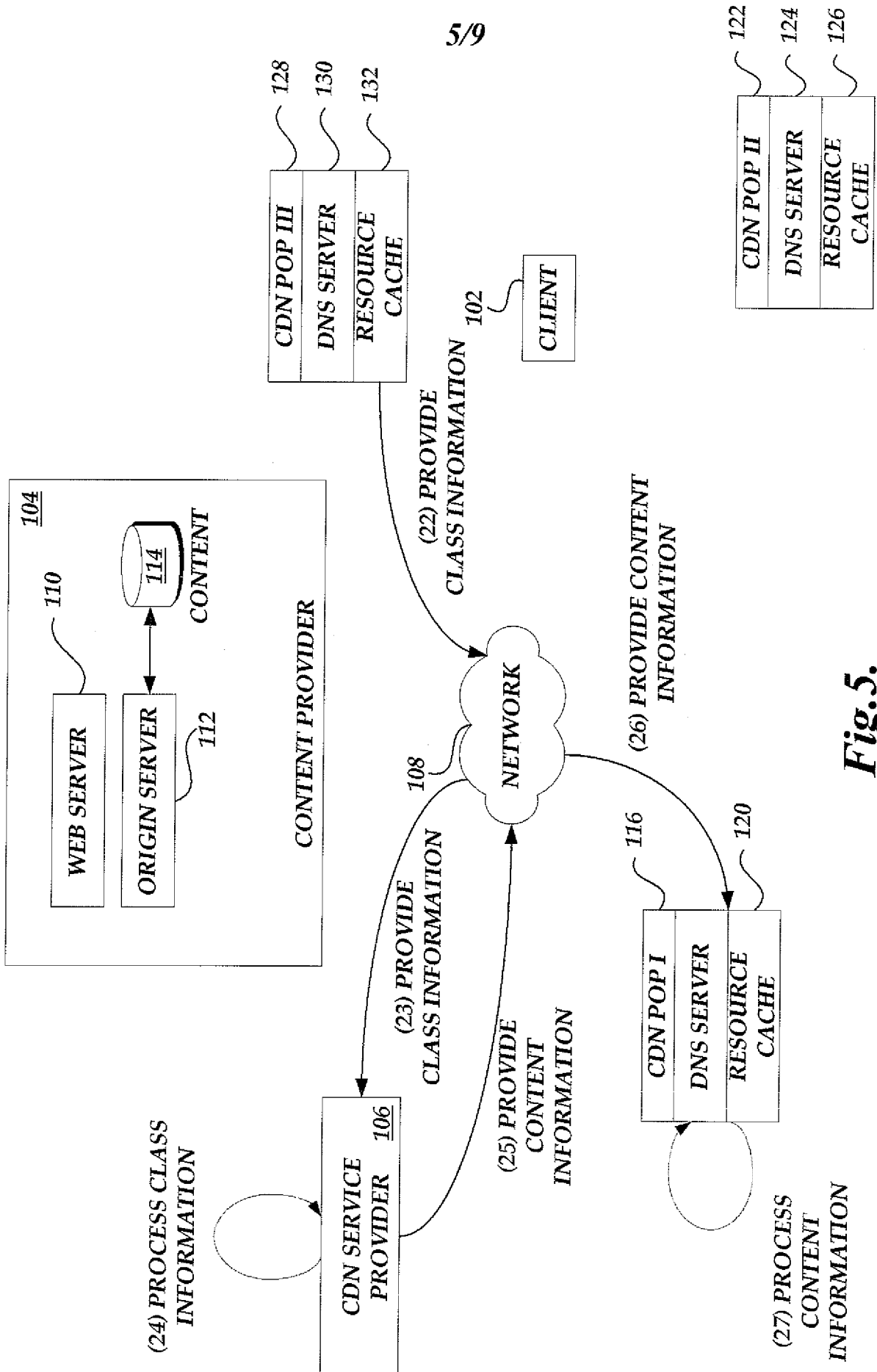


Fig. 5.

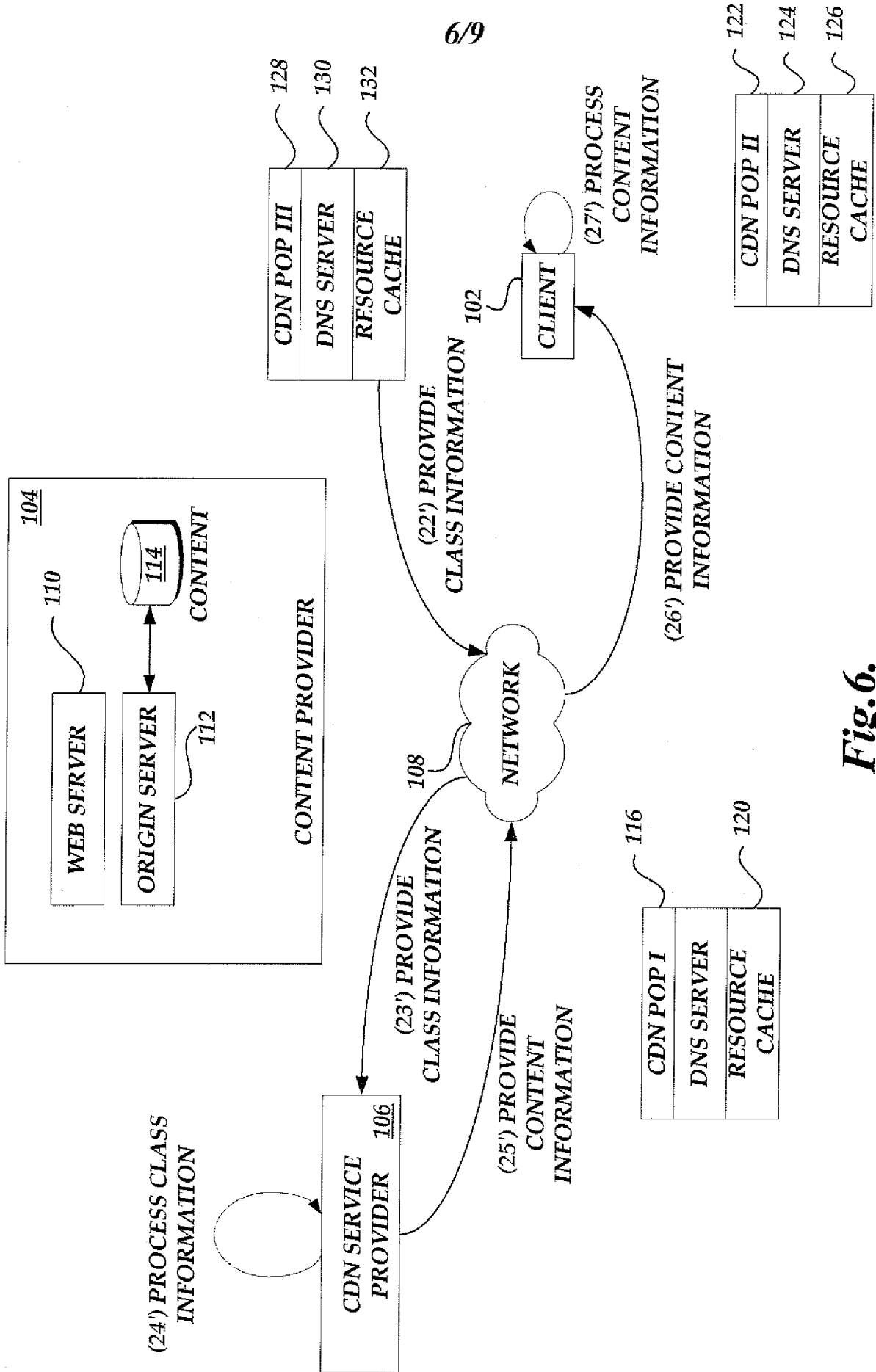


Fig. 6.

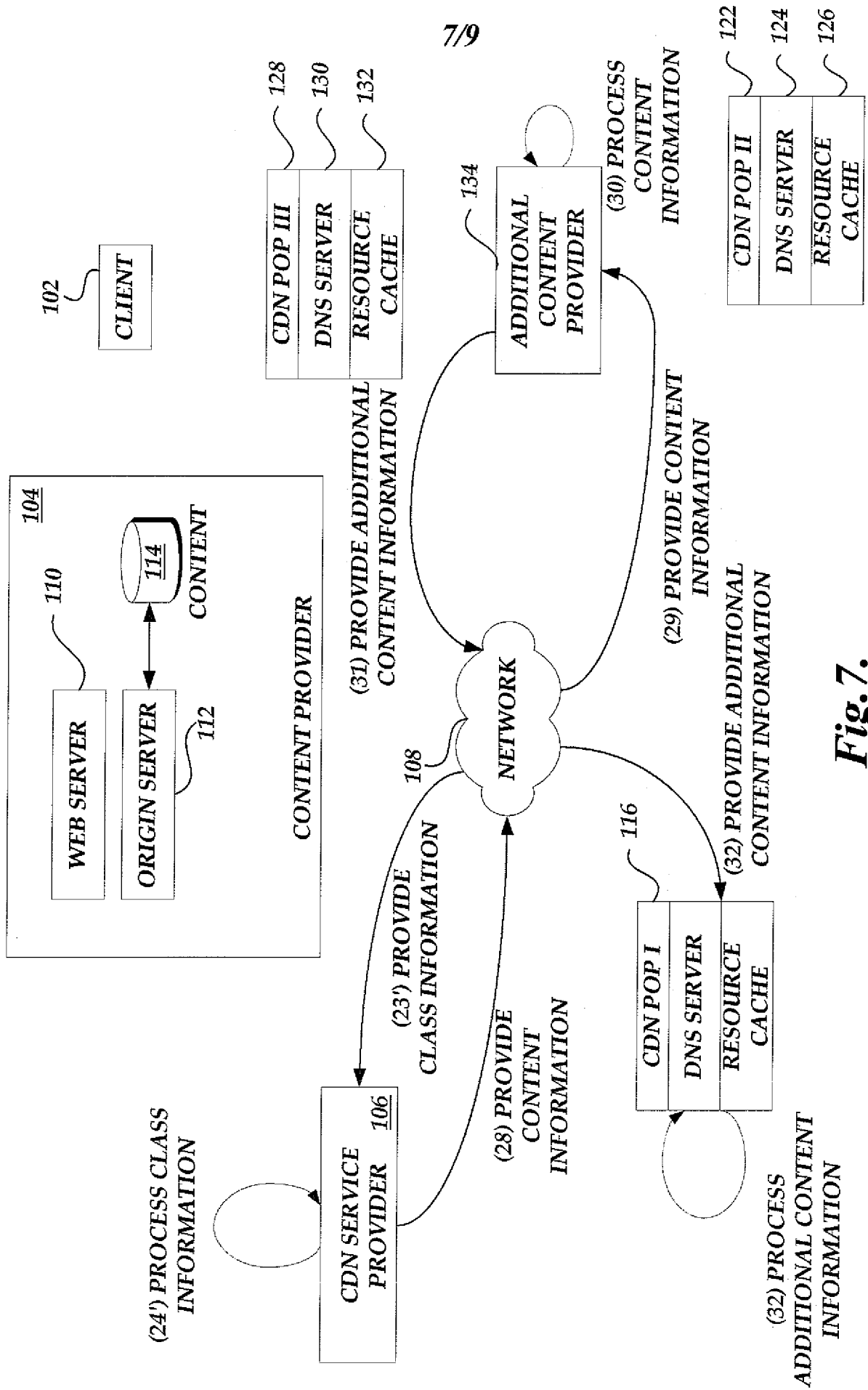


Fig. 7.

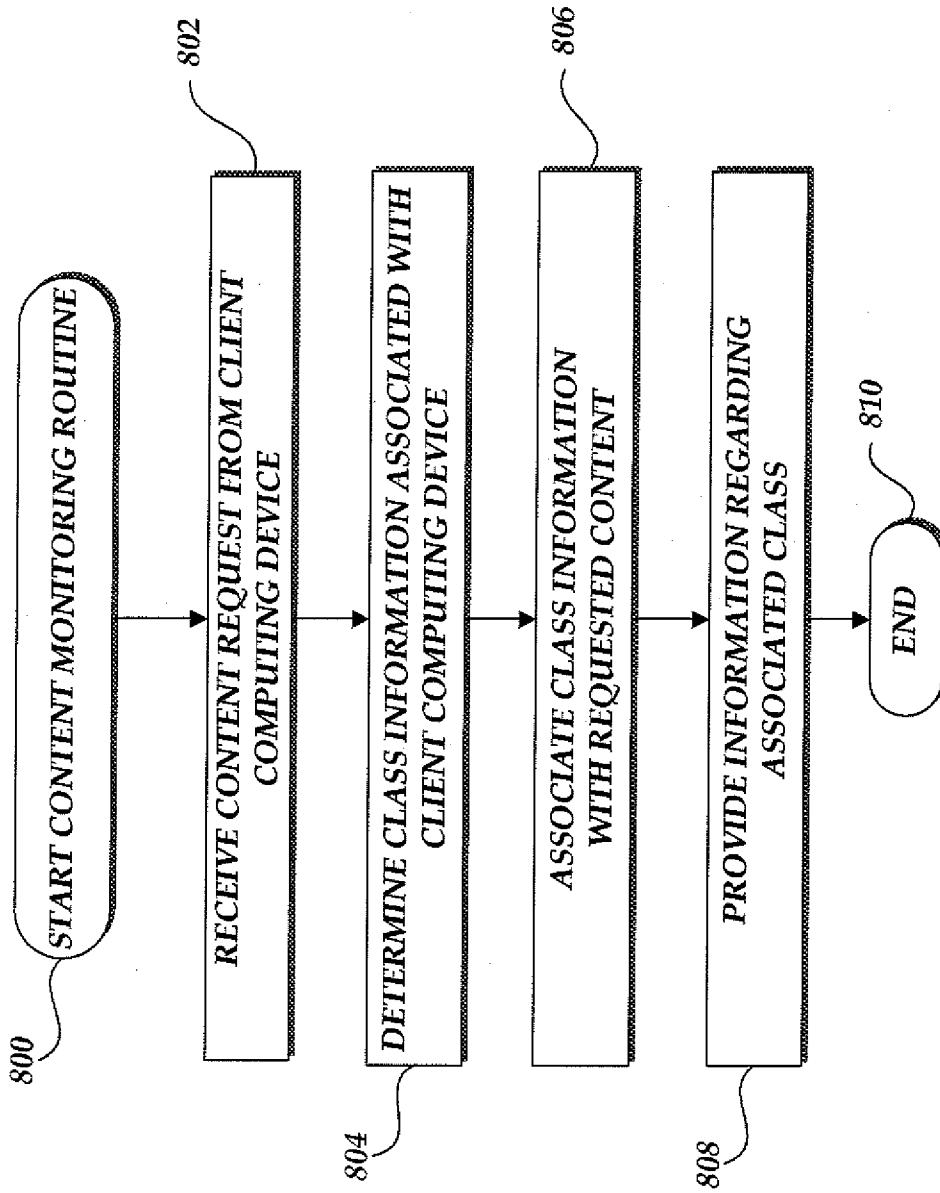


Fig. 8.

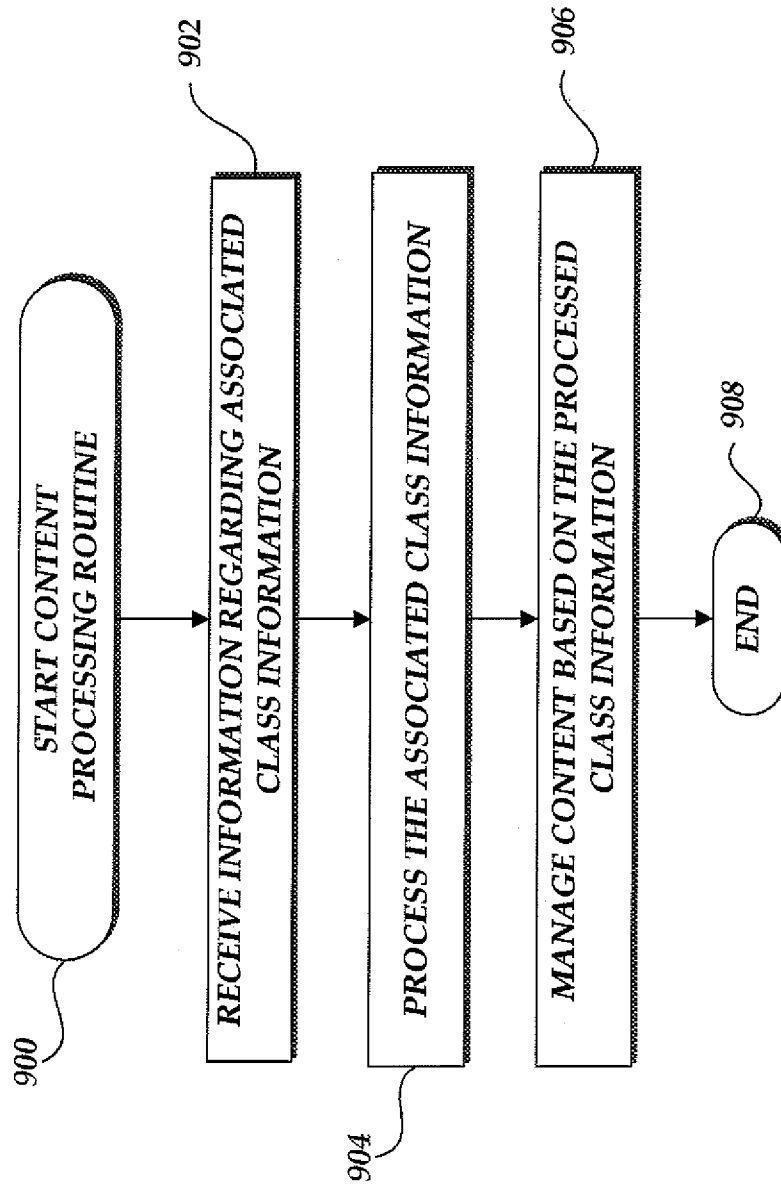


Fig. 9.